WP4/03 SEARCH WORKING PAPER

A country-level knowledge production analysis with parametric and non parametric methods

Stefano Usai, Barbara Dettori, Elisa Gagliardini

January 2013

SEARCH
SHARING KNOWLEDGE ASSETS: INTERREGIONALLY COHESIVE NEIGHBORHOODS

EUROPEAN COMMISSION
European Research Area
SEVENTH FRAMEWORK PROGRAMME
Funded under Socio-economic Sciences & Humanities

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2010-2.2-1) under grant agreement n° 266834

A country-level knowledge production analysis with parametric and non parametric methods

Stefano Usai, Barbara Dettori, Elisa Gagliardini

University of Cagliari, CRENoS

Abstract

Economic growth, according to both applied and theoretical economists, is not entirely dependent on traditional production factors endowment, such as physical capital and labour, but more and more related to the stock of intangible resources such as culture, competence, innovative capacity, knowledge.

Our main aim is, consequently, to provide an exploratory analysis on the phenomena of knowledge creation and diffusion in Europe and its neighbours at the country level. We firstly describe the innovative activity across countries in order to make available a comprehensive picture of this phenomenon across and beyond Europe. We analyse both input (R&D expenditure) and output indicators (patent applications) for the 27 European countries and the 16 European Neighbouring Countries.

Moreover, we analyse the main factors influencing the innovation process. We pursue this aim by adopting both parametric and non-parametric methods to investigate about the knowledge production function at the country level. The analysis is mainly speculative because the absence of information about some potentially important phenomena, such as human capital, may hinder our results and conclusions. Nevertheless, main results are robust and confirm previous analysis at the country and regional level. Moreover, they add some original finding about the potential for catching up of European Neighbouring Countries.

1. Introduction

Economic growth, according to both applied and theoretical economists, is not entirely dependent on traditional production factors endowment, such as physical capital and labour, but more and more related to the stock of intangible resources such as culture, competence, innovative capacity, knowledge.

Policy makers at the European Union reached the same conclusion and have set several initiatives which put particular attention to the process of creation of such intangibles. In particular, Europe 2020 agenda has confirmed the previous Lisbon strategy's goal to make Europe more competitive thanks to knowledge and technological change. Moreover, the high heterogeneity displayed by countries and regions with regards to their capacity to create knowledge and innovation, but also in their ability to exploit knowledge diffusion across the European territory, has motivated a special focus on its territorial dimension. This focus extends beyond borders since the European Neighboring suffers from a relatively poor economic performance mainly due to a severe technological gap with respect to the European Union.

Our main aim is, consequently, to provide an exploratory analysis on the phenomena of knowledge creation and diffusion in Europe and its neighbors at the country level. We firstly describe for the first time national innovative activity in order to make available a comprehensive picture of this phenomenon across and beyond Europe. We analyse both input (R&D expenditure) and output indicators (patent applications) for 43 countries: the 27 European ones and the 16 European Neighboring Countries. Moreover, we analyse the main factors influencing the innovation process. We pursue this aim by adopting parametric and non-parametric methods to investigate the knowledge production function at the country level.

More specifically, the analysis is based on regression models and on Data Envelopment Analysis (DEA). While regression models are particularly suitable to measure central tendencies of a given phenomenon, DEA is more adequate for benchmarking analysis as it permits to identify the best performing units within a given set of entities. The DEA approach will allow us to single out the specific characteristics of each region and to determine how far they are in relative terms from the most efficient areas, so that we will provide an assessment of the potential productive gains not yet accrued by inefficient regions. Since, in general, the two methods provide different indications on the same object of analysis, we employ both of them in a complementary guise in order to gain wider and different insights for the comparison of European Union and European Neighbouring countries innovative performance.

The first part of the analysis, is, therefore, devoted to the investigation of the impact of intangible assets on the innovative capacity of a region. We present results for a standard knowledge production function (Griliches, 1979) with the R&D as the main input. The analysis is mainly speculative because the absence of information of some potentially important phenomena, such as human capital, may hinder our results and conclusions. Nevertheless, main results are robust and confirm previous analysis at the country and regional level. Moreover, they also add some original finding about the potential for catching up of European Neighbouring Countries. The second part of the paper is devoted to the application of DEA in order to assess the relative innovative performance of countries with respect to the technological frontier. Despite some drawbacks, again due to the quality of the dataset, results are particularly interesting since efficiency and inefficiency scores are not necessarily attributable to EU and ENC respectively.

The paper is structured as follows. In the following section a descriptive analysis of innovative activity is offered in order to understand the main features of the background scenario for our empirical analysis. The third and the fourth section present the econometric and the data envelopment analysis respectively. Results are then commented in the final section where some tentative and cautious policy conclusions are put forward.

2. Innovative activity: input and output indicator

This report provides an analysis of ENCs' technological and innovation activities where the output of the innovative activity is mainly measured by patent production whilst the main input is the expenditure in research and development. The analysis is based on the latest available international comparable data on patents made available by OECD and WIPO and on R&D published by the World Bank (see appendix for details). Tables below present the standard statistical indicators designed to illustrate the trends and structure of science, technology and innovation for the EN countries and compare them to the European Union ones.

Although patents do not cover every kind of innovation activity, they do include many of them. Patents have become one of the most widely used sources of data in the construction of indicators on inventive output, as they are closely linked to invention and they provide detailed information available in relatively long time-series. Nevertheless, patent indicators also have several shortcomings and therefore need to be combined with other output indicators in order to obtain a full picture of innovation activities in individual countries. In conclusion, we have to take into account that not all inventions are patented, not all patents transform into an innovation and that patents have not the same economic value.

Patent statistics have made rapid progress in recent years. A continuing effort is made to improve the quality and availability of IP statistics. Even if it is difficult to obtain data for all IP offices with all possible breakdowns, every effort is made to cover data for as many offices and countries as possible. We have to notice that when it is necessary and feasible, missing data are estimated. Intellectual property data published in this report are mostly taken from the OECD and WIPO Statistics Database that offer a unique tool for analysts and producers of patent data and indicators. The WIPO Statistics database offers information on patent counts based on the international filing date and the residence country of the first named applicant. It also allows us to have data for Azerbaijan, Libya and Syria which are not included in the OECD patent database.

Only for Palestinian territories data are not available in either the databases. We should notice that, therefore, the indicator obtained by WIPO is only imperfectly comparable with the one which is obtained from the OECD database. In the latter database, as a matter of fact, multiple applicants and inventors are taken into account by dividing patents into fractions according to their residence[1] country. Nonetheless, the divergence in the two databases in term of per capita values is negligible.

The main indicator for innovation activity is identified as the number of Patent Cooperation Treaty (PCT) and number of Patents Applications to the European Patent Organization (EPO). The PCT indicator enables an international patent application to have the same effect as the national one in each of the contracting states designated in the application. In every case where the EPO is designated, the patent is known as an Euro-PCT patent. The PCT system is superimposed on the national and European systems, but patents are always nationally and/or regionally granted. All PCT applications are centralized through the World Intellectual Property Organisation (WIPO). We have to take into account that the PCT patent data suffer (or suffer less) from the usual home bias effect of the EPO data. Since patents at EPO protect innovation within their respective geographical area, they are preferred by domestic firms, and thus their quota overestimates their innovative capability with respect to foreign firms.

We are going to present patent data both by applicant's and by inventor's residence country. In most cases the applicant is an institution[2] (a firm, a government body such as an university or a public laboratory), which is the legal owner of the patent at the time of application. Thus, counting patents according to the applicant's country of residence tends to measure the degree of control on patents by each country's residents. It reflects the firms' innovativeness of a given country, whatever the location of their research facilities. The inventor, instead, is always an individual, usually a researcher employed in the applicant firm. Since patent statistics by inventor's captures the national location where the invention is introduced, it better reflects the technological innovativeness of researchers and laboratories located in a given country, whatever their ownership.

Moreover, in order to measure inventive activity, patent are counted according to the priority year. Since it corresponds to the first filing worldwide, this date is considered the closest to the introduction of the invention. Finally, we refer to the stock of patents filed within the period starting from 2000 to 2008 in order to have a relatively long time period since the innovative activity of ENC is rather sporadic.

In Table 1 we compare the number of patents filed to the EPO per million population in the EU27 and ENC. It is clear that there are large disparities in patenting activity comparing EU27 and ENC. The EU27 is the most active world economy in patenting at the EPO with an annual average

[1] When a patent has more than one inventor/applicant, a proportional share is attributed to each inventor. If inventors reside in different country, the patent is attributed proportionally to each conuntry.

[2] For EPO patents the share of institutions in total applicants is usually estimated to be higher than 90%.

that overcomes 2 thousand patents per country. But we can observe a high level of heterogeneity within EU27 countries, too. The patent activity in the European Union is, in fact, highly concentrated in the so called Old Europe (EU15) which generate most patent applications: with more than 3.5 thousand patents per country every year, EU15 accounted for more than 90% of overall patent activity from 2000 to 2008 in the EU27. Among EU Member States, Germany is the leader, while among the non-EU leading countries in patent applications to the EPO we find also Israel, an outlier among ENCs. It is interesting to notice that Israel apply for a number of patents per capita which is even higher than the EU15 average. When we keep Israel separated, the neighboring countries show very low levels of patenting activity. The South and East area reveal 69 patents by inventor and 38 by applicant, which implies less than 1 patent per million inhabitants for both statistics. Israel applies for more than 95% of total ENC patents (both counting by inventors and applicants).

Another interesting aspect is to determine whether these observations are specific to EPO patents or apply also to PCT patents (bottom part of Table 1). Not surprisingly, ENCs prefer to apply for international patents at PCT with respect to EPO, reaching nearly 1.9 and 1.6 thousand patents by inventors and by applicants, respectively. Thus the number of PCT patents per capita, both by inventor and applicant, is remarkably higher: when including Israel in the ENC average, it reaches respectively 7 and 6 patents per million inhabitants, outperforming the NMS12 (the new entrant members, mostly from Eastern Europe) averages. Moreover, the Israeli share on total ENC PCT patents is slightly lower, although very high (almost 90% both by inventor and by applicant). When we exclude Israel, in fact, per capita PCT patents, allocated by inventor, for Southern ENCs are just around 0.4 while they are around 1.6 for Eastern ENCs. Results are not very different when we count patents in terms of applicant, even though the numbers are always lower.

Table 1. PCT and EPO patents by area, average 2000-2008

EPO	by inventor(s)	EU27	EU15	NMS12	ENC	South*	East	Israel
	total patents	54,662.2	54,057.8	604.4	1,212.5	31.7	37.5	1,143.3
	patents (million pop)	111.4	139.9	5.8	4.6	0.2	0.5	167.9
	by applicant(s)							
	total patents	53,207.8	52,742.0	465.9	959.4	19.9	18.1	921.4
	patents (million pop)	108.5	136.5	4.5	3.7	0.1	0.2	135.3
PCT	**by inventor(s)**	**EU27**	**EU15**	**NMS12**	**ENC**	**South***	**East**	**Israel**
	total patents	43,198.2	42,471.6	726.7	1,887.4	69.1	123.8	1,694.5
	patents (million pop)	88.1	109.9	7.0	7.2	0.4	1.6	248.9
	by applicant(s)							
	total patents	42,791.0	42,183.7	607.3	1,568.8	53.9	98.1	1,416.8
	patents (million pop)	87.2	109.2	5.8	6.0	0.3	1.3	208.1

** South ENC data for patents by inventor(s) and by applicant(s) excludes Israel*
Source: CRENoS calculation on OECD Data

The analysis of patenting activity is developed in Table 2, where patent activity is referred to each ENC country. Total ENCs EPO patents applications from 2000 to 2008 are slightly less than 8.5 thousand and more than 13 thousand those filed under the PCT. As we might expect, Israel is the country with the highest number of patents, which represents 96 and 90 per cent of total EPO and PCT, respectively.

Within the ENC group, with a much more modest number of EPO patents, we find Ukraine (with 106 total patents, just 0.25 per million inhabitants per year). In the same country PCT patents per million population are much higher and amount to 1.44. Due to the low value of population, Armenia reaches 1.90 patents per million inhabitants on average every year, even if its inhabitants and firms have filed just 52 PCT patents in the period under observation.

Table 2 EPO and PCT patents by applicant(s) residence, 2000-2008

country	EPO stock	EPO per million pop	PCT stock	PCT per million pop
Armenia	3	0.11	52	1.90
Azerbaijan	6	0.08	43	0.57
Belarus	37	0.42	133	1.50
Georgia	7	0.17	51	1.29
Moldova	10	0.31	31	0.94
Ukraine	106	0.25	616	1.44
ENC-East	168	0.24	926	1.34
Algeria	4	0.01	52	0.18
Egypt	43	0.07	254	0.39
Israel	8,293	135.35	12,751	208.11
Jordan	68	1.42	30	0.64
Lebanon	13	0.36	16	0.45
Libya	2	0.04	1	0.02
Morocco	26	0.10	88	0.33
Syria	4	0.02	25	0.16
Tunisia	25	0.28	45	0.51
ENC-South	8,478	5.09	13,262	7.97

Note: EPO data for AZ, LY, SY : CRENoS calculation on European patent bulletin dataset.
PCT data for AZ, LY, SY : WIPO data

Source: CRENoS calculation on OECD Data

When we exclude Israel, the yearly number of patents per million population in the Southern area decreases drastically from 5.1 to 0.1 for EPO statistics and from 8.0 to 0.3 for the PCT ones; in fact the patents per million inhabitants for countries such as Libya, Algeria and Syria are very limited; excluding the EPO statistics for Jordan, none of them reaches the unit.

In Table 3 we can study the same innovation process by another classification criteria, showing the EPO and PCT statistics counted by inventor's residence country. We observe a general increase of the phenomenon, since the number of patents is sensibly higher than when we count by applicant's residence country. This is what usually occurs in the case of small or less innovative countries, reflecting the higher level of internationalisation of their research activities with a foreign ownership of domestic inventions.

Table 3. EPO and PCT patents by inventor(s) residence, 2000-2008

country	EPO stock	EPO per million pop	PCT stock	PCT per million pop
Armenia	10	0.37	60	2.18
Azerbaijan	9	0.12	-	-
Belarus	61	0.69	173	1.96
Georgia	20	0.52	71	1.81
Moldova	14	0.43	39	1.19
Ukraine	231	0.54	771	1.80
ENC-East	346	0.50	1,114	1.81
Algeria	7	0.03	58	0.20
Egypt	79	0.12	305	0.46
Israel	10,289	167.94	15,250	248.91
Jordan	82	1.72	56	1.18
Lebanon	35	0.99	41	1.14
Libya	3	0.05	-	-
Morocco	50	0.18	110	0.41
Syria	8	0.05	-	-
Tunisia	32	0.36	52	0.58
ENC-South	10,585	6.36	15,873	10.93

Note: EPO data for AZ, LY, SY : CRENoS calculation on European patent bulletin dataset.
Source: CRENoS calculation on OECD Data

After Israel, with 168 patents per million inhabitants, Jordan presents the best performance per population (1.7), followed by Lebanon that reaches the unit per million inhabitants; then, below, we find Belarus (0.7), Ukraine (0.5) and Georgia (0.5), while all the other countries are under the 0.45 patents per million inhabitants. On the bottom we find Algeria, Syria and Libya where innovation activities are almost absent. The ENCs international innovative activity reaches the highest value when it is counted by inventor: almost 16 thousand patents in the period considered, equal to more than 8 patent per million population every year, but of course the data is again strongly biased by Israel, which files more than 15 thousand international patents and is close to 250 patents per million inhabitants. When we look beyond Israel's case, it is interesting to notice the (relative) good performance for Armenia (60 patents in the considered period allow it to reach the highest value per population) and for Belarus, with almost 2 patents per year per million inhabitants. Further down in the ranking we find Georgia and Ukraine (about 1.8 patents per million population) that stand off all the other countries.

Finally, another way to measure innovation activity is to refer to investments in research and development, an input measure rather than an output one such as patents. We have to point out that for ENCs data related to R&D are very limited and, consequently in some cases, values are estimated. Moreover, they are missing for several years and as a result the indicator cannot always refer to the same period.

As we can see from Table 4, the Southern group has an average of R&D expenditure of one point while the Eastern group reaches approximately 0.8 per cent. However, we have to take into account that the slightly higher average for Southern ENC is again biased by Israel; its Gross domestic expenditure on R&D (GERD) reaches the exceptional average of 4.50 per cent of GDP in the period between 2000 and 2008. An average which is much higher with respect to the same index for EU27 or for the whole of OECD countries. As for the other countries, the share of R&D in GDP is quite low, around 0.2 per cent of GDP, in several countries, such as Armenia (0.22), Azerbaijan (0.23), Egypt (0.25), Georgia (0.21) and Algeria (0.20). Jordan and Moldova have a slightly higher share (0.39 and 0.43 respectively). Finally, the countries which invest the most are Belarus and Ukraine among the Eastern countries (with shares of 0.71 and 0.99 respectively) and Morocco and Tunisia in the South (with a quota of 0.61 and 0.88 respectively).

Table 4 R&D expenditure, country average 2000-2008

country	**R&D exp PPP** (constant 2005, million $)	**R&D expenditure** (% GDP)
Armenia	25.39	0.22
Azerbaijan	88.19	0.23
Belarus	566.05	0.71
Georgia	27.76	0.21
Moldova	37.38	0.43
Ukraine	2461.82	0.99
ENC-East	3206.58	0.80
Algeria	421.02	0.20
Egypt	865.71	0.25
Israel	7192.88	4.50
Jordan	94.50	0.39
Lebanon	-	-
Libya	-	-
Morocco	624.72	0.62
Syria	-	-
Tunisia	614.81	0.88
ENC-South	9813.63	1.07

Source: CRENoS calculation on World Bank data

3. The production of new knowledge

In this section we present the analysis aimed at investigating the returns of R&D expenditures on each country innovative capacity. The analysis is carried out by employing both parametric and non parametric methods. The former is based on the theoretical framework of the knowledge production function (Griliches, 1979) and has been developed mainly with respect to the regional territorial dimension (see Marrocu et al, 2012, in this chapter for more details). The non parametric Data Envelopment Analysis (DEA), on the contrary, permits to single out the best practices among regions or countries (or other typologies of territories) in performing innovation activities and to identify the less efficient ones in converting R&D investments (and human capital if available) into the creation of new knowledge (see Foddi and Usai, 2012, in this chapter for more details).

3.1 Econometric estimation

Following the well-established literature on the estimation of knowledge production functions (for an extensive review see Marrocu et al, 2012, in this research project), the dependent variable, to proxy innovative performance, is given by the amount of patent activity in a country in a certain period (*pat*). The variable is built summing the available data for all the 8 years in order to ensure that the number of zero values is kept to a minimum. The other conventional indicator used in the literature is the expenditure in R&D which is considered the principal input in the KPF model. The research and development (*rd*) effort is measured by the total intramural R&D expenditure over GDP. Moreover, we also include the resident population as a control variable to account for the relative dimension of countries. It was not possible to include also a variable on human capital, based on degree attainments, since this statistics is not available for all ENCs. Finally, we include firstly two additive dummies to assess potential differences for three groups of countries, that is the so called Old Europe (EU15), the 12 new members countries (NMS12) and the ENCs. Secondly we also insert three multiplicative dummies to assess the sign and size of the coefficient of the R&D with respect to the groups of EU15, NMS12 and ENCs.

As a result, the general form of the empirical model for the KPF is as follows

$$inn = f(inputs, controls) \qquad (1)$$

Consequently, (1) can be formalized, as a result of the log transformation of a Cobb-Douglas function, as follows:

$$pat_{it} = \alpha + \beta_1 rd_{it} + \beta_3 pop_{it} + \varepsilon_{it} \qquad (2)$$

where i=1,…41 and t as explained above; lower letters indicate log-transformed variables.

Note that all the explanatory variables included in the model are not lagged because of the lack of data. For this reason we are aware to be in presence of endogeneity problems and that results have be considered with cautious.

When we include the additive dummies the function becomes

$$pat_{it} = \alpha + \alpha_1 ENC + \alpha_2 EU15 + \beta_1 rd_{it} + \beta_2 pop_{it} + \varepsilon_{it} \qquad (3)$$

And when we include the multiplicative dummies as follows

$$pat_{it} = \alpha + \beta_1 ENC * rd_{it} \quad + \beta_2 EU15 * rd_{it} \quad + \beta_3 NMS12 * rd_{it} \quad + \beta_4 pop_{it} + \varepsilon_{it} \tag{4}$$

Table 5, 6, 7 and 8 report the aggregate estimations based on four different dependent variables: EPO and PCT patents, both by applicant(s) and by inventor(s) country of residence. The information on inventors is more indicative of the location where the invention really occurred while the information on applicants, that is on the location of the owner of the patent, tell us where the invention is going to produce economic results, if any.

Main results do not offer any surprise. A first common result is that the population is always statistically significant with the expected sign for all four specifications. At the same time, the R&D variable shows the same behavior with a slightly higher coefficient when comparing the specification with EPO patents to the corresponding PCT one.

The interesting, even though expected, results comes from the analysis of the last two columns in each tabled, those where the equation 3 and 4 are estimated. First of all, column three shows that on average ENC produce less patents than both NMS12 and especially than EU15 countries. Moreover, results in the last column of each table show that R&D plays a different role in knowledge production with respect to the three groups considered: the multiplicative dummy is significant and different for EU15, NMS12 and ENCs in all estimates. In particular, the magnitude of the coefficients suggests a weaker contribution of the R&D expenditure in the ENCs with respect to the EU15, but surprisingly slightly stronger if compared to NMS12. Results show that there is a potential for improving ENC performance in patenting activity. The efficiency gap shown by these countries is further analysed in the following section.

Table 5. Knowledge Production Function estimation with EPO inventors

Dependent variable: EPO by inventor(s) residence				
R&D	2.26***	1.79***	1.42***	
	(0.23)	*(0.26)*	*(0.35)*	
Population	0.65***	0.74***	0.60***	0.63***
	(0.14)	*(0.11)*	*(0.08)*	*(0.18)*
Dummy ENC		-1.77***	-1.15***	
		(0.52)	*(0.50)*	
Dummy EU15			1.83***	
			(0.42)	
R&D*ENC				4.88***
				(0.78)
R&D*EU15				6.90***
				(1.36)
R&D*NMS12				4.29**
				(1.75)
Constant	4.29***	4.39***	3.97***	4.27***
	(0.57)	*(0.47)*	*(0.39)*	*(0.85)*
Observations	39	39	39	38
R-squared	0.79	0.85	0.90	0.75

Estimation methods: OLS robust

Robust standard errors in parentheses, significance ***1%, **5%, *10%

All variables are log-transformed. R&D is total expenditure over GDP. R&D*ENC/EU15/NMS12 is a multiplicative dummy

Table 6. Knowledge Production Function estimation with EPO applicants

Dependent variable: EPO by applicant(s) residence				
R&D	2.46***	1.89***	1.44***	
	(0.26)	*(0.29)*	*(0.37)*	
Population	0.41***	0.58***	0.41***	0.46**
	(0.19)	*(0.15)*	*(0.11)*	*(0.21)*
Dummy ENC		-2.16***	-1.42***	
		(0.56)	*(0.57)*	
Dummy EU15			2.20***	
			(0.47)	
R&D*ENC				5.36***
				(0.9)
R&D*EU15				*7.73****
				(1.58)
R&D*NMS12				*3.70**
				(2.04)
Constant	4.91***	4.97***	4.52***	4.76***
	(0.85)	*(0.71)*	*(0.61)*	*(0.99)*
Observations	39	39	39	38
R-squared	0.74	0.81	0.88	0.71

Estimation methods: OLS robust
Robust standard errors in parentheses, significance ***1%, **5%, *10%
All variables are log-tranformerd. R&D is total expenditure over GDP. R&D*ENC/EU15/NMS12 is a multiplicative dummy

Table 7. Knowledge Production Function estimation with PCT applicant

Dependent variable: PCT by applicant(s) residence				
R&D	1.95***	1.65***	1.29***	
	(0.14)	*(0.20)*	*(0.26)*	
Population	0.55***	0.60***	0.46***	0.53***
	(0.11)	(0.12)	(0.08)	(0.15)
Dummy ENC		-1.12**	-0.51	
		(0.51)	*(0.48)*	
Dummy EU15			1.80***	
			(0.38)	
R&D*ENC				3.92***
				(0.67)
R&D*EU15				*6.71****
				(1.17)
R&D*NMS12				*3.43***
				(1.52)
Constant	4.92***	4.97***	4.60***	4.75***
	(0.58)	*(0.56)*	*(0.44)*	*(0.73)*
Observations	39	39	39	38
R-squared	0.79	0.82	0.88	0.76

Estimation methods: OLS robust

Robust standard errors in parentheses, significance ***1%, **5%, *10%
All variables are log-tranformerd. R&D is total expenditure over GDP.
R&D*ENC/EU15/NMS12 is a multiplicative dummy

Table 8. Knowledge Production Function estimation with PCT inventor

Dependent variable: PCT by inventor(s) residence				
R&D	1.88***	1.62***	1.33***	
	(0.15)	*(0.21)*	*(0.27)*	
Population	0.74***	0.80***	0.68***	0.69***
	(0.09)	(0.09)	(0.08)	(0.14)
Dummy ENC		-1.05**	-0.52	
		(0.52)	*(0.49)*	
Dummy EU15			1.44***	
			(0.36)	
R&D*ENC				3.74***
				(0.61)
R&D*EU15				6.12***
				(1.06)
R&D*NMS12				*4.07***
				(1.37)
Constant	4.11***	4.06***	3.82***	4.14***
	(0.59)	*(0.36)*	*(0.33)*	*(0.66)*
Observations	38	38	38	38
R-squared	0.83	0.85	0.80	0.80

Estimation methods: OLS robust
Robust standard errors in parentheses, significance ***1%, **5%, *10%
All variables are log-tranformerd. R&D is total expenditure over GDP.
R&D*ENC/EU15/NMS12 is a multiplicative dummy

3.2 Data Envelopment Analysis (DEA)

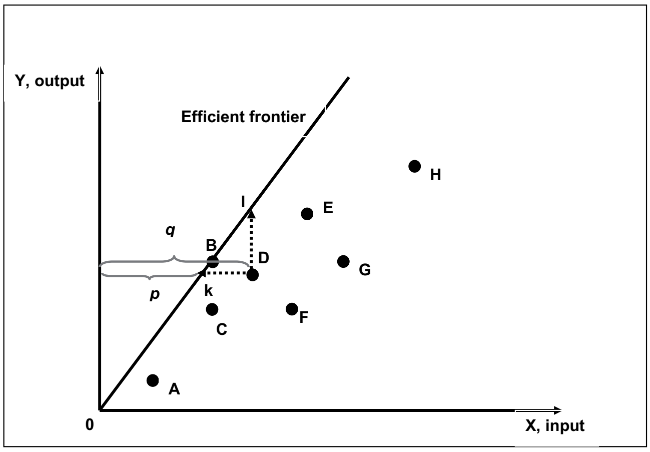
In this section we carry on our analysis on knowledge production by implementing a non parametric methodology: the DEA approach. This methodology was firstly developed by Farrell (1957) and based on mathematical programming techniques. One of the essential features of DEA, which makes this tool particularly suitable in this kind of analysis, is that it does not require to choose a specific functional form for the relation linking inputs to outputs. It aims at identifying the best performing units (countries in our case) among a set of entities whose objective is to convert multiple inputs into multiple outputs. Such inputs and outputs can be multiple and can be expressed in different units of measurement, as long as they are the same for all the decision making units (DMUs). Thanks to its high flexibility DEA has been proved successful in identifying various sources of inefficiency, in particular in studying benchmarking practices.

The best performance is characterized in terms of efficiency, so that the most performing units define the efficient frontier, which “envelope” all the other units. The technology frontier (efficiency frontier) is then defined as the maximum output attainable from each input level (see Coelli et al., 2005) and regions may or may not be on the frontier of this technology. Regions are therefore evaluated by calculating their distance from the frontier. In the analysis presented in the subsequent sections we focus on “technical” efficiency.

To illustrate how the DEA approach[3] operates we consider Figure 1 where we report different units labeled from A to H. If we assume constant return to scale (CRS), the frontier is identified, on the basis of the available empirical information, by DMU B, which is fully efficient. According to Cooper et al. (2007) a DMU is said to be fully (100%) efficient if the performance of other DMUs does not show that some of its inputs or outputs can be improved without worsening some of its other inputs or outputs. Note that this notion refers to “technical” efficiency and it does not require a priori information on prices or weights accounting for the relative importance of inputs or outputs.

Accordingly, DMU D, which is not on the frontier, will have an efficiency level proportional to its distance to DMU D. This measure is given by the ratio *p/q*, which is equal to 0.75 and it implies that if DMU D proportionally reduces all the inputs to the 75% of the their actual amounts, it could still produce the same level of output. In this way DMU D would be projected horizontally towards the efficient frontier. Under the assumption of constant returns to scale (CRS), the same efficiency gain would be obtained by a vertical projection, in this case with the same input amount DMU D could produce a level of output 33% (1/0.75=1.33) greater with respect to the previous one and move vertically towards efficient frontier, at point I. DMU B is called the benchmark or reference unit for DMU D.

Figure 1. DEA-CRS model, one input-one output



Source: CRENoS

In the former case we talk of *input*-oriented measures of efficiency, whilst in the latter case the measure is an *output*-oriented one. Note that under the assumption of CRS the two orientation identify the same frontier and the same set of efficient DMUs, only the measures associated with

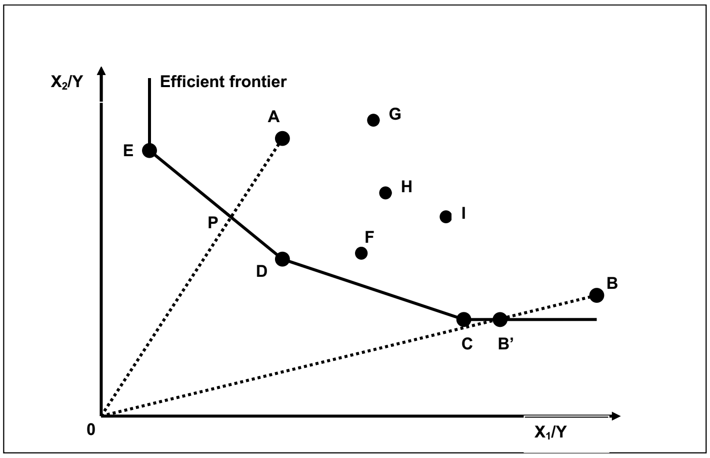
[3] This description is mainly based on Coelli (1996) and Cooper et al. (2007).

the inefficient DMU can be different. Note also that in the case of DMU D efficiency can be achieved by each movement in the area k-D-l.

Following Charnes et al. (1978) the maximization problem for each DMU is based on the ratio of outputs to inputs, which is used to measure the efficiency of a DMU with respect to all other DMUs. When the output to inputs ratio is maximized the model is referred to as input-oriented model; conversely, we have an output-oriented model when the ratio is inverted and a minimization problem is solved. The model we are going to use is also known as “Farrell model” and it can only provide measure of “weak” efficiency as it does not account for the presence of possible non-zero input or output slacks.

The case of *weak* efficiency (or mix inefficiency) is more easily described by referring to Figure 2, where the situation with two inputs (X_1, X_2) and one output (Y) is depicted.

Figure 2. DEA model, two inputs-one output



Source: CRENoS

The efficient frontier is now defined by DMU E, D and C. All others DMUs are inefficient. As explained above the efficiency measure for DMU A, for instance, is given by the ratio 0P/0A, and its benchmarks are DMU E and DMU D. The case of DMU B is different, its efficiency is calculated by 0B’/0B, but note that further gains could be obtained by moving leftwards along the efficient frontier from the weak efficient point B’ to point C. Differently from parametric methods which return smooth frontiers, this may happen because the dotted line crosses the piece-wise frontier in a straight trait, so that the same level of output can be obtained by using a smaller amount of input X_1. The distance CB’ is known as input slack, in this case for the X_1 input. In the DEA literature it is recommended to provide an accurate indication of technical efficiency by reporting both the Farrell measure of efficiency (such as B0/B’0 for DMU B) and any non-zero input and output slack. When there are more inputs and outputs it is not a simple task to identify the nearest efficient point (such as C in Figure 2).

Since the assumption of constant returns to scale is rarely attainable in real-world situations as it requires that each DMU is operating at an optimal scale, in what follows we brief describe the Varying Return to Scale (VRS) model, suggested by Banker et al. (1984). With respect to the CRS model the linear programming problem is augmented with an additional convexity constraint. The VRS approach allows to envelop the data more tightly so that technical efficiency measures are always greater or equal to the ones obtained under the assumption of CRS. The aim is to isolate "pure" technical inefficiency from "scale" inefficiency. Operationally this is done by carrying out both a CRS and VRS DEA, if for a given DMU there is a difference in the technical scores this is interpreted as evidence of scale inefficiency.

.

To investigate the innovative performance of European countries we follow Culliname et al. (2004) and adopt the *output*-oriented approach[4] since the objective of investment in R&D is to increase innovative output so as to improve national competitive position. As a result, this approach is more suitable when the analysis serves as the basis for defining planning and policy strategies, which is commonly the case for geographic units, such as areas, regions or countries. On the contrary, the input orientation is more adequate when operational and managerial objectives are involved.

The application of the DEA to the study of innovation performance is still quite rare. Nonetheless analogous KPF models have been, implemented in some studies which investigate knowledge production at the national level as in Wang (2007), Wang and Huang (2007) and Sharma and Thomas (2008) who have recently followed the pioneering contribution by Rousseau and Rousseau (1998). Other contributions are Schmidt-Ehmcke and Zloczysti (2009), who discriminate knowledge production across sectors and Cullmann et al. (2009) who distinguish the impact of private and public R&D and of different institutional and regulatory frameworks.

Based on the analysis above, in our empirical DEA R&D investments serve as internal inputs, while patent applications are used to approximate innovative output. Population controls for differences in the country dimensions.

We use the Farrell-type output oriented efficiency measure:

$$TE(inn_i, inputs_i) = \max\{\theta : (inn_i, \theta inputs_i) \in \Psi\}$$

θ measures the radial distance between the observation and the efficiency frontier. The efficiency score is the point on the frontier characterized by the level of inputs that can be reached if the DMU is efficient (Simar and Wilson, 1998). A value of θ = 1 indicates that a region is fully

[4] The DEAP software by Coelli (1996) is used throughout the analysis.

efficient and thus is located on the efficiency frontier based on the technology set ψ, which is unobserved and is thus estimated thanks to DEA.

The analysis is based on the same set of information used in the section above on KPF: inputs are represented not only by R&D but also by population as a control which provides a measure of the number of potential inventors in a region. So, our DEA model for evaluating inter-countries innovation efficiency includes one output and two inputs. Since we adopt an output orientation, we are implicitly assuming that countries aim to maximize the innovation output resulting from their inputs. Our approach is based both on the *VRS* assumption and *CRS* assumption. Summary statistics on relative scores are presented in tables below (see Appendix for the complete ranking).

The initial DEA is calculated by using as output EPO patents by inventor in the period from 2000-2008. Results, summarized in Table 9, show that only one country, that is Germany, is technically efficient with CRS assumption. Conversely, with the opposite assumption of variable returns to scale, the efficient countries are six, including some ENC such as Jordan and Malta. These countries represent the points on the frontier close to the origin and act as benchmarks for other similar countries. In addition, the average differs quite considerably: from 0.230 for CRS to 0.349 under VRS.

Table 9. Summary of DEA results, output: EPO patents by inventor, 2000-2008

	CRS efficiency	VRS efficiency	Scale efficiency
Min	0.001	0.001	0.002
Max	1	1	1
Mean	0.230	0.349	0.743
Std. Dev	0.322	0.398	0.347
N. efficient	1	6	1
Number of countries with efficiency score between 0.6 and 1	8	12	30
Number of countries with efficiency score less than 0.6	31	27	9

Obs: 39 countries . No data available for Syria and Palestinian territories
Source: CRENoS calculation

However, the *scale efficiency* confirmed that the only efficient country is Germany and the majority of countries under analysis shows values between 0.6 and 0.99, while only 9 of them present values below 0.6. Seven out of those 9 countries, as we could expect, are ENC from the Eastern or Southern group, while the other two are the small islands of Cyprus and Malta which pertain to the New Member States group.

In order to have a complete overview, Table 10 shows the efficiency scale averages for the three groups of countries (EU15, EU12 and ENC). As we could expect, there are important differences among observed groups. It can be easily seen that EU15 countries are significantly more efficient than the other two groups, both under CRS and VRS. Nonetheless, the results obtained in the KPF econometric analysis is somewhat confirmed: ENCs show an higher efficiency with respect to NMS.

Table 10. Average efficiency scale, output: EPO patents by inventors, 2000-2008

	CRS efficiency	VRS efficiency	Scale efficiency
EU15	0.496	0.526	0.939
NMS12	0.031	0.084	0.690
ENC	0.053	0.306	0.479

Source: CRENoS calculation

Other DEA results are summarized in Table 11 and 12. In fact, in order to test the robustness of our results we carried out our analysis with other output data. In Table 11 we present DEA results using PCT rather than EPO patents as output measure. In this second case two efficient countries emerge, that is Germany and Sweden, while the *scale efficiency* for all other countries is confirmed.

Table 11. Summary of DEA results, output: PCT patents by inventor, 2000-2008

	CRS efficiency	VRS efficiency	Scale efficiency
Min	0.003	0.003	0.010
Max	1	1	1
Mean	0.244	0.362	0.728
Std. Dev	0.333	0.391	0.332
N. efficient	2	6	2
Number of countries with efficiency score between 0.6 and 1	6	10	29
Number of countries with efficiency score less than 0.6	32	28	9

Obs: 38 countries . No data available for Azerbaijan, Syria and Palestinian territories
Source: CRENoS calculation

Table 12 reports the DEA summary considering EPO patents counted by applicant as output measure; the main result is the higher number of efficient countries: Germany, Luxemburg, and Sweden. All other results are unaffected.

Table 12. Summary of DEA results, output EPO patents by applicant, 2000-2008

	CRS efficiency	VRS efficiency	Scale efficiency
Min	0.000	0.001	0.002
Max	1	1	1
Mean	0.230	0.314	0.781
Std. Dev	0.337	0.390	0.319
N. efficient	3	6	3
Number of countries with efficiency score between 0.6 and 1	6	9	31
Number of countries with efficiency score less than 0.6	33	30	8

Obs: 39 countries . No data available for Syria and Palestinian territories
Source: CRENoS calculation

Finally it is interesting to note that DEA allows to consider not only multiple inputs but also multiple outputs. We have therefore performed a further robustness test with two outputs, that is the number of applications both at EPO and at PCT, at the same time. Results do not vary and confirm our main conclusions.

4. Conclusions

The main purpose of this analysis is to investigate the functioning of the knowledge economy at the country level in Europe and in its Neighbouring countries. In particular, we assess, by means of econometric techniques, the impact of intangible assets such as research and development activities on inventive performance measured by patent applications at either EPO or PCT. We also evaluate whether this impact is significantly different among countries in particular with respect to three areas: Old Europe, which includes the fifteen more developed Western economies, the New Europe, with the twelve countries which have entered the European Union in more recent times, and finally European Neighboring Countries. Moreover, we apply Data Envelopment Analysis as a benchmarking methodology which permits to identify the best performing countries within Europe and beyond. Since, in general, the two methods provide different indications on the same object of analysis, we employ both of them in a complementary guise in order to gain wider and different insights for the comparison of European Union and European Neighbouring countries innovative performance.

The analysis is preliminary and exploratory since the dataset is incomplete (no homogenous data on human capital were available) and the territorial unit, countries instead of regions, does not allow to have a sufficiently wide sample. Results, consequently have to be interpreted as mainly speculative.

Nonetheless some outcomes have proved quite robust and are worth being mentioned. As far as the returns of R&D expenditures on country innovative capacity are concerned, estimated by means of a Knowledge Production Function, there is robust evidence of the strong role played by this input in fostering innovation and knowledge creation. This role, however, is much stronger in EU15 countries whilst ENCs are relatively less effective in transforming R&D into patents. This difference in efficiency is further analyzed thanks to non parametric techniques, Data Envelopment Analysis, which allow to identify countries on the technological frontier. DEA shows that there is, again, a dualistic pattern in country innovation activities, with the highest efficient territories located in EU15 and the lowest efficiency scores in European peripheral areas, both new accession countries (NMS12) and European Neighboring Countries (ENC).

5. References

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

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

Coelli T.J. (1996) A Guide to DEAP Version 2.1: A Data Envelopment Analysis Computer Program. *Working Papers 8/96, CEPA*, The University of New England.

Coelli, T., Rao, D.S. P., O' Donnell, C. J., Battese, G. E. (2005), An introduction to efficiency and productivity analysis, 2nd edition, *New York et al.: Springer*.

Cooper W.W., Seiford L.M. and Tone K. (2007) Data Envelopment Analysis, *Springer,* second edition.

Cullinane K., Ji Ping and Wang T.-F. (2005) The relationship between privatization and DEAestimates of efficiency in the container port industry*, Journal of Economics and Business,* 57 433–462

Cullmann A., Schmidt-Ehmcke J. and Zloczysti P. (2009) Innovation, R&D Efficiency and the Impact of the Regulatory Environment: A Two-Stage Semi-Parametric DEA Approach, *Discussion Papers of DIW Berlin 883*, DIW Berlin.

Farrell M.J. (1957) The measurement of productive efficiency, *Journal of the Royal Statistical Society*, Series A, General 120, 253-281.

Foddi M. and Usai S. (2012), Regional Knowledge Performance in Europe, *Growth and Change, forthcoming*

Griliches, Z. (1979) Issues in assessing the contribution of research and development to productivity growth, *The Bell Journal of Economics*, 10, 92-116.

Marrocu E., Paci R. and Usai S. (2012), Productivity growth in the Old and New Europe: the role of agglomeration externalities, *Journal of Regional Science*, *forthcoming*.

Rousseau, S. and Rousseau, R. (1998) The scientific wealth of European nations: Taking Effectiveness into Account, *Scientometrics*, 42(1), 75–87.

Schmidt-Ehmcke J. and Zloczysti P. (2009) Research Efficiency in Manufacturing: An Application of DEA at the Industry Level, *Discussion Papers of DIW Berlin 884*, DIW Berlin.

Sharma, S. and Thomas, V. J. (2008) Inter-country R&D efficiency analysis: an application of data envelopment analysis, *Scientometrics*, 76(3), 483–501.

Simar, L., & Wilson, P. (1998). Sensitivity of efficiency scores: How to bootstrap in nonparametric frontier models. *Management Science*, 44, 49–61.

Wang, E. C. (2007) R&D efficiency and economic performance: a cross-country analysis using the stochastic frontier approach, Journal *of Policy Modelling*, 29(2), 345–360

Wang, E.C. and Huang, W. (2007) Relative efficiency of R&D activities: a cross –country study accounting for environmental factors in the DEA approach, *Research Policy*, 36(2), 260–273.

Appendix

Table A.1 EFFICIENCY SUMMARY (VRS)

	PCT by inventor(s)	PCT by applicant(s)	EPO by inventor(s)	EPO by applicant(s)	EPO by inv. and by appl.	PCT by inventor(s) and by applicant(s)
Algeria	0.047	0.044	0.004	0.002	0.004	0.047
Armenia	0.221	0.174	0.029	0.008	0.004	0.221
Austria	0.473	0.319	0.648	0.496	0.648	0.473
Belarus	0.01	0.008	0.003	0.002	0.003	0.01
Belgium	0.388	0.286	0.496	0.437	0.496	0.388
Bulgaria	0.023	0.02	0.01	0.007	0.009	0.023
Cyprus	0.326	1	0.234	0.668	0.243	1
Czech Republic	0.062	0.04	0.044	0.031	0.044	0.062
Denmark	0.728	0.576	0.751	0.629	0.751	0.728
Egypt	0.064	0.055	0.011	0.006	0.011	0.064
Estonia	0.078	0.046	0.047	0.024	0.047	0.078
Finland	0.992	0.975	0.935	0.962	0.97	0.992
France	0.486	0.478	0.462	0.471	0.471	0.486
Georgia	1	1	1	1	0.225	1
Germany	1	1	1	1	1	1
Greece	0.058	0.048	0.039	0.033	0.038	0.058
Hungary	0.098	0.065	0.052	0.032	0.052	0.098
Ireland	0.309	0.316	0.245	0.316	0.316	0.316
Israel	0.874	0.605	0.615	0.44	0.615	0.874
Italy	0.397	0.35	0.465	0.422	0.465	0.397
Jordan	1	1	1	1	1	1
Latvia	0.046	0.032	0.023	0.017	0.023	0.046
Lithuania	0.018	0.012	0.011	0.007	0.011	0.018
Luxembourg	1	1	1	1	1	1
Malta	1	1	1	1	1	1
Moldova	0.014	0.011	0.003	0.003	0.002	0.014
Morocco	0.011	0.009	0.003	0.002	0.003	0.011
Netherlands	0.964	1	0.8	0.992	0.992	1
Poland	0.043	0.035	0.027	0.02	0.027	0.043
Portugal	0.035	0.028	0.028	0.024	0.028	0.035
Romania	0.014	0.009	0.006	0.003	0.006	0.014
Slovak Republic	0.038	0.029	0.019	0.012	0.019	0.038
Slovenia	0.184	0.107	0.185	0.113	0.185	0.184
Spain	0.17	0.151	0.119	0.102	0.119	0.17
Sweden	1	1	0.956	1	1	1
Tunisia	0.003	0.002	0.001	0.001	0.001	0.003
Ukraine	0.015	0.012	0.003	0.001	0.003	0.015
United Kingdom	0.557	0.471	0.359	0.292	0.359	0.557

AQR UB IREA

Utrecht University

SEED CENTER
SOUTH & EAST EUROPEAN DEVELOPMENT CENTER

CRENoS

LSE

WU WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

Brunel UNIVERSITY LONDON

Université Jean Monnet SAINT-ETIENNE

UNIVERSITAS QUINQUEECCLESIENSIS

Leibniz Universität Hannover

UNIVERSITY of TARTU

STATE UNIVERSITY HIGHER SCHOOL OF ECONOMICS

Cadi Ayyad University

ICBSS INTERNATIONAL CENTRE FOR BLACK SEA STUDIES

IEMed. Institut Europeu de la Mediterrània

האוניברסיטה העברית בירושלים
The Hebrew University of Jerusalem

TÜBİTAK