Monica Mihaela MATEI, PhD
Email: matei.monicamihaela@gmail.com
Professor Liliana SPIRCU, PhD
Email: spircu2008@yahoo.com
Department of Economic Informatics and Cybernetics
The Bucharest Academy of Economic Studies

RANKING REGIONAL INNOVATION SYSTEMS ACCORDING TO THEIR TECHNICAL EFFICIENCY- A NONPARAMETRIC APPROACH

Abstract. *The new economic strategy of the European Union aims to improve EU` performance in innovation. An important step in achieving this goal is represented by the assessment of the innovation process at regional or national level. The purpose of this study is to measure and compare the performance of the Regional Innovation Systems using a nonparametric approach. Thus the efficiency of the decision making units represented by Regional Innovation Systems is estimated using a nonparametric frontier model: data envelopment analysis (DEA). Statistical inference for DEA estimators is based on bootstrap, a very well-known resampling method. The estimated efficiency scores obtained from DEA models help us identify best practice in the field of regional innovation.*

Keywords: *regional innovation systems, technical efficiency, DEA, bootstrap.*

JEL Classification P48, C14, C15, C67

1. Introduction

The new economic strategy launched by the European Commission for the coming decade in order to go out of the crisis is Europe 2020. One of the objectives associated to this strategy is to develop a smart growth by improving EU's performance in innovation. The reason behind this idea is based on the belief that innovation can be translated into new goods and services thus creating growth and jobs. We think that high performance in innovation can be achieved by the implementation of an assessment procedure dedicated to innovation systems. One concrete action that has been initiated in this regard is the development of a tool meant to help assessing the innovation performance in EU Member States. This tool is known as Innovation Union Scoreboard (IUS) and includes innovation indicators which capture the performance of the national innovation systems and trend analyses for the EU 27 Member States. The same tool was developed for the evaluation of regional innovation systems – RIS (Regional Innovation Scoreboard). Even if there are many different definitions of the innovation concept there are some features which are common to all of these: the systemic nature, the complexity and the importance for the economic development.

According to Buesa (2006) a system of innovation can be defined as ”the set of institutional and business organizations which, within a specific geographical area, interact with the aim of allotting resources to performing activities geared to generating and spreading knowledge which supports the innovations which are the basis of economic development”. This definition will help us understand which are the relevant inputs and outputs of the innovation process. We need to select from RIS database variables reflecting the inputs used to produce innovation outputs in order to estimate through nonparametric methods the efficiency of the regional innovation systems. In this paper Data Envelopment Analysis (DEA) models are used in order to rank regions according to the efficiency of their innovation systems. This analysis is useful for: development of effective policies, better use of resources in the field of innovation, designing regional innovation strategies. By this study we want to contribute to the process started by OECD and European Commission of assessing the regions innovation systems, finding good practices and benchmarks.

2. Literature review

An important step in the development of better innovation policies is the assessment of the innovation system performance. There are studies developed in this field based both on nonparametric and on parametric models where authors investigated the efficiency, the quality of the innovation systems and implicitly the dependence between variables describing this process.

In the parametric approaches, authors assume a functional form for the production function and they estimate its parameters showing which innovation inputs influence innovation outputs. When measuring the performance of a system we need to define the variables that describe its inputs and outputs. In order to estimate the production function Fritsch (2002) used as output the number of disclosed patent application and the input variable was given by private sector R&D employees .Its results show a strong impact of the number of private sector R&D employees on the number of patents. Another important result of this study is related to input/ output selection. They found that a time lag of three years is optimum when relating inputs to outputs.

Another study developed for the Spanish regions identified which are the main factors that have an impact on the regional innovation capacity measured by the number of patents (Buesa et al., 2006). They use information on patents as a tangible output indicator of the innovation process and they used regression analysis to find the factors that influence innovation capacity. Those factors are represented by variables related to science and technology. The variables they used describe: firms and their relationship with the regional innovation systems, support infrastructure for innovation, and the regional and national environment for innovation. (Buesa et al., 2006).

The determinants of the regional innovation in Europe were studied through a knowledge production function (Buesa et al., 2010) and the analysis reflected five important aspects of the innovation systems: the National environment, the Regional environment, Innovating firms, Universities and the R&D done by Public Administration. The results show that all factors have a statistically significant effect on the production of knowledge (patents), although they present very different impacts (Buesa et al., 2010).

There are also studies developed in this field that use nonparametric methods for the estimation of innovation systems efficiency. Most of them deal with regional innovation systems (Kutvonen, 2007). In his study, Antero Kutvonen used a DEA (Data Envelopment Analysis) model in order to identify best practice cases of regional innovation policies and he draw the conclusion that DEA provides means for benchmarking regional policies in different areas.

There is also a study that measures the performance of national innovation systems in Europe and Asia which applies Data Envelopment Analysis models. Their results show that Asian countries are generally better performers than European countries (Ta-Wei Pan et all, 2010). They proved that DEA is useful for estimating efficiency of the innovation systems.

3. Methodology

The analysis presented in this study is based on a nonparametric approach. The efficiency estimates are obtained using a Data Envelopment Analysis model. In order to rank regions according to their regional innovation systems efficiency we estimated an efficiency score for each region included in the analysis. In efficiency theory this means that each RIS is treated as a production unit or Decision making unit which transforms a number of *p* inputs into a number of *q* outputs. In the space the production set is defined as (Daraio et al., 2010):

,

where is feasible if it is physically possible to produce the output quantities *y* when the available quantities of inputs are given by *x*.

In the nonparametric approach context, the efficiency scores represent a distance from each unit to the production frontier which is estimated by a nonparametric method. This means no assumption is made regarding the functional form of the production function and only information regarding the quantities of inputs used and the quantities of outputs produced by each unit are needed (Coelli et al., 2005). The frontier of the production set is usually denoted by $\partial\psi$ and mathematically can be written as (Daraio et al., 2010):

$$\partial\psi(x,y) = \{(x,y) \mid (x,y) \in \psi(x,y), (\theta x, \lambda y) \notin \psi(x,y), \theta < 1, \lambda > 1\}$$

An important feature of the frontier is the return to scale. Depending on its shape, the frontier can exhibit: constant return to scale (CRS) or variable return to scale (VRS). As we will see below, this is an important aspect giving the restrictions of the linear problems that should be solved in order to compute the efficiency scores (Charnes et all., 1994).

If a decision making unit using the inputs given by vector *x* to produce the amount of outputs given by vector *y* lies on the frontier then it is rated as efficient otherwise the unit is inefficient and the level of inefficiency is measured through the distance from the point *(x,y)* to the frontier. This distance can be computed using Farrell`s (1957) radial measure given by the maximum proportionate increase of all outputs (in an output orientation case) or maximum proportionate reduction of all inputs (in an input orientation case) depending on the orientation chosen, with the condition of remaining in the production set. In this paper we investigated the decision making units using an output orientation approach. In this case the technical efficiency associated to the unit characterized by *(x,y)* is given by:

$$\lambda(x,y) = \sup\{\lambda \mid (x, \lambda y) \in \psi\}$$

According to the definition presented above the following is true:

$$\forall (x,y) \in \psi, \lambda(x,y) \geq 1.$$

Thus if the score is equal to 1 then the unit is efficient otherwise the larger the score the more inefficient the unit.

In order to find the values of the efficiency scores we first need to estimate the efficient frontier using the information from the sample:

$$\chi_n = \{(x_i, y_i) \mid i = 1,2..., n\}$$

representing the amounts of inputs and outputs associated to each decision making unit.

In this research we use a nonparametric technique known as Data envelopment analysis (DEA). The DEA estimator is built to measure the efficiency relative to the boundary of the production set (Charnes et all., 1994).

When assuming variable return to scale in the output orientation case, the efficiency estimator is computed by solving the following linear programming problem (Daraio et all., 2010):

$$\hat{\lambda}_{VRS}(x,y) = \max\{\lambda \mid \lambda y \leq \sum_{i=1}^{n} \gamma_i y_i ; x \geq \sum_{i=1}^{n} \gamma_i x_i ,$$
$$\sum_{i=1}^{n} \gamma_i = 1; \gamma_i \geq 0, i = 1,..., n\}$$

When assuming CRS, the problem becomes:

$$\hat{\lambda}_{CRS}(x,y) = \max\{\lambda \mid \lambda y \leq \sum_{i=1}^{n} \gamma_i y_i ; x \geq \sum_{i=1}^{n} \gamma_i x_i ,\ \gamma_i \geq 0, i = 1,..., n\}.$$

The main advantages of using envelopment models are given by the following aspects:

- Allow multiple input/output modelling;
- Do not require specifying a functional form of dependence between variables.

The major drawbacks are caused by the deterministic nature of these nonparametric models which creates problems in making statistical inference. One of the solutions found to eliminate this disadvantage is based on the application of the bootstrap techniques (Simar, Wilson 2000). Bootstrap is a resampling technique that can be used to estimate the statistical inference tools (standard deviation, bias, confidence intervals) when the distribution of estimators is unknown or when the sample size is reduced without introducing parametric assumptions.

In the context of DEA models the bootstrap technique has to be adapted to the particularities of these estimators and this is why a homogeneous bootstrap is used (Simar, Wilson, 1999). We only present here a simplified version of the algorithm. The bias, the confidence intervals for the DEA estimators are computed as follows:

1. The sample $\chi_n = \{(x_i, y_i), i = 1,...,n\}$ is used to estimate the efficiency scores $\hat{\lambda}_i(x_i, y_i)$, *i = 1,2..., n,* by solving *n* linear programming problems.

2. For each decision making unit, a bootstrap replica is computed using the homogeneous bootstrap algorithm. This yields to a generation based on a smooth estimate $\hat{f}(\cdot,\cdot)$ of the joint pdf on *(x,y)*. The algorithm proposed by Simar and Wilson (2000) to perform such a generation creates bootstrap sample containing units obtained after the projection of each observation (x_i, y_i) onto the estimated frontier is randomly projected away from the frontier. In order to project the initial point on the frontier, the estimated efficiency scores are used and for the second projection a bandwidth *h* and a variable randomly drawn from a standard normal distribution are employed (Simar, Wilson., 2000).

3. For the units in the bootstrap sample generated in stage 2 are computed the efficiency scores denoted by $\hat{\lambda}^*(x_i, y_i)$

By redoing the last two steps *B* times, for each data point *i=1…n* we will have *B* bootstrap estimations of the score . The final result consists in an empirical distribution of which is used to estimate the bias, the standard deviation and the confidence intervals of DEA estimators as follows (Daraio et al., 2010):

$-$

$$\hat{std}^2\left(\hat{\lambda}(x,y)\right) \approx \frac{1}{B}\sum_{b=1}^{B}\left[\hat{\lambda}_b^*(x,y) - \frac{1}{B}\sum_{b=1}^{B}\hat{\lambda}_b^*(x,y)\right]^2$$

$$\hat{\lambda}(x,y) - \hat{a}_{0.975} \leq \lambda(x,y) \leq \hat{\lambda}(x,y) - \hat{a}_{0.025}$$

where

$$Prob\left(\hat{a}_{0.025} \le \hat{\lambda}^{*}(x,y) - \hat{\lambda}(x,y) \le \hat{a}_{0.975}\right) = 0.95.$$

4. Empirical results

The last edition of the Pro Inno Europe report regarding regional innovation is that from 2009. The missing data problem is much more intense at regional level compared to national level (Matei, 2010) but the major advantage is given in this case by a considerable increase of the number of decision making units included in the sample. The investigation of the regional innovation systems brought to light the heterogeneity of the regions of the same country indicating the necessity of the implementation of some policies adapted to each region particularities.
In this paper we offer a quantitative tool dedicated to the assessment of the regional innovation performance across NUTS 2 regions. The first and may be the most important step in the implementation of an evaluation procedure is the selection of the input/output variables defining the production process. The variables included in this analysis are chosen form the data base RIS 2009 .

4.1 Data description

Due to the data missing problem we only investigated 116 decision making units represented by NUTS 2 regions from 13 countries. The distribution of regions by countries is represented in Table 1.

For the analysis developed in this paper, regional innovation systems are characterized by five input variables and two output variables described in Table 2. The input variables capture relevant aspects of the innovation process such as: human capital, research financing and number of patens as an indicator of intermediary results of innovation. The output variables available for many of the regions included in RIS 2009 reflect the share of persons employed in medium-high and high tech manufacturing and in knowledge intensive services in total workforce. This is why in this paper we measure the innovation efficiency using as outputs the economic effects of innovation on the labour market. Another argument supporting this approach comes from the fact that innovation is considered a key element of the Europe 2020 Strategy for creating new jobs. The philosophy behind is the following: the innovative ideas are a result of the innovation inputs (learning, research) but it is worthless unless these ideas are translated into new products and services which could create new jobs. Thus the final results of the innovation system designed in this paper make reference to the level of labour market modernisation and matching between supply and demand. There are many studies supporting our approach, but we mention here only two of them. Radosevic (2004) shows that a very well represented high tech sector represents an important factor for the development of a region and for its capacity of catching up. The second is a study at European level which proves that that employment in the knowledge based sectors has a positive impact on the regional development (Heidenreich, 2008).

Table 1. Distribution of regional innovation systems by countries

Country	Number of regions included in RIS 2009	Number of regions included in the investigation
Germany	29	22
Ireland	2	2
Greece	1	1
Spain	17	15
Italy	17	17
Hungary	6	6
Netherlands	12	12
Romania	8	0
Slovenia	2	0
Czech Republic	8	0
Poland	16	16
Portugal	5	5
Slovakia	4	4
Finland	4	4
Sweeden	8	8
Norway	7	4

As can be seen in Table 2, the first variables selected to define the regional innovation systems are reflecting the human capital which is one of the most important dimension of the innovation. The following arguments are underlying the selection of the variables building this first dimension of the innovation:

- highly skilled labour force represents a necessary condition for innovation and the number of tertiary education graduates represents a reliable indicator of this aspect;
- the knowledge economy functioning is based on lifelong learning.

Another important dimension of the innovation process is related to the financing of the research. The public R&D expenditures are considered one of the major factors of the economic development for a knowledge based economy and it is a good indicator of the future level of competitiveness and the welfare of the society. In order to emphasize the firms' efforts towards knowledge creation we introduced a variable that measures business R&D expenditures. The fifth input included in the system, EPO patents, represents an intermediary result of the innovation process. This variable is a proxy of the new innovative products development rate.

Table 2. Input/output variables

Variable	RIS 2009 Indicator	Year
I_1	Population with teriary education per 100 population aged 25-64	2004
I_2	participation in life long learning per 100 population aged 25-64	2004
I_3	Public R&D expenditures (% of GDP)	2004
I_4	Business R&D expenditures (% of GDP)	2004
I_5	EPO patents per million population	2004
O_1	Employment medium-high & high-tech manufacturing (% of total workforce)	2006
O_2	Employment knowledge-intensive services (% of total workforce)	2006

Therefore our objective is to measure the efficiency of 116 RIS by nonparametric techniques. In order to obtain a convergence rate close to the parametric estimations rate, considering we have a sample size of 116 units, the production process should be defined by 3 variables. Thus we first investigate the possibility of reducing the number of inputs or/and outputs in order to attenuate the negative effects of the dimensionality curse. The correlations coefficients show that input variables could be aggregated. We used the procedure described by Daraio (2010) which consists in computing an aggregated factor as a linear combination of the initial variables, following a method similar to principal component analysis.

In order to compute the aggregated factor *I*, an eigenvector associated to the largest eigenvalue (denoted a_1) of the matrix was used as follows:

$$I = 0.450 I_1^s + 0.448 I_2^s + 0.432 I_3^s + 0.444 I_4^s + 0.461 I_5^s,$$

where I_j^s represents the new variable obtained by dividing each input I_j by its mean, j=1,2..5 and X is the matrix of this scaled variables.

Table 3. Correlation coefficients among inputs variables

Correlation coefficient	I_1	I_2	I_3	I_4	I_5
I_1	1.000	0.479	0.424	0.570	0.495
I_2	0.479	1.000	0.442	0.608	0.598
I_3	0.424	0.442	1.000	0.501	0.385
I_4	0.570	0.608	0.501	1.000	0.850
I_5	0.495	0.598	0.385	0.850	1.000

We think that the aggregated factor summarizes well the information provided by the original input considering the values over 0.66 of the correlation coefficients among the new factor and the original variables. Another measure of the aggregation quality is the ratio $a_1/(a_1 + a_2 + a_3 + a_4 + a_5) = 0.93$, indicating that 93% of the inertia is explained by *I*. These results prove that it is suitable to use only one input variable instead of five. This new factor takes a considerable amount of information from the variables describing business R&D expenditures and the patents applications.

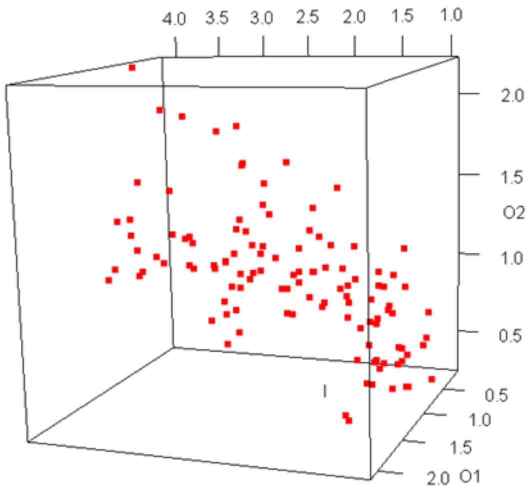

Figure 1. Decision making units in the input-output space

A simple analysis of the new input reveals the following aspects regarding the decision making units:

- Regions from Sweden, Finland, Norway register high levels meanwhile Slovakia, Portugal and Poland are characterized by low levels of input
- To be more specific, the difference between the maximum value (registered by Stockolm) and the minimum (Vychodne Slovensko from Slovakia) is 3.15, which is very high considering the average value of 2.24
- On average, the regions from Finland have the largest inputs in contrast to Hungary regions which register the lowest level.

We found a weak correlation among the output variables O_1 and O_2 and thus no possibility of aggregation. Hence, we investigate 116 units represented by three variables: *I, O1, O2* from the point of view of technical efficiency. As can be seen in Figure 1 this sample is rather homogeneous and we have no evidence of any outliers.

4.2 Testing the return to scale

With the aim of building a ranking of the 116 regional innovation systems based on the efficiency each unit transforms innovation inputs into innovation outputs, we used an output oriented DEA model assuming variable returns to scale. We employed a test based on the bootstrap algorithm discussed in the previous section, in order to select the type of return to scale of the production function. This test was developed by Simar and Wilson (2002). We went through the following steps in order to implement it:

1) Setting the null hypothesis and the alternative

2) Computing the test statistics denoted by T_{obs}:

$$T(\chi_n)=\frac{1}{n}\sum_{i=1}^{n}\frac{\hat{S}^{i}_{CRS,n}}{\hat{S}^{i}_{VRS,n}}=T_{obs}, \text{ where } \hat{S}_{CRS}=(\hat{\lambda}_{CRS})^{-1} \text{ and } \hat{S}_{VRS}=(\hat{\lambda}_{VRS})^{-1}.$$

which, for the original sample, takes the value 0.87.

3) Homogeneous bootstrap algorithm is used to generate B=1000 replicas. For each sample we computed the statistic $\hat{T}^{*b}$ using the same formula from step 2.

4) The 1000 values obtained at step 3 are used to approximate the *p-value* associated to this test by the formula:

$$p-value \approx \sum_{b=1}^{B}\frac{1(T^{*,b}\le T_{obs})}{B}$$

To be more precise, the p-value is given by the proportion of bootstrap samples with values $T^{*,b}$ less than the original observed value T_{obs}. In this case we rejected the null hypothesis of constant returns to scale given *p-value*= 0.004.

4.3 Estimating efficiency scores

Point estimates of the efficiency scores are computed by solving 116 linear programming problems (Charnes et al., 1994)). FEAR (Frontier Efficiency Analysis with R) library was used to compute the estimates presented below (Wilson, 2008).

Table 4. Efficiency scores distribution

Interval for $\hat{\lambda}$	Absolute frequency (Number of units)	Relative frequency (%)
$\hat{\lambda} = 1$	12	10.30%
(1 , 1.1)	11	9.50%
[1.1, 1.2)	10	8.60%
[1.2 , 1.3)	25	21.60%
[1.3 , 1.4)	12	10.30%
[1.4 , 1.5)	12	10.30%
[1.5 , 1.6)	3	2.60%
[1.6 , 1.7)	6	5.20%
[1.7 , 1.8)	9	7.80%
[1.8 , 1.9)	6	5.20%
[1.9 , 2.7)	10	8.60%

Farrell scores offer us the following information regarding the efficiency of the analysed innovation systems:

- The average efficiency is 1.41 with a standard deviation of 0.36;
- 10.3% of the regions received an efficiency score equal to one, showing efficiency. These regions belong to 7 countries from all 13 included in the analysis: three regions from Germany, two regions from Slovakia, two regions from Portugal, two regions from Italy, one from Hungary, one from Sweden and one from Poland;
- If we rank countries according to the average score of the regions investigated here, the most efficient countries will be: Slovakia, Italy and Germany with the averages 1.11, 1.18 and respectively 1.20;
- The regions within a country are very heterogeneous from this point of view. For example the German regions obtained scores lying in the interval (1, 1.65). Another example is that of the Slovak regions: two received efficiency score equal to 1 meanwhile the scores of the other two are 1.06 and 1.40.

The efficiency scores we analysed above were estimated on a small sample given the elimination of some region for which we confronted with missing data. This is the reason why countries like Romania are not represented in this sample. Hence we cannot draw the final conclusion without first estimating the bias, the variance of the efficiency scores and also the 95% confidence intervals. But in the nonparametric estimation framework statistical inference is not easy and the best solution is given by bootstrap. We generated B=2000 bootstrap replicas using the algorithm described in the previous section, in order to estimate the bias, the variance and the 95% confidence intervals associated to DEA scores. The correction of the estimators using their bias involves additional errors causing the growth of the variance. This means we have to decide which scores we should use for policymaking: the original scores or the bias corrected ones. According to Efron and Tibshirani (1993) rule, the correction is valid if the following condition is satisfied:

$$\widehat{bias}(\hat{\lambda}(x,y)) > \widehat{std}(\hat{\lambda}(x,y)) / 4$$

This restriction is accomplished by all units leading to a ranking of the innovation systems based on bias corrected scores. This decision involves the following:

- There are no decision making units with the same score (equal to 1) and it is easier to differentiate them;
- Bias correction increased the average efficiency to 1.50.

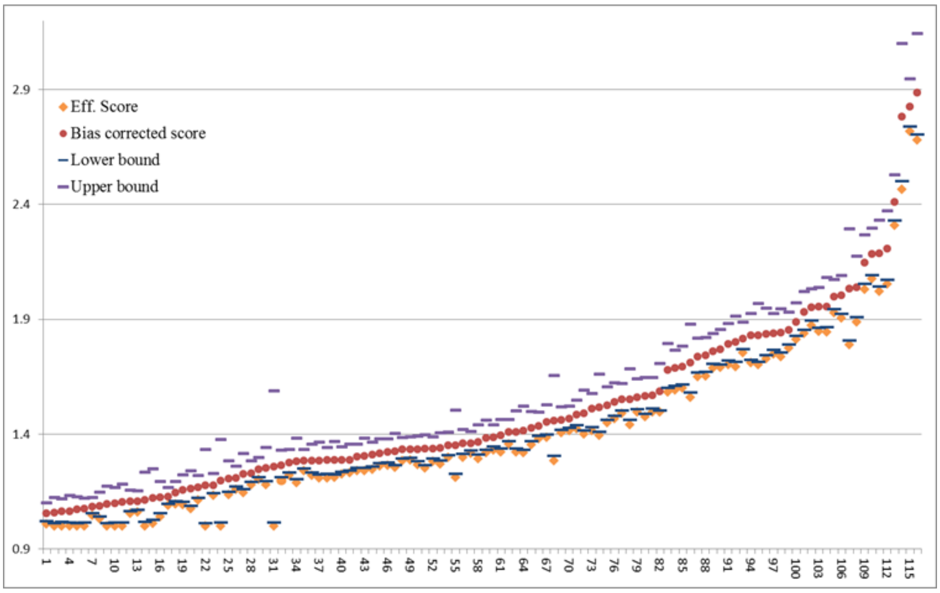

Figure 2. Efficiency estimation results at regional level

The length of the confidence intervals depends on the position of the decision making unit relative to the other units in the sample. The uncertainty regarding the technical efficiency is illustrated in Figure 2 which shows the confidence intervals length for all 116 units sorted by the value of the bias corrected efficiency score. We emphasize that the bias corrected scores lie inside the intervals meanwhile the initial scores lie outside, proving again that the original estimates were biased and we managed to correct it.

The ranking based on the new estimated measures of efficiency highlights the following aspects:

- The ranking is different from the one based on initial efficiency estimates;
- The most efficient region is Lazio from Italy. There is 95% chance that the score value associated to this region lies in the interval (1.019, 1.10). Even if Lazio is not among the regions with an estimated efficiency score equal to one, after the bias correction it ranks first in terms of technical efficiency. Very close in the ranking comes a region from Germany (Unterfranken) although the uncertainty in this case is higher: with a probability of 95% we can say that the efficiency score lies in the interval (1.012, 1.124);
- There are other two regions (Luneburg and Illes Balears) which despite the fact their initial scores were greater than one, are ranked above other regions rated according to the first estimates as 100% efficient.

For a better understanding and also for a validation of the classification presented in Table 5, we must keep in mind the variables we used to define the innovation systems. There is no surprise that Stockholm or Darmstadt are among the most efficient, considering the inputs that captured the human capital and the research investment and also the outputs which reflected the employment in high tech knowledge based sectors.

Table 5. First ten most efficient regional innovation systems

Region	Code	Eff. Score	Bias	Variance	Bias corrected score	95% Confidence interval		Rank (Eff. Score)	Rank (Bias corrected score)
						Lower bound	Upper bound		
Lazio	ite4	1.00888	-0.0462	0.00044	1.055039	1.019473	1.101065	13	1
Unterfranken	de26	1	-0.0581	0.00086	1.058063	1.011918	1.123611	1	2
Basilicata	itf5	1	-0.063	0.00065	1.063021	1.015555	1.118534	5	3
Zachodniopomorskie	pl42	1	-0.0643	0.00088	1.064285	1.015156	1.13273	7	4
Lisboa	pt17	1	-0.0715	0.00087	1.07145	1.010707	1.127967	9	5
Lombardia	itc4	1	-0.0737	0.00074	1.073673	1.013351	1.120476	4	6
Lüneburg	de93	1.04668	-0.0366	0.00035	1.083235	1.053444	1.123902	17	7
Illes Balears	es53	1.03263	-0.055	0.00078	1.087596	1.040943	1.148433	15	8
Stockholm	se11	1	-0.0958	0.00151	1.095771	1.011891	1.173586	12	9
Darmstadt	de71	1	-0.0968	0.00148	1.096842	1.015376	1.168097	2	10

About Stockholm we know that includes the capital of Sweden, being characterized by impressive development of "high tech" considering the number of factories producing high-technology goods like electronic cards, semi-conductors and advanced pharmaceuticals, and also by research centres that provide employment for graduates. According to Eurostat, in Stockholm region, there are ample opportunities to study at universities and various university colleges.

Darmstadt region, on the other hand, is known for the innovative branches in the engineering (mechanical, electrical, environmental), biotechnology and computing industries. Also the European Space Agency's European Operations Centre (ESOC) is also based in Darmstadt. Like Stockholm, the higher education and research system in this region enjoy a reputation in Europe.

It was not so obvious for other regions presented in Table 5 to be rated in the high performance group. For example Lazio is a region that lies at the centre of Italy, including also the city of Rome which accommodates various government ministries and head offices of State-run bodies, national banks. We found some relevant features that may explain Lazio position in our ranking:

- Given the historical context of the region, we can say that it is a privileged region in terms of culture;
- According to Eurostat about 14% of Italian students are concentrated in this region, also known for the high percentage of people employed in companies from "high tech" sector;
- High-tech sectors such as optics, telematics, pharmaceuticals, processing the data are highly promoted.

The arguments justifying the efficiency of the Portugal region, Lisbon, are based on the fact that here are concentrated the main economic and political institutions of the country, the largest companies and financial groups in Portugal, and a large number of scientific and technological research institutes. As a consequence, the workforce in this region is highly qualified.

The region ranked the forth in our classification, is located in the north-western portion of Poland. According to the Portrait of regions offered by Eurostat there some strong points of this region which enhanced the economic development as follows:

- well-developed tertiary education and scientific base; large percentage of students; educated and well-qualified labour force;
- access to renewable sources of energy;
- cultural diversity of the population.

The most striking situations are those of Basilicata and Illes Balears. These regions use inputs that are below the average value which could explain their classification is the most efficient group. Illes Balears is a region from Spain where the share of population attending tertiary education on vocational training programs is very low. Basilicata is a region in southern Italy with low inputs, with a value of GDP per capita below the national average. These examples highlight the fact that in order to decide which regions should be declared best practice in the field of innovation we have to consider a complex set of investigation tools based on different approaches. Thus before we draw the final conclusion we compare our ranking based on the efficiency analysis with the classification presented in the Regional Innovation Scoreboard report from 2009. According to the methodology presented in this report, a region can be assigned to one of the following groups: high innovators, medium-high innovators, average innovators, medium-low innovators, low innovators as a result of a cluster analysis. The regions Unterfranken, Darmstadt, Stockholm are classified among the high innovators, Lombardia and Luneburg are assigned to the second group meanwhile Lazio and Lisboa belong to the average innovators group and Basilicata, Illes Baleares and, Zachodniopomorskie belong to the last two groups.

It is obvious that a perfect concordance between the two methods of assessing innovation performance could not be achieved, given the following:

- The method presented in RIS report is based on variables aggregation and the main principle used here ”much is better” is different from our approach based on the comparison between effort and effect
- The technical efficiency analysis investigates the inputs effects on labour market meanwhile the performance indicators from RIS incorporate much more variables (including variables presenting missing values after using imputation techniques).

5. Conclusions

The final results of the analysis developed in this paper, consist of ranking the regions according to the efficiency of their innovation systems. Like in the national systems analysis (Matei, 2010), the most efficient regions are not necessarily in the innovation leaders group proposed by RIS 2009. Therefore, taking into account the classifications outlined in the RIS 2009 of the European Commission, we concluded that we may consider that regions such as Unterfranken of Germany, Stockholm of Sweden and Darmstadt of Germany are best practices in terms of innovation policies. These are developed regions that have rich resources (inputs) which are also technically efficient. This remark is based on the following issues that characterize the three regions:

- These are among the regions with many patent applications at the European Patent Office;
- The value of business research - development expenditures is well above the average of the 116 regions;
- The share of people employed in knowledge-intensive services far exceeds the average calculated for the 116 regions.

Thus given the complexity of innovation, the development and the implementation of innovation related policies must take into account a range of measures both qualitative and quantitative. The results obtained in this research demonstrates once again that the nonparametric DEA techniques is one of the quantitative tools that can help the policy makers to develop regional innovation strategies and the application presented provides the following improvements in this area:

- Dimensionality reduction and elimination of multicollinearity;
- Implementation of an algorithm for testing returns to scale;
- Statistical inference using bootstrap technique.

In conclusion, the analysis developed in this paper completes the previous results obtained in the field of innovation systems assessment, through the use of homogeneous bootstrap algorithm. Of course that solving the missing data problem would offer the possibility of increasing the accuracy of the results that may be achieved in estimating the efficiency of these systems. The importance given in the last years to innovation process and the efforts to measure it guarantees that in the future we will have more comprehensive data base describing the systems of innovation. Thus in the future we intend to continue this analysis and to improve the results by using robust nonparametric estimators known as partial frontiers.

Acknowledgment

Financial support from the Romanian National Authority for Scientific Research, CNCS - UEFISCDI, Project PN-II-ID-PCE-2011-3-0893 is gratefully acknowledged.

REFERENCES

[1]**Aldea, A., Ciobanu, A. (2011),** ***Analysis of the Renewable Energy Development Using Bootstrap Efficiency Estimates***. *Economic Computation and Economic Cybernetics Studies and Research; ASE Publishing.* Vol. 45, 1,77-91;

[2]**Aristovnik A.(2012),** ***The Relative Efficiency of Education and R&D Expenditures in the New EU Member States.*** Journal of Business Economics and Management Volume 13, Issue ;*Taylor & Francis Ed*, Co-Published with Vilnius Gediminas Technical University, pages 832-848, ISSN 1611-1699 (Print), 2029-4433 (Online);

[3]**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*, Volume 39, 722-735;

[4]**Buesa M., Heijs J., Ramirez Pellitero M., Baumert T. (2006),** ***Regional Systems of Innovation and the Knowledge Production Function: The Spanish Case*****;** *Technovation* 26, 463-472;

[5]**Charnes A., Cooper, W.W., Seiford, L.M. (1994),** ***Data Envelopment Analysis: Theory, Methodology and Application****;* Edition: 5, illustrated, *Springer*, 1994;

[6]**Coelli, T.J., Prasada Rao D.S., O' Donnel C., Battesse G. E. (2005),** ***An Introduction to Efficiency and Productivity Analysis***; Second Edition, *Springer;*

[7]**Daraio, C., Simar, L., (2010),** ***Advanced Robust and Nonparametric Methods in Efficiency Analysis: Methodology and Applications****; Springer*;

[8]**Doloreux D., Parto, S. (2005),** ***Regional Innovation Systems: Current Discourse and Unresolved Issues***; *Technology in Society* 27, 133-153;

[9]**Efron, B., Tibshirani, R.J. (1993),** ***An Introduction to the Bootstrap***; *Chapman & Hall*, New York;

[10]**Fritsch M. (2002),** ***Measuring the Quality of Regional Innovation Systems: A Knowledge Production Function Approach***. *International Regional Science Review*, Volume: 25, Issue: 1, Pages: 86-101;

[11]**Farrell, M.J (1957),** ***The Measurement of Productive Efficiency***. *Journal of Royal statistical society*, Series A, CXX, Part 3, 253-290;

[12]**Heidenreich, M., Wunder, C. (2006),** ***Patterns of Regional Inequality in the Enlarged Europe*** . *European Sociological Review*, Volume 24, Issue 19-36;

[13] **Kim S.K., Lee B.G., Park B.S., Oh K.S.(2011),** ***The Effect of R&D, Technology Commercialization Capabilities and Innovation Performance***; Technological and Economic Development of Economy, Volume 17, Issue 4,

Taylor & Francis Ed, Co-Published with Vilnius Gediminas Technical University, pages 563-578, , **ISSN** 2029-4913 (Print), 2029-4921 (Online);
[14]**Kutvonen, A. (2007),** ***Ranking Regional Innovation Policies: DEA Based Benchmarking in a European Setting***. *Research Report* 193, ISBN 978-951-214-516;
[15]**Matei, M (2010),** ***Innovation Efficiency Analysis for Romania***. *Economic Computation and Economic Cybernetics Studies and Research*, *ASE* Publishing, Vol. 44, 3, page 193- 204;
[16] **Silva J, Simoes J., Sousa G. (2012),** ***Inovații în marketing şi capacitatea inovatoare a marketingului: studierea firmelor portugheze***; Amfiteatru Economic, no. 31, Academia de Studii Economice din Bucuresti, ISSN 1582 – 9146;
[17]**Simar, L; Wilson, PW (1999) ,** ***Sensitivity Analysis of Efficiency Scores: How to Bootstrap in Nonparametric Frontier Models***; *Management Science*, Vol. 44, No.1, 49-61;
[18]**Simar, L; Wilson, P.W. (2000, a),** ***Statistical Inference in Nonparametric Frontier Models: The State of the Art;*** *Journal of productivity analysis* , 13 , 49-78;
[19]**Simar, L, Wilson, PW. (2000, b) ,*****A General Methodology for Bootstrapping in Nonparametric Frontier Models***. *Journal of Applied Statistics*, Volume 27, 779-802;
[20]**Simar, L., Wilson, P.W. (2001),** ***Testing Restrictions in Nonparametric Efficiency Models***. *Communications in Statistics*, 30 (1), 159- 184;
[21]**Simar, L., Wilson, P.W. (2002),** ***Non- parametric Tests of Returns to Scale.*** *European Journal of Operational Research*, Vol. 139, 115- 132;
[22]**Slavo R. (2004),** ***A Two Tier or Multi-tier Europe? Assessing the Innovation Capacities of Central and East European Countries in the Enlarged EU***. *Journal of Common Market Studies,* Vol. 42, Issue 3, 641-666;
[23]**Spircu L., Spircu T., Albu C., Bujor, M. (2009),** ***Descriptive and Multidimensional Indicators for Ordering Educational Offers***. *Economic Computation and Economic Cybernetics Studies and Research; ASE Publishing*; Vol. 43, 1135- 147;
[24]**Ta-Wei Pan, Shiu-Wan Hung, Wen Min LU (2010),** ***DEA Performance Measurement of the National Innovation System in Asia and Europe. Asia-Pacific Journal of Operational Research***, Vol 27, No. 3, 2010, 369-392;
[25]Wilson, P.W. (2008), FEAR*:* ***A Software Package for Frontier Efficiency Analysis with R;*** *Socio Economic Planning Sciences*, Volume 42, Issue 4, 247-254;
[26]**Zabala-Iturriagagoitia, J.M.,Voigt, P.,Gutierrez- Gracia, A.,Jimenez-Saez, F. (2007),** ***Regional Innovation Systems: How to Assess Performance****;* *Regional Studies*, Vol. 41 Issue 5, 661-672;
[27]* http://circa.europa.eu/irc/dsis/regportraits/info/data/en/ite4.htm
[28]*http://www.proinno-europe.eu/.
[26]* *www.proinno-europe.eu/.../**RIS_2009-**Regional_Innovation_Scoreboard.pdf .*