Chapter 20
Are We Spending Our Scarce R&D Resources Adequately? Analyzing the Efficiency of EU's Regional Innovation Systems

Thomas Baumert and Cristián Gutiérrez

Abstract The purpose of this article is to measure the efficiency of regional innovation systems (RIS) of the European Union (EU-14) between 2000 and 2010 based on the Data Envelopment Analysis. To measure the efficiency, we used 29 input variables synthesized in 5 factors or composite indicators. The output is reflected by patents and scientific publications. The results obtained highlight that only a few European RIS are situated on or near the efficiency frontier and most regions present very low efficiency levels. We detected a broad dispersion in terms of efficiency, although the tendency over time is a reduction of the dispersion reflecting a process of convergence. Moreover, the results reveal that an important possible cause of the inefficiencies is a problem of scale rather than technical inefficiency.

Keywords Efficiency · Regional innovation systems · DEA · Scale efficiency

JEL O32 · R58 · C61

20.1 Introduction

The economics of innovation, driven mainly—though not exclusively—by the so-called 'evolutionary approach,' has made an important effort in analyzing the allocation of resources towards the generation of scientific and technological knowledge. In doing so, it has stressed the importance not only of the agents intervening in the innovation process, but also the interactions between them, the institutional framework in which they interact and the policies aimed to favor them. In conjunction with growth economics, it has also been possible to solidly establish that the innovations derived

---

T. Baumert (✉)
Economics and History, ESIC Business & Marketing Madrid, Madrid, Spain
e-mail: thomasbaumert@esic.edu

C. Gutiérrez
Economics, Depto. Tecnologías de Gestión, Facultad Tecnológica, Univ. de Santiago de Chile, Santiago, Chile

N. Tsounis and A. Vlachvei (eds.), *Advances in Cross-Section Data Methods in Applied Economic Research*, Springer Proceedings in Business and Economics,
https://doi.org/10.1007/978-3-030-38253-7_20

from new knowledge constitute the central fundament of economic development.[1] A result of this paradigm is the widespread social belief regarding the convenience of government support towards science and technology (S&T); and, among politicians, the need for maintaining a broad set of economic and political instruments aimed at assisting R&D activities without any apparent limit in expenditure.

Accordingly, only seldom have economists of innovation or the makers of S&T policies analyzed in depth the possible limits of resource-assignment in the generation of knowledge. In general, it has been assumed that any level of R&D expenditure is pertinent insofar as it will always have a positive effect on economic development.

This notwithstanding, the matter of efficiency used to be a central question in the discussions of economists regarding innovation. Schumpeter (1942) already referred to it when emphasizing the role carried out by innovation in the long-run economic expansion, allowing the multiplication of a level of production given a limited volume of resources. At the same time, the neoclassical authors—who might be considered pioneers in the economics of innovation—also highlighted efficiency-related aspects.

Our purpose in this article is to embed the efficiency analysis into the evolutionist theory of the economics of innovation and the systemic approach. In fact, the efficiency is one of the driving forces behind the evolutionary path of competitiveness because those firms, regions or countries that are able to adapt themselves to the new circumstances will be more successful and efficient and therefore increase their possibilities of surviving during the competitive struggle between enterprises or regions. More specifically, our work aims to measure the efficiency level reached by RIS of the European Union (EU-14) during the production of knowledge (patents and scientific publications). In order to approach the matter of efficiency of the regional innovation systems from an empirical standpoint, we combine two multivariate analysis techniques. The first one (Factor Analysis), is used to create the combined input variables that allow us to describe in a synthetic way the complexity of the regional innovation systems.[2] The second technique (Data Envelopment Analysis), is used to build the efficiency frontier and determine the relative position of each of the regional systems with reference to it, and also permits the study of the causes of their inefficiency. This last aspect allows us to draw some relevant conclusions and suggestions for the design of innovation policies.

The existing studies[3] offer only some description of the detected efficiency indexes and only one of them tries to explain to a certain extent the causes of the inefficiency. Also, none of them except one analyses the possible potential for improvement. On the other hand, none of the studies used such many input factors based on a systemic approach. In this study, we use 29 input variables while in the twelve identified studies only one to seven variables were used. In fact, for several of these studies,

[1] Including the concept of 'new combinations' (Schumpeter 1911, Chap. 2), which Schumpeter would later include in his description of the 'process of creative destruction' (Schumpeter 1942, Chap. 7).

[2] Following the conceptual approach proposed by Buesa et al. (2007, 2010).

[3] Fritsch (2004), Fritsch and Slavtchev (2007, 2010), Zabala-Iturriagagoitía et al. (2008), Broekel (2008), Broekel and Meder (2008), Chen et al. (2011), Bosco and Brugnoli (2011), Badiola-Sánchez and Coto-Millán (2012), Niu et al. (2013), Kaihua and Mingting (2014).

the efficiency index was just an intermediate result to analyze other aspects such as the level of competitiveness or commercial success. Moreover, some studies used very debatable output indicators as the GDP per capita or the growth of regional employment. Concluding this study implies a significant methodological advance and can be considered as novel within the international literature. However, the analysis of the efficiency is still a new field, our paper is just one step forwards and many future improvements are still required.

20.2 Measuring Efficiency: The Basic Concepts

Several authors have challenged the task of defining and measuring the efficiency of activities related to the production of goods and services. Among them, the most noted are Koopmans (1951) and Debreu (1951), although it was Farrell (1957) who, relying on the works of the former two authors, prepared the ground of modern efficiency measurement. Following those authors, a global efficiency coefficient will be calculated, which is the correct approach when the microeconomic instrument of reference is the production function. This coefficient will be decomposed by the *pure technical efficiency*: which refers to the optimal employment of the inputs related to the output production. In the case of an *output orientation*, it refers to the maximum output that can be obtained given a certain level of input. There will also be the *scale efficiency:* which indicates whether the decision-making unit (DMU) operates on an optimal scale or not.

The purpose of this study is to measure the global technical efficiency[4] which consists of estimating production frontiers in such a way that the most efficient regions are the references and shape the form to the frontier of efficiency. The leading regions will have a normalized efficiency coefficient or level of 100, while the non-efficient units will be calculated (or 'positioned') in relation to their distance in percentage with the most efficient ones. In other words, they receive a lower score reflecting their (in) efficiency (as a percentage) regarding the frontier.

We followed the non-parametric DEA approach to measure the technical efficiency of the innovative results of the European regions regarding the use of resources during the innovation process for a timespan reaching from 2000 to 2010 (cross-section analysis) by means of a DEA, as this method has certain advantages in analyzing the efficiency of RIS (Niu et al. 2013: 149). The results thus obtained will allow us to establish which regions make a more efficient use of their innovation-oriented resources and quantify the inefficiency of the other regions.

[4]Which is the most commonly used methodology; although a few authors use non-frontier methods such as the construction of productivity indexes and other related econometric models.

20.3 Dataset and Methodology

20.3.1 Variables and Dataset

Opting for a holistic approach (using composite indicators) and using the DEA method—in which no specific functional form of the production function of innovations is going to be specified—implies that the selection of variables acquires special relevance. The variables employed in the model are 29 input and 2 output indicators that reflect the most significant and/or available information referring to regional innovation systems (see Fig. 20.1), which have been successfully employed in previous studies (Buesa et al. 2007).

The use of patents as a proxy for innovation output has been extensively debated in the literature,[5] confronting their advantages and disadvantages, with the former always by far outweighing the latter. Besides, there is also a pragmatic reason, namely the assignment of patents to the region where the research, design or engineering activity has in fact taken place, thus overcoming the so-called 'headquarter effect' (Eurostat 2011, Chap. 2). However, this does not mean that we ignore the fact that an important part of the research output of the regional innovation systems, especially those of scientific research is left out of analysis including only patents. To solve this problem, we include the scientific publications[6] in our analysis taking advantage of fact that the Data Envelopment Analysis allows us to simultaneously estimate the efficiency with two or more outputs.

The data employed in the empirical part of the paper correspond to the IAIF-RIS(EU) database[7] that basically contains data obtained from Eurostat's REGIO. It should be noted, that for some cases and for specific years REGIO does not offer data. In those, we have taken the missing values from the corresponding National Statistics Offices. Only rarely, have missing values been estimated according to the common procedure. Consequently, the initial database consists of a panel with 60

---

[5] See, among others, Scherer (1965), Schmookler (1966), Pavitt (1985), Mansfield (1986), Griliches (1990), Trajtenberg (1990), Archibugi (1992), Schmoch (1999), European Commission (2001: 38), Smith (2005: 158–160), Rondé and Hussler (2005: 1156), Hu and Mathews (2005: 1470) and Li (2009: 345). Regarding the time lag between R&D-input and patent application, we consider it to be nearly contemporaneous, a decision that seems to fit the results recently obtained by Wang and Hagedoorn (2014).

[6] The literature also recognizes certain problems associated with the use of publications as output variables. On the one hand, there is the language bias, in the sense that most publications in the most prestigious scientific journals are published in the English language, generating a bias towards researchers whose native language is this. Another criticism is that many publications are written by multiple authors, often from different regions or countries, and it is almost impossible to distinguish the individual contribution to the publication. However, both problems lose strength at the regional level within a country, since language bias affects all regions equally, as does the problem of co-authorship. This last problem was treated using the complete counting method (see Winkler 2014). Therefore, we have chosen to use this variable in the study. Regarding the time lag between R&D-input and scientific publication, we consider it to be nearly contemporaneous too.

[7] Employed, among others, by Buesa et al. (2007, 2010).

variables by 1452 cases (132 regions for 11 years). After revising the database and applying the factor analysis and its statistical tests 31 variables were used to identify and characterize the regional innovation systems (see Fig. 20.1). Those variables refer to the region's economic and population size, human resources, its sophistication of demand (wealth), the R&D efforts (both in economic and personnel terms), the propensity to patent[8] and other aspects related to the economic environment. On

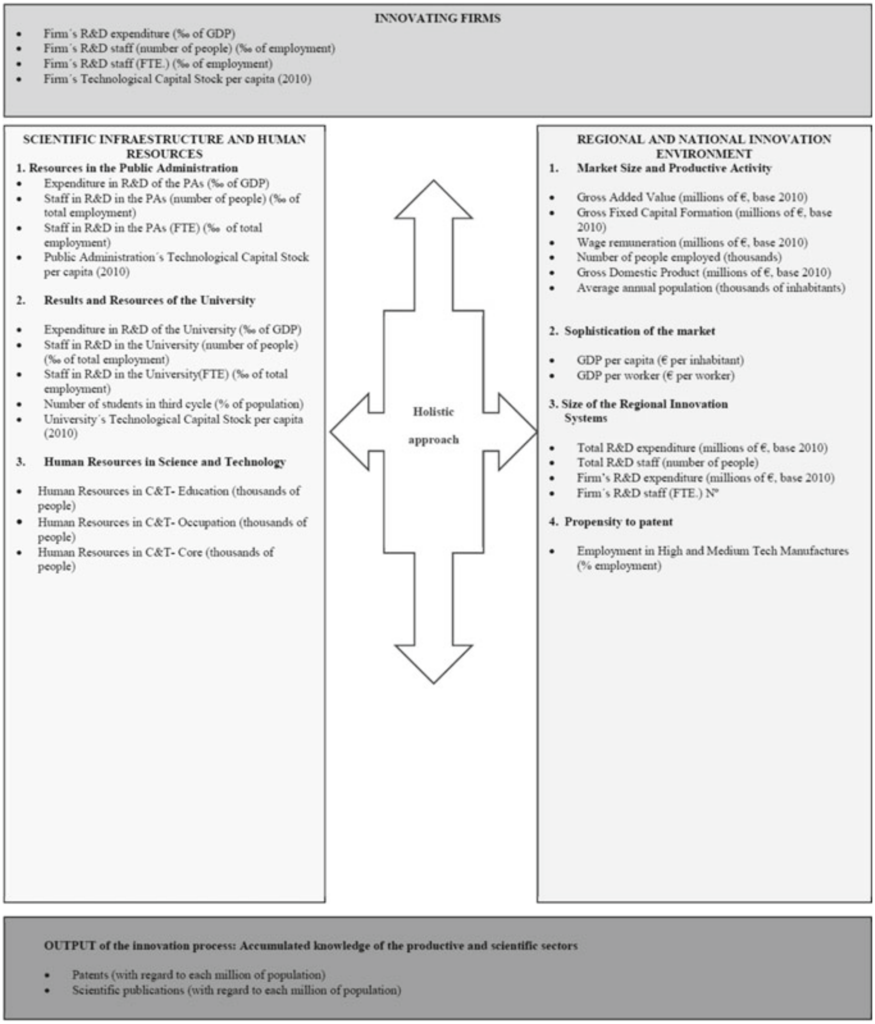


Fig. 20.1 Variables and indicators regarding regional innovation systems (own elaboration)

[8]Using Regional Employment in High-Medium Tech Manufactures (% of employment) as proxy.

the other side, regarding large several aspects considered as important by the ‘systemic approach’ no (statistical) data are available. In fact, no homogeneous publicly available data exist on aspects as the quality and quantity of technical infrastructures, R&D and innovation policies or institutional settings.

Another important limitation of the empirical research in the case of the regional innovation system is the limited availability of information that is equally measured for all the 132 European regions on the most appropriate level and moreover, those standardized indicators should be available for the whole period (11 years). Although our objective was to use a regional delimitation based on the region’s autonomy in the design and implementation of innovation policies, we had to work for some countries with other levels due to the absence of data. In fact, for some countries, no regional level data were available so we opted to use the national level data as in the case of Ireland, Denmark and Luxembourg. The NUTS level used for each country and the final used RIS (132) as DMU is presented in Fig. 20.2.

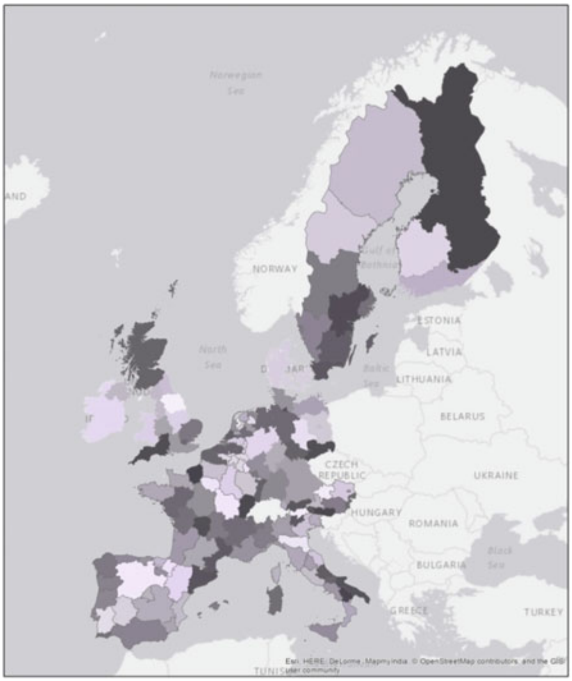


Fig. 20.2 Regional innovations systems in Europe (own elaboration)

20.3.2 Synthesizing the Elements of the Regional Innovation Systems in the European Union: Factor Analysis

As discussed before, regional innovation systems are complex realities composed by multiple actors whose institutional configuration can be very diverse and its interactions. This implies that, for a correct representation of these systems, a great variety of indicators is required.[9] As has been explained above, in the present case we are working with 29 input variables and two output variables (patents and publications) from the initial dataset with 60 variables. However, the set of input variables can be summarized in a smaller number of 'abstract', synthetic variables—called factors—which can be clearly identified with the elements that compose the RIS, while retaining most of the information (in terms of variance) contained in the original dataset. The use of the statistical technique of factor analysis turns out to be very appropriate for the study of such multidimensional economic realities as innovation systems,[10] as it does not only group together related variables taking into account their interaction but also considers at the same time the correlations with all the rest of the variables outside the specific factor. More specifically, the factorial scores are calculated using not only the correlations among the variables within each factor but also the correlations with all other variables/factors of the model. In this way, it implicitly measures somehow the interaction or interdependency between the subsystems. The relevance of these interactions in measuring an innovation system is pointed out, among others, by Niosi (2002).

The validation or quality of the factor analysis is based on the statistical tests and the inherent logic of the discovered factors. The different tests to confirm the quality of our factor analysis are all satisfying.[11] Moreover, the communalities (correlation of each variable regarding the set of the other variables making up this factor) of the variables are relatively high, most of them well above 0.75, which guarantees the reliability of the composite indicators, and indicates the high degree of preservation of their variance. The five factors obtained retain over 87% of the original variance, that is, there is scarcely a 13% loss of information originally contained in the variables. The second and maybe the most important criterion to judge the outcome of a factor analysis is that the extracted factors are consistent and interpretable in accordance with the theoretical or conceptual framework of the study, in our case, the regional innovation system. In other words, factor analysis is only useful if the results can be

[9] It should be noted, that using factors instead of a set of individual variables makes possible matters of collinearity irrelevant.

[10] Additionally, the working with factors as explanatory variables of an econometric model has a series of statistical advantages—such as the a priori maximization of orthogonally between factors when rotating them by the Varimax method, thus minimizing the possible collinearity between them—which have been detailed in previous studies (Buesa et al. 2007, 2010) and econometric manuals (Hair et al. 1999: 152).

[11] Thus, the Kaiser–Meyer–Olkin (KMO) test, which is based on the study of the partial correlation coefficients, gives a value a 0.8, within the upper limit of the recommended value of 0.6–0.8. Also, the Barlett Sphericity test, which tests for the null hypothesis that the correlation matrix is the identity matrix, is rejected at the 99% level.

interpreted correctly from a theoretical point of view. And such interpretation is only possible if simultaneously: (1) the included variables belong to the same component or subsystem of the overall regional innovation system; (2) the variables belonging to a certain subsystem are in only one factor; and (3) if each factor can be labeled with a 'name' which, without any reservation, clearly expresses its whole content.

On the other hand, comparing Fig. 20.1 and Table 20.1, it can be highlighted that the classification of the variables in five factors (based on the real correlations between the variables) doesn't differ from the initially a priori classification of Fig. 20.1 in which the variables were grouped by the theoretical arguments of the innovation system approach. We consider that the appropriateness of the model with five factors is supported by several facts, among others that our five factors (see Table 20.1) accomplish the three requirements mentioned above.

The composition of the factors and their interpretation ('names') and respective retained variance are given in Fig. 20.3. It should be observed that the resulting factors essentially coincide with the main determinants of a regional innovation system obtained by Buesa et al. (2010). Summing up, we can conclude that the factor analysis results in a coherent reduction of the original dataset, which fulfill all statistical and conceptual criteria and conveniently synthesize the main elements that constitute the European Union's regional innovation system. Thus, they seem suitable for further employment as independent variables in the study of the innovative efficiency of the European regions.

20.3.3 Data Envelopment Analysis

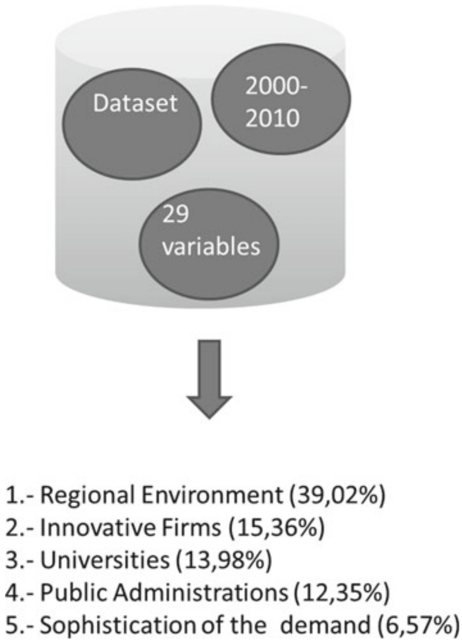
Once the five factors, which appropriately reflect the main elements of regional innovation systems, have been conveniently transformed,[12] we relate them with the output variable of the regional innovation systems (number of patents and publications—considering only technical fields—per capita) to analyze the efficiency in each year of period 2000–2010. Considering that—like in most real situations of economic analysis—we do not precisely know what the knowledge production function in terms of efficiency looks like. The Data Envelopment Analysis (DEA) allows the efficiency frontier to be drawn without the need for assuming a specific functional form under quite unrestrictive assumptions. This frontier is approached or estimated using the available data in which the frontier is drawn by the regions with the highest output level given a certain level of input (output orientation). It is not possible to draw the 'real' frontier which, in fact, is unknown but this approximation allows us to obtain a valid and quite useful measure of the *relative* level of efficiency of each singular case. For that reason, the DEA is the fundamental technique within the non-parametric approaches and has been much employed in microeconomic studies aimed to control and evaluate diverse units and actions of the public and private sector.

[12] As factor scores are calculated to follow a $N(0; 1)$ distribution, we use a linear transformation converting them into $N(4; 1)$ distributions in order to avoid any negative values in the input variables.

Table 20.1 Matrix of rotated components (own elaboration)

	Component				
	1	2	3	4	5
Wages (millions €2010)	0.977				
GAV (millions €2010)	0.976				
GDP (millions €2010)	0.975				
Number of people employed (thousand)	0.975				
Human resources in C&T—occupation (thousand)	0.969				
Annual average population (thousand)	0.964				
Human resources in C&T—core (thousand)	0.962				
Human resources in C&T—education (thousands of people)	0.950				
Gross fixed capital formation (millions €2010)	0.945				
Total R&D staff N°	0.900				
Total expenditure R&D (millions €2010)	0.860				
Firms R&D staff Nº	0.851				
Firms R&D expenditures (millions €2010)	0.818				
Firms R&D staff (HC) ‰ employment		0.881			
Firms R&D expenditures (‰ GDP)		0.877			
Firms R&D staff (HC) ‰ employment		0.861			
Stock of technological capital firms per capita (€2010)		0.852			
Regional employment hi-medium tech manufactures (% of employment)		0.587			
Universities R&D staff (HC) ‰ employment			0.909		
Universities R&D staff (FTE) ‰ employment			0.893		
Universities R&D expenditures (‰ GDP)			0.860		
Regional 3rd cycle students (% population)			0.833		
Stock of technological capital universities per capita (€2010)			0.829		
Public administration R&D staff (FTE) ‰ employment				0.944	
Public administration R&D staff (HC) ‰ employment				0.924	
Public administration R&D expenditures (‰ GDP)				0.921	
Stock of technological capital Public Administration per capita (€2010)				0.901	
GDP per worker (€2010)					0.799
GDP per capita (€2010)					0.793

Fig. 20.3 The factorial model (own elaboration)



There are two different models that can be implemented in the application of the technique: the model originally proposed by Charnes et al. (1978) (CCR model), which assumes *constant returns of scale* in the production function; and the modified version of this model proposed by Banker et al. (1984) (BCC model), that includes the possibility to consider *the efficiency of scale*. The model employed in the present work is the CCR one, as our aim is to make a comparative study among *all* the regions that compose the European Union and not only among those which present innovation systems with similar scale. However, we have also employed the BCC model as an instrument to calculate a measure of the efficiency of scale[13]: the coefficient between the CCR and the BCC model (multiplied by hundred), offers an index of the scale efficiency which indicates if a region is operating—or not—on its optimum scale. Accordingly, the inefficiencies of scale would be the result either of a region already operating on the stretch of the production function with decreasing returns to scale; or because it is still situated in the section of increasing returns of scale.[14]

The formulation of the DEA is based on a mathematical program that for each DMU—that is, for each RIS—calculates, from a perspective of input-reduction or from output-increase an index of pure technical efficiency. In the present paper, we have opted for an input orientation where the indexes reflect the reduction of the inputs that would be necessary for a region to become efficient.[15] The DEA also

[13]Comparing the efficiency level between those groups of regions with a similar input level.

[14]Or, in other words, because it is not situated on the section of constant returns of scale.

[15]The DEA analysis has been calculated using rDEA package for *R*.

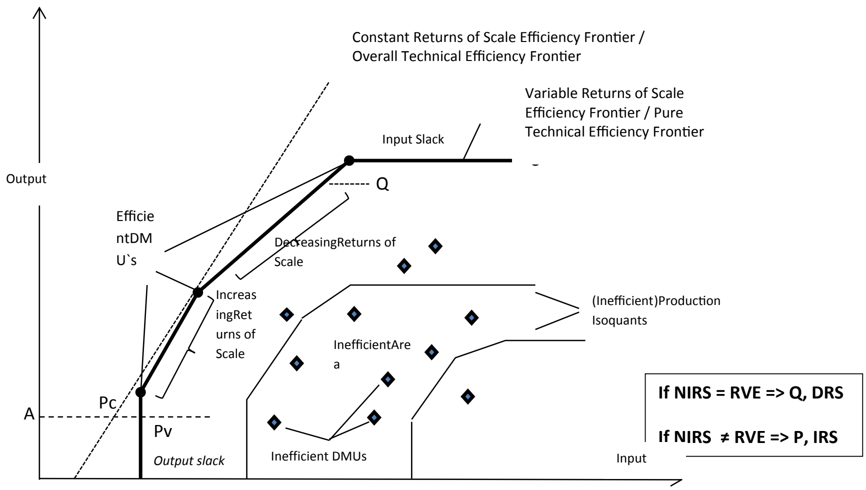


Fig. 20.4 Constant and variable returns of scale (Santin 2009)

allows other relevant information to be obtained: the volumes of input that a region could save by reaching a same output level (or the volume of output that a region could generate additionally, given a certain input) would it operate efficiently.[16]

These concepts are presented in Fig. 20.4, which illustrates different efficiency frontiers (isoquants) that can be estimated under constant and variables returns of scale, following the DEA-CCR model. This graph shows that the efficiency of a production unit *P*, under an input-orientation and constant returns to scale, is given by $\mathrm{ET}^{I}_{\mathrm{CRS}} = \mathrm{APv/AP}$. Taking into consideration these two measures, the scale efficiency (SE) is equal to $\mathrm{SE}^{\mathrm{I}} = \mathrm{APc/APv}$. When calculating some scale-inefficient units like P or *Q* are situated in the section of increasing returns to scale (*P*) or decreasing returns to scale (*Q*), an additional mathematical program must be computed if the scale returns are not increasing (NIRS). Thus, when NIRS = RVE, there are decreasing returns to scale; and when NIRS ≠ RVE, the returns to scale are considered constant (Santin 2009).

Summing it up, the aim of the DEA is to draw an evolving (hyper) plane that includes the efficient regions (and its linear combinations), that situates below it all inefficient units. As the envelopment plane represents the efficiency frontier, the distance of each region regarding this envelopment plane gives a value of its relative[17]

[16]The efficiency indexes present a measurement of radial efficiency, while these additional efficiencies, denominated slacks, provide a measure of the non-radial efficiency.

[17]It should be borne in mind, that DEA will measure a DMU's performance regarding its peers but—as it only very slowly converges to 'absolute' efficiency—not regarding the 'theoretical maximum' (Bhat et al. 2001: 32). However, we may assume that, given that some of the world's most innovation-efficient regions might be among the European ones, the 'maximum efficiency' might not be too far away from the one calculated in the present paper.

(in)efficiency that will be of one (100) if the region is situated on the border and less than one (100) if it is situated below it.

20.4 Results

Before developing the traditional DEA, we applied the super efficiency technique (Simar 2003; Banker and Chang 2006) with the objective to detect outliers in our dataset. The basic idea of this technique is that in the linear mathematical program used by DEA, each DMU is excluded from its own optimization, thus allowing some efficiency scores to reach values greater than one (100). The results of super-efficiency detect four regions that show super efficiency ratings in all the years of the series: Baden-Württemberg in Germany, Etelä-Suomi in Finland, Groningen in the Netherlands and Östra Mellansverige in Sweden. However, the only region with super efficiency ratings that would recommend its exclusion from the series is Noord-Brabant in the Netherlands, which in 2001, 2002, and 2003 obtained super efficiency scores higher than 2. Despite this, it was decided not to exclude this region from this dataset since more than an outlier is a region of high industrial development with a strong propensity to patent where one of the largest technology companies in the world as Philips is based and is entirely appropriate to consider it a benchmark to European level.

20.4.1 The Main Results in Terms of Efficiency Scores

In Tables 20.2, 20.3 and 20.4 we reflect the results for the efficiency scores for the years 2000, 2005 and 2010 using, as explained, the scientific publications and patents like outputs.

Calculating the efficiency using the two outputs separately (last two columns in Tables 20.2, 20.3 and 20.4 for the year 2010) is possible to divide the leading regions into three distinct groups. The leading technological regions (basically driven by patents) headed by Baden-Württemberg, Etelä-Suomi, Noord-Brabant, and Voralberg; secondly, the leading scientific regions (driven by publications) such as Groningen, Östra-Mellansverige, Övre-Norrland and Wien. A third group is formed by those regions that are jointly efficient (as well in patent and publications) such as Sydsverige, Stockholm and Bayern. One thing that should be emphasized is that the leading regions in some field are pushed to greater overall efficiency when considering the complementary field. For example, Baden-Württemberg and Etelä-Suomi, technology leaders in eight years have been global leaders (patents and publications) in 11 years.

Table 20.2 Efficiencies scores: Years 2000, 2005 and 2010 (0.6 till 1.0 in TE 2010) (own elaboration, using rDEA package from *R*)

		2000	2005	2010		
Regions	Countiy	TE	TE	TE	Tech E	Scient E
Baden-Württemberg	Germany	**1.000**	**1.000**	**1.000**	**1.000**	0.495
Eteä-Suomi (NUTS 2006)	Finland	**1.000**	**1.000**	**1.000**	0.998	0.646
Groningen	Nertherlands	**1.000**	**1.000**	**1.000**	0.161	**1.000**
Östra Mellansverige	Sweden	**1.000**	**1.000**	**1.000**	0.494	**1.000**
Övre Norrland	Sweden	**1.000**	0.994	**1.000**	0.310	**1.000**
Stockholm	Sweden	0.962	0.992	**1.000**	0.737	0.788
Vorarlberg	Austria	0.714	0.982	**1.000**	**1.000**	0.014
Sydsverige	Sweden	**1.000**	**1.000**	0.991	0.618	0.706
Noord-Brabant	Nertherlands	**1.000**	**1.000**	0.939	0.863	0.305
Bayern	Germany	0.912	0.873	0.913	0.786	0.473
Wien	Austria	**1.000**	0.912	0.869	0.215	0.869
Berlin	Germany	0.777	0.762	0.764	0.393	0.610
Nordrhein-Westfalen	Germany	0.751	0.816	0.749	0.592	0.319
Gelderland	Nertherlands	0.558	0.532	0.744	0.279	0.688
London	UK	0.813	0.726	0.705	0.228	0.569
Noord-Holland	Nertherlands	0.770	0.815	0.683	0.232	0.641
Denmark	Denmark	0.684	0.658	0.675	0.385	0.567
Utrecht	Nertherlands	**1.000**	**1.000**	0.672	0.333	0.567
Tirol	Austria	0.691	0.716	0.665	0.353	0.509
Vlaams Gewest	Belgium	0.648	0.724	0.656	0.283	0.582
Zuid-Holland	Nertherlands	0.681	0.643	0.649	0.300	0.583
Hessen	Germany	0.724	0.697	0.628	0.457	0.405
Rhône-Alpes	France	0.512	0.566	0.614	0.421	0.396
Länsi-Suomi	Finland	0.585	0.564	0.608	0.423	0.426
Scotland	UK	0.805	0.618	0.603	0.157	0.603
Västsverige	Sweden	0.825	0.838	0.603	0.432	0.380

The annual mean values are between 0.41 (2006) and 0.45 (2004) and the results are very heterogeneous with SD between minimum score 0.23 (2008) and 0.26 (2001), although the tendency over time is a reduction of the dispersion reflecting a process of convergence in terms of efficiency.

We used two ways of analyzing the level of dispersion among the regions of a specific country in 2010. The first is the calculation of distance of the values between

Table 20.3 Efficiencies scores: Years 2000, 2005 and 2010 (from average 0.41 till 0.6 in TE 2010) (own elaboration, using rDEA package from *R*)

		2000	2005	2010		
Regions	Country	TE	TE	TE	Tech E	Scient E
Île de France	France	0.588	0.586	0.595	0.367	0.462
Bremen	Germany	0.491	0.569	0.593	0.170	0.593
Rheinland-Pfalz	Germany	0.567	0.667	0.588	0.522	0.262
Pohjois- ja Itä-Suomi	Finland	0.628	0.553	0.588	0.269	0.576
Région de Bruxelles-Capitale/Brussels Ho	Belgium	0.806	0.718	0.584	0.240	0.516
Steiermark	Austria	0.650	0.556	0.578	0.366	0.444
Limburg (NL)	Nertherlands	0.460	0.544	0.578	0.254	0.503
Emilia-Romagna	Italy	0.650	0.693	0.567	0.273	0.481
South East (England)	UK	0.903	0.626	0.560	0.261	0.498
Centro (PT)	Portugal	0.230	0.374	0.532	0.024	0.532
Oberösterreich	Austria	0.344	0.382	0.532	0.482	0.152
Provincia Autonoma Trento	Italy	0.456	0.604	0.529	0.118	0.502
Comunidad Foral de Navarra	Spain	0.433	0.464	0.526	0.209	0.496
East of England	UK	0.934	0.643	0.526	0.209	0.487
Friuli-Venezia Giulia	Italy	0.568	0.572	0.516	0.290	0.425
Sachsen	Germany	0.411	0.458	0.512	0.293	0.474
Toscana	Italy	0.548	0.584	0.502	0.187	0.474
Alsace	France	0.454	0.495	0.490	0.309	0.360
North East (England)	UK	0.655	0.514	0.485	0.105	0.485
Salzburg	Austria	0.336	0.423	0.485	0.312	0.359
Niedersachsen	Germany	0.446	0.470	0.484	0.358	0.342
Saarland	Germany	0.481	0.516	0.474	0.302	0.355
Mecklenburg-Vorpommern	Germany	0.322	0.424	0.470	0.143	0.470
Overijs sel	Nertherlands	0.345	0.423	0.460	0.245	0.387
Thüringen	Germany	0.326	0.379	0.459	0.307	0.393
Hamburg	Germany	0.689	0.687	0.453	0.311	0.321
Wales	UK	0.577	0.427	0.450	0.083	0.450
Lombardia	Italy	0.472	0.498	0.450	0.245	0.369
Yorkshire and The Humber	UK	0.644	0.540	0.448	0.114	0.448
Norte	Portugal	0.154	0.265	0.447	0.022	0.447
Midi-Pyrénées	France	0.288	0.330	0.439	0.200	0.401
South West (England)	UK	0.588	0.451	0.437	0.196	0.389
Lazio	Italy	0.502	0.490	0.433	0.099	0.419

(continued)

Table 20.3 (continued)

		2000	2005	2010		
Regions	Country	TE	TE	TE	Tech E	Scient E
Schleswig-Holstein	Germany	0.516	0.428	0.431	0.283	0.294
Cataluña	Spain	0.379	0.422	0.422	0.128	0.404
Lisboa	Portugal	0.209	0.296	0.420	0.025	0.420
Aragón	Spain	0.350	0.419	0.416	0.107	0.401
Ireland	Ireland	0.353	0.390	0.412	0.128	0.383
East Midlands (England)	UK	0.643	0.466	0.410	0.153	0.395

the most efficient and least efficient region in a specific country. It should be noted that the greatest difference is observed for those countries with at least one very efficient region (Austria, Germany, Netherlands, Sweden and Finland). The second way to analyze the dispersion is to divide for each country its highest regional efficiency score with the lowest. In this case, the biggest differences in 2010 (Table 20.5) are observed in Italy (with the most efficient region being almost 30 times more efficient than the lowest), followed by Finland and Germany with a multiplier of 11 and 9, respectively.

About the efficiency distributions, applying normality test is easy to demonstrate that these are not normal, so using kernel density functions can reveal important features that would otherwise be hidden. This nonparametric approach requires the choice of a method to 'smooth' the data. In this paper, the kernel smoothing method has been chosen as this is one of the most commonly used in this type of work.[18] One of the advantages of kernel density functions is that they do not impose a priori functional forms on the distribution of data. We applied the kernel, and in particular estimating a Gaussian kernel with optimum bandwidth. Further, the differences among the European RIS and the dynamic perspective of the distribution of efficiency can be analyzed using stochastic kernel estimations that consider the probability of moving between any two levels in the range of values. A stochastic kernel is therefore conceptually equivalent to a transition matrix with the number of intervals tending to infinity (Quah 1993, 1996). The stochastic kernel can be approximated by estimating the density function of the distribution at a particular time $t + k$, conditioned by the values corresponding to a previous time t. For this, a nonparametric estimation of the joint density function of the distribution at times t and $t + k$ is carried out. Figure 20.5 shows the stochastic kernels estimated from the efficiency for time period of 11 years ($t = 2000$ and $t + k = 2010$). In this graph it is possible to appreciate a group of leading regions whose behavior is clearly different from the rest of the regions.

[18] As indicated by Tortosa-Ausina et al. (2005), authors such as Walter and Blum (1979) or Terrell and Scott (1992) note that virtually all non-parametric algorithms are asymptotically kernel methods (Suárez and de Jorge 2008).

Table 20.4 Efficiencies scores: Years 2000, 2005 and 2010 (below 0.41 in TE 2010) (own elaboration, using rDEA package from *R*)

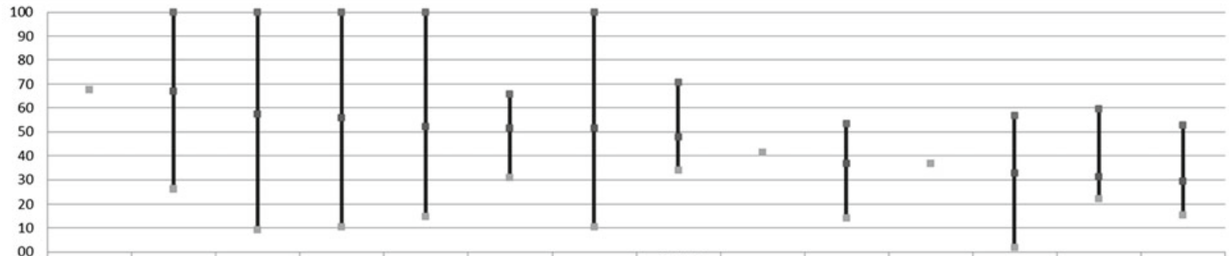
		2000	2005	2010		
Regions	Country	TE	TE	TE	Tech E	Scient E
Veneto	Italy	0.401	0.470	0.407	0.203	0.338
Umbria	Italy	0.485	0.482	0.404	0.074	0.404
Northern Ireland	UK	0.573	0.466	0.389	0.052	0.389
Picardie	France	0.371	0.347	0.384	0.135	0.335
North West (England)	UK	0.529	0.429	0.381	0.118	0.376
Liguria	Italy	0.489	0.461	0.370	0.150	0.327
Luxembourg	Luxembourg	0.362	0.472	0.366	0.215	0.191
Piemonte	Italy	0.332	0.379	0.358	0.192	0.286
Languedoc-Roussillon	France	0.355	0.343	0.354	0.130	0.335
Comunidad de Madrid	Spain	0.388	0.379	0.351	0.105	0.337
Provence-Alpes-Côte d'Azur	France	0.269	0.326	0.348	0.208	0.270
Galicia	Spain	0.278	0.337	0.347	0.028	0.347
Bretagne	France	0.273	0.347	0.342	0.221	0.241
West Midlands (England)	UK	0.480	0.365	0.339	0.144	0.313
Principado de Asturias	Spain	0.284	0.329	0.336	0.028	0.336
Aquitaine	France	0.328	0.313	0.334	0.132	0.300
Cantabria	Spain	0.362	0.377	0.333	0.054	0.333
Marche	Italy	0.355	0.390	0.332	0.168	0.288
Abruzzo	Italy	0.457	0.378	0.331	0.071	0.331
Franche-Comté	France	0.255	0.282	0.320	0.242	0.205
Comunidad Valenciana	Spain	0.298	0.332	0.309	0.053	0.309
Région Wallonne	Belgium	0.316	0.331	0.307	0.195	0.215
Brandenburg	Germany	0.179	0.281	0.303	0.220	0.160
Algarve	Portugal	0.206	0.416	0.291	0.012	0.291
Norra Mellansverige	Sweden	0.279	0.251	0.281	0.238	0.119
Pais Vasco	Spain	0.211	0.220	0.269	0.130	0.224
Auvergne	France	0.166	0.167	0.268	0.268	0.000
Småland med öarna	Sweden	0.186	0.259	0.260	0.214	0.108
Haute-Normandie	France	0.212	0.239	0.258	0.206	0.135
Región de Murcia	Spain	0.234	0.290	0.257	0.043	0.257
Sardegna	France	0.285	0.299	0.255	0.028	0.255
Pays de la Loire	France	0.172	0.211	0.248	0.156	0.190
Centre	France	0.218	0.243	0.247	0.164	0.167
Niederösterreich	Austria	0.194	0.280	0.247	0.243	0.017

(continued)

Table 20.4 (continued)

		2000	2005	2010		
Regions	Country	TE	TE	TE	Tech E	Scient E
Campania	Italy	0.243	0.262	0.245	0.029	0.245
Andalucia	Spain	0.210	0.243	0.240	0.028	0.240
Castilla y León	Spain	0.201	0.251	0.233	0.021	0.233
Sicilia	Italy	0.193	0.216	0.228	0.024	0.228
Molise	Italy	0.174	0.236	0.227	0.011	0.227
Bourgogne	France	0.246	0.232	0.226	0.119	0.186
Lorraine	France	0.283	0.242	0.225	0.093	0.210
Provincia Autonoma Bolzano-Bozen	Italy	0.148	0.195	0.224	0.210	0.052
La Rioja	Spain	0.199	0.232	0.221	0.051	0.207
Mellersta Norrland	Sweden	0.258	0.235	0.221	0.163	0.145
Limousin	France	0.125	0.191	0.220	0.132	0.169
Poitou-Charentes	France	0.205	0.202	0.213	0.088	0.189
Nord—Pas-de-Calais	France	0.188	0.185	0.209	0.086	0.190
Canarias (ES)	Spain	0.160	0.182	0.201	0.008	0.201
Puglia	Italy	0.176	0.217	0.196	0.033	0.196
Basse-Normandie	France	0.186	0.173	0.195	0.115	0.152
Calabria	Italy	0.160	0.194	0.193	0.013	0.193
Extremadura	Spain	0.142	0.219	0.180	0.009	0.180
Champagne-Ardenne	France	0.178	0.174	0.165	0.111	0.116
Castilla-la Mancha	Spain	0.086	0.146	0.153	0.021	0.150
Illes Balears	Spain	0.168	0.197	0.152	0.028	0.148
Kärnten	Austria	0.162	0.218	0.151	0.137	0.050
Burgenland	Austria	0.120	0.120	0.147	0.147	0.000
Alentejo	Portugal	0.045	0.096	0.142	0.013	0.142
Friesland (NL)	Nertherlands	0.087	0.117	0.129	0.129	0.000
Corse	France	0.154	0.162	0.126	0.008	0.126
Drenthe	Nertherlands	0.148	0.120	0.114	0.114	0.000
Flevoland	Nertherlands	0.205	0.125	0.111	0.111	0.000
Zeeland	Nertherlands	0.112	0.096	0.105	0.105	0.000
Sachsen-Anhalt	Germany	0.093	0.105	0.104	0.104	0.000
Åland	Finland	0.275	0.013	0.093	0.087	0.024
Valle d’Aosta/Vallée d’Aoste	Italy	0.131	0.171	0.091	0.091	0.000
Basilicata	Italy	0.007	0.025	0.019	0.018	0.002

Table 20.5 Dispersion of efficiencies, year 2010 (own elaboration)



	Denmark	Sweden	Finland	Germany	Austria	Belgium	Netherlands	United Kingdom	Ireland	Portugal	Luxembourg	Italy	France	Spain
EI Maximum	68	100	100	100	100	66	100	71	41	53	37	57	60	53
EI Average	68	67	57	56	52	52	52	48	41	37	37	33	31	29
EI Minimum	68	26	09	10	15	31	11	34	41	14	37	02	22	15

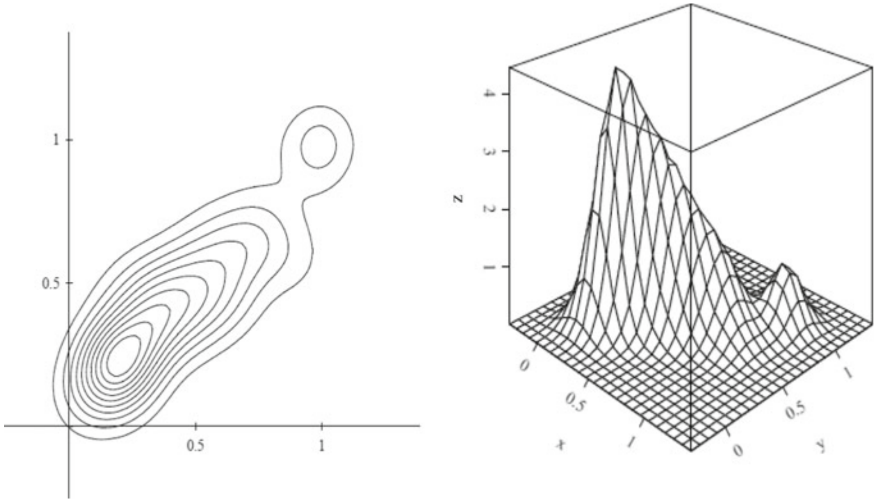


Fig. 20.5 Stochastic Kernel Efficiency for 2000–2010 period (*x* axis–*y* axis, respectively) in two (left) and three dimensions (right) (own elaboration using software Xtremes 4.1)

20.4.2 Changes Are the Efficiency Score Over Time; Convergence and Divergence

The inequality inefficiency was calculated for each of the models using the Gini index. This is, in fact, a concentration index is widely used in calculations of income inequality and takes values between 0 and 1, being 1 extreme inequality, only one is the most efficient, and 0 means total equality (all are equally efficient). The calculations were made for the years 2000, 2005 and 2010 and indicate that the Gini indexes are respectively 0.38, 0.34 and 0.33, showing a reduction of the inequality. Despite this process of convergence, the high level of heterogeneity (between and within countries) persists in the efficiency scores among European RIS. This heterogeneity is visualized in Table 20.5 for the year 2010.

Using the stochastic kernels again, we analyzed the evolution of the efficiencies distribution considering three years of our sample: 2000, 2005 and 2010. In Fig. 20.6, the stochastic kernels are estimated from the data of two years, 2000 and 2010, and then add the comparisons 2000–2005 and 2005–2010. This shows a convergence and a concentration process and displacement of the lines of level of the efficiency toward more values since the curves tend to concentrate reducing the group of leading regions as the comparisons are made between the greater years of the period. The in-depth analysis of the temporal dynamics and convergence goes beyond the scope of this paper.

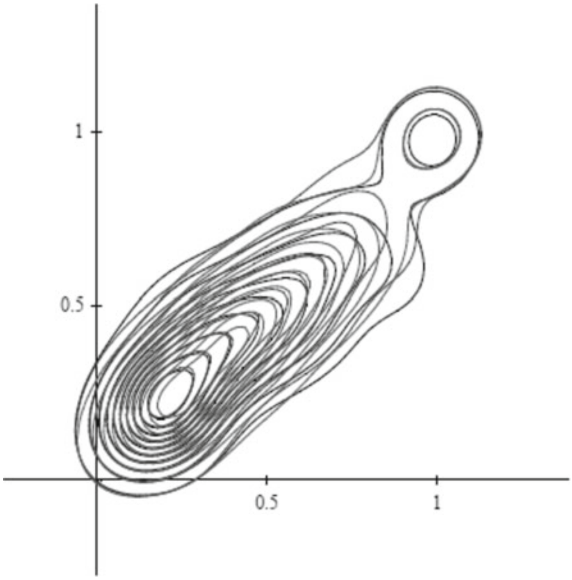


Fig. 20.6 Evaluation for periods of process of improvement of the efficiency in European RIS: 2000–2010 (black line), 2000–2005 (red line) and 2005–2010 (green line) (own elaboration using software Xtremes 4.1)

20.4.3 Scale Versus Technical Inefficiencies

Finally, the estimation of an index of scale efficiency for RIS as well as the test for returns to scale using bootstrap (Simar and Wilson 2002) reveals that much of the estimated inefficiencies in our model are caused by a dimension problem. Technical efficiency is high in many regions but its scale efficiencies[19] are very far from the frontier. This result highlights the fact that inefficiency maintains some relation with the need to reach a critical mass of economic and institutional resources of each region for the development of its innovation activities.

As mentioned above, the assumption of constant returns to scale, while useful for the determination of efficiency scores, is unrealistic. Therefore, in contrast to this hypothesis, we will try to confirm the greater relevance of scale problems in total inefficiency. We wish to test whether the technology set T from which our observations are sampled exhibits constant returns to scale (CRS). Formally, we

[19]In the case of the technical inefficiency it's about the technical capabilities of the regional agents to use their resources efficiently while in the case of scale advantages it is about the impact of the dimension of the regional innovation system on its efficiency. In fact, it analyzes what would happen with the efficiency of the regions with a similar input if they would have the same scale as the leading regions.

wish to test the hypothesis that the technology exhibits constant returns to scale (H_0) against the alternative (H_1), that it is variable returns to scale (VRS). If we reject H_0 then we can test if the technology set is decreasing returns to scale.

In accordance with Bogetoft and Otto (2011: 183): "If the hypothesis is true, then the efficiencies calculated from the VRS technology are the same as the efficiencies calculated from the CRS technology. If there is not CRS, then at least one of the efficiencies will be different; i.e., CRS efficiency will be smaller than VRS efficiency. One way to examine this is to see whether the scale efficiency,

$$\mathrm{SE}^k = \frac{E^k_{\mathrm{CRS}}}{E^k_{\mathrm{VRS}}}; \text{ with } k = 1, \ldots, K \tag{20.1}$$

is equal to 1 for all DMUs, meaning that the technology is CRS, or whether there is at least one firm where it is less than 1, meaning that the technology is VRS. For a given set of observations of K DMUs, we must therefore reject the hypothesis if at least one of the estimated SE has a value less than 1. However, as the connection between the technology set and the scale efficiencies is an uncertain or stochastic connection, we must reject the hypothesis if at least one of the estimated SE has a value significantly less than 1, i.e. if one of the estimated SE is less than a critical value."

We used the statistic defined by Bogetoft and Otto (2011),[20] but as we do not know the distribution of this statistic under H_0, therefore, we cannot calculate a critical value directly. One way to address this lack of distributional knowledge is to use a bootstrap method.

The results by each year are presented in the Table 20.6.

Applying the test with DRS as H_0 we rejected the null hypothesis at 95% of confidence and T (technology) would exhibit increasing returns to scale confirming the scale problems.

[20] In accordance with Bogetoft and Otto (2011: 183): "Instead of looking at the scale efficiencies individually, we could look at the test statistic $S^1 = \frac{1}{K}\sum_{k=1}^{K} \frac{E^k_{\mathrm{CRS}}}{E^k_{\mathrm{VRS}}}$; or the one that we are going to use in the following:

$S = \frac{\sum_{k=1}^{K} E^k_{\mathrm{CRS}}}{\sum_{k=1}^{K} E^k_{\mathrm{VRS}}}$; If the H_0 is true, then S will be close to 1, and if the alternative is true, then $S <$ 1. As $S \leq 1$ by construction, we will reject H_0 if S is significantly smaller than 1. We therefore seek a critical threshold for the statistic S; if it is smaller than this value, then we will reject the hypothesis. Thus, we seek a critical value c_α that will determine whether we reject H_0, the hypothesis of constant returns to scale, if $S < c_\alpha$ and Pr ($S < c_\alpha / H_0$) $= \alpha$ where α is the size of the test, typically 5% (α = 0.05). The size of the test, α is the probability of rejecting the hypothesis even though it is true (This is a type I error.)".

Table 20.6 Test of return to scale (own elaboration)

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
C_α	0.58	0.53	0.63	0.70	0.59	0.60	0.69	0.74	0.71	0.71	0.69
S	0.44	0.42	0.42	0.44	0.45	0.44	0.42	0.44	0.44	0.44	0.43
Test 5%	Reject H_0	Reject H_0	Reject H_0	Reject H_0	Reject H_0	Reject H_0	Reject H_0	Reject H_0	Reject H_0	Reject H_0	Reject H_0
Type *I* error	0.03	0.02	0.02	0.00	0.02	0.01	0.00	0.00	0.00	0.00	0.00

20.5 Conclusions

We start our study from the neoclassical and Schumpeterian framework of the economics of innovation using a holistic view in which all agents and organizations do interact and complement each other and add value to their mutual activities. Therefore, we decided to measure the efficiency of the RIS using a broad number of input variables (29) reflecting the broadest number of agents and factors as possible with the available statistical information. In this paper, we have explored the methods to measure the efficiency in which economic and institutional resources are used to obtain technologies useful to produce goods and services, as well as new scientific knowledge. The adopted approach is also linked to the evolutionary framework of this area of economic research, leaving the regional innovation systems in the center of the study, to calculate the level of innovative efficiency achieved by 132 regions from 14 countries belonging to the European Union. This is critically important as "the technical efficiency of a region largely reflects its ability to transform innovative investment into innovative output (and thereby transforming itself…) the key to this region to gain competitive advantage" (Chen and Guan 2012: 356).

The efficiency analysis carried out by the DEA technique allowed us to establish the efficient frontier by identifying those regions that maximize (minimize in an input orientation) the input/output relationship. In relation to this frontier, the DEA places the other regions by measuring their efficiency as a distance (in percentage) with respect to this border. The results obtained by this procedure allow us to point out, firstly, that only a few European regions are located on or very close to the efficiency frontier, with many regions obtaining systematically low efficiency scores. The dispersion of these levels of efficiency is very broad both within and between countries. Moreover, the differences in efficiency with which regions allocate their resources to innovation are a common feature of all multiregional nations, regardless of their level of income. In addition, RISs that are on or near the frontier belong to countries whose GDPs per capita are above the European average. On the other side, in all countries whose GDPs per capita are below the European average, the regions show efficiency levels below 20% of the frontier. Despite this, the tendency over time is a reduction of the dispersion reflecting a process of convergence in terms of efficiency.

The estimation of an index of scale efficiency for SRI as well as the test for returns to scale reveals that much of the estimated inefficiencies in our model are caused by a dimension problem. Technical efficiency is high in many regions, but its scale efficiencies are very far from the frontier. It points to the fact that inefficiency maintains some relation with the need to reach a critical mass of economic and institutional resources of each region for the development of its innovation activities. This last result should be considered by those responsible for designing and implementing innovation policies, aiming to economize resources employed with the highest possible returns. In other words, not any objective nor any actor is equally efficient developing R&D activities. According to this, there is no room for homogenous or 'coffee for all' policies; if not rather for 'tailor-made' innovation policies (see

Tödtling and Trippl 2005) implementing an improved personalized mix of science and technology instruments and R&D (see Chen and Guan 2012: 368), because at the end innovation activities differ strongly between regions in terms of their structural and institutional development.

Finally, the results support, to a certain extent, the organization of European R&D policy around two poles: one of promoting excellence through the Framework Program which would reward efficiency; And the other that of strengthening regional innovation systems through regional policy which would go in favor of critical mass and thus reduce the scale inefficiency.

References

Archibugi, D. (1992). Patenting as an indicator of technological innovation: A review. *Science and Public Policy, 19,* 357–368.

Badiola-Sánchez, A., & Coto-Millán, P. (2012). Determinants of technical efficiency and technological innovation of European regions in the period 2002–2006. *Journal of Business Management and Applied Economics*, *5*.

Banker, R., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science, 30*(9), 1078–1092.

Banker, R., & Chang, H. (2006). The Super-efficiency procedure for outlier identification, not for ranking efficient units. *European Journal of Operational Research, 175,* 1311–1320.

Bhat, R., Verma, B. B., & Reuben, E. (2001). *Hospital efficiency and data envelopment analysis*. Ahmedabad: Indian Institute of Management.

Bogetoft, P., & Otto, L. (2011). *Benchmarking with DEA*. SFA and R: Springer.

Bosco, M. G., Brugnoli, A. (2011). Regional efficiency, innovation and productivity. In Cached, RSA conference proceedings, Pécs.

Broekel, T. (2008). From average to the frontier: A nonparametric frontier approach to the analysis of externalities and regional innovation performance. In *Papers in evolutionary economic geography* (pp. 04). Utrecht University.

Broekel, T., & Meder, A. (2008). The bright and dark side of cooperation for regional innovation performance. In *Jena economic research papers* (pp. 053). Jena, Germany.

Buesa, M., Martínez Pellitero, M., Baumert, T., & Heijs, J. (2007). Novel applications of existing econometric instruments to analyse regional innovation systems: The Spanish case. In J. Suriñach, R. Moreno & E. Vayá, *Knowledge externalities, innovation clusters and regional development* (pp. 155–75). Cheltenham: Edward Elgar.

Buesa, M., Heijs, J., & Baumert, T. (2010). The determinants of regional innovation in Europe: A combined factorial and regression knowledge production function approach. *Research Policy, 39*(6), 722–735.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research, 2,* 429–444.

Chen, C. P., Hu, J. L., & Yang, C. H. (2011). An international comparison of R&D efficiency of multiple innovative outputs: The role of the national innovation system innovation: Management. *Policy and Practice, 13*(3), 341–360.

Chen, K., & Guan, J. (2012). Measuring the efficiency of China's regional innovation system: Application of network data envelopment analysis (DEA). *Regional Studies, 46*(3), 355–377.

Debreu, G. (1951). The coefficient of resource utilization. *Econometrica, 19*(3), 273–292.

European Commission. (2001). *Recherche et développement: statistiques annuelles*. Luxembourg.

Eurostat. (2011). *Eurostat yearbook 2011, Europe in figures*. Belgium.

Farrell, M. J. (1957). The measurement of efficiency productive. *Journal of the Royal Statistical Society*, *Series A (General)*, *120*(III), 253–66.
Fritsch, M. (2004). Cooperation and the efficiency of regional R&D activities. *Cambridge Journal of Economics, 28,* 829–846.
Fritsch, M., & Slavtchev, V. (2007). *How does agglomeration affect the efficiency of regional innovation systems?*. Kiel: Workshop Agglomeration and growth in knowledge based societies.
Fritsch, M., & Slavtchev, V. (2010). How does industry specialization affect the efficiency of regional innovation systems? *The Annals of Regional Science, 45*(1), 87–108.
Griliches, Z. (1990). Patent Statistics as economic Indicators: A Survey. *Journal of Economic Literature, 28,* 1661–1707.
Hair, J. F., Anderson, R., Tatham, R. & Black. (1999). *Análisis multivariante de datos'*. Madrid.
Hu, M. C., & Mathews, J. A. (2005). National innovative capacity in East Asia. *Research Policy, 34*(9), 1322–1349.
Kaihua, C., & Minting, K. (2014). Staged efficiency and its determinants of regional innovation systems: A two-step analytical procedure. *The Annals of Regional Science, 52*(2), 627–657.
Koopmans, T. C. (1951). An analysis of production as an efficient combination of activities. In T. C. Koopmans (Ed.), *Activity analysis of production and allocation* (pp. 33–97). New York: Jhon Wiley and Sons Inc.
Li, X. (2009). China's regional innovation capacity in transition: An empirical approach. *Research Policy, 38*(2), 338–357.
Mansfield, E. (1986). Patents and innovation: An empirical study. *Management Science, 32*(2), 173–181.
Niosi, Jorge. (2002). National systems of innovations are "x-efficient" (and x-effective): Why some are slow learners. *Research Policy, 31*(2), 291–302.
Niu, D., Wang, R., Zhang, X., & Song, W. (2013). Comparative studies on efficiency of China's regional innovation system on the basis of cooperation measurement model. *Journal of Emerging Trends in Economics and Management Sciences, 4*(2), 147–157.
Pavitt, K. (1985). Patent statistics as indicators of innovative activities: Possibilities and problems. *Scientometrics, 7*(1-2), 77–99.
Quah, D. T. (1993). Empirical cross-section dynamics in economic growth. *European Economic Review, 37,* 426–434.
Quah, D. T. (1996). Empirics for economics growth and convergence. *European Economic Review, 40,* 1353–1375.
Rondé, P. & Hussler, C. (2005). Innovation in regions: What does it really matter? *Research Policy*, *34*(8), 1150–72.
Santín, D. (2009). La Medición de la Eficiencia en el Sector Público. Técnicas Cuantitativas. http://www.sefin.fortaleza.ce.gov.br/apresentacoes/gerados/apresentacoes_madri/UD_EFICIENCIA_DANIEL_SANTIN.pdf.
Scherer, F. M. (1965). Firm size, market structure, opportunity and output patented inventions. *The American Economic Review, 55*(5), 1097–1125.
Schmoch, U. 1999. Eignen sich Patente als Innovationsindikatoren?. In: *Patentschutz und Innovation in Geschichte und Gegenwart*. In Frankfurt am Main et al: R. Boch.
Schmookler, J. (1966). *Invention and economic growth*. Cambridge Mass: Harvard University Press.
Schumpeter, J. A. (1911). *Theorie der wirtschaftlichen Entwicklung*. Leipzig: Duncker-Humboldt.
Schumpeter, J. A. (1942). *Capitalism, socialism and democracy*. Nueva York: Harper & Brothers.
Simar, L., & Wilson, P. (2002). Non-parametric tests of returns to scale. *European Journal of Operational Research, 139,* 115–132.
Simar, L. (2003). Detecting outliers in frontier models: A simple approach. *Journal of Productivity Analysis, 20,* 391–424.
Smith, K. (2005). Measuring innovation. In J. Fagerberg, D. Mowery & R. Nelson (Eds), *The oxford handbook of innovation* (pp. 148–77). New York: Oxford.
Suárez, C. & de Jorge, J. (2008). Efficiency convergence processes and effects of regulation in the nonspecialized retail sector in Spain. *The Annals of Reg Science.*

Terrell, G. R., & Scott, W. (1992). Variable kernel density estimation. *Annals of Statistics, 20*(3), 1236–1265.

Tödtling, F., & Trippl, M. (2005). One size fits All? Towards a differentiated regional innovation policy approach. *Research Policy, 34,* 1203–1219.

Tortosa-Ausina, E., Perez, F., Mas, T., & Goerlich, F. J. (2005). Growth and covergence profile in the Spanish provinces 1965–1997. *J Reg Sci, 45*(1), 147–182.

Trajtenberg, M. (1990). *Patents as indicators of Innovation.In Economic Analysis of Product Innovation?* MA: Cambridge.

Walter, G., & Blum, J. (1979). Probability density estimation using delta sequences. *Annals of Statistics, 7*(2), 328–340.

Wang, N., & Hagedoorn, J. (forthcoming—2014 online). The lag structure of the relationship between patenting and internal R&D revisited. *Research Policy* http://dx.doi.org/10.1016/j.respol.2014.03.010.

Winkler, W. E. (2014). Matching and record linkage. *WIREs Computational Statistics, 6*(5), 313–325.