Deutsches Institut für
Wirtschaftsforschung

DIW BERLIN

Discussion Papers

883

Astrid Cullmann • Jens Schmidt-Ehmcke • Petra Zloczysti

Innovation, R&D Efficiency and the Impact of the Regulatory Environment – A Two Stage Semi-Parametric DEA Approach

Berlin, May 2009

Opinions expressed in this paper are those of the author and do not necessarily reflect views of the institute.

IMPRESSUM

DIW Berlin
German Institute for Economic Research
Mohrenstr. 58
10117 Berlin
Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
http://www.diw.de

ISSN print edition 1433-0210
ISSN electronic edition 1619-4535

Innovation, R&D Efficiency and the Impact of the Regulatory Environment

A Two Stage Semi-Parametric DEA Approach[1]

by

Astrid Cullmann (DIW Berlin)*

Jens Schmidt-Ehmcke (DIW Berlin)**

Petra Zloczysti (Freie Universität Berlin)***

May 2009

* acullmann@diw.de

** jschmidtehmcke@diw.de

*** petra.zloczysti@fu-berlin.de

[1] This paper is a result from a research program on innovation economics at the DIW Berlin. We thank without implying Astrid Cullmann, Irwin Collier, Christian von Hirschhausen, Alexander Kritikos, Michael Meehan and Andreas Stephan for their support and helpful comments.

Abstract

This paper assesses the relative efficiency of knowledge production in the OECD using a nonparametric DEA approach. Resources allocated to R&D are limited and should therefore be used efficiently given the institutional and legal constraints. This paper presents efficiency scores based on an intertemporal frontier estimation for the period 1995 to 2004 and analyzes the impact of the regulatory environment using the single bootstrap procedure suggested by Simar and Wilson (2007). The empirical evidence supports the hypothesis that barriers to entry, aimed at reducing competition, lower research efficiency by attenuating the incentive to innovate and to allocate resources efficiently.

Keywords: R&D efficiency, data envelopment analysis, truncated regression, regulation

JEL Classification: C14, C24, L50, O31, O57

1 Introduction

The notion of a knowledge production function is central to endogenous growth models in which innovation (ideas' productivity growth) is a main driver of sustainable long-term growth (Porter and Stern, 2000). True innovation, in contrast to imitation, becomes even more important for productivity growth when a country approaches the world technology frontier because less room is left for copying. The empirical literature affirms the importance of the level and dynamics of R&D expenditures for economic growth (e.g. Guellec and van Pottelsberghe de la Potterie, 2004). Therefore, the efficient usage of the scarce resources devoted to R&D becomes increasingly important, especially in a globalized world. Countries are exposed to high levels of competition in domestic and foreign markets for innovative products and future technologies. This process forces nations to continuously update their technological capabilities. Countries utilizing their R&D resources inefficiently will be penalized with a growth discount.

Since the resources allocated to the generation of new knowledge are limited, they should be used as efficiently as possible given the local institutional, organizational and legal constraints. Hereby government policies aimed to encourage R&D play a major role in ensuring a sufficient level of R&D spending in the research process. Such policies ensuring a high level of competition by reducing market entry barriers are likely to affect innovation and research efficiency. Among others, Acs and Audretsch (1990) and Geroski (1991) found a positive link between the rates of entry and innovation. Studies by Baldwin and Gorecki (1991) and Geroski (1989) document a productivity enhancing effect of market entry on the industry level and recently Aghion et al. (2009) claim that entry encourages incumbent innovation and productivity growth.

The influence of market entry on research efficiency is twofold: first, high entry rates increase the incentives to innovate and thereby the overall level of research and development expenditures in a country. Market entry is often used as a vehicle for introducing new innovations (Geroski 1995). New innovative firms challenge incumbents that are often more interested in protecting their existing position than in seeking new business opportunities. Incumbents are then forced to increase their R&D investment in order to acquire a lead over their rivals due to a more competitive environment. Thus, more resources are allocated to R&D via growing incentives to innovate. Second, increasing competition by new entries forces firms to improve their R&D process. In competitive markets, firms are punished more severely for being inefficient (Boone, 2008). Competitive pressure induced by entrants increases the incentives to allocate the scarce resources optimally to ensure survival. Thus, high entry rates are associated with higher rates of innovation and increases in efficiency.

In light of this, the degree of governmental regulation plays a crucial role in ensuring low barriers to entry by altering market structures. A strict regulatory environment might hamper the entry of new competitors, like innovative entrepreneurs, and thereby reduce efficiency in the production and research processes. Hence in our empirical analysis, we test the hypothesis that governmental barriers to competition lower research efficiency by distorting the incentive to innovate.

Our model specification follows the "knowledge production function" framework, developed by Griliches (1979) and implemented by Pakes and Griliches (1984), Jaffe (1986), and Hall and Ziedonis (2001). According to Griliches (1979), innovative output is the product of knowledge generating inputs, similar to the production of physical goods. Some observable measures of inputs, such as R&D expenditures and the number of researchers, are invested in the knowledge production process and directed toward producing economically valuable knowledge. The process is seen as a continuum leading from R&D and human capital as inputs to some observable measure of innovative activity. Formally, it can be summarized using a knowledge production function:

$$I_c = f(R\&D_c, R_c)$$

where I is innovative output, R&D denotes R&D expenditures and R is the number of researchers engaged. The unit of observation is the country (c) level.

Innovative output as the result of knowledge production is hard to capture. We argue in favor of patent applications as a measure of valuable output of the knowledge production process. The use of patents as an indicator of innovative output has without a doubt some drawbacks. First of all, patent applications are often criticized as measuring just one component of the innovative output since inventors may choose other protectionist strategies like secrecy. The use of patents would thus underestimate real innovative activity. Second, research has shown that the value of patents is skewed to the right, with only some patents being highly valuable. This observation has been discussed by numerous authors, e.g. Scherer (1965), Pakes and Schankerman (1984), Pakes (1986), and Griliches (1990). Despite this criticism, patents are probably the most important indicator of research output. They are by definition related to inventiveness and based on an objective and relatively stable standard. Furthermore, data on patent application is widely available and provides additional information about the origin of the inventor and a detailed technological classification of the underlying invention. Therefore, patent applications are extensively used in the literature (e.g. Hausman, Hall, and Griliches, 1984 and Kortum, 1997).

The empirical literature using a knowledge production function framework affirms the importance of level and dynamics of research personnel and R&D expenditures as input factors. However, only recently the empirical literature has put more emphasis on the efficient usage of scarce resources. The relevant studies on research efficiency in this field that motivated our approach are summarized in Table 1.

We contribute in the following three aspects to the existing literature: We measure research efficiency in OECD countries and consider R&D expenditures distinguishing between public and private sources on the input side as well as accounting for the possibility of multiple inventors on the output side. In addition, we study the impact of product market regulation on research efficiency by applying a consistent two stage truncated regression approach proposed by Simar and Wilson (2007).

Table 1: Literature Review of R&D efficiency studies

Authors	Data Sets	Methodology	Specification	Key results
Sharma and Thomas, (2008)	UNESCO Institute of Statistics data base, SCI Expanded data base of the web of science, WIPO Statistics data base	DEA approach with constant (CRS) as well as variable returns to scale (VRS).	Inputs: R&D expenditures, researchers, gross domestic product, population Output: patents granted, publications counts	Japan, Republic of Korea, China lie on the efficiency frontier with CRS, Japan, Republic of Korea, China, India, Slovenia and Hungary are found to be efficient with VRS
Wang and Huang, (2007)	WIPO Statistics data, MSTI data base, SCI expanded data base	DEA approach (VRS) and second stage Tobit Regression, Three stage approach according to Fried et al. (1999)	Inputs: R&D net capital stock, researchers, technicians, Output: patents granted, publications counts Environmental Variables: like the enrollment rate of tertiary education, the PC density and the English proficiency	About half of the countries are efficient in their R&D activities, higher education can explain variations in R&D input slacks, increasing returns to scale for two thirds of the countries
Wang, (2007)	WIPO Statistics data, MSTI data base, SCI expanded data base, World development indicators, economic freedom index	Stochastic frontier analysis (SFA), Battese and Coelli (1992, 1995) specification	Inputs: R&D net capital stock, researchers, technicians, Output: patents granted, publications counts Environmental Variables: the PC density, economic freedom index, percentage of R&D performed by the government	External factors affect R&D achievements, PC density and economic freedom index have a significant impact on efficiency differences
Rousseau and Rousseau, (1998)	EPO Patents, Science citation index, UNITED NATIONS, Statistical Yearbook,	DEA approach with CRS, different output and input weights	Inputs: GDP, active population and R&D expenditure Outputs: publications and patents	Switzerland was in 1993 the most efficient and effective country of Europe, closely followed by the Netherlands.
Rousseau and Rousseau, (1997)	EPO Patents, Science citation index, UNITED NATIONS, Statistical Yearbook,	DEA approach with CRS	Inputs : GDP, active population and R&D expenditure Outputs: publications and patents	DEA can be used as a tool to construct performance indicators for governments.

The empirical analysis is conducted in two steps. First, to measure R&D efficiency we follow the nonparametric DEA approach and assume a constant intertemporal frontier. Second, we analyze the influence of product market regulation on the differences in R&D efficiencies on the country level by applying the recently developed single bootstrap procedures proposed by Simar and Wilson (2007). Due to unknown serial correlation among the estimated efficiencies, conventional approaches for drawing inferences are invalid.

The paper is organized as follows: section 2 introduces the methodology of the two stage efficiency analysis is explained while section 3 presents our model specification and the data set. The empirical results of the efficiency analysis and the truncated regression are summarized in section 4. Section 5 recapitulates the findings and concludes.

2 Efficiency Analysis with DEA

To measure the relative R&D efficiency and to provide a ranking of countries with regard to their achieved performance we apply a concept of nonparametric efficiency analysis: data envelopment analysis (DEA)[2]. The DEA approach assumes that decision making units within a sample (of our case countries) have access to the same technology of converting a vector of p inputs $x \in \Re_{+}^{p}$ into a vector of q outputs $y \in \Re_{+}^{q}$. The technology set ψ is then defined as according to Simar and Wilson (2007):

$$\psi = \{(x, y) \in \Re_{+}^{p+q} \,|\, x \text{ can produce } y\}$$

The R&D technology frontier (efficiency frontier) is then defined as the maximum output attainable from each input level (see Coelli et al., 2005) and countries may or may not be on the frontier of this technology. A particular county's distance from the technology frontier may depend on a mixture of different country specific factors. These factors may be exogenous, such as governmental regulatory policies and barriers to entrepreneurship, which in turn affect performance and therefore the distance to the frontier. Thus, the distance from the actual input/output combination to the frontier of the technology set ψ is assumed to correspond to the inefficiency caused by country specific exogenous factors of governmental regulatory policies and some unexplained statistical noise (see Barros and Dieke, 2008). The objective of this paper is to assess in a first stage such inefficiency and then investigate in a second stage its dependency on various indicators of the regulatory environment in each country.

[2] For a survey on the theoretical literature see Cooper et al. (2004).

2.1 Stage 1: Estimation of relative R&D efficiency scores

In the first stage we use the Farrell/Debreu-type output oriented efficiency measure[3]:

$$TE(x^j, y^j) = \max\{\theta : (x^j, \theta y^j) \in \psi\}$$

θ measures the radial distance between the observation x_i, y_i and the efficiency frontier. The efficiency score is the point on the frontier characterized by the level of inputs that should be reached to be efficient (Simar and Wilson, 1998). A value of $\theta = 1$ indicates that a country is fully efficient and thus is located on the efficiency frontier. As in practice the technology set ψ is unobserved and we replace it with its DEA-estimate (see Simar and Wilson, 2007 and Barros and Dieke, 2008). [4]

Calculations can be made using either an input-orientation where the output vector is held fixed and inputs are minimized to be efficient. Contrary to the case of output-orientation the input vector is fixed and outputs are maximized to be efficient. We apply output orientation since it is reasonable to assume that countries aim to optimize and maximize the research output with a given level of R&D expenditures and the number researchers. In the variable returns to scale model, the determination of the efficiency score of the *i*-th firm in a sample of *N* firms is equivalent to the following optimization (see Coelli et al., 2005):

$$\hat{\psi} = \{(x, y) \in \Re_{+}^{p+q} :$$
$$\sum_{k=1}^{n} \gamma_k y_q^k \geq y_q,\ q = 1,...,Q,\ \sum_{k=1}^{n} \gamma_k x_p^k \leq x_p,\ p = 1,...,P; \gamma_k \geq 0;\ \sum \gamma_k = 1,\ k = 1,...,n\}$$

The identified efficient countries could serve as peers to help improve performance of less efficient ones via technology transfer or detailed process analysis.

[3] Farrell (1957) originally proposed estimating production efficiency scores in a nonparametric framework. He drew upon the work on activity analysis by Koopmans (1951) and Debreu (1951). Charnes et al. (1978) and Banker et al. (1984) extended Farrell's ideas by imposing returns to scale properties.

[4] Different assumptions regarding the frontier can be made: the underlying technology determined either by constant returns to scale (CRS), (see Charnes et al., 1978, who first derived the DEA under CRS); or by variable returns to scale (VRS) which assume that scale inefficiencies are present (see Banker et al., 1984, who first allow for VRS). To determine efficiency measures under the variable returns to scale (VRS) assumption, a further convexity constraint ∑λ=1 must be considered. Within this framework countries of similar sizes concerning the input requirements are compared.

The DEA estimator belongs to the deterministic frontier models, which imply that all observations are assumed to be technically attainable. They are highly sensitive to outliers and extreme values in the data (Simar and Wilson, 2000, 2007). It is therefore important to assess ex ante if outliers in the data inappropriately influence the estimation of the performance of other countries in the sample. This paper uses the method of super-efficiency (see Banker and Chang, 2006 and Andersen and Peterson, 1993) to identify and delete extreme values ex-ante. Within the super-efficiency approach, decision-making units within the efficiency frontier might obtain an efficiency score greater than one because the observation itself cannot be used as a peer (see Coelli et al., 2005) and therefore cannot form part of its reference frontier.[5]

2.2 Stage 2: Regulatory environmental indicators as determinants of efficiency?

In addition to the relative R&D performance of OECD countries we assess the impact of regulatory indicators on efficiency differences. This represents an important step when deriving policy implications with regard to a favourable regulatory, competitive and administrative environment while assuring research efficiency. Thus, after the determination of the individual efficiencies in a first stage we regress in a second stage the efficiency scores on the country specific exogenous regulatory indicators provided by the OECD (see section 3).

The econometric model is based on Simar and Wilson (2007) who propose and derive a bootstrap procedure, which permits valid inference in the second-stage truncated regression. They show that conventional approaches for drawing inference in truncated Tobit regressions, which have been widely applied in the past, are invalid when regressing non-parametric DEA scores on environmental variables in the second stage. The inconsistency of simple second stage Tobit regressions is due to complicated, unknown serial

[5] According to Banker and Chang (2006) countries obtaining in a specific point in time efficiency score larger than 1.2 are supposed to be an outlier and therefore deleted from the sample.

correlation among the estimated efficiencies.[6] The econometric model is specified as follows:

$$\hat{TE}_i = Z_i \beta + \varepsilon_i \text{ with } i = 1, \ldots, n$$

where $\hat{TE}_i$ represents the estimated technical average efficiencies on the country level; Z_i a vector of country specific variables, which we expect to have an impact on the technical efficiencies; and β the coefficients to be estimated. Both sides are bounded by unity (see Simar and Wilson, 2007 and Barros and Dieke, 2008), thus ε_i is restricted by the condition $\varepsilon_i \geq 1 - Z_i \beta$. Therefore a truncated normal distribution for ε_i with a left truncation point at $1 - Z_i \beta$ is assumed. The truncated regression model is estimated by means of maximum likelihood. A parametric bootstrap procedure is used to estimate standard errors and confidence intervals for the estimated coefficients (for a detailed description of the estimation algorithm see Simar and Wilson, 2007).

3 Model Specification and Data

The empirical DEA model is specified as follows: based on the notion of a knowledge production function we use R&D expenditures and labor invested in R&D on the input side. Hereby, we distinguish between R&D expenditures conducted by business enterprises[7], by the government[8] and by the higher education sector[9]. This differentiation provides a more detailed picture compared to the conventional use of aggregate R&D[10] because the distribution of R&D expenditures over sources varies remarkably across countries. The importance of public vs. private R&D is country-specific and should therefore be taken into account when measuring research efficiency. Furthermore, the productivity of R&D may vary

[6] They argue that the serial correlation arises due to the fact that perturbations of observations lying on the frontier will often cause changes in efficiencies estimated for other observations. The semi-parametric two-stage model has been used already in other sectors and applications (see e.g. Barros and Dieke, 2008 for an evaluation of airports and Barros and Peypoch, 2007 for a measurement of technical efficiency in thermoelectric power plants).

[7] BERD in R&D terminology of MSTI

[8] GOVERD in R&D terminology of MSTI

[9] HERD in R&D terminology of MSTI

[10] GERD in R&D terminology of MSTI

across sectors-- a dollar invested in private R&D might increase a country's patent output more than a dollar invested in public R&D (see Wang, 2007). The distinction between private and public R&D is especially useful since the question of whether these are complements or substitutes has not yet been satisfactorily answered in the literature (David et al., 2000).

Another ongoing discussion in specifying knowledge production is the distinction between R&D stocks and R&D expenditures (see e.g. Wang and Huang, 2007 using R&D stocks as an input). From a theoretical point of view R&D stocks are preferable since they encompass the stock of knowledge available in an economy. In practice, assumptions need to be made for calculation due to missing data problems. R&D stocks[11] are built using the perpetual inventory method suggested by Guellec and van Pottelsberghe de la Potterie (2001). We tested both approaches by running separate DEA linear programming for each specification and found comparable results. This is not surprising because of high correlation between stocks and expenditures. Hence we follow a pragmatic approach and focus on R&D expenditures.

Data on human capital and R&D expenditures which serve as inputs are taken from the Main Science Technology Indicators published by the OECD. Manpower invested into R&D equals the number of researchers[12] per country. Patents serve as our indicator of inventive output. A number of applications of DEA on research efficiency in the past also suggested the use of scientific publications as an additional output (see Table 1). However, recent studies revealed a number of measurement problems inherent in the publication counts like co-authoring[13] and language bias (Rousseau and Rousseau, 1997) and therefore reject its usage (Sharma and Thomas, 2008).

This study analyzes research efficiency based on a sample of 26 OECD member countries and and two non-member countries (Argentina, China). The European Patent Office's Worldwide

[11] In line with the literature we assume a depreciation rate of 15%.
[12] measured in full time equivalents.
[13] The usage of all-author publication counts tends to overestimate the output of a country due to double counting when authors come from the same country.

Patent Statistical Database (PATSTAT[14]) serves as the base of information on patent applications.[15]

Central to our exercise is the construction of patent aggregates by country and year. We build this variable by using all patent applications filed with the European Patent Office according to their priority date between 1995 and 2004. We focus on EPO applications since an application to an international authority, in contrast to one made to a national authority, can be taken as a signal that the patentee believes the invention to be of high enough value to justify the expense of in international application. The term priority date refers to the date where the given invention was covered by a patent for the first time. However, this first filing of a given invention mainly occurs at the national level and therefore the majority of patent applications at the EPO are second stage filings. Accordingly, in this study we date patent applications using the priority instead of the usual application date because it closest to the date of invention and the decision to apply for a patent protecting the given invention (de Rassenfosse and van Pottelsberghe de la Potterie, 2007).

In the event that the country of the inventor and that of the applicant vary, (as with multinationals) patent applications are assigned to the country of the inventor, which compared to the country of the applicant, is closer to the location of invention. The literature has until now usually considered only the first inventor's country of residence (e.g. Wang 2007, WIPO 2008) and thereby ignores research cooperations across country borders. To overcome this problem, we construct patent aggregates based on all inventors' countries of residence and compare them with the conventional first inventor approach. The aggregation based on multiple inventors is conducted in two different ways:

- First, an unweighted sum over all inventors' countries of residence is calculated. This is by definition at least as large as the sum of all first inventors since patents with more than one inventor count more than once. Therefore, such an aggregation procedure might induce a bias due to double counting.

[14] Version 1/2008

[15] This database, maintained by the European Patent Office, contains all national and international patent applications including inventors, applicants and their location, priority date and technological classification.

- Second, we derive a weighted sum where all patent applications are assigned the reciprocal of the number of inventor countries in the original patent application as weights, meaning that an application with three inventor countries only contributes a third to each country's aggregate. Empirical testing of the correlation between the first inventor and the multiple inventor output measures leads to the conclusion that all can be used as an approxiamtion of inventive output and will behave rather similar in the empirical application. However, in the case of small countries the conventional first inventor approach could lead to an underestimation of patent output when countries engange extensively in cross-border research cooperations. Therefore, we argue in favor of weighted patent aggregates as the appropriate output for the DEA application.

Consistent with recent literature on research efficiency (Sharma and Thomas, 2008 and Wang and Huang, 2007), we impose a lag structure on inputs to account for the fact that R&D efforts do not immediately lead to innovative output (Hall et al., 1986). Therefore, inputs are lagged by two years in the DEA application. The different model specifications summarizing the input-output combinations are provided in Table 2.

Table 2: Model Specifications

Variables	**Model 1**	**Model 2**	**Model 3**
Inputs			
GERD			●
BERD	●	●	
HERD	●	●	
GOVERD	●	●	
Researchers	●	●	●
Outputs			
Weighted Patents	●		●
Unweighted Patents		●	

In the second stage of our empirical analysis we evaluate the impact of barriers to entry caused by regulation on research efficiency. The regulatory environment is captured using the product market regulation indicators provided by the OECD in 1998 and 2003 (Conway et al., 2005). These indicators focus on the regulations which are potentially able to reduce competition in the areas of product markets. Information on regulation is collected on a questionnaire basis aiming at specific policies applied by the government. The information on regulation in coded between 0 and 6 and increases with the restrictiveness of regulation. From this information a product market indicator system is derived based on 16 low-level indicators to cover various policy options. By means of principal component analysis, the low-level indicators are aggregated to sub-domain and domain-levels with the three domains being

- state control (extent of government control over business),
- barriers to trade and investment and
- barriers to entrepreneurship.

In our analysis about the influence of regulation on research efficiency, we focus on the domain barriers to entrepreneurship. In case of research efficiency, the regulations of considerable interest are those that influence the amount of competitive pressure by raising or lowering barriers to entry. A substantial amount of potential competitors are entrepreneurs which are either encouraged or deterrred from the prevalent degree of product market regulation. We find these aspects being reflected best in the barriers to entrepreneurship domain of the indicator (Table 3). In 1998, the countries with the highest level of regulation in this area were France, Italy and Poland while the Czech Republic ranked highest in 2003. Nearly all countries showed some improvement in the regulatory environment between 1998 and 2003.

Table 3: Product Market Regulation: Domain Barriers to Entrepreneurship

Country	1998	2003
Australia	1.4	1.1
Belgium	1.9	1.6
Canada	1	0.8
Czech Republic	2	1.9
Denmark	1.4	1.2
Finland	2.1	1.1
France	2.8	1.6
Germany	2	1.6
Greece	2.1	1.6
Hungary	1.6	1.4
Iceland	1.8	1.6
Ireland	1.2	0.9
Italy	2.7	1.4
Japan	2.4	1.4
Korea	2.5	1.7
Mexico	2.7	2.2
Netherlands	1.9	1.6
New Zealand	1.2	1.2
Norway	1.5	1
Poland	2.8	2.3
Portugal	1.8	1.3
Slovak Republic	-	1.2
Spain	2.3	1.6
Sweden	1.9	1.1
United Kingdom	1.1	0.8
United States	1.5	1.2

The domain indicator barriers to entrepreneurship is a composite indicator and is calculated in two steps: first, the following seven low-level indicators are derived by summarizing the information from the questionnaires:

- Licenses and permit system: reflecting rules for obtaining and issuing licenses and permits (z1),
- Communication and simplification of rules and procedures: reflecting government's communication strategy to reduce administrative burdens (z2),
- Administrative burdens for corporations: depicts administrative burdens on corporation creation (z3),

- Administrative burdens for sole proprietor firms: depicts administrative burdens on sole proprietor firm creation (z4),
- Sector-specific administrative burdens: measures administrative burdens in transport and retail distribution (z5),
- Legal barriers: measures legal limitations on the number of competitors (z6),
- Antitrust exemptions: measures the scope for exceptions to competition law for public enterprises (z7).

Second, these low-level indicators are aggregated by means of principal component analysis to the three sub-domain indicators:

- Regulatory and administrative opacity: z1 and z2,
- Administrative burdens on startups: z3, z4 and z5,
- Barriers to competition: z6 and z7.

Table 4: Product Market Regulation: low-level indicators

Indicator	1998 min	1998 max	1998 mean	2003 min	2003 max	2003 mean
Licenses and permit system	0.0	6.0	3.4	0.0	6.0	2.1
Communication and simplification of rules and procedures	0.3	2.6	1.0	0.0	2.6	0.5
Administrative burdens for corporations	0.5	5.5	2.2	0.8	4.3	1.8
Administrative burdens for sole proprietor firms	0.3	4.3	2.2	0.0	4.0	2.8
Sector-specific administrative burdens	0.0	4.7	1.9	0.3	4.1	1.6
Legal barriers	0.3	3.5	1.8	0.3	2.3	1.5
Antitrust exemptions	0.0	3.7	0.6	0.0	3.5	0.5

The summary statistics for the years 1998 and 2003 of the low-level indicators are given in Table 4. In 1998, product market regulation via the license and permit system played a dominant role while administrative burdens became relatively more important in 2003. Nevertheless, all indicators declined on average during the covered period.

4 Empirical Results

The empirical analysis is divided into two main sections. First the relative R&D efficiency is determined using DEA to identify the OECD countries that perform efficiently with respect to R&D efforts. Based on a ranking we assess countries that could serve as peers to help improve performance of less efficient countries. We estimate an intertemporal frontier, more precisely a cross section pooled frontier, where each observation is accounted for as a single unit without considering any panel structure of the data. Country averages are then calculated over the observation period.

In the second part we assess the impact of regulatory and administrative opacity, administrative burdens and barriers to competition on R&D efficiency by means of the truncated two-stage semi parametric regression proposed by Simar and Wilson (2007).

4.1 Relative R&D efficiency

We assume output orientation, thus countries aim to maximize the R&D output resulting from their inputs. In this context, inputs are exogenous. We estimate both, the constant returns to scale model (CRS, Charnes et al. 1978) and the variable returns to scale model (VRS, Banker et al.). Within the CRS model, technical and scale efficiency are aggregated, whereas the VRS model measures the pure technical efficiency. Scale efficiency can therefore be determined by the difference between the results obtained from both specifications. The scale efficiency indicates if size and magnitude of the research production process in the countries is optimal.

Our sample includes East European countries like Poland, Czech Republic and Slovakia which underwent a transition period after 1989. To leave room for changes towards market-oriented structures, we start our observation period in 1995. To ensure comparability across countries and years, we exclude countries for which less than four years are available from our sample.[16] In total, we end up with 217 observations which are representative for nonparametric estimation of relative efficiency by means of DEA under both (VRS and CRS) assumptions.

The underlying model for nonparametric efficiency analysis has to be robust against outliers and extreme values in the sample. To ensure a consistent and robust technology frontier we conduct ex ante outlier detection by means of super-efficiency analysis. We apply the criterion outlined in Banker and Chang (2006) and define outliers by an efficiency score of larger than 1.2. Only two observations obtain an efficiency score larger than 1.2 and are excluded from further calculations.[17] The small amount of observations revealing an efficiency score above 1.2 indicates that our frontier is not spanned by a number of unrealistic and extreme data points. Therefore, we claim the frontier being robust and consistent for the relative efficiency measurement of the remaining countries within the sample (214 observations).

[16] This is the case for Switzerland, Austria and Luxembourg, which are observed only for one and two years respectively.

[17] The deleted observations are Iceland (1996, 1999) and Slovak Republic (1996). Due to significantly lower efficiencies in the rest of the time period we assume data problems for both countries in these years.

We test three model specifications as outlined in section 3 (Table 2). The difference between model 1 and model 2 is the weighting scheme applied when deriving the patent aggregates. Model 1 uses weights for multiple inventors while model 2 involves double counting. As expected the results are highly similar due to strong correlation and a rank correlation of about 0.97.

Table 5: Results for different model specifications (VRS)

Model 1		Model 2		Model 3	
Sweden	0.976	Sweden	0.982	Germany	0.945
Germany	0.966	Germany	0.957	United States	0.874
United States	0.874	United States	0.883	Netherlands	0.699
Belgium	0.854	Iceland	0.874	Finland	0.606
Netherlands	0.780	Belgium	0.870	Iceland	0.565
Finland	0.692	Netherlands	0.685	Japan	0.557
New Zealand	0.685	Ireland	0.679	Italy	0.540
Iceland	0.658	New Zealand	0.632	Belgium	0.487
Italy	0.650	Finland	0.620	Denmark	0.483
Ireland	0.573	Slovak Republic	0.613	Sweden	0.464
Denmark	0.565	Japan	0.608	France	0.373
Japan	0.557	Hungary	0.541	United Kingdom	0.331
Slovak Republic	0.556	Italy	0.509	Ireland	0.320
France	0.400	Denmark	0.497	New Zealand	0.314
United Kingdom	0.379	France	0.350	Norway	0.248
Hungary	0.339	United Kingdom	0.337	Hungary	0.209
Norway	0.289	Korea	0.288	Spain	0.196
Greece	0.274	Norway	0.248	Australia	0.169
Spain	0.260	Spain	0.233	Canada	0.167
Korea	0.259	Greece	0.211	Korea	0.156
Australia	0.238	Canada	0.207	Greece	0.119
Canada	0.202	Australia	0.205	Slovak Republic	0.089
Portugal	0.174	Portugal	0.144	Czech Republic	0.079
Argentina	0.145	Czech Republic	0.132	Portugal	0.063
Czech Republic	0.130	Argentina	0.127	Argentina	0.058
Poland	0.089	Poland	0.103	Poland	0.042
Mexico	0.069	Mexico	0.068	China	0.026
China	0.046	China	0.046	Mexico	0.023

The ranking of the countries only changes slightly in the midfield (see for instance Italy and Ireland) which could be caused by the different degree of engagement in cross country research projects and country size.

In Model 3 we use aggregated R&D expenditures as inputs instead of R&D expenditures by source. Compared to model 1 we find a somewhat lower rank correlation (0.90) and slight changes in the ranking with the main difference being Sweden losing its top position.[18]

Table 6: Efficiency scores for model 1 according to different approaches (CRS, VRS, scale efficiency)

Country	Average Efficiency CRS	Average Efficiency VRS	Average Scale efficiency	Returns to scale[19]
Argentina	0.139	0.145	0.958	irs
Australia	0.237	0.238	0.996	irs
Belgium	0.839	0.854	0.982	irs
Canada	0.201	0.202	0.995	irs
China	0.046	0.046	0.994	irs
Czech Republic	0.114	0.130	0.878	irs
Denmark	0.552	0.565	0.977	irs
Finland	0.671	0.692	0.969	irs
France	0.400	0.400	1.000	crs
Germany	0.965	0.966	0.999	crs
Greece	0.258	0.274	0.943	irs
Hungary	0.324	0.339	0.957	irs
Iceland	0.369	0.658	0.561	irs
Ireland	0.441	0.573	0.770	irs
Italy	0.649	0.650	0.998	irs
Japan	0.431	0.557	0.774	drs
Korea	0.257	0.259	0.991	irs
Mexico	0.067	0.069	0.973	irs
Netherlands	0.777	0.780	0.996	irs
New Zealand	0.640	0.685	0.935	irs
Norway	0.285	0.289	0.989	irs
Poland	0.087	0.089	0.978	irs
Portugal	0.163	0.174	0.936	irs
Slovak Republic	0.165	0.556	0.296	irs
Spain	0.259	0.260	0.996	irs
Sweden	0.960	0.976	0.983	drs
United Kingdom	0.375	0.379	0.989	crs
United States	0.280	0.874	0.320	drs
Mean	0.391	0.453	0.898	
Median	0.305	0.389	0.978	
Standard deviation	0.268	0.286	0.192	

[18] Sweden is in particular efficient with respect to government expenditures on R&D. Aggregating over R&D by source eliminates the unique features with respect to different sources, thereby reducing Sweden's efficiency.

[19] returns to scale are calculated for each observation at each point in time; Exhibiting a property more than five times is our criterion for determining country-specific returns to scales.

We argue in favor of model 1 since we believe that disaggregating the inputs provides a more detailed picture of the research process in countries and therefore adds useful information to our analysis. Furthermore as it is known from the literature from author publication counts, double counting of outputs overestimates efficiency. Hence, we prefer Model 1 to Model 2. The relative R&D and scale efficiency scores of our benchmark model 1 are provided in Table 6.

The difference between the CRS and VRS scores indicates scale efficiency. Table 6 shows that the majority of countries are not characterized by an optimal size of the research production process with respect to input allocation. Only Germany, France and the United Kingdom feature constant returns to scale while Sweden, the United States and Japan show decreasing returns to scale.

The intertemporal frontier estimation exhibits an average technical efficiency of 0.39 in the CRS specification and 0.45 in the VRS specification. This is relatively low compared to other empirical work. It indicates that large inefficiencies are present within the knowledge production process. The low mean efficiency might also be explained by the fact that the sample includes low innovation intensive countries like China or Korea from 1995 onwards. As shown below, these countries only started recently to adapt their R&D expenditures to increase patent output. Furthermore, the intertemporal frontier is defined by the latest years in our sample, indicating that technological progress took place over time. Hence, it is not surprising that covering a larger time span lowers mean efficiency.

We calculate the mean annual efficiency from 1995 to 2004 by averaging over the individual efficiency scores of the countries per year. Implicitly we make the assumption of a constant intertemporal frontier and thereby consider the relative changes of the countries' positions towards the estimated DEA technology frontier. This is motivated by two aspects: first, we face a small annual sample size (of less than 30 observations) which makes it difficult to obtain robust and meaningful results. Second, we do not have a balanced panel data set, which prevents us from comparisons of different frontiers for different years, by means of e.g. Malmquist Indices (see Coelli, 2005).

Germany and Sweden are the most efficient OECD countries in providing R&D research output, followed by the United States and smaller countries like Belgium, the Netherlands and Finland. These countries could serve as peers to help improve performance of the least efficient countries. Compared to other European regions, most Scandinavian countries are located among the top third of the performance ranking. In the case of the United States the high performance is remarkable since European Patent Data are used which usually lead to a home bias that would benefit European countries. Therefore we find the United States is one of the leading and most efficient countries in research and development worldwide. In light of this estimation bias, the position of Japan is also worth mentioning since its performance is above average and it is - as expected - the leading Asian country. This is probably due to their leading role in communication and electronics as well as in the research intensive pharmaceutical industry.

The innovative capacity of advanced industrial countries is their most important source of prosperity and growth. Overall, our results suggest that a matured economic system leads to higher research efficiency compared to countries still developing their industry and technology pattern. Therefore, it is not surprising that the red lantern goes to Poland, Mexico and China which are characterized by a very low capacity of knowledge production, suggesting that they are still in the phase of imitating and replicating existing technologies, while only little effort is made on innovating at the world technology frontier.

4.2 The impact of regulatory environmental factors

In the second part of our empirical analysis, we test the influence of the regulatory environment on research efficiency according to the semi-parametric two stage approach suggested by Simar and Wilson (2007). We argue that regulation reduces competition by raising barriers to entry and thereby lowering competitive pressure and the incentives to innovate efficiently.

Our econometric model is specified as follows: we begin by regressing output oriented VRS efficiency scores obtained in the first stage on the sub-domain level (regulatory and administrative opacity, administrative burdens on startups and barriers to competition).

$$\hat{TE}_i = \beta_0 + \beta_1(w_1 z_1 + w_2 z_2) + \beta_2(w_3 z_3 + w_4 z_4 + w_5 z_5) + \beta_3(w_6 z_6 + w_7 z_7) + \varepsilon_i$$

$\hat{TE}_i$ represents the Farrell output efficiency scores, ranging from 1 to infinity with a value of 1 revealing full efficiency. Hence, a positive beta-coefficient indicates an efficiency loss caused by the corresponding variable.

Since the sub-domain level indicators are obtained by aggregating over the low level information, we further test for specific influence of the low-level indicators (licenses and permits system, communication and simplification of rules and procedures, administrative burdens for corporation, administrative burdens for sole proprietor firms, sector specific administrative burdens, legal barriers and antitrust exemptions).

$$\hat{TE}_i = \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \beta_3 z_3 + \beta_4 z_4 + \beta_5 z_5 + \beta_6 z_6 + \beta_7 z_7 + \varepsilon_i$$

In a third step we conduct a robustness check, identify the statistically significant disaggregated indicator from the previous estimation and test their influence in a separate estimation.

$$\hat{TE}_i = \beta_0 + \beta_2 z_2 + \beta_5 z_5 + \beta_7 z_7 + \varepsilon_i$$

Our estimation results are provided in Table 7. We find that the aggregated sub-domain indicators do not have a significant impact on research efficiency as can be seen from the bootstrapped confidence intervals. However, this cannot be interpreted as regulation being irrelevant for innovation since these indicators encompass various aspects providing an average image of barriers to entry. To obtain a more detailed picture regarding the different components, the effects of the aggregated indicators are disentangled by assessing their influences separately. Our estimation results suggest that three low-level indicators, namely communication and simplification of rules and procedures, sector specific administrative burdens have a significant positive impact on efficiency scores as shown by the bootstrapped confidence intervals. A positive impact implies that lowering the degree of regulation in these specific areas lowers barriers to entry and thus significantly increases research efficiency.

The low-level indicator on communication and simplification of rules and procedures can be interpreted as summarizing stumbling blocks related to the collection of information on start-up requirements, the enforcement of regulation and the treatment of administrative burdens. Therefore, less regulation in this field suggests an emphasis by the government on activities that facilitate innovation and entrepreneurship. This could be interpreted as a relevant factor stimulating competition by encouraging potential entrants to start a business.

In case of sector specific burdens, our results suggest that specific burdens being levied on the sector-level reduce research efficiency significantly. This result is probably mainly driven by country-specific heterogeneity since it depends on the economic importance and size of the sectors being regulated in an economy. Therefore, it implies that competitive barriers may play a larger role in specific sectors of the economy.

The third low-level indicator exhibiting a significant impact in our study covers antitrust exemptions for public enterprises. This is not surprising since the incentive of public enterprises to strengthen their position by innovation is reduced when they are protected by governmental regulations. Hence, antitrust exemptions are accompanied by lower research efficiency since there is less pressure on companies to innovate and patent efficiently.

The robustness check which evaluates solely the significant low-level indicators corroborates our findings from the previous estimations with slightly larger point estimates and confidence intervals.[20]

Table 7: Estimation results

[20] Due to large confidence intervals caused by the parametric bootstrap procedure we limit ourselves to interpreting the direction of the influence instead of the size of the point estimate.

The PMR Indicators	Variable	Lower bound	Estimated Coefficient	Upper bound
	weighted sum[21]:			
Regulatory and administrative opacity	z1+z2	-7.350	2.484	8.944
Administrative burdens on startups	z3+z4+z5	-18.018	15.152	29.311
Barriers to competition	z6+z7	-9.878	3.577	18.669
Licences and permits system	z1	-2.396	-0.558	1.049
Communication and simplification of rules and procedures	z2	1.986	**8.446***	16.319
Administrative burdens for corporation	z3	-1.071	4.426	12.107
Administrative burdens for sole proprietor firms	z4	-12.756	-5.734	1.485
Sector specific administrative burdens	z5	1.211	**7.526***	15.893
Legal barriers	z6	-7.803	-3.193	2.821
Antitrust exemptions	z7	4.930	**8.494***	15.011
Communication and simplification of rules and procedures	z2	2.684	13.201	24.232
Sector specific administrative burdens	z5	2.102	11.204	19.078
Antitrust exemptions	z7	0.865	11.078	20.933
All estimation with constant, * significant at 10% level				

Overall, our results can be summarized as follows: the decision of potential entrants to start a business depends considerably on their regulatory environment. A highly regulated product market might dissuade people from entering which reduces competition and thereby the incentive to innovate and allocate the resources devoted to R&D efficiently.

5 Conclusions

This paper assesses the relative efficiency of public and private research expenditures in the OECD using nonparametric efficiency analysis approaches, a data envelopment analysis (DEA) technique. In times of globalization the efficient usage of the scarce resources a country invests in R&D becomes increasingly important. Therefore, this paper sheds light on

[21] The weights are taken from Conway et al. (2005) and are derived from principal component analysis.

the research efficiency differences among OECD countries and its relationship to a country's regulatory environment.

The empirical analysis is conducted in two steps: in the first stage, an intertemporal knowledge production frontier is estimated. Our results suggest that Sweden, Germany and the United States belong to the best performing countries, located on or close to the world technology frontier. These countries could serve as peers to improve efficiency for less efficient ones. The innovative capacity of advanced industrial countries is their most important source of prosperity and growth. Thus, our results confirm the idea that a mature economic system leads to higher research efficiency compared to countries still developing their industry and technology pattern. The red lantern in case of research efficiency goes to Mexico and China which are characterized by a very low rate of knowledge production, suggesting that they are still in the phase of imitating and replicating existing technologies, while only little effort is made to innovate at the world technology frontier.

Government policies aimed at encouraging R&D play a major role in ensuring a sufficient level of R&D spending. We hypothesize that regulation reduces competition by raising barriers to entry, thereby lowering competitive pressure and the incentives to innovate efficiently. In the second stage of the analysis we assess the impact of the regulatory environment on research efficiency, using the recently developed single bootstrap procedures developed by Simar and Wilson (2007). The regulatory environment is described using the indicator of product market regulation provided by the OECD.

Our estimation results show that the low-level indicators on communication and simplification of rules and procedures, antitrust exemptions and sector specific burdens have a significant impact, suggesting that larger degrees of regulation in these fields lowers research efficiency. Overall, our results confirm our hypothesis that high regulation in product markets dissuades potential entrants, especially entrepreneurs, by imposing barriers to entry, thereby reduces the competitive pressure for existing firms, and thus lowers research efficiency in the economy.

6 References

Acs, Z. and Audretsch, D. (1990) *Innovation and small firms*, MIT Press, Boston.

Aghion, P., Blundell, R., Griffith, R., Howitt, P. and Prantl, S. (2009) The effects of entry on incumbent innovation and productivity, *Review of Economics and Statistics*, 91(1), 20–32.

Aghion, P., Harris, C., Howitt, P. and Vickers, J. (2001) Competition, imitation and growth with step-by-step innovation, *Review of Economic Studies*, **68**(3), 467–492.

Andersen, P. and Petersen, N. (1993) A procedure for ranking efficient units in data envelopment analysis, *Management Science*, **39**(10), 1261–1264.

Baldwin, J. and Gorecki, P. (1991) Firm entry and exit in Canadian manufacturing sector, *Canadian Journal of Economics*, **24**(2), 200–323.

Banker, R. D. and Chang, H. (2006) The super-efficiency procedure for outlier identification, not for ranking efficient units, *European Journal of Operational Research*, **175**(2), 1311–1321.

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

Barros, C. P. and Dieke, P. U. C. (2008) Measuring the economic efficiency of airports: Simar – Wilson methodology analysis, *Transportation Research Part E: Logistics and Transportation*, **44**(6), 1039–1051.

Barros, C.P. and Peypoch, N. (2008) Technical efficiency of thermoelectric power plants, Energy Economics, **30**(6), 3118–3127.

Battese, G. E. and Coelli, T. J. (1992) Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India, *Journal of Productivity Analysis*, **3**(1 – 2), 153–169.

Battese, G. E. and Coelli, T. J. (1995) A model for technical inefficiency effects in a stochastic frontier production function for panel data, *Empirical Economics*, **20**(2), 325–332.

Boone, J. (2008) A new way to measure competition, *Economic Journal*, **118**(531), 1245–1261.

Charnes, A., Cooper, W. W. and Rhodes, E. (1978) Measuring the inefficiency of decision making units', *European Journal of Operational Research*, **2**(6), 429–444.

Coelli, T., Rao, P., O'Donnell, C. J. and Battese, G. E. (2005) *An introduction to efficiency and productivity analysis*, Springer, New York.

Conway, P., Janod, V. and Nicoletti, G. (2005) Product market regulation in OECD countries: 1998 to 2003. OECD Economics Department Working Papers, No 419

Cooper, W., Seiford, L. M. and Zhu, J. (2004) *Handbook on data envelopment analysis*, Kluwer Academic Publishers, Boston.

David, P. A., Hall, B. and Toole, A. A. (2000) Is public R&D a complement or substitute for

private R&D? A review of the econometric evidence, *Research Policy*, **29**(4-5). 497–529.

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

De Rassenfosse, G. and van Pottelsberghe de la Potterie, B. (2007) Per un pugno di dollari: A first look at the price elasticity of patents, *Oxford Review of Economic Policy*, **23**(4), 588–604.

Farrell, M. (1957) The measurement of productive efficiency, *Journal of the Royal Statistical Society*, **120**(3), 253–281.

Fried H. O, Schmidt S. S. and Yaisawarng S. (1999) Incorporating the operating environment into a nonparametric measure of technical efficiency, *Journal of Productivity Analysis*, **12**(3), 249–267.

Geroski, P. A. (1989) Entry, innovation and productivity growth, *Review of Economics and Statistics*, **71**(4), 527–578.

Geroski, P. A. (1991) Entry and the rate of innovation, *Economics of Innovation and New Technology*, **1**(1), 203–214.

Geroski, P. A. (1995) What do we know about entry?, *International Journal of Industrial Organization*, **13**(4), 421–440.

Griliches, Z. (1979) Issues in assessing the contribution of R&D to productivity growth, *Bell Journal of Economics*, **10**(1). 92–116.

Griliches. Z. (1990) Patent statistics as economic indicators: a survey, *Journal of Economic Literature*, **28**(4), 1661–1707.

Guellec, D. and van Pottelsberghe de la Potterie, B. (2001) R&D and productivity growth, Working Paper No. 2001/3, OECD, Paris.

Guellec, D. and van Pottelsberghe de la Potterie, B. (2004) From R&D to productivity growth: Do the institutional settings and the source of funds of R&D matter?, *Oxford Bulletin of Economics and Statistics*, **66**(3), 353–378.

Hall, B. H., Griliches, Z. and Hausman, J. A. (1986) Patents and R&D: Is there a lag?, *International Economic Review*, **27**(2). 265–283.

Hall, B. H. and Ziedonis, R. H. (2001) The patent paradox revisited: an empirical study of patenting in the U.S. semiconductor industry, 1979–1995, *RAND Journal of Economics*, **32**(1), 101–128.

Hausman, J., Hall, B. and Griliches, Z. (1984) Econometric models for count data with an application to the patents-R&D relationship, *Econometrica*, **52**(4), 909–983.

Jaffe, A. (1986) Technology opportunity and spillovers of R&D: Evidence from firms' patents, profits, and market value, *American Economic Review*, **76**(5), 984–1001.

Koopmans, T. C. (1951) An analysis of production as an efficient combination of activities, in T. C. Koopmans (ed), *Activity analysis of production and allocation*, Wiley, New York.

Kortum, J. (1997) Research, patenting, and technological change, *Econometrica*, **65**(6), 1389–1419.

OECD (2008a) *Main Science and Technology Indicators*, OECD, Paris.

OECD (2008b) *Compendium of Patent Statistics*, OECD, Paris.

Pakes, A. (1986) Patents as options: some estimates of the value of holding European patent stocks, *Econometrica*, **54**(4), 755–784.

Pakes, A. and Griliches, Z. (1984) Patents and R&D at the firm level: A first look, in Z. Griliches (ed): *R&D, Patents and Productivity*, University of Chicago Press, Chicago and London.

Pakes, A. and Schankerman, M. (1984) The rate of obsolescence of patents, research gestation lags, and the private rate of return to research resources, in Z. Griliches (ed), *R&D, Patents and Productivity*, University of Chicago Press, Chicago and London.

Porter, M.E. and Stern, S. (2000) Measuring the „ideas“ production function: evidence from international patent output, Working Paper Series No. 7891, NBER, Cambridge, MA.

Rousseau, S. and Rousseau, R. (1997.) Data analysis as a tool for constructing scientometric indicators, *Scientometrics*, **40**(1), 45–46.

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

Scherer, F.M. (1965) Firm size, market structure, opportunity, and the output of patented inventions, *American Economic Review*, **55**(5), 1097–1125.

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. and Wilson, P. (1998) Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models, *Management Science*, **44**(1), 49–61.

Simar, L. and Wilson, P. (2000) A general methodology for bootstrapping in nonparametric frontier models, *Journal of Applied Statistics*, **27**(6), 779–802.

Simar, L. and Wilson, P. (2007) Estimation and inferences in two-stage, semi-parametric models of production process, *Journal of Econometrics*, **136**(1), 31–64.

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.

WIPO (2008) World patent report – a statistical review, WIPO, Geneva.