The innovation efficiency of German regions – a shared-input DEA approach

08.13

Tom Broekel, Nicky Rogge and Thomas Brenner

Philipps Universität Marburg

Impressum:

Working Papers on Innovation and Space
Philipps-Universität Marburg

Herausgeber:

Prof. Dr. Dr. Thomas Brenner
Deutschhausstraße 10
35032 Marburg
E-Mail: thomas.brenner@staff.uni-marburg.de

Erschienen: 2013

The innovation efficiency of German regions – a shared-input DEA approach

Tom Broekel[1]

Institute of Economic and Cultural Geography, Leibnitz University, Hannover.

Nicky Rogge

Faculty of Business and Economics Naamsestraat 69, Katholieke Universiteit Leuven.

Thomas Brenner

Economic Geography and Location Research, Philipps-University, Marburg.

Abstract:

The paper contributes to the debate on how to measure regions' innovation performance. On the basis of the concept of regional innovation efficiency, we propose a new measure that eases the issue of choosing between industry-specific or global measures. We argue for the use of a robust shared-input DEA-model to estimate regions' innovation efficiency in a global manner, while it can be disaggregated into industry-specific innovation efficiency measures. The latter is particularly useful when relating the innovative output to the R&D input involves the use of blurry matching procedures.

We illustrate the use of the method by investigating the innovation efficiency as well as its change in time of German labor market regions. It is shown that the method treats regions that have industry structures skewed towards industries with high and low innovation intensities more fairly than traditional approaches.

Keywords: regional innovation efficiency, shared-input DEA, nonparametric efficiency analysis, regional innovation.

JEL Classifications: R12, O18, O31

[1] Corresponding Author: Tom Broekel, Institute of Economic and Cultural Geography, Leibnitz University of Hannover, Schneiderberg 50, 30167 Hannover, Germany. E-Mail: broekel@wigeo.uni-hannover.de.

1 Introduction

The innovation performance of spatial units, e.g. regions, is frequently measured quantitatively and empirically (Jaffe, 1989; Audretsch 1998; Autant-Bernard & LeSage, 2011). Simply stated, the absolute count of innovations generated by regional organizations within a certain time period can be used as an indicator of regions' innovative success. While characterized by severe uncertainty, innovation generation processes involve the utilization of valuable resources such as human and financial capital. Hence, from an economic standpoint it is commonly more interesting to evaluate innovation success in light of the invested resources (cf. Chen and Guan, 2010; Fritsch and Slavtchev, 2011; Brenner and Broekel, 2012).

A raft of empirical approaches has been put forward aiming at capturing this relation between invested resources and innovative outcomes. Popular approaches are, for instance, patents per capita (Audretsch, 1998) or patent intensity, e.g. the number of patents per employee (Deyle & Grupp, 2005). Fritsch (2000) refines these approaches arguing in favor of measuring *regional innovation efficiency* whereby regions are compared with respect to their organizations' abilities to transform knowledge input factors into innovative output.

While the efficiency approach can be seen as a logical extension of the widely accepted *knowledge production function* approach by Griliches (1979), Brenner & Broekel (2011) argue that the empirical estimation of meaningful regional innovation efficiency measures is far from being easy. Amongst other things, there is the choice between global innovation efficiency measures (a measure reflecting the innovation efficiency of regions) and industry-specific approaches (a measure considering only one specific industry in a region). While both measures have their merits and problems, the first is more desirable in many instances. By way of contrast, the second might be more informative for industry representatives.

The paper has three objectives. First, it discusses differences between global and industry-specific innovation efficiency measures. Second, we propose the robust shared-input DEA-model as a new method to construct a global regional innovation efficiency measure that explicitly takes into account inter-regional variations in regions' industrial structure. An additional feature of this model is the computation of industry-specific innovation efficiency measures within the assessed regions. Third, we apply this method to the 150 German labor market regions for which we estimate innovation efficiency measures for multiple years.

The paper is structured as follows. The subsequent section discusses the empirical measurement of regional innovation efficiency with a particular focus on global and industry-specific approaches. The third section presents empirical data on German labor market regions, which represents data commonly used in such approaches. Section four introduces the robust shared-input DEA-model as a new method to estimate regions' innovation efficiency. The empirical results of its illustrative application to German labor market regions are presented and discussed in section five. The sixth section concludes the paper.

2 Towards regional innovation efficiency

2.1 The production allegory

Regional innovation performance has been investigated in the literature for some time (Jaffe, 1989; Audretsch, 1998; Autant-Bernard & LeSage, 2011). Nevertheless, there is no common definition of what regional innovation performance means and how it should be measured empirically. Recently, a number of papers have systematically discussed this issue (Carlsson, et al. 2002; Bonaccorsi & Daraio, 2006; Zabala-Iturriagagoitia et al., 2007; Brenner & Broekel, 2011). In particular, Brenner & Broekel (2011) discuss a number of approaches to the measurement of regional innovation performance from a theoretical perspective. The present paper particularly builds upon this work.

Despite the many different measures used in the literature, conceptualizing regional innovation performance as regional innovation efficiency has gained in popularity in recent years (Broekel, 2012; Chen and Guan, 2011; Fritsch, 2003; Fritsch & Slavtchev, 2011; Fritsch, 2000). This approach originates from production theory and implies that performance is defined as the achievements (output) in comparison to the involved costs (input).

While this idea is unchallenged in a production context it is not so easily translated to (regional) innovation processes. On the one hand this is because maximizing innovation (while ignorant of the costs) can be desirable from a normative perspective when speeding up technological progress is perceived as something positive. On the other hand, it is less clear what constitute inputs, throughputs, and outputs in innovation processes (see Brenner & Broekel, 2011 for a discussion).

Nevertheless, it is well accepted that innovations do not "*fall from heaven*". On the contrary, creative actors and a wide range of resources are needed for their creation (Nelson, 1959). Accordingly, there is agreement on the allegory of production processes being applicable in this context. This allegory became famous with the knowledge production function approach by Griliches (1979), which became particularly popular in economic geography with the work of Jaffe (1989). Nevertheless, it is essential to point out that this allegory has its limitations, which are pointed out by Brenner & Broekel (2011): "*there is at least one fundamental difference between innovation and production processes: innovation processes are by their very nature non-deterministic while production processes are largely deterministic*" (p. 11). To highlight this difference we follow these authors and rather speak of *input factors* instead of inputs and *innovative output* instead of output. In the remainder of the paper we assume this allegory to be appropriate.[1]

Setting innovative output and input factors into a relation on the regional level implies that both are known and can be meaningfully measured in the context of regions. Again, a wide range of approaches and definitions is applied in the literature. The variation on the output side thereby is relatively small as data availability leaves patents and patent applications as the dominant approximation of the innovative output.[2] The input side is a different story. A wide range of input factors have been considered, including the number of inhabitants (Greif & Schmiedl, 2002; Greif et al., 2006; Stern et al., 2002), the number of regional employees (Deyle & Grupp, 2005), the number of R&D employees (Fritsch, 2003), R&D employees in combination with the level of highly qualified employees in a region (Broekel, 2012), and a wide set of regional factors including R&D employees (Chen & Guan, 2011).

While there's no consensus about which of these input factor set definitions is the most appropriate (Brenner and Broekel, 2011), using the number of R&D employees as input factor has become the most frequent approach when such data is available. The rationale is that R&D employees provide the most accurate approximation (given the availability of data) to the true level of resources invested by organizations in innovation processes, which clearly are the most important driver of innovation processes.

For this reason we follow this approach labeled by Brenner & Broekel (2011) as the "R&D employees' innovation efficiency" approach. ``*In such an approach* [regional innovation efficiency], *we define the innovation performance of a spatial unit by the contribution of this unit to the innovation efficiency of the innovation generators present in the unit. Empirically we would have to measure the number of innovation generators, mainly*

[1] Brenner & Broekel (2011) discuss a number of additional differences, which we refrain from discussing to economize on space.
[2] However, there are also alternative measures based on the community innovation survey.

the R&D employees or activities in firms …, and relate this to the innovation output" (Brenner & Broekel, 2011, p. 24-25).[3]

The resulting measure can ``*be regarded as an indication of the quality, particularly the efficiency and workability of the … regional, or industry-specific innovation system''* (Fritsch, 2003, p. 85). Of course, regions are not innovative - the R&D generators located within a region are the creative actors. Their aggregated productivity/efficiency constitutes a region's innovation efficiency. This efficiency may be impacted by a wide range of regional factors such as urbanization economies, knowledge spillover, the presence of universities, regional cooperation intensity, etc. (Fritsch & Slavtchev, 2011; Broekel, 2012).

2.2 Global vs. industry-specific measures

Straightforward as the estimation of R&D employees' innovation efficiency approach appears to be, there are still different ways in which this efficiency can be measured. Most importantly, this concerns the choice of the measure's industrial dimension: shall it cover the entire regional economy (**global approach**) or is it to be estimated with respect to a specific sector/industry (**industry-specific approach**)? This is not just a theoretical question but also matters in terms of empirical results and potential political conclusions.

Differences between the two approaches are caused by industries varying to a considerable extent in terms of their innovation intensity (Pavitt, 1984; Arundel & Kabla, 1998; Malerba et al., 2000). Industry-specific regional innovation efficiency measures take this explicitly into account by establishing the relation between input factors and innovative output separately for each industry (see Broekel, 2012 for such an approach). Since inter-industrial variations in innovation activities also tend to correlate with the reliability of common approximations for innovations (e.g. patents), industry-specific measures will yield more precise and informative results than approaches that do not model such differences explicitly. For example, a region might show low regional innovation efficiency when estimated in a global approach only because the structure of its industry is dominated by industries with comparatively low innovation efficiencies.

However, industry-specific measures have their flaws as well. First of all, they usually require a higher quality of data, in particular with respect to the matching of innovative output to considered input factors that are frequently differently organized (e.g., NACE for employment and IPC for patents).

[3]A similar definition is put forward by Fritsch (2000) on page 415.

Secondly, the researcher has to decide about what the appropriate level of industrial aggregation is, i.e. how can industries be defined in a meaningful way? The decision usually involves many trade-offs implying that this question can seldom be answered in a completely satisfactory way.

Thirdly, estimating innovation efficiency in an industry-specific manner naturally results in an innovation efficiency index that applies only to one industry. However, interest will commonly be focused upon the performance of the entire region. Accordingly, to get a picture of an entire region's situation when using industry-specific measures one has to look at multiple measures, with the number of measures being determined by the number of industries separately investigated. To evaluate the global innovation efficiency of a region, these would need to be aggregated into a single index, which again involves trade-offs and information losses.

Despite these theoretical and practical differences, it is often data availability that determines what measure is being used. This implies that if industry-specific data are available, an industry-specific approach is chosen because of its higher scientific precision (cf. Brenner & Broekel, 2011).

This paper aims at presenting a methodology that provides a convenient and scientifically sound way of estimating a global regional innovation efficiency measure and minimizes the potential bias induced by variations in a region's industrial structure. At the same time, it is decomposable into industry-specific measures that are little influenced by matching procedures between input factors and innovative output data. Accordingly, this methodology combines the advantages of both approaches.

2.3 An empirical challenge and shared-inputs

When constructing industry-specific innovation efficiency measures, the industry-specific input factors are matched to the industry-specific innovative output. The level to which the quantities of input factors dedicated to the generation of a particular innovation are unknown determines the extent to which this is problematic. Frequently, even firms themselves do not exactly know how many R&D resources, e.g. laboratories, R&D staff, etc., are utilized in R&D activities related to a particular innovation. In the context of regional innovation efficiency this implies that even if data are available for industry-specific input-factors (R&D employment) and innovative output (patents) these numbers are (even at the firm and plant level) rough approximations. Moreover, at the regional level usually only regions' (aggregated) patent and R&D employment portfolios are observed. While the former is organized according to technologies, i.e. patents are classified by the international patent classification (IPC), some sort of economic sector classification organizes employment data (in Europe this is the NACE).[4] For this reason, approximated matching concordances are employed, such as the one proposed by Schmoch et al. (2003). This additionally introduces severe (and unknown) biases into the estimation.

Hence, we are dealing with two empirical problems. The first is the **unobserved resource allocation**: it is unknown to what extent regional innovation generators of a particular industry (NACE-code) dedicate their resources to R&D in a particular technology (patent IPC-code). Second, and this is what we are interested in, the efficiency with which these resources are then transformed into innovative output is not observed (**unknown efficiency**). Accordingly, the unobserved resource allocation blurs the efficiency estimate.

As we pointed out above, innovations (and patents for that matter) in different technologies vary considerably in their structure and in the resources needed for their realization (cf. Malerba & Orsenigo, 1993). When this differentiation is applied to the innovative output (e.g., patents per IPC-code), we are facing a so-called *shared-resource*, or *shared-input* factor problem because we do not know the exact allocation of R&D employees in the various industries among technologies. In the following section we will propose an empirical approach designed to overcome this problem.

3 Empirical approach

3.1 Data Envelopment Analysis

The suggested approach builds upon the Data Envelopment Analysis (DEA) methodology, which is very popular in the Operations Research and Management Science literature where it was originally developed by Farrell (1957) and further popularized by Charnes et al. (1978). In essence, DEA is a non-parametric efficiency measurement technique that estimates the efficiency of n ($j=1,\dots,n$) units, which use certain levels of p different inputs $x_{j,i}$ ($i=1,\dots,p$) to produce q different outputs $y_{j,r}$ ($r=1,\dots,q$) outputs typically characterized by no reliable information on the prices (weights) of inputs and outputs and/or no (exact) knowledge about the 'functional form' of the production function. The DEA-model computes

[4] Nomenclature Generale des Activites Economiques dans I`Union Europeenne (NACE).

for each unit the efficiency measure e_k^a as a ratio of a weighted sum of outputs over a weighted sum of inputs. DEA estimates this efficiency score in relation to the efficiencies of the other activity units in the sample. The estimation of the innovation efficiency for the evaluated region *k* can be made with linear programming:[5]

$$\underset{w_{k,r},v_{k,r}}{Max}\ e_k^a = \sum_{r=1}^{q} w_{k,r} y_{k,r} \qquad (1)$$

$$s.t.$$

$$\sum_{r=1}^{q} w_{k,r} y_{j,r} - \sum_{i=1}^{p} v_{k,i} x_{j,i} \le 0 \qquad \forall\ j = 1,...,n \qquad (1a)$$

$$\sum_{i=1}^{p} v_{k,i} x_{k,i} = 1 \qquad (1b)$$

$$v_{k,i}, w_{k,r} \ge \varepsilon \qquad \forall i = 1,..., p; \forall r = 1,...,q \qquad (1c)$$

Whereby, the input weights and output weights are $v_{k,i}\ (i = 1,..., p)$ and $w_{k,r}\ (r = 1,..., q)$, respectively. The key feature of the DEA-model is that the input and output weights are endogenously estimated (hence, the non-parametric nature of DEA). As information on the true values of the input and output weights is usually lacking, DEA looks for plausible weights by letting the data speak for themselves. In the context of our paper, this means that in the evaluation of regions' innovation efficiencies DEA looks (for each region) for the input factor and innovative output weights such that the region is evaluated optimally relative to the other regions in the sample set (i.e., with the highest possible global innovation efficiency score e_k^a).

By the normalization constraint, it holds that the global efficiency score e_k^a is situated between 0 and 1. An efficiency score for the evaluated region k below one implies that there is at least one region (but very probably multiple regions) in the dataset, which realize a better level of global innovation efficiency. Moreover, in the interpretation of this global innovation efficiency score e_k^a, note that the difference $1-e_k^a$ can be perceived as a measure of the overall inefficiency in a region's innovation performance, which quantifies its room for innovation efficiency improvements.

Before concluding this section, note that we argued for disaggregating the innovative output into multiple technologies to take the technology-specificity of regional innovation efficiency into account. By way of contrast, there is only a single, but *shared*, input 'R&D employment'' because we don't know its exact distribution across the technologies. This implies that in this

[5] For a more elaborate introduction see Daraio & Simar (2007).

set-up the traditional DEA-estimate will mix the unknown allocation of R&D employment with the unknown efficiency distribution.

3.2 Shared input DEA-analysis

The standard DEA-model can be adapted so that the focus will be on the estimation of partial innovation efficiency scores. The basic question to be addressed is: "*how should one split the shared input factor 'R&D employment' between the several technologies in the DEA model?*".

The solution proposed is to determine the distribution of the shared factor input among the different technologies (output dimensions) endogenously, which is similar to the definition of the input factor and innovative output weights in the traditional DEA-model. For this we make use of the **shared input DEA-model**, which is based on the contributions of, among others, Beasley (1995) and Cook et al. (2000, 2004). In essence, it adapts the basic DEA-model so that the key DEA-feature of endogenously determining the unknowns is not only applied in the definition of the input and output weights but also in the definition of the distribution of the shared-input. The shared input DEA-model thus determines for each evaluated unit the input and output weights as well as the distribution of the shared input from a relative perspective to the other regions in the sample. Formally,

$$\underset{w_{k,r},v_{k,r}}{Max}\ e_k^a = \sum_{r=1}^{q} w_{k,r} y_{k,r} \qquad (2)$$

$$s.t.$$

$$\sum_{r=1}^{q} w_{k,r} y_{j,r} - \sum_{i=1}^{p} v_{k,i} x_{j,i} \le 0 \qquad \forall\ j = 1,...,n \qquad (2a)$$

$$w_{k,r} y_{j,r} - \alpha_{k,r} v_{k,i} x_{j,i} \le 0 \qquad \forall\ r = 1,...,q; \forall\ j = 1,...,n \qquad (2b)$$

$$\sum_{i=1}^{p} v_{k,i} x_{k,i} = 1 \qquad (2c)$$

$$v_{k,i}, w_{k,r} \ge 0 \qquad \forall i = 1,..., p; \forall r = 1,...,q \qquad (2d)$$

$$\alpha_{k,r} \ge \varepsilon \qquad \forall r = 1,...,q \qquad (2e)$$

$$\sum_{r=1}^{q} \alpha_{k,r} = 1 \qquad (2f)$$

In the context of the paper, e_k^a is the global innovation efficiency score for the evaluated region k; $\alpha_{k,r}$ the DEA-estimated input factor shares, that is to say, the DEA-estimated shares of R&D employment across all technologies for the evaluated region k. The assumption is that all R&D employment is accounted for by the considered technologies: $\sum_{r=1}^{q}\alpha_{k,r}=1$ (constraint $(2f)$). As with the input factor and innovative output weights, the shared input DEA-model determines the R&D employment shares for the evaluated region such that the global innovation efficiency score e_k^a is optimal. As noted by Beasley (1995) in a different context, the advantage of letting the DEA-model decide on the input factor shares is that there is no need to determine an a priori distribution of the shared-input factor across the technologies. Clearly, in the context of ambiguity concerning the true R&D employment distribution, this feature is an advantage (or at least, an appealing second-best route).

3.3 Restricting the employment shares

The shared input DEA-model as in $(2)-(2f)$ is very flexible in the definition of the optimal employment shares. As with the input and output weights, only a normalization and non-negativity constraint applies to the optimal $\alpha_{k,r}$-values (i.e., $\alpha_{k,r}\geq\varepsilon$). The non-negativity constraint $(2e)$ imposes that at least a very small fraction ($\varepsilon=0.001$ or 0.1%) of the shared-input should be allocated to each technology in each region. The advantage of this flexibility is that the employment shares are optimally chosen for each evaluated region. This implies that eventual poor innovation efficiency scores cannot be blamed on the estimated $\alpha_{k,r}$-values because any other shares than the ones estimated will by definition lower a region's innovation efficiency. However, this flexibility also has disadvantages: in the attempt to maximize the region's innovation efficiency, it is perfectly allowable for the model to assign employment shares that are unrealistically low or high as the non-negativity constraint only imposes that $\alpha_{k,r}\geq 0.1\%$.

This potential problem can be overcome by fine-tuning the model so that the probability of the model selecting improper employment shares is lowered. More precisely, restrictions of the type $a_r\leq\alpha_{k,r}\leq b_r$ can be added to the model, which ensure that the optimal employment shares are fitted with the boundary values a_r and b_r. This requires information about the potential magnitude and range of those shares. Fortunately, as will be shown later on, such information is available in our case.

3.4 Measuring innovation efficiency change

To measure the change in innovation efficiency, we employ the Malmquist Productivity Index (MPI, hereafter). This frontier method was originally introduced by Malmquist (1953) and further popularized by, among others, Färe et al. (1994, 1998). The MPI measures the change in innovation efficiency between a period t and a subsequent $t+1$, denoted as IC_k, by calculating the ratio of innovation efficiency scores computed at each time relative to a common transformation technology.

$$IC_k = \left(\frac{\sum_{r=1}^{q} w_{k,r}^{t} y_{k,r}^{t+1}}{\sum_{r=1}^{q} w_{k,r}^{t} y_{k,r}^{t}} \right)^{1/2} \left(\frac{\sum_{r=1}^{q} w_{k,r}^{t+1} y_{k,r}^{t+1}}{\sum_{r=1}^{q} w_{k,r}^{t+1} y_{k,r}^{t}} \right)^{1/2} \qquad (3)$$

Each of the innovation efficiency scores[6] in (3) is measured by the shared input version of the DEA-model as described above. In the interpretation of the MPI-scores, IC_k-values above one indicate an improvement, and IC_k-values below one show a decrease, in the global innovation efficiency of the assessed region k during the observed period.

The MPI can be disaggregated into an "**environmental change**"-component (EC_k) and a "**catching-up**"-component (CU_k). The catching-up component reflects a region's idiosyncratic improvement and helps answering the question of how much closer it got to its 'contemporaneous' benchmark region(s). The environmental change component represents the change in the general innovation environment in which a region operates. This component focuses on the conduct of the benchmark region(s) (i.e., the change of the best practice region(s) between two periods). Mathematically, this disaggregation boils down to the following (for more on the decomposition, see Färe et al., 1994):

$$CU_k = \left(\frac{\sum_{r=1}^{q} w_{k,r}^{t+1} y_{k,r}^{t+1}}{\sum_{r=1}^{q} w_{k,r}^{t} y_{k,r}^{t}} \right) \qquad (4)$$

[6] Note that, for instance, $w_{k,r}^{t+1} y_{k,r}^{t}$ represents the efficiency estimate of region k in t benchmarked against the transformation technology (best-practice) in $t+1$.

$$EC_k = \left(\frac{\sum_{r=1}^{q} w_{k,r}^{t} y_{k,r}^{t+1}}{\sum_{r=1}^{q} w_{k,r}^{t+1} y_{k,r}^{t+1}} \right)^{1/2} \left(\frac{\sum_{r=1}^{q} w_{k,r}^{t} y_{k,r}^{t}}{\sum_{r=1}^{q} w_{k,r}^{t+1} y_{k,r}^{t}} \right)^{1/2} \qquad (5)$$

A CU-value above one indicates progress in a region's innovation efficiency relative to the benchmark regions between the two periods due to own effort, while the opposite interpretation holds for CU_k -values below one. In the same vein, values for the environmental change-component above (below) one imply that the general innovation environment has improved (worsened), i.e. it generally takes fewer input factors to create the same quantity of innovative output.

Two components of the innovation efficiency change may move in opposite directions. For instance, an increase (decrease) in innovation efficiency may be observed because of an improvement in a region's own efforts even if there is simultaneously a less (more) favorable general innovation environment than in the original period.

3.5 Robust efficiency analysis

An important drawback of the non-parametric DEA-model and MPI-approach is their sensitivity to the influences of outliers and measurement errors (or other data irregularities). This dependency results from two features: the deterministic nature of the DEA-model by which efficiency scores are taken to be perfect reflections of actual efficiency without considering the potential for any noise or other irregularities in the data. In addition, the input-output combinations of all regions are considered in the computation of each region's efficiency.

To overcome this drawback, the DEA-model and the MPI-analysis are adjusted to the insights of the robust order-m efficiency model of Cazals et al. (2002). Without going into detail, this approach models the transformation of inputs into outputs as a stochastic process and evaluates a region's output level (output-orientation) against the expected maximal value of output achieved by regions with equal or lower input levels. This considerably minimizes a single region's impact on the evaluations of other regions and hence the potential influence of statistical noise. In practice, Cazals et al. (2002) propose a Monte Carlo simulation approach in which each region's efficiency is estimated in a large number of DEA-based computation rounds (*in casu*, 1,000 rounds) in each of which its innovation efficiency is evaluated relative to a subsample of m randomly selected regions with equal or fewer inputs levels.[7] The robust innovation efficiency estimate is then computed as the average innovation efficiency score defined over the rounds (see Daraio & Simar (2007) for a detailed discussion on robust efficiency analyses and the role of the parameter *m*).

3.6 Data on regional input factors and innovative output

The 150 German labor market regions as defined by Eckey et al. (2006) are chosen as units of analysis because they seem to fit best the theoretical arguments for a regional dimension of innovation processes (Broekel & Binder, 2007).

As is common in this type of research, regions' innovative output is approximated by patent applications.[8] The data are taken from the German Patent and Trademark Office (DPMA) within the period from 1999 to 2008. Patents are organized into a multi-digit classification, the International Patent Classification (IPC). The inventor principle is applied to regionalize the patent data, i.e. each patent is assigned to the labor market region where its inventor is located. In the event that a patent is developed by multiple inventors located in different regions, it is equally assigned to each region.

We obtain data on R&D employees from the German labor market statistics. Following Bade (1987) the R&D personnel are defined as the sum of the occupational groups: agrarian engineers (032), engineers (60), physicists, chemists, mathematicians (61), and other natural scientists (883). In addition, the employees are available on an industry-specific basis as they are classified according to the NACE-classification.

The shared-input DEA implies that it is sufficient to consider the technology-specificity exclusively in the definition of the innovative output. On the input side we treat total R&D employment as shared-input, i.e. a single variable. We nevertheless make the two data sources comparable for three reasons. Firstly, the IPC captures the technological dimension of patented inventions while we are actually more interested in industry-specificities (i.e. the differentiation on the input side). Secondly, in order to evaluate the shared-input DEA-model it is useful to construct some industry-specific measures that require a matching of patent and employment data. Thirdly, we argued above that the quality of the shared-input DEA results greatly improves when upper and lower bounds for R&D

[7] We use $m = 50$. Sensitivity analysis points out that the results are relatively robust with respect to alternative choices of value of m (i.e., we also considered m-values of 20, 30, 40, 60, 70, 80, 90, and 100). The outcomes of the DEA-models and the MPI-analysis based on other values of m are available upon request.

[8] We acknowledge that patents capture inventions rather than innovation. However, in order to stay consistent with the literature we will use the term "innovation".

employment shares (respectively the a_r- and b_r-values as described in Section 3.3) for each technology (i.e. each dimension of the output variable) are specified. This is easily possible when total R&D employment is disaggregated according to the same dimensions as the innovative output.

For these three reasons, we employ the "standard" concordance developed by Schmoch et al. (2003) that relates the 3-digit NACE-classification of the R&D employees and the patents organized by IPC codes. Note that by making the input and output dimensions comparable in this way, we actually transform our original technology-specific innovation efficiency measure into an industry-specific measure. We therefore use the latter in the remainder of the paper.

The conversion results in 43 industries[9] for which corresponding R&D employment and patent data are estimated for each of the 150 German labor market regions. This means that we have 43 different innovative outputs, which correspond to the patent counts of regional organizations in the 43 industries. These are related to the total R&D employment numbers in the shared input data. However, in order to specify the upper and lower bounds for R&D employment shares in each of the 43 industries, we estimated for each region the 43 shares of industry-specific R&D employment. Subsequently, the minimum level of (maximum) employment shares across all regions for each industry is used as global lower (upper) bound for the respective industry in the shared-input DEA-estimations. For example, for the industry I9 (*basic chemicals*), the estimated minimum and maximum shares suggest that the optimal R&D employment share should be situated between 0.1% and 46.18%.

4 Empirical results

4.1 *Evaluating the efficiency measure*

4.1.1 Correlation with ratio measure

In order to put the results of the shared-input DEA-model, which are denoted with EFF in the following, into perspective, they are compared to the simplest type of "innovation efficiency measure", namely the ratio between a region's total patent output and the total number of regional R&D employees. The ratio measure represents the weighted average of

[9] Schmoch et al. (2003) identify 44 sectors, however, no patents are recorded for one sector.

the 43 industry-specific ratios of patent counts and R&D employment numbers.[10] The implicit weights correspond to industries' employment shares of total employment (relative input). The ratio is chosen as a benchmark as it strongly relates to the industrial structure of a region.

EFF and the ratio measure strongly correlate at R=0.847*** (1999-2003) and R=0.837*** (2004-2008), as visualized in **Figure 1**.[11] Accordingly, the two measures are relatively similar, which seemingly signals a relatively weak impact of regions' industry structures on their global innovation efficiency. It is, however, important to keep in mind that there are just two scenarios in which the two measures can be expected to yield very large differences. In the first one, the ratio measure will underestimate regions' global innovation efficiency if they are dominated by industries with low patent intensities, but which are (relatively) highly innovation efficient (**underestimation scenario**). In the second scenario, regions' innovation efficiency will be overestimated by the ratio if they are dominated by patent intensive industries that are (relatively) innovation inefficient in comparison to other regions (**overestimation scenario**).

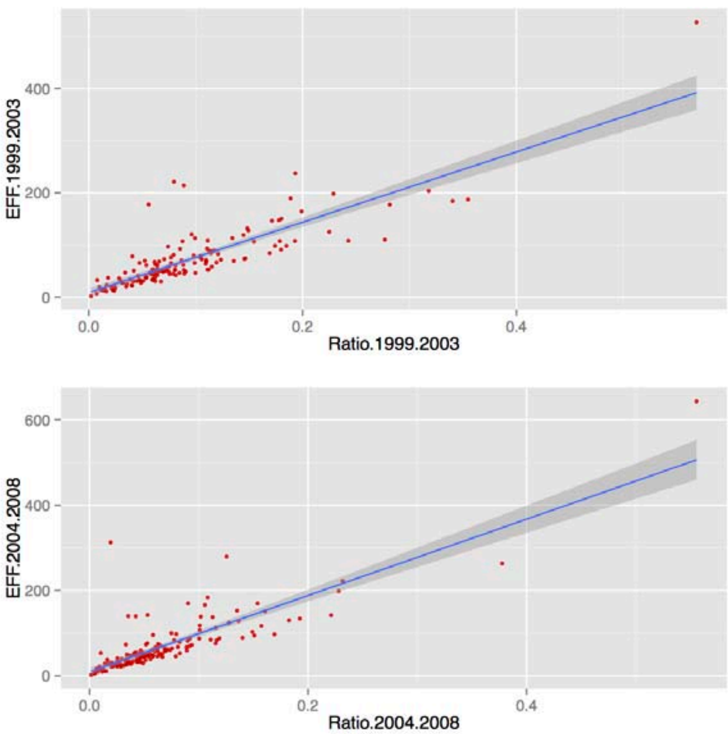

Figure 1: Ratio measure vs. efficiency

[10] The ratio measure strongly rank correlates to the residuals of a linear regression of logarithmized R&D employment and logarithmized innovation output ($R_s = 0.99$***). This method is frequently used when assessing the impact of regional factors' on regions' innovation performance.

[11] The rank correlations are similar in magnitude: 0.866*** (1999-2003) and 0.848*** (2004-2008).

In any other than these two scenarios, the two measures are likely to produce similar efficiency rankings of regions. While these two scenarios are not very common, they represent a considerable number of cases, which is indicated by the two measures' correlation remaining well below one. In this respect, it is also important to point out that the likelihood of observing one of the two scenarios is subject to the disaggregation of the input and output dimension. It is generally less likely that one will observe regions being dominated by an industry when the industrial dimension is strongly disaggregated, i.e. when the average relative size of industries is smaller. Reducing the dimensionality on the innovative output side, i.e. disaggregating the patents into fewer industries, tends to decreases the correlation between ratio and efficiency measure. This is related to the general tendency of weighting schemes to lose their importance when the number of weighted items increases (Wang & Stanley, 1970).

The disaggregation of the innovative output (and innovation input) also relates to a problem, which in its extreme is known as "*sparsity bias*" (Pedraja-Chaparro et al., 1999, p. 638). A growing dimensionality of the input-output space tends to decrease the number of regions that are comparable with another particular region, i.e. that may serve as benchmarks. This in turn will increase the regions' efficiency levels. A lower input-output dimensionality will accordingly allow the efficiency analysis to be more discriminating, which is likely to further reduce the correlation with the ratio measure. When applying the presented method, researchers therefore face a trade-off between the degree of industrial disaggregation and the extent with which the industrial structure influences the empirical results.

In summary, the relatively large correlation between the efficiency estimates and the ratio measure are explained by the chosen disaggregation of the innovative output (43 industries), the comparatively small sample size (150 regions), and the relative rareness of regions being highly specialized in industries with high (low) patent intensities but regionally low (high) innovation efficiencies.

The correlation in the levels of innovation efficiency translates into a significant, however comparatively smaller, correlation in their rates of change. The rank-correlation between the relative change in the ratio measure and the change of EFF between the two periods is R_S=0.71***.[12] Hence, the average difference between the two approaches' results is more severe in the dynamic perspective than when looking at one particular time period.

[12] Here, we only consider the catching-up component of this change, denoted as CU.

4.1.2 Comparison of individual cases

While being strong, the correlation between the two measures is far from the maximum value of one and hence there are considerable differences between the two. These are highlighted in the following on the basis of particular illustrative cases.

Take for instance the region Rottweil in the period 1999-2003. The region is located in the far southwest of Germany close to Freiburg im Breisgau. In terms of R&D employment and patent output Rottweil is of average size. It ranks 72th in patent output (1046.34 patents) and 60th in R&D employment (18,690 employees). According to the ratio of patent output and R&D employment (0.056) it has the 105th highest ratio among the 150 regions. However, its innovation efficiency value is 177.7%, which ranks 10th. In other words, we find a discrepancy of 95 ranks between the two measures.

Hence, Rottweil seems be characterized by an unfavorable industrial structure. This is confirmed by the data. The three industries with the largest patent output in Rottweil are I19 (*fabricated metal products*), I23 (*machine tools*), and I25 (*weapons and ammunition*). They account for about 81% of the region's total patent output. The three industries' industry-specific ratios of patent output to R&D employees are 0.057 (I19), 0.067 (I23), and 0.006 (I25). They rank 25th, 21st, and 42th among the 43 industries (see Table 3). Hence, Rottweil is strongly specialized in industries with low patent to R&D employment ratios. It is therefore strongly discriminated against when measured according to the simple ratio analysis. Our innovation efficiency measure reveals, moreover, that the region does outstandingly well in industry I25 (efficiency of 469%), while it is relatively inefficient in industries I19 (35%) and I23 (33%).

Another illustrative example is Darmstadt, a region in the south of Hessia, in the period 2004-2008. The region ranks much higher than Rottweil in terms of patent output (17th) and R&D employment (50th) with values of 3,833.4 patents and about 22,644 R&D employees. With a ratio of 0.169 it is among the top-ten German regions (rank 8). However, when taking its industrial structure into account it drops to rank 30 in the innovation efficiency ranking with a value of 97.04%. Hence, the difference is again marked with 22 ranks. However, in contrast to Rottweil, Darmstadt benefits from its industrial structure in the ratio estimation. About 43% of its patent outputs are attributed to industries I12 (*pharmaceuticals*) with 29% and I9 (*basic chemicals*) with 13%. These are the top-two industries in terms of patents to R&D employment ratios (I12: 0.545 and I9: 0.326).

Besides these two regions a great number of similar examples exist in the period 1999-2003. For instance, Wiesbaden is 7th when considering the ratio while it ranks at 25th in the

DEA-based innovation efficiency analysis in 1999-2003. Prignitz has the 66th highest ratio but ranks 3rd in terms of innovation efficiency. Trier is 62nd with respect to the ratio but 105th in the innovation efficiency value. Munich and Hannover both profit from their favorable industrial structures: Munich ranks 21st and Hannover 48th when looking at the ratio, while the corresponding values in the innovation efficiency analysis are 39th and 96th, respectively. Comparable patterns are also found in 2004-2008. A notable additional example is Stuttgart dropping from 83rd place in the ratio to the 121st place in the innovation efficiency analysis, highlighting its comparatively favorable industrial structure.

These examples and the mean difference between regions' ranks according to the two measures, which amounts to approximately 16 positions in both periods, highlight the fact that ignoring the industrial structure can be very misleading when measuring regions' innovation efficiency.

4.1.3 Regions' size and innovation efficiency

Another important quality aspect of a measure of regional innovation efficiency is its uncorrelatedness to regions' size. Table 1 reports the rank of a region in terms of total patent output. The innovative output is chosen over the input factor because it can be interpreted as a measure of a region's strength in innovation activities inasmuch as it combines the magnitude of R&D efforts and a region's efficiency in generating patentable innovation. The table shows that the top spots of the efficiency ranks are regions with relatively small innovation outputs. However, the fourth (1999-2003) and fifth (20034-2008) spots are regions (Jena and Aachen) already rank fairly high in terms of innovative output. The consecutive spots are also inhabited by regions with medium to large innovation output. Stuttgart in 2004-2008 is also an illustrative example for the difference between regions being innovative in terms of total number of innovations, and regions being able to efficiently transform their inputs into innovation. While it holds the top spot in patented innovations, it ranks 121st in innovation efficiency with a value of EFF = 29.58.

This underlines the size-independence of the efficiency measure. The rank correlations between the output and efficiency measure confirm this with values of $R_S = 0.33^{***}$ (1999-2003) and $R_S = 0.30^{***}$ (2004-2008). The efficiency measure is also doing a much better job of controlling for regions' size than the ratio measure for which the corresponding correlations amount to $R_S = 0.52^{***}$ in 1999-2003 and 2004-2008.

1999-2003				2004-2008			
Region	Efficiency	Patents	Output rank	Region	Efficiency	Patents	Output rank
Garmisch-Patenkirchen	527.22	326.3	112	Garmisch-Patenkirchen	643.55	230.8	113
Nordfriesland	237.36	83.6	137	Prignitz	312.87	18.5	146
Saalfeld	221.39	208.6	123	Nordfriesland	279.43	55.9	137
Rügen	214.1	10.5	150	Aachen	263.3	7783.5	8
Jena	203.92	1967.9	51	Jena	221.83	1491.5	49
Lörrach	198.42	2690.5	38	Altötting	198.39	1374.1	53
Kempten	189.58	748.072	87	Bodensee	183.5	2348.3	29
Aachen	187.2	9313.1	10	Osterode	169.76	151.2	123
Mainz	184.38	5582.3	17	Rügen	169.68	18.3	147
Rottweil	177.71	1046.3	72	Bremerhaven	166.16	189.2	115
Altötting	177.34	1664.3	57	Traunstein	152.92	1296.6	55
Pirmasens	164.22	554.7	98	Kempten	151.05	642.1	82
Traunstein	150.38	1800.5	54	Saalfeld	142.41	137.2	127
Würzburg	147.81	3268.4	29	Mainz	141.89	3576.5	18
Regensburg	146.7	5013.4	19	Amberg	139.76	656.3	80

Table 1: Top fifteen innovation efficient region

In summary, we find that the proposed efficiency measures yield results of significantly higher quality than a simple ratio approach. This is shown particularly when evaluated regions are highly specialized in industries with very low or very high patent intensities. Accordingly, it is superior to most existing measures and allows for a scientifically reliable evaluation of regions' innovation efficiency.

4.2 The innovation efficiency of German regions

The mean efficiency is 71.43% in 1999-2003 and 70.89% in 2004-2008. The median is 52.9% and 51.09% in 1999-2003 and 2004-2008, respectively, illustrating a left-skewed distribution, which is visualized in Figure 2. The obtained mean values are substantially larger than comparable scores reported in Broekel (2010) that range between 15 and 50.5.[13] Accordingly, regions' overall efficiency is higher than comparable estimates of Broekel (2010). The difference may in part be explained by the inflation of Broekel's efficiency scores

[13] The mean values of regional innovation efficiency reported in Broekel (2010) for different industries are actually 1.98, 2.34, 3.72, and 6.7. They need to be converted with 1/x*100% to become comparable to the values reported in the present paper.

due to some very extreme input / output combinations in his data that are caused by his purely industry-specific estimation approach.

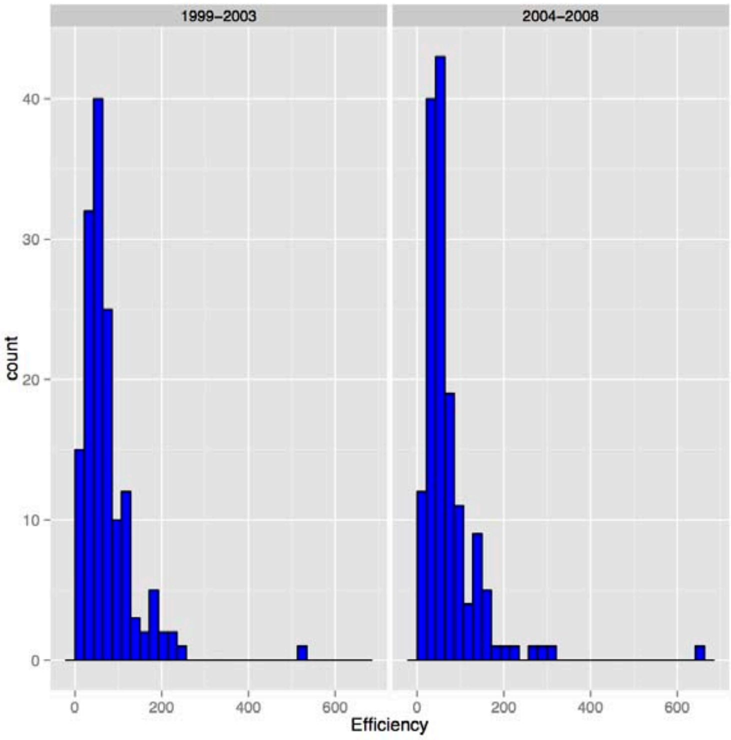

Figure 2: Distribution of efficiency values

Our analysis identifies 29 regions (19% of regions) as significant in generating patentable innovation in the period 1999-2003, and 27 (18% of regions) in 2004-2008. Their median efficiency score is 146.7% (1999-2003) and 141.89% (2004-2008) indicating a highly efficient innovation performance. The comparison with the median scores of the inefficient regions (1999-2003: 48.61%; 2003-2008: 46.44%) highlights that there is substantial potential for increasing regions' innovation efficiency in Germany.

We report the top fifteen innovation efficient regions in Table 1. Garmisch-Partenkirchen holds the top spot in both periods. Garmisch-Partenkirchen is highly efficient in a number of industries (11) with industries I4 (*wearing apparel*) and I6 (*wood production*) particularly obtaining dominating weights in the estimation. The region's outstanding performance is, however, explained by the fact that it has the largest number of efficient industries (EFF >= 100) in both periods (1999-2003: 9; 2004-2008: 16).

In 1999-2003, just 24 regions (23 in 2004-2008) are characterized by more than one innovation efficient industry, with the mean number of efficient industries per region being 3.5 (3.8 in 2004-2008). These figures underline the fact that outstanding innovation efficiency is by and large related to regions being highly efficient in a small number of industries.

Aachen and Jena, which hold the top spots for the most innovation efficient regions among regions with more than average patent output, are interesting cases. Both regions' good innovation efficiencies come from being efficient in multiple industries. Aachen particularly profits from excellent innovation efficiencies in industry I27 (*office machinery and computers*), I30 (*accumulators, battery*), I34 (*signal transmission, telecommunications*). In Jena, industries I36 (*medical equipment*), I37 (*measuring instruments*), and particularly I39 (*optical instruments*) explain its outstanding performance. All these industries represent well-known strengths of the particular regions, as exemplified by the relevance of the optical instruments industry in Jena with the headquarters of ZEISS and Jenoptik being located there.

The three cases of Garmisch-Partenkirchen, Aachen, and Jena confirm the idea that excellence in just one industry is not enough to achieve outstanding innovation efficiency. It takes a number of (potentially related) industries to boost a regions' performance (cf. Cooke et al., 1997; Frenken et al., 2007).

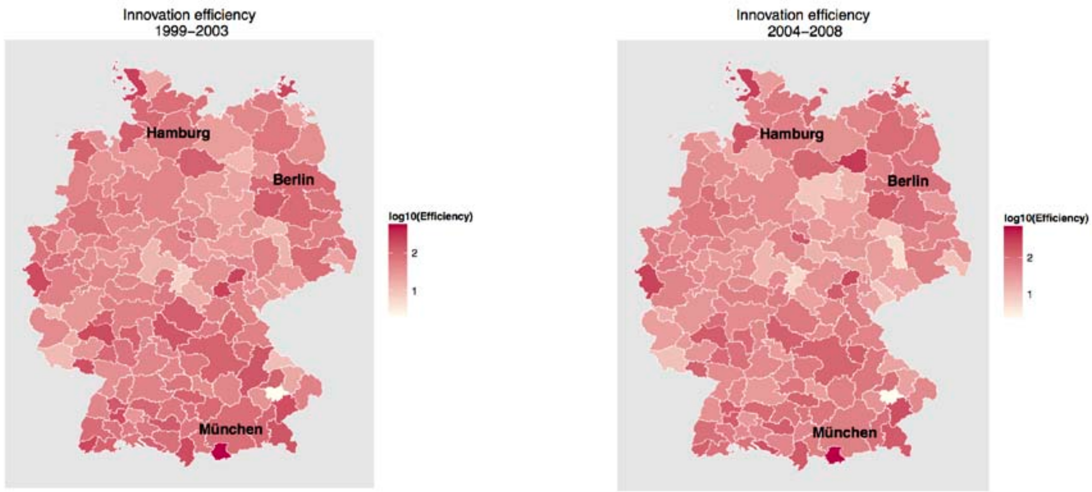

Figure 3: Geographic distribution of regional innovation efficiency

The maps in Figure 3 show the spatial distribution of regional innovation efficiency. While the visual inspection does not suggest the generally lower innovation efficiency of East German regions as reported in Fritsch & Slavtchev (2008) and Broekel (2012), East German regions have a mean efficiency of 56.2% in contrast to West German regions' 76.4% in 1999-2003. The difference drops somewhat in the following period (West: 74.7%, East: 59.34%). Both differences are significant at the 0.01 level.[14] We therefore confirm previous findings in the literature concerning the existence of an innovativeness gap between the two parts of

[14] Wilcoxon rank sum test 1999-2003: W = 2788, p-value = 0.002376; Wilcoxon rank sum test 2004-2008: W = 2696, p-value = 0.008348.

Germany, which not only shows in absolute patent numbers but also in terms of efficient innovation generation. We also show that the inefficiency of East German regions cannot be attributed solely to the presence of unfavorable industrial structures.

The second impression derived from the maps in Figure 3 is that regional innovation efficiency appears to be geographically clustered. While the Moran's I test on the efficiency scores (1999-2003 I = 0.064*; 2004-2008 I = 0.025) provides only little support, the same test applied to the ranks of the efficiency values is confirmative (1999-2003 I = 0.18***; 2004-2008 I = 0.22***).[15] We thus find regional innovation efficiency to be significantly spatially (rank-) autocorrelated.

4.3 Temporal dynamics of regional innovation efficiency

The two innovation efficiency measures computed for the regions for the periods 1999-2003 and 2004-2008 show similar patterns. Their correlation is, however, just 0.87***, which suggests that some regions improved and others declined in innovation efficiency.

We argued above that temporal change in innovation efficiency can be disaggregated into two components representing different types of change. The first type of dynamic concerns the "environmental change"-component (EC), i.e. how does the overall (region external) environment for generating innovations develop between two time periods. The mean value of the EC-component is just 0.76 and thereby well below one. Accordingly, the overall conditions for innovation generation deteriorated between the two periods. In other words, in 2004-2008 it takes on average significantly more R&D workers to generate the same number of patentable innovation than in 1999-2003. A possible explanation might be the "*.com*" bubble in 2000-2001 that boosted patent numbers in the earlier period but by its bursting led to strong decline in patenting in the second period. This clearly deserves more attention in future research.

The second component of innovation efficiency change reflects whether a region is able to catch up to its contemporaneous comparison regions. **Table 2** lists the top fifteen regions that were able to catch up. The region of Prignitz holds the top spot with a CU-value of 23.67. It clearly represents an extreme value caused by its very low level of output making extreme growth rates more likely (output rank: 148 in 1999-2003). This pattern of regions with low levels of output growing more in terms of efficiency than regions with large output values dominates **Table 2**. However, it is surprisingly more or less restricted to the listed regions. The correlation between patent output in 1999-2003 and efficiency growth is just about

[15] For the estimation of Moran's I we used direct neighborhood as basis for the spatial weight matrix.

R = -0.05 (rank correlation: $R_S = 0.01$). Stuttgart, the number one region in terms of patent output in both periods, is an example in this respect. While it ranks average at place 82 in 1999-2003, its rank in 2004-2008 is 121. Its efficiency dropped from 50.93% to 29.58%, which corresponds to a decline in the CU-component efficiency of 0.58%. The decline is caused by a decrease in patentable output of -28% from 35,871 patents to 25,713. The input, i.e. the number of R&D employees, remained more or less the same.

In addition to the regions' level of innovative output, the level of innovation efficiency is another factor that does not predict growth in innovation efficiency well. The relevant correlation is R = -0.09 (rank correlation: $R_S = -0.14$*). Other factors than regions' initial output and innovation efficiency level are accordingly driving the development of regional innovation efficiency.

Region	Efficiency change (CU)	Efficiency rank 1999-2003	Efficiency rank 2004-2008	Output rank 1999-2003	Output rank 2004-2008
Prignitz	23.67	145	2	148	146
Greifswald	2.62	133	58	124	114
Cham	2.07	148	134	127	117
Uckermark	1.98	111	46	143	130
Amberg	1.79	45	15	82	80
Osterode	1.75	32	8	128	123
Birkenfeld	1.71	139	118	131	133
Ostprignitz	1.64	119	69	149	148
Stralsund	1.63	68	28	142	139
Bodensee	1.62	21	7	35	29
Annaberg	1.47	141	136	140	143
Neubrandenburg	1.45	61	31	138	129
Suhl	1.45	140	130	118	111
Osnabrück	1.43	71	39	59	52
Flensburg	1.42	138	124	119	121

Table 2: Change in regional innovation efficiency from 1999-2003 to 2004-2008

In **Figure 4** we show the geographic distribution of innovation efficiency change (i.e. the extent to which regions are catching up). No particular neighborhood patterns are visible, something that is confirmed by a very small Moran's I: 0.01 (Moran's I of ranks: 0.03). This is about the same magnitude of spatial autocorrelation as Broekel (2010) reports for the change in industry-specific innovation efficiency measures. Innovation efficiency growth processes of neighboring labor market regions are more or less unrelated, a fact that can be interpreted to mean that the chosen delineation of the spatial units captures the (outer) spatial dimension of innovation processes well. This absence of spatial autocorrelation indicates non-spatially structured innovation efficiency growth processes.

Lastly, we look at the development of innovation efficiency in the two parts of Germany. It appears to be the case that the average level of innovation efficiency has decreased: The mean innovation efficiency of East German is 56.2% in 1999-2003 and increases to 59.34% in 2004-2008. The relevant levels are 76.4% in 1999-2003 and 74.7% in 2004-2008 for West German regions. However, a shrinking difference between the two parts of Germany cannot be statistically backed. While there is a difference in mean innovation efficiency growth of West and East German regions (West: 0.99%; East: 1.58%) it is not significant in any test set-up (t-Test, Wilcoxon test, log-transformed growth rates). Accordingly, we do not find statistically robust signs for a convergence between the two parts of Germany in terms of regional innovation efficiency.

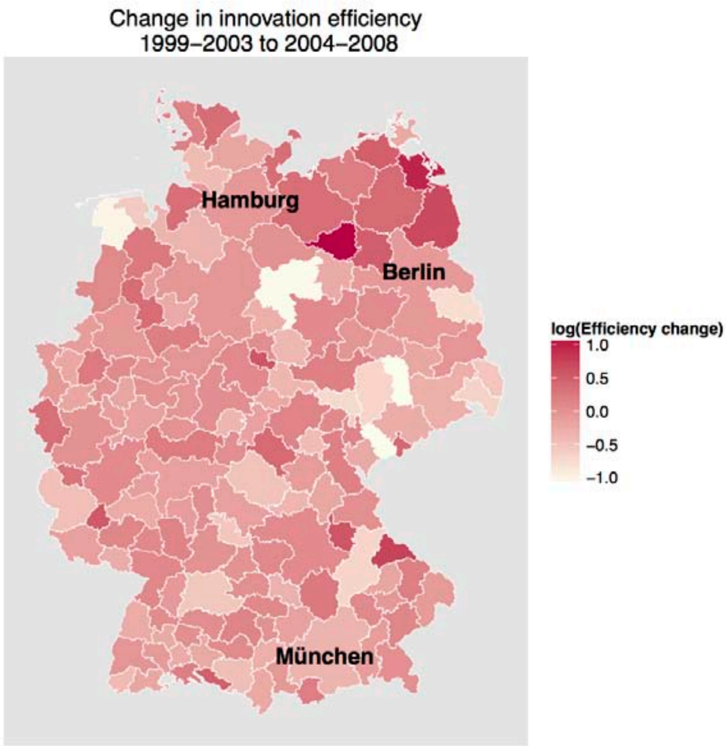

Figure 4: Geographic distribution of innovation efficiency change

5. Conclusion

In the paper we proposed the use of a recently developed method to estimate regions' innovation efficiency: namely the robust shared-input Data Envelopment Analysis. We argued that this method is particularly advantageous when analyzing regional innovation efficiency using employment data to approximate input factors and patent numbers as proxies for innovative output, i.e. in the scenario most common in this type of literature.

Amongst the more obvious advantages of the method are the sparsity of theoretical and empirical assumptions and the requirement of using only publicly available data, while it still allows for considering differences in regions' industrial structures in the estimation. As shown in the paper, the latter are biasing traditional measures leading to the overestimation or underestimation of regions' innovation efficiency when regions are dominated by industries with very low or very high patent intensities. By means of a comparison with the number of patented innovations per R&D employee, it has been shown that the new measure yields similar ranking structures but significantly reduces the bias induced by these over- and underestimation scenarios. Other advantages of the proposed method include differentiating between innovation efficient and innovation inefficient regions, its clear indications concerning the level of inefficiency in regions, and the possibility of disaggregating the obtained global measure of regional innovation efficiency into industry-specific innovation efficiency scores with limited use of matching concordances between employment and patent data. Moreover, when studying the dynamics of the measure by means of a Malmquist-index approach, different components of innovation efficiency changes can be identified.

We illustrated the usefulness of the method by investigating the innovation efficiency of German labor market regions in two periods, 1999-2003 and 2004-2008. Summarizing the empirical findings of our analysis, we firstly showed that there is considerable variance in regional innovation efficiencies among German regions, which cannot exclusively be explained by the location of innovation intensive industries. Garmisch-Partenkirchen was identified as Germany's most innovation efficient region. Aachen and Jena represent the top performers among regions with more than average absolute patent output. All these regions are characterized by being innovation efficient in multiple industries. We moreover confirmed the existence of a gap in innovation efficiency between East and West German regions in both periods. The comparison of the two time periods reveals that this gap persists and that on average no significant signs of convergence in terms of innovation efficiency are observable.

Despite the method's advantages and interesting empirical findings, a number of issues need to be put into perspective. Most importantly, in order to construct an innovation efficiency measure that takes into account regions' industrial structures, the method principally is applicable when information on industry (or technology) structures is exclusively available for the innovative output and not for the input factors. In the case of patent data, such information is publicly available. It is the matching of these industry/technology-specific patent numbers to the relevant industry/technology-specific employment numbers that is either unavailable or comes in the form of approximate associations. In both cases, the matching introduces significant inaccuracies into the empirical results obtained by traditional approaches designed to assess regions' innovation performance.

We showed that such matching procedures become almost obsolete when using the shared-input DEA-model. However, while the method requires very limited information on the matching of patents and employment, it cannot do entirely without it. In order to obtain useful results, we employed information on the maximum and minimum regional share of R&D employees in each industry observed across all regions in order to specify restrictions on the employment shares in the DEA-computations. Alternative approaches designed to define restrictions when lacking such information are yet to be explored.

The employed method can also be extended by means of using conditional efficiency approaches (cf., Daraio & Simar, 2007). These allow for nonparametrical consideration of external factors that may additionally alter regions' innovation efficiency (e.g., regional degrees of urbanization or specialization).

In addition to further methodological advances, the empirical specification of regional innovation efficiency remains an issue deserving more research. So far, the innovative output has been exclusively calculated by using patent data, which gives a restricted picture of organizations' innovative output. In the future, other forms of quantifiable innovation indicators should be considered. Academic publications are obvious candidates, although information on industry-specific shares of new and improved products is also desirable. A similar approach applies to the input side. While the argument has been made that R&D employees are the most important input into innovation processes, other inputs cannot be easily ignored. Financial and non-employment related assets (laboratories, equipment, etc.) are most important in this respect. Due to missing data they are widely ignored. Accordingly, the present study is just one further step in the direction of an academically sound and practically meaningful measure of regional innovation efficiency.

References

Arundel, A., & Kabla, I. (1998). What Percentage of Innovations are Patented? Empirical Estimates for European Firms. Research Policy.

Audretsch, D. (1998). Agglomeration and the Location of Innovative Activity. Oxford Review of Economic Policy, 14(2): 18-29.

Autant-Bernard, C. and LeSage, J. P. (2011). Quantifying knowledge spillovers using spatial econometric models. Journal of Regional Science, 51(3):471-496.

Bade, F.-J. (1987). Regionale Beschäftigungsentwicklung und produktionsorientierte Dienstleistungen. Sonderheft 143. Deutsches Institut für Wirtschaftsforschung, Berlin.

Beasley, J.E. (1995). Determining teaching and research efficiencies. Journal of the Operational Research Society, 46, 441–452.

Bonaccorsi, A., & Daraio, C. (2006). Econometric Approaches to the Analysis of Productivity of R{&}D Systems. Handbook of Quantitative Science and Technology Research Handbook of Quantitative Science and Technology Research - The Use of Publication and Patent Statistics in Studies of S&T Systems (pp. 51–74). Springer Netherland.

Brenner, T, & Broekel, T. (2011). Methodological issues in measuring innovation performance of spatial units. Industry and Innovation, 18(1), 7–37.

Brenner, Thomas, & Broekel, T. (2011). Methodological Issues in Measuring Innovation Performance of Spatial Units. Industry & Innovation, 18(1), 7–37. doi:10.1080/13662716.2010.528927

Broekel, T. (2007). A Concordance between Industries and Technologies - Matching the Technological Fields of the Patentatlas to the German Industry Classification. Jenaer Economic Research Papers.

Broekel, T. (2012). Collaboration intensity and regional innovation efficiency in Germany - A conditional efficiency approach. Industry and Innovation, 19(3), 155–179.

Broekel, T, & Brenner, T. (2007). Measuring Regional Innovativeness - A Methodological Discussion and an Application to One German Industry. DIME Working Paper, 2007-13.

Broekel, Tom, & Binder, M. (2007). The Regional Dimension of Knowledge Transfers—A Behavioral Approach. Industry & Innovation, 14(2), 151–175. doi:10.1080/13662710701252500

Carlsson, B., Jacobsson, S., Holmen, M., & Rickne, A. (2002). Innovation Systems: Analytical and Methodological Issues. Research Policy.

Cazals, C., Florens, J.P., & Simar, L. (2002). Nonparametric Frontier Estimation: A Robust Approach. Journal of Econometrics, 106, 1-25.

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

Chen, K., & Guan, J. (2011). Measuring the Efficiency of China's Regional Innovation Systems: Application of Network Data Envelopment Analysis (DEA). Regional Studies, (907217954), 1–23. doi:10.1080/00343404.2010.497479

Cohen, W., & Levinthal, D. (1989). Innovation and Learning: The two Faces of R&D. Economic Journal, 99, 569–596.

Cohen, W., & Levinthal, D. (1990). Absorptive Capacity: A New Perspective on Learning and Innovation. Administrative Science Quarterly, 35(1), 128–152.

Cook, W.D., and Green, R.H. (2004). Multicomponent efficiency measurement and core business identification in multiplant forms: A DEA model. European Journal of Operational Research, 157, 540–551.

Cook, W.D., Hababou, M., & Tuenter, H.J.H. (2000). Multicomponent efficiency measurement and shared inputs in Data Envelopment Analysis: An application to sales and service performance in bank branches. Journal of Productivity Analysis, 14, 209–224.

Cooke, P. (1998). Introduction: Origins of the Concept. In H.-J. Braczyk, P. Cooke, & M. Heidenreich (Eds.), Regional Innovation Systems - The Role of Governances in a Globalized World (pp. 2–25). UCL Press, London.

Cooke, P., Uranga, M. G., and Etxebarria, G. (1997). Regional innovation systems: Institutional and organisational dimensions. Research Policy, 26(4-5):475–491.

Daraio, C., & Simar, L. 2007. Advanced robust and nonparametric methods in efficiency analysis: Methodology and applications. Series: Studies in Productivity and Efficiency. Springer, New york.

Deyle, H., & Grupp, H. (2005). Commuters and the Regional Assignment of Innovative Activities: A Methodological Patent Study of German Districts. Research Policy.

Eckey, H.F., Kosfeld, R., and Türck, M. (2006). Abgrenzung deutscher Arbeitsmarktregionen. Volkswirtschaftliche Diskussionsbeiträge Universität Kassel.

Färe, R., Grosskopf, S., Norris, M., and Zhang, Z. (1994), Productivity growth, technical progress, and efficiency change in industrialized countries, American Economic Review, 81(1), pp. 66-83.

Färe, R., Grosskopf, S., and Roos, P. (1998). Malmquist productivity indexes: a survey of theory and practice, In: Färe, R., Grosskopf, S., Russell, R.R. (Eds.), Index Numbers: Essays in Honour of Sten Malmquist. Kluwer Academic Publishers, Boston.

Farrell, M.J., 1957. The measurement of productive efficiency. Journal of the Royal Statistical Society, Series A, CXX, Part 3, 253-290.

Frenken, K., van Oort, F. G., and Verburg, T. (2007). Related variety, unrelated variety and regional economic growth. Regional Studies, 41(5):685–697.

Fritsch, M. (2003). How and Why Does the Efficiency of Regional Innovation Systems Differ. In J. Br"ocker, D. Dohse, & R. Soltwedel (Eds.), Innovation Clusters and Interregional Competition. Springer, Berlin.

Fritsch, M, & Slavtchev, V. (2011). Determinants of the Efficiency of Regional Innovation Systems. Regional Studies, 45(7), 905–918.

Fritsch, Michael. (2000). Interregional Differences in R&D Activities—An Empirical Investigation. European Planning Studies, 8(4), 409–427. doi:10.1080/713666416

Greif, S., & Schmiedl, D. (2002). Patentatlas 2002 Dynamik und Strukturen der Erfindungstätigkeit. Deutsches Patent- und Markenamt, München.

Greif, S., Schmiedl, D., & Niedermeyer, G. (2006). Patentatlas 2006. Regionaldaten der Erfindungstätigkeit. Deutsches Patent- und Markenamt, München.

Griliches, Z. (1979). Issues in Assessing the Contribution of Research and Development to Productivity Growth. The Bell Journal of Economics, 10(1), 92. doi:10.2307/3003321

Jaffe, A. (1989). Real Effects of Academic Research. American Economic Review, 79(5), 957–970.

Malerba, F., & Orsenigo, L. (1993). Technological Regimes and Firm Behavior. Industrial and Corporate Change, 2(1), 45–74.

Malerba, F., Orsenigo, L., & Breschi, S. (2000). Technological Regimes and Schumpeterian Patterns of Innovation. Economic Journal, 110(463), 388–410.

Nelson, R. R. (1959). The Simple Economics of Basic Scientific Research. The Journal of Political Economy, 67(5), 297–306.

Pedraja-Chaparro, F., Salinas-Jiménez, J, and Smith, P. (1999). On the Quality of the Data Envelopment Analysis Model. The Journal of the Operational Research Society, 50(6): 636- 644

Pavitt, K. (1984). Sectoral Patterns of Technical Change: Towards a Taxonomy and a Theory. Research Policy, 13(6), 343–373.

Schmoch, U., Laville, F., Patel, P., & Frietsch, R. (2003). Linking Technology Areas to Industrial Sectors. Final Report to the European Commission, DG Research, Karlsruhe, Paris, Brighton.

Stern, S., Porter, M. E., & Furman, J. L. (2002). The Determinants of National Innovative Capacity. Research Policy.

Wang, M.W., & Stanley, J.C. (1970). Differential weighting: A review of methods and empirical studies. Review of Educational Research, 40, 663–705.

Zabala-Iturriagagoitia, J. M., Voigt, P., Gutiérrez-Gracia, A., & Jiménez-Sáez, F. (2007). Regional Innovation Systems: How to Assess Performance. Regional Studies, 41(5), 661–672. doi:10.1080/00343400601120270

Industry	Short	Ratio 2004.2008	Rank 2004.2008
Pharmaceuticals	IO12	0.545	1
Basic chemicals	IO9	0.326	2
Medical equipment	IO36	0.286	3
Office machinery and computers	IO27	0.218	4
Signal transmission, telecommunications	IO34	0.179	5
Special purpose machinery	IO24	0.172	6
Measuring instruments	IO37	0.107	7
Electronic components	IO33	0.1	8
Domestic appliances	IO26	0.083	9
Rubber and plastics products	IO16	0.082	10
Motor vehicles	IO41	0.078	11
Soaps, detergents, toilet preparations	IO13	0.072	12
Optical instruments	IO39	0.068	13
Non-metallic mineral products	IO17	0.065	14
Non-specific purpose machinery	IO21	0.063	15
Food, beverages	IO1	0.063	16
Agricultural and forestry machinery	IO22	0.053	17
Energy machinery	IO20	0.052	18
Pesticides, agro-chemical products	IO10	0.051	19
Paper	IO7	0.048	20
Television and radio receivers, audiovisual electronics	IO35	0.047	21
Machine tools	IO23	0.047	22
Furniture, consumer goods	IO43	0.044	23
Other chemicals	IO14	0.043	24
Fabricated metal products	IO19	0.041	25
Electric distribution, control, wire, cable	IO29	0.028	26
Basic metals	IO18	0.026	27
Other transport equipment	IO42	0.024	28
Other electrical equipment	IO32	0.02	29
Industrial process control equipment	IO38	0.018	30
Accumulators, battery	IO30	0.017	31
Lighting equipment	IO31	0.017	32
Textiles	IO3	0.017	33
Electric motors, generators, transformers	IO28	0.016	34
Petroleum products, nuclear fuel	IO8	0.016	35
Wearing apparel	IO4	0.015	36
Tobacco products	IO2	0.011	37
Wood products	IO6	0.008	38
Leather articles	IO5	0.007	39
Weapons and ammunition	IO25	0.005	40
Watches, clocks	IO40	0.005	41
Man-made fibers	IO15	0.003	42
Paints, varnishes	IO11	0.002	43

Table 3: Ratio.2004.2008 for industries