I.S.S.N.: 0213-7585

REVISTA DE ESTUDIOS REGIONALES

2ª EPOCA Septiembre-Diciembre 2018

113

SUMARIO

UNIVERSIDADES DE ANDALUCÍA

REVISTA DE ESTUDIOS REGIONALES Nº 113, I.S.S.N.: 0213-7585 (2018), PP. 221-241

A regional analysis of Spanish manufacturing firms for the period 2004-2009: A northeast/southwest pattern[1]

Un análisis regional de las empresas manufactureras españolas para el periodo 2004-2009: Un patrón Noreste/Suroeste

M. Isabel Pisa
Rosario Sánchez
Universidad de Valencia

Recibido, Junio de 2017; Versión final aceptada, Marzo de 2018.

PALABRAS CLAVE: Eficiencia, Formación, Investigación y desarrollo, Estudios regionales, Frontera estocástica.

KEY WORDS: Efficiency, Training, R&D, Regional studies, Stochastic frontier.

Clasificación JEL: J24, R11

ABSTRACT:

This paper analyses the technical efficiency of manufacturing firms located in different regions of Spain for the period 2004-2009. Spanish manufacturing firms were grouped using cluster methods. Firms located in north-eastern regions such as Catalonia or the Basque Country show higher levels of R&D, pay higher wages, and are larger than firms located in southern or western regions such as Andalusia or Extremadura. Furthermore, we found that larger firms with more than 500 workers are grouped with the highest levels of R&D, the highest labour costs, and the largest size in terms of number of workers. The results obtained with the Stochastic Frontier methodology allow us to analyse the parameters for those regions whose firms are more efficient. The results reinforce the idea of north-east manufacturing firms a higher level of technical efficiency with respect to those firms located in the south-west.

RESUMEN:

Este trabajo analiza la eficiencia técnica de las empresas manufactureras españolas para el período 2004-2009. Las empresas ubicadas en regiones del noreste muestran niveles más altos de

1 The authors wish to express their gratitude to the participants of the XLI International Conference on Regional Studies-AECR: Innovation and geographical spillover, New Approach and Evidence, for helpful comments, held at the Faculty of Business and Economics, Universitat Rovira y Virgili, November 2015.

I + D, pagan salarios más altos y presentan mayor tamaño que las empresas ubicadas en regiones meridionales u occidentales. Las empresas más grandes, se relacionan con los niveles más altos de I + D y mayores costes laborales. Los resultados obtenidos con la metodología frontera estocástica nos permiten analizar aspectos claves en la eficiencia. Los resultados verifican que empresas manufactureras del noreste tienen mayor nivel de eficiencia técnica que las ubicadas en el sudoeste.

1. INTRODUCTION

The economic disparities among regions has been analysed focusing on different aspect of the characteristics of the regions. Mas et al. (1996) provide empirical evidence, using panel data techniques, about the importance of certain forms of public capital in the improvements of productivity within the Spanish regions. In this sense, infrastructure investment was considered the key for economic growth and was one of the sources of regional development strategies of the European Union. The productive capital environment analyzed by Mancha, Moscoso and Santos (2017), set the region of Madrid in the first place for the period 2001-2014, followed by a second group of regions formed by Catalonia, Navarre, Basque Country and Aragon. Extremadura, Canary Islands and Castilla la Mancha present more modest values for this index. Pablo-Romero and Gil-Delgado (2011) obtain, by estimating a Cobb-Douglas production function from a panel data of 50 Spanish provinces, that the elasticity of human capital estimated in the Andalusian provinces is 0.22, much lower than the rest Spanish provinces whose value is 0.31.

For several years the highest portion of EU regional development funds was selected to transport structures (Rodriguez-Pose and Fratesi, 2004). The popular highway development schemes which have been at the centre of development strategies for Spain and Portugal, among others southern countries, have not obtained the expected outcome. The seminal work of Krugman (1991) gave a new justification for the phenomenon of accumulation of economic activities in space given by using general equilibrium models beached in microeconomic decision where the crucial element was the existence of increasing returns at the firm level and transportation cost. This means that what is important to see agglomeration dynamics is how far a location is from its consumer markets. More central locations in Europe offer higher wages for skilled labour which increases the individual's incentives for human capital accumulation (Lopez-Rodriguez et al., 2007). According to this argument resident workers in remote locations have lower incentives to invest in human capital (Redding and Schott 2003). Additionally, Lopez-Rodriguez et al. (2011) obtain that remoteness explains the differences in convergence in GDP pc in Romanian regions. The basic idea is the relevance of trade costs in reducing per capita income. In a world where countries or regions specialize in particular goods and exports, firms in remote locations further away from consumers will incur in higher trade costs and consequently obtain lower net revenues from their export sales. Hence, it is clear that being located in the economic periphery can reduce the return to skill and then also reduce the investment in human capital accumulation. The New Economic Geography (NEG) explains how remoteness (distance to consumer markets and sources of inputs) may prevent convergence and provoke the emergence of a heterogeneous economic space. In a recent paper Faiña et al (2017) show how these peripheral Spanish regions with low access to the central markets suffer a penalty for their growth and convergence. Also, Hervas-Oliver et al. (2018), with a sample of 6697 Spanish manufacturing firms, show that agglomeration gains exist but not all firms benefit equally.

In this paper we try to go further analysing the differences in productivity that arise for those firms located in different autonomous regions in Spain. To undertake this analysis, we use the stochastic frontier methodology as a new technic to be used in this field.

The stochastic frontier methodology has been used to measure the technical inefficiency/efficiency of a production unit as the ratio of a firm's production over its optimal level of production (given their levels of inputs and technology). What it means is that firms are compared with those firms in the sample, with the same characteristics (in terms of inputs and technology), that show the best results in terms of production. With the Inefficiency model (that is estimated simultaneously) we can analyse the variables that make firm close or far away from the frontier of the most efficient firms. The characteristic of this methodology is quite useful to analyse the differences among manufacturing firms located in different regions of Spain and how it conditions their performance in terms of technical efficiency.

Using frontier techniques, several studies have analysed which are the sources of technical inefficiency. Caves and Barton (1990) examine technical inefficiency of the manufacturing industry in United States, while Green and Mayes (1991) analyse technical inefficiency for United Kingdom firms. Caves et al. (1992) compare inefficiency and its determinants between developed countries. Other studies focus on particular determinants of inefficiency, such as the Hay and Liu study (1997), which focuses on the relevance of a competitive environment on efficiency; Patibandla (1998), who shows the relevance of capital market imperfections on the structure of an industry; and Dilling-Hansen et al. (2003), who analysed whether relative efficiency is due to R&D investment. In Spain, among others, Díaz and Sánchez (2008, 2013), use this methodology to analyses differences in efficiency due to size and investment on R&D or Pisa and Sánchez (2016), that obtain empirical evidence in favour of the idea that those firms that pay higher wages gain in productivity.

This paper differs from the previous Spanish regional analysis literature due to the connection between technical efficiency and location. With both the cluster technique and the stochastic frontier methodology, we obtain the classification of regions based on the type of firms located in each region. The behaviour of these firms, in terms of investment on R&D, proportion of permanent worker over temporal, and the wages that they are willing to pay, are conditioned for their geographical situation in Spain.

The paper is organized as follows. Section 2 provides information about the Spanish industrial survey on Business Strategies (ESEE). Section 3 examines the regional grouping of Spanish firms given their characteristics. Section 4 introduces the stochastic frontier methodology. The main results concerning location and other internal factors of firms are presented and discussed in section 5. Section 6 summarizes the main conclusions.

2. THE DATA: THE SPANISH INDUSTRIAL SURVEY ON BUSINESS STRATEGIES (ESEE)

The information source is published in the Spanish Industrial Survey on Business Strategies (Encuesta Sobre Estrategias Empresariales, ESEE). The data is collected by the Fundacion SEPI and supported by the Spanish Ministry of Industry. This information is provided as a panel of firms representing twenty industrial sectors. A characteristic of the survey is that companies were chosen through a careful selection structure. The sample of firms includes almost all Spanish manufacturing firms with more than two hundred employees. Firms with a number of employees between ten and two hundred employees were chosen according to a stratified random sample representative of the population of small firms. Given the procedure used to select the firms participating in the survey, both samples of small and large firms allow the estimation of the distribution of any of the characteristics of the population of Spanish manufacturing firms with information available from the ESEE data set. Each year a number of additional firms were selected using a random sampling procedure with the whole population of firms. This selection is conducted using the same proportion as in the original sample (see Fariñas and Jaumandreu, 2004, for technical details about the sample).

From the original sample, a number of firms were eliminated, either because of a lack of relevant data, or because they reported a value-added annual growth rate per worker of more than 500% (in absolute value). Either of these two cases would distort the analysis. Moreover, firms after a merger or division process were excluded from the sample. The sample contains 2,117 firms from the ESEE Survey, and it refers to an unbalanced panel where firms without two consecutive years of data were eliminated. The analysis comprises a time period from 2004 to 2009, with 9435 observations in all.

3. REGIONAL GROUPING OF SPANISH MANUFACTURING FIRMS

Summary of relevant data of the seventeen Spanish Regions (Autonomous Communities) are presented in Table 1 and 2.

TABLE 1
SPANISH REGIONS VALUES

Spanish Regions	Number of Firms	(*)Average Expenditure in R&D per worker	Proportion of permanent workers	**(*)Average Level of export per worker**
Andalusia	827	6.20	0.72	296.11
Aragon	351	15.40	0.85	706.25
Asturias	231	13.87	0.78	383.9
Balearic Islands	121	0.18	0.82	64.84
Basque Country	779	25.78	0.87	743.88
Canary Island	144	0.30	0.86	4.14
Cantabria	112	15.01	0.88	664.90
Castile-La Mancha	520	8.12	0.8	387.52
Castile-Leon	515	15.77	0.85	506.90
Catalonia	1881	22.24	0.9	573.35
Extremadura	135	4.41	0.81	312.98
Galicia	570	10.95	0.78	1362.72
La Rioja	90	7.57	0.81	357.91
Madrid	1381	13.92	0.87	282.55
Murcia	245	6.76	0.75	827.73
Navarre	232	17.62	0.87	687.78
Valencia	1301	8.14	0.82	536.26
Spain	**9435**	14.08	0.84	441.1

Source: (*) In real terms, own calculation from ESEE.

Table 1 shows differences among regions in the number of manufacturing firms included in the sample of ESEE. Catalonia, Madrid and Valencian' community are the regions with a higher number of manufacturing firms. These communities are the most densely populated areas, together with the Basque country. In contrast, the less densely populated areas are Castile-La Mancha, Castile –Leon and Extremadura.

The second column of Table 1 shows the average expenditure in R&D per worker. The highest investment on R&D is found in The Basque country followed by Catalonia. The bigger percentage of permanent workers corresponds to Catalonia followed by Cantabria, Basque Country, Madrid and Navarre, while the lowest percentage is found in Andalusia.

In Table 2, are the distribution of firms' size by regions. The average proportion of small firms in the industrial sector in Spain is also higher than in other European countries, with around 50% having between 10 and 50 workers. The highest proportion of this type of firms is located in the Balearic Islands (89%), whereas the lowest is in Navarra (31%). Moreover, Navarra is the Spanish region that accounts for the highest number of size 5 manufacturing firms, with more than 500 workers, whereas the lowest number of these firms corresponds to the Canary Islands and La Rioja (2%), followed by the Balearic Islands (3%).

TABLE 2
SPANISH REGIONS SIZE

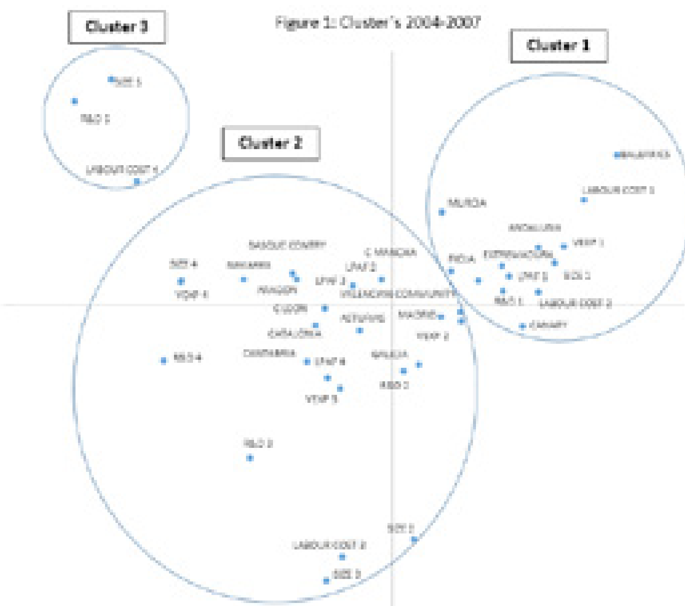
Spanish Regions	Size 1: From 10 up to 50 workers	Size 2: From 51 up to 100 workers	Size 3: From 101 up to 200 workers	Size 4: From 201 up to 500 workers	Size 5: Higher than 500 workers
Andalusia	0.63	0.12	0.08	0.14	0.04
Aragon	0.34	0.07	0.15	0.3	0.14
Asturias	0.39	0.17	0.13	0.19	0.12
Balearic Islands	0.89	0.05	0.02	0.01	0.03
Basque Country	0.36	0.09	0.15	0.29	0.11
Canary Island	0.52	0.09	0.15	0.21	0.02
Cantabria	0.32	0.18	0.14	0.3	0.05
Castile-La Mancha	0.52	0.08	0.11	0.22	0.07
Castile-Leon	0.38	0.15	0.12	0.26	0.09
Catalonia	0.41	0.12	0.11	0.24	0.11
Extremadura	0.54	0.16	0.07	0.14	0.09
Galicia	0.4	0.18	0.15	0.17	0.09
La Rioja	0.75	0.06	0	0.17	0.02
Madrid	0.56	0.11	0.12	0.13	0.08
Murcia	0.56	0.07	0.08	0.13	0.15
Navarre	0.31	0.09	0.16	0.21	0.23
Valencia	0.6	0.12	0.1	0.1	0.08
Spain	0.49	0.12	0.12	0.19	0.09

Source: Own calculation from ESEE

In this section our objective is to analyse the performance of the manufacturing firms and how it is related to their location, to this end, the Spanish regions are grouped using a sample of variables that can define their characteristic. These variables are the firm's exporting value, investment in Research and Development (R&D), labour cost per worker, firm size and the proportion of permanent workers. To make the clusters we have split the sample into two periods to distinguish the boom and the recession period. The precise definitions for the categorization of these variables are found in the Appendix section.

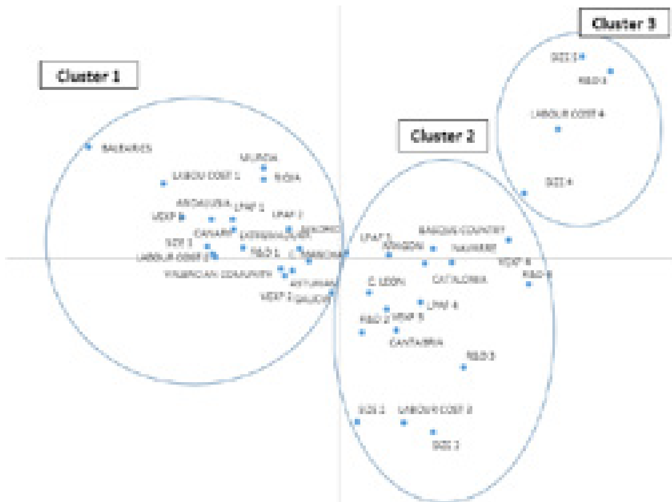
Figure 1 displays the three clusters obtained for the period 2004-2007. Cluster 1, show the regions with lower levels on R&D investment, exports, labour cost per worker and proportion of permanent workers. These regions are Andalusia, Balearic Islands, Canary Island, Extremadura, La Rioja and Murcia. In cluster 2, are grouped the remaining eleven regions with higher values of R&D, exports, labour cost per worker, higher size and proportion of permanent workers. In cluster 3, do not appear any autonomous community indicating that the manufacturing firms with the highest values of investment on R&D, size and labour cost per worker could be located in everywhere.

FIGURE 1
CLUSTER´S 2004-2007



In Figure 2, are grouped the regions for the recession period from 2008 to 2009. In this grouping the cluster 3 is again unrelated with any particular region and show the highest values of size, R&D and labour cost per worker. The most important difference with respect to the boom period is that some regions move from cluster 2 to cluster 1. These regions are Asturias, Castile-La Mancha, Valencian Community, Galicia and Madrid.

FIGURE 2
CLUSTER´S 2008-2009



This type of analysis is quite useful for the exploration of data and the construction of homogeneous groups within the whole sample. Sole and Ríos (2008) analyse the socio-economic convergence of the Spanish regions between 1990 and 2000. In 2000 samples, they found that Madrid, the Basque Country, and Catalonia were the regions placed in the cluster with the highest level of socio-economic welfare. Here we obtain the same result for the boom period but not for the recession period where Madrid move to the low performance cluster.

Table 3, shows the values of the labour cost per worker for both periods. If we analyse the differences in terms of labour cost per worker between these two periods we can remark that Catalonia is the regions where higher is the increment of this variable, widening the difference with respect to Madrid. In the period of boom, the Basque country is the region with the highest value of labour cost per worker followed by Navarre and Catalonia. In the recession period, still the Basque country is the region with the highest labour cost per worker, but now, the second position is for Catalonia that is the region with the highest increment in the recession period with respect to the boom period. In general, the labour cost per worker increased for the period 2008-2009.

With this information, we will carry on the estimation of the stochastic frontier to obtain the parametric information to distinguish the significant variables that discriminate firms located in different regions.

TABLE 3
LABOUR COST PER WORKER

	Period 2004-2007				Period 2008-2009			
Regions	Mean	Min	Max	Standard Deviation	Mean	Min	Max	Standard Deviation
Andalusia	24.278.12	6.006,14	75.376,45	10.881,88	25.471,98	5.800,50	108.251,09	12.424,63
Aragon	31.780,19	13.909,10	59.384,03	10.562,84	31.926,25	7.493,72	70.961,82	10.101,32
Asturias	35.173,21	11.587,82	79.198,84	13.746,41	32.902,83	12.467,50	67.117,24	11.078,41
Balearic Islands	24.298,61	8.800,05	51.739,96	9.753,51	24.549,63	12.015,16	61.566,20	8.300,59
Canary Islands	32.296,61	11.875,92	63.358,49	12.793,06	31.223,00	12.524,07	79.611,78	13.703,63
Cantabria	34.474,54	12.868,71	57.882,41	12.798,15	31.864,20	12.001,03	66.343,50	11.979,04
Castile-Mancha	25.504,47	11.028,09	90.568,50	11.475,38	29.361,68	10.058,47	151.377,76	13.129,72
Castile- León	32.374,02	12.721,85	59.830,05	11.456,95	31.215,18	10.516,61	71.407,48	9.970,32
Catalonia	35.729,69	9.901,60	106.713,95	13.208,77	38.281,20	11.656,29	284.705,58	16.690,96
Valencian' Community	26.603,46	5.779,21	77.741,49	9.579,58	29.168,00	7.113,98	195.340,21	12.665,17
Extremadura	20.994,15	7.186,62	60.550,54	10.333,12	23.873,46	5.620,60	58.281,36	10.591,12
Galicia	27.146,07	8.742,29	75.61,50	12.795,14	28.069,58	10.046,52	87.911,61	12.816,09
Madrid	33.696,13	10.865,16	373.429,50	18.961,85	34.114,52	9.880,04	134.976,02	13.234,09
Murcia	23.829,11	9.474,17	57.982,65	10.239,82	25.240,02	10.871,79	54.631,89	9.722,19
Navarre	36.813,01	15.599,50	71.180,91	11.350,83	38.107,44	17.421,27	102.889,23	12.864,13
Basque Country	38.497,23	11.934,70	109.548,88	11.976,04	38.831,86	14.813,44	96.276,55	11.583,93
La-Rioja	28.206,84	15.364,33	43.368,24	7.445,73	29.844,22	14.600,11	68.664,06	9.613,18

Source: Own calculation from ESEE

4. THE STOCHASTIC FRONTIER ANALYSIS

Differences in productivity among regions could be viewed as the result of some features. One of them could be technical innovations in processes and/or products that drive the frontier of potential production upward for regions that make this type of investment. Others are related to efficiency changes that reflect firms' ability to increase production with given inputs and existing technology. Given that productivity and technical efficiency are not completely separable, we use different elements that allow us to analyse the performance of firms located in different regions of Spain.

In previous section we have grouped the firms located in different regions to obtain the relations between location and some specific factors that could improve the production of manufacturing firms. Now we will use the stochastic frontier methodology to discriminate among location and technical efficiency performance. Explicitly, a panel data version of the Aigner et al. (1977) approach, following Kumbhakar and Lovell (2000) specification, is developed, in which technical inefficiency is estimated from the stochastic frontier and simultaneously explained by a set of variables representative of the firms' features. This approach avoids the inconsistency problems of the two-stage approach used in previous empirical studies when analysing the inefficiency determinants[2].

The model can be expressed as:

$$Y_{it} = f(X_{it};\beta)\exp(v_{it} - u_i) \tag{1}$$

Where *i* indicates firms and *t* represents the period, *X* is the set of inputs; β is the set of parameters, v_{it} is a two-sided term representing the random error, assumed to be *iid* $N(0,\sigma_v^2)$; u_i is a non-negative random variable representing the inefficiency, which is assumed to be distributed independently and obtained by truncation at zero of $N(\mu,\sigma_u^2)$. The average of this distribution is assumed to be a function of a set of explanatory variables.

Given that technical efficiency is the ratio of observed production over the maximum technical output obtainable for a firm (when there is no inefficiency), the efficiency index (TE) for firm *i* in year *t* could be written as[3]:

2 In a two-stage procedure, first of all a stochastic frontier production function is estimated and the inefficiency scores are obtained under the assumption of independently and identically distributed inefficiency effects. But in the second step, inefficiency effects are assumed to be a function of some firm-specific variables, which contradicts the assumption of identically distributed inefficiency effects.

3 Individual efficiency scores u_i, which are unobservable, can be predicted by the mean or the mode of the conditional distribution of u_i given the value of $(v_i - u_i)$ using the technique suggested by Jondrow et al (1982).

$$TE = \frac{f(X_{it};\beta)\exp(v_{it} - u_i)}{f(X_{it};\beta)\exp(v_{it})} = \exp(-u_i) \tag{2}$$

The efficiency scores obtained from expression (2) takes a value of one when the firm is efficient, and less than one otherwise.

The stochastic trans-logarithmic panel data production function is estimated by adding a term of inefficiency, whose mean is the function of a set of inefficiency determinants[4]. For estimation purposes, the econometric software Limdep version 8.0 (Green, 2002) was used.

$$\ln Y_{it} = \beta_0 + \sum_{j=1}^{J} \beta_j \ln X_{ijt} + \frac{1}{2}\sum_{j=1}^{J}\sum_{k=1}^{K} \beta_{jk} \ln X_{ijt} \ln X_{ikt} + \sum_{m=1}^{M} \phi_m S_{im} + v_{it} - u_i \tag{3}$$

$$\mu_{(u)} = \mu_{(u)} \exp(\delta' Z) \tag{4}$$

The variables used to estimate the production frontier are the value-added (Y_{it}), as the output variable, and the number of employees in the firm, capital stock and trend, as input variables (X_{it}), industrial sector dummies (S_i), and two dummies that indicate whether firms have undertaken product (INPR) or process innovation (INP). A more precise definition of the variables used to estimate and define the inefficiency determinants is introduced in the Appendix section.

The function coefficients (β) and inefficiency model parameters (δ) were estimated using the panel data technique to control for unobserved heterogeneity.

5. THE RESULTS

The estimation of the SFA provides the value of a firm's technical efficiency associated with the best observations of the sample. The value of the estimates allows us to explain the differences in the inefficiency effects in firms belonging to different regions in Spain. This section presents the estimations of value-added stochastic frontier to discriminate between the two groups of region obtained in the cluster analysis.

From the frontier approach, the measures of inefficiency of firms were obtained compared to the best observation of the sample. Technological and market conditions can change across the sector. Therefore, to control for this, we included sector dummy variables in the value-added production function.

4 The usual symmetry conditions are imposed on the trans-logarithmic function.

TABLE 4
STOCHASTIC FRONTIER: TRANS LOGARITHMIC VALUE ADDED PRODUCTION FUNCTION

Variables	Parameters	Coefficients	t-value
Constant	β_0	**7.036***	14.864
T	β_1	**0.1549***	9.068
L	β_2	**1.081***	20.998
K	β_3	-0.012	-0.657
K^2	β_{11}	**0.0269***	13.211
L^2	β_{22}	**0.0343***	4.052
T^2	β_{33}	**-0.0147***	-8.480
KxL	β_{23}	**-0.0133***	-8.189
LxT	β_{12}	**0.0214***	5.671
KxT	β_{13}	**-0.0165***	-6.863
Product Innovation	θ_1	**0.0255***	1.683
Process Innovation	θ_2	**0.0297***	2.541
Sector Classification. Category of Reference: Other Manufacturing Products			
Meat and manufacturing of meat; food industry and tobacco drinks….	γ_1	**-0.1449***	-4.648
Wood and derived, paper and derived	γ_2	-0.0696	-1.600
Chemical products; couch and plastic; non-metallic mineral products.	γ_3	0.0489	1.488
Basic Metal Products; fabricated metal products; industrial equipment	γ_4	**0.0938***	3.011
Office machinery and others; electric materials;	γ_5	**0.0703***	1.693
Cars and engine; other material transport	γ_6	-0.0531	-1.288
Inefficiency Model: Model I			
Constant	δ_0	**6.885***	5.449

TABLE 4
STOCHASTIC FRONTIER: TRANS LOGARITHMIC VALUE ADDED PRODUCTION FUNCTION (CONCLUSIÓN)

Variables	Parameters	Coefficients	t-value
Spanish Regions. Category of Reference: Andalusia, Balearic Islands, Canary Islands, Extremadura, Murcia and La Rioja			
Aragon	δ_1	-0.0987	-0.708
Asturias	δ_2	**-0.3126****	-1.898
Cantabria	δ_3	-0.0383	-0.136
Castile-La Mancha	δ_4	-0.1840	-1.400
Castile-Leon	δ_5	**-0.3646*****	-2.681
Catalonia	δ_6	**-0.4847*****	-5.743
C-Valencia	δ_7	0.1442	-1.614
Galicia	δ_8	-0.0330	-0.308
Madrid	δ_9	**-0.6129*****	-6.687
Navarra	δ_{10}	-0.2566	-1.247
Basque Country	δ_{11}	**-0.3181*****	-2.887
Recession. Category of Reference boom	δ_{12}	**-0.4120*****	-5.997
Exports per worker	δ_{13}	**-0.0002*****	-11.992
Firms 'expenditure in R&D per worker	δ_{14}	**-0.0021*****	-3.842
Proportion of permanent workers over total labour force	δ_{15}	**-0.4955*****	-3.052
Size of Firm. Category of Reference: Firms with a number of workers from 10 up to 50.			
From 51 up to 100	δ_{16}	**-0.4721 *****	-4.963
From 101 up to 200	δ_{17}	**-0.5060*****	-4.662
From 201 up to 500	δ_{18}	**-0.8811*****	-7.240
More than 500	δ_{19}	**-1.5093*****	-9.075
Variance Components			
Lambda	λ	**0.9764*****	57.189
Sigma(u)	σ_u^2	**0.3578*****	59.202

Note: (***) Significant at 1%; (**) Significant at 5%; (*) Significant at 10%;

Table 4 shows the maximum-likelihood estimates of the production frontier parameters defined in equation (3), given the inefficiency specification defined in equation (4).

The estimation offers the value of a firm's technical inefficiency associated with the best observations of the sample. These values allow us to explain the differences in the inefficiency effects among firms located in different geographical areas in Spain. With the classification obtained from the cluster technique, we estimate the firms located in different regions using as category of reference those regions that belongs to the clusters of worse performance in terms of investment on R&D, exporting outcomes, size or proportion of permanent workers. In The first part of the estimation of Table 4 we have the results of the trans logarithmic value added production function. With the estimation of the production function, we obtained the values to build the frontier of the most efficient firms, allowing us to analyse, with the inefficiency model, the factors that could bring the firm closer to the frontier (reducing inefficiency) depending on the region where they are located.

The coefficients for the parameters of the value-added production function are as expected. The coefficient for capital is not significant. This result is similar to what was obtained by Green and Mayes (1991) and Díaz and Sánchez (2008). Capital-intensive businesses require greater sunk costs, and so they can find it more difficult to modify performance as demands and technology change. This means that even if an increase in the stock of capital improves efficiency, doing this with a different timing from the rest of the firms could cause losses in productivity stemming from the short-term capital adjustment.

Regarding the results obtained in the non-parametric analysis, we decide to include in the inefficiency model for the whole sample, the regions grouped in cluster 2 in the period (2004-2007), taking as category of reference the regions grouped in cluster 1 also in the period (2004-2007). Also we have introduced a variable that takes into account the differences between the boom and recession period. This is a dummy that measures the recession period with respect to the boom. Our main goal is to discriminate among the regions with the most efficient firms in terms of technical efficiency. A negative and significant coefficient reduces the distance to the frontier of the best firms, in terms of value added given the same level of inputs. It was found that being located in Asturias, Castile-Leon, Catalonia, Madrid and Basque country brings firms closer to the efficient frontier of the best firms with respect to those firms placed in Andalusia, Balearic Islands, Canary Islands, Extremadura, Murcia and La Rioja. While those firms located in Aragon, Cantabria, Castile-La Mancha, Valencian Community, Galicia and Navarre are the furthest away from the frontier of best firms of the sample. Faiña et al. (2017) claim that Basque Country, Catalonia, Madrid and Navarre, are the most developed regions and are also the regions of the highest productivity levels. Here we obtained the same results with the exception of Navarre in terms of technical efficiency.

The coefficient of the recession variable is negative and significant which means that firms that survived in the recession period were closer to the frontier. The higher the value of the export per worker, the closer the firm is to the frontier, as are the firms that invest in R&D and those that have a higher proportion of permanent workers. Size, in terms of the number of workers in manufacturing firms, is an important determinant of technical inefficiency. These coefficients are negative and significant, and the impact is higher as firm size increases. In fact, there is less technical inefficiency in firms with more than 50 and less than 500 workers. This result suggests that the smaller a business is, the higher the distance to the frontier is compared with the most efficient firms in the sample. This result may be partially explained by the fact that large businesses spend much more on research and development than medium and small-sized businesses in Spain.

As we showed in previous sections, the highest values of these variables were found in firms placed in the North-Eastern regions.

A large body of literature has analysed the effect of innovation on productivity[5]. In addition, the effect of size on innovation activities has been widely investigated. Size has been found to be one of the factors that explain firms' differences in innovation activities and in their returns on R&D expenditures[6]. Most studies establish that large firms are more innovative than small and medium-sized firms. Large firms can benefit from scale economies, better qualified workers, greater superior access to external financial funds, and a greater capacity to exploit innovation and develop new production methods. On the one hand, some empirical evidence shows that, to some extent, there is a linear relationship between R&D investment and size. Large firms innovate more and obtain higher returns on their investment (Diaz and Sanchez, 2013). In a recent paper Capello and Lenzi (2018) suggest that diversification of economic/innovative activities can be a good option to pursue change, but it is not enough if it is an isolated activity.

A large number of studies have analysed the relationship between innovation and firms' productivity growth (Cohen and Keppler, 1996; Crépon et al., 1998; Griliches, 1979 and Hall and Mairesse, 1995), and between exports and productivity (De Loecker, 2007; Kim et al., 2009; Delgado et al., 2002 and Fariñas and Martin-Marcos, 2007). Some papers have considered the relationship between innovation activities and exports as complementary factors that enhance firms' productivity (López Rodriguez and García Rodriguez, (2005) and Baldwin and Gu (2004)). Golovko and Valentini, (2011) hypothesize that innovation and exports are two complementary determinants of firms' growth. Using a sample of Spanish manufacturing firms during the 1990-1999 periods, they obtained empirical evidence supporting

5 Griliches (1979), Crépon et al. (1998) and Huergo and Jaumandreu (2004).

6 See Cohen and Klepper (1996).

the hypothesis that the positive effect of innovative investment on firms' growth is greater for firms that participate in export markets. Sánchez and Díaz (2013) found that innovation was an important determinant of efficiency for large firms, but not for small and medium-sized firms. Large firms are better able to obtain external financial funds and finance their R&D activities, obtaining product and process innovation that allows them to gain competitiveness in foreign markets. Size is also related to firms' ability to compete in foreign markets.

These results reinforce the idea that a good industrial policy designed to improve the level of investment of Spanish industrial firms will increase the levels of efficiency of those firms located in the south-west of the country compared to those situated in the north-east. Goecke and Hünther (2016), obtain that the positive effect of the relative size of the manufacturing sector makes clear the importance of manufacturing for the convergence process in the EU from 2000 to 2011.

In summary, empirical evidence has been provided in this section supporting the idea that there are differences in productivity among firms belonging to different regions in Spain. Regions in the northern and north-eastern parts of Spain have a better pool of firms with higher proportion of exporting per worker, higher investment in R&D, higher proportion of permanent workers and larger sizes than firms located in regions mostly in the southern and western part of Spain.

Further research is needed in this area to improve the performance of those firms located in areas with difficult access to markets and with more difficulties in R&D investment. Furthermore, a policy of higher wages in this area could attract workers with the necessary skills to improve productivity.

5.1. Testing for Robustness

The lambda parameter ($\lambda = \sigma u/\sigma v$) is positive and significant, which means that inefficiency is stochastic; therefore, the frontier model cannot be reduced to a mean-response production function value-added equation (OLS estimation). In the estimated value-added of the frontier equation, the variance parameter for "u" (σu^2) is a significant component of the total error term variance for the stochastic frontier analysed; thus, deviation from the potential value-added production is not only due to random factors. The results show that technical inefficiency is stochastic and relevant in obtaining an adequate representation of the data.

6. CONCLUDING REMARKS

This paper has analysed the regional locations of Spanish manufacturing firms and how this geographical situation can be associated with specific firm characteristics, such as investment in R&D, proportion of permanent contracts, labour costs, or size in terms of number of workers.

We have split the sample in two periods to observe differences in the performance of firms between the boom and recession period. The first period was from 2004 to 2007 and the second period from 2008-2009. Three clusters were obtained for both periods using cluster techniques. Regions are grouped in two clusters by different firm characteristics, such as levels of investment in R&D, proportion of permanent contracts, labour costs, and size judged by number of workers. For the boom period, we obtain the following-trends. The regions with low performance are clustered in cluster 1. These include Andalusia, Balearic Islands, Canary Island, Extremadura, Murcia and La Rioja in terms of the variables mentioned before. The regions with better performance are grouped in cluster 2, which is composed of the remaining of the eleven regions. Cluster 3 represents the best performance firms that correspond to the largest size, highest R&D, and highest labour cost per worker, but whose high performance is not linked with regions. The differences between the boom and recession period arise from the movement of Madrid, among other regions, from cluster 2 to cluster 1, indicating that the recession worsened the performance of the manufacturing firms located in Madrid.

To go further in our analysis, we estimate the stochastic frontier for the whole sample of firms, using as category of reference regions located in the cluster with low performance firms. We obtain that Asturias, Basque Country, Castile- Leon, Catalonia, and Madrid are the regions with manufacturing firms closer to the frontier of the best performance firms. The firms that belong to these regions are those that pay higher wages, spend more in R&D, have a high proportion of permanent workers and are larger than the category of reference that were the regions belonging to the low performance cluster. The results indicate differences between the profiles of firms located in the northeast with respect to those firms located in the southwest. So, these results reinforce the idea of a north-east/south-west pattern in terms of technical efficiency.

The agglomeration of firms in these areas could also be related to the proximity to central markets in Europe. Here, the findings show that firms located in the north-eastern part of Spain pay higher wages to their workers and invest more in R&D than firms located in western and southern areas of Spain. This latter result could be compatible with the NEG theory.

Further research is needed in this area to improve the performance of firms located in the south-west regions. Also an industrial policy directed to facilitate the investment in R&D could help to reduce the distance between northeast and southwest regions.

REFERENCES

AIGNER D., LOVELL K. and SCHMIDT P. (1977): "*Formulation and estimation of stochastic frontier production function models"*, *Journal of Econometrics,* 6, 21–37.

BALDWIN, J.R. and GU, W. (2004): "*Trade liberalization: export-market participation, productivity growth, and innovation*", *Oxford Review of Economic Policy*, 20-3, 372-392

BREINLICH, H. (2006): "*The spatial income structure in the European Union: what role for economic geography?*", *Journal of Economic Geography*, 6, 593-617.

CAPELLO, R. and LENZI, C. (2018): "*Regional innovation patterns from an evolutionary perspective", Regional Studies*, 52:2, 159-171.

CAVES, R.E. and BARTON, D.R (1990): "*Efficiency in US manufacturing industries"*. MIT Press, Cambridge, MA.

CAVES, R.E. (1992): "*Industrial Efficiency in six nations*". MIT Press, Cambridge, MA.

COHEN, W. M. and KLEPPER. S. (1996): "*Firm Size and the Nature of Innovation within Industries: The Case of Process and Product R&D*", *Review of Economics and Statistics*, 78, 232–243.

CRÉPON, B., DUGUET, E. and MAIRESSE, J. (1998): "*Research, innovation, and productivity an econometric analysis at the firm level*", *NBER Working* Paper 6696

DE LOECKER, J. (2007): "*Do exports generate higher productivity? Evidence from Slovenia"*, *Journal of International Economics*, 73, 69–98

DELGADO, M.A., FARIÑAS, J.C. and RUANO, S. (2002): "*Firm productivity and export markets: a non-parametric approach", Journal of International Economics*, 57, 397–422

DÍAZ, M. A. and SÁNCHEZ R. (2008): "*Firm size and productivity in Spain: a stochastic frontier analysis*", *Small Business Economics*, 30, 315-323**.**

DILLING, M., MADSEN E.S and SMITH, V. (2003): "*Efficiency, R&D and ownership-some empirical evidence*", *International Journal of Production Economics*, 83: 85-94

FAIÑA, J.A., LÓPEZ, J., MONTES, P., ROMERO, I. and VARELA-CANDAMIO, L. (2017): "*Regional Development and Structural Change. The productivity paradox of Spanish Peripheral Regions",* In *Regional upgrading in Southern Europe (Fonseca and Fratesi, Eds) Springer, Advanced in Spatial Sciences.*

FARIÑAS, J.C. and JAUMANDREU, J. (2004): "*Diez años de encuesta sobre estrategias empresariales (ESEE)*", *Economía Industrial***,** 329, 29-42

FARIÑAS, J.C. and MARTÍN, A. (2007): "*Exporting and economic performance: Firm-level evidence of Spanish manufacturing*", *The world Economy,* 3, 618-629.

GOECKE, H. and HÜNTHER, M. (2016): "*Regional convergence in Europe"*, *Intereconomics, Springer, Heidelberg*, 51-3, 165-171.

GOLOVKO, E. and VALENTINI, G. (2011): "*Exploring the complementarity between innovation and export for SMEs' growth", Journal of International Business Studies* 42, 362-380

GREENE, A., and MAYES, D.G. (1991): *"Technical efficiency in manufacturing industries", The Economic Journal,* 101, 523-538

GREENE, W., (1997): "*Frontier Production Functions*". *In Pesaran, M.H., Schmidt, P. (Eds), Handbook of Applied Econometrics*, Microeconometrics Blackwell Publishers, Oxford

GREENE W., (2002): *"Econometric modelling guide: Limdep Version 8.0"*, *Econometric Software, Inc., Castle Hill*.

GRILICHES, Z. (1979): "*Issues in Assessing the Contribution of Research and Development to Productivity Growth. Bell Journal of Economics*", *The RAND Corporation,* 10-1: 92-116.

HALL, B.H. and MAIRESSE, J. (1995): "*Exploring the relationship between R&D and productivity in French manufacturing firms", Journal of Econometrics* 65-1, 263-293

HAY, D. and LIU, G. (1997): "*The efficiency of firms: what difference does competition make?", The Economic Journal* 107, 597-61

HEAD, K. and MAYER, T. (2006): "*Regional Wage and Employment Response to Market Potential in the EU*", *Regional Science and Urban Economics* 36, 573-594.

HERVAS-OLIVER, J.L; SEMPERE-RIPOLL, F; ROJAS-ALVARADO, R and ESTELLES-MIGUEL, S. (2018): "*Agglomerations and firm performance: who benefits and how much?", Regional Studies* 52-3, 338-349.

HUERGO, E. & JAUMANDREU, J. (2004): "*Firms' age, process innovation and productivity growth*", *Journal of Industrial Organization*, 22, 541-559.

JONDROW J., LOVELL K., MATEROV I.S. and SCHMIDT T P. (1982): "*On the estimation of technical inefficiency in the stochastic frontier production function model", Journal of Econometrics*, 19, 233–238.

KIM, S., M. GOPINATH and KIM, H. (2009): "*High Productivity Before or After Exports? An Empirical Analysis of Korean Manufacturing Firms", Journal of Asian Economics*, 20-4, 410-418.

KRUGMAN, P. (1991): "*Increasing Returns and Economic Geography", Journal of Political Economy*, 99 -3, 483-499.

KUMBHAKAR and LOVELL (2000): "*Stochastic Frontier Analysis", Cambridge University Press.*

LEVINE, D., (1992): "*Can wage increases pay for themselves? Test with a production function", The Economic Journal*, 102, 1102-15.

LOPEZ, J and FAIÑA, J. A (2007): "*Human Capital Accumulation and Geography: Empirical Evidence from the European Union*", *Regional Studies*, 41-2, 217-234.

LÓPEZ J., FAIÑA J.A and BOLEA G. (2011): "*Economic remoteness and wage disparities in Romania*", *Tijdschrift voor Economische en Sociale Geografie*, 102-5, 594-606.

LÓPEZ, J. and GARCÍA, R. (2005): "*Technology and export behaviour: A resource-based view approach*", *International Business Review* 14, 539–557

MANCHA-NAVARRO, T., MOSCOSO-DURAN, F. and SANTOS, J.L. (2017): *"Un índice de competitividad regional para España"*. *Revista de Estudios Regionales, nº 109, 67-94.*

MAS, M., MAUDOS, J., PEREZ, F., and URIEL E. (1996): "*Infrastructures and productivity in Spanish Regions*", *Regional Studies*, 30 -7, 641-649

PABLO-ROMERO GIL-DELGADO, M.P. and GOMEZ-CALERO VALDÉS, M.P. (2011): *"Efecto del Capital Humano sobre la productividad: Andalucía y el resto de España", Revista de Estudios Regionales, nº 90, 45-70.*

PATIBANDLA, M. (1998): "*Structure, organizational behaviour, and technical efficiency: The case of an Indian industry", Journal of Economic Behaviour and Organization* 34, 419-434.

PERAITA, C. (2005): "*Firm-sponsored training in regulated labour markets: evidence from Spain*", *Applied Economics*, 37-16, 1885-1898

PISA, M.I., SÁNCHEZ, R. (2016): "*Work incentive and productivity in Spain*", *Prague Economic Papers*, 25-01, 99-111.

REDDING, S. and SCHOTT, P. (2003): "*Distance, skill deepening and development: will peripheral countries ever get rich?*", *Journal of Development Economics*, 72-2, 515-541.

RODRIGUEZ, A and FRATESI, U, (2004): "*Between Development and social Policies; the impact of European structural funds in objective 1 regions*", Regional *Studies*; 88, 97-113

SÁNCHEZ, R., URBANO, A and ORTÍ, A. (1995): "*Wage premium of the industrial sector of the Spanish economy: empirical evidence", Review of labour economics and industrial relations*, 9, 253-74

SANCHEZ, R. and TOHARIA, L. (2000): "*Temporary workers and productivity: the case of Spain", Applied Economics*, 32, 583-91

SÁNCHEZ, R. and DÍAZ, M. A. (2013): **"***Are large innovative firms more efficient?*", *Theoretical and Practical Research in Economic Fields*, 4- 1, 89-97

SOLÉ, T. T., & RÍOS, M. S. (2008): "*La convergencia socio-económica de las comunidades autónomas españolas en la década de los noventa*", *Estudios de Economía Aplicada*, 26-3, 1-28

STENVENSON R., (1980): "*Likelihood Functions for Generalized Stochastic Frontier Functions", Journal of Econometrics,* 13, 57-66

WADHWANI, S. and WALL, M. (1991): "*A direct test of the efficiency wage model using UK microdata", Oxford Economics Papers,* 43, 529-48

APPENDIX: DESCRIPTION OF VARIABLES.

Variables used for the cluster technique:
Labour cost 1: Firms with a labour cost per worker below 22.366, 35
Euros per year
Labour cost 2: Firms with a labour cost per worker from 22.366,35 to 29.683,19 Euros per year.
Labour cost 3: Firms with a labour cost per worker from 29.683,2 to 38.888,25 Euros per year.
Labour cost 4: Firms with a labour cost per worker above 38.888,26 Euros per year.
LPAF1: Firms with less than 17 workers with permanent contracts.
LPAF2: Firms that have a number of permanent workers between 18 and 45.
LPAF3: Firms that have a number of permanent workers between 46 and 198.
LPAF4: Firms with more than 199 permanent workers.
R&D 1: Firms with any investment in R&D.
R&D 2: Firms with an investment in R&D below 81.651,5 Euros per year.
R&D 3: Firms with an investment in R&D between 81.651,51 and 296.915 Euros per year.
R&D 4: Firms with an investment in R&D between 296.915,1 and 1.008.261,5 Euros per year.
R&D 5: Firms with an investment in R&D of more than 1.008.261.5 Euros per year.
VExp1: Firms without exportations.
VExp2: Firms with export per worker less than 6555,52 Euros per year
VExp3: Firms with a value of export per worker between 6555.53 and 54428.58 Euros per year
VExp4: Firms with a value of export per worker above 54428.59 Euros per year

Variables of Stochastic Frontier estimations:
VA: The value added in real terms. This is a dependent variable.
CAPITAL STOCK (K): Inventory value of fixed assets excluding grounds and buildings; in logs.
L: Total employment by firm in logs.
T: This is the time trend.
Innovation of Product (INPR): This is a dummy variable that takes a value of 1 if the firm has obtained a product innovation, otherwise 0.
Innovation of Process (INP): This is a dummy variable that takes a value of 1 if the firm has obtained a process innovation, otherwise 0.
Sector classification: There are seven dummy variables that take a value of one when the firm belongs to the corresponding sector of activity; otherwise, this value is zero.
SEC1: Meat and manufacturing of meat; food industry and tobacco drinks; textiles, clothing and shoes; leather, shoes and derivatives.
SEC2: Wood and derivatives, paper and derivatives.
SEC3: Chemical products; cork and plastic; non-metallic mineral products.
SEC4: Basic metal products; manufactured metal products; industrial equipment.
SEC5: Office machinery and others; electrical materials.
SEC6: Cars and engines; other material transport.
SEC7: Other manufactured products. Category of Reference
Determinants of inefficiency:
Recession: This is a dummy variable that takes value 1 for 2008-2009 and zero otherwise.
Export per workers: The value of export over the total number of workers.
Firms 'expenditure in R&D per worker: Total spending on R&D of the firm deflated by the consumer price index and divided by the number of workers in the firm
Proportion of permanent workers: The number of workers with a permanent contract over the total amount of workers in the firm.
SIZE: There are five dummy variables that take a value of one when the firm belongs to the corresponding interval of workers, otherwise 0:

- **SIZE 1**: Firms with 10 to 50 workers.
- **SIZE 2**: from 51 to 100.
- **SIZE 3**: from 101 to 200.
- **SIZE 4**: from 201 to 500.
- **SIZE 5**: Firms with more than 500 workers