Space-time analysis of GDP disparities among European regions: A Markov chains approach

Julie Le Gallo*

March 2001

University of Burgundy
LATEC UMR-CNRS 5118
Pôle d'Economie et de Gestion
B.P. 26611, 21066 Dijon Cedex
FRANCE

E-mail: JLeGallo@aol .com

* I would like to thank C. Baumont, L. Bertinelli, C. Ertur and M.-C. Pichery and for their helpful comments. Any errors or omissions remain my responsibility.

Space-time analysis of GDP disparities among European regions: A Markov chains approach

Abstract

The purpose of this paper is to study the evolution of the disparities between 138 European regions over the 1980-1995 period. We characterize the regional per capita GDP cross-sectional distribution by means of nonparametric estimations of density functions and we model the growth process as a first-order stationary Markov chain. Spatial effects are then introduced within the Markov chain framework using regional conditioning and spatial Markov chains. The results of the analysis indicate the persistence of regional disparities, a progressive bias toward a poverty trap and the importance of geography to explain growth and convergence processes.

Keywords: convergence, regional disparities, spatial autocorrelation, spatial Markov chains

JEL Classification: C14, O52, R11, R15

Introduction

Numerous recent studies[1] have reported the persistence of per capita Gross Domestic Product (GDP) or income disparities among European regions, despite the high degree of openness between these regions and in contradiction with the predictions of the neoclassical growth model. From an empirical point of view, the analysis of economic disparities is often linked to two concepts of convergence, called, respectively b- and s-convergence (Barro and Sala-I-Martin, 1995).

Empirical evidence on b-convergence has usually been investigated by regressing growth rates of GDP on initial levels, sometimes after other variables maintaining constant the steady-state of each region have been added (conditional b-convergence). A negative regression coefficient is interpreted as an indication of b-convergence, which implies that poor regions tend to grow faster than rich regions, so that the poor regions catch up in the long run the level of per capita GDP of the rich regions. s-convergence refers to a reduction of the dispersion within the per capita GDP cross-sectional distribution over time. However, both coefficients raise several problems. In particular, Friedman (1992) and Quah (1993b) show that an increase of the dispersion (i.e. no s-convergence) is consistent with a negative b-convergence regression coefficient. Furthermore, Quah (1993a,b) argues that dispersion indicators do not provide any information on the behavior of the entire regional per capita GDP distribution.

In addition, a common problem to all these methods concerns the role of space. At the regional scale, spatial effects and particularly spatial autocorrelation cannot be neglected for the analysis of convergence processes. Indeed, several factors, like trade between regions, technology

[1] Empirical studies are numerous. See among others Baumol (1986), Barro (1991), Barro and Sala-i-Martin (1991, 1995), Armstrong and Vickerman (1995), Sala-i-Martin (1996), Beine and Docquier (2000).

and knowledge diffusion and more generally regional externalities and spillovers, lead to geographically dependent regions. Because of spatial interactions between regions, the geographical location plays an important role for explaining the economic performances of the regions. Despite their importance, the role of spatial effects in convergence processes has only recently been examined using appropriate spatial statistics and spatial econometric methods (Armstrong, 1995; Moreno and Trehan, 1997; López-Bazo et al. 1999; Fingleton, 1999; Rey and Montouri, 1999; Le Gallo and Ertur, 2000; Baumont et al., 2001).

This paper is related to the work of Quah (1993a, b) and deals with an alternative form of convergence, which is then measured from the evolution of the shape of the per capita GDP cross-sectional distribution and from the changes of the regions' relative positions inside this distribution[2]. Based on a sample of 138 European regions over the 1980-1995 period, the paper is run in two steps.

First, we characterize the evolution of the disparities between the European regions by examining the per capita GDP cross-sectional distribution over the 1980-1995 period. In that purpose, we use the tools developed by Quah (1993a, b, 1996a-c), i.e. non-parametric estimation of density functions and modeling of the growth process as a stationary first-order Markov chain. Second, the spatial dimension is explicitly considered within the Markov chain framework using regional conditioning (Quah, 1996b) and spatial Markov chains (Rey, 2001). These tools allow studying how the economic performances of a region can be explained by its geographical

[2] This kind of study may lead to the identification of convergence clubs and has been applied to various groups of regions or countries (Bianchi, 1997; Desdoigts, 1999; Paap and Van Dijk, 1998; or Johnson, 2000).

environment, the extent to which this environment influences the regions' relative position inside the GDP cross-sectional distribution and the role of space in the constitution of convergence clubs.

1. The evolution of the per capita GDP distribution

The analysis of the evolution of the regional GDP distribution is carried out on the *Europe-relative GDP distribution* over the 1980-1995 period. The Europe-relative GDP is defined as the ratio of the regional GDP to the European wide average GDP. It is preferable to work on relative GDP, opposed to absolute GDP, so that co-movements due to the European wide business cycle and trends in the average regional GDP are removed.

The analysis of the Europe-relative regional GDP distribution is based on two main lines. First, non-parametric density estimation methods allow studying the external shape of the GDP distribution for each year, as well as the changes in this shape during the period. Second, the temporal dynamics within the GDP distribution is examined with the estimation of probability transition matrices (or Markov chains) and the associated long-run distributions.

The data are extracted from the EUROSTAT-REGIO databank[3]. The sample includes 138 regions for 11 countries (Denmark, Luxembourg and United Kingdom in NUTS1 level and Belgium, Spain, France, Germany, Greece, Italy, Netherlands and Portugal in NUTS2 level) over the 1980-1995 period[4].

11. The evolution of the shape of the per capita GDP distribution

[3] Series E2GDP measured in Ecu_hab units.
[4] We exclude Groningen in the Netherlands from the sample due to some anomalies related to North Sea Oil revenues, which increase notably its per capita GDP. We exclude also Canary Islands and Ceuta y Mellila, which are geographically isolated. Corse, Austria, Finland, Ireland and Sweeden are excluded due to data non-availability over the 1980-1995 period in the EUROSTAT-REGIO databank. Berlin and East Germany are also excluded due to well-known historical and political reasons.

In cross-country studies, a polarization or stratification process in several "convergence clubs" is often observed (Quah, 1996c; Bianchi, 1997; Desdoigts, 1999). These convergence clubs mean an increase in the homogeneity within regions or countries of the same groups, and also an increase in the differences between groups. A first technique aimed at detecting such convergence clubs is the estimation of the density function for the regional per capita GDP distribution and the analysis of its mono- or multimodality characteristics. Quah (1993b, 1996c) and Bianchi (1997) detect bimodality for international income distribution, i.e. the existence of two convergence clubs, but Quah (1996b) obtains no evidence of it for a subsample of European regions in per capita GDP.

In order to characterize the evolution of regional GDP, we have examined the per capita GDP distribution (relative to the European average) in 1980 and the way this distribution has changed in time until 1995. Figure 1 plots two estimated density functions for regional relative GDP for the initial year 1980 and the final year 1995[5]. These density plots can be interpreted as the continuous equivalent of a histogram, in which the number of intervals has been let to infinity and then to the continuum. By definition of the data, 1 on the horizontal axis indicates the European average GDP, 2 indicates twice this average, and so on.

[Figure 1 about here]

Compared to 1980, more regions have regional GDP less than the European average or twice the European average in 1995. Moreover, besides the main mode, a second persistent mode is situated around 50% of the European average. This may reflect the existence of an important group

[5] All densities are calculated nonparametrically using a Gaussian kernel with bandwith set as proposed in Silverman (1986, section 3.4.2.).

of regions, with per capita GDP levels below the average, and which converge towards a lower GDP level than the rest of the regions. This result contradicts those in Quah (1996b) but is similar to those in López-Bazo et al. (1999). This difference may be explained by the samples used in this study and in López-Bazo et al. (1999), where all poor regions of Portugal and Greece are included. On the contrary, the sample used in Quah (1996b), does not include these poor regions (78 regions from 1980 to 1989). Let us note as well that the distance between the two modes in our case is far below that the distance detected between the two peaks in the cross-country distribution. Finally, the little peak, situated around 200% of the European average on the 1995 density plot, fluctuates over the period. Therefore it is not possible to surely identify a third mode concerning the very rich regions.

The density plots suggest a persistent polarization of European regional GDP. However the density plots alone cannot support this interpretation. It is true that there are more very rich or very poor regions in 1995 compared to 1980, but we could wonder as well what were their relative positions in previous years. In other words, these density functions do not inform if the right tail of the initial distribution (1980) contains the same regions that the right tail in the final distribution (1995). Finally, while these functions allow characterizing the evolution of the global distribution, they do not provide any information on the movements of the regions inside this distribution.

12. Markov chains

A possible way to answer these questions is to track the evolution of each region's relative GDP over time by constructing transition probability matrices or Markov chains[6]. In the case of the

[6] Presentations of Markov chains can be found in Chung (1960), Kemeny and Snell (1976).

European regions, different studies, based on different samples have been carried out: Neven and Gouyette (1995), Quah (1996a-c), Fingleton (1997, 1999), López-Bazo et al. (1999), Magrini (1999).

Denote F_t the cross-sectional distribution of regional per capita GDP at time t relative to the European average. Define a set of K different GDP classes, which provide a discrete approximation of the per capita GDP distribution.

We suppose that the frequency of the distribution follows a first-order stationary Markov process. In this case, the evolution of the regional GDP distribution is represented by a transition probability matrix, M, in which each element (i,j) indicates the probability that a region that was in state i in time period t ends up in state j the following period.

The $(K,1)$ vector F_t, indicating the frequency of the regions in each class in time t, is described by the following equation:

$$F_{t+1} = MF_t \qquad [1]$$

where M is the (K,K) transition probability matrix representing the transition between the two distributions.

If the transition probabilities are stationary, i.e. if the probabilities between two classes are time-invariant, then:

$$F_{t+s} = M^s F_t \qquad [2]$$

The transition probability matrix has a number of properties that can be exploited to study the evolution of regional income distributions.

1/ The first property is the propensity of the regions in each class to move in other classes and the average time required for a region to move between any pair of states i and j. These information are provided by the estimation of transition probabilities for our sample and by the determination of mean first time passage matrix (cf below: empirical results).

2/ The second property is the determination of the ergodic distribution (or the long-term distribution) of F_t, characterized when s tends to infinity in [2]. Such a distribution exists if the Markov chain is regular, i.e. if and only if for some N, M^N has no zero entries. In this case, the transition probability matrix converges to limiting matrix M^* of rank 1:

$$M^{T^*} = M^* \qquad [3]$$

where T^* is the number of years required to reach this steady state. The existence of an ergodic distribution, F^* is then characterized when:

$$F^* M = F^* \qquad [4]$$

Each row of M^t tends to the limit distribution as $t \rightarrow \infty$. According to [4], this limit distribution is therefore given by the eigenvector associated to the unit eigenvalue of M.

3/ Finally, the second eigenvalue (in absolute value) of M, l_2, is a measure of mobility and allows characterizing the speed with which the steady-state is approached. The half-life, which is the amount of time taken to cover half the distance from the stationary distribution, is defined as (Shorrocks, 1978):

$$dm = -\frac{\log 2}{\log\left|l_2\right|} \qquad [5]$$

This indicator ranges between infinity –when the second eigenvalue is equal to 1 and a stationary distribution does not exist- and 0- when l_2 is equal to 0 and the system has already reached its stationary equilibrium.

13. Empirical results

We distinguish between five different states: 1/ less than 65% of the European average 2/ between 65% and 95% of the European average 3/ between 95% and 110% of the European average 4/ between 110% and 125% of the European average 5/ more than 125% of the European average. As advised by Quah (1993a), followed by López-Bazo et al. (1999) or Neven and Gouyette (1995), the discretization has been chosen so that the initial classes include a similar number of individuals. Markov chains with other grid points and other number of states have also been estimated but the main results found in this paper are qualitatively the same with a different discretization.

Table 1 contains the transition probability matrix between 1980 and 1995 with the maximum likelihood estimates of the transition probabilities. The estimation of any element p_{ij}, is the total number of regions moving from class i in year t to class j in year $t + 1$ over all 15 years of transitions divided by the total sum of regions ever in i over the 15 years. For example, during the 15-year-period, there were 416 instances of a region having a GDP lower than 65% of the European average. The majority of these regions (95.9%) remained in that GDP class at the end of the year, while 4.1% moved up one class by the end of the year.

[Table 1 about here]

López-Bazo et al. (1999) and Neven and Gouyette (1995) compute Markov chains for two sub-periods due to the changes in the convergence process over time detected in other works[7]. On the contrary, in this paper, stationarity of transition probabilities is formally tested against non-stationarity (different year-to-year probabilities) for the whole period 1980-1995 and also for two subtime periods 1980-1985 and 1985-1995. The tests indicate that the null hypothesis of stationary transition probabilities cannot be rejected in both cases. Therefore, the reliability of all subsequent interpretations is strengthened and in the remainder of the paper, we will analyze the convergence process for the whole period[8].

Several comments can be made about this matrix.

1/ First, the transition probabilities on the main diagonal are relatively high. Indeed, if a region is in the i^{th} class, the probability of being in this same class the year after is at least 79,6%. Since the diagonal elements dominate, especially among the extreme classes, these results indicate that the poorest and the richest regions do not seem to modify their relative position over time. Furthermore,

[7] Subperiods (1980-1985) and (1985-1992) for López-Bazo et al. (1999), subperiods (1980-1985) and (1985-1989) for Neven and Gouyette (1995).

[8] The χ^2 statistic is (Anderson et Goodman, 1957 ; Kullback et al., 1962):

$$-2\log\left\{\prod_t \prod_j \prod_j \left[\frac{\hat{p}_{ij}}{\hat{p}_{ij}(t)}\right]^{M_{ij}(t)}\right\}$$

with $(T-1)K(K\text{-}1)$ degrees of freedom. $\hat{p}_{ij}$ is the stationary estimate; $\hat{p}_{ij}(t)$ are the year-to-year estimates; $m_{ij}(t)$ is the number of regions moving from i to j in year t, T is the total number of years and K is the number of cells in the distribution. The test has been computed for the whole period and for two subperiods. The p-value is 0,99 in the first case and 0,91 in the second case.

there is no spectacular move from year to year as strictly positive elements are only observed around the diagonal.

Movements average 8%. The second eigenvalue of the transition matrix is close to unity (0,97) and implies a half-life of 22 years. All these elements indicate a very low inter-class mobility and an important persistence of the regions within each class.

2/ In order to precise the speed with which the regions move in the distribution, we consider mean first passage time, for a process starting at time zero. If p_{jk}^{t} is the probability that a city in state j first visits t periods later the state k, then the mean first passage time mp_{jk} from j to k is:

$$mp_{jk} = \sum_{t=1}^{\infty} t p_{jk}^{t} \qquad [6]$$

Table 2 shows the matrix of mean first time passage that allows examining the issue of fluidity[9]. The diagonal transition probabilities are mean first return times, where first return means staying in the own cell for one year or first returning to that cell if a region leaves it in the first period.

If we concentrate on the elements outside the main diagonal, it seems that the transitions are relatively high. Indeed, the lower passage time is 13,7 year and the higher is 162,7 years. Globally, movements up are slower than movements down.

[Table 2 about here]

3/ The ergodic distribution can be interpreted as the long run equilibrium regional GDP distribution in the regional system. If the distribution collapses into a single class, there is convergence. Returning to the density functions, this case corresponds to an unimodal per capita

[9] See Kemeny and Snell (1976) for the computation of this matrix by means of the so-called fundamental matrix for regular Markov chains (chap.4).

GDP distribution. However, a concentration of the regions in some of the classes, i.e. a multimodal limit distribution, can be interpreted as a tendency towards stratification in different convergence clubs. Finally, a dispersion of this distribution is interpreted as divergence.

Here (table 1), the characteristics of the ergodic distribution indicate a poverty trap: the probability that a region leaves the poorer class increases relative to the initial distribution. Combined with the very weak mobility observed, we can conclude that the poorer regions will probably remain poor. On the other side of the distribution, the two big classes are less important.

Finally, if these observed tendencies remain, the per capita GDP distribution will progressively be biased towards the relative poor regions. Globally, the situation is remarkably stable and persistent: there is neither important changes for the external shapes of the distribution, nor important intra-distribution mobility.

2. Integrating the spatial dimension in Markov chains

The data used in this study are spatial data, which combine attribute information with locational information. Spatial data often have special properties, and need to be analyzed in different ways from aspatial data. However, this spatial dimension has not been taken into account in the previous analysis even though some recent papers, dealing with regional GDP patterns, have recognized the need to consider spatial effects when growth and convergence processes are analyzed (Armstrong, 1995; Fingleton, 1999; Lopez-Bazo et al.; Rey and Montouri, 1999; Le Gallo and Ertur, 2000; Baumont et al., 2001). This section is therefore devoted to the integration of the spatial dimension of the data in Markov chain analysis.

21. Geographic patterns of the transitions

Spatial data are often characterized by spatial autocorrelation, which can be defined as the coincidence of value similarity with locational similarity (Anselin, 2000). Therefore there is positive spatial autocorrelation when high or low values of a random variable tend to cluster in space and there is negative spatial autocorrelation when geographical areas tend to be surrounded by neighbors with very dissimilar values.

To illustrate the potential importance of space in the explanation of convergence patterns, we examine the extent to which the regions that have moved up or down in the distribution are geographically concentrated. In other words, we study the level of spatial autocorrelation in per capita GDP transitions[10]. Figure 2 displays the regions' upward or downward moves between 1980 and 1995.

[Figure 2 about here]

Movements up (12 regions) are located in south Germany whereas movements down mainly concern French or English regions. Let us underline that these regions correspond to spatial clusters detected by Le Gallo and Ertur (2000) having respectively high and low growth rates. This visual impression of positive spatial autocorrelation of these transitions must be confirmed by a formal spatial autocorrelation test.

[10] Rey (2001) performed a similar analysis for the United States.

Denote W^*, the spatial weight matrix of dimension (138, 138). This matrix contains the information about the relative spatial dependence between the 138 regions *i*. The elements w^*_{ii} on the diagonal are set to zero whereas the elements w^*_{ij} indicate the way region *i* is spatially connected to the region *j*. In this study, we define the elements of W^* in the following way:

$$w^*_{ij} = \begin{cases} 0 & \text{if} \quad i = j \\ d_{ij}^{-2} & \text{if} \quad d_{ij} < \overline{d} \\ 0 & \text{if} \quad d_{ij} > \overline{d} \end{cases} \qquad [7]$$

where d_{ij} is the great circle distance between the centroids of region *i* and region *j* and $\overline{d}$ is the cutoff, equal to the lower quartile of the great circle distance distribution (321 miles)[11]. The spatial weight matrix is row-standardized such that the elements in each row sum up to one: $w_{ij} = w^*_{ij} / \sum_j w^*_{ij}$. This particular weight matrix has been preferred to a simple contiguity matrix, which is not really appropriate for our sample of European regions for two reasons. First, the islands would be isolated and unconnected with the other regions[12]. Second, a contiguity indicator may imply a block-diagonal pattern if some regions do not share a common border with any other region in the sample considered (it is indeed the case of Great-Britain and Greece).

Using this distance-based weight matrix, spatial autocorrelation of upward transitions between 1980 and 1995 is formally evaluated using the following joint-count test (Cliff et Ord, 1981):

$$NN = \frac{1}{2} \sum_i \sum_j w_{ij} (d_i d_j) \qquad [8]$$

[11] This cutoff has been determined with a correlogram in order to maximise the level of spatial autocorrelation.

where w_{ij} is the element of the weight matrix, $d_i = 1$ if region i experiences an upward move in the distribution, otherwise $d_i = 0$. The NN statistic is a count of the number of joins for which two neighboring regions both experienced upward moves in the GDP distribution (the neighborhood is defined by the weight matrix). Similarly, to test for spatial autocorrelation for downward transitions, we define $d_i = 1$ if region i experiences a downward move in the distribution, otherwise $d_i = 0$.

Statistical inference is based on random permutations of the regions on the map[13]. For both upward and downward moves, the null hypothesis of spatial autocorrelation is always rejected (*p*-values are respectively 0.001 and 0.005). It is therefore unjustified to consider each region and its transitions in the different GDP classes as if the regions were geographically independent. Consequently, the spatial dimension in the analysis of regional GDP transition dynamics should explicitly be taken into account so that the role of spatial effects in growth and convergence processes can be examined.

22. Spatial conditioning

To determine the factors explaining some of the features of the density plots and of the probability transition matrix, Quah (1996b) has suggested to "condition" the per capita GDP distribution. The general idea of this approach is to study how closely the evolution of each region's GDP has followed that of some group of regions, which are expected to behave similarly. Quah (1996b) considered two kinds of references, either the neighboring regions (geographical criterion),

---

[12] Consequently, the rows and columns in the weight matrix corresponding to these observations would consist of zero values.

[13] In this approach, it is assumed that, under the null hypothesis, each observed value could have occurred at all locations with equal likelihood. But instead of using the theoretical mean and standard deviation (given by Cliff and Ord 1981), a reference distribution is empirically generated for NN, from which the mean and standard deviation are computed. In practice this is carried out by permuting the observed values over all locations and by re-computing NN for each new sample. The mean and standard deviation for NN are then the computed moments

or the regions belonging to the same country (national criterion). The results of his work suggest that physical location factors seem to matter more than do macro national factors for explaining regional GDP inequality in Europe.

Consequently, we group the regions by the geographic criterion. In this purpose, we construct now a new GDP series: *neighbor-relative per capita GDP* where each region's per capita GDP is normalized by the average per capita GDP of the neighboring regions. Denote y the vector containing the regions' per capita GDP. Since W is a standardized weight matrix, the weighted average of the neighboring regions' GDP is given by the vector Wy, which is usually called the spatial lag in the spatial statistical/econometric literature. Quah suggests that the neighbor-relative per capita GDP can be interpreted as the part unexplained by physical-location factors. Consequently, if the physical location explains everything, what is left over vanishes, or is small. If, on the other hand, physical location explains nothing, what is left over is what we begin with. Conditional density functions and conditional probability matrices can now be constructed with this new GDP relative distribution.

Figure 3 plots two density functions for neighbor-relative per capita GDP distribution, one for the initial year 1980 and the other for the final year 1995. Comparing these densities to the formerly computed Europe-relative GDPs density functions (figure 1) indicate that the second mode, which was situated at around 50% of the European average, has disappeared and that the majority of the density is symmetrically much more concentrated around the mean. The economic

---

for the reference distribution for all permutations. This test has been computed using the software Spacestat 1.90 (Anselin, 1999).

performance of the regions is consequently well explained by the neighboring regions' performances, except maybe for regions with very high per capita GDP.

[Figure 3 about here]

Consider now the conditional transition probability matrix (table 3). This matrix contains the transitions between the Europe-relative GDP distribution and the neighbor-relative GDP distribution *for a given year*. Compared to the previous Markov chain (table 1), where we established the transitions between the same distribution for two points in time, we establish in this conditional matrix the transitions between two different distributions at the same moment in time. Therefore, as pointed out by Quah (1996b), these transition probabilities do not describe transitions over time, but rather quantify the effects of conditioning.

For example, there were 450 instances of a region having a GDP below than 65% of the European average but only 22% of these same regions had a GDP below than 65% of their neighbors' average GDP for the same year. If the regional context did not matter, each region could be considered as an island independent of its neighbors: "If conditioning explained nothing (…), these transition probability matrices should be the identity matrix: the distributions are invariant and, in addition, no intra-distribution movements occurs" (Quah, 1996b). On the other hand, if regional conditioning explained all the regional GDP variations, then all the elements of the column for the interval containing 100% should be equal to 1 (third class).

Here, none of these two extreme cases is relevant. Indeed, all diagonal elements are below or equal to approximately 50%. The regional conditioning accounts therefore for a large part of the observed regional inequality. This result conforms to Quah's study (1996b) and geographic spillovers seem to be an important factor of the regional inequality dynamics in Europe.

[Table 3 about here]

The regional conditioning allows capturing the geographical dimension in regional GDP variation. However, let us note again that these transitions do not represent transitions over time, but transitions between two different GDP distributions for a given year. On the contrary, the Markov chains presented in the following section explicitly take into account space without losing the temporal dynamics of the regional GDP evolution in Europe.

23. Temporal and spatial dynamics

1/ The first way to study explicitly the role of space while keeping an information on the temporal dynamics of the transition is to estimate a transition probability matrix similar to the traditional matrix (table 1), where the GDP is not Europe-relative anymore but neighbor-relative. Table 4 reports this matrix for our sample. For example, there were 43 instances of a region having a GDP lower than 65% of its neighbors' average at the beginning of the year. The majority of these regions (83,7%) remained in the same GDP class at the end of the year (their GDP was still below 65% of their neighbors' average), while 16,3% of the regions had a GDP between 65% and 95% of their neighbors' average.

Compared with the first matrix with European conditioning, it turns out that for the same grid points, the tails of the distribution have become much smaller, for example 416 rich and 438 poor with European conditioning compared to 43 poor and 159 rich for regional conditioning. Furthermore, middle GDP classes are much more important and concentrate almost all the regions.

This result suggests that neighboring regions evolve in the same way and do not differentiate from each other. In other words, there is an important positive spatial autocorrelation phenomenon between European regions. The exceptions to this general feature are the regions that stay a lot poorer than their neighbors over the whole period, i.e. regions in the first cell (1,1) of the transition probability matrix. These regions are Vlaams Brabant from 1980 to 1989, the German region Lüneburg for all the years, the Portuguese regions Norte from 1980 to 1985, Centro and Alentejo from 87 to 89. The other exceptions are the regions that stay much richer than their neighbors over the period (cell (5,5)). These regions are mainly the capital-regions (Bruxelles, Ile de France, Lisbonne over the period and Madrid from 85) and some German regions (Oberbayern, Bremen, Hamburg, Darmstadt over the whole period).

[Table 4 about here]

At first look, regional conditioning seems therefore to point towards a more important convergence than does conditioning on the European average. However, it is worth mentioning than conditioning on the neighbor average only provide information on *local* or *intraregional* convergence, i.e. the way regions catch up with their geographical neighbors. Consequently, this method considers the role of space but doesn't allow anymore studying the regions' position in the entire cross-sectional distribution, i.e. the analysis of *interregional* convergence.

2/ The second way of simultaneously considering spatial and temporal dynamics has been proposed by Rey (2001) and applied to US data. The spatial Markov chain estimated in his study provides insights to the role of spatial clustering in the dynamics of the GDP distribution over time and furthermore addresses the issue of interregional convergence (and not intraregional convergence like in the previous matrix).

The traditional Markov matrix (table 1) is modified in such a way that the transition probabilities of a region are conditioned on the initial GDP class of its spatial lag (i.e. GDP class in 1980). This particular conditioning implies a spatial transition matrix, which is a traditional (K,K) matrix decomposed in K conditional matrices of dimension (K,K)[14]. Therefore, if we consider the k^{th} of these conditional matrices, then an element $m_{i,j|k}$ of this matrix is the probability that a region in class i at the time period t goes in j at the end of the period, given that the spatial lag was in class k in 1980.

The spatial Markov matrix allows examining the positive or negative influence of the neighbors on the transitions of a region. Indeed, the influence of spatial dependence is reflected in the differences existing between the initial transition values (not conditioned) computed in the first section (table 1), and the various conditional transition values. In our example with five classes, the first class contains poor regions, the third class contains the median GDP regions and the final class contains the rich regions. Therefore, if $m_{35} > m_{35|1}$, then median GDP regions with poor neighbors have a lower probability of moving upwards than median GDP regions on average. Conversely, if $m_{13} < m_{13|5}$, then poor regions with rich neighbors have a higher probability of moving upwards than poor regions on average. Formally, if regional context did not matter for transition probabilities, then the conditional probabilities should be equal to the initial probabilities:

$$m_{ij|1} = m_{ij|2} = ... = m_{ij|5} = m_{ij} \quad \forall i = 1,...5 \quad j = 1,...5 \qquad [9]$$

[14] Here $K = 5$.

This spatial Markov chain does not contain the same information that the conditional matrices in Quah (1996b): it gives the probability for a region to experience upward or downward moves in the distribution, conditional to the past or present movements of its neighbors and therefore it allows studying the possible correlation between the direction and probability of the transition of a region and the regional context faced by each region.

[Table 5 about here]

Table 5 reports the spatial Markov matrix for our sample of European regions. It turns out that the spatial lag of a region influences the transitions over time of this region. For example, the richest regions are negatively affected when poorer regions surround them. Indeed, the probability of moving down one class increases as the GDP level of the neighbors decreases. On average, the richest regions move down one class with a probability of 7,3% (cell (5,4) of table 1). If other rich regions (class 5) surround these regions, the probability of moving down is only 5,7% but if the neighboring regions are poorer (class 4), the probability increases to 9,2% and reaches 12,5% if the neighboring regions are in the middle class. The poorest regions are also negatively affected when they are surrounded by other poor regions. For example, on average, the probability that a poor region moves up one class is 4.1% (cell (1,2) of table 1). If these regions are surrounded by other poor regions (class 1), the probability drops to 2.3% whereas it reaches 40% if these regions are surrounded by richer regions.

For each conditional matrix, an ergodic distribution has been computed. Like the initial distributions, the long-run distributions are strongly biased. Indeed, when the economies are surrounded by richer regions, the final distribution is more and more skewed upwards: the probability of staying or remaining rich on the long run is strong. Alternatively when the economies are surrounded by poorer regions, the ergodic distribution is more and more negatively skewed: the probability of staying or becoming poor is very strong.

Finally, in order to summarize all the information contained in the spatial Markov chain, we study the relationship between the direction of a region's transition in the GDP distribution and its spatial lag by computing the probability of a particular transition (down, none or up) conditioned on the GDPs of the region's neighbors in 1980 (table 6), as suggested by Rey (2001). As expected, the regional context has a strong influence on the probability of moving downward or upwards. For example, the probability of moving down is twice as large when the regions are surrounded by poorer regions than richer regions (12,8% vs 6%). Alternatively, the probability of a region moving to a higher GDP class is 11,4% if the region is surrounded by richer regions, but it is only 2,2% if the neighbors are poorer. These probabilities indicate the long-run influence of the neighbors on a region's transitions in the GDP distribution. Alternatively, instead of conditioning on the initial spatial lag, we could condition on the spatial lag at the beginning of each year, so that the short-run influence of the regional context is captured. The probabilities of a particular transition conditioned on the GDP of the region's neighbor at the beginning of each year are reported in table 7. We can see that results are very similar[15].

All these results therefore highlight the strong spatial dimension associated to the features detected in the (aspatial) analysis of interregional convergence conducted in the first section. For example, the progressive bias towards the poverty trap mainly has a spatial explanation since poor regions are negatively influenced by being surrounded by other poor regions and since the long-run distribution is negatively skewed downwards when the neighbors are poor. Also, the relative absence of intra-distribution mobility can be explained by persistent spatial clusters of high and low GDP regions over the period. More generally, from an interregional convergence perspective, spatial Markov chains indicate that the changes of the relative position of a region in the cross-sectional distribution are highly constrained by its geographical environment.

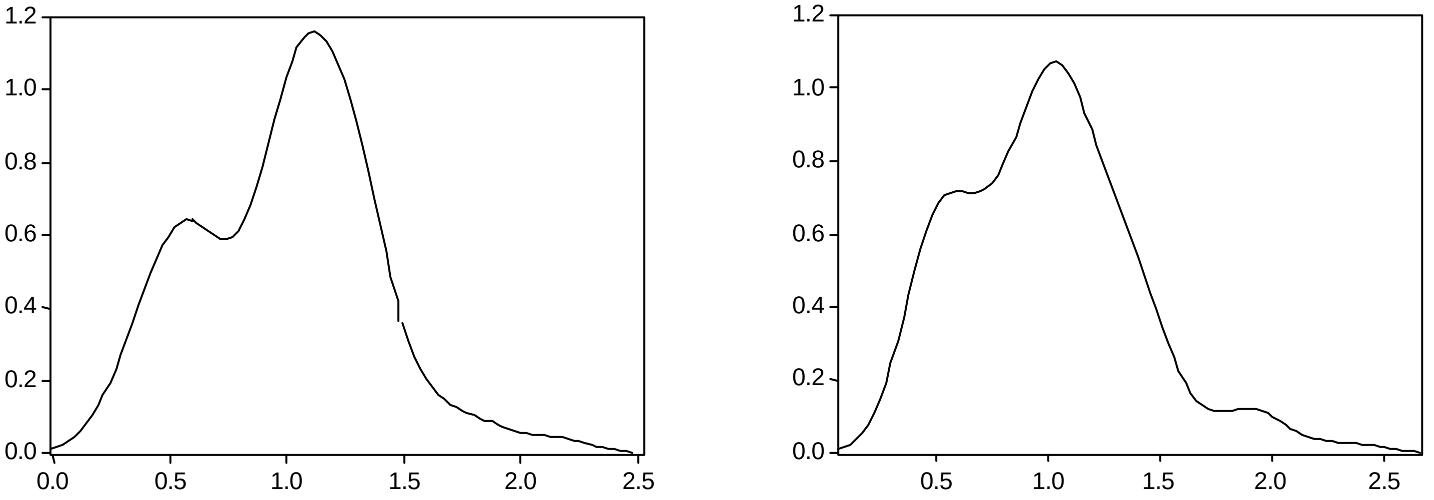
[Table 6 and table 7 about here]

[15] In both cases, a χ^2 test for the independence of direction of move and neighbor's GDP has been computed and the null hypothesis is always rejected at $p < 0.01$: the type of the transition experienced by a region is dependent with its geographical environment.

Conclusion

The aim of this paper is to analyze the evolution of regional GDP disparities and the convergence process among European regions over the 1980-1995 period. The methodology adopted here follows Quah's methodology: convergence is measured from the evolution of the shape of the per capita GDP cross-sectional distribution and from the changes of the regions' relative positions inside this distribution. In order to study the entire GDP cross-sectional distribution, nonparametric estimation of density functions are computed, the growth process is modeled as a first-order stationary Markov chain and the role of space is explicitly considered, using Quah's regional conditioning and spatial Markov chains.

The results of the analysis, based on a data set for 138 European regions over the 1980-1995 period, suggest that the process of economic growth at work in the European Union during this period has globally been characterized by the persistence of regional disparities, a relative absence of mobility of the regions in the GDP distribution as well as a progressive bias toward a poverty trap. Regional conditioning and spatial Markov chains clearly indicate that location and physical geography still matter in the European Union to explain growth and convergence processes. Indeed, intraregional convergence is very strong and, from an interregional convergence perspective, the changes of the relative position of a region in the cross-sectional distribution are highly constrained by its geographical environment.



Europe-relative GDP 1980

Europe-relative GDP 1995

Fig.1. Densities of Europe-relative regional GDP

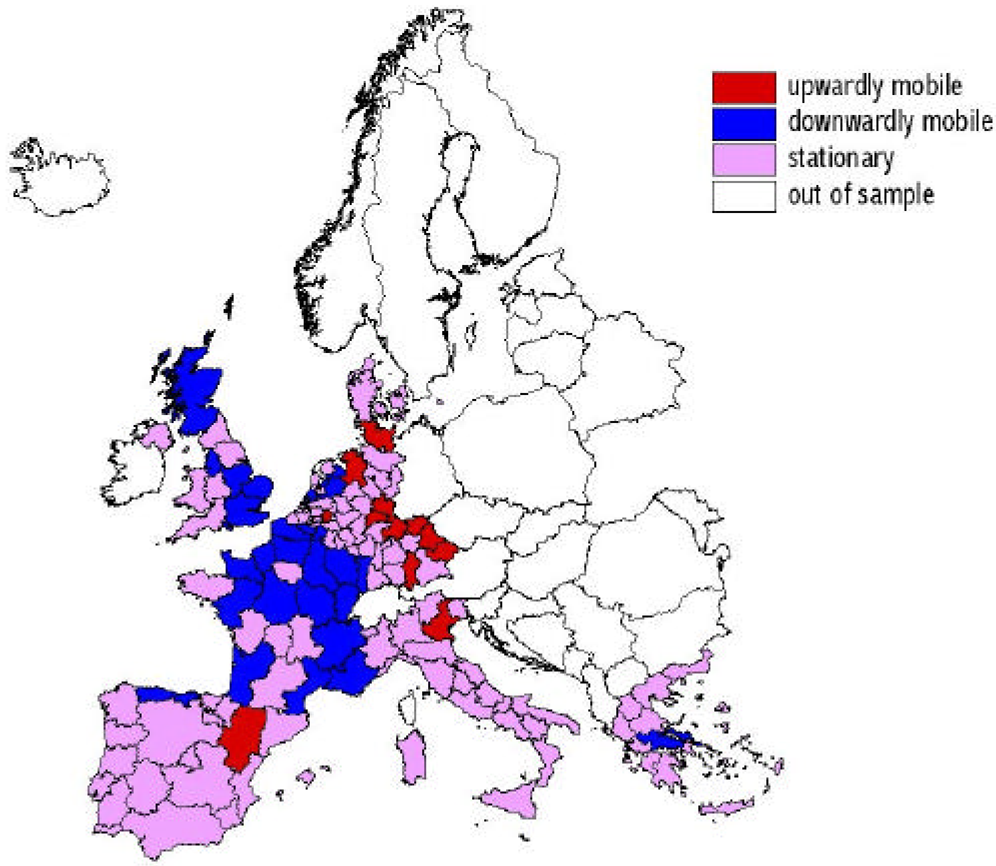
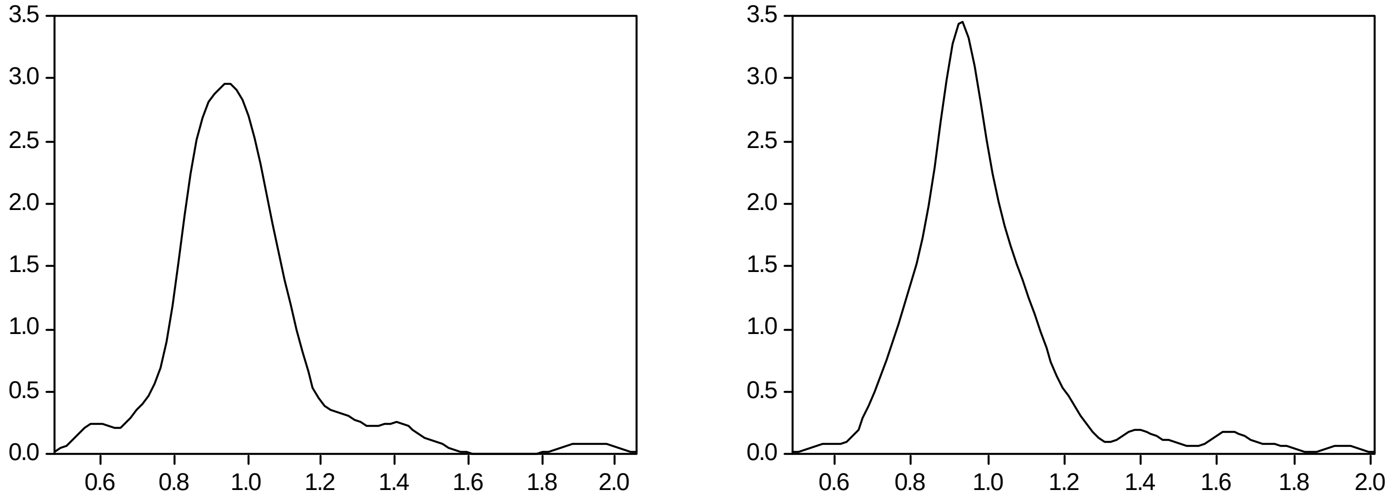


Fig.2. Region GDP class transitions 1980-1995



Neighbor-relative GDP 1980

Neighbor -relative GDP 1995

Fig.3. Densities of neighbor-relative regional GDP

ti \ ti+1		Europe-relative 1 <65%	2 <95%	3 <110%	4 <125%	5 >125%	Number of observations
Europe relative	1	0,959	0,041	0,000	0,000	0,000	416
	2	0,047	0,865	0,087	0,000	0,000	401
	3	0,000	0,095	0,822	0,083	0,000	433
	4	0,000	0,000	0,123	0,796	0,081	382
	5	0,000	0,000	0,000	0,073	0,927	438
Initial dist.		0,201	0,194	0,209	0,185	0,212	
Ergodic dist.		0,264	0,227	0,210	0,142	0,157	

Tab.1. Probability transition matrix 1980-1995; Europe-relative per-capita GDP

(3,788)	24,456	49,006	89,000	162,720
68,146	(4,405)	24,655	64,604	138,161
93,501	25,579	(4,762)	40,099	113,648
110,547	42,777	17,123	(7,042)	73,459
124,162	56,512	30,800	13,702	(6,369)

Tab.2. Mean first passage time matrix

ti \ ti		Neighbor-relative 1 <65%	2 <95%	3 <110%	4 <125%	5 >125%	Number of observations
Europe relative	1	0,038	0,498	0,347	0,076	0,042	450
	2	0,031	0,533	0,336	0,061	0,040	426
	3	0,030	0,652	0,286	0,026	0,006	465
	4	0,000	0,428	0,504	0,043	0,025	397
	5	0,000	0,049	0,451	0,245	0,255	470

Tab.3. Regional conditioning

ti \ ti+1		Neighbor-relative 1 <65%	2 <95%	3 <110%	4 <125%	5 >125%	Number of observations
Neighbor relative	1	0,837	0,163	0,000	0,000	0,000	43
	2	0,005	0,924	0,072	0,000	0,000	878
	3	0,000	0,085	0,878	0,038	0,000	800
	4	0,000	0,000	0,153	0,805	0,042	190
	5	0,000	0,000	0,000	0,050	0,950	159
Initial dist.		0,021	0,424	0,386	0,092	0,077	
Ergodic dist.		0,012	0,444	0,375	0,092	0,077	

Tab.4. Probability transition matrix 1980-1995; Neighbor-relative per-capita GDP

Spatial lag	t_i \ t_{i+1}	Europe-relative 1 <65%	2 <95%	3 <110%	4 <125%	5 >125%	Number of observations
1	1	0,977	0,023	0,000	0,000	0,000	390
	2	0,164	0,821	0,015	0,000	0,000	58
	3	0,000	0,500	0,500	0,000	0,000	2
	4	0,000	0,000	0,000	0,000	0,000	0
	5	0,000	0,000	0,000	0,000	0,000	0
	Initial dist.	0,852	0,144	0,004	0,000	0,000	
	Ergodic dist.	0,875	0,121	0,004	0,000	0,000	
2	1	0,600	0,400	0,000	0,000	0,000	26
	2	0,044	0,918	0,038	0,000	0,000	167
	3	0,000	0,167	0,771	0,063	0,000	48
	4	0,000	0,000	0,200	0,800	0,000	29
	5	0,000	0,000	0,000	0,000	0,000	0
	Initial dist.	0,074	0,674	0,178	0,074	0,000	
	Ergodic dist.	0,078	0,708	0,163	0,051	0,000	
3	1	0,000	0,000	0,000	0,000	0,000	0
	2	0,000	0,803	0,197	0,000	0,000	85
	3	0,000	0,171	0,768	0,061	0,000	82
	4	0,000	0,000	0,100	0,820	0,080	44
	5	0,000	0,000	0,000	0,125	0,875	29
	Initial dist.	0,000	0,271	0,364	0,222	0,142	
	Ergodic dist.	0,000	0,303	0,349	0,213	0,136	
4	1	0,000	0,000	0,000	0,000	0,000	0
	2	0,000	0,795	0,205	0,000	0,000	50
	3	0,000	0,072	0,878	0,050	0,000	165
	4	0,000	0,000	0,163	0,721	0,116	82
	5	0,000	0,000	0,000	0,092	0,908	123
	Initial dist.	0,000	0,096	0,343	0,212	0,348	
	Ergodic dist.	0,000	0,171	0,488	0,151	0,190	
5	1	0,000	0,000	0,000	0,000	0,000	0
	2	0,000	0,865	0,135	0,000	0,000	41
	3	0,000	0,049	0,821	0,130	0,000	136
	4	0,000	0,000	0,106	0,819	0,075	227
	5	0,000	0,000	0,000	0,057	0,943	286
	Initial dist.	0,000	0,074	0,230	0,321	0,376	
	Ergodic dist.	0,000	0,087	0,238	0,290	0,385	

Tab.5. Spatial Markov chain; conditioning on the spatial lag in 1980

Spatial lag	N	Movement Down	None	Up
poorer	360	0,128	0,850	0,022
same	1011	0,050	0,919	0,031
richer	699	0,060	0,825	0,114

Tab.6. Transition probabilities conditioned on the spatial lag of GDP in 1980

Spatial lag	N	Movement Down	None	Up
poorer	277	0,119	0,866	0,014
same	1202	0,065	0,909	0,026
richer	591	0,047	0,810	0,142

Tab.7. Transition probabilities conditioned on the spatial lag of GDP at the beginning of each year

References

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

Anselin L. (1999) SpaceStat, a software package for the analysis of spatial data, Version 1.90, Ann Arbor, BioMedware.

Anselin L. (2000) Spatial econometrics, in: Baltagi B. (Ed.), *Companion to Econometrics,* Oxford, Basil Blackwell.

Armstrong H.W., Vickerman R. (1995) *Convergence and Divergence among European Regions,* London, Pion.

Armstrong H.W. (1995) Convergence among the regions of the European Union, *Papers in Regional Science*, 74, 143-152.

Barro R. (1991) Economic growth in a cross section of countries, *Quarterly Journal of Economics*, 106, 407-443.

Barro R., Sala-i-Martin X. (1991) Convergence across states and regions, *Brookings Papers in Economic Activity,* 1, 107-122.

Barro R., Sala-i-Martin X. (1995) *Economic Growth*, London, McGraw Hill.

Baumol W.J. (1986) Productivity growth, convergence, and welfare: what the long-run data show, *American Economic Review*, 54, 1072-1085.

Baumont C., Ertur C., Le Gallo J. (2001) A spatial econometric analysis of geographic spillovers and growth for European regions, 1980-1995, working paper, LATEC, n°2001-04, University of Burgundy, Dijon.

Beine M., Docquier F. (Eds.) (2000) *Croissance et Convergence Economique des Régions. Théories, Faits et Déterminants*, De Boeck Université.

Bianchi M. (1997) Testing for convergence: evidence from nonparametric multimodality tests, *Journal of Applied Econometrics,* 12, 393-409.

Chung K.L. (1960) *Markov Chains with Stationary Transition Probabilities*, Berlin, Springer-Verlag.

Cliff A.D., Ord J.K (1981) *Spatial Processes: Models and Application*s, London, Pion.

Desdoigts A. (1999) Patterns of economic development and the formation of clubs, *Journal of Economic Growth*, 4, 305-330.

Fingleton B. (1997) Specification and testing of Markov chain models: an application to convergence in the European Union, *Oxford Bulletin of Economics and Statistics,* 59, 385-403.

Fingleton B. (1999) Estimates of time to economic convergence: an analysis of regions of the European Union, *International Regional Science Review*, 22, 5-34.

Friedman M. (1992) Do old fallacies ever die ?, *Journal of Economic Litterature*, 30, 2129-2132.

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

Kemeny J., Snell L. (1976) *Finite Markov Chains*, New York, Springer-Verlag.

Kullback S., Kupperman M., Ku H.H. (1962) Tests for contingency tables and Markov chains, *Technometrics*, 4, 573-608.

Le Gallo J., Ertur C. (2000) Exploratory spatial data analysis of the distribution of regional per capita GDP in Europe, 1980-1995, working paper, LATEC, n°2000-09, University of Burgundy, Dijon.

López-Bazo E., Vayá E., Mora A.J., Suriñach J. (1999) Regional economic dynamics and convergence in the European Union, *Annals of Regional Science*, 33, 343-370.

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

Moreno R., Trehan B. (1997) Location and the growth of nations, *Journal of Economic Growth*, 2, 399-418.

Neven D., Gouyette C. (1995) Regional convergence in the European Community, *Journal of Common Market Studies*, 33, 47-65.

Paap R., Van Dijk H. (1998) Distribution and mobility of the wealth of nations, *European Economic Review*, 42, 1269-1293.

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

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

Quah D. (1996a) Empirics for economic growth and convergence, *European Economic Review*, 40, 1353-1375.

Quah D. (1996b) Regional convergence clusters across Europe, *European Economic Review*, 40, 951-958.

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

Rey S. (2001) Spatial empirics for economic growth and convergence, forthcoming in *Geographical Analysis*.

Rey S., Montouri B. (1999) US regional income convergence: a spatial econometric perspective, *Regional Studies*, 33, 143-156.

Sala-i-Martin X. (1996) Regional cohesion: evidence and theories of regional growth and convergence, *European Economic Review*, 40, 1325-1352.

Shorrocks A. (1978) The measurement of mobility, *Econometrica*, 46, 1013-1024.

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