Spatial disparities in European agriculture: A regional analysis

Roberto Ezcurra, Belén Iraizoz, Pedro Pascual and Manuel Rapún*

Department of Economics

Universidad Pública de Navarra

Abstract

This paper examines the territorial imbalances in European agriculture during the period 1980-2001, by means of the information provided by various methodological instruments with which it is possible to overcome the drawbacks of conventional convergence analysis. The results obtained reveal that the regional distribution of productivity in the agricultural sector is characterized by the presence of positive spatial dependence. This fact implies that the European regions in close spatial proximity register similar levels of the variable under study. The empirical evidence, meanwhile, also shows that regional disparities have remained almost constant during the time interval considered. However, the increase in density around the European mean explains the observed reduction in bipolarization, while intradistributional mobility is relatively limited. Finally, the analysis carried out allows us to assess the role of variables such as country of origin, investment per worker in the agricultural sector, regional per capita income or the size of the agrifood industry in explaining the dynamics of the distribution under analysis.

Key words: Agricultural productivity, regions, European Union.

JEL code: Q10, R11, R12.

*Postal address: Roberto Ezcurra, Department of Economics, Universidad Pública de Navarra, Campus de Arrosadia s/n. 31006 Pamplona (Spain). E-mail address: roberto.ezcurra@unavarra.es.

1 Introduction

Recent years have seen the publication of a great a number of works using a variety of approaches to analyze territorial imbalances in per capita income or aggregate productivity within the European Union[1]. Various factors contribute to the level of interest raised by this issue. Not least among them are the major developments that have taken place in economic growth theory over the last twenty years, coinciding with the advent of endogenous growth models during the eighties. Another is the need to reduce development gaps across the various European regions, an issue that inspired some of the basic principles upon which the Union was founded, especially since the signing of the Single Act and the Maastricht agreements. Indeed, one of the specific assumptions of the European integration program is that it will drive the growth of all Member States, and thereby lead to economic and social cohesion[2].

However, sectoral analysis at regional level for the European Union as a whole are fewer in number, probably as a result of problems arising from the scarcity of adequate statistical data[3]. Especially surprising within this context is the fact that there has as yet appeared so little in-depth analysis of existing territorial imbalances between the European regional agricultures[4], taking into account the large amount of funds that have been allocated to the Common Agricultural Policy (CAP) over the last four decades[5]. In

[1]A review of the main conclusions reported in this literature can be found in Armstrong (2002) or Terrasi (2002).

[2]Article 2 of the Treaty of the European Union specifically states that: "the Community mission will be to promote (...) the harmonious, balanced and sustainable development of economic activities, sustained growth (...), a high degree of convergence of economic performance (...)".

[3]In relation to this, see Paci (1997), Paci and Pigliaru (1999) or Gil, Pascual and Rapún (2002).

[4]The only exception, to our knowledge, is the work by Colino and Noguera (2002). There are, however, several examples of research efforts using national level data to investigate this issue within the European setting, among which we could mention Schimmelpfenning and Thirtle (1999), Gutierrez (2000) or Ball *et al.* (2001).

[5]The resources used to finance this policy accounted, in 2002, for 46.5 per cent of the community budget.

light of these circumstances, this paper aims to perform a fairly detailed analysis of the regional distribution of output per worker in the European Union agricultural sector. Our ultimate goal, in doing so, is to draw some kind of inference that might be applied in the design of community policy within the framework of the recent enlargement of the Union towards Central and Eastern Europe.

The few existing studies of regional disparities at sectoral level in the European context apply the concepts of sigma convergence and beta convergence introduced by Barro and Sala-i-Martin (1991, 1992), combining the information provided by various dispersion statistics with the estimation of convergence equations. However, as Quah (1993, 1996a, 1997) has repeatedly pointed out, not only does this approach raise a number of econometric problems[6], it also fails to capture a series of potentially interesting issues relating to the dynamics of the distribution in question. In particular, this type of approach offers only a partial view of the distribution under analysis, since, not only does it ignore the issue of intradistributional mobility, it also fails to consider the possibility of the appearance of clusters of regions with specific characteristics distinguishing them from the rest of the population[7].

In order to overcome the limitations of conventional convergence analysis, we have opted in this paper to use the non-parametric approach proposed by Quah (1996a, b; 1997) to examine the dynamics of a distribution over time. We have also applied a set of techniques adopted from spatial econometrics in order to examine the role played by the space variable in this context.

The statistical data employed in this study was supplied by the Cambridge Econo-

[6] For further details, see Canova and Marcet (1995).

[7] Readers interested in a more detailed discussion of these issues might consult, in addition to the above-mentioned works of Quah, the panoramic study by Magrini (2004).

metrics regional database. The use of this source allowed us to work with data on gross value added at market prices and employment corresponding to the agricultural sector for 194 NUTS-2 regions during the period 1980-2001[8].

The content of the paper is organized as follows. Following on from this introduction, section 2 presents an exploratory analysis of the spatial distribution of productivity in the agricultural sector in the European Union. Section 3 examines the evolution of regional disparities in output per worker in the European agriculture between 1980 and 2001. Then, in section 4, we explore the dynamics of the distribution of interest, paying particular attention to polarization and regional mobility. This done, and in order to complete the results obtained, we investigate in section 5 the role played by a range of variables in explaining existing territorial imbalances in the European primary sector. We finish by presenting the main conclusions in section 6.

2 Regional distribution of agricultural productivity: An exploratory spatial data analysis

We will commence our study by carrying out a preliminary analysis of the spatial distribution of agricultural productivity in the European Union between 1980 and 2001. Thus, close observation of Figures 1 and 2 reveals the existence in the two years in question of an apparent spatial non-stationarity in the distribution that concerns us. Indeed the presence in this context of some degree of spatial heterogeneity is quite evident.

[8]Lack of information for the whole of the time interval considered has obliged us to omit from our analysis the countries newly incorporated into the European Union in May 2004, as well as the former East German *Länder* and the French Overseas departments. We also decided not to include data for Brussels and Inner London, two regions with practically negligible levels of farm sector employment, but which, nevertheless register over time major fluctuations in the variable of interest, thus affecting the interpretation of our results.

Thus, notwithstanding the changes that have taken place over time, the regions with the highest levels of output per worker in the primary sector are mainly those of the northern and central areas of the Union. On the contrary, and save for a few exceptions, the lowest values of the variable under study are to be found in regions situated in the southern periphery. This initial analysis, therefore, suggests that agricultural productivity is not randomly distributed across European space. By contrast, there appears to be positive spatial association between adjacent areas, in so far as output per worker in the primary sectors of adjacent regions tends to exhibit similar levels. This is consistent with the dualized view of the European agriculture reported by, among others, Kearney (1992) and Gutierrez (2000).

[INSERT FIGURE 1]

[INSERT FIGURE 2]

It is necessary to exercise caution when interpreting the information supplied by Figures 1 and 2, however, because any conclusions drawn from them will be highly sensitive to the number of intervals selected to represent the variable in question. Bearing this in mind, therefore, and in order to check formally the existence of spatial autocorrelation in the distribution of productivity in agricultural sector in the European Union, we proceeded by computing Moran's I and Geary's C global tests, defined respectively as (Cliff and Ord, 1973, 1981; Haining, 1990):

$$I = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(y_i - \overline{y})(y_j - \overline{y})}{S_0 \sum_{i=1}^{n} (y_i - \overline{y})^2} \tag{1}$$

and

$$C = \frac{(n-1)\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(y_i - y_j)^2}{2S_0 \sum_{i=1}^{n}(y_i - \overline{y})^2} \tag{2}$$

where y_i denotes the agricultural productivity in region i, while $\overline{y}$ is the sample mean. Likewise, w_{ij} denotes the corresponding element of the spatial weight matrix, W, with $\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} = S_0$. As far as interpretation is concerned, it should be noted that after standardization, a significant and positive (negative) value of Moran's I (Geary's C) will indicate the presence of positive spatial autocorrelation, while a significant and negative (positive) value of Moran's I (Geary's C) will tell us of the existence of a pattern of spatial association between dissimilar values.

As can be seen from (1) and (2), before performing these tests, it is first necessary to construct a spatial weight matrix to capture the strength of interdependence between each pair of regions i and j. A first option is to use the concept of first order contiguity, according to which $w_{ij} = 1$ if regions i and j are physically adjacent and 0 otherwise[9]. However, the use of this type of matrix may raise problems in the European context, given that the presence of islands means that W will include rows and columns containing only zeros. This implies that the observations in question are not considered in the analysis, which in turn has an effect on the interpretation of the results obtained. For this reason, in this paper we have opted for the alternative of using a spatial weight matrix that takes into account the interactions existing beyond the adjacent regions. In particular, following the proposal made by Le Gallo and Ertur (2003), we have considered a row-standardized matrix W based on the ten nearest neighbors, calculated using the geographical distance among the corresponding regional centroids (Pace and Barry, 1997;

[9]This is in fact the option taken, for example, by López-Bazo *et al.* (1999) or Rey and Montouri (1999).

Pinkse and Slade, 1998)[10].

The results of the two global spatial autocorrelation tests just mentioned are presented in Table 1, where it can be seen that the standardized values of Moran's I (Geary's C) are in all cases statistically significant and positive (negative). This reveals the presence of a clear pattern of positive spatial association, which in turn supports the first impression drawn from observation of Figures 1 and 2[11]. We are able to conclude, therefore, that, within the European Union, spatially adjacent regions tend to be characterized by having similar levels of agricultural productivity[12]. Leaving aside all other considerations, this situation is consistent with the results obtained for several geographical settings in a series of studies highlighting the presence of important R&D spillover effects in the agricultural sector (Johnson and Evenson, 1999; Schimmelpfenning and Thirtle, 1999; McCunn and Huffman, 2000). In any event, Table 1 suggests that spatial factors have decreased in importance in the European setting over the course of the twenty-two years considered.

[INSERT TABLE 1]

[10] In order to check the robustness of the conclusions obtained, we considered various spatial weight matrices. In particular, we constructed additionally two new matrices W based on the fifteen and twenty nearest neighbors. Nevertheless, the results are in all cases very similar to those reported in the text.

[11] This result is of particular importance if, following the approach adopted by Paci (1997), Colino and Noguera (2002) or Gil, Pascual and Rapún (2002), the aim of the analysis is to examine regional disparities in European agriculture through the estimation of convergence equations. Indeed, the information supplied by Table 1 raises serious doubts as to the consistency, unbiasedness and/or efficiency of estimations obtained without taking into account the existence of spatial dependence. A more detailed analysis of this issue can be found in Cliff and Ord (1981) or Anselin and Griffith (1988).

[12] To further confirm this finding, we also constructed the Moran's scatterplots for the distribution under analysis. In this type of graph, the standardized values of the variable to be analyzed are plotted on the horizontal axis and the spatial lag of the same variable on the vertical axis. Thus each of the quadrants represents a different type of spatial association. As can be seen from Figures A1 and A2, in both 1980 and 2001, there is a noticeable concentration of regions in quadrants I and III, thus confirming a predominating tendency in the European Union towards spatial clustering involving regions with similar values of farm productivity, while there are relatively few cases of major discrepancy between the agricultural productivity of any given region and the mean of its neighbors.

It is important to keep in mind, however, that both Moran's I and Geary's C are calculated on a global basis for the entire sample. Hence, we do not know whether, irrespective of the overall dependence pattern, there exist clusters of regions in which the concentration of high or low agricultural productivity levels is significantly greater than would be predicted in a spatially homogeneous distribution. In addition, it is not possible with these tests to detect the existence of groups of regions with dissimilar values of the study variable, that is, areas in which output per worker in the primary sector is significantly lower or higher than in the adjacent areas. To overcome both these shortcomings, we calculated the local Moran's I, I_i, by means of the following expression (Anselin, 1995):

$$I_i = \frac{n(y_i - \overline{y})}{\sum_{i=1}^{n}(y_i - \overline{y})^2} \sum_{j \in J_i} w_{ij}(y_j - \overline{y}) \tag{3}$$

where J_i denotes the set of neighboring regions of i[13]. After standardization, a significant positive (negative) value of I_i will indicate clustering around region i of regions with similar (dissimilar) values of the study variable.

As in the global tests considered earlier, the usual practice in the case of I_i is to assume a normal asymptotic distribution. Anselin (1995), however, has shown that the first and second order moments used in the standardization of the statistic I_i are obtained under the null hypothesis of no global spatial autocorrelation. For this reason, and following the proposal of this author, in this paper we calculated the pseudo-significance levels obtained by means of an empirical distribution derived from 1,000 random permutations.

[13]Note that $I = \frac{\sum_{i=1}^{n} I_i}{S_0}$, given that $\sum_{i=1}^{n} I_i = \frac{n \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(y_i-\overline{y})(y_j-\overline{y})}{\sum_{i=1}^{n}(y_i-\overline{y})^2}$.

Figures 3 and 4 show the significant regional clusterings detected in 1980 and 2001, and indicate furthermore, whether or not they concentrate similar levels of agricultural productivity[14]. As can be checked, the conclusions that emerge from the analysis carried out are consistent with the results obtained earlier. One point worth noting is that, in 1980, the clusters of regions with high values of the study variable are situated in Finland, the Netherlands, Belgium, the British Midlands, northern Sweden and France. By contrast, the regional groupings characterized by low levels of farm productivity are located in Portugal, most of Spain and Austria, and the southern parts of Germany and Italy. The situation at the end of the sample period is similar to that just described. However, in 2001 the I_i statistic for the Netherlands and large parts of Spain and Austria ceases to be statistically significant. Nevertheless, in that same year, there emerges a new cluster formed by the whole of Greece, grouping regions with a lower level of output per worker in the primary sector than those adjacent to them.

[INSERT FIGURE 3]

[INSERT FIGURE 4]

It is also worth noting the fact that the various regional clusters detected in the above analysis are comprised mainly by regions with similar levels of agricultural productivity. In any event, it is interesting to observe the slight increase that has taken place in the number of regions with significant and negative values of the local Moran's I coefficient during the period 1980-2001. However, it must be borne in mind that of the nine regions

[14] In addition, Figures A3 and A4 report the significance level corresponding to each region.

that in 1977 presented a value of the variable under study significantly distinct from that of their neighbors, only Luxembourg, Auvergne, Tees Valley and Durham remain in the same situation in 2001.

3 Regional disparities in European agriculture

Following this preliminary analysis of the spatial distribution of the variable of interest, we continue our study by examining the evolution of regional disparities in output per worker in the European Union primary sector. To this end, we consider the information supplied by the two dispersion measures commonly used to capture the concept of sigma convergence (Barro and Sala-i-Martin, 1992, 1992): the standard deviation of the logs ($SDlog$) and the coefficient of variation (CV)[15].

First of all, it should be noted that the magnitude of the territorial imbalances observed in European agriculture is considerable greater than in the secondary and tertiary sectors (Table 2)[16]. In any case, as Figure 5 shows, inequality in the regional distribution of agricultural productivity within the European setting registered little relevant changes between 1980 and 2001. In fact, the absolute variation in the values of $SDlog$ and CV never exceeded 2 percent during the whole of the twenty-two years contemplated. However, the degree of dispersion in the distribution that concerns us did not remain constant throughout that period. In fact, it is possible to identify three distinct stages, each with its own distinguishing features. Thus, during the eighties there was an increase in inequality, reaching its maximum level in 1990. However, in the early

[15] Dalgaard and Vastrup (2001) have demonstrated that the joint use of the standard deviation of the logs and the coefficient of variation does not prove redundant in this setting, since these two dispersion statistics could yield different conclusions.

[16] A similar result is found by Paci (1997) in a more limited geographical and temporal setting than that covered in the present study. Indeed, similar conclusions have been reached in national analyses. In the Spanish context, see, for example, Mas *et al.* (1994) or Raymond and García-Greciano (1994).

nineties it is possible to detect a process of productivity convergence in the European regional agricultures. This situation changes in 1997, when inequality again starts to increase.

[INSERT TABLE 2]

[INSERT FIGURE 5]

When assessing the implications of these findings, it is important to bear in mind that there exist various possible explanations for the lack of convergence in output per worker in European agriculture throughout the entire period of analysis. Thus, it should be remembered that, in spite of the major structural reform that has taken place in the primary sector over the last few decades, agricultural productivity continues to be closely linked to the natural setting and climatic conditions of the various regions. This, in turn, helps to account for the heterogeneity of farming systems and the differences in their remunerative capacity, which constitute a considerable hindrance when it comes to reducing regional inequality over time (Jollivet and Eizner, 1996; Limouzin, 1996).

It should not be forgotten, however, that the evolutionary pattern of agricultural productivity is ultimately a consequence of changes in value added and/or employment in the different regions. Following the approach adopted by several authors within the framework of the literature dealing with the study of territorial imbalances in per capita income, to explore the role of each of these variables in the evolution of regional disparities in European agriculture, a first option is to characterize the spatial distribution of value added and employment by applying a range of concentration measures (Villaverde, 2003). For obvious reasons, however, it is risky to relate the results of an analysis of this nature to the evolution of regional disparities in output per worker in the primary sector. For a better understanding of this idea, let us consider a hypothetical situation in which two regions, each with a very different level of value added, exchange their respective employment levels. In such a context, none of the concentration measures proposed in the literature will register any variation over time, despite the obvious fact that changes in the spatial distribution of employment have an effect on observed inequality.

To overcome this problem, therefore, we perform an alternative analysis, in which we estimate the degree of dispersion in two virtual distributions. In the first, regional value added is kept constant at the level for the start of the period and only employment levels are allowed to vary. On the contrary, in the second virtual distribution employment is kept constant and only regional value added is allowed to vary over time. The results of this exercise are reported in Figure 6, where it can be seen that dispersion in the first virtual distribution remained practically constant throughout the period 1980 to 2001, while in the second it can be seen to increase. This shows that changes in the distribution of employment prevented an increase in regional disparities in the primary sector during the twenty-two years considered. Specifically, the empirical evidence provided in this respect suggests that the higher growth in value added registered by the regions located at the upper end of the distribution has in turn been offset by a greater loss of employment in regions with low agricultural productivity levels in 1977. This conclusion is in fact consistent with the findings of Paci (1997) or Gil, Pascual and Rapún (2002) with respect to the process of structural change that has taken place in the European Union over the last few decades.

[INSERT FIGURE 6]

4 The dynamics of the regional distribution of productivity in agricultural sector

In the section above we examined the regional disparities in output per worker in the European agriculture. It is obvious, however, that the various dispersion measures calculated so far do not give a precise picture of the regional distribution of agricultural productivity in the European context. For this reason, we will now estimate the density functions of the distribution analyzed. Following common practice in the literature, we will use non-parametric estimation techniques, thus avoiding the need to specify any particular functional form beforehand. This kind of approach undoubtedly offers major advantages in the present context, given the lack of generality and flexibility associated with parametric approximations.

Figure 7 shows the density functions for the regional distribution of output per worker in the European farm sector[17]. Normalized productivity levels are plotted on the horizontal axis (sample mean equal to 100), while the associated density estimates are plotted on the vertical axis. All estimates are based on calculations using Gaussian kernel functions and the optimal smoothing parameter value is determined in each case following Silverman (1986, p. 47).

[INSERT FIGURE 7]

As can be checked in Figure 7, the unimodal form of the distribution analyzed is a constant throughout the study period. Nevertheless, the results reveal certain differences in the shape of the estimated density during the twenty-two years considered, allowing

[17] Though density functions were estimated for each year of the period analyzed, in order to save space, we present only those for 1980, 1990 and 2001. The rest are available from the authors upon request.

us to conclude that the initial situation did not remain stable during the period. Thus, the density concentrated around the mean increased between 1980 and 2001, largely due to the weight loss of the upper end of the distribution over the nineties. These results complete and further the findings obtained earlier in relation to the evolution of regional inequality in the productivity of agricultural sector throughout the time interval contemplated.

However, there is major limitation to the non-parametric approach used in this section so far, since it does not permit us to obtain a precise quantification of the variations in the degree of polarization that have taken place over time. To address this problem, we applied the methodology proposed by Esteban and Ray (1994), along with an extension by Esteban, Gradín and Ray (1999).

According to Esteban and Ray (1994), it is possible to measure the degree of polarization of a distribution f partitioned into a number of groups exogenously determined by means of the following expression:

$$P^{ER}(f,\alpha,\rho) = \sum_{j=1}^{m}\sum_{k=1}^{m} p_j^{1+\alpha} p_k \left|\mu_j - \mu_k\right| \qquad (4)$$

where, for the purposes of the present study, μ_j and p_j denote, respectively, the mean level of productivity in the primary sector and the population share of group j. Likewise, $\alpha \in [1, 1.6]$ is a parameter that captures the degree of sensitivity of P^{ER} to polarization. Nevertheless, before applying this measure, it is necessary to define a simplified representation of the original distribution in a series of exhaustive and mutually exclusive groups, $\rho = (z_0, z_1, \ldots, z_m, \mu_1, \ldots, \mu_m, p_1, \ , \ldots, p_m)$, the boundaries of which are given by intervals of the form $[z_{j-1}, z_j]$, for $j = 1, \ldots, m$. This will involve some degree of error, however, as this grouping will generate certain loss of information, depending on the degree of dispersion within each of the groups considered. Taking into account this idea, the measure of generalized polarization proposed by Esteban, Gradín and Ray (1999) is obtained after correcting the P^{ER} index applied to the simplified representation of the original distribution with a measure of the grouping error, $\varepsilon(f, \rho)$.

It is important to bear in mind, meanwhile, that there are no unanimous criteria for establishing the precise demarcation between the various groups. To address this problem, Esteban, Gradín and Ray (1999) use the methodology proposed by Aghevli and Mehran (1981) and Davies and Shorrocks (1989) in order to find the optimal partition of the original distribution, ρ^*. This means selecting the partition that minimizes the Gini index value attributable to within-group inequality, $G(f) - G(\rho^*)$. The measure of generalized polarization proposed by Esteban, Gradín and Ray (1999), therefore, is given by:

$$P^{EGR}(f, \alpha, \rho^*, \beta) = P^{ER}(f, \alpha, \rho^*) - \beta [G(f) - G(\rho^*)] \tag{5}$$

where $\beta \geq 0$ is a parameter that informs about the weight assigned to the error term in expression (5).

We proceeded by applying this methodology to the study of the evolution of polarization in the regional distribution of agricultural productivity in the European Union for the two-group case (bipolarization). In order to check the robustness of the results, we consider in our analysis different values of the parameter of sensitivity to polarization. Specifically, $\alpha = 1, 1.6$. Likewise, as in Esteban, Gradín and Ray (1999), $\beta = 1$ in all cases.

As shown in Figure 8, the results obtained reveal a reduction in generalized bipolarization over the period analyzed, which was particularly intense between 1985 and 1995. Indeed, the value of P^{EGR} decreased by between 12 and 22 percent, depending on the degree of sensitivity to polarization considered in each case[18].

[INSERT FIGURE 8]

Nevertheless, expression (5) highlights the fact that the evolution of P^{EGR} depends on two factors: the polarization observed in the simplified representation of the original distribution and the degree of internal dispersion within each of the various groups. Figure 9, therefore, provides additional information regarding these two components of the generalized polarization measure. Thus, it is possible to observe a decrease in bipolarization in the simplified representation during the time interval considered. Specifically, the value of P^{ER} fell by 4 per cent for the various levels of sensitivity to polarization considered in the analysis. With respect to the evolution followed by the error term, Figure 9 reveals that the degree of internal cohesion in both groups decreased, since the value of ε increased by 8 per cent throughout the twenty-two years contemplated. In any event, this evolution of the error term contributes to explain the reduction in generalized bipolarization observed between 1980 and 2001, since it worsened in absolute terms the prediction of bipolarization based on the simplified representation of the original distribution.

[INSERT FIGURE 9]

The analysis carried out so far is based exclusively on the information obtained from a series of cross-sectional observations of the distribution under study. Therefore, it

[18]The optimal partition for the two-group case is characterized by the fact that the productivity level that separates the two groups coincides with the sample mean. By adopting this criterion for the classification of the various regions considered, it is possible to explain an average of 71 per cent of total inequality, measured in terms of the Gini index. Thus, the within-group inequality left unexplained by this partition is about 29 per cent.

does not take into account that the different economies may modify over time their relative positions in terms of output per worker in the primary sector. To address this shortcoming and to complete the results obtained so far, we have examined the degree of mobility in the regional distribution of agricultural productivity within the European Union during the period 1980-2001.

Most of the studies that have addressed this issue are based on the information provided by discrete transition matrices, obtained by dividing the distribution into a series of exhaustive and mutually exclusive classes[19]. This approach entails a problem, however, since the results it yields are sensitive to the way in which the observed distribution is divided up. Nevertheless, since there is no a single procedure for determining the optimum number of classes in each case, the researcher's decision is inevitably arbitrary (Bulli, 2001; Kremer, Onatski and Stock, 2001). To address this problem, Quah (1996a, 1997) suggests substituting the transition matrix with a stochastic kernel to reflect the probabilities of transition between a hypothetically infinite number of classes, reducing their size infinitesimally[20]. The stochastic kernel can be reached by estimating the density function of the distribution over a given period, $t + k$, conditioned on the values corresponding to a previous period, t. In other words, the joint density function at moments t and $t + k$ is estimated and then divided by the implicit marginal distribution in order to obtain the corresponding conditional probabilities.

Figure 10 shows the stochastic kernel estimated from the regional distribution of output per worker in the primary sector for a time interval of twenty-two years ($t = 1980$ to $t + k = 2001$)[21]. This three-dimensional graph can be interpreted as a transition matrix

[19] For the European case López-Bazo *et al.* (1999) and Cuadrado, Mancha and Garrido (2002) estimate various transition matrices to analyze regional mobility in terms of per capita income and aggregate productivity.

[20] For a formal definition, see Durlauf and Quah (1999).

[21] Gaussian kernel functions are used again, while the smoothing parameter has been selected following

with an infinite number of classes, that informs about the probabilities associated with each pair of values in the first and last years of the study period. In other words, the stochastic kernel gives us, as does a discrete transition matrix, the probability distribution of 2001 agricultural productivity for regions with a given value in 1980. Thus, if the probability mass is concentrated around the main diagonal, the intradistributional dynamics are characterized by a high level of persistence in the relative positions of the regions over time and, therefore, low mobility. If, on the other hand, the density is located mainly on the opposite diagonal to the main diagonal, this would indicate that regions situated at both extremes of the distribution exchange their relative positions. Finally, the probability mass could, in theory, accumulate parallel to the t axis. This would reflect the convergence of regional output per worker in the primary sector throughout the study period. In order to aid interpretation of the graph, Figure 10 also includes contour plots on which the lines connect points at the same height on the three-dimensional kernel.

[INSERT FIGURE 10]

As Figure 10 shows, the probability mass is concentrated around the main diagonal. As we know, this can be taken as evidence of low mobility in the regional distribution of output per worker in the primary sector during the period 1980-2001. The European regions, therefore, tended on the whole to maintain their relative positions over the twenty-two years contemplated. In fact, this finding is not affected by the right turn observed at the upper end of the distribution, which is entirely due to the behavior of a

Silverman (1986, p. 86).

small number of regions (Champagne-Ardenne, Ovre Norrland and Outer London) with high values of the study variable that act as outliers.

We then estimated the corresponding ergodic distribution by iteration of the stochastic kernel until to reach the convergence of the process. Given that this is, by definition, a continuous distribution, it can be represented graphically (Figura 11). As shown, the ergodic distribution obtained is characterized by a single local maximum located around the sample mean, which is, in turn, consistent with the information yielded by the density functions estimated in Figure 7 for various years of the period 1980-2001[22]. This points out that the existing territorial imbalances in European agriculture will persist into the future. Nevertheless, it is worth mentioning that there is not evidence to suggest that in the long term the distribution under study will fragment into various separate groups of regions[23].

[INSERT FIGURE 11]

5 Determinants of the dynamics of the regional distribution of agricultural productivity

In order to complete the results obtained so far, in this section we will examine the role played by a series of variables in the dynamics of the regional distribution of output per worker in the European agricultural sector. We will tackle this issue using

[22]Having reached this point, however, it should be noted that any comparison of Figures 7 and 11 must be based exclusively on the shape of the distributions, since there is no point in comparing the density levels that appear on the vertical axes.

[23]It is obvious, however, that the findings obtained from the analysis of Figures 10 and 11 are determined by the dynamics of the regional distribution of output per worker in the farm sector throughout the whole of time interval analyzed. We therefore decided to repeat the analysis using only data for the subperiods 1980-1990 and 1990-2001. The results, which are very similar to those already discussed, are available from the authors upon request.

a series of instruments proposed by Quah (1996a, 1997) and already presented in the preceding pages. These will enable us to obtain a fairly accurate assessment of how far the distribution varies when factors relating to issues other than the productivity in the primary sector are introduced into the analysis.

Ever since the pioneer work by Molle, Van Holst and Smit (1980), authors dealing with spatial disparities in per capita income or aggregate productivity within the European setting have placed repeated emphasis on the importance of country-specific factors in regional growth processes (Quah, 1996b; Rodríguez-Pose, 1999; Ezcurra *et al.*, 2005a). It will be of interest, therefore, to explore the role played by the national component in the evolution of the regional distribution of output per worker in the European agriculture throughout the period considered. With this idea in mind, we constructed a conditioned distribution by normalizing the productivity in the agricultural sector of each region according to the mean in the rest of the country to which it belongs, excluding the region in question.

Thus, leaving aside any political and administrative factors, we also considered the possible influence on the regional distribution of agricultural productivity of variables such as the investment per worker in the primary sector, the regional per capita income, or the role of industries directly related with farming activity[24]. For our purposes, we classified the various regions by deciles, taking as reference the average values of the different variables over the whole of the sample period. We then constructed three more conditioned distributions by normalizing in this case the output per worker in the primary sector of each region according to the average agricultural productivity level of

[24]This last variable was proxied by the share in regional value added corresponding to the food, beverages and tobacco industry (sectoral classification NACE-CLIO R17). It should be noted in this respect that, despite our best efforts, we were unable to obtain data at a higher level of sectoral disaggregation to cover the entire geographical and temporal scope of the present study.

those regions in the same decil.

The various conditioned distributions just defined can be intuitively interpreted as that part of the original distribution that remains unexplained by the set of variables considered. For a better understanding of this idea, let us imagine a situation in which the *country effect* had no influence whatsoever on the evolution of the distribution under analysis, so regions where the level of output per worker in the primary sector were lower (higher) than the sample mean, would also be less (more) productive than regions belonging to the same country. In this hypothetical setting, the original distribution would coincide with the conditioned distribution. If, on the other hand, the national component were to play a relevant role, it would be reasonable to expect that less (more) productive regions would register a similar value of the variable analyzed as the mean for the group to which they belong, defined, this time, in political-administrative terms.

To overcome the problems involved in using discrete transition matrices in this setting, we have opted in this paper to use a non-parametric approach based on the estimation of stochastic kernels and contour plots[25]. Before going on to discuss the outcomes obtained, it might be worth clarifying a few points relating to the interpretation of stochastic kernels and contour plots in this context. Within this framework, these instruments provide information concerning the probabilities of transition between the initial distribution and the conditioned distribution, and not between two moments of time as in the previous case. Thus, if the factors considered do not contribute to explain the distribution dynamics, the probability mass should cluster around the main diagonal. If, on the other hand, the variables selected are determinant in explaining the evolution of the distribution analyzed, the density will tend to cluster parallel to the

[25] This type of methodology has been used by Magrini (2004) and Ezcurra *et al.* (2005b) to investigate the causes of regional differences in per capital income in the European Union.

axis corresponding to the initial distribution and around the mean.

[INSERT FIGURE 12]

[INSERT FIGURE 13]

[INSERT FIGURE 14]

[INSERT FIGURE 15]

Figures 12, 13, 14 and 15 present the results obtained when this methodology is used to examine the role of the country to which a region belongs, the investment in the primary sector, and the size of the agrifood sector, in explaining the dynamics of the regional distribution of output per worker in the European agriculture. The empirical evidence provided by the various graphs clearly shows that, unlike the rest of the variables analyzed, the national component and the investment per worker in agriculture play an important role in this context. Thus, the analysis carried out reveals considerable differences between the regional distribution of agricultural productivity in any member State and in the European Union as a whole. Close observation of the graphs in Figure 12, however, allow us to qualify this finding, at least in part. The *country effect* does indeed appear to be more significant among regions with low or medium values of the variable analyzed. At the same time, however, the probability mass at the upper end of the distribution appears to be approaching the main diagonal.

This suggests that the productivity in the agricultural sector is closer to the national mean in regions where output per worker in agriculture is low.

Figure 15, meanwhile, highlights the predominant role played by farm investment in this context. Unlike the national component, however, this variable proves more relevant in regions located at the upper end of the distribution that concerns us, given that, generally speaking, the productivity of agriculture in these regions tends to be on a par with that of other regions making a similar level of investment in the sector.

6 Conclusions

Throughout the preceding pages, we have examined the territorial imbalances in European agriculture between 1980 and 2001. In order to overcome the limitations of conventional convergence analysis, in this paper we have combined the non-parametric approach proposed by Quah (1996a, b; 1997) with a set of techniques adopted from spatial econometrics.

The results obtained show the presence of positive spatial dependence in agricultural productivity. This fact reveals that, with a few specific exceptions, spatially adjacent regions in the European setting register similar values of the variable analyzed. In particular, we have detected the existence of several clusters of regions with similar levels of output per worker in the primary sector, which distinguishes them from the adjacent zones. However, these clusters are not spread throughout the European territory. In particular, they are located in specific areas of the North, Centre and South of the Union.

The level of spatial dispersion in agricultural sector productivity is also found to have remained practically constant between 1980 and 2001. This is due to the fact that higher growth in value added registered by the regions at the upper end of the distribution was offset by a greater loss of employment in the regions less productive.

At the same time, the various density functions estimated show that the probability mass concentrated around the sample mean increased over the course of the twenty-two years considered, largely due to weight loss experienced by the upper end of the distribution. As a consequence, regional bipolarization decreased during the time interval considered, irrespective of the value taken by the parameter of sensitivity to polarization.

We have also analyzed the level of mobility in the distribution under study. The results obtained in this respect suggest a low degree of intradistributional mobility. Therefore, save for a few exceptions, the European regions tend to have maintained their relative positions in terms of agricultural productivity between 1980 and 2001.

Finally, we have explored the role played in the dynamics of the regional distribution of output per worker in the European agriculture by variables such as the country to which a region belongs, the investment per worker in the farming sector, the regional per capital income, or the impact of industries directly related to agricultural activity. The empirical evidence provided in this respect reveals the importance in this context of the national component and the investment in the primary sector. The analysis carried out shows, in particular, that farm productivity appears to be closer to the national mean in regions with relatively low levels of output per worker in the primary sector. Agricultural investment, on the other hand, is more important in those regions at the upper end of the distribution, where farm sector productivity tends to be similar to that of regions with matching levels of investment in the primary sector.

References

Aghevli, B. B. and Mehran, F. (1981): "Optimal Grouping of Income Distribution Data", *Journal of the American Statistical Association* 76, pp. 22-26.

Anselin, L. (1995): "Local indicators of spatial association-LISA", *Geographical Analysis* 27, pp. 93-115.

Anselin, L. and Griffith, D. A. (1988): "Do spatial effects really matter in regression analysis?", *Papers Regional Science Association* 65, pp. 11-34.

Armstrong, H. W. (2002): "European Union Regional Policy: Reconciling the Convergence and Evaluation Evidence", in Cuadrado, J. R. and Parellada, M. (Eds.): *Regional Convergence in the European Union, Facts, Prospects and Policies.* Berlin: Springer-Verlag.

Ball, V. E., Bureau, J. C., Butault, J. P. and Nehring, R. (2001): "Levels of Farm Sector Productivity: An International Comparison", *Journal of Productivity Analysis* 15, pp. 5-29.

Barro, R. y Sala-i-Martin, X. (1991): "Convergence across States and Regions", *Brooking Papers on Economic Activity* 1, pp. 107-182.

Barro, R. and Sala-i-Martin, X. (1992): "Convergence", *Journal of Political Economy* 100, pp. 407-443.

Bulli, S. (2001): "Distribution dynamics and cross-country convergence: a new approach", *Scottish Journal of Political Economy* 48, pp. 226-243.

Canova, F. and Marcet, A. (1995): "The poor stay poor: Non-convergence across countries and regions", Centre for Economic Policy Research, Discussion Paper Series 1265.

Cliff, A. and Ord, J. (1973): *Spatial Autocorrelation.* London: Pion.

Cliff, A. and Ord, J. (1981): *Spatial Process.* London: Pion.

Colino, J. and Noguera, P. (2002): "Patrones estructurales y convergencia interregional en la agricultura europea", *Libro Blanco sobre Agricultura y Desarrollo Rural.* Madrid: Ministerio de Agricultura y Pesca.

Cuadrado, J. R., Mancha, T. and Garrido, R. (2002): "Regional Dynamics in the European Union: Winners and Losers", in Cuadrado, J. R. and Parellada, M. (Eds.): *Regional Convergence in the European Union, Facts, Prospects and Policies.* Berlin: Springer-Verlag.

Dalgaard, C. J. and Vastrup, J. (2001): "On the measurement of σ-convergence", *Economics Letters* 70, pp. 283-287.

Davies, J. B. and Shorrocks, A. F. (1989): "Optimal Grouping of Income and Wealth Data", *Journal of Econometrics* 42, pp. 97-108.

Durlauf, S. N. and Quah, D. (1999): "The new empirics of economic growth", in *Handbook of Macroeconomics*, Volume 1A. Amsterdam: North-Holland.

Esteban, J. M. and Ray, D. (1994): "On the Measurement of Polarization", *Econometrica* 62, pp. 819-851.

Esteban, J. M., Gradín, C. and Ray, D. (1999): "Extension of a Measure of Polarization with an application to the income distributions of five OECD countries", Luxembourg Income Study Working Paper Series 218. Maxwell School of Citizenship and Public Affairs, Syracuse University. (ftp://lissy.ceps.lu/218.pdf).

Ezcurra, R., Gil, C., Pascual, P. and Rapún, P. (2005a): "Regional inequality in the European Union: does industry-mix matter?", *Regional Studies*, forthcoming.

Ezcurra, R., Gil, C., Pascual, P. and Rapún, P. (2005b): "Inequality, polarization and regional mobility in the European Union", *Urban Studies*, forthcoming.

Gil, C., Pascual, P. and Rapún, M. (2002): "Structural change, infrastructure and convergence in the regions of the European Union", *European Urban and Regional Studies* 9, pp. 115-135.

Gutierrez, L. (2000): "Convergence in US and EU agriculture", *European Review of Agricultural Economics* 27, pp. 187-206.

Haining, R. (1990): *Spatial data analysis in the social and environmental sciences.* Cambridge: Cambridge University Press.

Johnson, D. and Evenson, R. E. (1999): "R&D Spillovers to Agriculture Measurement and Application", *Contemporary Economic Policy* 17, pp. 432-456.

Jollivet, M. and Eizner, N. (1996): *LEurope et ses campagnes.* Paris: Presses de Sciences Po.

Kearney, B. (1992): "Rural society, disparities in incomes and alternative policies", in Marsh, J., Green, B., Kearney, B. and Mahé, L. (Eds.): *The Changing Role of the Common Agricultural Policy: the Future of Farming in Europe.* London: Belhaven Press.

Kremer, M., Onatski, A. and Stock, J. (2001): "Searching for prosperity", *Carnegie-Rochester Conference Series on Public Policy* 55, pp. 275-303.

Le Gallo, J. and Ertur, C. (2003): "Exploratory spatial data analysis of the distribution of regional per capita GDP in Europe, 1980-1995", *Papers in Regional Science* 82, 175-201.

Limouzin, P. (1996): *Les agricultures de lUnion europenne.* Paris: Armand Colin.

López-Bazo, E., Vayá, E., Mora, A. and Suriñach, J. (1999): "Regional Economic Dynamics and Convergence in the European Union", *The Annals of Regional Science*

33, pp. 343-370.

Magrini, S. (2004): "Regional (Di)Convergence", in V. Henderson and J. Thisse (Eds.): *Handbook of Urban and Regional Economics*, vol. IV. Amsterdam: North Holland.

Mas, M., Maudos, J., Pérez, F. and Uriel, E. (1994): "Disparidades regionales y convergencia en las Comunidades Autnomas", *Revista de Economa Aplicada* 4, pp. 129-148.

McCunn, A. and Huffman, W. (2000): "Convergence in U.S. Productivity Growth for Agriculture: Implications of Interstate Research Spillovers for Funding Agricultural Research", *American Journal of Agricultural Economics.*

Molle, W., Van Holst, B. and Smit, B. (1980): *Regional Disparity and Economic Development in the Euroepan Community.* Farnborough: Saxon House.

Pace, R. K. and Barry, R. (1997): "Quick computation of spatial autoregressive estimators", *Geographical Analysis* 29, pp. 232-246.

Paci, R. (1997): "More Similar and Less Equal: Economic Growth in the European Regions", *Weltwirtschaftliches Archiv* 133, pp. 609-634.

Paci, R. y Pigliaru, F. (1999): "European Regional Growth: Do Sectors Matter?", in Adams, J. and Pigliaru, F. (Eds.): *Economic Growth and Change. National and Regional Patterns of Convergence and Divergence.* Cheltenham: Edward Elgar.

Pinkse, J. and Slade, E. (1998): "Contracting in space: an application of spatial statistics to discrete-choice models", *Journal of Econometrics* 85, pp. 125-154.

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

Quah, D. (1996a): "Empirics for Economic Growth and Convergence", *European Economic Review* 40, pp. 1353-1375.

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

Quah, D. (1997): "Empirics for Growth and Distribution: Stratification, Polarization and Convergence Clubs", *Journal of Economic Growth* 2, pp. 27-59.

Raymond, J. L. and García-Greciano, B. (1994): "Las disparidades en el PIB per cápita y la hiptesis de convergencia", *Papeles de Economa Espaola* 59, pp. 37-58.

Rey, S. and Montouri, B. (1999): "US Regional Income Convergence: A Spatial Econometric Perspective", *Regional Studies* 33, pp. 143-156.

Rodríguez-Pose, A. (1999): "Convergence or Divergence? Types of Regional Responses to Socio-Economic Change in Western Europe", *Tijdschrift voor Economische en Sociale Geografie* 90, pp. 363-378.

Schimmelpfenning, D. and Thirtle, C. (1999): "The internationalization of farm technology: patents, R&D spillovers, and their effects on productivity in the European Union and United States", *Contemporary Economic Policy* 17, pp. 457-468.

Silverman, B. (1986): *Density Estimation for Statistics and Data Analysis*, Monographs on Statistics and Applied Probability 26, London: Chapman and Hall.

Terrasi, M. (2002): "National and Spatial Factors in EU Regional Convergence", in Cuadrado, J.R. and Parellada, M. (Eds.): *Regional Convergence in the European Union: Facts, Prospects and Policies*, Berlin: Springer-Verlag.

Villaverde, J. (2003) Regional convergence, polarisation and mobility in the European Union, European Integration, 25, pp. 73-86.

Appendix

[INSERT FIGURE A1]

[INSERT FIGURE A2]

[INSERTAR FIGURA A3]

[INSERT FIGURE A4]

Tables and Figures

Table 1: Global spatial autocorrelation tests (standardized values).

Year	Moran's I	p-value	Geary's C	p-value
1980	16.769	0.000	-15.192	0.000
1981	19.873	0.000	-17.311	0.000
1982	20.424	0.000	-18.039	0.000
1983	20.761	0.000	-18.113	0.000
1984	20.627	0.000	-18.374	0.000
1985	20.757	0.000	-18.204	0.000
1986	21.130	0.000	-17.963	0.000
1987	21.348	0.000	-18.509	0.000
1988	19.793	0.000	-17.383	0.000
1989	18.672	0.000	-16.337	0.000
1990	16.258	0.000	-14.542	0.000
1991	14.802	0.000	-13.225	0.000
1992	17.835	0.000	-15.478	0.000
1993	18.009	0.000	-15.114	0.000
1994	17.884	0.000	-15.676	0.000
1995	18.448	0.000	-15.996	0.000
1996	16.918	0.000	-14.427	0.000
1997	15.316	0.000	-13.660	0.000
1998	14.474	0.000	-12.993	0.000
1999	14.322	0.000	-12.773	0.000
2000	14.625	0.000	-12.962	0.000
2001	12.921	0.000	-11.448	0.000

Table 2: Regional disparities in the various sectors.

Sector	*SDlog*	*CV*
Agriculture	0.540	0.462
Industry	0.379	0.429
Services	0.299	0.285
Total	0.328	0.301

Note: Means for the period 1980-2001.

Figure 1: Spatial distribution of agricultural productivity, 1980 (box map).

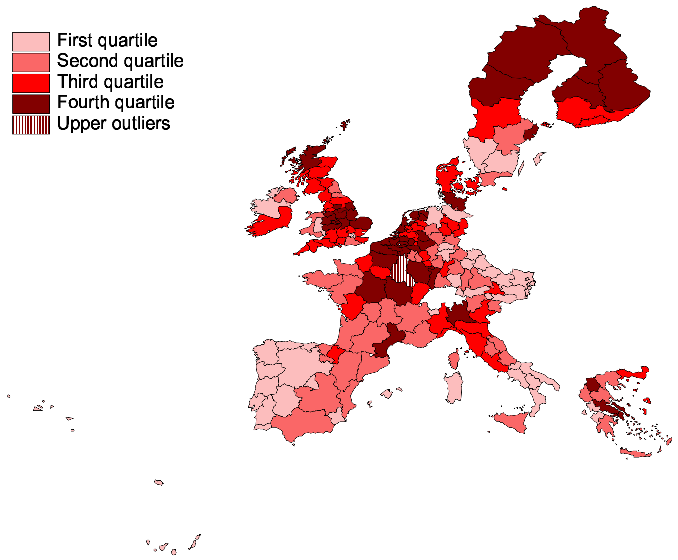

Figure 2: Spatial distribution of agricultural productivity, 2001 (box map).

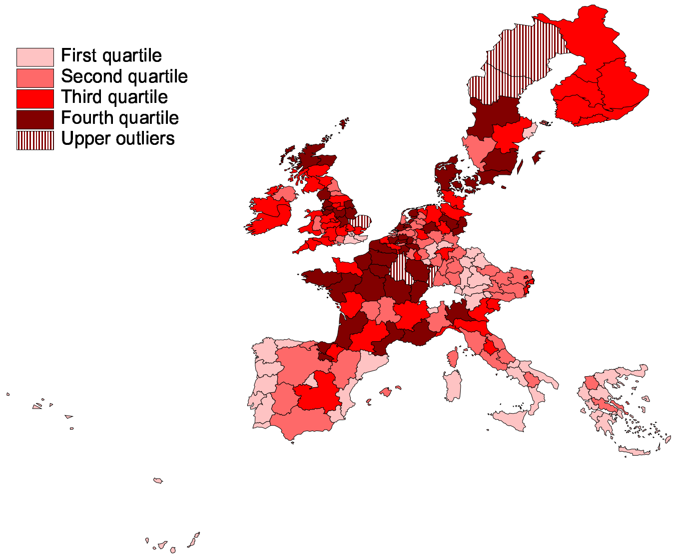

Figure 3: Spatial distribution of local Moran's I, 1980.

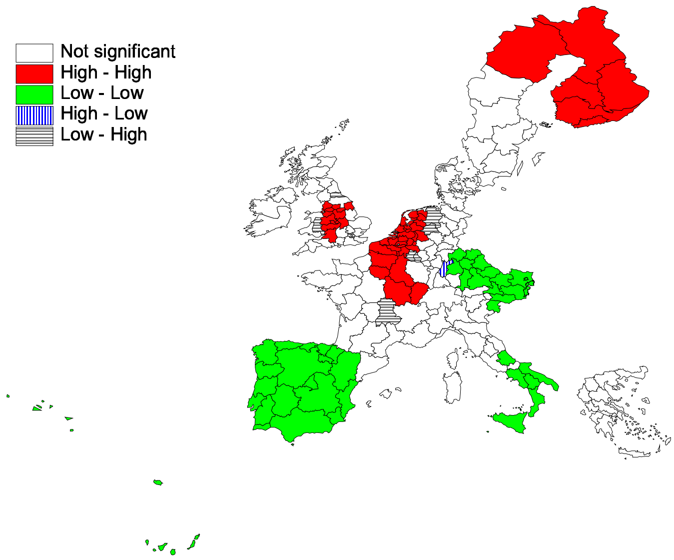

Figure 4: Spatial distribution of local Moran's I, 2001.

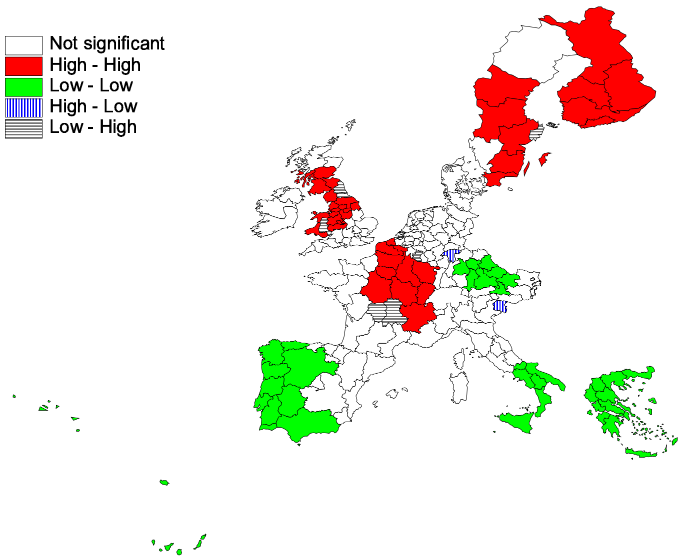

Figure 5: Regional disparities in farm sector productivity.

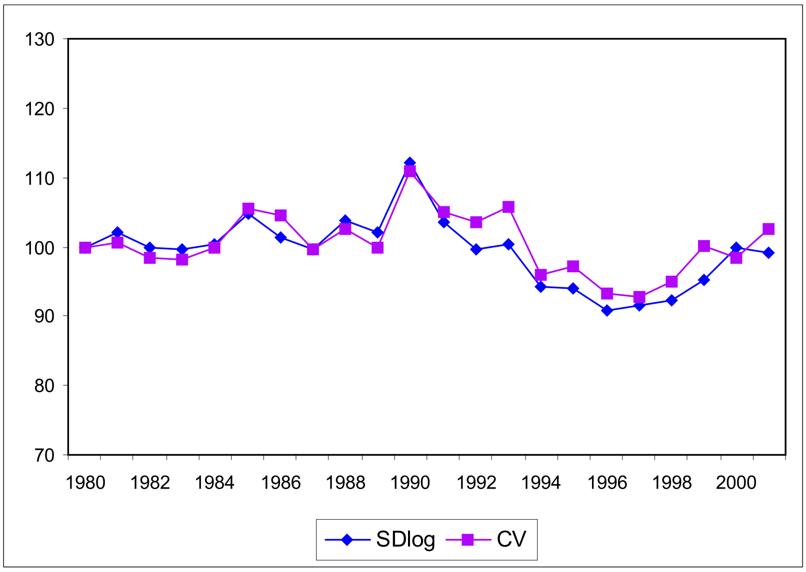

Figure 6: Regional disparities: the role of variations in value added and employment.

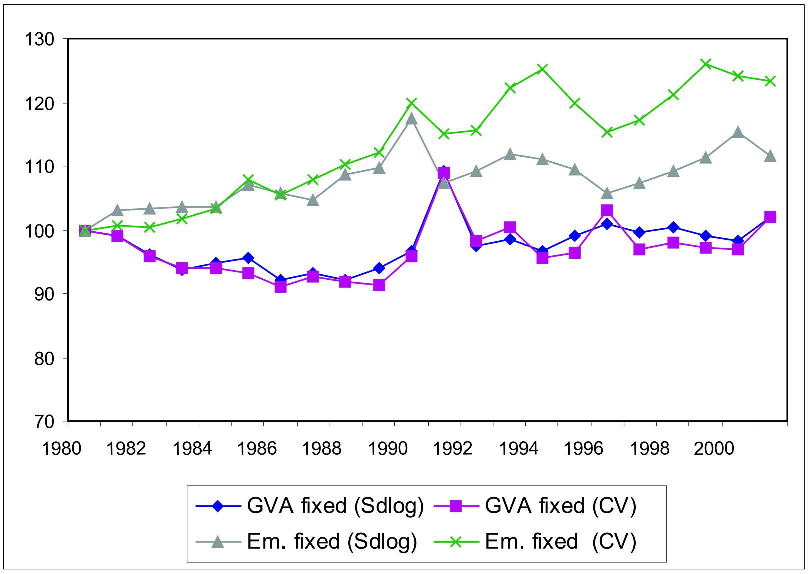

Figure 7: Density functions of the regional distribution of agricultural productivity.

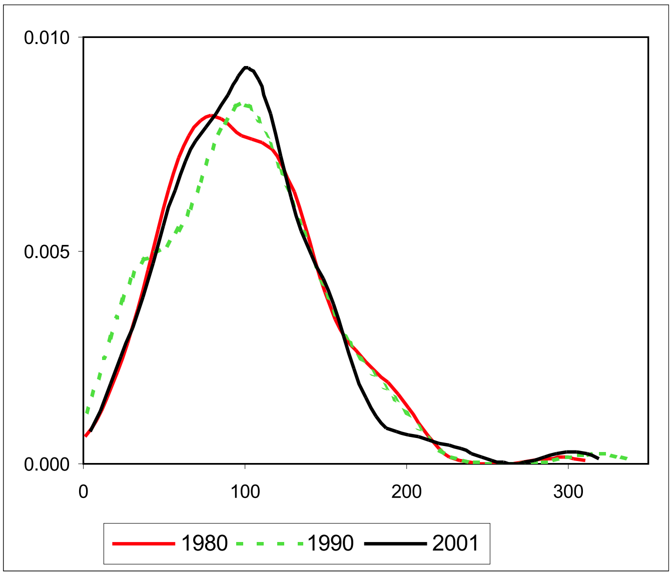

Figure 8: Generalized bipolarization.

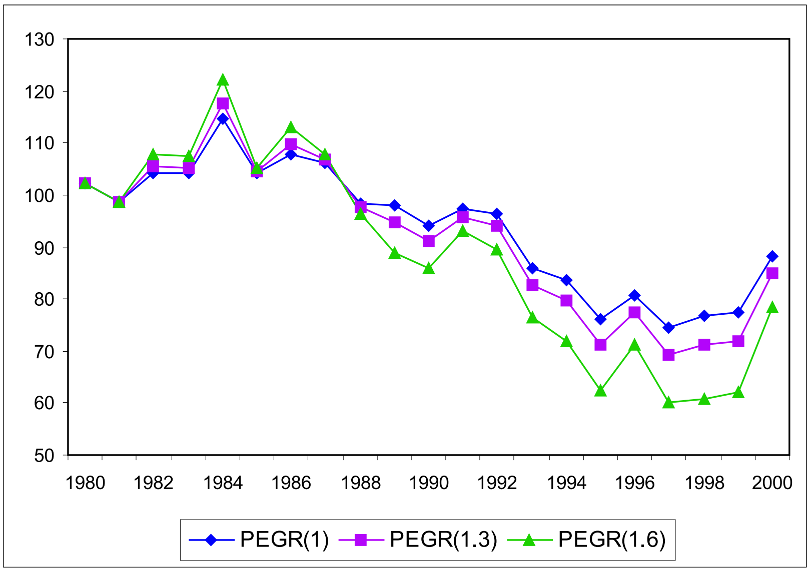

Figure 9: Bipolarization of the simplified representation and internal dispersion.

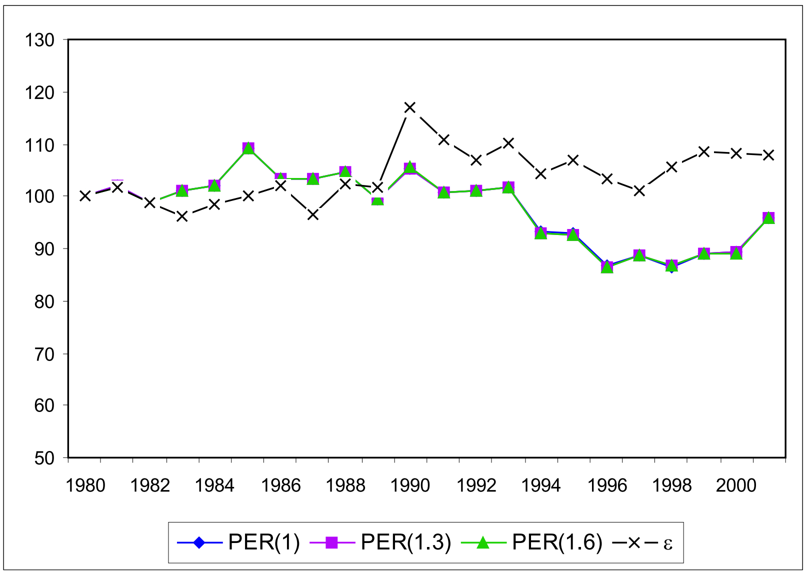

Figure 10: Stochastic kernel and contour plot of the regional distribution of agricultural productivity, 1980-2001.

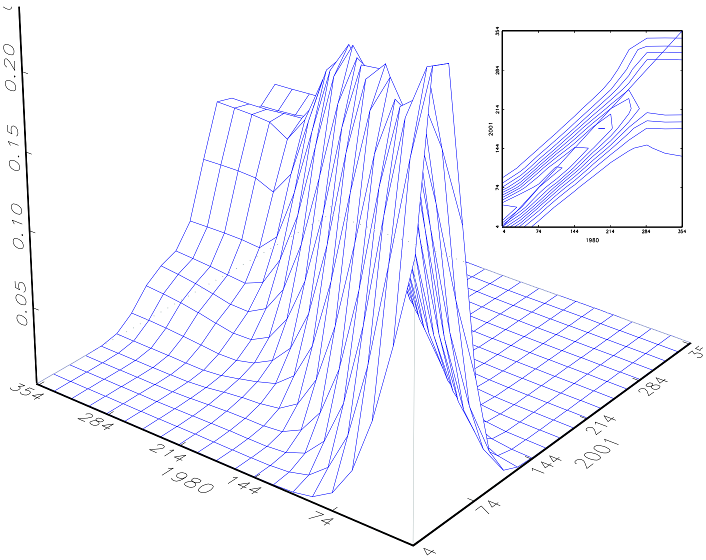

Figure 11: Ergodic distribution of agricultural productivity.

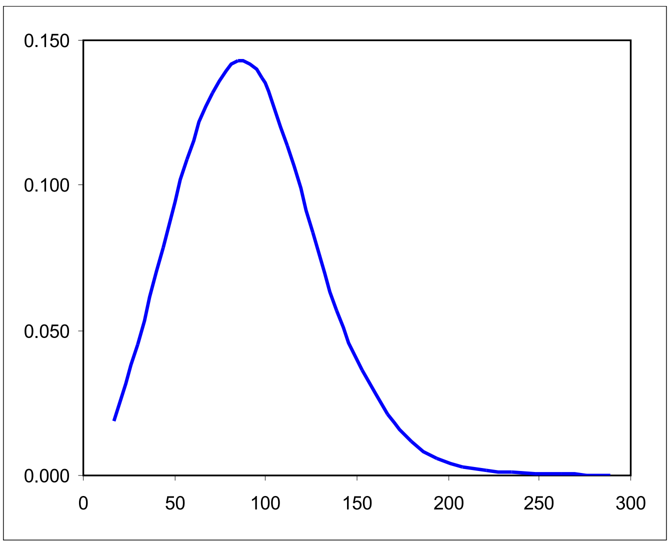

Figure 12: The national component and the dynamics of the regional distribution of agricultural productivity.

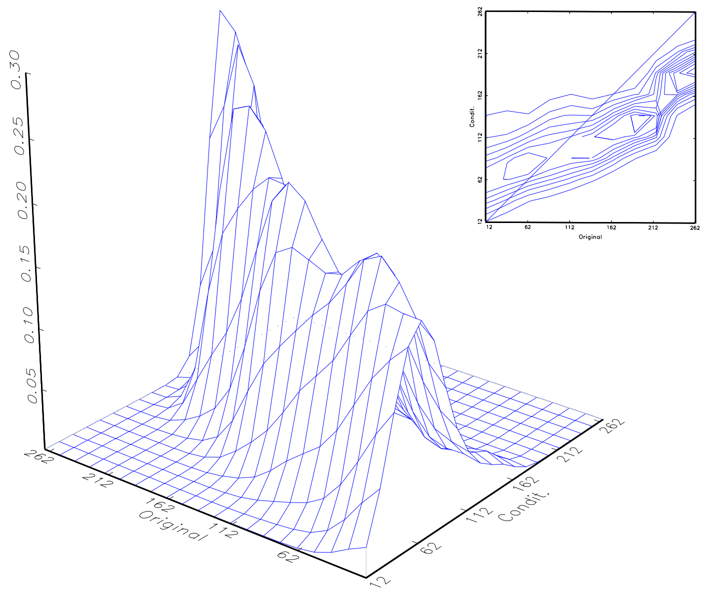

Figure 13: Investment per worker in the sector and the dynamics of the regional distribution of agricultural productivity.

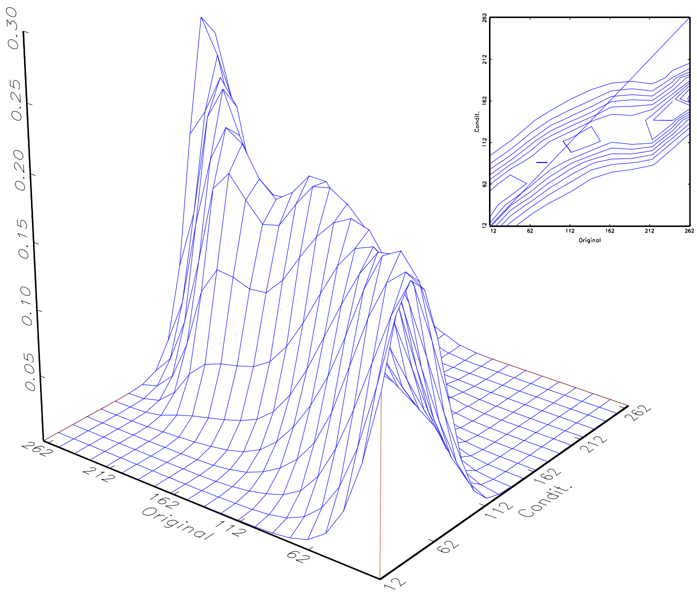

Figure 14: Development level and the dynamics of the regional distribution of agricultural productivity.

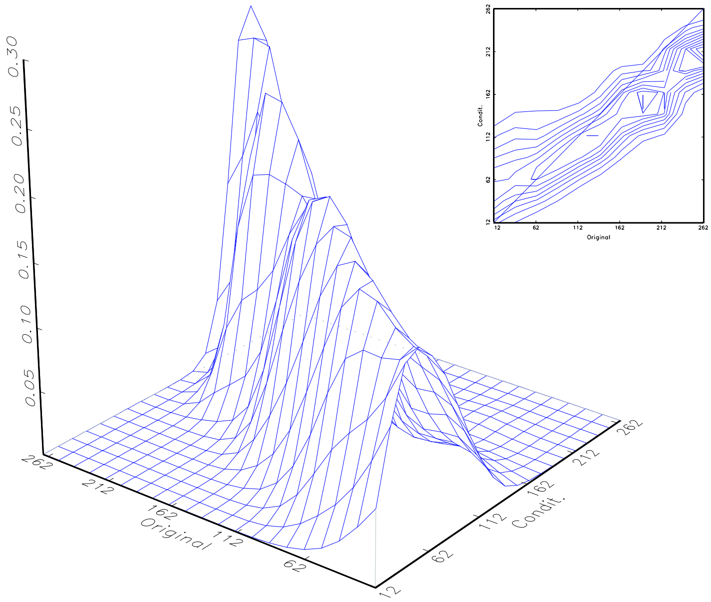

Figure 15: The agrifood industry and the dynamics of the regional distribution of agricultural productivity.

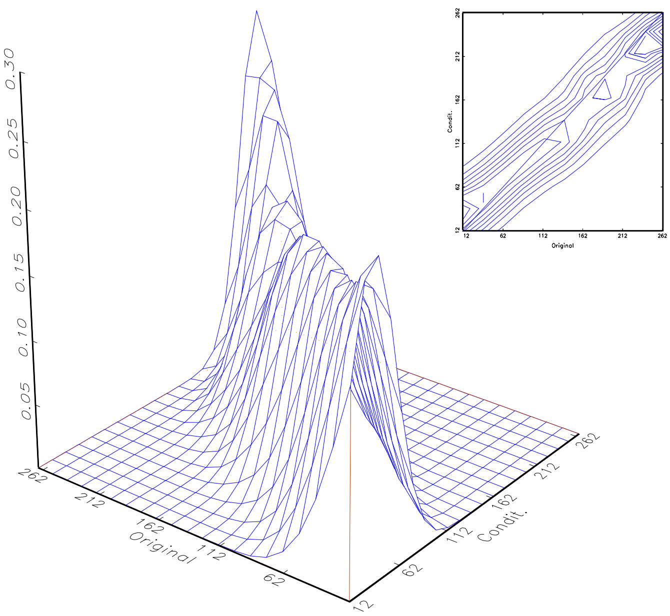

Figure A1: Moran's scatterplot, 1980.

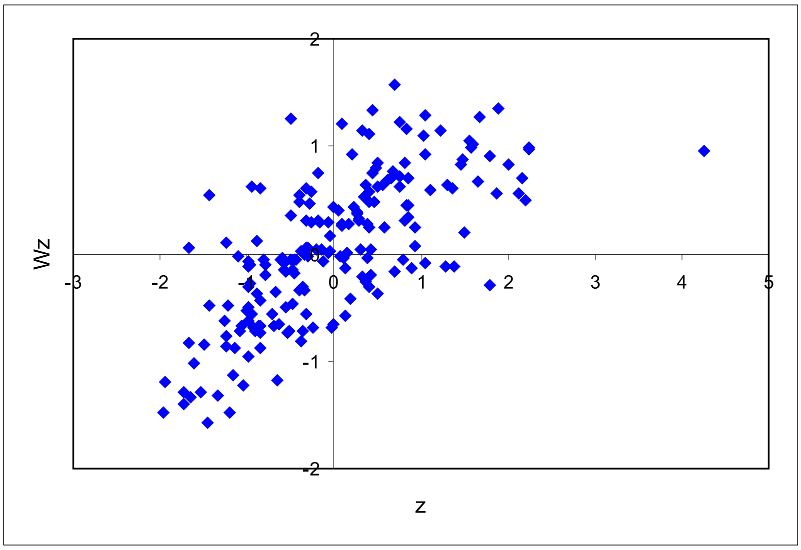

Figure A2: Moran's scatterplot, 2001.

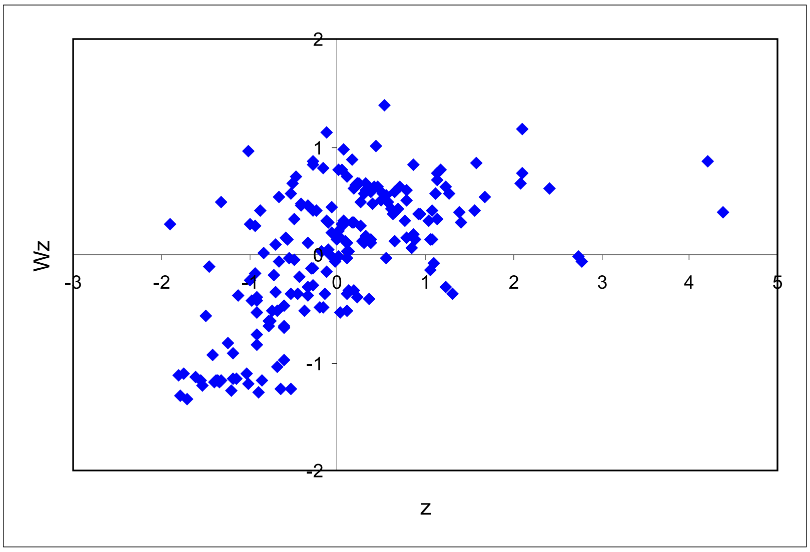

Figure A3: Significance of local Moran's I test, 1980.

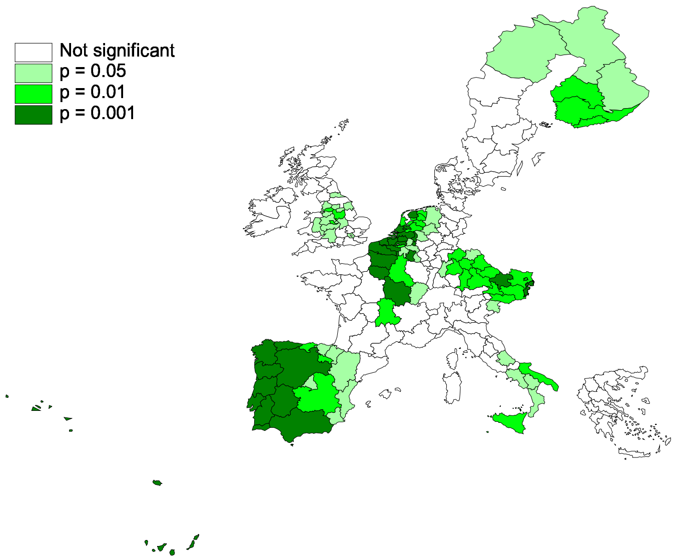

Figure A4: Significance of local Moran's I test, 2001.