EXTENDED ABSTRACT

Title: Long-run technological convergence and divergence of American cities (1860–2010)

Authors and e-mail of all:
Jesús Peiró-Palomino, e-mail: peiroj@uji.es
Joan Crespo, e-mail: jcrespo@ut-capitole.fr
Pierre-Alexandre Balland, e-mail: p.balland@uu.nl

Department: Jesús Peiró-Palomino, Department of Economics. Joan Crespo, LEREPS. Pierre-Alexandre Balland, Department of Economic Geography.

University: Jesús Peiró-Palomino, Jaume I University (Spain). Joan Crespo, University of Toulouse (France). Pierre-Alexandre Balland, Utrecht University (Netherlands).

Subject area: S08 – The geography of innovation and knowledge spillovers.

Abstract:
This paper studies the patterns of technological convergence and divergence of 302 American cities (Metropolitan Statistical Areas, MSA). We collect data on patents from the HistPat database (see Petralia et al., 2016) and the United States Patent and Trademark Office (UPSTO) and evaluate long run patterns (1860-2010) of patenting activity. Convergence is approached using nonparametric methods based on the distribution dynamics and conditional density estimation. We also assess the evolution of the diversity of the portfolio of patents.

Keywords: convergence, diversity, patents
JEL codes: O3; N0

Long-run technological convergence and divergence of American cities (1860–2010)

Preliminary and incomplete manuscript

[Please do not quote]

1 Introduction

Innovation is an essential factor in the generation of economic dynamism, progress and development. Although there are numerous studies considering innovation at the national level, innovation processes usually take place locally, and this is one of the reasons why it is important to focus on smaller geographical settings. Moreover, the current innovation patterns might be the result of long-run processes started many decades ago, being the historical component essential to understand today's reality. In this paper we analyse the evolution of the innovation across 302 Metropolitan Statistical Areas (MSA) of the USA over a period of 150 years (1860–2010). This is possible by considering historical data on patents, provided by the HistPat dataset (Petralia et al. 2016), combined with data from the United States Patent and Trademark Office (UPSTO).

The study is approached using nonparametric methods based on the study of the evolution of the kernel distributions of patents per capita. In a first step, we assess the shape of the distributions over time and their dynamics, as they show important changes over the decades which are the result of different processes of convergence and divergence. In a second step, we proceed to analyse the potential determinants of the patterns found. As it is fairly difficult to find historical data apart from population size (available at the American Census), we could not analyse factors other than the structure of the patent portfolio. In that regard, we focused on the diversity of the portfolio, considering 36 technological subcategories.

The preliminary results reveal that innovation is concentrated in highly innovative MSA. In addition, we found four differentiated periods of convergence and divergence. The analysis of the portfolio diversity suggests that there is some correspondence with the observed periods. In particular, convergence periods are accompanied by an increase in diversity and vice versa. Results by geographical areas show that West and South regions have traditionally had a more concentrated portfolio, although there have been a progressive catch up process that concluded during the 1970s and the 1980s. Finally, our next step is focusing on the complexity of the innovation and how it might have affected the observed patterns. Complexity can be calculated at the patent level and we expect that it could further clarify our results so far.

The rest of this preliminary working document is structured as follows. Section 2 contains some notes on the methodology used to analyse convergence. In Section 3, information on the sample, the data and some descriptive statistics are provided. Section 4 is devoted to present the first results of our research and, finally, Section 5 concludes with a brief summary of our main findings.

2 Methodological notes

Distribution dynamics has been widely used in the convergence literature, especially to analyse country and regional convergence in income per capita or productivity levels. In practice, however, its great support in this field and its flexibility have made the technique a useful tool to analyse other issues in economics.

The approach consists of estimating the following kernel density function for the variable of interest at different periods t:

$$\hat{f}_t(y) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{1}{h}\left\| y - Y_i \right\|\right) \quad (1)$$

where n is the number of geographical units (MSA in our case), Y_i is the target variable (patents per capita), K is a kernel function and h is the bandwidth parameter. Finally, $\| . \|$ is the Euclidean distance.

The kernel selected is the Gaussian kernel, defined as:

$$K(x) = \left(\sqrt{2\pi}\right)^{-1} exp\left(-\frac{1}{2}x^2\right) \tag{2}$$

In practice, the choice of the kernel is not an essential issue in non-parametric methodologies. In fact, differences between other competing alternatives such as the rectangular, the triangular or the Epanechkinov kernel are marginal. Of much greater importance is the selection of the bandwidth (h), which determines the amplitude of the bumps (Silverman,1986). In the case h is too small it produces an excessive number of bumps (undersmoothing), which severely hinders the understanding of the data structure. Oppositely, a too large h (oversmoothing) produces an excessive degree of smoothing and some data particularities might remain hidden. Among the different automated procedures for bandwidth selection we use the method proposed by Silverman (1986), which is for computability reasons the default option in most statistical softwares.

From Equation 1, evidence in favor of convergence is found when the probability mass in the distribution of the variable under analysis accumulates around a certain value. It is common practice to standardise the data before the analysis, for instance, dividing each observation by the sample mean. This implies dealing with relative data instead of their original units of measure. If that is the case, MSA converge "to the mean" when the probability mass concentrates around the unity.

Analysing the distributions at different t periods is a good strategy to study disparities in given years. However, this *static* approach does not permit the study of the internal *dynamics* of the distribution, i.e. whether the MSA have remained stable in their respective relative positions or, on the contrary, if they have moved to a different stage of innovation over time. This issue is tackled by means of conditional density estimation (CDE). This allows for the examination of the law of motion that describes how F_t (the distribution at time t), converts into F_{t+s} after s periods, subject to the relative position of each economy in t. The transition is defined by a n-th order Markov process:

$$\forall s \geq 1 : F_{t+s} = M^s F_t \tag{3}$$

where M is a representation of a stochastic kernel mapping the transition.

Let us denote as $f_t(y)$ and $f_{t+s}(m)$ the cross-section distribution of patents per capita at time t and $t+s$, respectively. Their time evolution is represented by:

$$f_{t+s}(m) = \int_0^\infty v_s(m \mid y) f_t(y) dy \tag{4}$$

where $v_s(m \mid y)$ is the conditional density which shows the probability of a MSA transiting between two particular innovation states, given its relative innovation level at period t. Following Rosenblatt (1969), the conditional density is estimated dividing the bivariate density function by the implied marginal:

$$\hat{v}_s(m \mid y) = \frac{\hat{f}_{t,t+s}(y,m)}{\hat{f}_t(y)} \tag{5}$$

where

$$\hat{f}_{t,t+s}(y,m) = \frac{1}{nh_y h_m} \sum_{i=1}^{n} K\left(\frac{1}{h_y} \| y - Y_i \|_y\right) K\left(\frac{1}{h_m} \| m - M_i \|_m\right) \tag{6}$$

is the join density of (M,Y) and $\hat{f}_t(y)$ is the marginal density of Y, being h_y and h_m the respective associated bandwidths.

The results are entirely graphical and provided using Hyndman et al.'s (1996) visualization tools. They consist of three-dimensional plots that show the stacked conditional densities for a grid of values of the conditioning variable, that is, the relative position at year t. The stacked densities plot the intra-distribution mobility of the MSA in the sample, conditioned on their initial relative position. The interpretation of the results from the plots is also straightforward. If the probability mass in the graphs concentrates along the main diagonal it means that MSA have remained stable in the same relative position, i.e. there are no signs of convergence. On the contrary, if the probability mass was widespread in year t (Y axis) but accumulates around a certain value in year $t+s$ (X axis) then convergence took place. Additional details on the methodology can be found in the seminal papers (Quah, 1993, 1996, 1997), and in recent applications (Ezcurra et al., 2005; Poletti Laurini and Valls Pereira, 2009; Peiró-Palomino, 2016).

3 Empirical framework and descriptive statistics

The sample comprises 302 MSA, for which we have information on both the number of patents and population. Information on patents is provided by the HistPat dataset (Petralia et al. 2016) and the United States Patent and Trademark Office (UPSTO)[1]. The former offers information on patenting activity at the county level from 1836 to 1975, whereas the latter provides information from 1975 to the current days. We control by population size by computing patents per capita using population data from the American Census. Information is aggregated to the MSA level in order to deal with meaningful statistical units in terms of innovation and population. MSA are regions with relatively high population density at its core and close economic ties throughout the area. They are normally made of a large city with great influence over its region. There are 382 MSA, although there are some data constrains in some years so

[1] Accessible at https://www.uspto.gov/

that the final sample covers 302 MSA for the period 1860–2010, for which we were able to obtain complete information. The sample is representative in terms of both patenting activity and population. Depending on the decade, our sample of MSA captures from 84% to 93% of the total patents in USA. In terms of population, they represent from 60% to 76% of the total population in the USA.

As patenting activity typically exhibits high volatility, yearly information is aggregated per decades in order to obtain more representative figures. In doing so, data attributed to 1860 is actually the average of the period 1860–1869. The last year for which we have data is 2015. Then, year 2010 corresponds to data from 2010 to 2015.

Table 1 provides descriptive information for the ratio of patents per capita (defined as patents per 1,000 inhabitants) in each decade. The mean is also displayed by the left panel in Figure 1, where the horizontal dotted line represents the average for the whole period. As can be observed, the mean ratio increased in the first decades until 1900, when the tendency becomes negative until 1940. From this decade a positive tendency is observable until 2000, with a downward movement in the 80s. In almost all decades we find MSA registering no innovation or remarkably low levels, whereas the maximum levels are achieved, by far, in the last decades. In contrast, the period 1930-1980 is characterized by a reduction of the extreme observations. Considering statistics that account for the entire distribution as the kurtosis and the skewness coefficients, the maximum levels are for 2010, while the minimum are found in 1910, being the decades with the greatest and smallest overall disparities, respectively.

We finally consider the percentage of MSA with the ratio of patents per capita below and above the mean. As can be observed, there is a great proportion of MSA below the mean over the entire period. Again the minimum (58.61%) and the maximum (70.53%) correspond to 1910 and 2010, respectively. As we find such important disparities, it might be insightful to look at the proportion of MSA above 2, 4 and 6 times the average. In that regard a notable part of the sample (between 8% and 14%) innovates between twice and four times the average, thus giving an initial idea on how great disparities are. Logically, smaller shares are found above the interval 4–6 times the average and more than 6 times the average. This latter category, however, is increasing in the last decades, suggesting that some MSA are raising dramatically global disparities.

4 Results

4.1 Identifying periods of convergence and divergence

We first examine the evolution of a concentration index as the Gini index for patents per capita, displayed in Figure 1, panel b). A stable pattern is not observable. In contrast, the data reveal four well-defined periods of convergence and divergence. The first one corresponds to years 1860–1910, characterized by the end of the mechanical revolution and by a reduction of the Gini index of 37%, which represents a period of convergence. The second period comprises the years 1910–1930 and shows the opposite tendency. This period witnessed the electrical revolution and in only 20 years the Gini index raised by 26%, similar to levels in 1870. Then, this short period was of great divergence. The third period is longer, covering years 1930–1980. It is a period characterized by a 34% decline of the Gini index, which is in 1980 similar to that in 1910, then showing that there was remarkable convergence in those 50 years. Finally, the fourth period starts in 1980 and it lasts until nowadays. This period corresponds with the electronic revolution and great divergence is clearly observed, the Gini increasing by 25% in 30 years.

After the analysis of the Gini index over our period of 150 years we identify five relevant decades. The initial and final ones: 1860 and 2010; and three intermediate turning points: 1910, 1930 and 1980. Then, the subsequent analyses will be based on these periods. To start explaining the dynamics of innovation, Figure 2 displays the density functions for the distribution of patents per capita in the decades of interest. To ease the interpretation and following the common practice in the literature of convergence based on the analysis of distributions, data have been standardized to the mean, i.e. the ratio of patents per capita in each MSA and decade was divided by the sample mean in that decade, thus showing how many times a given MSA is above or below the mean in a particular temporal point. Taking into account that the average is represented by 1, densities show a long right tail in all periods of interest, implying that in all decades there are MSA far above the mean. Yet there are notable differences determined by the nature of the period (convergence or divergence). For instance, in 1860 the most innovative MSA was 10 times above the mean. In 1910, after 50 years of convergence, no MSA was above 4 times the average, being all 302 MSA much more similar in innovative terms. However, in the subsequent period of divergence, the right tail became much longer, with MSA again above 8 times the average. The tail shortened moderately in 1980 after 50 years of reduction of the Gini index and, finally, in 2010 we observe MSA which are around 20 times above the average. In all cases, the flatness of the tail indicates that it is formed by very few observations, thus meaning that the convergence and divergence processes are driven by a reduced group of MSA, which we identify latter on.

Apart from the long right tail, the distributions show that most of the MSA concentrates in the interval 0–5, being this feature highly persistent over decades. Having such a great share of the sample with levels twice the average, or even 3, 4 or 5 times the average means that innovation is nonexistent or very low in a substantial proportion of MSA, which contrasts with the high levels shown by others. This feature is particularly clear for 2010, the one with the longest tail and with the sharpest mode. In any case, the shape of the distributions reveal that, regardless of the dynamics of convergence or divergence, innovation takes place in a reduced group of MSA, whereas the largest proportion of the sample show levels below the mean.

4.2 Intradistribution mobility: identifying the actors behind convergence and divergence

This section focuses on the analysis of the intradistribution movements of the densities, which enables to identify the actors driving the observed convergence and divergence processes. To do so, we apply conditional density estimation, explained in detail in Section 2. Figure 3 shows the stacked densities for the four periods. To easily identify the MSA in each relative position, Figure 4 displays results in two dimensions, including a grid and labels for the MSA. Considering the first period, we observe two important shifts in the probability mass. On the one hand, a group of regions with no innovation or with very low levels in 1860 started to innovate in the period, and were in 1910 on the average or even above the average in some cases, as for instance Denver-Aurora-Broomfield (CO), San Diego-Carlsband-San Marcos (CA) or Los Angeles (CA). On the other hand, a relatively large group of MSA worsened their relative position. Some of them are Miami-Fort Lauderdale-Pompano Beach (FL), Tampa-St. Petersburg-Clearwater (FL), Boston-Cambridge-Quincy (MA) or Worcester (MA), all below the main diagonal of the scatterplot of Figure 6. However, the most notable worsening is that for New Heaven-Milford (CT), which was more than 10 times the average in 1860 but is only 3 times the average in 1910. Then, this first period of convergence was leaded by a relative improvement of some MSA in the West coast, while others from the East performed relatively worse.

Regarding the second period, defined by overall divergence, we again observe two opposite forces. On the one hand, while some MSA from the West coast such as San Diego-Carlsband-San Marcos (CA) or Fresno (CA) worsened their relative position significantly, a larger group of MSA witnessed a remarkable improvement. This can be easily observable from the stacked densities, as there are two spots of probability mass above 6 and 8 times the average in 1930 for MSA which were only between twice and 3 times the average in 1910. Apart from these particular cases, we observe that the area between 3 and 6 times the average in 1930 is nourished by MSA which were between 2 and 3 times the average at the beginning of this period. These are, to name the most notable, Columbus (IN), Toledo (OH), Dayton (OH), Racine (WI) and especially South Bend-Mishawaka (IN). The two most extreme cases corresponds to Denver (DE) and Kankakee-Bradley (IL). Again, there seems to be some spatial pattern. In this occasion the MSA driving the overall divergence process are mostly located in States from the North and East of the USA.

Focusing on the period 1930–1980, characterized by convergence, the greatest force comes from the severe downturn of some MSA which leaded the divergence process in the preceding period. These are Denver (DE) and Kankakee-Bradley (IL), which in 1980 are back to average levels. This is perfectly observable in the stacked densities, which experienced a general clockwise movement and showing that most of the probability mass is in 1980 in the interval 0–4. This interval contains several MSA such as Michigan City-La Porte (IN), Chicago-Joliet-Naperville (IL), Rockford (IL), Racine (WI) and South Bend-Mishawaka (IN). The convergence process was also spurred by the relative improvement of MSA such as Austin-Round Rock-San Marcos (TX), Huntsville, (AL), Springfield (MA), San Jose-Sunnyvale-Santa Clara (CA) and Trenton-Ewing (NJ), which substantially improved their relative positions.

Finally, the analysis of the last period 1980–2010, featured by great divergence, reveals that once more only a few MSA are responsible for the general trend. In these years, a great part of the MSA remained stable at their relative position. For a reduced group we observe a remarkable relative improvement, including Seattle-Tacoma-Bellevue (WA), Santa Cruz-Watsonville (CA), Burlington (NC), Rochester (MN) or San Francisco-Oakland-Fremont (CA), which were all between 0 and 5 times the mean in 1980 and in 2010 are between 5 and 10 times. In the stacked densities, these correspond to the bulk of observations in the middle part of the plot. Especially outstanding is the case of San Jose-Sunnyvale-Santa Clara (CA), which rocketed from below 5 times the mean to almost 20 times in this last period, becoming the main responsible of the divergence process. In is also interesting the case of Trenton-Ewing (NJ), which performed pretty well in relative terms in the previous period but in this one witnesses a relative worsening, although it is still clearly above the mean.

4.3 Portfolio analysis

This section analyses the diversity of the patents portfolio. Diversity is measured by means of the Shannon diversity index, widely used in other scientific disciplines such as ecology to study the biodiversity of ecosystems. To analyse portfolios, we consider 36 technological subcategories, which are defined at the patent level. As each patent can be attributed to several subcategories we selected the main one. In our case, patents are considered as individuals and technological subcategories as species. The index simultaneously accounts for the variety of subcategories in each MSA and how evenly distributed are patents among those subcategories. Typical values are generally between 1.5 and 3.5, higher scores representing higher diversity. It is rarely greater than 4.

The top left panel in Figure 5 displays the results for all the sample. Considering the entire period, diversity shows a crescent patter over time. The greatest improvements are seen in the periods of convergence in terms of patents per capita (1860–1910; 1930–1980), whereas declines are observed in periods of divergence (1910–1930; 1980–2010). Panel b) differentiates diversity by MSA above and below the average innovation, taking also into account if they show a convergence or divergence trend. In general terms, we observe a clear tendency of convergence in terms of diversity. However, in 1860 there were remarkable differences between those above the mean (high diversity) and those below the mean (low diversity). One may think that more innovation necessarily leads to higher diversity. Correlations between innovation intensity and diversity are, however, rather low, and are in all decades below 20%. Then, we can argue that the processes of convergence in terms of patents per capita were accompanied by convergence in terms of diversity. Interestingly, in the latest period MSA below mean continued the diversification process, whereas those above mean show a tendency sift (concentration).

We also report results for geographical areas in panels c) and d) (regions and divisions). Patterns are similar to those found for the general case. In particular, MSA from the West and South show lower diversity. These mostly hold low innovation levels, the opposite from MSA from the Midwest and North West. Considering divisions, the greatest diversity takes place in Middle Atlantic and New England (in general in all divisions from the East coast). Contrarily, the Mountains division has the lowest diversity along most of the period, although it catched up during the 1960s and has the highest level in 2010. Finally, Figure 6 shows the diversity patterns for some particular MSA.

5 Concluding remarks

We have analysed the long run patterns of technological convergence and divergence of 302 MSA in the USA. We used the HistPat and UPSTO databases, which provided information on patents for the period 1860-2010. Our main results are:

Four differentiated periods of technological convergence and divergence are found. These periods roughly correspond with technological revolutions.

We find that whereas the performance in patenting terms is relatively similar across most MSA (innovation below the mean), there are some clear outperformers that drive the main tendencies.

The portfolio analysis suggests that periods of convergence are characterised by an increase in the diversity, whereas in periods of divergence the innovation tends to concentrate in fewer technological subcategories.

Results by regions and divisions indicate a convergence process in terms of diversity, the gap being closed mostly in the convergence periods.

Despite the correlation between patents per capita and diversity is rather low (around 20%), MSA that innovate above the average show more diversity, especially before 1980. After that decade, convergence in terms of diversity is observed.
More in particular, for the most recent periods the outperformers are specialising (lower diversity) while those below the average are diversifying more.

References

Ezcurra, Roberto and Gil, Carlos and Pascual, Pedro. Regional welfare disparities: the case of the European Union. *Applied Economics*, 37(12):1423–1437, 2005.

Peiró-Palomino, Jesús. European regional convergence revisited: the role of intangible assets. *The Annals of Regional Science*, 57:165-194, 2016.

Poletti Laurini, Márcio and Valls Pereira, Pedro L. Conditional stochastic kernel estimation by nonparametric methods. *Economics Letters*, 105(3):234–238, 2009.

Quah, Danny T. Empirics for growth and distribution: stratification, polarization, and convergence clubs. *Journal of Economic Growth*, 2:27–59, 1997.

Quah, Danny T. Galton's fallacy and tests of the convergence hypothesis. *Scandinavian Journal of Economics*, 95:427–443, 1993.

Quah, Danny T. Twin peaks: growth and convergence in models of distributional dynamics. *Economic Journal*, 106:1045–1055, 1996.

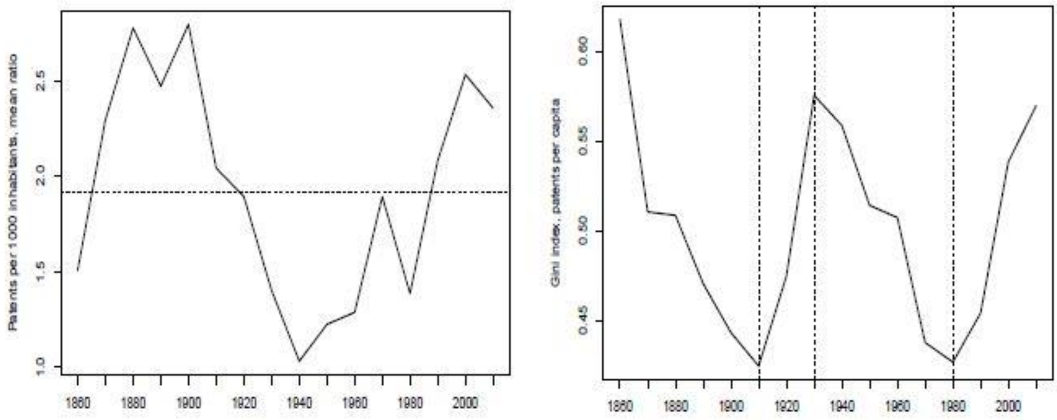
Murray Rosenblatt. Conditional probability density and regression estimators. In P. R. Krishnaiah, editors, *Multivariate Analysis*, pages 25–31. Academic Press, New York, 1969.

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

Table 1: Descriptive statistics, patents per capita

	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000	2010
N	302	302	302	302	302	302	302	302	302	302	302	302	302	302	302	302
Mean	1.51	2.30	2.78	2.47	2.80	2.04	1.89	1.40	1.03	1.22	1.29	1.60	1.39	2.08	2.53	2.36
S.d.	1.95	2.36	2.89	2.26	2.38	1.62	1.75	1.72	1.18	1.28	1.29	1.40	1.18	2.02	3.61	3.67
Min.	0.00	0.00	0.00	0.00	0.00	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.05	0.05	0.06	0.07
Max.	15.38	16.38	18.73	13.20	14.80	7.58	10.13	11.58	6.78	7.14	7.50	8.86	9.20	16.45	37.03	44.45
Kurtosis	9.48	5.89	6.90	3.03	3.24	1.15	3.31	7.35	3.38	4.01	2.97	5.69	7.76	12.09	38.60	59.63
Skewness	2.42	2.03	2.21	1.60	1.55	1.22	1.65	2.35	1.84	1.90	1.67	1.98	2.13	2.77	5.19	6.19
Below mean (%)	62.91	62.25	64.24	62.25	59.60	58.61	61.92	68.21	67.55	68.54	67.55	61.92	62.58	63.58	70.20	70.53
Above 2 times the mean (%)	12.58	9.93	12.58	11.92	10.60	11.59	12.25	11.26	13.58	11.59	14.24	11.26	9.27	9.60	7.95	8.28
Above 4 times the mean (%)	3.31	1.99	1.32	1.32	0.99	0.00	1.32	2.32	2.65	2.98	1.66	1.32	0.66	0.99	1.66	1.99
Above 6 times the mean (%)	0.33	0.33	0.66	0.00	0.00	0.00	0.00	0.99	0.33	0.00	0.00	0.00	0.33	0.66	1.32	1.66

Figure 1: Evolution of patents per capita



(a) Patents per capita

(b) Patents per capita, Gini index

Figure 2: Density functions for selected decades, patents per capita

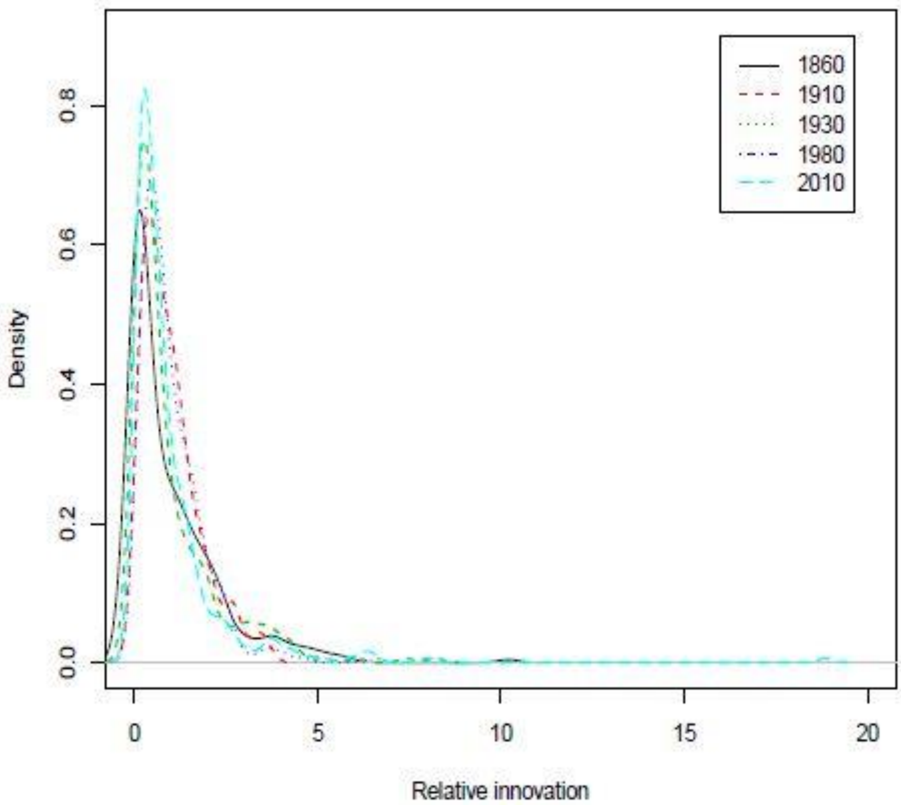


Figure 3: Conditional density estimation, main periods

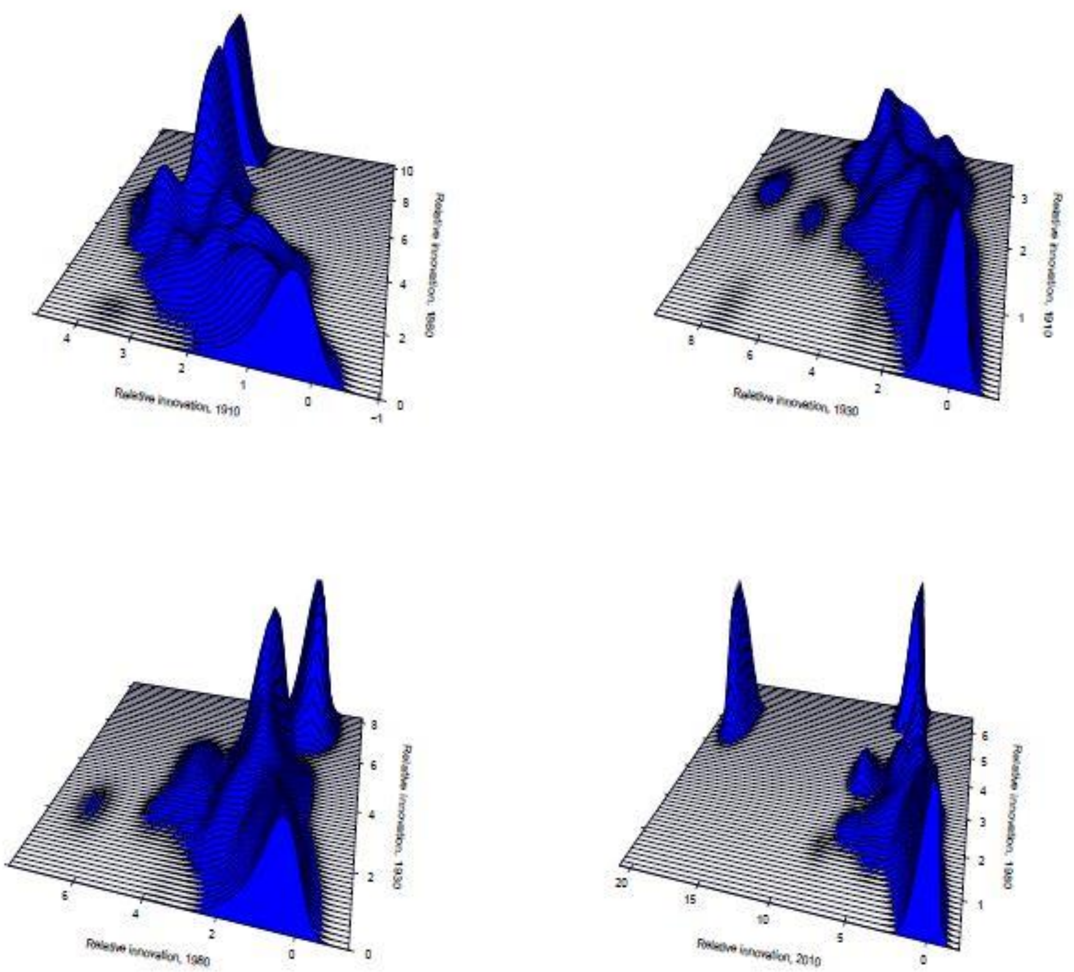


Figure 4: Intradistribution mobility, main periods

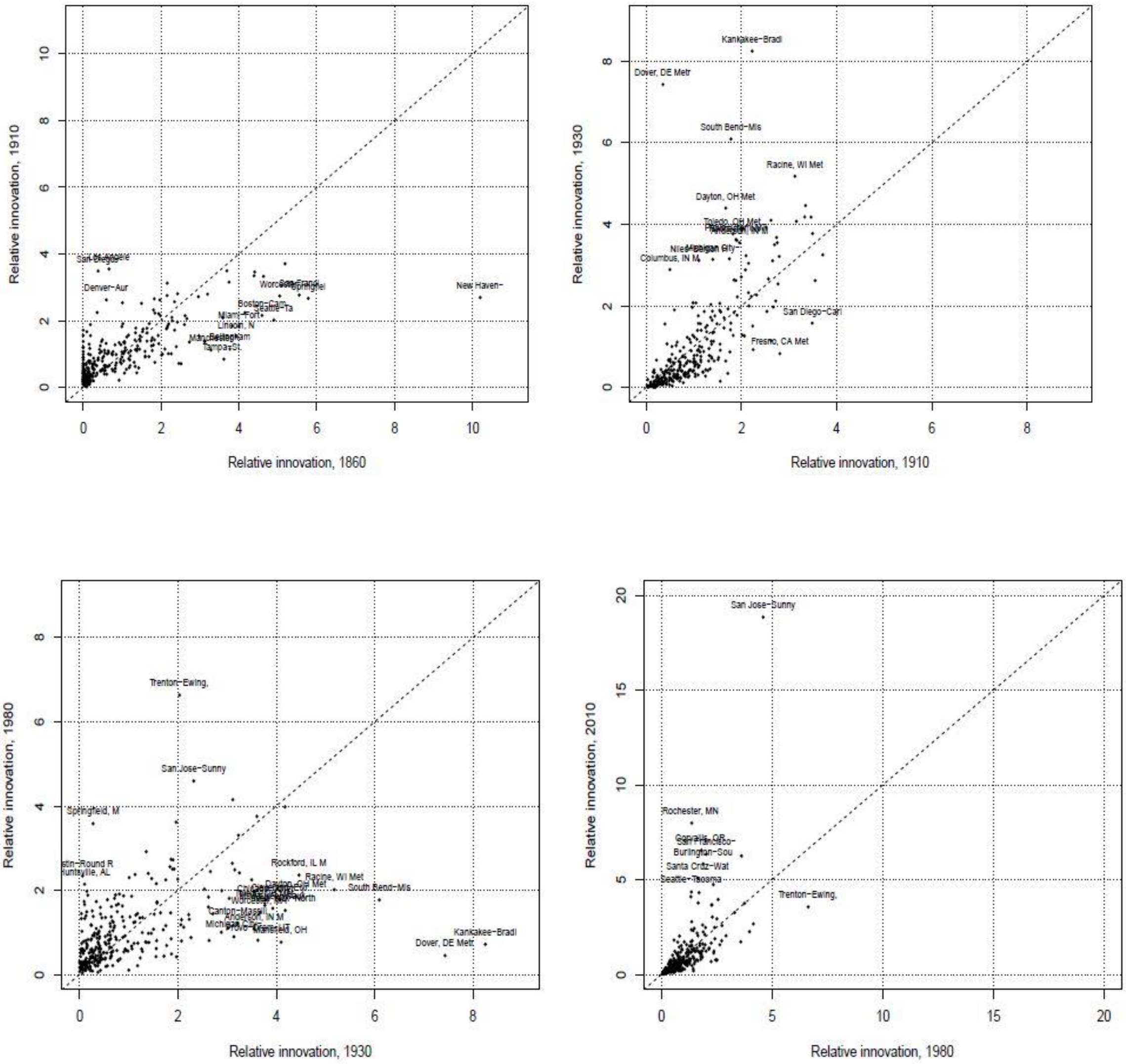
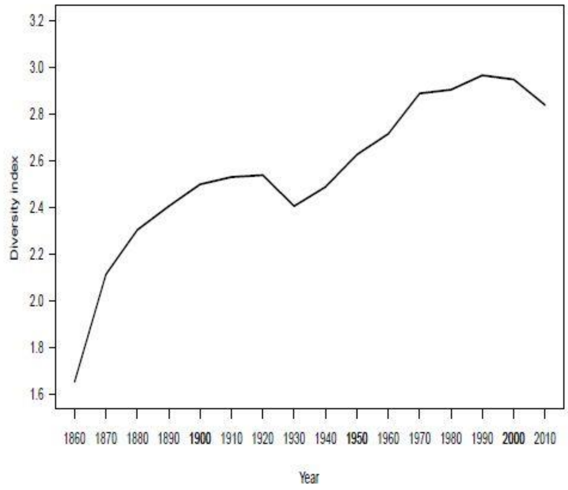
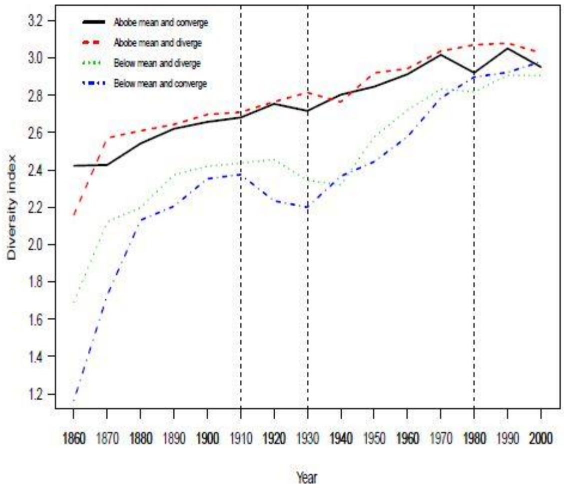


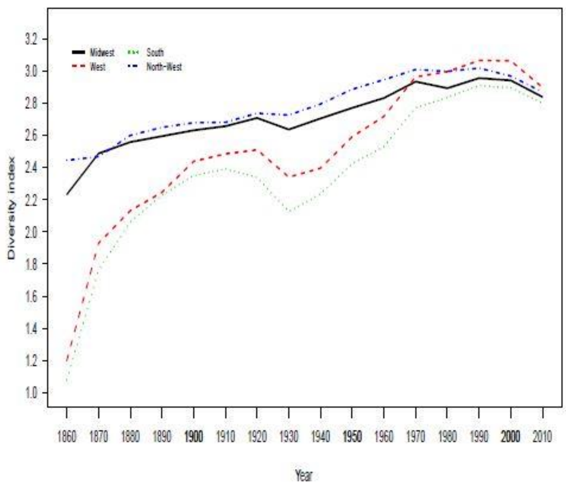
Figure 5: Portfolio analysis, Shannon diversity index



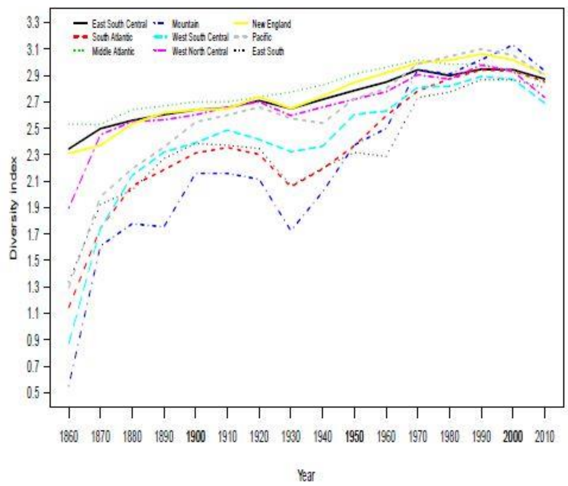
(a) All sample



(b) Converging and diverging MSA

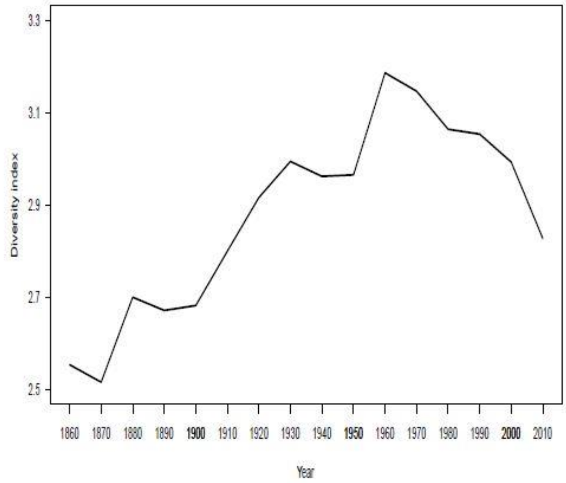


(c) Regions

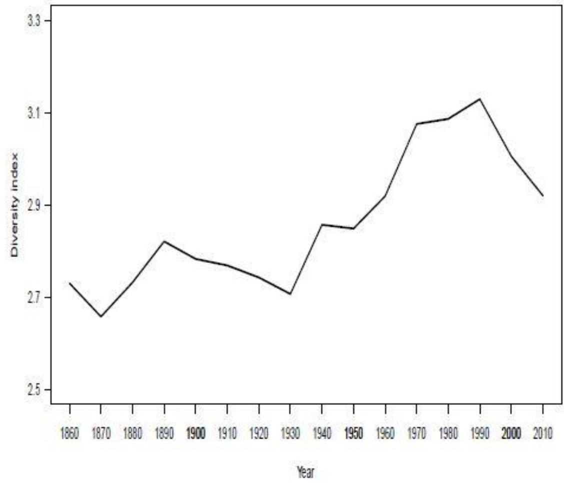


(d) Divisions

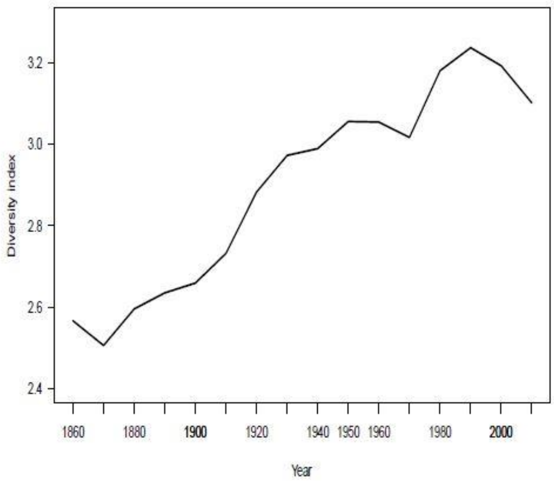
Figure 6: Portfolio analysis, Shannon diversity index for selected cases



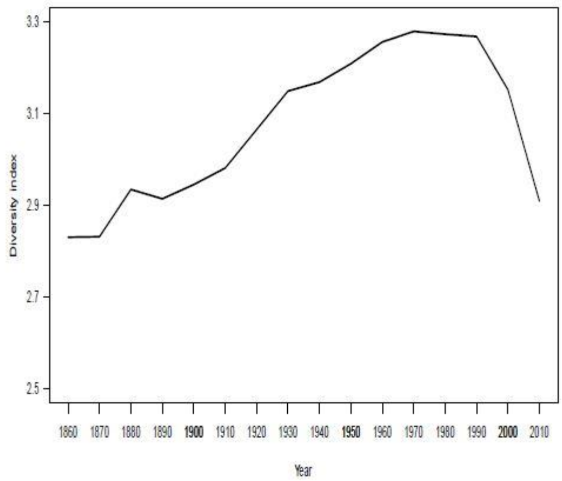
(a) San Jose



(b) Detroit



(c) New Heaven



(d) New York