Papers in Evolutionary Economic Geography

08.11

The Bright and Dark Side of Cooperation for Regional Innovation Performance

Tom Broekel and Andreas Meder

Utrecht University
Urban & Regional research centre Utrecht

The Bright and Dark Side of Cooperation for Regional Innovation Performance

- revised version -

Tom Broekel [$], [§]   Andreas Meder [#]

[$] Utrecht University, Utrecht, The Netherlands
[#] Friedrich Schiller University, Jena, Germany

9th December 2010

Abstract

Taking a regional perspective the study analyzes the relationship between collaboration intensity and innovation efficiency. For the empirical assessment, data is employed covering 270 German labor market regions and the Electrics & Electronics industry. Patent co-applications approximate collaboration intensities that are related to innovation efficiency using conditional nonparametric efficiency analysis. The investigation shows that regions with medium levels of intra- and inter-regional collaboration intensities outperform regions characterized by low and very high intensities.

[§] *Corresponding author:* Tom Broekel, Department of Economic Geography, Urban & Regional Research Centre Utrecht (URU), Faculty of Geosciences, Utrecht University, Heidelberglaan 2, 3584 CS Utrecht, The Netherlands. *Phone*: +31 30 253 3206. *Fax*: +31 30 253 2037. *E-mail*: t.broekel@geo.uu.nl

JEL codes: R12, O18, O31
Keywords: regional innovation efficiency, overembeddedness, regional networks, collaboration

1 Introduction

It is frequently emphasized that intensive collaboration can foster innovation performance by triggering collective learning processes (see, e.g., Cooke et al., 1997). Backed by evidence from numerous case studies the innovation-promoting role of collaboration is widely accepted (Cooke and Morgan, 1994; Saxenian, 1998; Asheim and Isaksen, 2002).

However, it is well known that extensive regional collaboration can yield low innovation performance if inter-regional linkages are lacking (Bathelt et al., 2004). Accordingly, intensive collaboration can promote innovation activities while beyond a certain intensity level it may be related to negative effects. The latter becomes especially visible in regional overembeddedness scenarios, which characterize regions with dense intra-regional collaboration but missing outside links (Camagni, 1991).

The empirical picture on this issue is still mixed. The majority of studies suggest a positive relation between high collaboration intensity and innovative success (Arndt and Sternberg, 2000). However, there are also studies that do not find collaboration to play a conductive role for innovation (Fritsch, 2004; Oerlemans and Meeus, 2005).

The paper aims at shedding more light on this issue by taking a regional perspective and investigating how the level of intra- and inter-regional collaboration intensity relates to regional innovation efficiency. For the empirical assessment a conditional nonparametric efficiency analysis is employed. It is particularly suitable for exploring this empirical relation in an unrestricted manner. Moreover, it allows assessing region's particular situations in a very detailed manner providing a useful basis for designing regional policy. The investigation is based on a unique data set covering 270 German labor market regions and the Electrics & Electronics industry in 1999-2002.

The study confirms the existence of an inverted-u shape relationship between the regional levels of collaboration intensity and regional innovation efficiency. Medium collaboration intensity and a balance between intra- and inter-regional collaboration characterize regions with the highest innovation efficiency. Low and very high intensity levels as well as an unbalanced mix of intra- and inter-regional collabo-

ration are found for regions with comparatively low innovation efficiencies.
The paper is organized as follows. Section 2 gives an overview of the literature on collaboration and regional innovation performance. In Section 3, the empirical methodology is described in detail. The employed data approximating intra- and inter-regional collaboration intensities, regional characteristics, and innovation are presented in Section 4. Section 5 describes and discusses the results. Section 6 concludes.

2 Collaboration and innovation

2.1 The relationship between collaboration and innovation

Inter-organizational collaboration in the field of research and development (R&D) is an important supplement to internal R&D activities as it is shown to increase the probability of innovative success (Oerleman and Meeus, 2000; Hagedoorn, 2002). Benefits of collaboration activities are the sharing of risk and costs (Cassiman and Veugelers, 2002), access to complementary knowledge and assets (Teece, 1986), as well as the transfer of knowledge (Eisenhardt and Schoonhoven, 1996). Not surprisingly, there is a clear conclusion in recent literature that firms improve their innovative capabilities by engaging in collaborative R&D projects (Faems et al., 2005).

However, this does not mean that collaboration is always beneficial. Establishment and maintenance of collaborative agreements require efforts and success is not guaranteed (Bleeke and Ernst, 1993). Many collaboration fail, which often goes along with lost investments. On the one hand, imperfect appropriability of knowledge can increases the benefits of collaborative R&D projects. On the other hand, it simultaneously raises the incentives to free ride on partners' R&D efforts (Kesteloot and Veugelers, 1995). Also "learning races between the partners [...], diverging opinions on intended benefits [...] and a lack of flexibility and adaptability" (Faems et al., 2005, p. 240) can induce negative effects of collaboration. Hence, whether the benefits of collaboration are realized, depends on the complementarity of the partners' resources, aims, and working routines.

In particular the positive effects associated to collaborating have also been recognized in Economic Geography. It has been shown that the outstanding innovation performance achieved by a number of regions can be attributed to a location-specific collaboration cultures and regional collective learning processes (see, e.g., Aydalot and Keeble, 1985; Saxenian, 1998; Asheim and Isaksen, 2002). This motivated the development of concepts like the 'innovative milieu' (Camagni, 1991), 'learning region' (Florida, 1995), and 'regional innovation system' (Cooke, 1992). Accordingly, regions characterized by dense regional collaboration tend to realize higher innovation performance than regions with less dense regional knowledge networks (Fritsch, 2004).

However, intensifying collaboration among members of a relatively small group can lead to very dense networks. Eventually this can result in "overembeddedness". Overembeddedness, which is a known problem in network research, refers to a situation where organizations' relations became long-lasting, trust-rich, thick, and eventually redundant (see, e.g., Granovetter, 1985; Uzzi, 1996). In other words, overembeddedness can occur when social aspects "supersede the economic imperatives" underlying collaboration (Uzzi, 1997, p. 59). Social proximity outweighs economic reasoning, which leads to economically sub-optimal collaboration decisions. Accordingly, the economic effects of collaborating are not maximized or can even become disadvantages.

The overembeddedness problem translates to the regional level because individuals can develop a tendency to concentrate on their home region for knowledge exchange (Broekel and Binder, 2007). If this applies to a sufficiently large number of regional organizations, regional knowledge networks may become "too closed and [...] too rigid" (Isaksen, 2001, p. 110). Or as Grabher (1993) puts it: the "ties that bind" may become "ties that blind" (p. 24). Being over-embedded into regional networks, organizations are likely to miss technologies and innovations developed outside their region. Their innovation performance is reduced because they lack access to non-regional knowledge and skills (Camagni, 1991). In the following, such a set-up is referred to as "regional overembeddedness". It is closely related to a regional lock-in, which describes the situation of regional organizations being unable to leave a particular development trajectory, which delivers suboptimal economic results (see, e.g., Grabher, 1993; Martin and Sunley, 2006). While regional overembeddedness can be part of a regional lock-in it only refers to the static description of a particular configuration of collaboration activities.

The same rationale can be applied to organizations' embeddedness into inter-regional knowledge networks. Knowledge networks can span regional boundaries and do not develop automatically when organizations are co-located (Isaksen, 2001). If regional organizations don't provide valuable knowledge or they distrust each other, inter-regional collaboration is the remaining option for collaborative research. Accordingly, regional knowledge networks do not develop when regional organizations are not willing or capable of realizing the efforts needed to for their establishment and maintenance. In the remainder of the paper 'supra-regional overembeddedness' is referred to as a situation in which regional organizations are well embedded into inter-regional knowledge networks but fail to develop intensive regional collaboration. In a related manner, Bathelt et al. (2004) argues that a "cluster which is more or less empty because its important organizations are constantly traveling the world in order to build and maintain an extensive pipeline system [global knowledge networks] will run an obvious risk of becoming less vibrant" (p. 48). Not collaborating regionally implies that firms do not take advantage of the benefits geographic proximity yields for interacting and learning. This includes lower transport and travel costs, easier development of trust, facilitating of collective learning, and the exchange of non-codifiable knowledge (see, e.g., Bathelt et al., 2004). For this reason, regions with low intra-regional collaboration and underdeveloped regional knowledge networks are likely to be outperformed by regions that organizations exploit these benefits.

In summary, in order to sustain high innovation performance firms need to participate in "local buzz" while having simultaneously access to "global pipelines" of knowledge (Bathelt et al., 2004).

A rich empirical literature analyzes the effects of collaboration and knowledge networks on firms' performance (see, e.g., Powell et al., 1996; Uzzi, 1996; Faems et al., 2005; Tsai, 2009). These studies generally find that active collaboration has a stimulating impact. Overembeddeness, which relates to lower innovation performance, is however also found to exist (Uzzi, 1996).

Many studies exploring the effects related to the level of regional collaboration intensity, i.e. the aggregate collaboration behavior of regional organizations, are of qualitative nature. Nevertheless, they offer rich evidence for a positive relationship between this intensity and innovation performance (see, e.g., Saxenian, 1998; Asheim and Isaksen, 2002). However, regional lock-in situations are also observed (Grabher, 1993; Hassink, 2007; Cho and Hassink, 2009), which may imply regional over-embeddedness situations. Fewer quantitative approaches exist that consider multiple regions. For instance, Fritsch (2004) uses data on eleven European regions but finds no "support for the suggestion that collaboration or a relatively pronounced cooperative attitude in a region is conducive to innovation activity" (p. 844). Fritsch and Franke (2004) add to this with their study on three German regions.

In light of this mixed picture, it is the objective of the paper to shed more light on this issue from a regional perspective. More precise, the study investigates the relationship between the level of regional collaboration intensity and regional innovation efficiency. Accordingly, collaboration intensity, regional overembeddedness, and supra-regional overembeddedness are treated as *regional* phenomena referring to the collective behavior of regional organizations.

In light of the above, high regional collaboration intensity can be expected to stimulate regional innovation performance. However, if overembeddedness situations exist very high intensities are related to negative effects. Hence, similar to what has been found at the firm level by Uzzi (1996), the relationship between collaboration intensity and innovation performance is likely to follow an inverted u-shape at the regional level.

3 Estimating regional innovation efficiency

3.1 Nonparametric conditional frontier analysis

The innovation performance of regions is commonly evaluated in a knowledge production function framework (see, e.g., Griliches, 1979; Jaffe, 1989). In this framework, variables representing knowledge inputs are set into a functional relationship with knowledge outputs generated by regional organizations. On this basis, their innovation performance can be perceived of as the efficiency with which knowledge inputs are transformed into innovative outputs (Fritsch, 2003; Brenner and Broekel, 2010).

In this context Bonaccorsi and Daraio (2006); Broekel (2007) advocate the use of nonparametric approaches for the estimation of this efficiency. For this reason regional innovation efficiency is calculated using the robust version of the traditional Free Disposal Hull (called *order-m* in the following) as introduced by Cazals et al. (2002). The principal idea of nonparametric efficiency analysis is that observations are evaluated with respect to a frontier function. The frontier function consists of *best-practice* observations, i.e. observations showing a maximum of output given a certain level of input.[1] Compared to parametric approach these techniques yield a number of advantages. Nonparametric efficiency analyses relax a number of critical assumptions that are inherent to parametric production function approaches (regressions). For example, the risk of model mis-specification is reduced by their nonparametric nature. The estimations also do not assume the existence of a universal (pre-defined) functional relationship between knowledge inputs and innovative output. Instead, the constructed best-practice frontier functions are allowed to differ between regions, which accounts for unique situations characterizing some regions. Efficiency analyses are also useful for the design and evaluation of regional policy. This is firstly, because regions are compared to *best-practice* regions, while in traditional production function approaches the evaluations are done on the basis of *average practice*. Secondly, the analyses give detailed information on a region's particular situation, which can be a valuable input for the design of region-specific policies.

Robust nonparametric frontier approaches conceive of the transformation of inputs into outputs as a probabilistic process. The interest is in the probability an observation (x_0,y_0) is dominated by other observations.[2] According to Cazals et al. (2002) an observation's benchmark (frontier) can be the expected maximal value of output of m randomly drawn (with replacement) observations with equal or less levels of input (output-orientation).[3]

Practically, the efficiency measure of order-m can be computed the following.

[1]This corresponds to an output-oriented version, which has been argued to be more appropriate in this context than the input-orientation (Broekel and Brenner, 2007).

[2]x represents the input and y the output.

[3]The value of m is specified by the researcher. It is a trimming parameter defining the sensibility of the estimation with respect to outliers in the data. Following Bonaccorsi et al. (2005) the level of robustness is set to below ten percent. This means that ten percent of the observations have efficiency values smaller than one.

$Y_1, ..., Y_m$ are m random variables drawn from the conditional distribution function of Y given $X \leq x_0$. The output-oriented order-m efficiency measure $\tilde{\lambda}_m(x_0, y_0)$ is defined for observation (x_0,y_0) as

$$\tilde{\lambda}_m(x_0, y_0) = \max_{i=1,\dots,m} \left\{ \min_{j,\dots,q} \left(\frac{Y_i^j}{y_0^j} \right) \right\} \tag{1}$$

with $Y_i^j(y_0^j)$ being the jth component of Y_i (of y_0 respectively). $\tilde{\lambda}(x_0, y_0)$ is a random variable because the Y_i are randomly drawn. In order to obtain the final $\hat{\lambda}_m(x_0, y_0)$, a simple Monte-Carlo algorithm in which $\tilde{\lambda}_m(x_0, y_0)$ is estimated B times (B being large) (Cazals et al., 2002). The order-m efficiency measure of observation (x_0, y_0) is then defined by

$$\hat{\lambda}_m(x_0, y_0) = E[\tilde{\lambda}_m(x_0, y_0)|X \leq x_0] = \frac{1}{B} \sum_{b=1}^{B} \tilde{\lambda}_m^b(x_0, y_0) \quad . \tag{2}$$

The order-m frontier function is a partial frontier because not all observations are enveloped. This feature makes it less sensible to outliers and statistical noise.
The result of the order-m frontier analysis is a measure of relative efficiency. In the context of the paper, it indicates by how much a region's innovative output has to increase in order for this region to become best practice (efficient) given its level of knowledge inputs. Values of $\hat{\lambda}_m$ smaller or equal to one indicate efficiency and larger values represent inefficiency.
Intuitively, this efficiency measure could serve as dependent variable in a regression approach and by this means its relationship to other variables ('external factors') can be investigated. However, such two-stage procedure, a efficiency analysis that is followed by a regression, is methodologically problematic (Simar and Wilson, 2007). Conditional frontier analyses have been developed for this reason and represent a way to explore the relationship between external factors and efficiency measures (Daraio and Simar, 2007b,a; Badin et al., 2008). These conditional frontier approaches can be used for an unrestricted exploration of variables' influences on efficiency without ex-ante imposing a particular type of relationship. This is particularly important if non-linear relationships are to be expected, which is the case in the context of the present paper.

In practice, two efficiency measures have to be estimated: a *conditional* and an *unconditional*. The unconditional measure has been described above. It compares the relation between an observation's output to the best practice found among observations with equal or less input levels. The same applies to the conditional measure. However, the comparison is done under consideration of (conditional on) one or more external factors z. More precise, an observation's evaluation is biased towards comparisons with observations having similar values in these external factors. The likelihood that an observation is compared to another, therefore depends negatively on the difference between their values of z.
The output-oriented conditional order-m efficiency measure $\tilde{\lambda}^c_m(x, y)$ is defined for observation (x_0, y_0) by:

$$\tilde{\lambda}^c_m(x_0, y_0) = \underset{i|X_i \leq x|K_z(z,z_i),i=1,\ldots,m}{max} \left\{ \underset{j,\ldots,q}{min}\left(\frac{Y^j_i}{y^j_0}\right) \right\} \quad . \tag{3}$$

$K_z(z, z_i)$ is the generalized product kernel proposed by De Witte and Kortelainen (2009). These authors refine the approach by Badin et al. (2008) allowing the estimation of significance levels for the relationship between external factors and the efficiency measures. An appropriate bandwidth for this kernel is selected in accordance to their approach. Similar to the unconditional $\hat{\lambda}_m(x_0, y_0)$, $\hat{\lambda}^c_m(x_0, y_0)$ can be estimated with a Monte-Carlo algorithm or by numerically solving an integral (see De Witte and Kortelainen, 2009).
The central variable in a conditional frontier analysis is Q_z representing the ratio between conditional and unconditional efficiency. Q_z can be set in relation to the values of the external factors. The significance of the relationships between Q_z and the external factors is estimated according to De Witte and Kortelainen (2009) using a bootstrap approach with a mixed kernel function and a data-driven bandwidth selection procedure.
To assess the impact of a variable on efficiency, the relation between Q_z and the external factors are presented in three-dimensional and two-dimensional scatter plots. Nonparametric regressions highlight existing trends in these plots, which detailed picture the relation between the external factors and Q_z (for details see Daraio and Simar, 2007a). In the present setting an increasing regression curve

indicates a positive association between efficiency and the external factor, while a decreasing curve hints at a negative association. Following De Witte and Kortelainen (2009), the local-linear estimator by Li and Racine (2004) is used for the nonparametric regressions. Bootstrapped error bands are constructed to validate the robustness of the regression trends. In this way, regional innovation efficiency is related to the levels of regional collaboration intensity, with the latter being defined as 'external factors'.

Opposed to regression analyses, a feature of this type of analysis is that two efficiency measures are obtained for each region. Accordingly, it is possible to evaluate a particular region's situation in detail with respect to the relationship between the considered external factor and its level of efficiency. This will be demonstrated later in the paper.

The regional level of collaboration intensity is a result of a long development process and it is influenced by many factors, of which many have not been discussed above (i.e. culture, location of region, etc.). It seems reasonable to assume that this intensity changes little over time. The data employed for the empirical analyses covers only four years. Any observed change within this time span is therefore unlikely to reflect 'real' changes in regional collaboration intensity.[4] While the conditional and unconditional efficiency measures are separately estimated for each year the resulting data is pooled and a cross-sectional analysis is conducted for this reason. This significantly reduces the effects of outliers and statistical noise. Moreover, large numbers of observation increase the robustness of the employed nonparametric regressions. However, such approach does not allow inferring on the causal relation between the variables, though. While most theoretical arguments suggest collaboration intensity to impact innovation performance, the opposite is possible as well. For example, innovative firms are more attractive partners for R&D collaboration than firms without a record of successful R&D. The same can apply to firms located in regions that are well known for highly innovative products. In these instances, innovative success induces higher levels of collaboration. The primary focus of the investigation is therefore on the shape of the relationship between collaboration intensity and innovation efficiency.

[4] Indeed, the two later introduced collaboration intensity measures show considerable variance between the years: from year to year they are correlated less than $r = 0.25^{***}$.

3.2 Regional innovation efficiency

For the estimation of regional innovation efficiency the number of innovations regional organizations generate can easily be defined by the innovative output. More problematic is the choice of input factors. This is discussed at great length in Brenner and Broekel (2010). These authors conclude that there is no optimal solution and argue in favor of using a range of efficiency measures, which can be used for assessing the robustness of the results. Following their suggestion, three measures of regional innovation efficiency are defined that differ with respect to the considered set of input factors. First, the innovation efficiency of R&D employees is estimated. Here only two R&D employment variables are used. This approach is most often applied in the literature (see, e.g., Fritsch, 2002; Fritsch and Slavtchev, 2006; Broekel, 2008).

In the second set-up, variables are considered as input factors that are significantly related to R&D employees' innovation efficiency (EFF) estimated in the first set-up. These variables are identified using the conditional frontier analysis in analogy to a stepwise approach in in regression frameworks. In practice this means that all combinations of external factors (variables approximating regional characteristics) are tested for their significance with EFF. The combination is chosen in which the largest set of variables is simultaneously significant. In addition, it is checked if the relationship between each significant variable and EFF is monotone and positive, which is a necessary requirement for a variable to be considered an input factor in an efficiency analysis (see, e.g., Coelli et al., 1998). This second set-up is primarily econometrically motivated and is expected to deliver the statistically most reliable results.

In the final set-up, the set of input variables is extended by all variables that are significantly correlated to the innovation efficiency of R&D employees estimated in the first set-up (EFF). Here, it is controlled for a large range of regional characteristics but some statistically redundant variables might be considered.

4 Data on innovation, R&D, and collaboration

4.1 Patent applications and R&D employment

It is well known that innovation activities differ strongly between industries and sectors (see, e.g., Oerlemans and Meeus, 2005). For this reason the focuses is on a single industry: the manufactures of electrical and electronic equipment (ELEC). The industry is chosen because it is highly innovative and relatively R&D intensive (Pavitt, 1984). Moreover, patenting represents an important property rights protection mechanism for this industry (Arundel and Kabla, 1998). This is crucial because following a common approach in innovation research, the outputs of innovation activities are approximated with patent applications.[5] This ensures that the indicator captures most, or at least a significant share, of this industry's innovations. This industry is also second in terms of patent applications in Germany, which guarantees a sufficient number of patents in most German region.

The units of analyses are the 270 German labor market regions that have been used in related studies (see, e.g., Buerger et al., 2010). These regions are defined by the German Institute for Labor and Employment and reflect the spatial dimension of labor mobility in Germany (Haas, 2000). Moreover, they correspond to spatial constraints in firms' search for cooperation partners (Broekel and Binder, 2007). Hence, a significant portion of knowledge spillovers is captured by this level of spatial disaggregation. For this use of patent data it is also important that an inventor's residence and work place tend to be located in the same labor market region (Greif and Schmiedl, 2002).

The patent application data for the years 2000-2003 approximate the innovative output. It is published by the German Patent Office in Greif and Schmiedl (2002) and Greif et al. (2006). Applications by public research institutes, e.g., universities and research societies (e.g. Max Planck Society), as well as patent applications by private inventors are not included because the data on R&D employment covers only industrial R&D.

R&D efforts are approximated by R&D employment data, which are obtained from the German labor market statistic provided by the German Federal Employment

[5]It is acknowledged that patents rather capture inventions than innovations. However, in order to stay consistent with the literature, the term 'innovation' is used in the paper.

Agency. It covers all employees subject to social insurance contribution. The R&D personnel is organized according to the NACE classification. It is match with patent data organized in 31 technological fields on the basis of the concordance by Broekel (2007). The latter is based on the concordance of Schmoch et al. (2003) and adapts it to the data used here.

Table 3 presents the matched technological fields and NACE codes. The patent applications of the five technological fields assigned to ELEC are summed resulting in a single innovation output measure. This is motivated by the existence of a great number of zeros in most of these fields. A time lag of two years to the R&D employment data is moreover assumed.

The R&D employment of ELEC covers the three two-digit NACE codes DL30, DL31, and DL32. In particular DL30 shows a great number of zero values (147 out of 270 observations in 1999). For this reason, the R&D employees of DL30 are summed with DL31 obtaining two variables approximating R&D employment. To construct a meaningful efficiency measure, all regions with zero R&D employment are excluded, which leaves 258 valid observations per year.[6]

4.2 Regional characteristics

The literature suggests a wide range of regional characteristics that influence firms' innovation and collaboration activities (see, e.g., Feldman and Florida, 1994; Broekel and Brenner, 2010). To control for their effects on regional innovation efficiency, the following variables are created.

Among others the benefits of urbanization show as rich local labor markets, well developed infrastructure, strong local demand, as well as the presence of private and public research facilities (Burger et al., 2007). Urbanization advantages are approximated by population density (POP_DEN). The gross domestic product per capita (GDP) captures the regional demand and the public financial situation. The share of employees with high qualifications (EMP_HIGH) is considered because it measures the quality of local human capital (Weibert, 1999). It also approximates the presence of other high-tech industries. The data for these variables are obtained from the German Federal Institute for Research on Building

[6]For this and for the estimation of correlations the two R&D variables are summed.

(INKAR, 2005).

Moreover, the literature highlights the importance of business services for innovation and collaboration activities (see, e.g., Feldman and Florida, 1994), which is why the variable SERVICE is constructed. It represents location coefficient of the employees assigned to NACE code 74 in a region. Following Laursen (1998) the coefficient is made symmetric by calculating $\frac{SERVICE-1}{SERVICE+1} + 1$. This index ranges from 0 to +2, with one indicating average specialization.

SPEC accounts for the specialization of a region with respect to ELEC. Industrial agglomeration is, among others, argued to stimulate knowledge spill-overs, which in turn fosters innovation performance (Greunz, 2004). It is approximated by the location coefficient of the employees of ELEC, which is also made symmetric as described above.

Large multinational firms often centralize their patenting in their headquarter's regions, implying a potential bias of the regional innovation efficiency. However, the data at hand does not include disaggregated size information on firms with more than 500 employees. For this reason, the average firm size (SIZE) in ELEC is estimated. Contrasting intuition it is good indicator when interpreting the smallest values to indicate the presence of large firms. For example, Erlangen with the huge Siemens plants has the second smallest value, Dresden with the large semiconductor plants of AMD the sixths smallest values, and Munich with the Siemens headquarter shows the 34th smallest value. The top-ten regions with the largest numbers in this variable are all small rural areas in East Germany.

The data used to construct the above variables are taken from the German labor market statistics.

The six regional characteristics presented so far are argued to influence only firms located in one region. In contrast, the effects of the following variables are less regionally bounded. Foremost, this concerns knowledge spill-overs from public research facilities that are sensitive to, but not bounded by geographic distance. Two variables are included that account for the geographic mobility of university graduates because this captures most of the non-collaboration related spillovers between research institutes and firms (Faggian and McCann, 2006). University graduates of engineering (GRAD_ENG) and natural sciences & math (GRAD_NAT) are relevant in this context. The graduates of each German university and technical colleague are obtained from the German Statistical Office (DESTATIS, 2005). Their numbers are aggregated at the regional level.

After obtaining their degree, a certain share of graduates leaves the region in which they studied and move to other regions. Accordingly, the receiving regions benefit from the knowledge created in universities' host regions that "spills over". However, graduates are solely assigned to the host regions in regional statistics. Following the procedure proposed by Broekel and Brenner (2007), the numbers of graduates are distributed across the regions such that a region's probability to obtain another regions' graduates depends positively on its population and hyperbolic negatively on the geographic distance between the regions. In addition, a certain share of the graduates stays in the region of their university. The parameters of the hyperbolic function used for estimating the probabilities are fitted by a maximum likelihood calculation. Data on population counts for five digit postal code areas and empirical findings on the mobility of graduates from Legler et al. (2001) are employed in the procedure. Table 2 summarizes the resulting parameters. To control for size effects, the distributed graduate counts enter the analysis as ratios of regions' total employment.

4.3 Two collaboration intensity measures

For modeling regional collaboration intensity, the analysis relies on the measure developed by Cantner and Meder (2008). This measures aims to capture the regionally aggregated collaboration behavior of organizations abstracted from external restrictions and opportunities to collaborate.[7] Accordingly, the measure is characterized by the absence of what one may call *collaboration potential* effect. This refers to the simple fact that some regions offer more possibilities to collaborate because a greater number of potential collaboration partners are located in it. In this context, potential collaboration partners are those that are active in similar technological fields. This effect should be considered when approximating regional collaboration intensity.

Data on collaboration are obtained from German patent data published by the

[7]However, the data covers only collaboration counts. It is acknowledged though that the (unobserved) quality is a crucial aspect as well.

German Patent Office in the "Patentblatt", which includes data from the German patent office as well as data from the European patent office (DPMA, 2005). On the basis of the published patent applications, two organizations are argued to collaborate if they jointly apply for a patent within the IPCs classes assigned to ELEC. Such measure is a commonly used indicator of inter-organizational collaboration (Guellec and van Pottelsberghe de la Potterie, 2005).

Patents are by no means perfect measures of innovation as well as of collaboration activities (see, e.g., Griliches, 1990; Desrochers, 1998). In ELEC patenting is however important and the majority of innovations are patented (Arundel and Kabla, 1998). Co-application agreements are complex and reasons for firms to engage in patent co-application are manifold (see on this Hagedoorn, 2003). How much of the actual collaboration activities are captured by patent co-applications is unknown and remains a weakness of present study. It seems however reasonable to assume that co-applications represent a lower bound of collaboration activities in a region. If the share of collaboration resulting in co-applications is nearly constant in all regions, co-application based indicators deliver representative results (Cantner and Meder, 2008). This is likely the case if the focus is on a single industry in one national state, which is the case in this paper. The presence of large firms may nevertheless systematically bias this measure as large firms tend to be more engaged in collaboration (Colombo, 1995; Becker and Dietz, 2004).

The collaboration intensity measures are constructed as proposed by Cantner and Meder (2008). In a first step, the national collaboration propensity of ELEC is calculated by dividing the total number of collaborations by the total number of innovations within this industry. In a second step, the patents are assigned to labor market regions according to the inventor principle. They reflect regions' industry-specific technological endowment. Next, the number of collaboration is calculated that can be expected according to this technological endowment. It is estimated by multiplying the number of innovations in ELEC (technological endowment) and the collaboration propensity of ELEC, which has been calculated in step one. This can be regarded as the *collaboration potential*, i.e. the number of expected collaboration given the number of collaboration possibilities and average collaboration intensity. In a final step, for each region the number of observed collaboration is divided by the number of expected collaboration.[8] The resulting index I is made symmetric by $\frac{I-1}{I+1}+1$. The idea is that when observed and expected collaboration measures are equal implying that the index is equal to one, it means that the regional collaboration corresponds to the national average. Lower values indicate intensities below the expected level and values above the opposite. Intra-regional collaboration intensity (INTRA) is constructed on the basis of collaboration between organizations located within the same region. Inter-regional collaboration intensity (INTER) is indicated by the collaborating organizations being located in different German regions. The latter implies that inter-regional collaboration refer to national collaboration. International collaboration are not considered because of too many zero values in the resulting measure. A time lag of two years with the R&D employment data is considered because the collaboration intensity measures are based on patent data.

The intensities of intra- and inter-regional collaboration are only weakly correlated ($r = 0.01^{***}$).[9] Accordingly, organizations' intra- and inter-regional collaboration behavior seem to be independent of each other. Figure 1 and 2 in the Appendix show the two intensities' densities as well as the corresponding contour plot.[10] In addition, both variables' histograms are shown in Figure 6 and 7 in the Appendix. The plots illustrate that the mass of observations are characterized by values around one in INTRA as well as in INTER. In addition, a considerable number of observations show close to zero values in INTRA and values close to one in INTER. The latter ones are regions with comparatively low R&D employment in ELEC. This is not surprising because lacking regional alternatives for collaboration, firms located in such regions need to collaborate across regional boundaries.

For the pooled data of 1999-2002, the correlation between the two collaboration intensities and the numbers of patent applications in ELEC is $r = 0.29^{***}$ for INTRA and $r = 0.08^{***}$ in the case of INTER. Accordingly, they are not biased by the magnitude of regions' innovative output. Similar applies to a potential head-

[8] To ensure proper estimations, a small constant is added to the expected collaboration intensity.

[9] *** denotes a significance level of 0.01, ** of 0.05, and * of 0.1.

[10] For the estimating the density a truncated gaussian kernel is used as proposed by Daraio and Simar (2007a).

quarter effect as the correlation with SIZE is marginal.

In line with the results of Cantner and Meder (2008), East German regions tend to be characterized by lower intra-regional collaboration intensities. They do show slightly higher inter-regional collaboration intensities than West German regions, though. The mean difference for the first is 0.30^{***} and for the second -0.10^{***}.[11] The lower propensity of East German firms to engage in region-spanning collaboration is probably a relict of the reunification. These firms still seem to struggle to connect to global (or at least inter-regional) knowledge networks and tend to search locally for collaboration partners (Beise and Stahl, 1999). The missing of large multinational firms in East Germany might be another reason for the lack of inter-regional connections (Licht and Zoz, 2000).

Table 4 shows some basic descriptives and Table 5 presents the correlation structure of all variables considered in the empirical investigation.

5 Regional innovation efficiency and collaboration intensities

5.1 The estimation of regional innovation efficiency

As pointed out before regional innovation efficiency is estimated in three distinct set-ups, which differ in the set of considered inputs. In the first setup, the only inputs are the two R&D employment variables. The stepwise procedure is employed in the second set-up. Here, GDP is the first regional factor tested for its relation with the innovation efficiency estimated in the first set-up (EFF). While significant at first, it becomes insignificant when EMP_HIGH is added to the external factor set. The latter has the lower p-value, which is why GDP is subsequently replaced with SIZE, SERVICE, POP_DEN, GRAD_ENG, GRAD_NAT, and SPEC. None of these variables gains significance when EMP_HIGH is part of the external factor set. In addition, none undercuts EMP_HIGH's p-value. In fact, all p-values of two simultaneously included variables are very large or even one in most instances, see Table 6 in the Appendix. It suggests that these variables explain more or less the

[11] Significance is based on Wilcoxon rank sum test.

same portion of variance in the efficiency measure, which is supported by their comparatively high correlations (see Table 5 in the Appendix).

Except for few extremely high values, EMP_HIGH is positively monotonic related to innovation efficiency, see Figure 3. Accordingly, the selection criteria are met and this variable is added to the input factor set of the second set-up, which also includes R&D employment.

The share of highly educated employees is a reasonable factor to be considered. First, it accounts for a region's human capital and for knowledge spillovers from other industries. Second and more important, it approximates the presence of other high-tech industries that might also patent into patent classes used to construct the output variable. Considering this variable should therefore increase the reliability and robustness of the empirical results.

It may seem surprising to find only one regional variable being significantly related to R&D employees' innovation efficiency. Using a parametric approach Fritsch and Slavtchev (2008) for example find nine significant regional characteristics. However, contrasting their general innovation efficiency measure, industry-specific data is used here, which reflects only regional disparities in one industry's innovation efficiency. Such industry-specific approaches tend to deliver smaller numbers of significant regional factors as variations in the regional industrial structures are eliminated (see, e.g., Brenner and Broekel, 2010; Broekel and Brenner, 2010). Moreover, the conditional frontier analysis seems to be rather conservative in detecting significances.

In the third set-up, the variables SIZE, GDP, and EMP_HIGH join the two R&D employment variables to form the input factor set because they show significant correlations to the efficiency scores obtained in the first set-up (EFF), see Table 5 in the Appendix.

5.2 The innovation efficiency of German regions

The estimated order-m efficiency scores for the three different set-ups vary little. The lowest rank correlation exists between the first and the third set-up ($r = 0.68^{***}$). Accordingly, the inclusion of additional input variables impacts the spatial distribution of the efficiency scores to a limited extent.

- Figure 5 about here -

For this reason, only the innovation efficiency scores obtained in the second set-up are discussed in more detail because it is empirically most sound and delivers very robust results.

Figure 5 gives an impression on the regional distribution of ELEC's innovation efficiency when considering EMP_HIGH as additional input (EFF_H). The corresponding histogram is shown in Figure 4 and some basic descriptives are included in Table 4 in the Appendix. When looking at the descriptives, the first thing to notice is the magnitude of some (in-)efficiency values: EFF_H's mean is 943 while the median is just 5.02. This distortion is caused by extremely high efficiency scores. Values of this magnitude are induced by zero output but positive input observations.[12] Less than five percent of the observations show values of EFF_H > 50, though. This distortion therefore should not impact the later analysis of the relationship between innovation efficiency and collaboration intensity.

The map in Figure 5 reveals that agglomerations tend to show above average efficiency (e.g., Berlin, Munich), while rural areas are rather inefficient (e.g., Lingen, Cloppenburg). There are however significant exceptions. For example, Hamburg and Cologne are not efficient. Leipzig is even highly inefficient. In contrast, similarly sized regions like Dresden and Nurnberg are efficient. This is not surprising as the latter two are known to be regions with strong industrial activities in Electrics and Electronics.

In average, West German regions are more efficient than East German regions (median difference of -13.95^{***}). While Fritsch and Slavtchev (2008) suspect that primarily differences the industrial structures determine this gap in innovation efficiency, the industry-specific results of the present study suggest that additional factors play a role.

5.3 Robustness of results

Table 7 shows that based on the conditional efficiency analysis both variables, INTRA and INTER, are significantly (at the 0.1 level) related to regional innovation

[12] In the estimation, a constant of 0.001 is added to all regions' output to ensure proper estimations.

efficiency in all three set-ups. Their significance levels are particularly better in the second set-up with EMP_HIGH being an addition input. This meets the expectation that considering EMP_HIGH will reduce some distortion of the efficiency measure potentially caused by the matching of R&D employment and patent data.

- Table 7 about here -

Figure 8, 9, and 10 in the Appendix show the estimated relationships of intra-regional (INTRA) and inter-regional (INTER) collaboration intensity for the three set-ups. The shapes of the obtained surfaces differ little, which indicates that the relationship between collaboration intensity and regional innovation efficiency is quite robust and not very sensitive to the definition of the efficiency measures. In other words, the additionally considered regional factors in the second and third set-up do not significantly impact this relationship. For this reason, in the following only the results of the second set-up are presented because it is the soundest empirical approach.

The three-dimensional plot in Figure 9 in the Appendix illustrates the simultaneous relation of both intensity measures, which show to be curve-linear. More precise, the association has as an inverted u-shape. In addition to the above presented significance levels, the relationships' robustness is further assessed with bootstrapped error bands plotted in Figure 11 in the Appendix. The bands indicate that most parts of the curves are highly robust. Only collaboration intensity below 0.1 and above 1.8 are not reliable in case of INTER, which is a result of few observations falling into this interval.

It has been pointed out before that the collaboration intensities are only weekly correlated with innovative output. Nevertheless, regions with low number of applications may still bias the results. Therefore, the analyses are conducted a second time excluding all observations with less than 10 patent applications.[13] This reduces the number of observations per year to 115. The results do not change by and large, though. Although both measures remain significant at the 0.1 level, the inverted u-shape persists in case of INTRA (Figure 12 in the Appendix). In case of INTER, the downward trend for large collaboration intensities is less pronounced.

[13]The same results are obtained with cut-off points at 5 and 20.

Accordingly, high levels of INTER do not show a negative relation with innovation efficiency in case of regions with few patents. However, few regions correspond to this category implying that this observation might not be very reliable.

5.4 Collaboration intensity and innovation efficiency

While the three-dimensional nicely illustrated the relationship between collaboration intensity and innovation efficiency, a more detailed view can be obtained by slicing the surface and estimating partial regressions. In a similar manner as Daraio and Simar (2007a) the results are divided into six sub-samples on the basis of their INTRA and INTER values. Figure 14 shows the (partial) non-parametric regressions for INTRA and the third of regions with the lowest (dashed), medium (solid), and highest (dotted) values of INTER. In a similar manner in Figure 15 the three regressions for INTER are depicted for low, medium, and high levels of INTRA.

- Figure 14 and Figure 15 about here -

The results clearly indicate that regions with best innovation efficiency are characterized by medium levels of intra- and inter-regional collaboration intensities. In contrast, regions with low innovation efficiency tend to show low or very high collaboration intensities. More precise, efficient regions show balanced intra- and inter-regional collaboration intensities. Regional overembeddedness situations with high levels of intra- but low intensities of inter-regional collaboration are more frequently observed among less performing regions. The same applies to supra-regional overembeddedness situations in which inter-regional collaboration intensity is high but intra-regional collaboration is underdeveloped.

These observations seem to support the idea of collaboration intensity impacting regional innovation efficiency. In particular, situations of overembeddedness are shown to be associated with low innovation efficiency. Accordingly, local buzz and access to global pipelines of knowledge characterize regions with high innovation efficiencies, which implies that firms need to exploit local collaboration in order to benefit from the advantages geographic proximity with collaboration partners. Access to global pipelines of knowledge is similarly crucial to "overcome identified shortcomings in the local knowledge base" (Bathelt et al., 2004, p. 44). Negative effects of regional overembeddedness then might be due to missing rivalry between regional firms, which is crucial for the motivation of firms to innovate (Porter, 1990). In addition, in this situation, firms may over-invest in collaboration activities and eventually suffer from free riding (Cassiman and Veugelers, 2002) or learning races (Faems et al., 2005). Bathelt et al. (2004) also suggest the possibility of "buzz congestion", which is caused by information overload. As a result of very intense intra-regional interactions organizations get exposed to too much information such that they are unable to separate useful from useless information. However, it has been pointed out before that the empirical approach does not allow making inference on the causal relationship between collaboration intensity and innovation efficiency. It just represents the empirical association between the empirical variables. For this reason, differences in innovation efficiency may also induce distinct collaboration intensities. Being highly innovative (innovation efficient) can imply that regional organizations have developed rich and vibrant knowledge bases that make them attractive partners for collaborative projects. For the empirical analysis this means, though, that innovation efficient regions should be characterized by above average collaboration intensities. In contrast, the empirical analysis suggests that highly efficient regions tend to show medium collaboration intensities. For this reason, it might rather be the case that organizations in innovation efficient regions have better capabilities balancing their collaboration activities or that, given their attractiveness as collaboration partners, they can be more discriminating in their choice of collaboration partners.

Accordingly, the observed association can also be caused by regional innovation efficiency inducing particular collaboration intensities or that differences in collaboration intensity impact innovation efficiency. Most likely the two mechanisms are simultaneously effective. The study shows, though, that innovation efficiency and collaboration intensities are related, which confirms firm-level findings by Arndt and Sternberg (2000) but contrasts results by Fritsch (2004), Sternberg and Arndt (2001), and Oerlemans and Meeus (2005). Table 8 summarizes the findings.

- Table 8 about here -

It has been pointed out in Section 3 that in contrast to standard regression approaches the conditional frontier analysis allows for a detailed investigation of regions' particular situations. For instance, Stuttgart is highly efficient with an (unconditional) order-m score of 0.818. In contrast, the region of Minden, is deemed fairly inefficient with a score of 1.569. It means, that Minden needs to increase its innovative output to about 157% its level for becoming innovation efficient. However, when considering these regions' conditional order-m efficiencies, which take into account their intra- and inter-regional collaboration intensities, the opposite picture emerges. Stuttgart's conditional efficiency of 1.975 is outperformed by Minden's 1.01. What does that mean? For the estimation of the conditional efficiency score, regions are more likely compared to other regions with similar collaboration intensities, while the unconditional efficiency compares regions to a random sample. Accordingly, given its level of collaboration intensity Minden is performing well. In contrast, Stuttgart is doing worse compared to regions with similar collaboration intensities. If for the sake of the illustration, collaboration intensity is assumed to influence regional innovation efficiency, this means that Stuttgart's high (unconditional) efficiency is a result of its organizations' favorable collaboration behavior. In Minden, the situation is the opposite. To increase efficiency, changes in regional collaboration intensities are necessary. Indeed, Minden is characterized by low INTRA and INTER values (0.02 & 0.66), while Stuttgart's are comparatively higher (1.02 & 1.05). Moreover, within this framework, benchmark regions could be identified for each region, which may help policy to design adequate regional initiatives to improve both regions' positions.
Another interesting observation concerns differences between East and West German regions. It was previously emphasized that East and West German regions differ in terms of regional innovation efficiency and collaboration intensities. For this reason, it can be expected that parts of East German regions' lower innovation efficiency are related to the higher intra-regional collaboration intensities, which have been shown to go along with lower innovation efficiency. To shed light on this issue, the analysis is conducted once more. However, this time a dummy for regions in East Germany (EAST) is included as external factor joining the two collaboration intensities. Contradicting the above expectations, it gains significance with two collaboration intensity measures being significant as well (p-values: IN-

TRA: 0.018; INTER: 0.003; EAST: 0.001). This implies that the lower innovation efficiency of East German regions is not solely related to higher collaboration intensities. Hence, despite differences between East and West Germany, Figure 13highlights that the inverted u-shape relation between collaboration intensity on innovation efficiency holds in both parts of Germany. However, in East Germany the negative association between INTER and innovation efficiency is much less pronounced suggesting that high levels of inter-regional collaboration are an issue of West German regions.

6 Conclusion

The paper contributed to the literature by providing a quantitative empirical analysis analyzing the relationship between collaboration intensities and innovation efficiency at the regional level. On the basis of 270 German regions and the Electrics & Electronics industry it was shown that a significant relationship exists between the two, which follows an inverted u-shape. It was found that medium intensities of intra- and inter-regional collaboration intensities characterize innovation efficient regions. In contrast, regions with unbalanced intra- and inter-regional collaboration are rather found to be inefficient. More precise, situations in which organizations are over-embedded in regional networks or fail to connect to inter-regional knowledge networks are more frequently found in regions with low innovation efficiencies. Moreover, the observed gap in innovation efficiency between East and West German regions does not seem to be related to higher collaboration intensities in East German regions.

The study has a number of shortcomings that may lead the track for future research. First of all, the study relies on patent data to approximate innovation as well as collaboration behavior. In particular the latter is not unproblematic as there are many (not considered) factors that influence firms' engagement in collaboration and in collaborative patenting (see, e.g., Giuri and Mariani, 2005). Measures based on collaborative patent data should therefore be confronted with collaboration indicators constructed from alternative data sources in future research.

The availability of data also restricted the analysis to a cross-sectional approach that does not allow disentangling the causal relationship between collaboration intensity and regional innovation efficiency. While the theoretical arguments rather suggest that collaboration intensity impacts regional innovation efficiency also mechanism exist that justify the opposite. The extension of the data base to cover a longer time periods may allow for shedding more light on this issue. It is also interesting to take a dynamic perspective because the importance of intra- and inter-regional collaboration intensity are likely to vary between the different stages of regions' and industries' life-cycles (see, e.g., Neffke et al., 2008).

The paper moreover focuses on a single industry: namely, the German Electrics & Electronics industry. Hence, the findings remain restricted in their generality, as the observed patterns might be a characteristic of this particular industry. Future analyses need to be extended covering a wider range of industries.

Collaboration intensity was moreover quantitatively defined without taking into account differences in the qualitative of collaborations. It is however well known that it matters with whom firms collaborate as not all partners may offer valuable and complementary knowledge (Boschma and Iammarino, 2009).

Despite these limitations the study points towards an important issue for the design of regional policy, which aims to foster innovation by stimulating regional collaboration. In light of the study's findings, such initiatives are likely to fail in situations of already high intra-regional collaboration intensity, or in the absence of inter-regional links. It is therefore a prerequisite to carefully analyze the situation of regional organizations and support the type of collaboration they are missing. Stimulating the wrong type of collaboration (intra- or inter-regional) may not only yield ineffective efforts, but can even lead to inferior situations. The presented empirical approach is a good empirical tool for such analyses as it allows for a detailed evaluation of each region's particular situation.

7 Tables & Figures

Distance	$< 50km$ #	$50 < 200km$ #	Share 1999	Share 2000
GRAD_ENG (university)	48.8%	29.8%	41.6%	40.7%
GRAD_ENG (tc)	42.3%	35.5%	58.4%	59.3%
GRAD_NAT (university)	61.2%	14.9.9%	89.7%	89.8%
GRAD_NAT* (tc)	45.4%	36.0%	10.3%	10.2%
* No data available, the shares of all technical graduates taken together are used.				
# Data based on Legler et al. (2001) but adjusted for inner Germany mobility				

Table 1: Graduates Mobility

	Empirical values		Estimation		
Spill-over source	$< 50km$	$200km <$	$< 50km$	$200km <$	α
GRAD_ENG	45,1%	33,1%	44.70%	34.35%	1.4851
GRAD_NAT	60.0%	17.0%	56.34%	29.29%	1.6358
Estimation based on sum of 1999 and 2000 data.					

Table 2: Range of spill-overs of graduates, hyperbolic distribution

Technological fields *	NACE industries**
Time measurement, controls, computing (TF27)	Manufacture of office machinery and computers (DL3
Acoustics, electronic data storage (TF28)	Manufacture of electrical machinery
Nuclear physics (TF29)	and apparatus n.e.c. (DL31)
Electrical engineering (TF30)	Manufacture of radio, television and
Electronics, communication technology (TF31)	communication equipment and apparatus (DL32)
* As defined in Greif and Schmiedl (2002) ** According to the GIC DESTATIS (2002)	

Table 3: Definition of the Electrics & Electronics industry according to Broekel (2007)

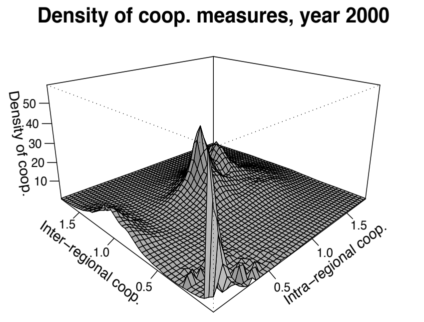

Figure 1: Density of coop. measures

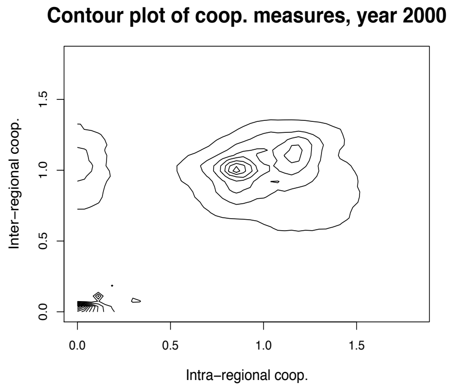

Figure 2: Contour of coop. measures

VAR	mean	sd	median	min	max	skew
PAT	30.30	107.88	8.06	0.00	1763.33	11.01
R&D	656.77	1510.53	196.00	0.00	16548.00	6.41
SIZE	63.00	75.40	38.70	1.67	605.83	3.69
SERV	0.09	0.23	0.05	0.01	3.93	12.89
GRAND_ENG	0.00	0.00	0.00	0.00	0.01	1.45
GRAD_NAT	0.00	0.00	0.00	0.00	0.01	1.23
GDP	37.83	31.67	24.90	12.20	277.90	3.65
EMP_HIGH	11.29	10.40	7.50	2.40	83.50	3.13
SPEC	0.94	1.05	0.56	0.01	7.43	2.58
POP_DEN	849.71	1277.86	247.00	43.00	8495.00	3.03
INTRA	0.69	0.58	0.84	0.00	1.86	0.00
INTER	0.88	0.44	0.98	0.00	1.86	-0.67
EFF	1412.82	7121.45	5.02	0.34	72217.98	6.38
EFF_H	943.41	4798.52	2.60	0.37	53199.73	6.11
EFF_FACT	402.11	3076.25	1.70	0.34	51165.69	9.83
Qz	0.81	0.39	0.98	0.00	2.76	-0.33
Qz_H	0.88	0.33	1.00	0.00	2.41	-0.92
Qz_FACT	0.91	0.30	1.00	0.00	2.59	-0.97
Number of observations: 1036 (pooled data for 1999-2002).						

Table 4: Descriptives

	R&D	SIZE	SERVICE	GRAD_ ENG	GRAD_ NAT	GDP GDP	EMP_ HIGH	SPEC
R&D								
SIZE	0.74***							
SERVICE	0.09	0.04						
GRAD_ENG	−0.21***	−0.08**	−0.14***					
GRAD_NAT	−0.13*	−0.04	−0.08	0.8				
GDP	0.5***	0.29***	0.22**	−0.19**	−0.11*			
EMP_HIGH	0.38***	0.05	0.21**	−0.33***	0.23**	0.69***		
SPEC	0.81***	0.84***	0.02	−0.01	0	0.21**	0.04	
POP_DEN	0.51***	0.22**	0.29***	−0.34***	−0.24**	0.8	0.78***	0.18*
INTRA	0.2**	0.09	0.05	0.03	0.07	0.18**	0.09	0.12*
INTER	0.08	0	0.02	−0.07	−0.04	0.09	0.17**	0.04
EFF	−0.33***	−0.11**	−0.06	−0.01	−0.09	−0.43***	−0.25***	−0.14*
EFF_H	−0.33***	−0.21***	0	−0.08	−0.08	−0.28***	0	−0.24***
EFF_FACT	−0.19**	−0.06	0.08	−0.05	−0.13**	0.04	0.06	−0.15**
Qz	0.37***	0.19**	0.05	−0.02	0.09	0.34***	0.21***	0.24***
Qz_H	0.33***	0.18**	0.02	0.03	−0.09	0.28***	0.16**	0.21***
Qz_FACT	0.25**	0.13*	0	0	0.05	0.17***	0.14***	0.14***
	POP_ DEN	INTRA	INTER	EFF	EFF_ H	EFF_ FACT	Qz	Qz_ EMP
INTRA	0.17***							
INTER	0.15***	0.07**						
EFF	−0.45***	−0.24***	−0.05*					
EFF_H	−0.26***	−0.22***	0.02	0.9***				
EFF_4FACT	−0.04	−0.1***	0	0.68***	0.77***			
Qz	0.35***	0.37***	0.18***	−0.49***	−0.42***	−0.28***		
Qz_H	0.28***	0.26***	0.09***	−0.48***	−0.5***	−0.39***	0.47***	
Qz_FACT	0.21***	0.22***	0.08**	−0.41***	−0.41***	−0.48***	0.34***	0.44***

Table 5: Correlation matrix based on year 2000

1st variable	p-value	2nd variable	p-value
1. GDP	0.001		
2. GDP	1	0.09	
3. EMP_HIGH	0.005		
3. EMP_HIGH	0.859	SIZE	1
4. EMP_HIGH	0.901	POP_DEN	1
5. EMP_HIGH	0.745	SPEC	0.881
6. EMP_HIGH	0.745	SERVICE	1
7. EMP_HIGH	0.001	GRAD_ENG	0.341
8. EMP_HIGH	0.004	GRAD_NAT	0.248

Table 6: P-values of variables tested in stepwise procedure.

Set-up	INTRA p-value	INTER p-value	Inputs
1.	0.073	0.07	R&D
2.	0.006	0.001	R&D, EMP_HIGH
3.	0.001	0.068	R&D, EMP_HIGH, SIZE, GDP

Table 7: Significance of INTRA and INTER

		Intra-regional collaboration intensity		
		low	**medium**	**high**
Inter-regional collaboration intensity	**low**	isolation (-)	A (+)	regional overembeddedness (-)
	medium	B (-)	balanced coop. (++)	C (+)
	high	supra-regional overembeddedness (-)	D (+)	overembeddedness (-)
(+) indicates a positive, and (-) a negative relationship				

Table 8: Relationship between innovation efficiency and collaboration intensity

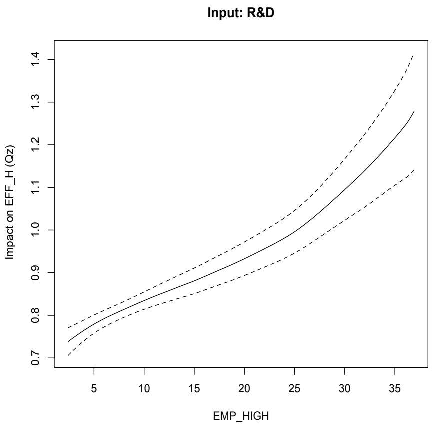

Figure 3: Relation between EFF and EMP_HIGH

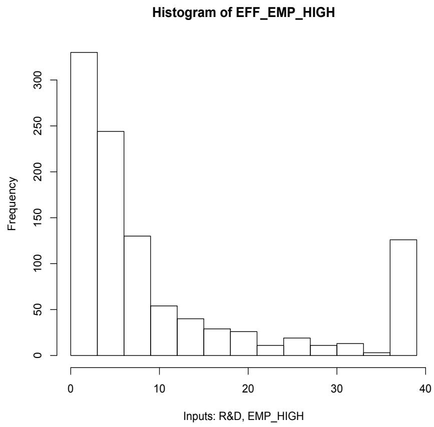

Figure 4: Histogram EFF_H

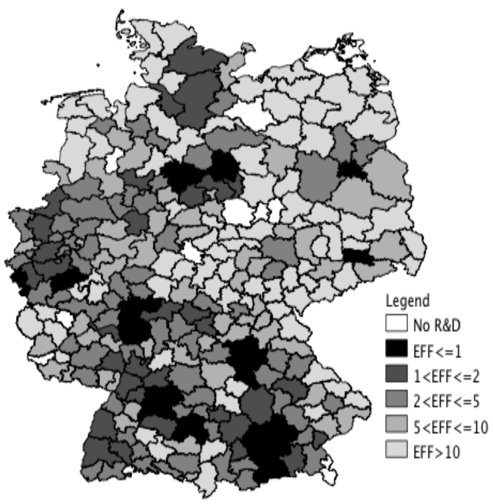

Figure 5: Efficiencies of German LMR

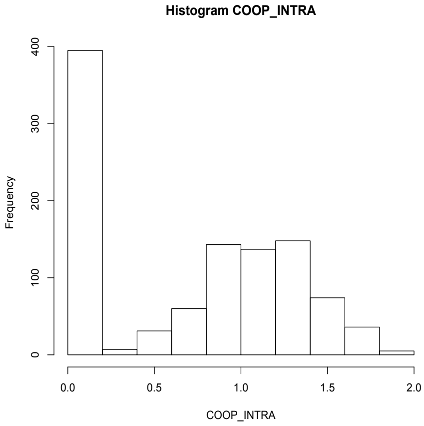

Figure 6: Histogram of INTRA

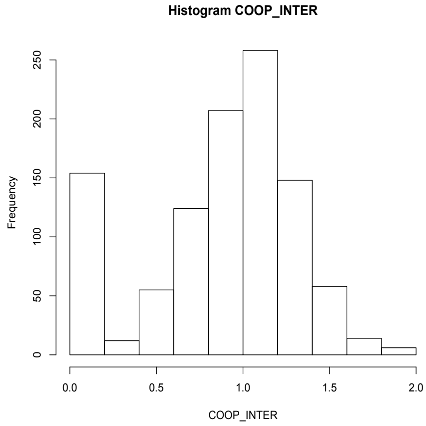

Figure 7: Histogram of INTER

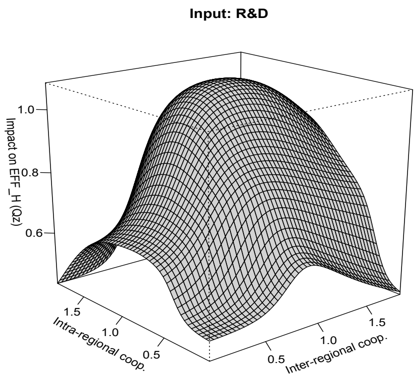

Figure 8: 3d surface: 1. set-up:

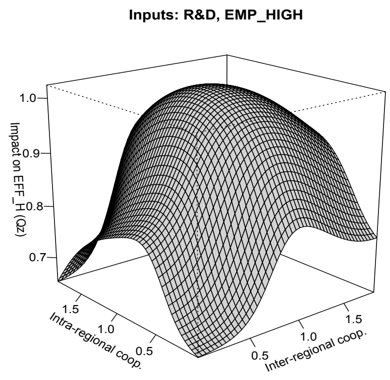

Figure 9: 3d surface: 2. set-up

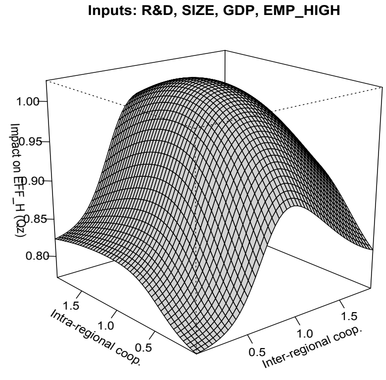

Figure 10: 3d surface: 3. set-up:

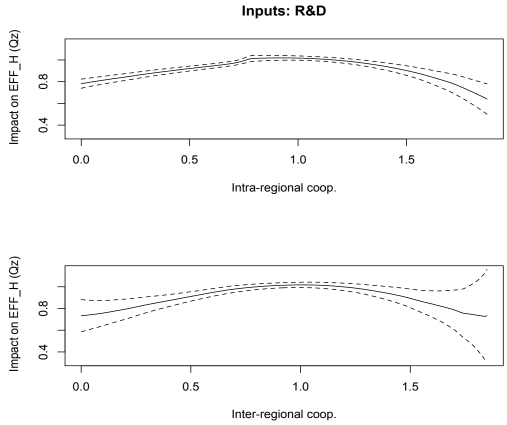

Figure 11: Results: error bands 2. set-up

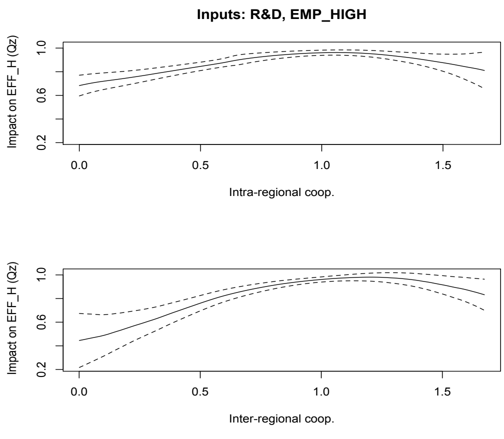

Figure 12: Results: restricted sample

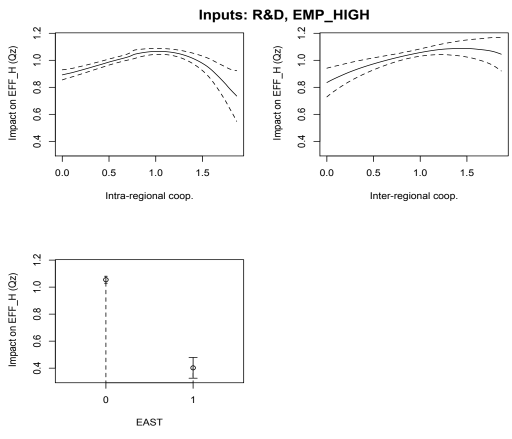

Figure 13: Results: including EAST

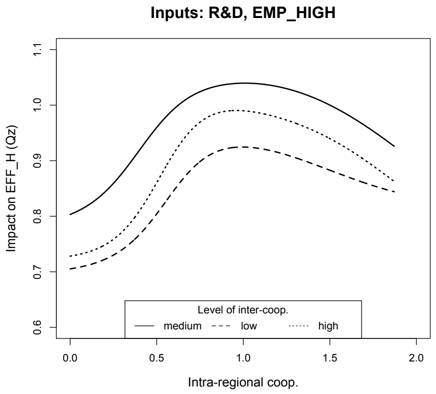

Figure 14: Results: EFF_H relation with INTRA

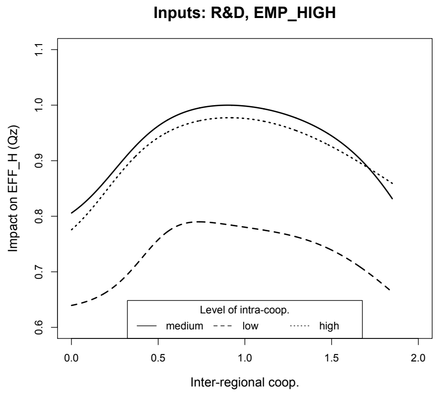

Figure 15: Results: EFF_H relation with INTER

References

Arndt, O. and Sternberg, R. (2000). Do Manufacturing Firms Profit from Intraregional Innovation Linkages? An Empirical Based Answer. *European Planning Studies*, 8(4):465–485.

Arundel, A. and Kabla, I. (1998). What Percentage of Innovations are Patented? Empirical Estimates for European Firms. *Research Policy*, 27(2):127–141.

Asheim, B. T. and Isaksen, A. (2002). Regional Innovation Systems: The Integration of Local 'Sticky' and Global 'Ubiquitous' Knowledge. *Journal of Technology Transfer*, 27:77–86.

Aydalot, P. and Keeble, D., editors (1985). *High Technology Industry and Innovative Environments*. Routledge, London.

Badin, L., Daraio, C., and Simar, L. (2008). Optimal Bandwidth Selection for Conditional Efficiency Measures: A Data-Driven Approach. *Institut De Statistique Discussion Paper, No. 0828*.

Bathelt, H., Malmberg, A., and Maskell, P. (2004). Clusters and Knowledge: Local Buzz, Global Pipelines and the Process of Knowledge Creation. *Progress in Human Geography*, 28(1):31–56.

Becker, W. and Dietz, J. (2004). R&D Cooperation and Innovation Activities of Firms - Evidence for the German Manufacturing Industry. *Research Policy*, 33:209–223.

Beise, M. and Stahl, H. (1999). Public Research and Industrial Innovations in Germany. *Research Policy*, 28(4):397–422.

Bleeke, J. and Ernst, D. (1993). *Collaborating to Comepte: Using Strategic alliances and acquisitions in the global marketplace*. John Wiley.

Bonaccorsi, A. and Daraio, C. (2006). Econometric Approaches to the Analysis of Productivity of R&D Systems. In *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*, chapter 2, pages 51–74. Springer Netherland.

Bonaccorsi, A., Daraio, C., and Simar, L. (2005). Advanced Indicators of Productivity of Universites. An Application of Robust Nonparametric Methods to Italian Data. *Scientometrics*, 66(2):389–410.

Boschma, R. A. and Iammarino, S. (2009). Related Variety, Trade Links, and Regional Growth in Italy. *Economic Geography*, 85(3):289–311.

Brenner, T. and Broekel, T. (2010). Methodological issues in measuring innovation performance of spatial units. *Industry and Innovation*, forthcoming.

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*, 2007-013.

Broekel, T. (2008). From Average to the Frontier: A Nonparametric Frontier Approach to the Analysis of Externalities and Regional Innovation Performance. *Papers in Evolutionary Economic Geography*, 08.04.

Broekel, T. and Binder, M. (2007). The Regional Dimension of Knowledge Transfers - A Behavioral Approach. *Industry and Innovation*, 14(2):151–175.

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

Broekel, T. and Brenner, T. (2010). Regional Factors and Innovativeness - An Empirical Analysis of Four German Industries. *The Annals of Regional Science*, DOI: 10.1007/s00168-009-0364-x.

Buerger, M., Broekel, T., and Coad, A. (2010). Regional dynamics of innovation - investigating the co-evolution of patents, r&d, and employment. *Regional Studies*, 10.1080/00343404.2010.520693.

Burger, M., van Oort, F., and van der Knaap, B. (2007). A Treatise on the Scale-Dependency of Agglomeration Externalities and the Modifiable Areal Unit Problem. *Paper presented at the Workshop on Agglomeration and Growth in Knowledge-based Societies in Kiel, Germany, April 20-21, 2007.*

Camagni, R. (1991). Local "Milieu", Uncertainty and Innovation Networks: Towards a New Dynamic Theory of Economic Space. In Camagni, R., editor, *Innovation Networks: Spatial Perspectives*. Belhaven Stress. London, UK and New York, USA.

Cantner, U. and Meder, A. (2008). Regional and technological effects of cooperation behavior. *Jenaer Economic Research Papers*, 14.

Cassiman, B. and Veugelers, R. (2002). R&D cooperation and spillovers: Some empirical evidence from Belgium. *The American Economic Review*, 92:1169–1184.

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

Cho, M. and Hassink, R. (2009). Limits to Locking-out through Restructuring: The Textile Industry in Daegu, South Korea. *Regional Studies*, 43(9):1183–1198.

Coelli, T., Prasada Rao, D. S., and Battese, G. (1998). *An Introduction to Efficiency and Productivity Analysis*. Kluwer Academic Publishers, Boston.

Colombo, M. G. (1995). Firm Size and Cooperation: The Determinants of Cooperative Agreemetns in Information Technology Industries. *International Journal of the Economics of Business*, 2(1):3–30.

Cooke, P. (1992). Regional Innovation Sytems: Competitive Regulation in the New Europe. *GeoForum*, 23:356–382.

Cooke, P. and Morgan, K. (1994). The Regional Innovation System in Baden-Würtemberg. *International Journal of Information Technology Management*, 9(3-4):394–429.

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. and Simar, L. (2007a). *Advanced Robust and Nonparametric Methods in Efficiency Analysis - Methodology and Applications*. Kluwer Academic Publishers, Boston / Dordrecht / London.

Daraio, C. and Simar, L. (2007b). Conditional Nonparametric Frontier Models for Convex and Non Convex Technologies: a Unifying Approach. *Journal of Productivity Analysis*, 28(1):13–32.

De Witte, K. and Kortelainen, M. (2009). Blaming the exogenous environment? Conditional efficiency estimation with continuous and discrete exogenous variables. *MPRA Paper*, 14034.

Desrochers, P. (1998). On the Abuse of Patents as Economic Indicators. *The Quarterly Journal of Austrian Economics*, 1(4):51–74.

DESTATIS (2002). *Klassifikation der Wirtschaftszweige, Ausgabe 2003 (WZ2003)*. Statistisches Bundesamt, Wiesbaden.

DESTATIS (2005). *Absolventen der Universitäten und Fachhochschulen nach Fächergruppen geordnet*. German Federal Statistics Office, Wiesbaden, Germany.

DPMA (2005). *DEPAROM and Patentblatt CD.* Deutsches Patent- und Markenamt, München.

Eisenhardt, K. M. and Schoonhoven, C. B. (1996). Resource-based view of strategic alliance formation: Strategic and social effects in entrepreneurial firms. *Organization Science*, 7(2):136–150.

Faems, D., Looy, B. V., and Debackere, K. (2005). Interorganizational Collaboration and Innovation: Towards a Portfolio Approach. *Journal of Product Innovation Management*, 22(3):238–250.

Faggian, A. and McCann, P. (2006). Human Capital Flows and Regional Knowledge Assets: A Simultaneous Equation Approach. *Oxford Economic Papers*, 52:475–500.

Feldman, M. P. and Florida, R. (1994). The Geographic Sources of Innovation: Technological Infrastructure and Product Innovation in the United States. *Annals of the Association of American Geographers*, 84(2):210–229.

Florida, R. (1995). Toward the Learning Region. *Futures*, 27:527–536.

Fritsch, M. (2002). Measuring the Quality of Regional Innovation Systems - A Knowledge Production Function Approach. *International Regional Science Review*, 25:86–101.

Fritsch, M. (2003). How and Why Does the Efficiency of Regional Innovation Systems Differ. In Bröcker, J., Dohse, D., and Soltwedel, R., editors, *Innovation Clusters and Interregional Competition*. Springer, Berlin.

Fritsch, M. (2004). Cooperation and the efficiency of regional R&D activities. *Cambridge Journal of Economics*, 28:829–846.

Fritsch, M. and Franke, G. (2004). Innovation, Regional Knowledge Spillovers and R&D cooperation. *Research Policy*, 33:245–255.

Fritsch, M. and Slavtchev, V. (2006). Measuring the Efficiency of Regional Innovation Systems: An Empirical Assessment. *Freiberg Working Papers*, 2006-6.

Fritsch, M. and Slavtchev, V. (2008). Determinants of the Efficiency of Regional Innovation Systems. *Regional Studies*, 10.1080/00343400802251494.

Giuri, P. and Mariani, M. (2005). Everything you always wanted to know about inventors (but never asked): Evidence from the Patval-EU survey. *LEM Working Paper Series*, 2020.

Grabher, G. (1993). The Weakness of Strong Ties: The Lock-in of Regional Development in the Ruhr Area. In Grabher, G., editor, *The Embedded Firm - On the Socioeconomics of Industrial Networks*, pages 255–277. Routledge, London, New York, Reprinted in 1994.

Granovetter, M. (1985). Economic Action and Social Structure: The Problem of Embeddedness. *American Journal of Sociology*, 91(3).

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

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

Greunz, L. (2004). Industrial Structure and Innovation - Evidence from European Regions. *Journal of Evolutionary Economics*, 14(5):563–592.

Griliches, Z. (1979). Issues in Assessing the Contribution of R&D to Productivity Growth. *Bell Journal of Economics*, 10:92–116.

Griliches, Z. (1990). Patent Statistics as Economic Indicators: A Survey. *Journal of Economic Literature*, 28:1661–1701.

Guellec, D. and van Pottelsberghe de la Potterie, B. (2005). Measuring the Internationalisation of the Generation of Knowledge: An Approach based on Patent Data. In *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*, pages 645–662. H. F. Moed and W. Gänzel and U. Schmoch.

Haas, A. (2000). Regionale Mobilität gestiegen. *IAB Kurzbericht*, 4/200:1–7.

Hagedoorn, J. (2002). Inter-firm R&D partnerships: an overview of major trends and patterns since 1960. *Research Policy*, 31:477–492.

Hagedoorn, J. (2003). Sharing intellectual Property rights - An Exploratory Study of Joint Patenting Amongst Companies. *Industrial and Corporate Change*, 12(5):1035–1050.

Hassink, R. (2007). The Strength of Weak Lock-Ins: The Renewal of the Westmünster Textile Industry. *Environment and Planning A*, 39:1147–2265.

INKAR (2005). *Aktuelle Daten zur Entwicklung der Städte, Kreise und Gemeinden, CD-Rom, Berichte 8*, volume 14. Bundesministerium fur Bauwesen und Raumplanung, Berlin.

Isaksen, A. (2001). Building Regional Innovation Systems: Is Endogenous Industrial Development Possible in the Global Economy? *Canadian Journal of Regional Science*, 24(1):101–120.

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

Kesteloot, K. and Veugelers, R. (1995). Stable r&d cooperation with spillovers. *Journal of Economics and Management*, 4:651–672.

Laursen, K. (1998). Revealed Comparative Advantage and the Alternatives as Measures of International Specialisation. *DRUID Working Papers*, 98-30.

Legler, H., Licht, G., and Egeln, J. (2001). *Zur Technologischen Leistungsfähigkeit Deutschlands - Zusammenfassender Endbericht 2000*. Gutachten im Auftrag des Bundesministeriums für Bildung und Forschung.

Li, Q. and Racine, J. S. (2004). Cross-Validated Local Linear Nonparametric Regression. *Statistica Sinica*, 14(2):485–512.

Licht, G. and Zoz, K. (2000). Patents and R&D: An Econometric Investigation Using Applications for German, European and US Patents by German Companies. In Encaoua, D., Hall, B. H., Laisney, F., and Mairesse, J., editors, *The Economics and Econometrics of Innovation*. Kluwer Academic Publishers, Dordrecht.

Martin, R. and Sunley, P. (2006). Path Dependence and Regional Economic Evolution. *Journal of Economic Geography*, 6:395–437.

Neffke, F., Henning, M. S., Boschma, R. A., Lundquist, K.-J., and Olander, L.-O. (2008). Who needs agglomeration? varying agglomeration externalities and the industry life cycle. *Papers in Evolutionary Economic Geography*, 08.08.

Oerleman, L. and Meeus, M. (2000). Firm behaviour and innovative performance: An empirical exploration of the selection-adaptation debate. *Research Policy*, 29:41–58.

Oerlemans, L. A. G. and Meeus, M. T. H. (2005). Do Organizational and Spatial Proximity Impact on Firm Performance. *Regional Studies*, 39(1):89–104.

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

Porter, M. (1990). *The Competitive Advantage of Nations*. The Free Press, New York.

Powell, W. W., Walter, W., Koput, K. W., and Smith-Doerr, L. (1996). Interorganizational Collaboration and the Locus of Innovation: Networks of Learning in Biotechnology. *Administrative Science Quarterly*, 41(1).

Saxenian, A. (1998). *Regional Advantage - Cluture and Competition in Silicon Valley and Route 128.* Cambridge, Mass.

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

Simar, L. and Wilson, P. W. (2007). Estimation and Inference in Two-Stage Semi-Parametric Models of Production Processes. *Journal of Econometrics*, 136:31–64.

Sternberg, R. and Arndt, O. (2001). The Firm or the Region: What Determines the Innovation Behavior of European Firms? *Economic Geography*, 77(4):364–382.

Teece, D. (1986). Profiting from technological innovation: Implication for integration, collaboration, licensing and public policy. *Research Policy*, 15:285–305.

Tsai, K.-H. (2009). Collaborative Networks and Product Innovation Performance: Towards a Contingency Perspective. *Research Policy*, 38(5):765–778.

Uzzi, B. (1996). The Sources and Consequences of Embeddedness for the Economic Performance of Organizations: The Network Effect. *American Sociological Review*, 61(4):674–698.

Uzzi, B. (1997). Social Structure and Competition in Interfirm Networks: The Paradox of Embeddedness. *Administrative Science Quarterly*, 42(1):35–67.

Weibert, W. (1999). *Regionale Determinanten der Innovation.* Nomos Verlagsgesellschaft, Baden-Baden.