




# The relationship of policy induced R&D networks and inter-regional knowledge diffusion

Marcel Bednarz<sup>1</sup> · Tom Broekel<sup>1</sup> 

Published online: 10 May 2019  
© The Author(s) 2019

## Abstract

Knowledge diffusion is argued to be strongly influenced by knowledge networks and spatial structures. However, empirical studies primarily apply an indirect approach in measuring their impact. Moreover, little is known about how policy can influence the spatial diffusion of knowledge. This paper seeks to fill this gap by testing empirically the effects of policy induced knowledge networks on the propensity of inter-regional patent citations. We use patent citation data for 141 labor market regions in Germany between 2000 to 2009, which is merged with information on subsidized joint R&D projects. Based on the latter, we construct a network of subsidized R&D collaboration. Its impact on inter-regional patent citations is evaluated with binomial and negative binomial regression models. Our findings do not indicate that inter-regional network links created by public R&D subsidies facilitate patent citations and, hence, inter-regional knowledge diffusion.

**Keywords** Knowledge diffusion · Subsidized R&D-networks · Gravity model · Inter-regional spillover · Proximity · Patent citations

**JEL classification** L14 · O18 · O33 · O38 · C31 · D83 · O18

## 1 Introduction

To sustain innovation and gain competitive advantage, firms and regions need to innovate constantly, which requires the utilization of internal and external knowledge (Lundvall and Johnson 1994; Sternberg 2000). The utilization of external knowledge requires the diffusion of knowledge between organizations and across space. Since the pioneering work of Hågerstrand (1967) it is known that knowledge does not diffuse frictionlessly within socio-economic systems. Despite technological advantages,

---

✉ Tom Broekel  
t.broekel@uu.nl

<sup>1</sup> Faculty of Geosciences, Utrecht University, Heidelberglaan 2, 3584 CS Utrecht, The Netherlands

geographic distance remains a significant obstacle. Other factors such as cognitive, institutional, organizational, and social distances add to an unequal, selective, and potentially too low diffusion of knowledge (Buisseret et al. 1995; Boschma 2005). Policy has recognized this and established measures seeking to stimulate knowledge diffusion, foremost by subsidizing joint R&D projects (Buisseret et al. 1995; Broekel and Graf 2012). The EU-Framework Programmes are a well-known example for such an initiative (Schemgell and Barber 2009).

Despite such initiatives existing for decades, evidence for a positive contribution to knowledge diffusion remains scarce. While their effects on firms' and regions' innovation activities has been frequently documented (Hewitt-Dundas and Roper 2010; Maggioni et al. 2014; Broekel 2015), little to no evidence exists on their effectiveness for inter-organizational or inter-regional knowledge diffusion. The present paper seeks to fill this gap by using an empirical approach frequently applied to study the diffusion of knowledge in space. Following the work of Jaffe et al. (1993), we used patent citations as indications of inter-regional knowledge diffusion and tested for the effect of joint R&D projects subsidized by the German Federal Government. Hence, besides modelling knowledge diffusion directly, we contribute to the literature by studying the effect of subsidized knowledge networks among regions of a single country, as most existing studies focus on knowledge diffusion at the European level (Maggioni et al. 2011).

In the empirical analysis, we use information on patents and patent citations for 141 German labor market regions between 2000 and 2009. On this basis, technology-specific knowledge diffusion, regions technological relatedness and co-inventor relations are established. Following Broekel and Graf (2012), we construct policy induced networks emerging from the subsidization of joint R&D projects by the German federal government, which are matched to the patent data.

The paper is structured as follows. The second section discusses the mechanism of knowledge diffusion between organizations and in space. The empirical data and variables are described in section 3. Section 4 introduces the empirical approach. The results of the analyses are presented in section 5. Section 6 concludes the paper with a short summary, and discussion on its limitations, future research prospects and political implications.

## 2 Theoretical considerations

### 2.1 Knowledge diffusion, networks, and proximities

It has long been established that knowledge creation and innovation are closely linked to knowledge diffusion. Kline and Rosenberg (1986) emphasize the central role of frequent knowledge exchange between economic actors in their "chain-linked model". Here, innovation processes consist of feedback loops, forward and backward linkages, as well as other interactions between actors internal and external to an organization. In line with this, Schrader (1991) confirms a positive correlation between firms' economic performance and the frequency of informal inter-firm interactions. Zucker et al. (1998) add to this by finding cooperation between biotechnology firms and scientists positively impacting product innovation.

In general, knowledge diffusion can be described by individuals (senders and recipients) (un)intendedly sharing their knowledge via (in)direct communication (Witt et al. 2007). Successful knowledge diffusion depends on the senders' intentions and capacities to communicate (Ibid.) as well as on the recipients' willingness and abilities to recognize and to absorb knowledge (Cohen and Levinthal 1990; Lundvall and Johnson 1994). The ability to assimilate new knowledge tends to be greater "when the object of learning is related to what is already known" (Cohen and Levinthal 1990: p. 131). Therefore, it is positively influenced by the cognitive overlap of two actors' cognitive "interpretation system[s]" (Nooteboom et al. 2007 p. 1017). These arguments are frequently summarized as cognitive (alternatively technological) proximity between actors positively influencing the likelihood and effectiveness of knowledge diffusion. In addition to cognitive proximity, Boschma (2005) argues in favor of four additional proximities that make knowledge diffusion more likely and effective: organizational, institutional, social and geographic proximity. Organizational proximity is a control-related dimension: relations with high organizational proximity are embedded in the same business routines, hierarchies, and value systems, e.g., being part of the same firm (Balland 2012; Loasby 2001). Such circumstances tend to help dealing with uncertainty and opportunism, which in turn facilitates knowledge diffusion (Boschma 2005). Similar conclusion can be drawn when actors are embedded in the same institutional framework, which implies that they share the same norms and formal rules (Boschma 2005).

Hägerstrand (1952, 1965, 1967) promotes the view of geographic distance being an (and maybe "the") explanatory factor for knowledge diffusion. Geographic proximity tends to facilitate face-to-face interactions and mutual trust, which are necessary conditions for successful knowledge exchanges (Howells 2002). However, Boschma (2005) points out that geographical proximity is neither necessary nor sufficient for knowledge diffusion, as other forms of proximity are similarly crucial. Besides geographic proximity, Hägerstrand (1965) also suggests knowledge diffusion being closely related to the embeddedness of actors into social relations, e.g. kinship or friendship. Social interactions consist of multilateral, interdependent and multilevel network relations that may serve as (in)direct knowledge channels (e.g. Tijssen 1998; Maggioni et al. 2007). This is commonly referred to as social proximity (Boschma 2005).

## 2.2 Knowledge diffusion and proximities – What about R&D policy?

Given that one or multiple of these proximities are frequently absent or weakly developed, inter-organizational knowledge diffusion can be expected to be below a social optimum. As this may reduce innovation, policy intervention can be justified (Buisseret et al. 1995). While different approaches exist to deal with this, policy mostly provides monetary incentives in the form of subsidizing joint R&D projects to increase inter-organizational learning and knowledge diffusion (Breschi and Cusmano 2004; Broekel and Graf 2012; Broekel et al. 2015). Usually, organizations can apply for project funding within the scope of policy-defined calls and by providing information on their projects' aims, required resources, partners, and expected outcomes. A well-known example of such an initiative is the EU-Framework Programme (EU-FRP). The EU-FRP has become the most important R&D policy tool of the European Union. For the period 2014 to 2020 it involves about 80 billion €, which are granted in the form of subsidies to R&D projects. For a more detailed discussion see e.g. Breschi and

Cusmano (2004). Here, the subsidies are exclusively granted to projects in which organizations conduct R&D in a collaborative manner. Similar programs also exist at the national level. For instance, in Germany, the Federal Ministry of Education and Research as well as the Federal Ministry of Economics and Energy invest about 3–4 billion € each year as subsidies for project-based R&D. About 30% of the subsidies are granted to collaborative R&D projects (Broekel and Graf 2012). Crucially, cooperating organizations are obligated to write a cooperation agreement in which they grant access to their intellectual property rights that are necessary to conduct the project. In these cases, financial support is only granted when all participants agree on exchanging knowledge within the scope of the project. This includes the use of property rights, access to technical expertise and regular face-to-face meetings (BMBF 2008). Accordingly, subsidized joint R&D projects bear the potential of facilitating if not enabling inter-organizational knowledge diffusion (Broekel and Graf 2012). The rather limited existing research on this issue shows that knowledge generation of regions and organizations is positively associated with participating in subsidized joint R&D projects. For example, Fornahl et al. (2011) discover that biotechnology firms engaged in subsidized collaborative R&D have higher patent output than firms not receiving such funds. Maggioni et al. (2014) and Broekel (2015) confirm this for the regional level. However, we argue that these studies do not provide direct empirical evidence for a direct knowledge diffusion-enhancing effect of subsidized joint R&D projects.

### 2.3 The indirect and direct approaches of analyzing spatial knowledge diffusion

Traditionally, knowledge diffusion studies have investigated the inter-organizational<sup>1</sup> knowledge diffusion by empirically tracking inventions, applications, or products over time and space (see e.g. Hägerstrand 1952, 1965, 1967; Rogers 2003). The primary interest of these studies is to find re-occurring patterns of (spatial) knowledge diffusion and to analyze the extent to which these dimensions represent significant obstacles to knowledge diffusion. More recently, this research tradition, among others, has stimulated the analysis of publication and patent citation patterns, which are seen as indications of knowledge diffusion (Jaffe et al. 1993). By studying the spatial structure of patent citations, these authors show that patents are more likely to cite other patents if their inventors are located in geographic proximity. Breschi and Lissoni (2009) extend this research by using the same approach and identify limited geographic mobility of inventors being primarily responsible for this finding, as they tend to cite their older patents assigned to their previous employers. Since then, the evaluation of patent (and publication) citation has been used frequently to evaluate spatial knowledge diffusion (Peri 2005; Maggioni et al. 2007; Paci and Usai 2009; Hoekman et al. 2009). Interestingly, this, what we call, the “direct approach”, has not been used to study the effectiveness to subsidized joint R&D projects, as vehicles of knowledge diffusion.

When testing the effect of subsidies for R&D projects on inter-organizational knowledge diffusion, researchers seem to be primarily inspired by another literature that focuses on the relevance of spatial knowledge spillovers. In this line of research, studies relate the innovation output of individuals, organizations, and regions to the

<sup>1</sup> For the sake of readability, we exclusively use the term inter-organizational knowledge diffusion. However, the arguments apply in an identical fashion to knowledge diffusion between individuals or regions.

knowledge potentially available to them (e.g. Bode 2004). Inspired by the classic knowledge production function approach, Griliches (1979) and Jaffe (1986) argue that the difference between observed innovation output and knowledge of an innovation-generating entity (inventors, organizations, regions) can partly be explained by the entities use of external knowledge that was absorbed through various mechanisms, i.e. by knowledge diffusing between the entities. In the spatial knowledge spillover literature, researchers extend the individual knowledge input of entities by their potential access to external knowledge. For instance, studies assessing the innovation performance of regions consider a wide range of regional characteristics approximating their knowledge endowment, e.g. presence of high-tech industries, universities, etc. (Broekel and Brenner 2011). This set of factors is extended by variables describing the knowledge endowment of neighboring regions (usually the spatial lag of their innovation output). If the latter show a positive correlation with the innovation output of the focal region (while controlling for its own knowledge endowment), it will be interpreted as a confirmation of knowledge having diffused from the neighboring regions into the focal region and thereby having contributed to its innovation generation (Bottazzi and Peri 2003).

The same approach can be transferred to test for the effect of knowledge networks. In this case, the positions of entities in inter-organizational knowledge networks are plugged into the knowledge production function approach replacing or complementing the spatial lag of other entities' innovation output (Maggioni et al. 2014). The latter is usually described in terms of their local (degree) or global (betweenness / eigenvector) centrality. Larger centralities imply a "better" access to the knowledge potentially diffusing in the network. Hence, if the innovation output (or its change) is found to be statistically related to the centrality measures, it is inferred that the correlation is caused by the better access and ultimately use of the diffusing knowledge.

The before mentioned studies by Fornahl et al. (2011) and Maggioni et al. (2014) use this approach to assess the significance of subsidized knowledge networks for knowledge diffusion. While this "indirect approach" is elegant and has its merits, it has one major flaw in this context: whether knowledge is actually diffusing between entities and is utilized, remains unobserved. This implies that the evidence for an enhancing effect of R&D subsidies on knowledge diffusion, which is largely based on this approach, remains subject to interpretation.

The present study therefore seeks to complement the existing evidence on the effectiveness of R&D subsidies using the "direct approach". Following Jaffe et al. (1993), we use patent citations to approximate knowledge diffusion. While it is also troubled by the unobserved knowledge sourcing and utilization, it is a much more direct approach and avoids many issues of spurious correlation that are likely troubling the indirect approach.

## 3 Data and empirical approach

### 3.1 Modelling knowledge diffusion

We follow Jaffe et al. (1993) and Breschi and Lissoni (2009) and rely on patent citations to model knowledge diffusion. Within the knowledge spillover literature,

patent citations are an often used “paper trail” that “track” knowledge diffusion (e.g. Jaffe et al. 1993; Peri 2005). Patent citations are argued to be an indicator of knowledge transfer and accumulation, as the citing patent builds its knowledge upon a piece of the existing knowledge of the referred patent (Jaffe et al. 1993). According to Howells (2002), patents are a proxy for tacit knowledge flows, since it is necessary to recognize, understand, and recombine the codified knowledge imprinted in the patent.

However, making use of patent citations brings along some issues and limitations that must be carefully considered (Breschi and Lissoni 2004). First, patent citations may only account for knowledge flows that are encodable, commercially exploitable and legally patentable (Criscuoloa and Verspagen 2008). Moreover, the propensity to patent in the first place significantly differs between industries; e.g. only 8% of product innovations are patented in the textile industry, while the share is almost 80% in pharmaceuticals (Arundel and Kabla 1998). Therefore, patents capture only a portion of the created knowledge and correspondingly, their citations only a fraction of the knowledge circulating between innovators (Criscuoloa and Verspagen 2008). Second, patent citations may be a noisy indicator as usually not all citations are made by the inventor himself. Submitted patents will be checked by examiners who probably add further citations where appropriate (Jaffe et al. 1993). Fortunately, the European Patent Office (EPO) specifies which citations have been made by the inventor and which have been added by the examiner. Moreover, it has been argued that, at the regional level (which we will use in the later analysis), these issues are of smaller relevance and patent citations may be a good indicator of possible spillovers, even though it is not sure that these potentials have been realized in all cases (Paci and Usai 2009).

The data on patents and citations is taken from the REGPAT and Citations Database provided by the OECD. In contrast to other patent data sources, patents have already been assigned to the regions of inventors’ residence. Patents that have been developed by more than one inventor have been weighted accordingly (fractional counting). Considering the beforementioned limitations, we exclusively rely on patent citations identified as ‘inventor citations’. Moreover, to diminish biases stemming from intra-organizational citations or inventors’ self-citations, we remove all intra-regional citations. We also remove citations from different inventors and regions that share the same applicant (company) and so will further reduce the likelihood of these citations representing intra-organizational knowledge flows.<sup>2</sup> To reduce the biases stemming from variations in industries’ patent and citation activities, we restrict the analysis to individual technologies, i.e. we only consider technology-specific knowledge flows, that is, we exclusively consider citations between patents belonging to the same IPC-subclass (4-digit level).

The empirical analysis is conducted at the regional level because we are interested in spatial knowledge diffusion (e.g. Maggioni et al. 2014; Broekel 2015). The approach avoids potential biases caused by unclear assignments of patents to applicants. We use the 141 German labor market regions classified by Kosfeld and Werner (2012). Labor market regions are an often-applied regional unit of analysis (e.g. Frenken et al. 2007; Broekel et al. 2015), especially when patent data are used. They ensure that an inventor

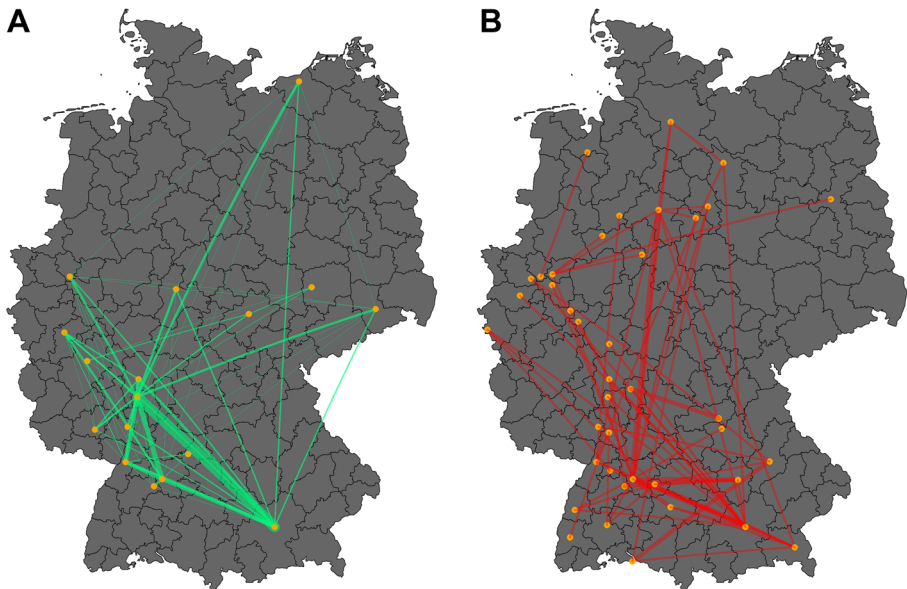
<sup>2</sup> Note that there is no direct assignment of applicants to inventors. This is particularly true, when multiple inventors are given on an applicant’s patents.

's workplace and residence are most likely located in the same region and, hence, in the same unit of analysis (Broekel 2015).

The most important variable is the citation frequency between two regions. The variable, denoted as  $CIT_{i,j,t,s}$  is constructed as follows. In year  $t$ , a knowledge flow exists from region  $i$  to  $j$  when at least one inventor living in region  $i$  cites at least one patent assigned to an inventor in region  $j$  within technology  $s$ . See exemplary Fig. 1a for the patent citations in technology class H04L. More precise,  $CIT_{i,j,t,s}$  is the count of citations from region  $i$  to  $j$ . We employ a five-year moving window for the patent and citation counts to control for yearly fluctuations. For example, the moving window of region  $i$ 's inventors citing region  $j$ 's inventors in 2005 includes citations made from 2005 to 2009. The five-year moving window is consistent with the literature since most patents lose their economic impact within this time frame (Griliches 1979).

### 3.2 Knowledge diffusion channels and regional characteristics

The variable, in which we are most interested is the intensity of subsidized R&D projects, in that organizations of two regions jointly participate. Data on subsidized collaborative R&D projects is obtained from the subsidies catalogue of the German federal government ("Förderkatalog"). According to Broekel and Graf (2012) and Broekel (2015), joint R&D projects listed in this database are intended to stimulate collective learning and knowledge diffusion. For instance, to acquire financial support for a joint R&D project, all participants must agree on certain rules that facilitate collaboration and knowledge exchange. Moreover, all intellectual property rights lying within the scope of the project and that existed before project start must be revealed if demanded by the partners (BMBF 2008). The partners moreover have frequent face-to-



**Fig. 1** Spatial networks of technology H04L "Transmission of digital information" **A** Inter-regional patent citations in 2003 and **B** Subsidized R&D network in 2000. Thickness of lines indicates number of citations or cooperation

face interactions in consortia meetings. They usually have a proven track record of R&D and have convinced application reviewers that their projects (and their partnership) have the potential to become successful.

All projects listed in the subsidies catalogue have been assigned to a “Leistungsplansystematik” (LPS), which is a thematic classification scheme similar to the IPC. The classification includes 22 main classes, for example, biotechnology or nanotechnology. These main classes are further divided into more differentiated subclasses (e.g. photonics (class: I25020)) (Broekel and Graf 2012). Based on these subclasses and the project descriptions, we manually matched 87 4-digit IPC subclasses to about 400 thematic subclasses of the subsidies catalog. Accordingly, we can differentiate between 87 technologies for which information on R&D subsidies and patenting are available (the technologies are listed in the Appendix, Table 3). In line with Broekel and Graf (2012), we constructed the variables  $SUBS.NET_{ij,(t-3),s}$  and  $SUBS.NET_{ij,(t-5),s}$  as follows: a direct subsidized link between two regions  $i$  and  $j$  exists if at least two of their organizations participated in the same R&D project in year  $t$  (see exemplary Fig. 1b for the R&D network of technology class H04L). To model the intensity of their joint R&D work, the variable  $SUBS.NET_{ij,t}$  captures the count of the region’s co-involvement in subsidized R&D projects. To avoid biases in the coding of joint projects, we restricted the analysis to projects with at least three participants. Moreover, we neglected all intra-regional relations as these are of no interest in the context of the paper.

In accordance with the patent data, the subsidies data include information on the projects’ exact starting and ending dates making the construction of a moving window unnecessary. We consider a three- and five-year time lag to the citation variable  $CIT_{ij,t,s}$  as we suspect a time lag given between the cooperation, i.e. the potential knowledge exchange, and the patent citation, i.e. the usage of this knowledge. To make sure that our results are robust, we decide to test for two time lags.

Different types of proximity influence the likelihood of knowledge diffusion between two regions (see Section 2.2).<sup>3</sup> We mainly included these into the models as control variables. The first considered knowledge diffusion channel approximates social proximity. The intensity of social proximity is measured by the strength of co-inventor ties. The co-inventor network has been constructed in two versions denoted as  $CO.INV_{ij,(t-3),s}$  and  $CO.INV_{ij,(t-5),s}$ . A co-inventor relation between regions  $i$  and  $j$  exist when two inventors, one from each region, have worked together on the same patent in technology  $s$ . It is usually interpreted as these two inventors having met personally and hence having established a personal relationship. Accordingly, the frequency of such co-inventor relations signals the strength of the social relations (of inventors) between two regions. We included this variable in a three and alternatively a five-year time lag to the citation variable, to avoid problems of endogeneity and the possibility that citations directly emerge from co-invented patents.

The other proximity types are defined as follows. Geographic proximity is captured by the variable  $DIST_{ij}$ , which is specified by the great circle distance between the centroids of regions  $i$  and  $j$  in kilometers and divided by 100 in order to scale the variable.

<sup>3</sup> Note that the data at hand does not allow to capture institutional proximity.



The construction of technological proximity ( $TECH_{ij}$ ) is somewhat more complex. We follow the early work of Jaffe (1986), Engelsman and van Raan (1994) and Verspagen (1997) and estimate the cosine similarity between regions' technological profiles. In practice, for each of the 141 German labor market regions, we count the applied patents of each IPC subclass resulting in a 141 (regions)  $\times$  629 (technologies) matrix. In the next step, we calculated the cosine similarity between each region and obtain the 141  $\times$  141 technological relatedness matrix for 2000 to 2009, which serves as a technological proximity measure.

Boschma (2005) describes organizational proximity as a control-related dimension: high organizational proximity means that relations are embedded in the same organizational routines and hierarchies. We sought to capture this by the organizational proximity ( $ORG_{ij}$ ) measure. It represents the number of patents that are invented by the same organization with inventors from both regions  $i$  and  $j$ .

The likelihood of inter-regional knowledge flows is also dependent on regions' individual characteristics. The first variable accounts for regions' innovative output ( $PAT_{i,t,s}$  and  $PAT_{j,t,s}$ ), i.e. the sum of technology specific patents created by inventors (factual counting) located in the respective regions  $i$  or  $j$  in time period  $t$  and technology  $s$ . The more a region patents, the more likely it will receive citations and cite other regions' patents.

Moreover, we consider the number of subsidies that regional actors acquired in period  $t$ , ( $SUBS_{i,t,s}$  and  $SUBS_{j,t,s}$ ). More precisely,  $SUBS_{i,t,s}$  and  $SUBS_{j,t,s}$  are the sum of granted subsidized projects to organizations located in regions  $i$  and  $j$  in technology  $s$  and year  $t$ .  $SUBS_{i,t,s}$  and  $SUBS_{j,t,s}$  are included for two reasons; first, large sums of subsidies may indicate the presence of extensive R&D activities, which in turn make citations more likely and second, while a significant coefficient of  $SUBS.NET$  is the most direct confirmation of R&D subsidies enhancing inter-regional knowledge diffusion, the coefficients of  $SUBS$  can still give some indication as to whether the subsidization of R&D activities will lead to larger citation activities or to research that is more frequently cited.

Descriptive statistics for all variables are listed in Table 1.

**Table 1** Descriptive statistics

| Variables               | Mean  | SD    | Median | Min  | Max   |
|-------------------------|-------|-------|--------|------|-------|
| $CIT_{ij,t,s}$          | 0.226 | 0.816 | 0      | 0    | 28    |
| $CO.INV_{ij,(t-3),s}$   | 0.207 | 1.771 | 0      | 0    | 81    |
| $CO.INV_{ij,(t-5),s}$   | 0.23  | 1.668 | 0      | 0    | 84    |
| $SUBS.NET_{ij,(t-3),s}$ | 0.014 | 0.351 | 0      | 0    | 47    |
| $SUBS.NET_{ij,(t-5),s}$ | 0.012 | 0.282 | 0      | 0    | 45    |
| $DIST_{ij}$             | 3.56  | 1.86  | 3.43   | 0.16 | 9.52  |
| $TECH_{ij}$             | 0.410 | 0.173 | 0.41   | 0    | 0.92  |
| $ORG_{ij}$              | 0.007 | 0.03  | 0      | 0    | 6     |
| $PAT_{i,t}$             | 6.846 | 15.01 | 0      | 0    | 207.1 |
| $PAT_{j,t}$             | 7.159 | 15.75 | 0      | 0    | 207.1 |
| $SUBS_{i,t}$            | 0.571 | 1.305 | 0      | 0    | 46    |
| $SUBS_{j,t}$            | 0.597 | 1.268 | 0      | 0    | 46    |
| $CIT.LAG_{ij,t}$        | 0.109 | 0.586 | 0      | 0    | 28    |

### 3.3 Empirical modelling

Our aim is to explain the intensity of knowledge flows ( $CIT_{i,j,t,s}$ ) between regions with the variables introduced above and particular emphasis on the impact of R&D subsidies. In a common manner, we employed a gravity model approach. Based on the work of Isard (1954) and Tinbergen (1962), the gravity approach is a conceptual framework frequently used in investigations of trade and migration flows as well as tourism and commuting interactions (Burger et al. 2009). It has also previously been used to estimate knowledge flows between regions (see Peri 2005; Maggioni et al. 2007; Paci and Usai 2009; Hoekman et al. 2009).

The basic expression of a gravity model can be written as follows (Burger et al. 2009; Hoekman et al. 2009):

$$I_{i,j} = \frac{K^{\beta_1} MASS_i^{\beta_2} MASS_j^{\beta_3}}{DISTANCE_{i,j}^{\beta_4}}$$

where  $I_{i,j}$  is the intensity of interaction between two entities, i.e. the amount of knowledge flows between regions  $i$  and  $j$ . The “masses” of origin  $i$  and destination  $j$  are individual regional characteristics potentially influencing this interaction intensity. In this paper, the number of regional patents ( $PAT$ ) and subsidies ( $SUB$ ) represent the “MASS” variables.

$DISTANCE_{i,j}$  represents the dyadic relations (e.g. geographical distance) between regions  $i$  and  $j$ . Crucially, the model is not restricted to two masses and one distance variable, as in practice; it is logarithmized, transforming the equation into a standard (log)linear regression equation, which can include many more explanatory variables (see for a detailed discussion Broekel et al. 2014).

Our data sample consists of patent citation relations among 141 German labor market regions in 87 IPC classes over a period of ten years (2000–2009). This gives us slightly more than 16 million observations. 34,487 of these are positive ( $CIT_{i,j,t,s} > 0$ ). Notably, we defined a positive case if at least one citation occurred between a pair of regions in any of the ten years. Fitting a regression model to this vast amount of data with just 0.2% non-zero observations would lead to methodological as well as computational difficulties. We therefore reduce the data by randomly matching one positive case with two negative cases. More precise, each positive case of  $CIT_{i,j,t,s} > 0$  is matched with two zero-cases of  $CIT_{i,j,t,s} = 0$ , whereby one of these zero-cases is randomly drawn from the zero-cases of the citing region  $i$  and one from the set of zero-cases of the cited region  $j$ . Hence, our zero-cases share many characteristics of the positive cases.

The matched sampling has 255,493 observations of which 34,487 are positive ( $CIT_{i,j,t,s} > 0$ ), i.e. 13.5% of the sample. The ratio is less than 1:3 because the figure is estimated considering all years, while  $CIT_{i,j,t,s} > 0$  might only be positive in one.

Given that our dependent variable,  $CIT_{i,j,t,s}$ , is a count variable, Poisson or negative binomial regression models are generally appropriate. We decided to calculate a binomial and a negative binomial regression. The binomial regression will provide insights into the probability of inventors in two regions citing each other at all. The negative binomial part seeks to explain the variance in the citation intensities for the sample of observations with at least one citation between 2000 and 2009.

The regressions utilize the panel structure of our data, which allows for including technology- and time-fixed effects controlling for unobserved time invariant effects. Moreover, we also included the lagged version of our dependent variable ( $CIT.LAG_{i,j,t,s}$ ) to capture any remaining unobserved and time-variant structures. It is lagged by six years to the dependent variable ensuring no overlap in the citations used in its construction.

## 4 Empirical results

We estimated four regression models in total: two binomial models distinguishing between inter-regional citations and no inter-regional citations and two negative binomial models explaining the magnitude of inter-regional citations, both with varying time lags of three and five years. The results are presented in Table 2 (3-year time lag)

**Table 2** Results of the fixed-effects, binomial and negative binomial (count) regression models (3-year time lag)

| Variables                      | Binomial model       | Count model          |
|--------------------------------|----------------------|----------------------|
|                                | Estimate (SE)        | Estimate (SE)        |
| Policy                         |                      |                      |
| SUBS.NET <sub>ij,(t-3),s</sub> | 0.011<br>(0.013)     | -0.007<br>(0.007)    |
| SUBS <sub>i,t,s</sub>          | 0.013***<br>(0.003)  | 0.005***<br>(0.002)  |
| SUBS <sub>j,t,s</sub>          | 0.021***<br>(0.003)  | 0.002<br>(0.002)     |
| Proximities                    |                      |                      |
| CO_INV <sub>ij, (t-3),s</sub>  | 0.109***<br>(0.005)  | 0.012***<br>(0.001)  |
| DIST <sub>ij</sub>             | -0.098***<br>(0.004) | -0.017***<br>(0.003) |
| TECH <sub>ij</sub>             | 1.816***<br>(0.038)  | 0.124***<br>(0.027)  |
| ORG <sub>ij</sub>              | -0.015***<br>(0.004) | 0.005***<br>(0.001)  |
| Regional properties            |                      |                      |
| PAT <sub>i,t,s</sub>           | 0.017***<br>(0.0005) | 0.004***<br>(0.0003) |
| PAT <sub>j,t,s</sub>           | 0.014***<br>(0.0005) | 0.003***<br>(0.0002) |
| Control                        |                      |                      |
| CIT_lag <sub>ij,t</sub>        | -0.231***<br>(0.011) | 0.046***<br>(0.003)  |
| Year Dummies                   | Yes                  | Yes                  |
| Technology Dummies             | Yes                  | Yes                  |
| AIC                            | 163,794              | 102,695              |
| 2x Log-likelihood              |                      | -102,526             |
| Observations                   | 255,493              | 34,487               |
| Non-zero obs                   | 34,487               | -                    |

\* significant at 90% level, \*\* significant at 95% level, \*\*\* significant at 99% level

Table 5 in the Appendix (5-year time lag). The results are very robust across all models. We will therefore interpret all outcomes at once.

Mostly, our control variables behave according to our expectations. Increasing geographical distance ( $DIST_{ij}$ ) between two regions decreases the probability of positive citation counts (binominal model) and the magnitude of inter-regional citations (count model). This confirms the geographically localized nature of knowledge spillover (Jaffe et al. 1993; Storper and Venables 2004). Technological proximity ( $TECH_{ij}$ ) has a positive and significant impact on the existence of inter-regional citation in general (binominal model) and on their intensity (count model). The finding is in line with those of Peri (2005) and supports the idea of actors being more likely to seek and absorb new knowledge in cognitive proximity to what they already know (Cohen and Levinthal 1990). Contrary to our expectations, organizational proximity ( $ORG_{i,j,t}$ ) mostly obtains a significantly negative coefficient in the binomial model. It means that region pairs with strong organizational linkages are less likely to cite each other. We suspect that this finding might be caused by our data cleaning in which we removed all citations associated with patents of the same applicant and with the inventors residing in two or more regions. Accordingly, in the construction of the dependent variable, we removed intra-organizational citations, which are closely related to patents with multiple inventors and one applicant, which provide the underlying information for  $ORG_{i,j,t}$ . In light of this, it is somewhat surprising to see  $ORG_{i,j,t}$  gaining a significantly positive coefficient in the count model, suggesting that once citation links are established between regions, organizational proximity facilitates these and supports knowledge diffusion.

We interpret this contradictory finding as an indication of large organizations acting as gatekeepers (Giuliani and Bell 2005). Sometimes they help other organizations in establishing inter-regional linkages, which they, for various reasons, are not able to build otherwise. In network science, this corresponds to the triadic closure argument: partners are likely to connect to partners of their partner (e.g. Granovetter 1973). Accordingly, large organizations with facilities in multiple regions may stimulate other actors in these regions to interact with each other. However, our results suggest that this effect only comes into play when at least one organization other than the large one has managed to establish such a link. As  $ORG_{i,j,t}$  is primarily included as a control variable, we refrained from exploring this issue in greater detail and leave it to future research.

The findings for our measure of social proximity clearly confirms the crucial role this plays for spatial knowledge diffusion.  $CO.INV_{i,j,(t-3),s}$  and  $CO.INV_{i,j,(t-5),s}$  are both significantly positive in the binomial and count model. Hence, as Hägerstrand (1965) proclaimed social networks are crucial for diffusing information about new knowledge and for subsequently spreading this knowledge.

The results for the individual patenting activities of citing ( $PAT_{i,t,s}$ ) and cited regions ( $PAT_{j,t,s}$ ) are highly significant and have a positive relationship with citations in the binomial and count model. Thus, confirming a size effect, regions with a larger patent output have a higher likelihood to cite and to be cited.

The last control variable,  $CIT.LAG_{i,j,t,s}$  measures the citation intensity between regions in the previous period. It is significantly negative in the binomial model and significantly positive in the count model. Accordingly, if inventors in region  $i$  cite patents of inventors in region  $j$ , it is unlikely that they cite each other again in the subsequent period. However, in case they do, their citation intensity will grow. With

some speculation, we can interpret this as a kind of “probing” behavior. Inventors will tap into other regions’ knowledge bases to solve specific problems. However, if the problem is solved, their interest in keeping these relations vanishes. In some cases, though, the regions’ knowledge bases turn out to be complementary, leading to increasing cross-citations and intensifying knowledge exchanges. The observation might also be related to the vary erratic patenting numbers characterizing many (and particularly smaller) regions (Burger et al. 2009), which translate into erratic citation numbers. In the case of regions with few patents, a decrease in these likely translates into the vanishing of many inter-regional linkages created by previous patents’ citations. While the decrease in patents will also lower inter-regional citation frequencies of regions with many patents, their cited patents’ set of regions is likely to remain intact due to the higher levels of citation frequencies.

Our results clearly support the idea of proximities and general economic as well as technological structures shaping the spatial diffusion of knowledge. The question remains, if policy can impact the diffusion by subsidizing R&D projects, which provide a framework for inter-organizational learning.

The variables  $SUBS.NET_{i,j,(t-3),s}$  as well as  $SUBS_{i,t,s}$  and  $SUBS_{j,t,s}$  answer this question.  $SUBS.NET_{i,j,(t-3),s}$  (and  $SUBS.NET_{i,j,(t-5),s}$ ) remain insignificant in all models. Accordingly, joint projects with participants from multiple regions do not facilitate inter-regional knowledge flows, at least when these are measured by patent citations three or five years after the project. This clearly contradicts our expectations and contrasts with the findings of Fornahl et al. (2011) and Broekel (2015), which, however, employ an indirect empirical approach. They investigate the relation between organizations and regions probable exposure to knowledge flowing through subsidized joint R&D projects and their innovation output. With the more direct approach used in this paper, we fail to replicate their findings. There are multiple reasons that may cause this result. First, partners in subsidized joint R&D projects may use these subsidies as windfall gains by using subsidies to strengthen already existing collaboration, which knowledge exploitation potential has already depleted. Second, many subsidized R&D projects may also not result in any patents and, accordingly, citations. Third, subsidized joint projects may fail in establishing relationships that outlast the length of the project. While the knowledge exchange might happen during the project, the inventive process might exceed projects’ durations (which is, on average, close to 36 months) and our time lags of three and five years. Fourth, the collaborative feature of joint R&D projects covered by our data may simply be insufficient for significant knowledge exchange. Clearly, further work, which may have more detailed data available, may shed additional light on these issues.

More promising results (from a policy perspective) are obtained for the number of subsidized projects acquired by regional organizations. The variables  $SUBS_{i,t,s}$  and  $SUBS_{j,t,s}$  are significantly positive in all binominal models, i.e., the more projects acquired by regional organizations, the more likely are these regions’ inventors citing and getting cited. However, this is not consistently true in the count models. Here, only  $SUBS_{i,t,s}$  is positively significant, i.e., the number of subsidized projects in the citing region. The number of projects acquired by the cited region ( $SUBS_{j,t,s}$ ) is mostly insignificant. Accordingly, organizations in regions that are more successful in project acquisition are more likely to create (patent) output that refers to the work of inventors outside their region. Hence, the subsidization of projects seems to stimulate external

knowledge sourcing. Unfortunately, whether this sourcing goes beyond already existing contacts cannot be tested. The insignificance of  $SUBS_{j,t,s}$  implies that (subsidized) research of cited regions does not produce findings that are recognized and utilized by actors outside the region at an above average rate. Put differently, if citations are interpreted as an impact indicator, their research's impact does not seem to exceed the average. Hence, subsidies are granted to organizations in regions, which research impact is not outstanding.<sup>4</sup> This is in line with Broekel et al. (2015) who discovered that, in contrast to the EU-Framework Programmes, research excellence at the regional level does not seem to be the primary allocation determinant of R&D subsidies granted by the German federal government.

## 5 Conclusion

Despite being one of its central aims, few attempts have been made at evaluating the contribution of subsidies for joint R&D projects on inter-organizational and inter-regional knowledge diffusion. Most existing studies follow a knowledge production function approach to test for such policy's effectiveness (Maggioni et al. 2014; Broekel 2015; Broekel et al. 2015). We argued that the traditional literature on knowledge diffusion offers a more direct approach for assessing the contribution of R&D subsidies to inter-regional knowledge diffusion. By following the work of Jaffe et al. (1993) and others, we used patent citations as indication of knowledge diffusion, which are related to subsidies for joint R&D projects by the German national government. We aggregated the data to the level of German labor market regions during the years 2000 to 2009. While our results confirmed the diffusion hampering effects of different kinds of distances (cognitive, geographical, social, organizational), we did not find clear evidence of subsidies significantly stimulating inter-regional knowledge diffusion. While they seem to enhance organizations' ability to source knowledge in other regions, subsidizing inter-regional collaboration does not increase the intensity of subsequent patent citation intensity. Accordingly, we fail to confirm existing evidence obtained by the indirect, knowledge production function type approach (Ponds et al. 2009; Fornahl et al. 2011; Broekel 2015).

While our results fail to deliver support for project-based R&D subsidies achieving one of their primary objectives, there are a number of shortcomings that put our results into perspective. These may also serve as possible starting points for future research. First, by aggregating the data at the regional level, we reduced some of the inherent limitations of patent data. However, when data are available, the organizational level is certainly more appropriate for such kind of analyses. Consequently, the findings must be interpreted with caution, as there is the threat of aggregation biases or "ecological fallacy" (Downs and Mohr 1976: p. 707). Second, while providing more general results, this and similar studies also lack the detailed insights, which can be obtained with qualitative research. For example, the "innovation biographies" approach by Butzin (2009) may be used to study the impact of R&D subsidies on knowledge diffusion in more depth. Third, the study also shares the well-known and frequently discussed limitations of patent data (Griliches 1998). Particularly,

<sup>4</sup> Note that one must be very careful with this interpretation as it is prone to an ecological fallacy trap. We do not say (and find) that the research of organizations receiving R&D subsidies is "just" average or not excellent. Our results exclusively refer to the regional level.

patent citations may only account for knowledge flows that are encodable, commercially exploitable and legally patentable (Criscuolo and Verspagen 2008). Thus, our paper only considered knowledge flows that led to new patents and cite the corresponding ones. However, an organization might develop a new product without patenting it, as the organization uses other ways of securing their intellectual property (Cohen et al. 2000). In this case, knowledge diffuses from one region to another without leaving a trail in patent data. The extent to which innovations and thereby knowledge diffusion are covered by patent data varies significantly between industries (Arundel and Kabla 1998). This makes a comparison of patent-based observations difficult across industries. In order to minimize the likelihood of an inter-industry bias in the use of patent data as indication of knowledge diffusion, we decided to consider exclusively technology specific patent citations excluding citations between technologies. Future studies might be able to tackle this issue in other ways and, hence, might be able to exploit the full breath of patent citations. In addition, future research should make use of alternative data sources to measure knowledge diffusion. One way could be to relate firms' product portfolio diversification to their cooperation activities, as their diversification might be the consequence of obtaining and transforming new knowledge. The study's fourth shortcoming is the consideration of only one specific type of subsidy, i.e., that granted by two ministries of the German federal government. Hence, our findings remain restricted to these programs, as other subsidization schemes such as the EU-Framework Programmes or local policy initiatives might induce different processes and be more effective. Clearly, more research is needed in this direction.

In addition to these shortcomings, there might be other processes that might be responsible for our finding of R&D subsidies not enhancing inter-regional knowledge diffusion. It has been frequently shown that joint R&D projects tend to bring together organizations in geographic, cognitive, organizational, social, and institutional proximity (Breschi and Cusmano 2004; Scherngell and Barber 2009, 2011; Balland 2012, Broekel and Hartog 2013a, 2013b). Yet, these are precisely those constellations that are most likely to emerge without subsidization. It can therefore be argued that to stimulate knowledge diffusion, policy should try to stimulate interactions among organizations that are rather unlikely to interact in the first place. For instance, if subsidized collaboration primarily connect partners at low cognitive distances, their learning potentials are rather small, implying relatively low possibilities for mutual referencing (e.g. in form of patent citations) because they are already familiar with each other's work. Similar arguments apply in the case of organizations sharing a long history of collaboration, which also seem to team up frequently in subsidized joint R&D projects (Breschi and Cusmano 2004).

In summary, our paper marks an additional step towards a better understanding of the effects and effectiveness of subsidizing joint R&D projects. Yet, as the discussion shows, the study opens a barrage of additional issues that hopefully will be addressed by future research.

**Acknowledgments** The authors are thankful for the helpful comments of two anonymous reviewers. Moreover, special thanks go to Lars Mewes for his continuous support and help.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest.

## Appendix

**Table 3** Included IPC subclasses and their description

| IPC-subclass | Description   |
|--------------|---|
| A01C         | Planting; Sowing; Fertilising   |
| A01G         | Horticulture; Cultivation Of Vegetables, Flowers, Rice, Fruit, Vines, Hops, Or Seaweed; Forestry; Watering  |
| A01H         | New Plants Or Processes For Obtaining Them; Plant Reproduction By Tissue Culture Techniques   |
| A01K         | Animal Husbandry; Care Of Birds, Fishes, Insects; Fishing; Rearing Or Breeding Animals, Not Otherwise Provided For; New Breeds Of Animals           |
| A01N         | Preservation Of Bodies Of Humans Or Animals Or Parts Or Parts Thereof   |
| A61F         | Filters Implantable Into Blood Vessels; Prostheses; Devices Providing Patency To, Or Preventing Collapsing Of, Tubular Structures Of The Body       |
| A61K         | Preparations For Medical, Dental, Or Toilet Purposes  |
| A61L         | Methods Or Apparatus For Sterilising Materials Or Objects In General  |
| A62C         | Fire-Fighting   |
| B01D         | Separation  |
| B01J         | Chemical Or Physical Processes, e.g., Catalysis, Colloid Chemistry; Their Relevant Apparatus  |
| B03C         | Magnetic Or Electrostatic Separation Of Solid Materials From Solid Materials Or Fluids; Separation By High-Voltage Electric Fields                  |
| B04B         | Centrifuges   |
| B25J         | Manipulators; Chambers Provided With Manipulation Devices   |
| B60L         | Propulsion Of Electrically-Propelled Vehicles   |
| B60P         | Vehicles Adapted For Load Transportation Or To Transport, To Carry, Or To Comprise Special Loads Or Objects   |
| B61B         | Railway Systems; Equipment Therefor Not Otherwise Provided For  |
| B63B         | Ships Or Other Waterborne Vessels; Equipment For Shipping   |
| B63H         | Marine Propulsion Or Steering   |
| B64C         | Aeroplanes; Helicopters   |
| B64D         | Equipment For Fitting In Or To Aircraft; Flying Suits; Parachutes; Arrangements Or Mounting Of Power Plants Or Propulsion Transmissions In Aircraft |
| B64G         | Cosmonautics; Vehicles Or Equipment Therefor  |



Table 3 (continued)

| IPC-subclass | Description  |
|--------------|--|
| B65D         | Containers For Storage Or Transport Of Articles Or Materials   |
| B81B         | Micro-Structural Devices Or Systems  |
| C01B         | Non-Metallic Elements; Compounds Thereof   |
| C02F         | Treatment Of Water, Waste Water, Sewage, Or Sludge   |
| C09K         | Materials For Applications Not Otherwise Provided For; Applications Of Materials Not Otherwise Provided For  |
| C10B         | Destructive Distillation Of Carbonaceous Materials For Production Of Gas, Coke, Tar, Or Similar Materials  |
| C10G         | Cracking Hydrocarbon Oils; Production Of Liquid Hydrocarbon Mixtures   |
| C10J         | Production Of Gases Containing Carbon Monoxide And Hydrogen From Solid Carbonaceous Materials By Partial Oxidation Processes Involving Oxygen Or Steam                   |
| C10L         | Fuels Not Otherwise Provided For; Natural Gas; Synthetic Natural Gas Obtained By Processes Not Covered By Subclasses C10g Or C10k  |
| C12N         | Micro-Organisms Or Enzymes; Compositions Thereof   |
| C21B         | Manufacture Of Iron Or Steel   |
| C22B         | Production Or Refining Of Metals, Pretreatment Of Raw Materials  |
| C25B         | Electrolytic Or Electrophoretic Processes For The Production Of Compounds Or Non- Metals; Apparatus Therefor   |
| E01F         | Additional Work, Such As Equipping Roads Or The Construction Of Platforms, Helicopter Landing Stages, Signs, Snow Fences, Or The Like                                    |
| E02B         | Hydraulic Engineering  |
| E04B         | General Building Constructions; Walls  |
| E21B         | Earth Or Rock Drilling; Obtaining Oil, Gas, Water, Soluble Or Meltable Materials Or A Slurry Of Minerals From Wells  |
| E21F         | Safety Devices, Transport, Filling-Up, Rescue, Ventilation, Or Drainage In Or Of Mines Or Tunnels  |
| F01 N        | Gas-Flow Silencers Or Exhaust Apparatus For Machines Or Engines In General; Gas-Flow Silencers Or Exhaust Apparatus For Internal-Combustion Engines                      |
| F03G         | Spring, Weight, Inertia, Or Like Motors; Mechanical-Power-Producing Devices Or Mechanisms, Not Otherwise Provided For Or Using Energy Sources Not Otherwise Provided For |
| F16F         | Springs; Shock-Absorbers; Means For Damping Vibration  |
| F16 L        | Pipes; Joints Or Fittings For Pipes; Supports For Pipes, Cables Or Protective Tubing; Means For Thermal Insulation In General  |
| F16P         | Safety Devices In General  |
| F22B         | Methods Of Steam Generation; Steam Boilers   |

Table 3 (continued)

| IPC-subclass | Description   |
|--------------|---|
| F24D         | Domestic- Or Space-Heating Systems  |
| F24F         | Air-Conditioning; Air-Humidification; Ventilation; Use Of Air Currents For Screening  |
| F24 J        | Production Or Use Of Heat Not Otherwise Provided For  |
| F25B         | Refrigeration Machines, Plants, Or Systems; Combined Heating And Refrigeration Systems; Heat Pump Systems   |
| F26B         | Drying Solid Materials Or Objects By Removing Liquid Therefrom  |
| F27B         | Furnaces, Kilns, Ovens, Or Retorts In General; Open Sintering Or Like Apparatus   |
| F28D         | Heat-Exchange Apparatus, Not Provided For In Another Subclass, In Which The Heat-Exchange Media Do Not Come Into Direct Contact   |
| G01 M        | Testing Static Or Dynamic Balance Of Machines Or Structures; Testing Of Structures Or Apparatus, Not Otherwise Provided For   |
| G01 N        | Investigating Or Analysing Materials By Determining Their Chemical Or Physical Properties   |
| G01R         | Measuring Electric Variables; Measuring Magnetic Variables  |
| G01 T        | Measurement Of Nuclear Or X-Radiation   |
| G01 V        | Geophysics; Gravitational Measurements; Detecting Masses Or Objects; Tags   |
| G01 W        | Meteorology   |
| G02B         | Optical Elements, Systems, Or Apparatus   |
| G02F         | Devices Or Arrangements, The Optical Operation Of Which Is Modified By Changing The Optical Properties Of The Medium Of The Devices Or Arrangements For The Control Of The Intensity, Colour, Phase, Polarisation Or Direction Of Light,  |
| G05B         | Control Or Regulating Systems In General; Functional Elements Of Such Systems; Monitoring Or Testing Arrangements For Such Systems Or Elements  |
| G05D         | Systems For Controlling Or Regulating Non-Electric Variables  |
| G06F         | Electric Digital Data Processing  |
| G06K         | Recognition Of Data; Presentation Of Data; Record Carriers; Handling Record Carriers  |
| G06 N        | Computer Systems Based On Specific Computational Models   |
| G06Q         | Data Processing Systems Or Methods, Specially Adapted For Administrative, Commercial, Financial, Managerial, Supervisory Or Forecasting Purposes; Systems Or Methods Specially Adapted For Administrative, Commercial, Financial, Managerial, Supervisory Or Forecasting Purposes, Not Otherwise Provided For |
| G10K         | Sound-Producing Devices; Methods Or Devices For Protecting Against, Or For Damping, Noise Or Other Acoustic Waves In General; Acoustics Not Otherwise Provided For  |
| G10 L        | Speech Analysis Or Synthesis; Speech Recognition; Speech Or Voice Processing; Speech Or Audio Coding Or Decoding  |

Table 3 (continued)

| IPC-subclass | Description   |
|--------------|---|
| G21C         | Nuclear Reactors  |
| G21F         | Protection Against X-Radiation, Gamma Radiation, Corpuscular Radiation Or Particle Bombardment; Treating Radioactively Contaminated Material; Decontamination Arrangements Therefor |
| H01B         | Cables; Conductors; Insulators; Selection Of Materials For Their Conductive, Insulating Or Dielectric Properties  |
| H01F         | Magnets; Inductances; Transformers; Selection Of Materials For Their Magnetic Properties  |
| H01G         | Capacitors; Capacitors, Rectifiers, Detectors, Switching Devices, Light-Sensitive Or Temperature-Sensitive Devices Of The Electrolytic Type   |
| H01J         | Electric Discharge Tubes Or Discharge Lamps   |
| H01L         | Semiconductor Devices; Electric Solid State Devices Not Otherwise Provided For  |
| H01M         | Processes Or Means  |
| H02B         | Boards, Substations, Or Switching Arrangements For The Supply Or Distribution Of Electric Power   |
| H02G         | Installation Of Electric Cables Or Lines, Or Of Combined Optical And Electric Cables Or Lines   |
| H02J         | Circuit Arrangements Or Systems For Supplying Or Distributing Electric Power; Systems For Storing Electric Energy   |
| H04B         | Transmission  |
| H04J         | Multiplex Communication   |
| H04L         | Transmission Of Digital Information, E.G. Telegraphic Communication   |
| H04M         | Telephonic Communication  |
| H04N         | Pictorial Communication   |
| H04W         | Wireless Communication Networks   |
| H05G         | X-Ray Technique   |
| H05H         | Plasma Technique; Production Of Accelerated Electrically-Charged Particles Or Of Neutrons; Production Or Acceleration Of Neutral Molecular Or Atomic Beams                          |

**Table 4** Variable operationalization and source

|                                 | Variable                | Operationalization  | Source                            |
|---------------------------------|-------------------------|---|-----------------------------------|
| $CTT_{i,t,s}$                   | Knowledge flow          | Sum of citations made by inventors living in region $j$ of patents created in region $i$ (within the respective IPC classes) in year $t$ .                | OECD REGPAT and Citation database |
| $SUBS_{,j}NET_{-i,j,t,(t-3),s}$ | Direct network links    | 3-year time lagged weighted sum of links within a joint R&D project (in the respective IPC classes) between organizations located in region $i$ and $j$ . | Subsidies catalogue               |
| $SUBS_{,j}NET_{-i,j,t,(t-5),s}$ | Direct network links    | 5-year time lagged weighted sum of links within a joint R&D project (in the respective IPC classes) between organizations located in region $i$ and $j$ . | Subsidies catalogue               |
| $DIST_{i,j}$                    | Geographical distance   | Great circle distance in kilometers between the centroids of region $i$ and $j$ .   |                                   |
| $TECH_{i,j}$                    | Technological proximity | Cosine similarity between region $i$ and $j$ based on the knowledge space of both regions.  | OECD REGPAT database              |
| $CO.INV_{i,j,t}$                | Social proximity        | Sum of patents developed in co-inventorship by inventors living in region $i$ and $j$ in year $t$ .   | OECD REGPAT database              |
| $PAT_{i,t,s}$                   | Innovative masses       | Sum of patents applied at EPO (in the respective IPC classes) in region $i$ and $j$ in year $t$ .   | OECD REGPAT database              |
| $PAT_{j,t,s}$                   | Innovative masses       | Sum of granted subsidized projects (in the respective IPC classes) in region $i$ and $j$ in year $t$ .  | Subsidies catalogue               |
| $SUBS_{,i,t,s}$                 |                         |   |                                   |
| $SUBS_{,j,t,s}$                 |                         |   |                                   |

**Table 5** Results of the fixed-effects, binomial and negative binomial (count) regression models (5-year time lag)

|                                | Binomial model 6     | Count model 6        |
|--------------------------------|----------------------|----------------------|
| Variables                      | Estimate (SE)        | Estimate (SE)        |
| Policy                         |                      |                      |
| SUBS.NET <sub>ij,(t-5),s</sub> | 0.024<br>(0.018)     | -0.009<br>(0.009)    |
| SUBS <sub>i,t,s</sub>          | 0.013***<br>(0.003)  | 0.005***<br>(0.002)  |
| SUBS <sub>j,t,s</sub>          | 0.02***<br>(0.003)   | 0.002<br>(0.002)     |
| Proximities                    |                      |                      |
| CO_INV <sub>ij, (t-5),s</sub>  | 0.11***<br>(0.005)   | 0.012***<br>(0.001)  |
| DIST <sub>ij</sub>             | -0.098***<br>(0.004) | -0.017***<br>(0.003) |
| TECH <sub>ij</sub>             | 1.809***<br>(0.038)  | 0.124***<br>(0.027)  |
| ORG <sub>ij</sub>              | -0.002***<br>(0.004) | 0.006***<br>(0.001)  |
| Regional properties            |                      |                      |
| PAT <sub>i,t,s</sub>           | 0.018***<br>(0.0005) | 0.004***<br>(0.0003) |
| PAT <sub>j,t,s</sub>           | 0.015***<br>(0.0005) | 0.003***<br>(0.0003) |
| Control                        |                      |                      |
| CIT_lag <sub>ij,t</sub>        | -0.236***<br>(0.011) | 0.047***<br>(0.003)  |
| Year Dummies                   | Yes                  | Yes                  |
| Technology Dummies             | Yes                  | Yes                  |
| AIC                            | 163,808              | 102,700              |
| 2x Log-likelihood              |                      | -102,532             |
| Observations                   | 255,493              | 34,487               |
| Non-zero obs.                  | 34,487               | -                    |

\* significant at 90% level, \*\* significant at 95% level, \*\*\* significant at 99% level

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Arundel A, Kabla I (1998) What percentage of innovations are patented? Empirical estimates for European firms. *Res Policy* 27:127–141
- Balland PA (2012) Proximity and the evolution of collaboration networks: evidence from research and development projects within the global navigation satellite system (GNSS) industry. *Reg Stud* 46:741–756

- Bode E (2004) The spatial pattern of localized R&D spillovers: an empirical investigation for Germany. *J Econ Geogr* 4:43–64
- Boschma R (2005) Proximity and innovation: a critical assessment. *Reg Stud* 39:61–74
- Bottazzi L, Peri G (2003) Innovation and spillovers in regions: evidence from European patent data. *Eur Econ Rev* 47:687–710
- Breschi S, Cusmano L (2004) Unveiling the texture of a European research area: emergence of oligarchic networks under EU framework Programmes. *Int J Technol Manag* 27:747–772
- Breschi S, Lissoni F (2004) Knowledge networks from patent data methodological issues and research targets. In: Henk FM, Glänzel W, Schmoch U (eds) *Handbook of quantitative science and technology research*. Springer, The Netherlands, pp 613–643
- Breschi S, Lissoni F (2009) Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows. *J Econ Geogr* 9:1–30
- Broekel T (2015) Do cooperative R&D subsidies stimulate regional innovation efficiency? Evidence from Germany. *Reg Stud* 49:1087–1110
- Broekel T, Brenner T (2011) Regional factors and innovativeness: an empirical analysis of four German industries. *Ann Reg Sci* 47:169–194
- Broekel T, Graf H (2012) Public research intensity and the structure of German R&D networks: a comparison of 10 technologies. *Econ Innov New Technol* 21:345–372
- Broekel T, Hartog M (2013a) Explaining the structure of inter-organizational networks using exponential random graph models. *Ind Innov* 20:277–295
- Broekel T, Hartog M (2013b) Determinants of cross-regional R&D collaboration networks: an application of Exponential Random Graph Models. In: Scherngell T (ed) *The geography of networks and R&D collaborations*, Springer, Cham, pp 49–70
- Broekel T, Balland PA, Burger M, van Oort F (2014) Modeling knowledge networks in economic geography: a discussion of four methods. *Annals of Regional Science* 53:423–452
- Broekel T, Brenner T, Buerger M (2015) An investigation of the relation between cooperation intensity and the innovative success of German regions. *Spat Econ Anal* 10:52–78
- Buisseret TJ, Cameron HM, Georgiou L (1995) What difference does it make? Additionality in the public support of R&D in large firms. *Int J Technol Manag* 10:587–600
- Bundesministeriums für Bildung und Forschung (BMBF) (2008) Merkblatt für Antragsteller/Zuwendungsempfänger zur Zusammenarbeit der Partner von Verbundprojekten. Bundesministerium für Bildung und Forschung, BMBF Publik. Berlin: Federal Ministry of Education and Research. [http://www.kooperation-international.de/fileadmin/public/Vordruck\\_0110.pdf](http://www.kooperation-international.de/fileadmin/public/Vordruck_0110.pdf). Accessed 30.01.2016
- Burger M, van Oort F, Linders GJ (2009) On the specification of the gravity model of trade: zeros, excess zeros and zero-inflated estimation. *Spat Econ Anal* 4:167–190
- Butzin, A (2009) Innovationsbiographien als Methode der raum-zeitlichen Erfassung von Innovationsprozessen. In: Dannenberg P, Köhler H, Lang T, Utz J, Zakirova B, Zimmermann T (ed.) *Innovationen im Raum-Raum für Innovationen: 11. Junges Forum der ARL, 21. bis 23. Mai 2008 in Berlin*, Verlag der ARL-Akademie für Raumforschung und Landesplanung, Hannover, pp 189–198
- Cohen W, Levinthal D (1990) Absorptive capacity: a new perspective on learning and innovation. *Adm Sci Q* 35:128–152
- Cohen W, Nelson R, Walsh JP (2000) Protecting their intellectual assets: appropriability conditions and why US manufacturing firms patent (or not) (no. w7552). National Bureau of Economic Research
- Criscuoloa P, Verspagen B (2008) Does it matter where patent citations come from? Inventor vs. examiner citations in European patents. *Res Policy* 37:1892–1908
- Downs GW, Mohr LB (1976) Conceptual issues in the study of innovation. *Adm Sci Q* 21:700–714
- Engelsman EC, van Raan AF (1994) A patent-based cartography of technology. *Res Policy* 23:1–26
- Fornahl D, Broekel T, Boschma R (2011) What drives patent performance of German biotech firms? The impact of R&D subsidies, knowledge networks and their location. *Pap Reg Sci* 90:395–418
- Frenken K, van Oort F, Verburg T (2007) Related variety, unrelated variety and regional economic growth. *Reg Stud* 41:685–697
- Giuliani E, Bell M (2005) The micro-determinants of meso-level learning and innovation: evidence from a Chilean wine cluster. *Res Policy* 34:47–68
- Granovetter MS (1973) The Strength of weak ties. *Am J Sociol* 78:1360–1380
- Griliches Z (1979) Issues in assessing the contribution of research and development to productivity growth. *Bell J Econ*:92–116
- Griliches Z (1998) Patent statistics as economic indicators: a survey. In: Elliott C (ed) *R&D and productivity: the econometric evidence*. University of Chicago Press, Chicago, pp 287–343

- Hägerstrand T (1952) The Propagation of innovation waves (Lund studies in geography: Series B, Human geography, 4). Lund: Royal University of Lund, Dept. of Geography
- Hägerstrand T (1965) Aspects of the spatial structure of social communication and the diffusion of information. *Papers of the Regional Science Association* 16:27–42
- Hägerstrand T (1967) Innovation diffusion as a spatial process. University of Chicago Press, Chicago
- Hewitt-Dundas N, Roper S (2010) Output additionality of public support for innovation: evidence for Irish manufacturing plants. *Eur Plan Stud* 18:107–122
- Hoekman J, Frenken K, van Oort F (2009) The geography of collaborative knowledge production in Europe. *Annals in Regional Science* 43:721–738
- Howells JRL (2002) Tacit knowledge, innovation and economic geography. *Urban Stud* 39:871–884
- Isard W (1954) Location theory and trade theory: short run analysis. *Q J Econ* 68:305–322
- Jaffe AB (1986) Technological opportunity and spillovers of R&D: evidence from firms' patents, profits and market value. *Am Econ Rev* 76:984–1001
- Jaffe AB, Trajtenberg M, Henderson R (1993) Geographic localisation of knowledge spillovers as evidenced by patent citations. *Q J Econ* 108:577–598
- Kline S, Rosenberg N (1986) An overview of innovation. In: Landau R, Rosenberg N (eds) *The positive sum strategy*. National Academy Press, Washington, pp 275–305
- Kosfeld R, Werner A (2012) Deutsche Arbeitsmarktregionen – Neuabgrenzung nach Kreisgebietsreform 2007–2011. *Raumforsch Raumordn* 70:49–64
- Loasby BJ (2001) Organisation as interpretative systems. *Rev Econ Ind* 97:17–34
- Lundvall BA, Johnson B (1994) The learning economy. *Journal of Industrial Studies* 1:23–42
- Maggioni MA, Nosvelli M, Uberti TE (2007) Space versus networks in the geography of innovation: a European analysis. *Pap Reg Sci* 86:471–493
- Maggioni MA, Uberti TE, Usai S (2011) Treating patents as relational data: knowledge transfer and spillovers across Italian provinces. *Ind Innov* 18:39–67
- Maggioni MA, Uberti TE, Nosvelli M (2014) Does intentional mean hierarchical? Knowledge flows and innovative performance of European regions. *Ann Reg Sci* 53:453–485
- Nooteboom B, Van Haverbeke W, Duysters G, Gilsing V, Van den Oord A (2007) Optimal cognitive distance and absorptive capacity. *Res Policy* 36:1016–1034
- Paci R, Usai S (2009) Knowledge flows across European regions. *Annual of Regional Science* 43:669–690
- Peri G (2005) Determinants of knowledge flows and their effect on innovation. *Rev Econ Stat* 87:308–322
- Ponds R, Oort FV, Frenken K (2009) Innovation, spillovers and university–industry collaboration: an extended knowledge production function approach. *J Econ Geogr* 10:231–255
- Rogers EM (2003) *Diffusion of innovations*, 5th edn. Free Press, New York
- Scherngell T, Barber MJ (2009) Spatial interaction modelling of cross-region R&D collaborations: empirical evidence from the 5th EU framework programme. *Pap Reg Sci* 88:531–546
- Scherngell T, Barber MJ (2011) Distinct spatial characteristics of industrial and public research collaborations: evidence from the fifth EU Framework Programme. *Ann Reg Sci* 46(2):247–266
- Schrader S (1991) Informal technology transfer between firms: cooperation through information trading. *Res Policy* 20:153–170
- Sternberg R (2000) Knowledge networks and regional development – evidence from the European regional innovation survey (ERIS): theoretical concepts, methodological approach, empirical basis and introduction to the theme issue. *Eur Plan Stud* 8:389–407
- Storper M, Venables AJ (2004) Buzz: face-to-face contact and the urban economy. *J Econ Geogr* 4:351–370
- Tijssen RJW (1998) Quantitative assessment of large heterogeneous R&D networks: the case of process engineering in the Netherlands. *Res Policy* 26:791–809
- Tinbergen J (1962) *Shaping the world economy*. Twentieth Century Fund, New York
- Verspagen B (1997) Measuring intersectoral technology spillovers: estimates from the European and US patent office databases. *Econ Syst Res* 9:47–65
- Witt U, Broeckel T, Brenner T (2007) Knowledge and its economic characteristics – a conceptual clarification. In: Arena R, Festré A, N Lazaric (eds.) *Handbook of Economics and Knowledge*. Edward Elgar, Cheltenham, Northampton, pp 369–382
- Zucker LG, Darby MR, Armstrong J (1998) Geographically localized knowledge: spillovers or markets? *Econ Inq* 36:65–86