Spatial location patterns of Spanish manufacturing firms*

José M. Albert, Marta R. Casanova and Vicente Orts*

Preliminary Version. Please do not quote without permission.

Abstract

In this paper, we evaluate the spatial location patterns of Spanish manufacturing firms and we assess the different tendencies to cluster in each industry. To do this, we use a distance-based method (Marcon and Puech, 2003, Duranton and Overman, 2005), more concretely the Ripley's *K* function, which allows us to treat space as continuous, measuring concentration by counting each firm's average number of neighbours within a circle of a given radius. In this way, $K(r)$ function is used to describe characteristics of the point patterns at different geographical scales at the same time, letting us detect the statistical significance of departures from randomness.

Our approach incorporates an additional improvement. By means of the software employed, 'R', we can apply border corrections adequately in any irregular polygonal shape; therefore, we introduce a polygonal envelope, which allows us to improve the delimitation of our area of study, Spain, avoiding the nuisance of empty spaces.

We apply this method to Spanish manufacturing sectors at two-digit level and we realise that results depend on the benchmark employed. In fact, whether we use 'complete spatial randomness' as benchmark, every sector analysed presents significant concentration whatever the distance of the radius we consider. However, as is well known, this type of analysis is sensitive to the benchmark considered, thus we also use as benchmark the whole of manufacturing sectors. In this case, we find patterns of localization completely different to the previous ones; appearing dispersion in some sectors relative to all manufacturing firms. Therefore, by means of this method we can know characteristic features of the pattern of localization of every manufacturing sector, like whether concentration or dispersion exists, which is its intensity and at which distance, or geographical scale, we obtain its highest level.

Keywords: distance-based method, polygonal envelope, Ripley's *K* function, Spanish manufacturing firms, spatial location patterns.

JEL classification: C15, C40, C60, R12.

---

* We gratefully acknowledge financial support from the Spanish Ministry of Education and Science, project SEJ2005-08764/ECON.

* Department of Economics, Universitat Jaume I.
Contact: Marta R. Casanova, Department of Economics, Universitat Jaume I, Avda. Sos Baynat s/n, 12071 Castellón, (mroig@eco.uji.es).

1. Introduction

Economic activity is heterogeneously distributed across space and this distribution does not take into account, necessarily, administrative frontiers or country boundaries. Some traditional examples of this geographic concentration are high-tech firms in Silicon Valley, the auto industry in Detroit, the carpet industry in Dalton or, dating back, the textile industry in Lancashire. We as well find, in Spain, the tile industry in Castellón and the leather and footwear industry in Elche.

Heterogeneity of spatial distribution of activity can be caused by multiple and very different factors and substantial literature has focused on this topic. The causes of concentration of economic activity have been evolving along time and the propensity of firms to agglomerate in space has changed as regions have become more integrated. Marshall (1890) introduced the notions of three positive externalities because of the fact that firms were located next to other firms: benefits coming from localized information and knowledge spillovers, from a specialised local labour market, as well as lower costs coming from a closer supply and demand (backward and forward linkages associated with large local markets).

The 'new economic geography', driven by Krugman[1], stressed the role of historical accidents and "agglomerative forces", emphasizing the increasing returns to scale. New economic geography underlined also the importance of trade costs towards the spatial concentration, the relationship between market integration and industrial concentration and considered the possibility of there being 'multiple equilibriums', as we can see in Fujita *et al*. (1999), Puga (1999, 2002) and Ottaviano and Puga (1998). Mobile factors, as firms or workers, locate for profiting from a higher productivity and this creates a positive feedback. Firms and workers go where productivity is high and tend to raise it, creating an uneven distribution of activity.[2]

Not only the *determinants* of geographic concentration worry the economists, but also another aspect, *how to measure* this mentioned spatial distribution. Therefore, this paper gets into this second aspect and its main objective will be to measure the spatial location patterns of Spanish manufacturing sectors.

Many are the employed techniques to assess geographic concentration in the literature, but the most usual measures are the indices of Herfindahl, Gini or Ellison and Glaeser (1997).

---

[1] See, for example, Krugman (1991a, 1991b).

[2] For further details and to go into the forces that are changing the spatial distribution of activity, see Overman, Redding and Venables (2003) and Venables (1995, 2006).

Herfindahl index is a measure of industry concentration equal to the sum of the squared market shares of the firms in the industry. The *Gini index* is a measure that studies whether distribution is or not concentrated, in other words, measures the uniformity of a distribution. It compares the geographic patterns of employment, income or wealth distribution for one industry and for the aggregate. This index cannot distinguish whether concentration of activity is due to a few huge firms in a specific area or due to many small firms in the same specific area. Krugman (1991a), Brülhart (2001) or Amiti (1997), among others, used this index to measure how the economic activity was distributed across space. Finally, *Ellison and Glaeser (1997)* tried to solve this inconvenient comparing concentration resulting from a random and independent location of firms with the real geographic concentration of an industry, taking into account, this time, the size of firms. Thus, this index let us compare concentration between industries or concentration of a given industry in different countries and, many authors as Devereux *et al.* (2004), Rosenthal and Strange (2001), Maurel and Sédillot (1999) or Callejón (1997) used it to measure geographic concentration of activity in their respective countries, UK, US , France and Spain.

Revising the literature about the spatial distribution of activity in Spain, we realise that many other authors, apart from Callejón, are interested in this topic, emphasizing Viladecans (2001), Paluzie *et al.* (2004) and Alonso-Villar *et al.* (2001, 2003 and 2004), among others. Other studies, as Paluzie *et al.* (2001) and Tirado *et al.* (2002), try to analyse, as well, the determinants of the localization of the industrial activity in Spain. Should be known that most of this Spanish studies use the Gini index to measure geographic concentration of activity, but Alonso-Villar *et al.* (2001) introduce in their analysis the index proposed in Maurel and Sédillot (1999)[3] and Viladecans (2001) uses an econometric spatial index, called Moran's I statistic of spatial association.

These methods used until now, in Spain and in other countries, to measure geographic distribution of activity, have a common characteristic: their conclusions may differ considerably according to the spatial scale chosen. In other words, empirical results remain, in some degree, inconclusive with patterns detected depending on the geographical scale considered. Thus, Viladecans (2001) uses more than one geographic level in her analysis[4], the same as Alonso-Villar *et al.* (2001) do[5], to assess which would be the administrative unit more suitable in every analysis.

---

[3] Moreover, compares it with other two concentration indices to obtain results that are more robust.
[4] She uses municipal and provincial level, corresponding to NUTS 5 and 3.
[5] They use provincial and regional level, corresponding to NUTS 3 and 2.

Moreover, all these traditional concentration indices *treat space as discrete*. In fact, these indices restrict the spatial distribution just to *one scale*, analysing the distribution of activity over discrete geographic units, not having why to coincide these ones with the relevant scale from the economic point of view.

The main purpose of this paper is to avoid the inconvenience of *geographic scale*, as Marcon and Puech (2003a) or Duranton and Overman (2005) did. In this way, we will *treat space as continuous* to obtain a proper analysis of spatial location patterns, instead of sticking to administrative scale data. To do this, we must employ a specific method that satisfies these essential requirements, letting us know and compare the concentration intensity for every spatial scale. Thus, just a *distance-based method* possesses the necessary requirements to achieve our aim, being also essential a suitable data to consider space as continuous.

Consequently, the distance-based method we are going to use to measure spatial distribution of activity in Spain is Ripley's *K* function[6] and presents great advantages compared to traditional concentration indices. In fact, by means of this method employed, we can know whether concentration exists, which is its intensity and at which distance we obtain its highest level.[7]

Finally, before showing how is structured our paper, we must emphasize other two aspects that differ our empirical analysis from others. First, our location of firms is precise, because we know the UTM[8] coordinates of every establishment. In this way, we employ these coordinates to situate firms, as a dot in the map, obviating administrative frontiers. On the contrary, Marcon and Puech (2003) and Duranton and Overman (2005) know the postcode of each manufacturing firm and associate geographic coordinates for all postcodes, obtaining a location error between 100 m and about 2 km. Second, we incorporate an innovative technique to improve the delimitation of our area of study, substituting the rectangular shape, used by other authors, for a polygonal shape. Thus, Marcon and Puech (2003) limited their empirical analysis to rectangular areas because they affirmed that '*complexity depends on the shape of the area under study*'. However, in our case, by means of the software employed, 'R'[9], we can apply border corrections adequately in any irregular polygonal shape.

The paper is organised as follows. In Section 2, we outline the methodology employed, Ripley's *K* function. In Section 3, we describe our data. In Section 4, we present the main results achieved, discuss them and outline the key

[6] For further details, see Ripley (1976, 1977, 1979).
[7] Marcon and Puech (2003) have previously used this concentration index, in France, to measure the spatial distribution of economic activity.
[8] Universal Transverse Mercator.
[9] This software is downloadable from the following website: http://www.r-project.org/.

improvements incorporated into the measuring method of spatial concentration used. Finally, Section 5 contains the main conclusions reached.

2. Methodology

The Ripley's *K* function is a distance-based method that measures concentration by counting each firm's average number of neighbours within a circle of a given radius, calling 'neighbours' all the points situated at a distance equal or lower than the radius (r).

$K(r)$ function describes characteristics of the point patterns at many and different scales, depending on the value of "r" we take into account, that is,

$$K(r) = \frac{1}{\lambda N} \sum_{i=1}^{N} \sum_{j=1, i \neq j}^{N} w_{ij} I(d_{ij})$$

being r the different values of the radius we consider, w_{ij} the weighting factor to correct for border effects, d_{ij} the distance between two firms, N the total number of points observed in the area of the study region and λ the density. This density is defined as follows, being A the area of the study region.

$$\lambda = \frac{N}{A}$$

If we substitute the value of λ in the previous $K(r)$ function, we obtain the definitive expression of $K(r)$ value. Therefore, this function can be defined as the average number of neighbours in a radius (r) divided by the density (λ) and describes the cumulative frequency distribution of all point-to-point distances.

$$K(r) = \frac{A}{N^2} \sum_{i=1}^{N} \sum_{j=1, i \neq j}^{N} w_{ij} I(d_{ij})$$

As we have said before, a point is considered 'neighbour' if the distance between i^{th} and j^{th} points is lower than r; then, the indicator function $I(d_{ij})$ takes value 1. Whereas if this distance, (d_{ij}), is higher than r the point is not considered neighbour and $I(d_{ij})$ takes value 0. Moreover, the weighting factor (w_{ij}) provides the border correction, having this factor the value of 1 when the circle centred at i and passing through the point j, with a radius of d_{ij}, is completely inside the study area. However, if part of the circle falls outside the study area, w_{ij} will be the proportion of that circle that falls in the area of study.

These border-effect corrections should be incorporated in our analysis to evaluate correctly the geographic concentration. In fact, if we remove the term w_{ij} from $K(r)$, the value of this function will be decreasing at long distances.

The use of a benchmark is necessary in our analysis because we need to have a reference when we study the location patterns of Spanish manufacturing sectors. Additionally, the knowledge of the characteristics of this reference let us identify how our point pattern differs from the benchmark and when is concentrated or dispersed relative to this one. Therefore, we are going to analyse two scenarios regarding the benchmark.

On the one hand, in first scenario we use C*omplete Spatial Randomness* (CSR) as a benchmark. This first analysis consists in calculating the K function of the real point pattern (empirical K value) and comparing it with the K function associated to a point pattern generated by a Poisson process with the same intensity, or number of points, (theoretical K value).[10] The theoretical value of $K(r)$ will be πr^2 as long as we assume CSR. In this way, we define $M_{CSR}(r)$ as the value that quantifies the difference between the empirical K value and the theoretical K value, and is characterized as follows,

$$M_{CSR}(r) = K(r) - \pi r^2$$

Whether the empirical K value, $K(r)$, is higher than the theoretical K value, πr^2, this indicates concentration of our point pattern distribution, since the real density is greater than the benchmark's, in the aggregates. Lower values indicate dispersion and whether $K(r)$ is equal to πr^2, it means that our points are independently distributed.

The following figures try to illustrate the above-mentioned affirmations, using CSR benchmark.

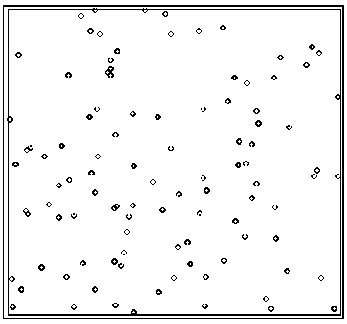


Figure 1. Independent distribution.

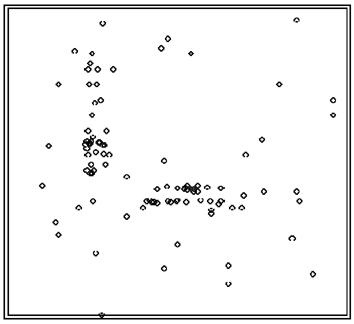


Figure 2. Concentrated distribution.

[10] Marcon and Puech (2003) used the normalised function, $L(r)$, because they considered that '*a practical limitation of Ripley's K function is the need to compare any value to* πr^2'. However, this is not an inconvenient for us, since the statistical software employed, 'R', calculates this theoretical value (πr^2) for each radius and we can compare it directly with real k-value.

Figure 1 shows an independent distribution of points and Figure 2 shows a concentrated distribution, both having one hundred points and the same area (1000 km^2). Figure 3 and Figure 4 give us useful information about the point patterns shown in Figure 1 and 2, as whether appears concentration or dispersion, which is its intensity and at which distance we obtain its highest level.

As we can see, two lines appear in Figure 3 (the graphic on the left). The dashed line represents the empirical *K* value, that is to say, the *K* value of the observed point pattern in Figure 1, and the continuous line takes a value of πr^2, since this represents the CSR benchmark. We realise that the values of these two lines are almost the same, whatever the distance of the radius we take into account ($K(r) \approx \pi r^2$). The graphic on the right of Figure 3 shows us the *M* curve, which is the difference between the empirical and the theoretical *K* value. Therefore, we can affirm that an independent and random distribution of points produces an almost flat *M* curve, around the zero value.

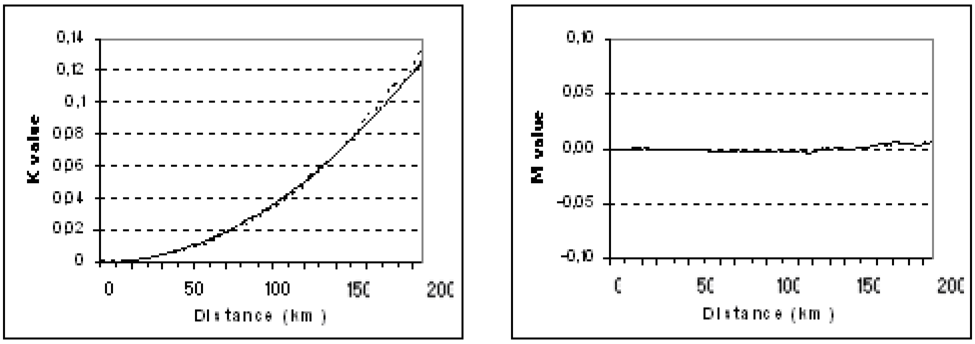


Figure 3. Ripley's *K* function (theoretical and empirical) and *M* function corresponding to the point pattern from Figure 1.

Graphs in Figure 4 give us the same information that the previous ones, the *K* values and the difference between them, the *M* value. However, the interpretation of them is completely different, since they are the result of a concentrated distribution of points. We observe, on the left graph, that $K(r) > \pi r^2$ at all distances of "*r*" considered, meaning this that the point pattern in Figure 2 presents concentration. Moreover, viewing *M* curve we can know at which distance the highest level of concentration is reached. In this way, *M* value indicates us whether the point pattern is concentrated or dispersed, depending its positive or negative value, adding information about at which distance positive or negative peaks are found, being possible to find more than one. Therefore, owing to the *M*-value graph provides us information that is more detailed, we will use this graph along the paper to analyse the spatial location patterns of Spanish manufacturing firms.

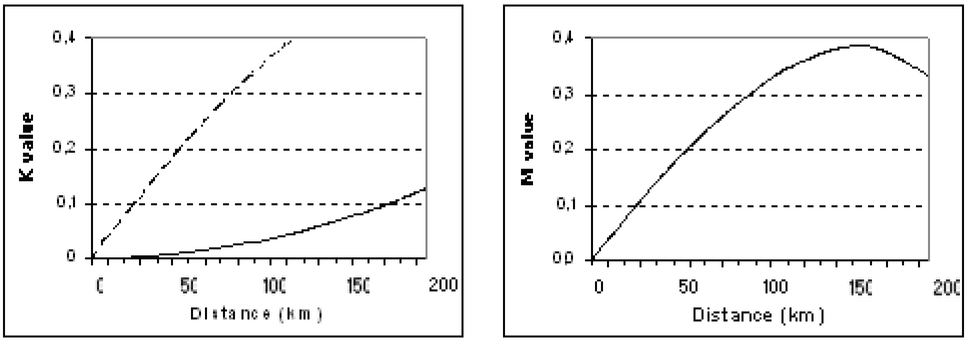


Figure 4. Ripley's *K* function (theoretical and empirical) and *M* function corresponding to the point pattern from Figure 2.

Once explained the first scenario, we realise that although CSR benchmark let us verify that firms are not distributed randomly and many authors use it, it is not a realistic benchmark because economic activity does not locate in a random and independent way. Economic activities are spatially concentrated because of dissimilarities in such natural features as mountains, rivers or harbours. In this way, CSR as a benchmark does not take into account the first nature or the second nature, understanding second nature as the endogenous mechanisms leading to agglomeration.

Consequently, in the second scenario, we use the whole of manufacturing industries as a benchmark, taking into account/considering the point pattern formed with the 43.087 establishments. Like this, we can compare the spatial distribution of every sector with the overall tendency of manufacturing to agglomerate.

As we already know, this benchmark takes into account the first nature and can explain some aspects of the second nature, thus it is a more appropriate benchmark than the CSR.

$$M_{TM}(r) = K(r) - K_{TM}(r)$$

In this way, $M_{TM}(r)$ is the difference between the *K*-value of every considered sector and the *K*-value of the total manufacturing, appearing concentration in a particular sector whether its *K*-value is higher than the total manufacture *K*-value. In such a case, we affirm that this sector is concentrated relative to the whole of manufacturing firms.

Regarding to the area of study, Marcon and Puech (2003) did not analyse the French whole country, but an industrial area of 40 x 40 km around Paris and a larger French rectangular area of 550 x 630 km. Their reason by not using the whole country was the increasing complexity when simulating random points inside the area and when correcting the border-effects on convex shapes[11]. Therefore, these inconveniences limited their empirical analysis to rectangular areas. Nevertheless, Duranton and Overman (2005) did not have this shortcoming, since border-effect corrections were not necessary in their test, simplifying this fact their empirical analysis.

In our analysis, we improve the delimitation of the area of study, substituting the rectangular shape, used by other authors, for a polygonal shape. Thanks to the statistical software employed, 'R', we can apply border corrections, adequately, in any irregular polygonal shape, simplifying the treatment of border effects. In Figure 5, we can observe the polygonal shape that delimits very accurately our territory and envelopes the area of study, avoiding the nuisance of empty spaces where no firms are found, represented by the oblique lines.

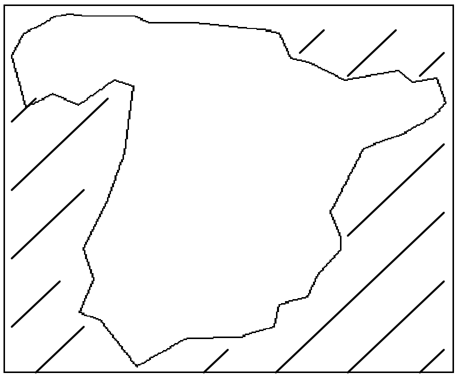


Figure 5. Polygonal shape envelope.

Now, to evaluate the statistical significance of departures from randomness, in a robust way, we should construct a confidence interval. The traditional technique to construct this confidence interval is the Monte Carlo method, generating a large number of independent random simulations. Should be highlighted that the construction of the confidence interval will be different in both scenarios. In first scenario, we simulate Poisson patterns with the same number of points than the real distribution of every sector. In addition, these simulations use the same area than the observed point pattern, being executed inside the above-mentioned polygonal area. In second scenario, we as well simulate random distributions with

[11] They explicitly said "It was impossible to use the whole of France because of border-effect corrections".

the same number of points than the real distribution but, this time, location of points is restricted to the sites where we can currently find firms from the whole manufacturing. Thus, to obtain the confidence interval in the second scenario we should take away the total manufacture *K*-value from the simulations *K*-value. Lastly, we just can add that both of them are generated by using 100 simulations and both allow us to reject the non-significant values, choosing a 95% confidence interval.

We should append that this methodology does not allow us to obtain robust results when the value of "*r*" is, approximately, higher than 25% of the largest distance of our area of study. This inconvenient is overcome in our analysis, since we restrict it to 200 km, a value of the radius that represents a quarter of the largest distance of the area of study.

Finally, must be emphasized that only Marcon and Puech (2003) and Duranton and Overman (2005) have used, until now, distance-based methods to assess the geographical concentration of activity. Methodology of both tests is similar, but not identical, and we can find some differences whether we analyse them in depth.

On the one hand, Duranton and Overman (2005) emphasized that their test fulfils all the five requirements that any concentration measure should satisfy[12]. However, the Duranton and Overman's method has a disadvantage, with regard to Ripley's *K* function, not being possible to quantify the concentration or dispersion but only detecting the proportion of sectors that are concentrated.

On the other hand, whether we revise Marcon and Puech (2003), we realise that *K* function does not satisfy two of the five above-mentioned requirements, as M_{CSR} function, since it does not control for the overall agglomeration of manufacturing or for the industrial concentration.

Now, looking over the properties of our concentration measure and once improved the benchmark employed, we realise that M_{TM} function fulfils all the five requirements established by Duranton and Overman and allows us to get a measure that quantifies the concentration or dispersion of all our point patterns.

We achieve the comparability across sectors and the control for the overall agglomeration of manufacturing considering as the benchmark the whole of manufacturing firms, not being our concentration measure sensitive to the number of firms and locating randomly the firms in our hypothetical sectors across the existing establishments of the total manufacture. The third requirement, the control for industrial concentration, is satisfied considering hypothetical sectors

[12] As underlined by D&O (2005), (1) it is comparable across industries, (2) controls for the overall agglomeration of manufacturing, (3) controls for industrial concentration, (4) is unbiased with respect to scale and aggregation and (5) gives an indication of the significance of the results.

with the same number of firms. Fourth, as we use a continuous distance to measure spatial concentration and not administrative scale, we can know characteristic features of the patterns of localization at different scales and our test is unbiased with respect to scale and aggregation. Lastly, the statistical significance is as well satisfied, since the confidence interval allows us to know that the observed distribution is significantly different from randomness.

3. Data

Our empirical analysis uses current establishment level data, for the year 2007, from Analysis System of Iberian Balances database[13], which contains detailed information about Spanish and Portuguese companies. We restrict our database to Spanish manufacturing establishments, using the National Classification of Economic Activities[14] and analysing sectors at two-digit level. Besides, we add another two requirements to our database. First, we make that our database contains only intercontinental Spanish manufacturing firms, without including firms from Canary and Balearic Islands, Ceuta and Melilla. Second, we restrict our analysis just to firms employing at least ten workers. Finally, once applied these requirements, our database contains exactly 43.087 observations, or firms.

The comparison between our restriction, related to the number of employees, and Marcon and Puech (2003)'s restriction must be considered. Actually, they use French manufacturing firms employing at least twenty workers. This difference, concerning the number of employees, is because of the fact that SME (small and medium-sized enterprises) predominate in Spain. Therefore, too many firms would be left out if we just considered those with twenty or more than twenty workers, as Marcon and Puech did.

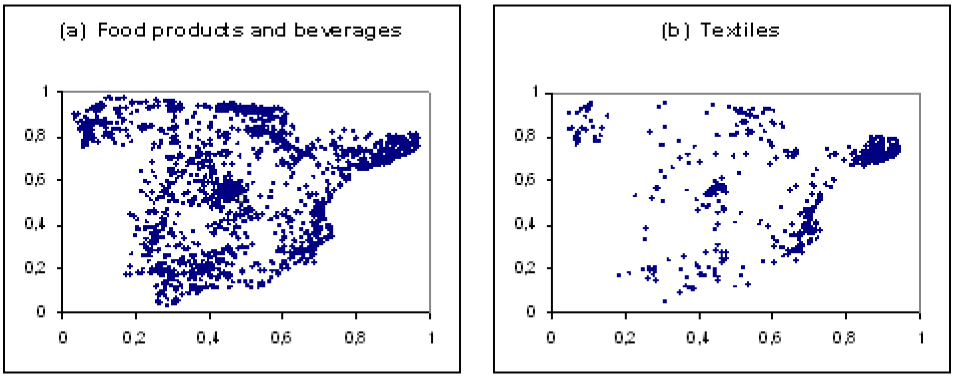
Spanish manufacturing activities are classified into 23 sectors according to 'NACE 93 - Rev. 1' and these are the following ones: (15) Food products and beverages, (16) Tobacco products, (17) Textiles, (18) Wearing apparel and dressing, (19) Tanning and dressing of leather, (20) Wood and products of wood, (21) Pulp, paper and paper products, (22) Publishing, printing and recorded media, (23) Coke, refined petroleum products, (24) Chemical and chemical products, (25) Rubber and plastic products, (26) Other non-metallic mineral products, (27) Basic metals, (28) Fabricated metal products, (29) Other machinery and equipment, (30) Office machinery and computers, (31) Electrical machinery, (32) Radio, televisions and other appliances, (33) Instruments, (34) Motor vehicles and trailers, (35) Other transport equipment, (36) Furniture and other products, (37) Recycling.

[13] SABI
[14] NACE 93 - Rev. 1

Table A1, situated in *Appendix 1*, shows us a brief descriptive analysis about the above-mentioned sectors including additional information, as the number of firms, the number of employees, or the technological intensity of each one. Thus, as we can see in this table, there are great differences in the number of firms, depending on the sector we refer. In fact, three of the twenty-three sectors considered, (16) Tobacco products, (23) Coke, refined petroleum products and (30) Office machinery and computers, will not be analysed because they are too small alluding to their number of establishments.

For every establishment we know its *precise location* through geographic coordinates (longitude and latitude). Each firm has been localized in the same place as in reality; in this way, our margin of error for any firm's location is minimum. These geographic coordinates are transformed into UTM coordinates or, also named, flat coordinates. This transformation has been carried out by means of the method proposed by Morton (2003). This procedure converts Latitude and Longitude coordinates to Easting and Northing coordinates, on a Transverse Mercator projection; being the UTM coordinates expressed in metres. We should highlight that the construction of this system let us move away from the Equator with hardly distortions, because any point is far away from the central meridian of its zone. For this reason, it appears the great advantage of the UTM system, being able to localize any firm in a very accurate way.

Then, in Figure 6, we show the spatial distribution of four Spanish manufacturing sectors (15, 17, 20 and 35). Here, each dot corresponds to an establishment and we can realise that come into view great differences in the spatial distributions of these sectors and in the total number of dots distributed.



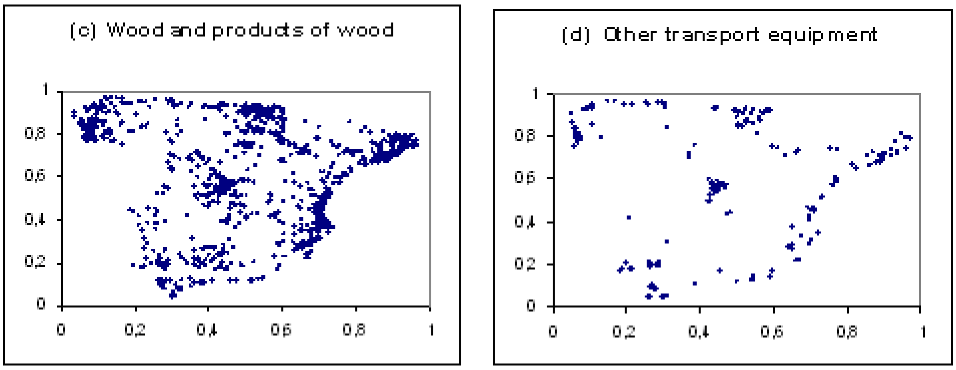


Figure 6. Maps of spatial distribution of firms.

Given that the spatial location of Spanish manufacturing firms is the key information to construct our database, it is obvious that our area of study should be quite similar to the Spanish shape. It is as well obvious that will be unproductive to analyse a rectangular shape, as in previous studies, because Spanish shape is not regular as a rectangle. Therefore, will be necessary the construction of a polygonal area of study by means of a polygonal envelope, already mentioned before. The area of this envelope has similar measures to Spanish territory, since has been built by means of the union of thirty-five points of the Spanish territory perimeter.

4. Empirical Results and Discussion

In the exposition of the methodology above, we have differentiated two scenarios. Consequently, this sequence will be the same in the discussion of our results, analysing first those results coming from 'CSR benchmark' and, after that, those coming from 'TM benchmark'.

In 'Table 1' appears summarised the results obtained from M_{CSR} function. All sectors are included in this table (excepting 16, 23 and 30), although we keep using the four sectors presented before to illustrate our results (see Figure 7).[15]

Graphics, in this figure, show us how activity is distributed in space. Whether we observe them, we can find different location patterns between the sectors presented (15, 17, 20 and 35). For example, sector 17 presents a higher degree of concentration than the rest, and this concentration is reached at a different spatial scale. Actually, sector 35 reaches its maximum concentration peak at a lower distance than sector 17. However, the location pattern of these two sectors is similar between them, and different from the other two sectors, presenting an

[15] Should be highlighted that, in *Appendix 2*, appear gathered the whole of sectors analysed.

initial increase in the activity concentration, reaching a maximum peak and decreasing when the distance of the radius makes higher.

The whole of the differences found in the location patterns of every sector will be analysed, in great detail, in the paragraphs.

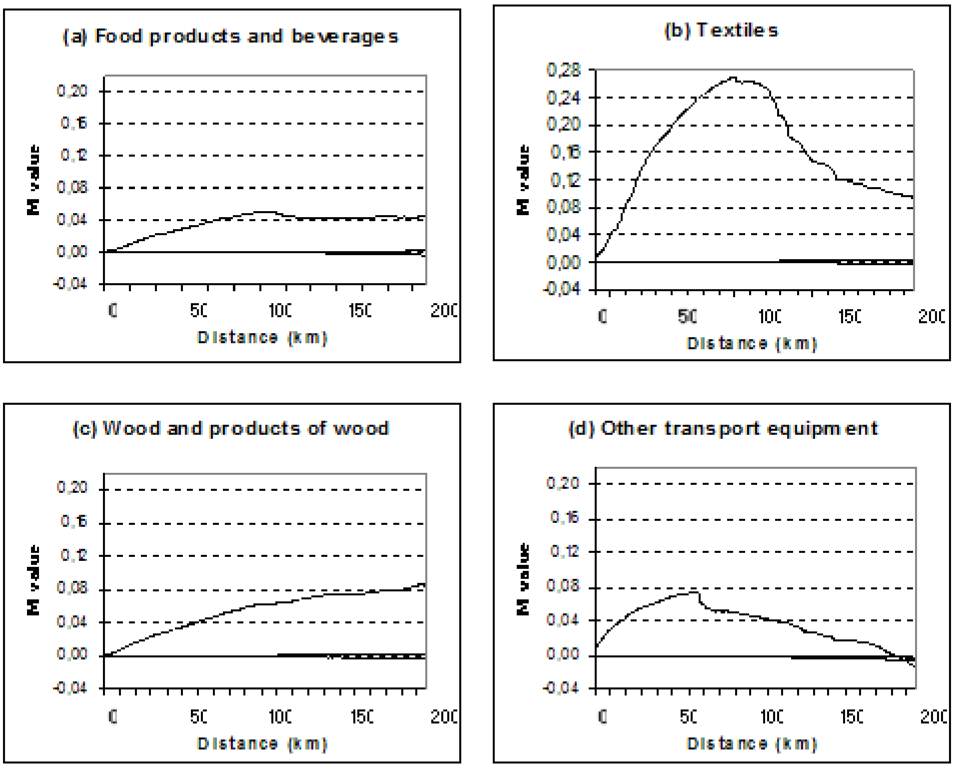


Figure 7. Spatial location patterns (M_{CSR}).

Firstly, we must focus on the second and third column of 'Table 1'. In this way, we realise that every Spanish manufacturing sector analysed, excepting (35) 'Other transport equipment', presents concentration relative to complete spatial randomness, whatever the distance of "*r*" we consider. Nevertheless, we should not forget that the highest radius considered is 200 km.

Secondly, we should pay attention to the *intensity* reached by the different sectors, that is to say, we should point out which are the highest and the lowest concentrated manufacturing sectors in Spain. Thus, the Spanish manufacturing sectors that reach a highest concentration level are, in the following order, (19) Tanning and dressing of leather, (17) Textiles, (32) Radio, televisions & other appliances, (33) Instruments, (22) Publishing, printing and recorded media, (24) Chemical and chemical products and (31) Electrical machinery. Should be highlighted that this result is independent from the *M* value taken into account, M_{CSR} or M_{TM}. On the contrary, the sectors with the lowest concentration level are

(15) Food products and beverages, (20) Wood and products of wood, (26) Other non-metallic mineral products and (35) Other transport equipment.

Table 1. Location patterns of Spanish manufacturing sectors (M_{CSR})

Sectors (NACE 93 - Rev. 1)	Significant concentration	Significant dispersion	Significant peak		Type of cluster
			M_{CSR} value	Distance (r)	
15 Food products and beverages	All distances	---	0,05	100 km	2
17 Textiles	All distances	---	0,27	87 km	1
18 Wearing apparel and dressing	All distances	---	0,09	87 km	2
19 Tanning and dressing of leather	All distances	---	0,50	156 km	2
20 Wood and products of wood	All distances	---	0,08	200 km	2
21 Pulp, paper and paper products	All distances	---	0,16	99 km	1
22 Publishing, printing & recorded media	All distances	---	0,20	85 km	1
24 Chemical and chemical products	All distances	---	0,17	90 km	1
25 Rubber and plastic products	All distances	---	0,13	97 km	1
26 Other non-metallic mineral products	All distances	---	0,08	200 km	2
27 Basic metals	All distances	---	0,10	80 km	1
28 Fabricated metal products	All distances	---	0,09	86 km	1
29 Other machinery and equipment	All distances	---	0,13	96 km	1
31 Electrical machinery	All distances	---	0,16	90 km	1
32 Radio, televisions & other appliances	All distances	---	0,21	91 km	1
33 Instruments	All distances	---	0,21	90 km	1
34 Motor vehicles and trailers	All distances	---	0,11	100 km	1
35 Other transport equipment	0-186 km	189-200 km	0,07	64 km	1
36 Furniture and other products	All distances	---	0,12	183 km	2
37 Recycling	All distances	---	0,09	86 km	1

It may be interesting to compare the above-mentioned results with those obtained by Duranton and Overman (2005) in UK. Surprisingly, they are very similar. On the one hand, they find that the localized sectors in UK are 17, 18, 19, 22, 30, 31 32 and 33, which almost coincide with the most concentrated sectors in Spain. On the other hand, the lowest concentrated sectors in Spain coincide perfectly with those non-localized sectors in UK, 15, 20 and 26. Therefore, we notice that manufacturing sectors tend to follow similar patterns of localization between countries; at least, it seems to be the case between Spain and UK. Besides, the most concentrated sectors in France and US are Textile (17) and Leather products

(19)[16], coinciding as well these results with the most concentrated sectors in Spain and UK.

Alonso-Villar *et al.* (2004) as well studied the geographical concentration of Spanish industry, between 1993 and 1999[17], and concluded that the most highly concentrated industries, according to Maurel and Sédillot index, are 19, 17, 32, 22, 33 and 24. As we can see, these results coincide perfectly with ours. Therefore, thanks to the results obtained by them, we can also deduce that the spatial location patterns of Spanish manufacturing sectors have not varied significantly in the last years, since the highest concentrated sectors are still being the same as in 1999.

Thirdly, we should valuate the *persistence* of this concentration in the space, in other words, the spatial scale dimensions of the cluster. If we pay attention to this characteristic of the spatial distribution, we can find two differentiated kinds of sectors, named 1 or 2 depending on the characteristic features of their location patterns.[18] Sectors named with number 1 present, in the beginning, an increasing in the activity concentration, reaching an absolute maximum point or peak and, finally, decreasing just as the distance of the radius makes higher. We should add that, although the evolution of these sectors is similar between them, the intensity and the distance at which is reached the maximum concentration are different; being similar, but not identical, the distribution patterns of these ones. Most of the Spanish manufacturing sectors belong to this 'Type 1' of location patterns[19] and the textile sector is a clear example (see Figure 7b).

However, just six sectors belong to the 'second type' (15, 18, 19, 20, 26 and 36). These sectors keep a constant or growing concentration, whatever the value of "*r*" considered, without reaching an absolute maximum peak. In this second kind of sectors, we assume that this maximum peak will be reached in a value of the radius higher than 200 km (see Figure 7c).

Last but not least, the *distance* (*r*-value) at which the highest concentration is reached or, in other words, the size of the cluster, is an important point to analyse. In fact, this distance may be altered by the scope of the agglomerative forces that act and characterise each sector. Therefore, the intensity and the distance at which the highest concentration is reached are distinctive of each sector and differ between them. That is to say, each Spanish manufacturing sector presents different location patterns, although, we can find some regularities between them.

---

[16] As can be deduced from Maurel and Sédillot (1999) and Ellison and Glaeser (1997) results.

[17] Although they just presented results for 1999, since not many differences were observed throughout the whole period.

[18] We can find this information in the sixth column of Table 1, 'Type of cluster'.

[19] Sectors 17, 21, 22, 24, 25, 27, 28, 29, 31, 32, 33, 34, 35 and 37.

For example, paying attention to the spatial scale at which the cluster is produced, we realise that this distance coincides in some sectors. In fact, as we can see in the fifth column of ‘Table 1’, nine of the twenty sectors analysed obtain their maximum concentration peak at a distance comprised between 85 and 91 km[20]. Once appreciated this coincidence, we should strive to find out whether there is any common characteristic between these sectors. However, we do not find similarities between them. Therefore, we are not capable to infer shared agglomerative forces that cause this regularity in the maximum concentration peak.

In this way, we realise that a common behaviour exists in the distance at which appears the highest concentration, but we do not know the cause. Therefore, we should look for a reason that could explain this fact, being able to be possible that our ‘benchmark’ can create regularities in the maximum concentration distance of every sector. Indeed, by means of a random and independent distribution (*CSR* benchmark), we do not take into account the “first nature” factors or the general tendency for economic activity to agglomerate. This fact can create a coincidence in the distance at which the highest concentration is reached, creating a fictitious regularity, non-attributable to economic factors, but to “first nature” factors.

In order to correct the lack of economic realism of our first benchmark, we consider a different one. As explained in methodology, this second benchmark is constructed using the whole of manufacturing industries and the results of this second analysis are summarised in ‘Table 2’.

We are going to use the sector 35 to exemplify the results coming from both benchmarks. In this manner, we will be able to become aware of the existing differences between both analysis.

In Figure 6d, we can see the spatial distribution of firms of sector 35, *Other transport equipment*, in Spanish territory, being able to observe that its establishments are distributed in small clusters around Madrid, Barcelona, Vigo and the Basque Country and along the coast.

The aim of our analysis is to obtain the spatial location patterns, at different scales, of every Spanish manufacturing sector. Therefore, the resulting patterns should correspond with reality. However, whether we compare the outcomes obtained using both benchmarks (Figure 7d and 8d), we realise that the resulting location patterns are very different. Actually, Figure 8d represents more accurately the reality; that is to say, the location pattern shown in this graph has characteristic features more similar to the real distribution of firms than Figure 7d. The apparent

[20] These sectors are 17, 18, 22, 24, 28, 31, 32, 33 and 37.

concentration in Figure 7d is, in part, due to an unreal and inappropriate benchmark. In fact, using M_{CSR} function the sector 35 presents significant concentration up to a distance of the radius of 186 km. However, using M_{TM} function we achieve more realistic results, appearing significant concentration up to a radius of 28 km and reaching the maximum significant peak at a distance of 9 km. Besides, we find significant dispersion from km 66 onwards. Thus, we realise that the results obtained from M_{TM} function describe in a more accurate way the real distribution of establishments in space, confirming the concentration of activity of this sector at low distances and the dispersion at large distances.

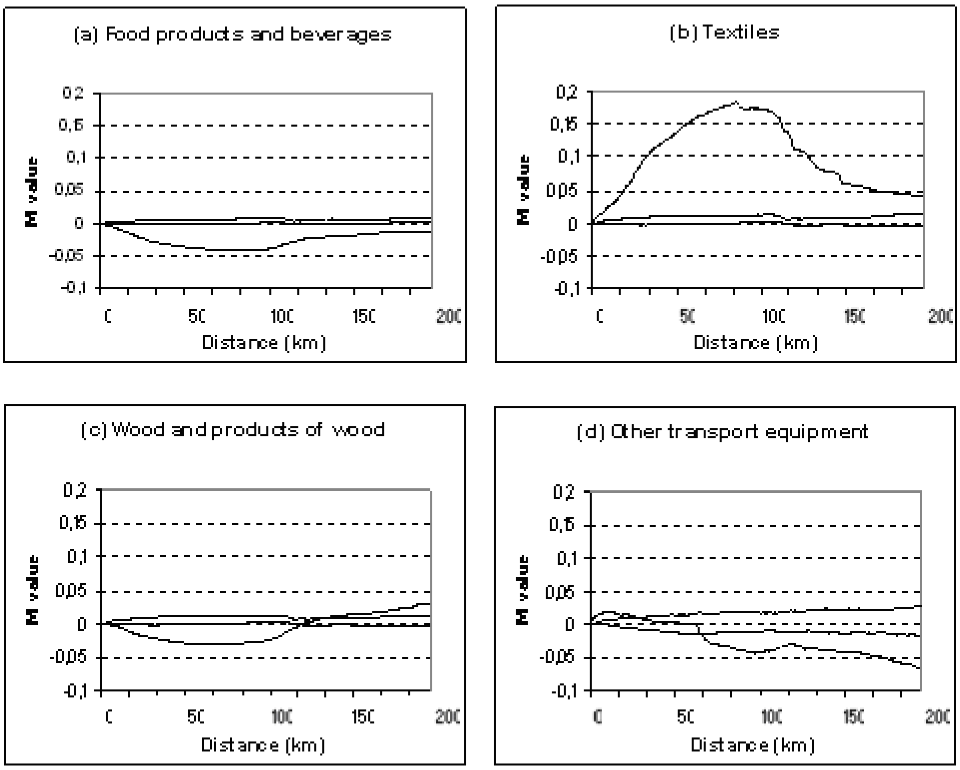


Figure 8. Spatial location patterns (M_{TM}).

Graphics in Figure 8 show how activity is distributed in space. We keep using the four above-presented sectors (15, 17, 20 and 35) to be able to compare the location patterns resulting from both benchmarks and, at first glance, we observe that these patterns differ considerably from first analysis.

The rest of the sectors analysed using the second benchmark, presented in 'Table 2', as well show different location patterns. A clear example of this fact is that not every sector presents concentration relative to overall manufacturing, whatever the distance of the radius considered, as happens with CSR benchmark. Nevertheless, as was to be expected, we find a coincidence. The highest concentrated sectors, those that reach a highest M value, are the same than in first analysis and in the same order, 19, 17, 32, 33, 22, 24 and 31. Obviously, the M_{TM} value of these sectors is not as high as the M_{CSR} value, because, in this case, we compare each sector with overall manufacturing and not with complete spatial randomness.

Table 2. Results of the M_{TM} function for every Spanish manufacturing sector

Sectors (NACE 93 - Rev. 1)	Significant concentration	Significant dispersion	Significant peak		Type of cluster
			M_{TM} value	Distance (r)	
15 Food products and beverages	---	All distances	-0,04	80 km	2
17 Textiles	All distances	---	0,18	87 km	1
18 Wearing apparel and dressing	---	---	---	---	---
19 Tanning and dressing of leather	All distances	---	0,45	157 km	1
20 Wood and products of wood	137-200 km	0-120 km	-0,03	64 km	4
21 Pulp, paper and paper products	All distances	---	0,07	103 km	1
22 Publishing, printing & recorded media	All distances	---	0,12	36 km	1
24 Chemical and chemical products	0-147 km	---	0,08	83 km	1
25 Rubber and plastic products	0-167 km	---	0,04	90 km	1
26 Other non-metallic mineral products	145-200 km	0-123 km	-0,03	60 km	4
27 Basic metals	42-84 km	127-200 km	0,02	63 km	3
28 Fabricated metal products	34-87 km	97-200 km	0,01	66 km	3
29 Other machinery and equipment	0-121 km	130-200 km	0,04	77 km	3
31 Electrical machinery	0-131 km	---	0,08	81 km	1
32 Radio, televisions & other appliances	All distances	---	0,13	83 km	1
33 Instruments	0-175 km	---	0,12	81 km	1
34 Motor vehicles and trailers	51-100 km	---	0,02	80 km	1
35 Other transport equipment	0-28 km	66-200 km	0,02	9 km	3
36 Furniture and other products	106-200 km	25-93 km	0,06	193 km	4
37 Recycling	---	---	---	---	---

The Spanish manufacturing sectors present different location patterns relative to the whole of manufacturing firms. In fact, the *persistence* or spatial dimension of the cluster varies depending on the sector considered. Therefore, we classify the whole of the sectors analysed in four groups: (1) those that show just concentration patterns, (2) those that show just dispersion patterns, (3) those that are concentrated at low distances and dispersed at large distances and, finally, (4) those that are dispersed at low distances and concentrated at large distances.

The classification of the different sectors appears recapitulated in last column of table 2, 'Type of sector'. On the one hand, we realize that 10 sectors, half of the total considered, present only concentration, whereas just 1 sector presents only dispersion relative to the total manufacturing. On the other hand, we can see that the rest of the sectors present different location patterns depending on the spatial scale chosen, that is to say, depending on the distance of the radius considered. So sectors 27, 28, 29 and 35 show concentration patterns, relative to the whole of manufacturing, at low distances and dispersion at large distances, that is to say, concentration takes place at small scales (type 3). Consequently, we may deduce that the establishments of these sectors are distributed in small clusters, presenting dispersion when the distances make larger. Fourth type of sectors, 20, 26 and 36, are dispersed at low distances and concentrated at large distances; in this way, we should consider these sectors as dispersed because we cannot interpret the only existence of concentration at large distances as a concentrated sector. In fact, as we see in the previous table, we just find significant concentration patterns in these sectors from 137, 145 and 106 km onwards, respectively. Finally, we should notice that sector 18 and 37 do not present significant results.

Once explained the resulting location patterns of every sector, we realise that most of them are concentrated at low distances, excepting 15, 20, 26 and 36. However, dispersion happens at different scales, appearing across all distances in sector 15, at low distances in sectors 20, 26, and 36, and at large distances of the radius in sectors 27, 28, 29 and 35.

From previous paragraph and taking into account the *M* value from fourth column of table 2, we can deduce that the highest dispersed sectors in Spain are (15) Food products and beverages, (20) Wood and products of wood and (26) Other non-metallic mineral products. These sectors, which show dispersion, have an elevated dependence on natural resources or are related to food. Moreover, it is as well likely that these sectors occupy specialised manufacturers that disperse their establishments to provide in the best way the different markets.

Nevertheless, whether we pay attention to the sectors that have been classified as the highest concentrated, we realise that these do not exhibit particular characteristics. The role of '*knowledge spillovers'* seems evident whether we refer to them as a source of concentration, although in our case it is not so obvious. In fact, the two first sectors that have a higher level of concentration, 19 and 17, are clearly low-tech. This indicates that not only the 'knowledge spillovers' determine the concentration of activity, but also other factors as local labour pooling, natural advantages, tradition, transport costs or the linkages[21], play an important role in the patterns of localization. Therefore, among the most highly concentrated Spanish manufacturing sectors we find those for which the geographic concentration is completely determined by *historical trends* [22] (17 and 19), those for which *technological spillovers* seem to be the main reason of their localization (31, 32 and 33) and finally, those for which the search of *skilled labour* is determinant in their decision to concentrate (22 and 24).

According to literature, high-tech sectors must be the most highly concentrated, but it does not happen in our analysis. However, it occurs neither in UK nor in France. In fact, Devereux *et al.* (2004), according to their results, underlined that '*the most geographically concentrated industries appear to be relatively low-tech*' and Maurel and Sédillot (1999) obtained that the most concentrated sectors are textile and leather products, two of the most traditional and low-tech sectors.

Relating to the percentage of sectors that show concentration or dispersion, we realise that the percentage of sectors that show concentration patterns is quite elevated (78%). In this way, we should ask ourselves whether the percentage of manufacturing workers employed in these sectors is higher or lower than the percentage of manufacturing workers employed in dispersed sectors. We find that 33% of the total manufacturing employees work in dispersed sectors whereas 67% work in concentrated sectors. This confirms that a larger proportion of manufacturing employees work in sectors with concentration patterns. However, we realise that the percentage of workers in dispersed sectors is higher than the percentage of dispersed sectors, so we can affirm that sectors that present dispersion patterns are larger than those that present concentration patterns.

Finally, we can now assure that whether exists any kind of coincidence in the *distance* at which the highest concentration is reached in some sectors, it will not depend on 'first nature' factors, because the benchmark employed in this second analysis takes into account this aspect. In this way, this coincidence will be produced by the agglomeration strengths belonging to each sector.

At last, just rest to add that we are conscious about the aggregation of the data of our analysis. In fact, we treat with sectors aggregated at two-digit level. This aspect may generate a clear effect of compensation between the different branches of each sector, since the most aggregated and the most dispersed ones can compensate between them. Therefore, try to find out whether the different branches of every sector present similar or different location patterns than the sector itself could be the next step in our analysis.

[21] So emphasized by Krugman (1991a).
[22] They are probably settled and clustered in the same area since the Industrial Revolution.

5. Conclusions

This paper analyses the spatial location patterns of manufacturing firms in Spain. To do this, we used a distance-based method, which allowed treating space as continuous and avoided the inconvenience of administrative scale, letting us to measure geographic concentration at different scales. Therefore, by means of this method, we can know the intensity of concentration or dispersion of every Spanish manufacturing sector and at which distance we obtain its maximum level. Moreover, we can detect whether the departures are statistically significant from randomness

Our purpose of managing to control for the overall distribution of manufacturing was accomplished, since we construct our confidence interval restricting the location of firms to the actual sites of manufacturing firms. Besides, we also improve the shape of our area of study. In fact, by means of the software employed, R, we can apply border corrections, adequately, in any irregular polygonal shape. Therefore, we introduce a new envelope technique, substituting the rectangular shape, used by Marcon and Puech (2003), for a polygonal shape that delimits more properly the Spanish territory.

Regarding to the benchmark employed, we use two different ones, complete spatial randomness and the whole of manufacturing, because this type of analysis is sensitive to the benchmark considered. However, we realise that first one has some shortcomings, since economic activity does not locate in a random and independent way, being able to affirm that this benchmark does not take into account the first or the second nature. Our test just satisfies the five requirements mentioned by Duranton and Overman (2005) whether we use as a benchmark the whole of manufacturing. In this way, our test is comparable across industries, controls for the overall agglomeration of manufacturing, controls for industrial concentration, is unbiased with respect to scale and aggregation and gives an indication of the significance of the results.

The application of our test provides us the following results. First of all, we realise that spatial location patterns in Spain are very different depending on the sector analysed. However, the highest concentrated sectors coincide in both analysis, independently of the benchmark used. Second, most of the sectors are concentrated at low distances, but dispersion happens at different scales. Third, whether we pay attention to the percentage of sectors that show concentration or dispersion, we obtain that 78% of them present concentration. Lastly, we find that just 66% of the total manufacturing employees work in concentrated sectors, so we can affirm that sectors that present concentration patterns are smaller (considering the number of workers) than those that present dispersion patterns.

This last finding coincides with Duranton and Overman (2005) results. Moreover, appears also coincidences between the highest and the lowest concentrated sectors, being possible to conclude that manufacturing sectors tend to follow similar patterns of localization between countries, at least between Spain and UK.

Finally, we just can add that many questions remain to be investigated and we will try to solve them in next papers. It is evident that the level of aggregation may be interesting and that would be motivating to analyse, in a nearer future, whether the different branches of every manufacturing sector present similar or different location patterns than the sector itself. In fact, we already have at our disposal the required data.

Appendix 1

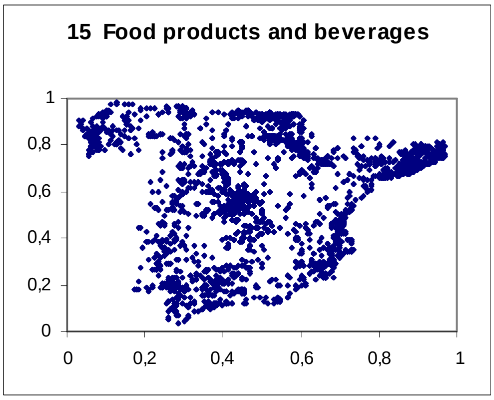
Table A1. Additional descriptive information about Spanish manufacturing sectors

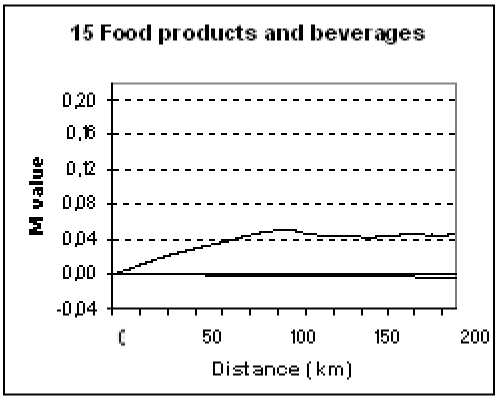
Sectors (NACE 93 - Rev. 1)	Number of firms	Number of employees	Technological intensity[23]
15 Food products and beverages	5761	356314	L
16 Tobacco products	6	1226	L
17 Textiles	1949	81818	L
18 Wearing apparel and dressing	1710	59286	L
19 Tanning and dressing of leather	1698	46708	L
20 Wood and products of wood	2340	75844	L
21 Pulp, paper and paper products	837	56890	L
22 Publishing, printing & recorded media	3004	130222	L
23 Coke, refined petroleum products	12	16417	M-L
24 Chemical and chemical products	1722	158238	H
25 Rubber and plastic products	2165	138488	M-L
26 Other non-metallic mineral products	3413	225792	M-L
27 Basic metals	986	137066	M-L
28 Fabricated metal products	8094	267568	M-L
29 Other machinery and equipment	3015	161407	M-H
30 Office machinery and computers	77	6374	H
31 Electrical machinery	1099	79357	M-H
32 Radio, televisions & other appliances	344	31593	H
33 Instruments	376	19528	H
34 Motor vehicles and trailers	876	192873	M-H
35 Other transport equipment	451	58274	M-H
36 Furniture and other products	2924	100084	L
37 Recycling	228	8095	L

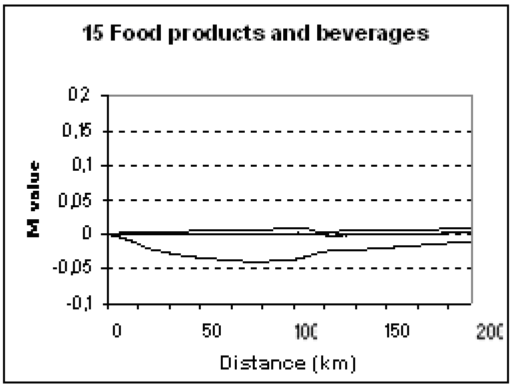
[23] This classification of sectors according to the technological intensity belongs to the National Statistics Institute, meaning H = high, M-H = medium high, M-L = medium low and L = low.

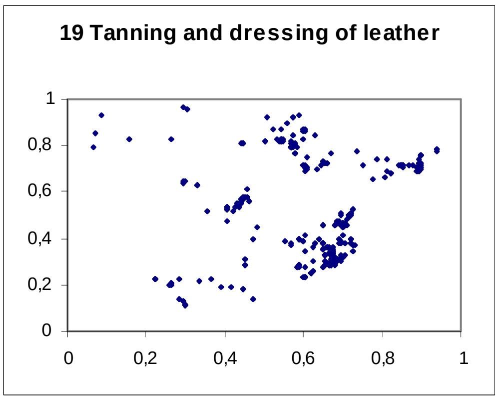
Appendix 2

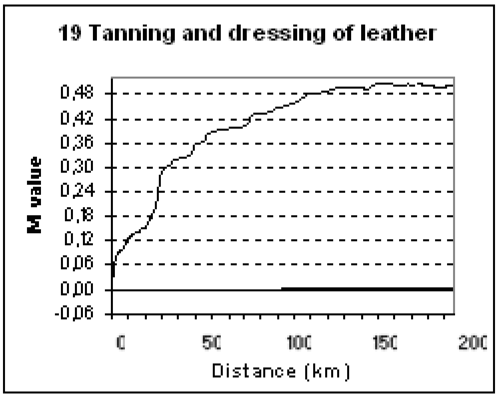
On the one hand, graphs situated on the left show the spatial distribution of firms of each Spanish manufacturing sector. On the other hand, graphs on the right-up show the spatial location patterns of these sectors measured by the M_{CSR} function and graphics situated on the right-down show the spatial location patterns measured by means of the M_{TM} function. These two last graphics allow us to detect the differences between geographical concentration scales according to the sector considered.

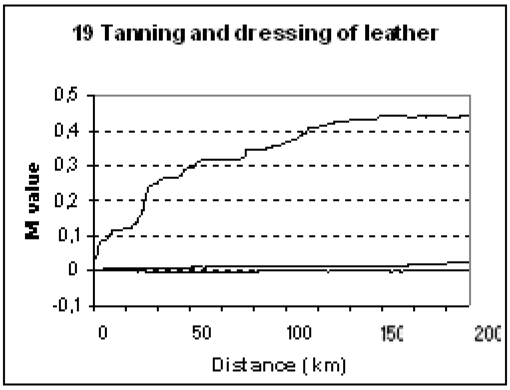


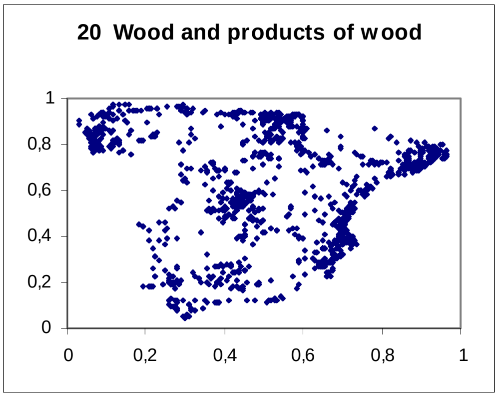


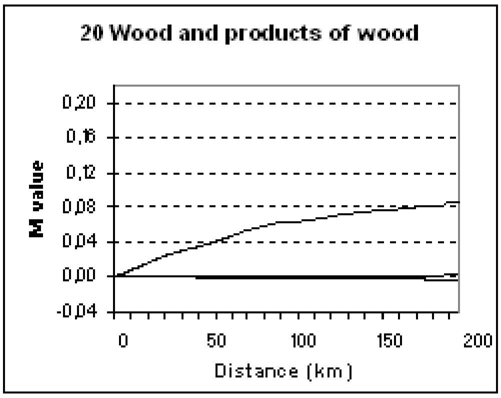


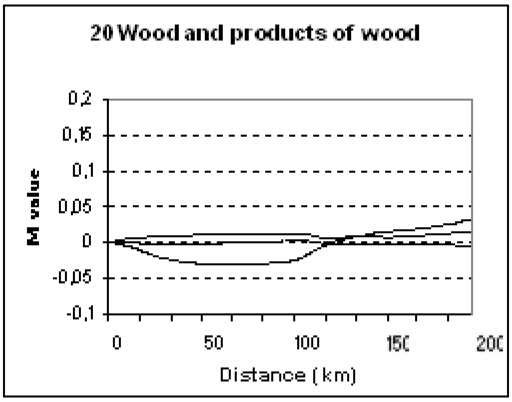


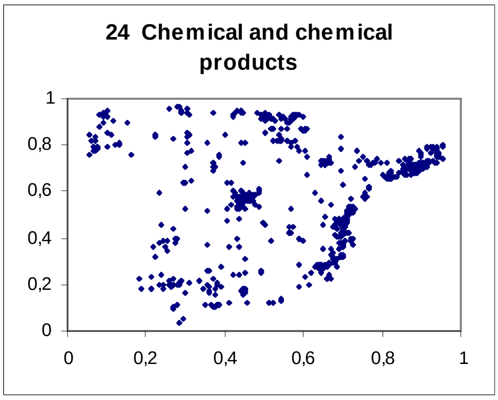


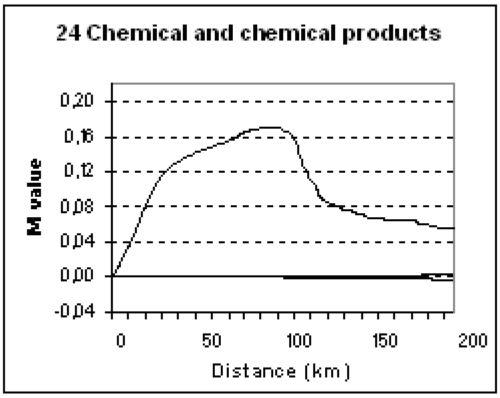

19 Tanning and dressing of leather
M value
0,5
0,4
0,3
0,2
0,1
0
-0,1
0
50
100
150
200
Distance (km)

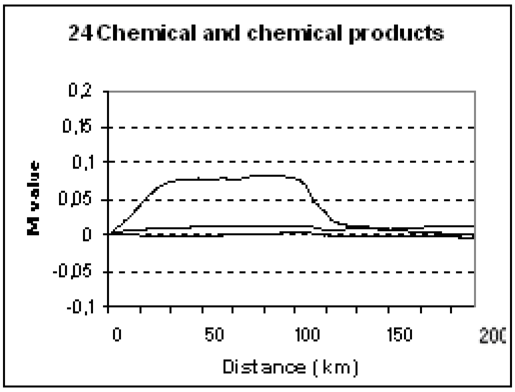

20 Wood and products of wood
1
0,8
0,6
0,4
0,2
0
0
0,2
0,4
0,6
0,8
1

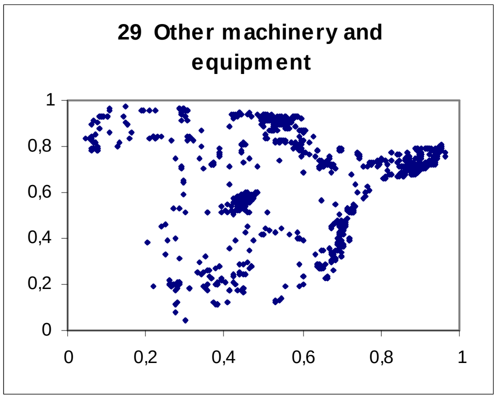

20 Wood and products of wood
M value
0,20
0,16
0,12
0,08
0,04
0,00
-0,04
0
50
100
150
200
Distance (km)

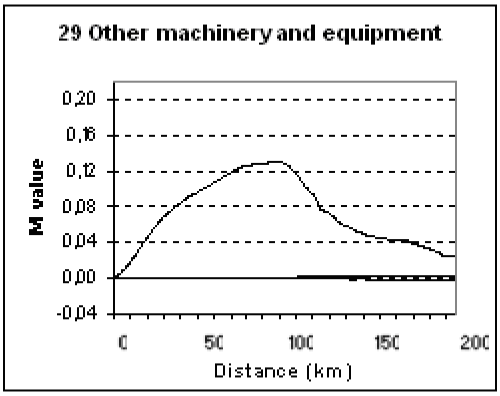

20 Wood and products of wood
M value
0,2
0,15
0,1
0,05
0
-0,05
-0,1
0
50
100
150
200
Distance (km)

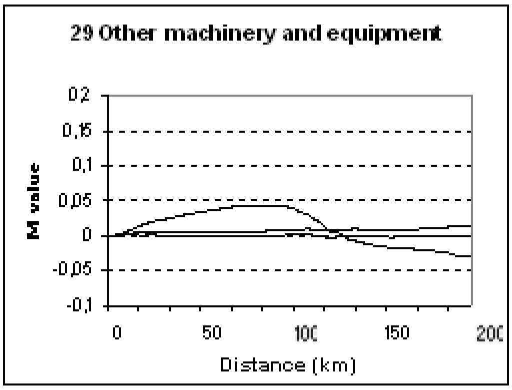

24 Chemical and chemical products
1
0,8
0,6
0,4
0,2
0
0
0,2
0,4
0,6
0,8
1

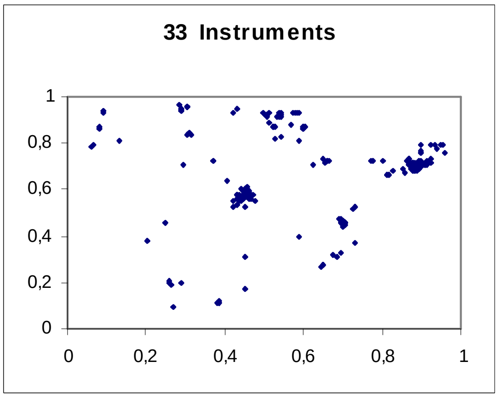

24 Chemical and chemical products
M value
0,20
0,16
0,12
0,08
0,04
0,00
-0,04
0
50
100
150
200
Distance (km)

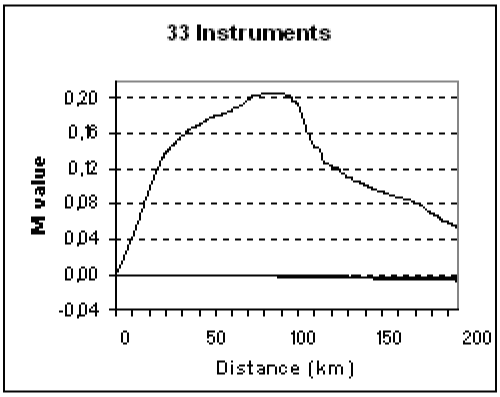

24 Chemical and chemical products
M value
0,2
0,15
0,1
0,05
0
-0,05
-0,1
0
50
100
150
200
Distance (km)

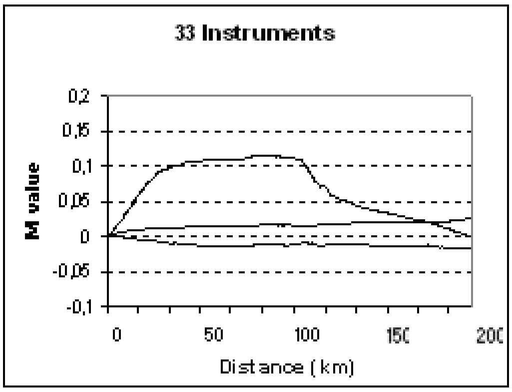

29 Other machinery and equipment
1
0,8
0,6
0,4
0,2
0
0
0,2
0,4
0,6
0,8
1


29 Other machinery and equipment
M value
0,20
0,16
0,12
0,08
0,04
0,00
-0,04
0
50
100
150
200
Distance (km)


29 Other machinery and equipment
M value
0,2
0,15
0,1
0,05
0
-0,05
-0,1
0
50
100
150
200
Distance (km)


33 Instruments
1
0,8
0,6
0,4
0,2
0
0
0,2
0,4
0,6
0,8
1


33 Instruments
M value
0,20
0,16
0,12
0,08
0,04
0,00
-0,04
0
50
100
150
200
Distance (km)


33 Instruments
M value
0,2
0,15
0,1
0,05
0
-0,05
-0,1
0
50
100
Distance (km)

References

Alonso-Villar, O., Chamorro-Rivas, J.M. and González-Cerdeira, X., (2001) An Analysis of the Geographic Concentration of Industry in Spain. Documento de Traballo 0103, Departamento de Economía Aplicada, Universidade de Vigo.

Alonso-Villar, O., Chamorro-Rivas, J.M. and González-Cerdeira, X., (2003) Spillovers geográficos y sectoriales de la industria española. Revista de Economía Aplicada, 32, 77-95.

Alonso-Villar, O., Chamorro-Rivas, J.M. and González-Cerdeira, X., (2004) Agglomeration economies in manufacturing industries: the case of Spain. Applied Economics, 36, 2103-2116.

Amiti, M., (1997) Specialisation patterns in Europe. Discussion paper 363, Centre for Economic Performance, London

Brülhart, M., (2001) Evolving geographical concentration of European manufacturing industries. Weltwirtschaftliches Archiv, 137 (2), 215-243.

Callejón, M., (1997) Concentración geográfica de la industria y economías de aglomeración. Economía Industrial, 317, 61-68.

Devereux, M.P., Griffith, R. and Simpson, H., (2004) The geographic distribution of production activity in the UK. Regional Science and Urban Economics, 34, 533-564.

Duranton, G., Overman, H.G., (2005) Testing for Localisation using Micro-Geographic Data. Review of Economic Studies, 72, 1077-1106.

Ellison, G., Glaeser, E., (1997) Geographic concentration in U.S. manufacturing industries: a dartboard approach. Journal of Political Economy, 105 (5), 889-927.

Ellison, G., Glaeser, E., (1999) The geographic concentration of industry: does natural advantage explain agglomeration? American Economic Review, 89 (2), 311-316.

Fujita, M., Krugman P. and Venables A.J., (1999) The Spatial Economy: Cities, Regions and International Trade. MIT Press, Cambridge, MA.

Krugman, P., (1980) Scale Economies, Product Differentiation and the Pattern of Trade. American Economic Review, 70, 950-959.

Krugman, P., (1991a) Geography and Trade. MIT Press, Cambridge, USA.

Krugman, P., (1991b) Increasing returns and economic geography. Journal of Political Economy, 99 (3), 413-499.

Marcon, E. and Puech, F., (2003a) Evaluating the Geographic Concentration of Industries using Distance-Based Methods. Journal of Economic Geography, 3 (4), 409-428.

Marcon, E. and Puech, F., (2003b) Measures of the Geographic Concentration of Industries: Improving Distance-Based Methods. Cahiers de la MSE, 2003.18. 22p.

Marshall, A., (1890) Principles of Economics. MacMillan, London.

Maurel, F., Sédillot, B., (1999) A measure of the geographic concentration in French manufacturing industries. Regional Science and Urban Economics, 29 (5), 575-604.

Morton, A., (2003) Workbook from Alan Morton. Electronic publication. Distribution mapping software (DMAP), http://www.dmap.co.uk.

Ottaviano, G.I.P., Puga, D., (1998) Agglomeration in the Global Economy: A Survey of the ‘New Economic Geography’. The World Economy

Ottaviano, G., Thisse, J.F., (2004) Agglomeration and Economic Geography. Handbook of Regional and Urban Economics, Volume 4, 2563-2608.

Overman, H.G., Redding S. and Venables A.J., (2003) The Economic Geography of Trade Production and Income: a Survey of Empirics. Handbook of International Trade. Choi, E.K. and Harrigan, J., Blackwell Publishing Ltd., 353-387.

Paluzie, E., Pons, J. and Tirado, D.A., (2001) Regional Integration and Specialization Patterns in Spain. Regional Studies, 35 (4), 285-296.

Paluzie, E., Pons, J. and Tirado, D.A., (2004) The geographical concentration of industry across Spanish regions, 1856-1995. Jahrbuch für Regionalwissenschaft (Review of Regional Research), 24, 2, 143-160.

Puga, D., (1999) The rise and fall of regional inequalities. European Economic Review, 43 (2), 303-334.

Puga, D., (2002) European regional policies in light of recent location theories. Journal of Economic Geography, 2, 373-406.

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

R Development Core Team (2007) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Ripley, B.D., (1976) The second-order analysis of stationary point processes. Journal of Applied Probability 13, 255-266.

Ripley, B.D., (1977) Modelling Spatial Patterns. Journal of the Royal Statistical Society - Series B (Methodological), 39 (2), 172-192.

Ripley, B.D., (1979) Test of ‘randomness’ for spatial patterns. Journal of the Royal Statistical Society - Series B (Methodological), 41, 368-374.

Rosenthal, S.S., Strange, W.C., (2001) The Determinants of Agglomeration. Journal of Urban Economics, 50 (2), 191-229.

SABI. System of Iberian Balances Analysis.

Tirado, D.A., Paluzie, E. and Pons, J., (2002) Economic integration and industrial location: the case of Spain before World War I. Journal of Economic Geography, 2, 343-363.

Venables, A.J., (1995) Economic Integration and the Location of Firms. American Economic Review, 85(2): 296-300.

Venables, A.J., (2006) Shifts in economic geography and their causes. Discussion Paper. Centre for Economic Performance, London School of Economics and Political Science, London, UK.

Viladecans, E., (2001) La concentración territorial de las empresas industriales: un estudio sobre la unidad geográfica de análisis mediante técnicas de econometría espacial. Document de treball 2001/2, Institut d’Economia de Barcelona.