INVESTIGACIONES ECONOMICAS. Volumen XX (2) - Mayo 1996, pp. 263-269

PROVINCIAL INCOME DISTRIBUTION DYNAMICS: SPAIN 1967-1991

Javier GARDEAZABAL*

Universidad del País Vasco

In this paper I analyze the dynamic evolution of Spanish provincial income distribution. I apply to Spanish provincial data Quah's idea of how to model the dynamics of the distribution of income among countries using a Markov process. The results show that, in contrast to Quah's finding that the distribution of income among countries tends to pile up in the upper and lower income classes, Spanish provinces tend towards an equilibrium distribution with more probability mass at the center. The 1991 income distribution among Spanish provinces is very similar to the equilibrium distribution. Therefore, Spanish provincial income distribution seems to have converged towards the equilibrium distribution.

1. Introduction

Quah (1993a, 1993b) has proposed a way of studyng the dynamic evolution of countries' *per capita* GDP over time[1]. He focuses on the dynamics of the distribution of income among countries and estimates Markov chains for the distribution of real *per capita* GDP among countries (Quah, 1993a) and the distribution of per worker output among countries (Quah, 1993b). His findings suggest that the cross-country income distribution evolves over time exhibiting a tendency for the rich remaining rich and the poor, poor, and for the gap to widen. This paper is a straight application of Quah's (1993a) paper to Spanish provincial data. The idea is to determine whether Quah's findings hold for Spain at provincial level. In other words, I am interested in answering the following question: Do poor Spanish provinces tend to catch up with richer ones and *viceversa?* Convergence among Spanish provinces will manifest in the distribution tending to an equilibrium distribution with its probability mass concentrated on a point or small region, i.e. a unimodal distribution.

* I would like to express my thanks to Eva Ferreira, Marta Regúlez and the referees for their comments, and to José Manuel Merodio for his able research assistance. I would also like to thank BBV and INE for providing me with the data and Gobierno Vasco (PI9442) for financial support.

[1] Barro and Sala-i-Martin (1992), Mankiw, Romer and Weil (1992) and many others have taken a different approach in modelling the evolution of contries' income along time, the so called β-convergence and σ-convergence. In few words, these concepts of convergence focus on the dynamic evolution of the mean and standard error of the income distribution. Quah's approach focuses on the dynamic evolution of the entire income distribution.

Divergence, on the contrary, will manifest in the distribution tending to an equilibrium distritution with its probability mass concentrated on two (or more) points or regions, i.e. a bimodal (or multimodal) distribution.

The analysis is a mere description of how Spanish provincial income distribution evolved during the 24 years from 1967 to 1991. As this is a descriptive analysis, no inference about causal patterns can be made. Therefore, no attempt is made to identify what the forces driving this dynamic process are. In recent work, however, Quah (1996) includes some variables to explain the dynamics of the cross-country distribution of per worker output. However, in this paper I do not follow this latter approach.

Following Quah (1993a, 1993b), I postulate a first order Markov process to describe the dynamic evolution of the distribution of personal income among Spanish provinces. Section 2 of this paper briefly summarizes the statiscal model. I estimate the Markov model using Spanish data on provincial personal income. Section 3 reports the results. The personal income distribution of Spanish provinces tends to pile up at the center of the distribution as time elapses.

2. The statiscal model

Let $x_{i,t}$ be the logarithm of a measure of per capita income of province $i = 1, \ldots, I$, at time $t = 1, \ldots, T$, $\bar{x}_t$ its time mean, i.e. $\bar{x}_t = (1/I)\, \Sigma_{i=1}\, x_{i,t}$, and $y_{i,t} = x_{i,t} - \bar{x}_t$. Partitioning the range of variation of $y_{i,t}$ on a set on n disjoint intervals $(-\infty, k_1]$, $(k_1, k_2]$, $\ldots$, (k_{n-1}, ∞), defines a set of n states. A Markov model can thus be defined. The statistical model that I use in this paper is well summarized in Amemiya (1985). Construct a set of n auxiliary variables

$$z_j\,(y_{i,t}) = \begin{cases} 1, \text{ if } y_{i,t} \in (k_{j-1}, k_j] \\ 0, \text{ otherwise} \end{cases}$$

$j = 1, \ldots, n$, and stack them in a $(n \times 1)$ vector $z_{i,t}$ whose j-th element is $z_j\,(y_{i,t})$. Define $p^j_{i,t}$ as the probability that province i is in state j at time t and $p_{i,t}$, as the $(n \times 1)$ vector of probabilities the j-th element of which is $p^j_{i,t}$.

Transition probabilities are denoted by $p^{j,k}_{i,t}$ whose meaning is the probability that province i moves to state k in period t given that it was in state j at time $t-1$. Let $P_{i,t}$ be the $(n \times n)$ transition matrix whose (j, k) element is $p^{j,k}_{i,t}$. From the previous definitions it follows that

$$p_{i,t} = P'_{i,t}\, p_{i,t-1}. \qquad [1]$$

Suppose now that the Markov process is homogeneous, i. e. $P_{i,t} = P_t$ and $p_{i,t} = p_t$ for all i and also that the Markov process is stationary, i.e. $P_t = \text{P}$ for all t, then [1] can be written as

$$p_t = P'\, p_{t-1} \qquad [2]$$

By repeated substitutions in [2] one gets $p_s = (P')^{s-1} p_1$. The equilibrium probabilities are defined as $\lim_{s \to \infty} (P')^{s-1} p_1$, if this limit exists.

Now I turn to the estimation of the vectors of probabilities p_t and the transition matrix P. For a given t, the maximum likelihood estimator of p_t^j is

$$\hat{p}_t^j = \frac{n_t^j}{\sum_{j=1}^{n} n_t^j} \qquad [3]$$

where $n_t^j = \Sigma_{i=1}^{I} z_j(y_{i,t})$ and $\Sigma_{j=1}^{n} n_t^j = I$. Maximum likelihood estimates of the transition probabilities $p^{j,k}$ of a first order, homogeneous and stationary Markov process are given by

$$\hat{p}^{j,k} = \frac{n^{j,k}}{\sum_k n^{j,k}} \qquad [4]$$

where $n^{j,k} = \Sigma_{i=1}^{I} \Sigma_{t=2}^{T} z_k(y_{i,t})\, z_j(y_{i,t-1})$; see Anderson and Goodman (1957).

3. Empirical evidence

In this paper I use Spanish provincial data on personal income at market prices and population figures from BBV (1995), and provincial data on the consumer price index from INE[2]. The data set consists of 13 estimates of personal income, population and consumer price index for 50 Spanish provinces from 1967 to 1991 sampled every other year[3].

The variable I analyze in this paper, $x_{i,t}$, is the logarithm of the ratio of *per capita* personal income to the consumer price index for each province at each time period. For each time period, I computed a non-parametric estimate of the probability density function of the distribution of *per capita* real personal income among Spanish provinces[4]. Four of those estimates are plotted in Figure 1. Two stylized facts can be detected visually. First, the distribution moves rightward and, second, it changes in shape as time

[2] Personal income (Renta Familiar Neta Disponible) is defined by Servicio de Estudios BBV (1995) as the flow of income that accrues to households and non-profit organizations (excluding the public sector) in a given time period once indirect taxes, social security payments and depreciation allowances have been deducted. That is, the amount that can be either consumed or saved. The Instituto Nacional de Estadística elaborates consumer price indexes at provincial level based on data from the capital city of every province, which, with very few exceptions, is the largest city in the province.

[3] Personal income figures for Ceuta and Melilla are only available from 1983 to 1991, so they were excluded from the analysis.

[4] I used a Gaussian kernel with a bandwidth parameter equal to 0.9 $\hat{\sigma}_t$ n, where $\hat{\sigma}_t$ is the standard deviation of the cross-province distribution of *per capita* real personal income at time t, and $n = 50$ is the number of provinces.

elapses. The first fact is due to the growth of provincial income during the period analyzed. The second fact reflects the process of convergence among Spanish provinces. In this paper I am interested in studying the process of convergence of provincial income distribution rather than analyzing provincial income growth. In order to eliminate the pure growth effect and concentrate on the convergence effect, I subtracted the time mean (over provinces) from each observation. The transformed data, $y_{i,t} = x_{i,t} - \bar{x}_t$, are analyzed hereinafter. Using the estimated standard error of $y_{i,t}$, $\hat{\sigma} = 0.1846$, I computed the values $k_1 = 0.9\,\hat{\sigma}$, $k_2 = -0.3\,\hat{\sigma}$, $k_3 = 0.2\,\hat{\sigma}$ and $k_4 = 0.9\,\hat{\sigma}$, which define the intervals $(-\infty, k_1]$, $(k_1, k_2]$, $(k_2, k_3]$, $(k_3, k_4]$, (k_4, ∞), reported in the first row of Tables 1 and 2 as *low, medium-low, medium, medium-high* and *high.* This choice of intervals is arbitrary, the numbers k_i were selected so as to divide the range of variation of $y_{i,t}$ into classes of roughly equal sizes. The last row of Table 1 reports the number of observations per interval in the whole sample. The sum of this row is 650, the total number of observations IT ($I = 50$, $T = 13$). Therefore, over the entire sample, 126 observations fell in the *low* income bracket, 138 fell in the *medium-low* income interval and so on.

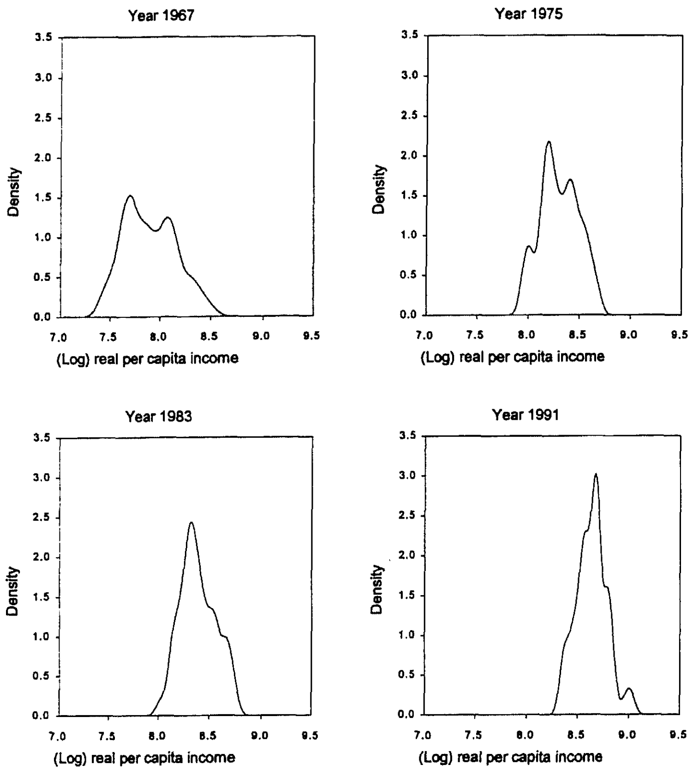

Figure 1
Density estimates

Table 1 reports the empirical distribution of $y_{i,t}$ among Spanish provinces for the years 1967-1991. Each row of Table 1 is an estimate of the distribution at one point in time. For a given row, each cell is computed using [3]. Both from Table 1 and Figure 1, there appears to be a tendency for Spanish provinces to pile up at the center of the distribution as we move forward in time. The top and bottom classes decrease in size. There also seem to be two different periods regarding the evolution of real *per capita* personal income distribution among Spanish provinces. During the 1967-1979 period the distribution changes from a bimodal distribution to a unimodal distribution, the probability mass tends to pile up at the center. The period 1979-1991, however, exhibits a less pronounced change in the distribution.

TABLE 1
Provincial personal income distribution

Year	Low	Medium-Low	Medium	Medium-High	High
1967	0.34	0.10	0.12	0.14	0.30
1969	0.30	0.16	0.12	0.10	0.32
1971	0.28	0.20	0.10	0.16	0.26
1973	0.26	0.24	0.06	0.20	0.24
1975	0.16	0.28	0.14	0.22	0.20
1977	0.20	0.18	0.18	0.26	0.18
1979	0.16	0.18	0.28	0.22	0.16
1981	0.18	0.22	0.24	0.22	0.14
1983	0.16	0.26	0.18	0.22	0.18
1985	0.12	0.26	0.26	0.26	0.10
1987	0.12	0.24	0.26	0.26	0.12
1989	0.10	0.22	0.28	0.32	0.08
1991	0.14	0.22	0.26	0.24	0.14
Observations per interval	126	138	124	141	121

In Figure 2 I plot the value of $y_{i,t}$ against that of $y_{i,t-1}$ for $i = 1, \ldots, 50$, $t = 2, \ldots, 13$. The plots show that there is some mobility between income classes. However, unless one could indentify each province, there is no way to determine whether the interclass mobility leads to an asymptotic distribution with more probability mass at the tails or at the center. In order to shed some light on this, I estimated the matrix of transition probabilities between classes.

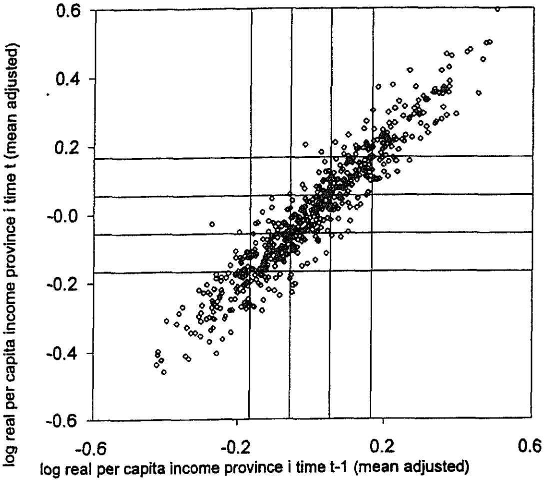

Figure 2
Mobility between classes

Table 2 presents the estimate of the transition matrix. A generic entry (*j*, *k*) of this array can be interpreted as the probability of moving from state *j* to state *k* and it is computed using [4]. Thus, a province in the *low* income interval at time *t*, had a 0.7731 probability of staying in the same income interval at time *t* + 1, a 0.1933 probability of moving to the *medium-low* income interval and so on. Provinces in the lowest and highest income classes could either stay in their classes or move towards the contiguous class. Provinces in the middle classes, however, could move in either direction. For provinces in the *medium-low* income bracket the probability of reaching a higher income class is larger than the probability of becoming poorer. However, for provinces in the *medium-high* income class, the probability of moving downwards in the distribution is higher than the probability of becoming richer. There is a tendency for provinces to move towards the middle income classes rather than to the tails of the distribution.

TABLE 2
Transition Matrix

Intervals	Low	Medium-Low	Medium	Medium-High	High	Equilibrium probabilities
Low	0.7731	0.1933	0.0336	0.0000	0.0000	0.1433
Mediun-Low	0.1181	0.6772	0.1811	0.0236	0.0000	0.2400
Medium	0.0180	0.1982	0.5858	0.1982	0.0000	0.2320
Medium-High	0.0000	0.0155	0.1860	0.6822	0.1163	0.2440
High	0.0000	0.0000	0.0175	0.1842	0.7982	0.1406

Iterating equation [2] I computed the equilibrium probabilities, reported in the last column of Table 2. These equilibrium probabilities confirm what we inferred from Table 1 and Figure 1: the asymptotic distribution tends to pile up a the center. In other words, there seems to be convergence in real *per capita* personal income among Spanish provinces. It is very interesting to see how the asymptotic distribution given by the equilibrium probabilities is similar to the actual distribution of real *per capita* personal income among Spanish provinces in 1991, and how it differs from what it was in 1967. The distribution of real *per capita* personal income among Spanish provinces seems to have converged to its equilibrium distribution.

Referencias

Amemiya, T. (1985): *Advanced Econometrics,* Harvard University Press, Cambridge, Massachusetts.

Anderson, T. W. and Goodman, L. A.(1957): «Statistical inference about Markov chains», *Annals of Mathematical Statistics* 28, pp. 89-110.

Servicio de Estudios del Banco Bilbao Vizcaya (1995): *Renta Nacional de España* 91, Banco Bilbao Vizcaya, Bilbao.

Barro, R. J. and X. Sala-i-Martín (1992): «Convergence», *Journal of Political Economy* 100, nº 2, pp. 223-251.

Mankiw, N. G., Romer, D. and Weil, D. N. (1992): «A contribution to the empirics of economic growth», *Quarterly Journal of Economics* 107, pp. 407-437.

Quah, D. (1993a): «Empirical cross-section dynamics in economic growth», *European Economic Review* 37, pp. 426-434.

Quah, D. (1993b): «Galton's fallacy and tests of the convergence hypothesis», *Scandinavian Journal of Economics* 95, pp. 427-443.

Quah, D. (1996): «Convergence empirics across economies with (some) capital mobility», *Journal of Economic Growth* 1, pp. 95-124.

Resumen

En este artículo analizo la evolución dinámica de la distribución de la renta entre las provincias españolas. Aplico a los datos de las provincias españolas el modelo de la dinámica de la distribución de la renta entre países de Quah. Los resultados muestran que, en contraste con los resultados obtenidos por Quah en los que la distribución de la renta entre países tiende a concentrarse en los intervalos de ingreso bajo y alto, las provincias españolas tienden hacia una distribución de equilibrio con más masa probabilística en el centro de la distribución. La distribución de la renta de 1991 es muy similar a la distribución de equilibrio. En consecuencia, parece ser que la distribución de la renta entre las provincias españolas ha convergido a la distribución de equilibrio.

Recepción del original, enero de 1996
Versión final, mayo de 1996