Spatial distribution of economic activities

Federico Pablo-Martí (♠): federico.pablo@uah.es
Josep-Maria Arauzo-Carod (♣, ♦): josepmaria.arauzo@urv.cat

Abstract

The aim of this paper is to analyse spatial distribution of economic activities and, concretely, to identify single industry clusters in Spanish manufactures. We study the spatial distribution of firms from XXXXXX at a microgeographic level and we avoid MAUP by using homogeneous cells that cover all mainlaind Spain. Dataset used comes from Mercantile Registers of Spanish firms (manufacturers and services).

(♠) Facultad de CC. Económicas y Empresariales
(Universidad de Alcalá)
Pl. de la Victoria, 2; 28802 - Alcalá de Henares
Phone: + 34 918 854 231, Fax + 34 918 854 201

(♣) CREIP - QURE
Department of Economics (Universitat Rovira i Virgili)
Av. Universitat,1; 43204 - Reus
Phone: + 34 977 758 902, Fax + 34 977 759 810

(♦) Institut d'Economia de Barcelona (IEB)
Av. Diagonal, 690; 08034 - Barcelona

Key words: clusters, microgeographic data, self-organizing maps, firm location
JEL classification: R10, R12, R34

*This research was partially funded by the grants SEJ2010-19088/ECON, SEJ2010-16934/ECON, the "Xarxa de Referència d'R+D+I en Economia i Polítiques Públiques" and the SGR Program (2009-SGR-322) of the Catalan Government and by the PGIR program N-2008PGIR/05 of the Rovira i Virgili University. We are grateful to seminar participants at the 51st European Congress of the Regional Science Association International (Barcelona, 2011). Any errors are, of course, our own.

1. Introduction

Analysis of spatial distribution of economic activity has plenty of implications in several areas like urban planning, infrastructures, firm supporting policies and land use, among others, and is receiving an increasing attention by researchers. Most of analyses of spatial distribution of economic activity have been carried out using extant administrative units (e.g. counties, regions, etc.), but unfortunately, these analyses suffer from the shortcoming that administrative units vary greatly in size and shape, do not always coincide with real economic areas and are sometimes arbitrary.

To deal with such constraints, recent research has started to use ad hoc units (usually smaller), as we do in this paper. These units are created by equally dividing a space into homogeneous squared cells and, therefore, do not exactly match any extant administrative unit.[1]

Mapping the spatial distribution of economic activity is of key importance for policy makers, but there is no agreement as to which technical approach is best for designing policy. Currently, there are two main approaches: Industrial Districts and Clusters. While the former is more popular mainly due to the Sforzi-ISTAT methodology, the latter is potentially easier to use because of its lower data requirements. Therefore, in this paper we will use the cluster approach due to both data availability and the shortcomings of Sforzi-ISTAT methodology (Boix and Galletto, 2008).

The methodology proposed in this paper aims to overcome previous constraints, to obtain more precise results and, as a result, to improve public policy design. Accordingly, in this paper we try to identify manufacturing and service clusters in Spain. Additionally, we classify these clusters according to the reasons behind the clusterization processes; that is, whether firms tend to locate together because they look for the same types of site (regardless of the

[1] There are also other approaches such as those that use the stochastic methodology of Point Pattern or those that use Neuronal Networks for pattern recognition. However, these approaches are not able to do the multisectorial analyses that are the goal of this paper.

industry to which they belong), or whether firms look to be located close to their suppliers / customers in order to optimise commercial exchanges. Concretelly, by dividing Spain into homogeneus cells we check whether each industry follows a concentrated or dispersed pattern and, later, whether co-localization exists for pairs of industries, so clusters made by different industries can alse be identified. And finally, once we have identified such industry location patterns we apply self-organizing maps to show the local microstructure of clusters.

This paper is organised as follows. In the next section we review the main literature on the spatial distribution of economic activity and the spatial units used in empirical analysis. In the third section we explain the data set, we describe and analyse the spatial distribution of firms in Spain and we define the methodology used for identifying clusters. In the fourth section we present and discuss our main empirical results. In the final section we present our conclusions.

2. Spatial distribution of economic activity: theories and empirical approaches

There is plenty of empirical evidence regarding the uneven spatial distribution of economic activity and the way how firms tend to cluster to each other, sometimes due to urbanization economies and sometimes due to localization economies. Among most relevant contributions that have reported and analysed this phenomena there are those of XXXX (2011), XXXX (2010), XXXX (2010), Duranton and Overman (2005), Devereux et al. (2004), Maurel and Sédillot (1999) and Ellison and Glaeser (1997). There is also a large list of contributions for the Spanish case (Boix and Galletto, 2008; Paluzie et al., 2004; Viladecans, 2004), where clusterization of economic activity is quite important in some regions.[2]

[2] Concretelly, Boix and Galletto (2008) identify four axes where specialised industrial district are of great importance: the main axis runs across the Mediterranean coast from the north of Catalonia to the south of Murcia; the second one links the south of Catalonia to the Basque Country and North-East of Castile and León; the third one goes South from Madrid to the

Therefore, firms look to be close to other firms. As we have said before, some firms prefer proximity with similar firms (localization economies), while others just want to be close to other firms, no matter their activities (urbanization economies). In any case, in order to better optimise external resources and to increase productivity firms need neighbours, and usually these neighbours are firms that have some types of common characteristics, as size, markets, industry, technological level, supply needs, type of workforce or use of infrastructures. So there are plenty of reasons to cluster with similar firms that, consequently, have similar characteristics.

There is a lot of empirical evidence showing that economic activities do not perfectly fit into extant administrative borders, rather they tend to spread accros neighbour areas. This phenomenon implies that contiguous areas could share a single agglomeration of firms without internal borders, making difficult to precisely identify where to analyse this phenomenon. At the lowest geographical level where the spillover effects dissolve internal borders ot at a higher level (combining smaller units) where the phenomenon exists only for a small part of the area considered?

Previous shortcomings illustrate that it is important to accurately analyse implications of spatial aggregation issues and which spatial areas are to be used, in view that using non appropriate areas could tend to biased results as several scholars like Arbia (2001, p. 414) ("(…) *any statistical measure based on spatial aggregates is sensitive to the scale and aggregation problems*") and Duranton and Overman (2005, p. 1079) ("(…) *any good measure of localization must avoid these aggregation problems*") point out.

This spatial aggregation problem is known as *Modifiable Area Unit Problem* (MAUP)[3], which is clearly illustrated by Arbia (2001) when showing that depending on how spatial borders are designed, the same spatial distribution of

---

provinces of Toledo, Ciudad Real, Jaen and Córdoba; and the last one is scattered across the provinces of Pontevedra and A Coruña (North-West of Spain).

[3] See Openshaw and Taylor (1979) for a detailed analysis and Wrigley (1995) for a further review.

, for instance, firms, could result in a minimum concentration pattern, in a maximum concentration pattern or an intermediate concentration pattern. Unfortunately, these issues have not been a major concern for scholars[4], usually due to the lack of sufficiently spatial disaggregated data, but this situation has started to change several years ago with the spread of spatially disaggregated datasets and the extended use of raster data with GIS packgages.

According to previous considerations, in this paper we aim to explicitly address such spatial aggregation issue when analysing cluster's formation in Spain. As we will explain in section 4, our spatial unit is not an administrative one, but a cell of 10 km * 10 km that covers all mainland Spain. This microgeographical approach has been used (with some variations) by other scholars like Duranton and Overman (2008, 2005), who used Britsh postcodes.

3. Data

Our data set refers to 2006 and comprises Spanish firms[5] from manufacturing, services and agriculture. The source of this data base is SABI (*Sistema de Análisis de Balances Ibéricos*), which uses data from the Mercantile Register including balance sheets and income and expenditure accounts. For each firm we also know the number of employees, the industry to which it belongs (the four digit NACE code), and its sales and assets, among other variables. We also have detailed information about the firm's geographical location; that is, information which is particularly relevant for the purposes of this paper. Nevertheless, the SABI dataset also has two important shortcomings. The first concerns the sample. Although the number of firms is very high (e.g. 581,712 service firms for the 2007 edition), microfirms and self-employed individuals are not taken into account, despite that fact that it is reasonable to assume that the

[4] See, nevertheless, papers by Arauzo-Carod and Manjón-Antolín (2004) and Arauzo-Carod (2008) about the implications for industrial location analysis. See, also Olsen (2002) for a borad discussion of the units to be used in geographical economics.
[5] It is important to notice that SABI data set is about firms, not establishments, so each firm could have more than one establishment, although most of firms have only one establishment.

spatial distribution of such activities is similar to that of the firms included. The second concerns the nature of the units; that is, SABI only covers firms, not establishments,[6] the latter being more appropriate for analyzing the spatial distribution of economic activity. In any case, since SABI covers most of the economic activity carried out in Spain, these disadvantages are easily overcome.[7]

4. Methodology of cluster identification

Our methodology departs partially from previous contributions based on distribution comparisons (Brenner, 2006 and 2004; Ellison and Glaser, 1997) and on distance distributions (Duranton and Overman, 2005) but we introduce some variations that allow us to better portrait single-industry clusters at a very detailed spatial disaggregation level. What do we do is to use homogeneous cells of 10 km * 10 km[8] instead of administrative spatial units, to use firm's georeferenced data to more precise on firm's location, to take into account total number of firms both at each cell and at a national level, to compare real distribution of firm with a random estimated distribution, to map both real and random distributions by using 2D and 3D maps and, finally, to identify specialised areas.

First, instead of using administrative units[9] (e.g., municipalities) we use homogeneous cells of 10 km * 10 km. By this way, we can overcome several limitations like (López-Bazo, 2006) the inability to take into account the precise

[6] Other alternative statistical sources such as *Censo de Locales* (INE) are not currently updated, although having firms as observation units instead of establishments also provides useful information since it highlights the role of municipalities when firms are choosing where to locate their headquarters.
[7] There are alternative datasets such as DIRCE (INE) but their data is presented only at 2-digit level and geographical location of the firms is also highly spatially aggregated.
[8] Cell's size of 10 km * 10 km was decided in terms to avoid compational constraints (smaller sizes implied a huge increase in computational capacity in order to deal with a larger unit of spatial units) and also trying to get a cell big enough to have several firms form different industries. Even if alternative sizes were also feasible (e.g., 5 km * 5 km, 20 km * 20 km) and, consequently, were also tested, we considered that the selected size was appropriate both from a computational and economic point of view.
[9] See, among others, Brenner (2006 and 2004) and Ellison and Glaeser (1997) for empirical applications with such administrative units.

location of firms, the limitations resulting from the special administrative aggregation levels in each country, the difficulties in comparing the results obtained for different levels of administrative aggregation, the non-economic nature of such administrative units, the size differences across administrative units, the modifiable areal unit problem (MAUP) and the existence of neighbour effects across units.

Second, as we have detailed before, using homogeneous cells allow us to more precisely identify location of firms, as Duranton and Overman (2005) do, although we only care about whether a cell is occupied or not, not about the exact location of the firms inside the cell.

Third, we built up conterfactuals by assuming that total number of firms in each industry remains constant and that total number of firms in each cell also remains constant.[10] This strategy allows us to compare the same number of firms but with different industry distributions (we expect to find the same industry distribution at each cell that at that of the whole sample). Thus, if the real data shows a cell with only one firm, our simulations will also show this cell with one firm, although the industry will appear as a random variable depending on industry distribution.

Fourth, we compare the actual number of cells with firms (real distribution) with the expected number of cells with firms (random distribution), and obtain a concentration index similar to that of Ellison and Glaeser (1997), excep that *i*) we focus on industry shares instead of agglomeration and *ii*) our index is centred at 1 (values below 1 indicate concentration and values over 1 indicate dispersion), while Ellison and Glaeser's (1997) index ranges between zero and infinite (i.e., they arbitrarily define the concentration threshold).

[10] This latter requirement implies that firms localise randomly inside "occupied" cells (i.e. areas where real firms are located) as stated by Duranton and Overman (2008). This approach means that firms are expected to be located only in those places that are available for economic activity (as firms do). Unfortunately, a major shortcoming of this approach is that it assumes that firms could be located elsewhere with other firms, regardless of the industry they are involved in, which is not as realistic (especially at a 2/3 digit level). An extension of this work (and a possible solution for this shortcoming) would be to regard manufacturing, services and agricultural firms as being located with other firms from the fields of manufacturing, services and agriculture respectively.

Fifth, we assume that there is a clusters when comparing real spatial distribution of firms with several computational simulations we get that the number of firms from an industry is significantly higher than the number obtained by simulation procedures.

Sixth, we make 3D cluster maps that easily allow to identify significant concentration of firms of the same.

5. Main results

Our main results show that the location of firms are driven by several industry-specific determinants (i.e. whether the firm belongs to a manufacturing or services activity or to a specific industry within these sectors) and also by their technological level. In some vertically integrated industries, reducing distance to providers / suppliers is a key issue, whereas other types of industries do not need such spatial proximity. Additionally, there are industries with no clear location patterns and which show a homogeneous firm distribution.

[INSERT TABLE 1]

Table 1 illustrates the expected spatial distributions of firms across regular cells[11] (according to the number of firms in each industry) and the real (observed) spatial distribution of such firms. In particular, it shows how many cells (*X*) contain firms from industry *y* (i.e. this is the “real” spatial distribution of firms); the expected number of cells (Mean) where firms from industry *y* should appear if they were randomly spatially distributed (according to the total number of firms in each industry); and a co-location index (Index) that relates these measurements to each other (i.e. Index = *X* / Mean). This index can be understood in the following way: if Index < 1, this means that the industry *y*

[11] These regular cells have an area of 100 km^2 (10 km * 10 km).

appears in fewer cells than expected (i.e. this industry is spatially concentrated in a smaller number of cells); and if Index > 1, this means that the industry *y* appears in more cells than expected (according to a random distribution), which means that this industry is spatially dispersed. This indicates that there is a certain location behaviour taking place that should be analyzed to determine whether or not it is a cluster (i.e. whether or not firms from industry *y* tend to locate together).

On a technological level, it seems that the lower the technological level of the industry, the higher the spatial dispersion (Table 1). Thus, high-tech firms tend to be more spatially concentrated than low-tech firms[12]. This appears to be logical since the markets and resources of such firms tend to be concentrated in a few areas, which means there is no logical reason for a dispersion pattern.

Our results regarding the differences between manufacturing and services, (Table 1) are even clearer than those of previous studies and show that whereas most services activities show high concentration levels (e.g. financial intermediation, education, business services, etc.), manufacturing activities are more dispersed (agriculture and fishing, food, beverages and tobacco, etc.). These results reflect the spatial distribution of population and economic activity and the production and distribution requirements of manufacturing and services. Specifically, most services need face-to-face interactions and thus their location decisions are strongly motivated by the locations of their customers (both firms and individuals). In contrast, manufacturers can transport their goods easily, which means that such interactions are not essential and that these firms can locate elsewhere.

So far we have analysed the spatial distribution of firms at single industry level and have shown that looking at certain industry specificities (i.e. manufacturing vs. services and high-tech vs. low-tech) helps us to understand such location patterns.

---

[12] As an example, indices of high-tech industries such as office machinery, computers and medical equipment, precision and optical instruments (0.644) and electrical machinery and apparatus (0.664) are clearly lower than those of some low-tech industries such as food, beverages and tobacco (1.452) and agriculture and fishing (1.424).

6. Conclusions

With this paper we have contributed to extant literature on cluster identification by designing a procedure to identify groups of industries that tend to cluster together and to analyse whether this behaviour can be explained in terms of vertical integration or by common location determinants shared by those industries. This distinction allows detailed analysis of firm location determinants and our results show that diversified clusters are not casual and are strongly determined by industry characteristics. In particular, it means that firms need “specific” neighbours in order to maximise their performance.

The methodology proposed in this paper allows the main reasons driving cluster formation to be better explained, but much more work needs to be done in this area, particularly to identify cluster size and thus better capture cluster borders. This methodology involves dividing spaces into homogeneous cells of equal size. This procedure must be handled with care because cell size influences the number and characteristics of the identified clusters. Specifically, bigger cells are more lilkely to contain a cluster, whereas smaller cells are more likely to have fewer inter-industrial clusters because the number of firms in each cell will be smaller. Given that in this paper we have assumed equal sizes for all the clusters, it would appear that using flexible sizes fits better with the real distribution of economic activity and is therefore a promising line for future research.

This is just a first attempt to better identify the forces driving cluster formation. Consequently, we have studied several types of clusters in order to provide a general overview of this phenomenon. However, this is just a starting point and further work needs to be done, in particular to cover industry specific characteristics that influence the location decisions of firms. We therefore plan to extend our analysis of specific types of clusters (both specialised and diversified) to cover several types of urban / rural environments that are hypothesised to influence such agglomerative behaviour. Finally, as we mentioned beforehand, industry aggregation is also important and, despite the computational constraints that make it unfeasible to work with such disaggregate industry-levels, we need to carry out further research to accurately determine whether our results are robust to different industry aggregation levels.

References

Arauzo-Carod, J.M. (2008): "Industrial Location at a Local Level: Comments on the Territorial Level of the Analysis", *Tijdschrift voor Economische en Sociale Geografie - Journal of Economic & Social Geography* **99**: 193-208.

Arauzo-Carod, J.M. and Manjón-Antolín, M. (2004): “Firm Size and Geographical Aggregation: An Empirical Appraisal in Industrial Location”, *Small Business Economics* **22**: 299-312

Arbia, G. (2001): “Modeling the Geography of Economic Activities on a Continuous Space”, *Papers in Regional Science* **80**: 411-424.

Barlet, M., Briant, A. and Crusson, L. (2009): “Location patterns of services in France: A distance-based approach”, Paris School of Economics Working Paper.

Boix, R. (2008): "Los distritos industriales en la Europa Mediterránea. Los mapas de Italia y España", in V. Soler (ed.), *Mediterráneo Económico*, Fundación Cajamar: Almería.

Boix, R. and Galletto, V. (2008): “Marshallian industrial districts in Spain”, *Scienze Regionali / Italian Journal of Regional Science* **7 (3)**: 29-52.

Brenner, T. (2006): “Identification of Local Industrial Clusters in Germany”, *Regional Studies* **40 (9)**: 991-1004.

Brenner, T. (2004): *Local industrial clusters: existence, emergence and evolution*, Routledge: London.

Chen, C. (1999): “Visualizing semantic spaces and author co-citation networks in digital libraries”, *Information Processing and Management* **35(3)**: 401-420.

Devereaux, M.; Griffith, R. and Simpson, H. (2004), “The geographic distribution of production activity in the UK”, *Regional Science and Urban Economics* **34 (5)**: 533-564.

Duranton, G. and Overman, H.G. (2008): “Exploring the Detailed Location Patterns of U.K. manufacturing Industries using Microgeographic Data”, *Journal of Regional Science* **48 (1)**: 213-243.

Duranton, G. and Overman, H.G. (2005): “Testing for Localization Using Microgeographic Data”, *Review of Economic Studies* **72**: 1077-1106.

Duranton, G. and Puga, D. (2004): “Micro-foundations of urban agglomeration economies”. In: Henderson, J.V., Thisse, J.-F. (Eds.), *Handbook of Regional and Urban Economics*, vol. IV. North-Holland.

Ellison, G. and Glaeser, E.L. (1997): “Geographic concentration in US manufacturing industries: A dartboard approach”, *Journal of Political Economy* **195**: 889-927.

Hoover (1936): “The measurement of industrial location”, *The Review of Economics and Statistics* **18**: 162-171.

Lambert, D.M.; McNamara, K.T. and Garrett, M.I. (2006): "An Application of Spatial Poisson Models to Manufacturing Investment Location Analysis", *Journal of Agricultural and Applied Economics* **38**: 105-121.

López-Bazo, E. (ed.) (2006): *Definición de la metodología de detección e identificación de clusters industriales en España*, Dirección General de la Pequeña y Mediana Empresa (DGPYME): Madrid.

Marshall, A. (1890): *Principles of Economics*, MacMillan: New York.

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

NWB Team (2006), Network Workbench Tool. Indiana University, Northeastern University, and University of Michigan, http://nwb.slis.indiana.edu

Olsen, J. (2002): “On the Units of Geographical Economics”, *Geoforum* **33**: 153-164.

Openshaw, S. and Taylor, P.J. (1979): “A Million or so Correlation Coefficients: Three Experiments on the Modifiable Areal Unit Problem”. In N. Wrigley, *Statistical Applications in the Spatial Sciences*, London, Pion: 127-144.

Pablo-Martí, F. and Muñoz-Yebra, C. (2009): “Localización empresarial y economías de aglomeración: el debate en torno a la agregación espacial”, *Investigaciones Regionales* **15**: 139-166.

Paluzie, E; Pons, J. and Tirado, D. (2004): “The geographical concentration of industry across Spanish regions, 1856-1995”, *Review of Regional Research* **24 (2)**: 143-160.

Porter, M. (1998): “Clusters and the new economics of competition”, *Harvard Business Review* **76 (6)**: 77-90.

Schvaneveldt, R.W. (ed.) (1990): *Pathfinder Associative Networks: Studies in Knowledge Organization*, Norwood, NJ: Ablex.

Schvaneveldt, R.W., Dearholt, D.W. and Durso, F.T. (1988): “Graph Theoretic Foundations of Pathfinder Networks”, *Computer Mathematics Applications* **15 (4)**: 337-345.

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

Sonis, M.; Hewings, G.J.D. and Guo, D. (2008): "Industrial clusters in the input-output economic system". In C. Karlsson, *Handbook of Research on Cluster Theory*, Cheltenham, Edward Elgar: 153-168.

Viladecans, E. (2004): "Agglomeration economies and industrial location: city-level evidence", *Journal of Economic Geography* **4/5**: 565-582.

Wrigley, N. (1995): "Revisiting the Modifiable Areal Unit Problem and the Ecological Fallacy". In A.D. Cliff, P.R. Gould, A.G. Hoare and N.J. Thrift (eds.), *Diffusing Geography*, Oxford, Blackwell: 49-71.

Table 1: Concentration patterns of firms at a single industry level

Code	Industry	*X*	Mean	STD	Index	X-2S	X+2S	Concentrated	Dispersed
22	Financial intermediation	882	1480,11	17,6811804	0,59590166	1444,74764	1515,47236	TRUE	FALSE
6	Paper and publishing	947	1494,58	17,9619013	0,63362282	1458,6562	1530,5038	TRUE	FALSE
13	Office machinery, computers and medical equipment, precision and optical instruments	324	502,86	13,3553001	0,64431452	476,1494	529,5706	TRUE	FALSE
26	Education	790	1209,17	17,5580164	0,65334072	1174,05397	1244,28603	TRUE	FALSE
14	Electrical machinery and apparatus	520	782,36	14,6890463	0,66465566	752,981907	811,738093	TRUE	FALSE
24	Business services	1360	1979,03	21,1557261	0,68720535	1936,71855	2021,34145	TRUE	FALSE
23	Real estate activities	1957	2803,29	18,8970069	0,69810829	2765,49599	2841,08401	TRUE	FALSE
28	Other services	1375	1819,52	21,5638493	0,75569381	1776,3923	1862,6477	TRUE	FALSE
12	Machinery and equipment	820	1076	17,2533118	0,76208178	1041,49338	1110,50662	TRUE	FALSE
4	Textiles, leather clothes and shoes	1169	1523,26	17,384319	0,76743301	1488,49136	1558,02864	TRUE	FALSE
27	Health and veterinary activities, social services	1122	1458,21	20,5029168	0,7694365	1417,20417	1499,21583	TRUE	FALSE
8	Rubber and plastic products	698	903,5	19,1498609	0,77255119	865,200278	941,799722	TRUE	FALSE
25	Public administration	141	179,3	7,24812759	0,78639152	164,803745	193,796255	TRUE	FALSE
7	Chemical products	734	837,17	14,8691634	0,87676338	807,431673	866,908327	TRUE	FALSE
10	Basic metals	567	629,55	16,7460986	0,90064332	596,057803	663,042197	TRUE	FALSE
15	Transport and communications	668	726,47	16,8111645	0,91951491	692,847671	760,092329	TRUE	FALSE
19	Trade and repair	2888	3035,78	16,6336521	0,95132058	3002,5127	3069,0473	TRUE	FALSE
16	Recycling	349	359,69	9,90020406	0,97027996	339,889592	379,490408	FALSE	FALSE
11	Fabricated metal products	1682	1701,7	19,8267751	0,98842334	1662,04645	1741,35355	FALSE	FALSE
21	Transport and communications	2090	2034,14	19,9479221	1,02746124	1994,24416	2074,03584	FALSE	TRUE
17	Construction	2706	2585,57	21,9674944	1,04657774	2541,63501	2629,50499	FALSE	TRUE
20	Hotels and restaurants	2238	2136,5	20,4181045	1,04750761	2095,66379	2177,33621	FALSE	TRUE
18	Electricity and water distribution	795	739,43	15,2674838	1,07515248	708,895032	769,964968	FALSE	TRUE
5	Wood, furniture and other manufactures	1734	1610,89	20,5956232	1,07642359	1569,69875	1652,08125	FALSE	TRUE
9	Non-metallic mineral products	1297	1125,88	18,1566027	1,15198778	1089,56679	1162,19321	FALSE	TRUE
2	Extractive activities	1152	823,16	15,7015858	1,39948491	791,756828	854,563172	FALSE	TRUE
1	Agriculture and fishing	2409	1691,54	20,5354682	1,42414604	1650,46906	1732,61094	FALSE	TRUE
3	Food, beverages and tobacco	2236	1540,31	20,5001577	1,45165584	1499,30968	1581,31032	FALSE	TRUE

Note: X-2S equals X minus 2 standard deviations and X+2S equals X plus 2 standard deviations.
Source: own calculations.

Annexes

Annex 1: List of industries

Code	Industry
1	Agriculture and fishing
2	Extractive activities
3	Food, beverages and tobacco
4	Textiles, leather clothes and shoes
5	Wood, furniture and other manufactures
6	Paper and publishing
7	Chemical products
8	Rubber and plastic products
9	Non-metallic mineral products
10	Basic metals
11	Fabricated metal products
12	Machinery and equipment
13	Office machinery, computers and medical equipment, precision and optical
14	Electrical machinery and apparatus
15	Transport materials
16	Recycling
17	Construction
18	Electricity and water distribution
19	Trade and repair
20	Hotels and restaurants
21	Transport and communications
22	Financial intermediation
23	Real estate activities
24	Business services
25	Public administration
26	Education
27	Health and veterinary activities, social services
28	Other services
	Source: SABI.