THE EVOLUTION OF U.S. REGIONAL INEQUALITY: A MIXTURE MODEL EXPLORATORY APPROACH

Chenxi Yu

REAL 12-T-04 August, 2012

The evolution of U.S. regional inequality: A mixture model exploratory approach

Chenxi Yu
Regional Economic Applications Laboratory
and Department of Urban and Regional Planning
University of Illinois at Urbana Champaign
Email: yu28@illinois.edu

Abstract: This paper studies the emergence of convergence clubs in the United States since the 1980s. The finite mixture normal model is used to identify the clubs based on the per capita personal income dataset for 700 U.S. labor market areas from 1969 to 2009. The results reveal that the collection of high income areas, termed the "rich places club," was formed in the 1980s, and the share of the rich places club stabilized at around 10-12% of total labor market areas for the 1990s and 2000s. We also find that the gap between the rich places club and the "everywhere else club" has been increasing since the 1990s.

Key words: finite mixture normal model, convergence clubs, rich places club

1. Introduction

Following Solow (1956), neo-classical economic theory predicts that, in the long run, the economic development levels in different regions within a country will tend to converge to a steady state; this is the basis of the economic convergence hypothesis. Many methods have being developed to test this economic convergence hypothesis since the 1980s (Durlauf, *et al.*, 2005). These methods can be divided into two categories: the regression analysis approach (Barro and Sala-i-Martin, 1992) and the distributional dynamics analysis approach (Quah, 1997). Conclusions reached by these two analytical methods differ from each other dramatically: the regression analysis approach usually points to convergence, while the distributional dynamics approach usually points to club convergence or divergence.

Barro and Sala-i-Martin (1992) propose the β-convergence model and find that the per capita personal income data for the 48 contiguous U.S. states provides clear evidence of the convergence hypothesis. However, as Quah (1993) points out, the results from the β-convergence model could be false because it could be a manifestation of a regression toward the mean. Thus Quah (1993a) concludes that the β-convergence model does not provide enough evidence to support the convergence argument.

Quah (1996, 1997) proposed the distributional dynamics approach to study economic growth. The distributional dynamics approach focuses on the evolution of the entire distribution over time. Quah proposes two distributional methods: the Markov chain transaction probability (Quah, 1993b) and the stochastic kernel density plot (Quah, 1997). By using both methods, Quah finds that the income distributions evolve from a unimodal "one peak" distribution toward bimodal "twin peaks" distribution: "Eventually, the middle-income group of economies vanish, and the rich continue to become richer, and the poor, poorer. Clustering occurs at high and low parts of the income distribution." (Quah, 1996)

The conflicting evidence from the regression analysis approach and the distributional dynamic approach led to the search for a new method that could help understand regional economic development patterns. In this paper, the finite mixture model is introduced as a new exploratory method to study the convergence or club convergence debate. The paper is organized as follows: the finite mixture model is introduced in section 2; data and descriptive statistics and the rationale for using the finite mixture model are in section 3; section 4 contains estimation results, section 5 is the robustness check; and section 6 provides some concluding commentary.

2. Finite mixture normal model

According to McLachlan and Peel (2000), the mixture model was first introduced to the statistical field by Karl Pearson in 1894 when he and his colleague Raphael Weldon discovered the asymmetry in the histogram of the crabs they sampled from the Bay of Naples. Pearson and Weldon suspected that the asymmetry in the histogram might be a signal that this crab population was evolving towards two new subspecies. Pearson fitted a mixture of two normal distributions with different means μ_1 and μ_2 and variance $\sigma\boldsymbol{\sigma}_1^2$ and σ_2^2 in proportions π_1 and π_2 to accommodate the apparent skewedness in the crab data. The model used by Pearson and Weldon was a mixture model with two normal distributions, thus it is a finite mixture normal model; in this paper it is also referred to as a mixture normal model. To provide readers an example of the mixture normal distribution, a two components mixture normal model was simulated with the means as 0 and 3, and standard deviations as 1 and 2. The proportion of the first normal distribution was 40%, while the proportion of the second normal distribution was 60%. In figure 1, the red line represents the first normal distribution and the green line represents the second normal distribution. Combining them together provides the overall distribution, shown as the black line in figure 1. The overall distribution is skewed toward the right side.

<<insert figure 1 here>>

A more general version of the mixture normal model can be presented as follows:

$$y_i \sim \pi_j N\left(\mu_j, \sigma_j^2\right), for\ j = 1, \dots, m$$

$$0 \leq \pi_j \leq 1, for\ j = 1, \dots, m$$

$$\sum_{j=1}^{m} \pi_j = 1$$

In our case, y_i denotes the per capita income for region i. Here it is assumed that the underlying data generating process is a mixture of m normal distributions, where each distribution has mean μ_j and variance σ_j^2. The key parameter, π_j, is the mixing proportion, or weight, of the j^{th} normal distribution. The sum of all the m normal distribution proportions $(\pi_1, \pi_2, \dots, \pi_m)$ is equal to one.

The mixture normal model provides a natural way to deal with the heterogeneity in a dataset that may contain two or more sub-populations. In the field of regional development, there has always been the debate about whether regions grow more like each other or whether they grow apart. In the language of the mixture normal model, the regional development debate can be presented as follows: for all regions, can they be classified into one normal distribution, or are the distinctions between them so great that they have to be treated as if they are drawn from different normal distributions?

To identify the convergence clubs using the mixture normal distribution, the likelihood ratio test is used to choose the number of components. Then, the Expectation-Maximization (EM) algorithm is used to estimate the means, the variances, and the mixing weights for each normal component. In the final step, a parametric bootstrap will be conducted to produce standard errors for all the parameters in the mixture normal model estimated in the second step.

Using the mixture normal approach to identify the convergence club has two significant advantages over the distributional density approach proposed by Quah (1997). First, the mixture normal model approach provides a more powerful test for the convergence club hypothesis (Pittau *et al.* 2010). To detect the convergence club, the distributional density approach relies on the detection of multimodality by observing the kernel density function, while the mixture approach relies on the detection of multiple components within the distribution. In the distributional density approach, a great deal of emphasis is placed on the researcher's personal judgment to detect convergence clubs based on the shape of the kernel density function. Compared to the distributional density approach, the mixture normal approach is a more powerful test for the convergence club hypothesis because the distribution does not have to be as sharply multimodal for this approach to detect the multiple components. Furthermore, it is possible to use the bootstrap technique to produce standard errors for the parameters in the mixture normal model. The convergence club test based on the mixture normal model is supported by statistical evidence.

The second advantage of the mixture normal approach is on the mobility analysis. Mobility is a measurement used to quantify the transitions out of and into distinct clubs. A low mobility implies stable convergence clubs, while a high mobility implies the convergence clubs are not so stable. Quah (1993b) first applied the Markov chain transitional probability approach to study the mobility of convergence clubs. The Markov chain approach has a drawback in this case because it relies on studying the transition matrix with arbitrarily defined cell boundaries—usually the entire group is equally divided into four quarters. In a later paper, Quah (1997) proposed using the stochastic kernel approach to study the mobility between the clubs. The stochastic kernel is an improved version of the Markov chain approach because it is built on a continuum transition matrix. However, the drawback of the stochastic kernel transition matrix is that it is usually represented in 3-D graphs. This approach does not provide a direct measurement for researchers to draw conclusions with respect to whether regions converge or diverge.

The mobility measurement for the mixture normal approach is derived from the conditional probability that can be calculated from the mixture normal model estimation result. The mixing weights, π_j, can be interpreted as the unconditional probability that region *i* comes from the normal component *j*. The conditional probability $\zeta_{j,i}$ for each region *i* is given by:

$$\zeta_{ji} = \frac{\pi_j N(\mu_j, \sigma_j^2)}{\sum_{j=1}^{m} \pi_j N(\mu_j, \sigma_j^2)}, for\ j = 1, \dots, m$$

$$\sum_{j=1}^{m} \zeta_{j,i} = 1$$

For each region *i*, there will be a *j* conditional probability. All the *j* conditional probabilities for region *i* sum up to one. These conditional probabilities can be used to assign region *i* to that component with the largest estimated $\zeta_{j,i}$. In this research, use is made of the regional income data from 1969 to 2009 for all the labor market areas in the continental United States; therefore it will be possible to study mobility by tracing the change of assignment of region *i* over time.

Given the advantages of the mixture normal approach over the distributional dynamics approach, the use of the mixture normal model for detecting convergence clubs is still limited. Paapaa and Van Dijk (1998) and Pittau, *et al*. (2010) used this approach to study the cross-country distribution of per capita income. In the European Union, Pittau (2005) and Pittau and Zelli (2006) used this approach to detect EU convergence clubs. Tsionas (2000) used the finite normal mixture model to study the distribution of per capita gross state product for the U.S. from 1977-1996. Tsionas found that there was a club of rich states and a club of poor states. However, his finding lacks validity, because he did not provide statistical evidence to support the significance of his estimations. Without knowing the statistical significance of these estimations, one cannot draw the conclusion that there is a division between a rich states club and a poor states club. Pittau *et al*. (2010) present the most recent development in utilizing the mixture normal model in identifying the convergence clubs. They find three categories within their data set of 102 countries: rich counties such as the U.S and many EU countries, median countries like China and Peru, and poor countries such as Nepal and Nigeria.

This paper uses the same method as Pittau *et al*. (2010); however, this paper is different in two significant ways. First, the focus is on the identification of convergence clubs for the labor markets within a country. The difference between labor markets within a country is likely to be much smaller than the difference between 102 countries. Therefore, if evidence is found to support convergence clubs within a country, the result would provide a very strong counter-argument to the β-convergence notion. Secondly, the paper utilizes spatial visualization methods to provide compelling information for the understanding of spatial development patterns of convergence clubs.

3. Data and descriptive analysis

3.1 Data

The population and personal income data used in this paper are derived from the Regional Economic Information System (REIS) provided by Bureau of Economic Analysis. The REIS provides state level data starting from the year 1929 and county level data starting from the year 1969. The first issue to address is to decide the appropriate spatial unit to use for the analysis. Most convergence studies have focused on the state level (e.g., Rey and Montouri, 1999; Tsionas 2000). However, the state level may be too large a unit to reflect local labor market dynamics. For example, Upstate New York has a totally different demographic and economic structure when compared with the New York Metropolitan area. The same situation happens in the State of Illinois: the Chicago Metropolitan area is completely different from Downstate Illinois. The second commonly used spatial unit is the county level. The county level analysis may also raise problems because the county boundary is merely an arbitrary political boundary. It does not reflect the economic structure of a region. A county may be only a part of an economic or labor market area. For example, DuPage County, Illinois, is only one part of the Chicago Metropolitan statistical area. In contrast, the third and most widely used spatial unit is the metropolitan statistical area (MSA). It is composed of one or more counties, with a relatively high population density at its core and close economic ties throughout the area. An MSA is a much more complete economic and labor market area. However, it is not defined in the rural parts of the United States.

Therefore, a more appropriate spatial unit for the study would be a system of economic and labor market areas that is defined all across the United States. The commuting zones and labor market areas classification system developed by the U.S. Department of Agriculture (USDA) fits these requirements. The USDA identified 741 commuting zones based on the 2000 census journey-to-work data. Compared with the relatively arbitrary county boundaries, commuting zones are much more useful for analysis because they represent the supply and demand of labor in the local area. This spatial unit has become more popular in recent years because it covers the entire U.S. (Tolbert and Sizer, 1996; Autor and Dorn, 2009; Molloy, *et al*., 2011; Feser and Sweeney, 2003).

In this paper, use is made of the crosswalk provided by the USDA to link and merge the REIS county level data to commuting zone level data. The per capita personal income data used are calculated simply as each commuting zone's total personal income divided by population. The per capita personal income data are then adjusted to constant 2005 dollars. In the rest of this paper, the commuting zones will be referred to as labor market areas. This study focuses on the continental part of the U.S. that includes 702 labor market areas for the period from 1969 to 2009.

3.2 Descriptive analysis

As noted earlier, a great deal of empirical research shows evidence to support the existence of convergence in the U.S. For example, Barro and Sala-i-Martin (1992) find statistically significant β-convergence effects by using U.S. state level data, while Higgins, *et al.* (2006) use U.S. county level data and also find statistically significant β-convergence effects across the U.S. On the other hand, many other studies challenge the notion of convergence. One very good example is provided by Bickenbach and Bode (2003) where the authors used the first order property Markov chains implemented with U.S. state level data. They found two structural breaks: one occurs after World War II and the other in the 1990s. The second structural break in the 1990s indicated that U.S. regional development was switching from convergence to divergence. The U.S. labor market areas data used in the current analysis reveals a similar pattern. In figure 2, the measure of σ-convergence, the coefficient of variation (CV), is plotted for the U.S. labor market area from 1969 to 2009. The coefficient of variation is decreasing from 1970 to the 1990s, while it is increasing from the mid-1990s to 2009.

<<insert figure 2 here>>

There are many reasons that income divergence could have happened in the late 1990s and 2000s. One of the most important reasons could be technological innovation. The empirical work of Galli (1997), based on a panel data set of labor productivity in 20 industrial sectors of the European Union for the period between 1960 and 1993,suggests that a period of convergence may be followed by a period of divergence as a consequence of radical technological and economic transformations. The new wave of technological innovation was led by the use of computers, semiconductors, data processing, and information and communication technologies. This technological innovation began to be adopted and implemented by the economic system in the 1980s. By the 1990s and 2000s, these new technologies spread quickly and changed not only production methods, but also almost all aspects of doing business and everyday life. Some of these changes include the use of personal computers; the development of the internet, wireless technology, and online commerce; the development of biotechnology; and the use of industrial robots. In the language of Schumpeter's technological innovation theory, this is a new long-run Kondratiev cycle (van Duijn, 1983) led by information technology (IT).

Technological innovations in the IT sector could generate two effects on the economic development levels of regions. First, because of the complementary nature of high-skilled workers and the new technologies (internet and computer), the IT development would bring more benefit to regions with more high-skilled workers (Autor and Dorn, 2009). Secondly, the positive externality from productive high-skilled workers in certain regions could attract more highly motivated workers to these regions. Therefore, a human capital accumulation and polarization process could occur because of the impact of the IT sectors (Berry and Glaeser, 2005; Florida, 2002; Moretti, 2012). The polarization of human capital naturally leads to the polarization of the economic development level.

Figure 3 shows the kernel income density functions of the 702 labor markets from 1969 to 2009. All the per capita personal income values are adjusted to the 2005 value. The distributions for the years 1969, 1979, 1989, 1999, and 2009 have been highlighted; it is clear that the income distribution is becoming more dispersed over time and the shape of the density functions are skewed to the right in many cases.

Now the kernel density functions of 1969 and 2009 will be used to further elaborate on the difference between the mixture normal approach and the distributional dynamic approach, and explain how multiple modes in the distribution do not necessarily guarantee multiple components. Using the terminology of the distributional dynamic approach, the income kernel density function of 1969 suggests a "twin-peak" distribution, while the income kernel density function of 2009 suggests a "single-peak" distribution. Therefore, if using the distributional dynamic approach, the year 1969 is more likely to have two convergence clubs than the year 2009. The mixture normal approach shows different results, as presented in the next section. According to the mixture normal approach, the evidence does not suggest that there were two convergence clubs in 1969, but there is statistically significant evidence to conclude that there were two convergence clubs in 2009.

<<insert figure 3 here>>

4. Results

One of the most efficient ways to estimate the mixture normal model is to use the Expectation-Maximization (EM)[1] algorithm (McLachlan and Peel, 2000). The initial test explored how many components should be included in the mixture model by following the likelihood ratio test procedure used in Pittau *et al*. (2010). Since most papers that utilize the mixture normal model to identify convergence clubs identify 2 or 3 clubs, the likelihood ratio was used to test the two components mixture normal model against the three components mixture normal model. For some years in the late 1980s, 1990s, and early 2000s, the likelihood ratio tests do not have enough evidence to reject the three components mixture normal model. In the case of a model with three components, in addition to the high income and middle income groups, a low income group is identified. However, this low income group does not persist over time. It appears and disappears in relation to economic conditions. Also the population share of this low income group is much smaller compare to the other two groups. Since this research focuses on identifying the regional inequality that was driven by the disproportional growth of the high income group, it does not change our analysis when we group the middle income group with the low income group. That is why, to facilitate comparisons of the results across years, a two-component model was estimated for all 41 years.

The model estimation results include the mixing proportion/weight for each component, the means for each component, the variance for each component, and also the conditional possibilities for each region. Then, a bootstrap procedure was performed to construct standard errors for the mixing proportions, the means, and the variances. In the case of the two components mixture model, it was only necessary to show one mixing proportion/weight because the two mixing proportions sum to one. The group of labor market areas with significantly higher per capita personal income will be referred to as the "rich places club" (hereafter RPC), and the group of labor market areas with significantly lower per capita personal income as the "everywhere else club" (hereafter EEC).

<<*insert figure 4 here*>>

The present use of the mixture normal model to identify the convergence clubs parallels the practice of biologists when they evaluate the evolution of a species and decide when differences are significant enough to deserve the classification of a new species. In figure 4, the mixing proportion/weight for the RPC with the higher per capita personal income from 1969 to 2009 are plotted: the black line with square marks is the proportion for the group of labor market areas with a higher per capita personal income, and

---

[1] In this paper, the EM algorithm from the "mixtools" package in R is used (Benaglia, *et al.* 2009)

the two grey dashed lines are the 90% confident interval for the mixing proportion parameter. The lower bound of the 90% confidence interval is zero or less than zero before the mid-1980s, but in the second half of the 1980s the lower bound increases to above the zero line. The interpretation follows that the RPC began to appear in the 1980s with the proportion of the rich places club increasing in this decade. In the 1990s and 2000s, the RPC accounts for about 10-15% of all the labor market areas. The proportion of RPC has not increased very much in the 1990s and 2000s.

<<*insert figure 5 here*>>

Figure 5 shows the plots of the average per capita personal income for the EEC and the RPC from 1985 to 2009. The RPC has a larger variance than the EEC. Figure 5 shows that the per capita personal income gaps between the two clubs grew over time. A linear regression model was estimated to see whether the gap between the two clubs was becoming larger over time using the per capita personal income ratio of the RPC to the EEC as a measure of the gap between these two clubs and time as an independent variable. Here, the time is from year 1 to year 25 instead of year 1985 to 2009. The estimated model is as follows:

$$\underset{}{\text{Ratio}} = \underset{(0.0139)}{1.30} + \underset{(0.0009)}{0.0024} * \text{Time}, \quad \text{adjusted } R^2=0.1899$$

On average, the RPC is 1.30 times richer than the EEC. From 1985 to 2000, the RPC became 0.24% richer each year when compared with the EEC.

In figure 6, the spatial distribution of the RPC for 1990-2009 has been plotted. The darkness of grey color represents how many times a labor market area is classified into the RPC between 1990-2009. All the RPC are classified into four quartiles. The top quartile consists of labor market areas that are classified as belonging to the RPC 19-20 times out of 20 time periods; it is the most stable rich club. The quartile consists of the labor market areas that are classified as belonging to the rich places club 9-18 times out of 20 time periods. They are also very stable as members of the rich club. The third quartile collects labor market areas that are classified as belonging to the rich place club 3-8 times out of 20 time periods; essentially, they are the emerging, stable rich club members. The fourth quartile is composed of labor market areas that are classified as belonging to the rich places club only 1-2 times out of 20 time periods; for the most part, these are areas that have not yet qualified as members of rich places club.

For this fourth quartile, labor markets usually move in to the rich places club in one year and move out of the rich places club in the second year. The most active years in which labor markets joined the RPC in the fourth quartile were 1991, 2002, 2005, and 2008, while the most active years of moving down to the EEC for the fourth quartile were 1992, 2003, 2007, and 2009. In the fourth quartile, of the 35 labor market areas, 12 of them are from North and South Dakota. An analogy to the situation of this group of labor market areas would be a "lucky" lower league English football team that finds itself occasionally promoted to the Premier League. However, the performance of this "lucky" football team is never very consistent. That is why, in the next season, the team drops back to its original league. Here, the question is, will this team be again promoted and eventually be able to stay in the higher league? One cannot answer this question without studying the "lucky" factors that allowed this football team to be promoted in the first place. The same thing applies for this current research: there is not enough information to judge whether the labor market areas that made it into the fourth quartile will enter the RPC again in the future. Further research is required.

It is important to distinguish between the third and the fourth quartile. As was just discussed, entering the RPC once or twice could simply be because of idiosyncratic factors. However, for the third quartile, there should be more consistent reasons to explain how these labor market areas were able to enter the RPC for 3-8 times in the last two decades. Therefore, future research should pay special attention to the third quartile and find out the reasons why this group of labor market areas was able to emerge, and establish themselves as stable members of the rich places club.

<<*insert figure 6 here*>>

Figure 6 also suggests there is a spatial pattern in terms of distribution of the RPC. Four spatial clusters stand out clearly: the Boston-New York-DC cluster, the southern Florida cluster, the Colorado cluster, and the California cluster. A hot spot analysis using the General *G* statistic (Getis and Ord, 2010) confirmed the existence of these four clusters.

The spatial distribution of the RPC suggests that the rich places are usually places with larger population masses; the average population size for each quartile is presented in table 1. The average population size for "always rich (quartile 1) labor market areas" is 2.6 million in 2009, while the average population size for the EEC is only 0.23 million. This pattern seems to amplify the effect of economic agglomeration: a labor market with a larger population mass is more productive than a labor market with smaller population mass (Fujita and Thisse, 2002). The average population size for the second quartile is very close to that of the first quartile, with 2.2 million people in 2009. In contrast with the first two quartiles, the average population sizes for the third and the fourth quartile are only 0.6 and 0.5 million, respectively.

<<*insert table 1 here*>>

5. Robustness Check: Mobility between components

In section 4, the analysis revealed that there is a RPC that emerged in the 1980s and the this club accounts for 10-12% of the 702 labor market areas in the Continental U.S. The emergence of the RPC is one necessary requirement for the existence of convergence clubs. The other requirement is low mobility between clubs. Mobility can be equated with transitional probabilities: it refers to the probability transfer into or out of distinct clubs over time. Low mobility will mean that most of the labor market areas that belong to the rich places club in time period *t* will also belong to the rich places club in time period *t+1*. If the low mobility condition holds over time, then it suggests that the classification system and the convergence clubs are stable.

The conditional probability, ζ_{ji}, is used for the identification of the labor market areas that belong to the RPC and the labor market areas that belong to the EEC; labor market areas are assigned to clubs according to their maximum estimated conditional probability. Since a panel of data for 41 years is available, it is possible to trace the mobility of the change of assignment of labor market areas over time.

We check the transitional probability for three stages: (1) 1970-1979, before the formation of the RPC; (2) 1980-1989, the formation of the RPC; and (3) 1990-2009, when the RPC is stable and the incomes diverge between the RPC and the EEC. Four cells in the transitional probability matrix correspond to four different situations: the diagonal refers to stability; the upper right hand cell indicates downward mobility (rich to non rich) while the lower left cell indicates upward mobility (non rich to rich).

The upward mobility plus downward mobility divided by the share of stable rich places club is used as a measure of stability of the rich places club. For the instability measurement, a smaller number means more stability because there is less upward and downward mobility. The other way to understand this stability measurement is to treat it as a mobility measure: a large number means higher mobility in between clubs. The upward mobility minus downward mobility is used as a measure of the formation of the rich places club. The three key indicators in summarized in table 3.

<<insert tables 2 and 3 here>>

Two indicators confirmed the classification of these three stages. First, the shares of stable RPC members were increasing from 1.51% in the 1970s, to 4.84% in the 1980s, and 7.12% in the 1990s-2000s. Secondly, the rates of formation of the RPC were negative in the 1970s and the 1990s-2000s, with values of -0.13% and -0.15%, respectively, while the rate of formation of the RPC was 0.88% in the 1980s. The positive RPC formation rate in the second stage confirmed the 1980s as the time when the RPC was established. The third indicator, the instability of the RPC measurement, is very important for the convergence club test. In the first stage, the 1970s, the RPC was very unstable; the instability measurement for this stage is 1.01, meaning high mobility between the two clubs. This corroborates the evidence shown in figure 4 from the previous section: the mixing parameter is insignificant in the 1970s. The RPC was relatively stable for the second and third stage, with instability measurements of 0.36 and 0.33, respectively. There is no clear cut point for determining whether a stage is stable or mobile. However, with only 0.33 instability measurement for stage 3 (about 15% inflow and 18% outflow), one might suggest that stage three is relatively stable. Hence, the suggestion can be made that the convergence clubs identified by the mixture normal model for U.S. labor market areas are stable clubs.

What is the difference in between this transitional probability approach and the Markov chain approach? In the Markov chain approach, it is assumed that the regional per capita income follows a first-order Markov Chain with stationary transition probabilities. However, in the research setting of an economic convergence test, the Markov chain might not be stationary over time, as shown in Bickenbach and Bode (2003). The results in this paper also show that the transition probability is not a stationary process.

6. Conclusion and future research

In this paper, the objective was to demonstrate that the mixture normal model can be used as a new way to identify convergence clubs, and in a more general way, to identify the heterogeneity in the dataset. Using this method, we are able to identify a group of labor market areas in the United States that have significantly different per capital personal income levels than other labor market areas. These high personal income level areas constitute the RPC that formed in the 1980s, and the size of this club has been stable through the 1990s and 2000s. Within the U.S., regions are becoming even more polarized and this polarization trend that began in the 1980s is also based on personal income levels (Autor and Dorn, 2009).

However, this paper only focuses on the identification of the convergence clubs. It does not provide an explanation for the possible mechanisms that created this divide between the clubs. The results support the recent findings of Moretti (2012) who explored the role of the accumulation of human capital as one mechanism that may have created this regional divide, since it seems that places with higher levels of human capital are more productive. There are two possible ways of enhancing the human capital level of a region: increasing the education level of its people, and attracting high-skilled migrants. Future research might profitably test the contribution of these two human capital enhancing channels to the creation and sustainability of the RPC and EEC.

7. References

Autor, D., & Dorn, D. (2009). The growth of low skill service jobs and the polarization of the US labor market. NBER Working Papers.

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

Benaglia, T., Chauveau, D., Hunter, D., Young, D., & others. (2009). mixtools: An R package for analyzing finite mixture models. Journal of Statistical Software, 32(6), 1–29.

Berry, C. R., & Glaeser, E. L. (2005). The Divergence of Human Capital Levels across Cities. Papers in Regional Science, 84(3), 407–444.

Bickenbach, F., & Bode, E. (2003). Evaluating the Markov property in studies of economic convergence. International Regional Science Review, 26(3), 363–392.

Chatterji, M., & Dewhurst, J. H. L. (1996). Convergence Clubs and Relative Economic Performance in Great Britain: 1977–1991. Regional Studies, 30(1), 31–39.

Durlauf, S. N., Johnson, P. A., & Temple, J. R. . (2005). Growth econometrics. Handbook of economic growth, 1, 555–677.

Feser, E., & Sweeney, S. (2003). Out-migration, depopulation, and the geography of US economic distress. International Regional Science Review, 26(1), 38–67.

Florida, R. L. (2002). The rise of the creative class: and how it's transforming work, leisure, community and everyday life. Basic Civitas Books.

Fujita, M., & Thisse, J. F. (2002). Economics of agglomeration: Cities, industrial location, and regional growth. Cambridge Univ Pr.

Galli, R. (1997). Is There Long Run Industrial Convergence in Europe? International Review of Applied Economics, 11(3), 333–368.

Getis, A., & Ord, J. K. (2010). The Analysis of Spatial Association by Use of Distance Statistics. In L. Anselin & S. J. Rey (Eds.), Perspectives on Spatial Data Analysis, Advances in Spatial Science (pp. 127–145). Springer Berlin Heidelberg.

Higgins, M. J., Levy, D., & Young, A. T. (2006). Growth and convergence across the United States: evidence from county-level data. The Review of Economics and Statistics, 88(4), 671–681.

McLachlan, G. J., & Peel, D. (2000). Finite Mixture Models. John Wiley & Sons.

Molloy, R., Smith, C. L., & Wozniak, A. K. (2011). Internal migration in the united states. NBER Working Papers.

Moretti, E. (2012). The New Geography of Jobs. Houghton Mifflin Harcourt (HMH).

Paapaa, R., & Van Dijk, H. K. (1998). Distribution and mobility of wealth of nations. European Economic Review, 42(7), 1269–1293.

Pittau, M. G., & Zelli, R. (2006). Empirical evidence of income dynamics across EU regions. Journal of Applied Econometrics, 21(5), 605–628.

Pittau, M. G., Zelli, R., & Johnson, P. A. (2010). Mixture models, convergence clubs, and polarization. Review of Income and Wealth, 56(1), 102–122.

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

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

Quah, D. T. (1996). Empirics for economic growth and convergence. European Economic Review, 40(6), 1353–1375.

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

Rey, S. J., & Montouri, B. D. (1999). US regional income convergence: A spatial econometric perspective. Regional Studies, 33(2), 143–156.

Solow, R. M. (1956). A contribution to the theory of economic growth. Quarterly Journal of Economics, 70(1), 65–94.

Tolbert, C. M., & Sizer, M. (1996). US commuting zones and labor market areas: A 1990 update. Rural Economy Division, Economic Research Service, U.S. Department of Agriculture, Staff Paper No. AGES-9614.

Tsionas, E. G. (2000). Regional growth and convergence: evidence from the United States. Regional Studies, 34(3), 231–238.

van Duijn, J. J. (1983). The long wave in economic life. Allen & Unwin.

Young, A. T., Higgins, M. J., & Levy, D. (2008). Sigma Convergence versus Beta Convergence: Evidence from US County-Level Data. Journal of Money, Credit and Banking, 40(5), 1083–1093.

8. Figures and tables

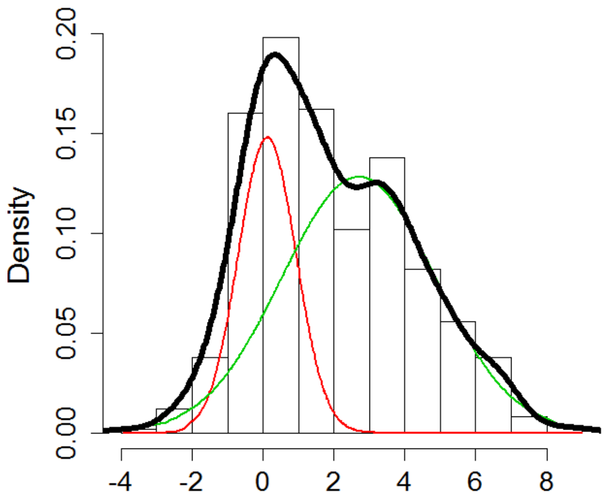

Figure 1: A simulated example of the mixture normal distribution with two components

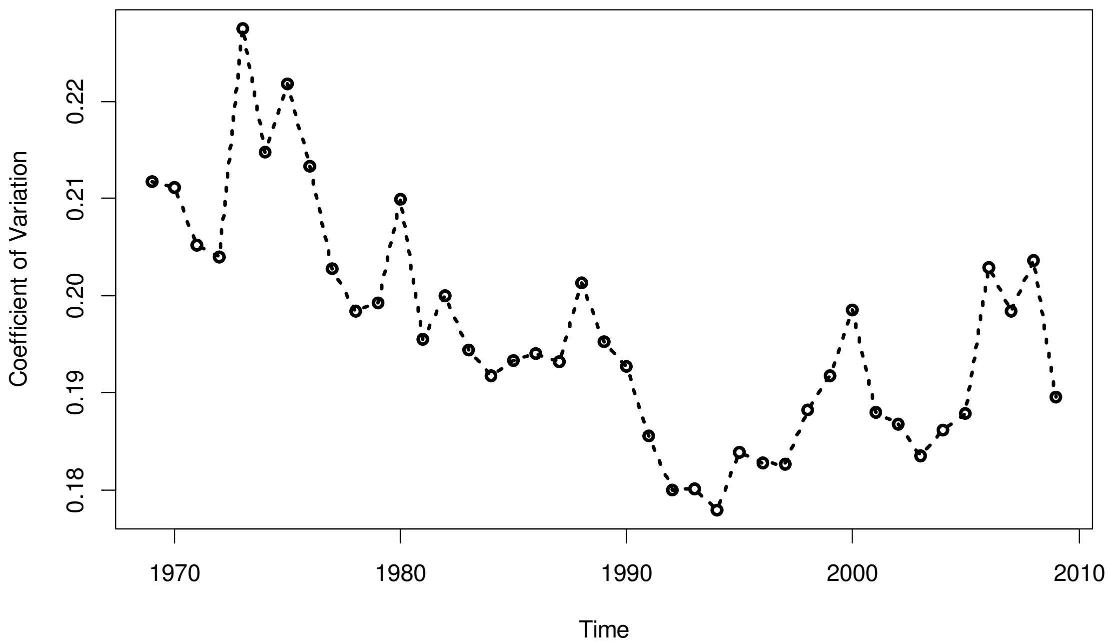

Figure 2: Coefficient of variation for labor market areas, 1969-2009

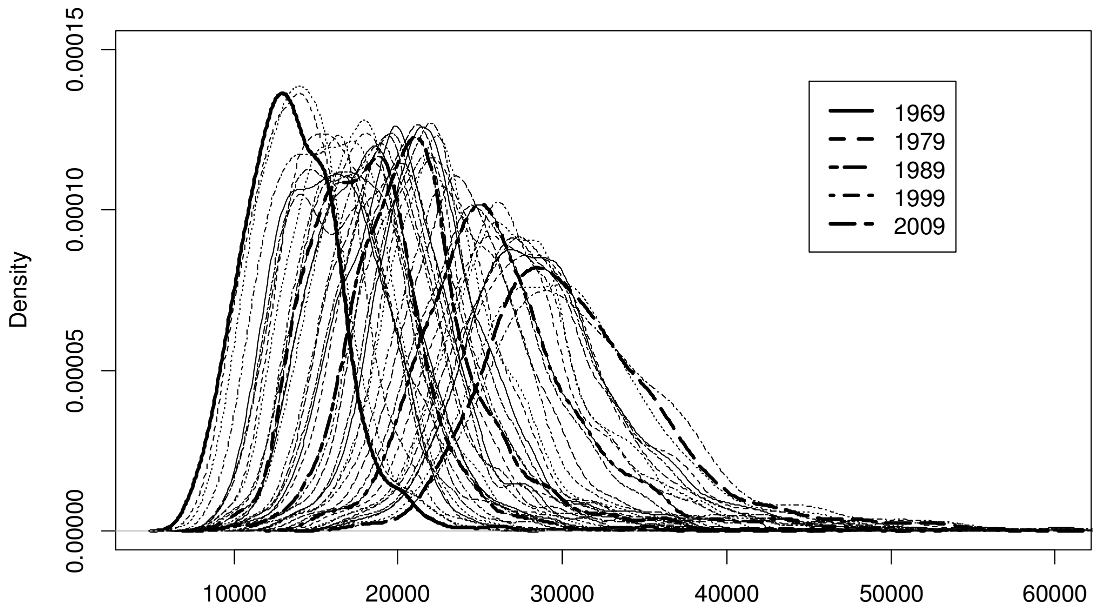

Figure 3: Per capita income distributions from 1969-2009

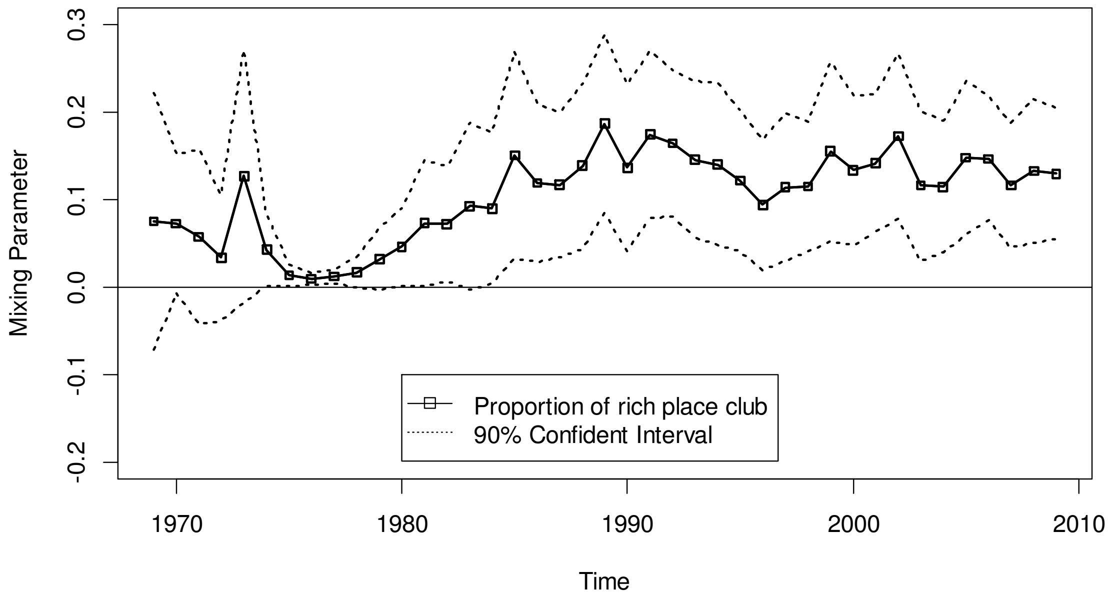

Figure 4: Mixing proportion of the rich places club 1969-2009

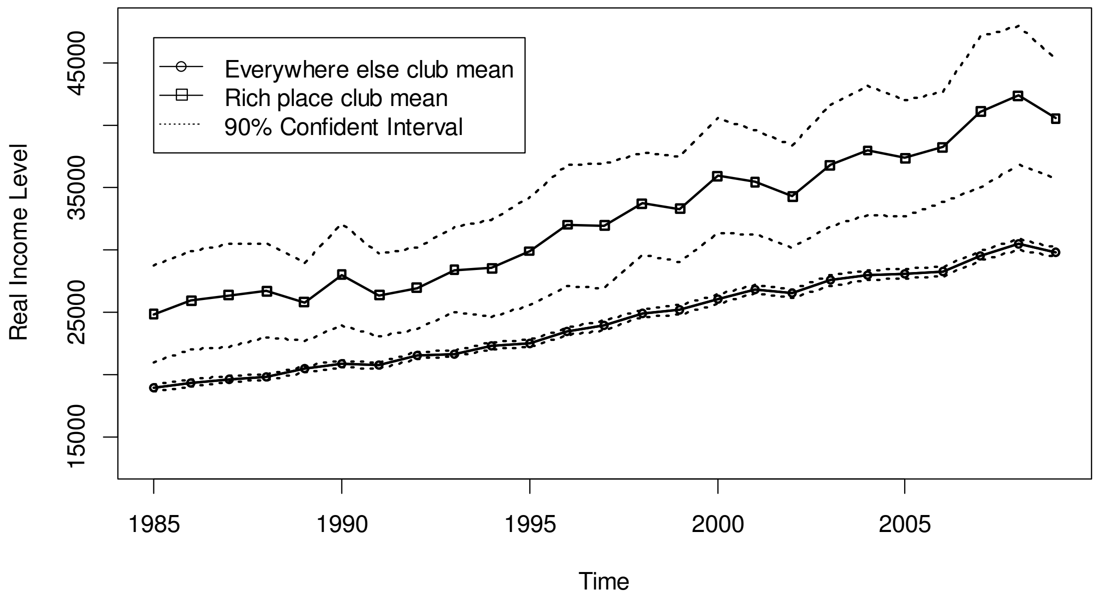

Figure 5: Average per capital income for the two clubs from 1985-2009

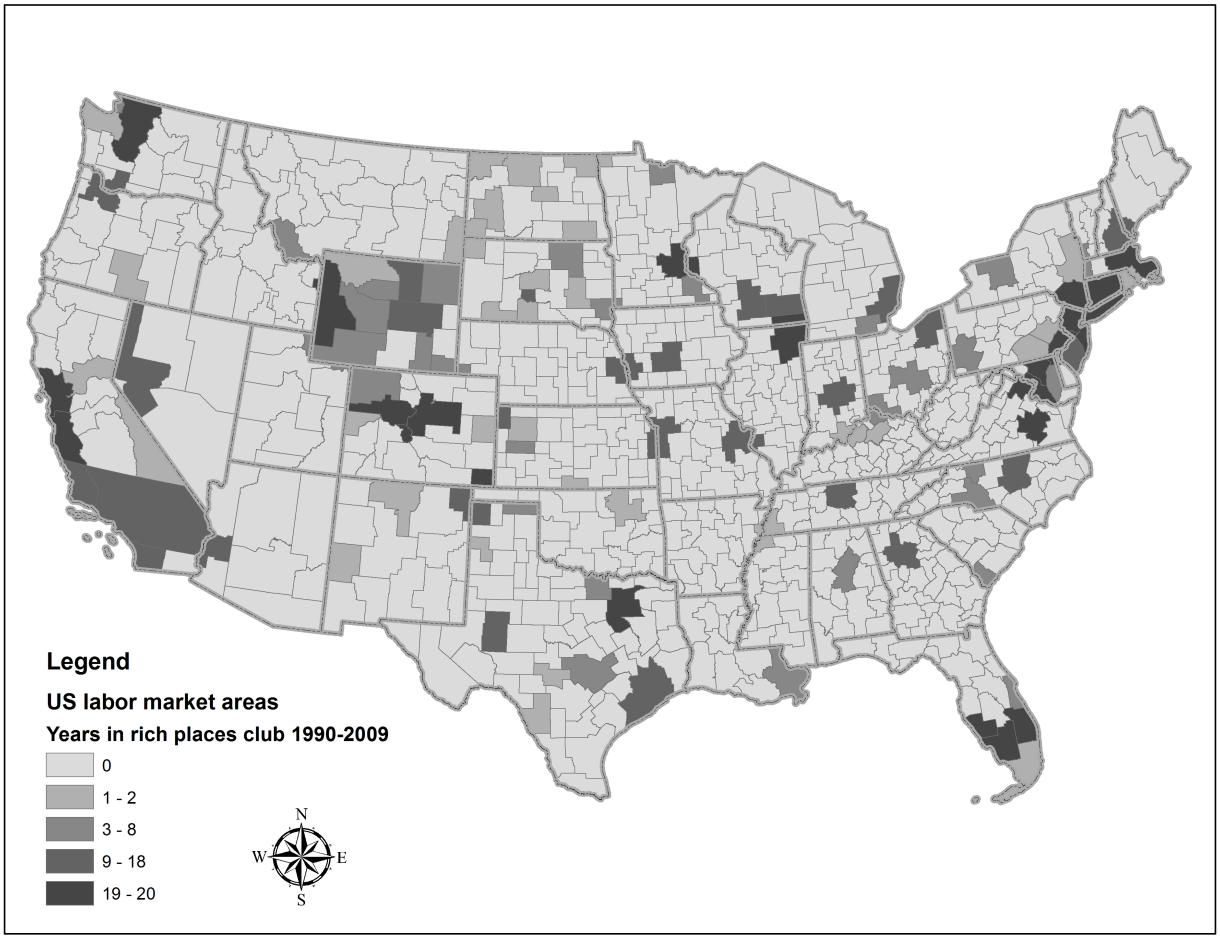

Figure 6: the spatial distribution of the rich places club, 1990-2009

Table 1: Average population size for four quartile of rich places club and everywhere else club, 2009

	LMA	Total population	Share of U.S. population	Average population size
Quartile 1: 19-20	29	75,820,125	25%	2,614,487
Quartile 2: 9-18	27	60,388,869	20%	2,236,625
Quartile 3: 3-8	31	17,753,180	6%	572,683
Quartile 4: 1-2	35	17,312,969	6%	494,656
Everywhere else club	580	132,209,828	44%	227,948
Total	702	303,484,971	100%	432,315

Table 2: transitional probability for three stages

Transitional Probability 1970-1979

		Time *T*	
		Rich place	Everywhere else
Time *T-1*	Rich place	1.51%	0.83%
	Everywhere else	0.70%	96.97%

Transitional Probability 1980-1989

		Time *T*	
		Rich place	Everywhere else
Time *T-1*	Rich place	4.84%	0.44%
	Everywhere else	1.32%	93.39%

Transitional Probability 1990-2009

		Time *T*	
		Rich place	Everywhere else
Time *T-1*	Rich place	7.12%	1.25%
	Everywhere else	1.10%	90.53%

Table 3: Measurement of the transitional probability

	Stable rich place club	Rate of forming rich place club	Instability of rich place club
1970-1979	1.51%	-0.13%	1.01
1980-1989	4.84%	0.88%	0.36
1990-2009	7.12%	-0.15%	0.33