

Fritsch, Michael; Titze, Mirko; Piontek, Matthias

**Article — Published Version**

## Identifying cooperation for innovation—a comparison of data sources

Industry and Innovation

**Provided in Cooperation with:**

Halle Institute for Economic Research (IWH) – Member of the Leibniz Association

*Suggested Citation:* Fritsch, Michael; Titze, Mirko; Piontek, Matthias (2020) : Identifying cooperation for innovation—a comparison of data sources, *Industry and Innovation*, ISSN 1469-8390, Routledge, London, Vol. 27, Iss. 6, pp. 630-659, <https://doi.org/10.1080/13662716.2019.1650253>

This Version is available at:

<https://hdl.handle.net/10419/225010>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<http://creativecommons.org/licenses/by-nc-nd/4.0/>

# Identifying cooperation for innovation—a comparison of data sources

Michael Fritsch <sup>a,b</sup>, Mirko Titze <sup>c</sup> and Matthias Piontek<sup>d</sup>

<sup>a</sup>Faculty of Economics and Business Administration, Friedrich Schiller University Jena, Jena, Germany; <sup>b</sup>Halle Institute for Economic Research (IWH), Halle (Saale), Germany; <sup>c</sup>Centre for Evidence-Based Policy Analysis (IWH-CEP), Halle Institute for Economic Research (IWH), Halle (Saale), Germany; <sup>d</sup>Service Center for Research and Transfer, Friedrich Schiller University Jena, Jena, Germany

## ABSTRACT

The value of social network analysis is critically dependent on the comprehensive and reliable identification of actors and their relationships. We compare regional knowledge networks based on different types of data sources, namely, co-patents, co-publications, and publicly subsidized collaborative R&D projects. Moreover, by combining these three data sources, we construct a multilayer network that provides a comprehensive picture of intraregional interactions. By comparing the networks based on the data sources, we address the problems of coverage and selection bias. We observe that using only one data source leads to a severe underestimation of regional knowledge interactions, especially those of private sector firms and independent researchers.

## KEYWORDS

Knowledge interactions; social network analysis; regional innovation systems; data sources, R&D cooperation

## JEL

O30; R12; R30

## 1. Introduction

There is strong indication that knowledge exchange and division of labor play an increasingly important role for innovation processes (Wuchty, Jones, and Uzzi 2007; Tomasello et al. 2017). Empirical analyses of innovation processes are faced with the problem of identifying the ties and knowledge flows between actors involved. Information on the relationships among innovating actors may come from sources such as patent statistics (Graf 2006), publications, and other forms by which research and knowledge become manifest (Ter Wal and Boschma 2009). Because each of those data sources is selective in the sense that it only records certain types of interactions and disregards others, analyses of a certain innovation system may show differing results depending on the data source used.<sup>1</sup> Consequently, actors that appear to be relatively important in a network constructed with a certain data source may appear to be unimportant or completely disregarded if a different source of data is used (Broekel and Graf 2010).

This paper compares three types of databases, namely, patent statistics, co-publications, and subsidized research collaborations that reflect different types of

**CONTACT** Mirko Titze  [mirko.titze@iwh-halle.de](mailto:mirko.titze@iwh-halle.de)  Halle Institute for Economic Research (IWH), Germany

<sup>1</sup>For example, patent data disregard cooperation for inventions that are not patented (e.g. Arundel and Kabla 1998).

© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

interactions. We describe the comprehensiveness and selectivity of these three types of data. We combine them to construct a multilayer network that provides a comprehensive picture of regional interactions and serves as a benchmark for assessing the measurement bias of the individual data sources.<sup>2</sup> This empirical exercise covers the period 2000–2010 and is performed for six regions in Germany with varying levels of innovation activity and population density.

The contribution of our paper is threefold. *First*, we want to stimulate the awareness of researchers regarding the fact that considering only one type of data yields a rather incomplete network of regional knowledge transfer. Our analyses show considerable differences among the networks constructed with the three types of data. Although a relatively high share of public research institutions is involved in all three forms of interactions, we observe many private sector firms that only participate in one specific form of knowledge transfer. According to our findings, only 40% of private sector firms involved in regional knowledge transfer are captured by the co-patent indicator, 30% by the co-publication indicator and 46% by subsidized research collaborations. Hence, investigating only one type of data neglects a considerable share of factual interactions, especially relationships between private firms and public research institutions. This point is of crucial relevance if this data is used to evaluate the success of cluster or network development programs.

*Second*, we introduce a technical tool – record-linkage techniques – to identify identical actors across networks. In doing so, we are capable to define key actors of regional knowledge transfer in a comprehensive and systematic way. *Finally*, we want to increase awareness that incomplete networks may affect the results of the analysis of (regional) innovation processes since our approach is capable to identify different types of knowledge transfer over the different stages of an innovation project, e.g. R&D collaboration followed by a co-patent application and/or co-publication.

The remainder of this paper is organized as follows. In section 2 we discuss the related literature. Section 3 introduces the case study regions and describes the data sources. Section 4 compares the networks constructed with the types of data. Moreover, we use the Dresden region as an example to provide a comprehensive picture of the regional innovation network by combining all three data sources. Section 5 summarizes and concludes.

## 2. Related literature

Although some scholars have indicated possible differences between networks investigated on the basis of different datasets (e.g. Broekel and Graf 2010; Ter Wal and Boschma 2009; Fritsch and Graf 2011; Broekel, Fornahl, and Morrison

---

<sup>2</sup>Few empirical analyses combine different data sources for the construction of networks because of limited data availability and the more technical problems of combining multiple sources such as data matching. For first approaches see, for example, Schmoch (1999), Meyer (2002), and Youtie and Shapira (2008). A study by Lata, Scherngell, and Brenner (2015) combines three different datasets (granted projects supported within the EU framework programs, co-patents, and co-publications). Although, these datasets are merged at the regional level and not at the level of actors.

2015), the empirical evidence on the actual extent of these differences is rare.<sup>3</sup> Most empirical analyses of innovative interactions are based on co-patents (e.g. Graf and Henning 2009; Hoekman, Frenken, and van Oort 2009)<sup>4</sup>, co-publications (e.g. Ponds, van Oort, and Frenken 2007; Hoekman, Frenken, and van Oort 2009; Hoekman, Frenken, and Tijssen 2010), or data on publicly supported collaborative R&D projects<sup>5</sup>. Few studies such as Lata, Scherngell, and Brenner (2015) or Broekel and Graf (2010) go one step further and consider different channels of regional knowledge transfer. Many of such analyses are carried out at the regional level<sup>6</sup> or at the level of specific technology fields rather than at the level of actors (organizations, institutions).

Analyses based on co-patents identify a key role of universities, extra-university public research institutions and large firms who tend to have the role of an important broker and gatekeeper in the network (e.g. Graf and Henning 2009; Graf 2011; Kauffeld-Monz and Fritsch 2013). A significant advantage of patents is that important related information such as the name and address of the applicant(s), the names and home addresses of inventors, the respective technological field etc. are easily available. A further advantage is that patenting requires a certain minimum level of newness, so that marginal inventions are excluded. This safeguards a certain minimum level of comparability. However, because inventions are just ideas many of which are never commercially applied, patent-based analyses can hardly say anything about later stages of the innovation process. Moreover, due to the requirement of a certain minimum level of newness, ideas and innovations that are commercially relevant but do not attain such a level of newness are not included. The same holds for results of basic research that cannot be patented in contrast to more applied and product-oriented research.

Ter Wal and Boschma (2009) mention a further potential deficiency of patents as indicators for innovation activities. They argue that co-patents represent relatively formal, often legal cooperation agreements. As a consequence, linkages that are more informal in nature are covered to a comparatively lesser degree in patent statistics. Another shortcoming is that patenting behavior differs considerably across economic sectors and firm sizes. Ter Wal and Boschma (2009) argue that sectors such as pharmaceuticals and semiconductor industries are more present in co-patenting networks than the software industry or services. Moreover, they claim that co-patent networks are biased in terms of firm size where large firms are more likely to file patents than small and medium-sized firms due to the costs of patenting. Finally, research cooperation among universities and public research institutes as well as cooperative relationships between public research and private firms are probably under-reported as public institutions have only comparatively weak incentives to patent.

---

<sup>3</sup>Micro-level analyses of R&D cooperation in the form of strategic alliances (Contractor and Reuer 2018; Link and Antonelli 2018) and joint ventures (Adams and Link 2018) focus on bilateral relationships and ego networks, so that they largely neglect indirect links between actors.

<sup>4</sup>Fischer and Griffith (2008) as well as Fischer, Scherngell, and Jansenberger (2006) use patent citations to depict innovative interactions with the help of the patent indicator.

<sup>5</sup>E.g. Maggioni, Nosvelli, and Uberti (2007), Broekel, Fornahl, and Morrison (2015), Scherngell and Barber (2011), Scherngell and Lata (2013), Barber and Scherngell (2013).

<sup>6</sup>Regional level means that the region is the unit of observation. Knowledge transfer is then modelled as knowledge transfer between regions.

Based on data of subsidized collaboration projects in Germany Broekel and Graf (2010) distinguish between joint projects in basic and in applied research. For basic research they find smaller and more centralized networks which results in a relatively pronounced concentration of interactions between a relatively small numbers of actors. Quite frequently, universities have central positions with many broker functions in these networks. In contrast, networks in applied research are more characterized by involvement of larger firms as central actors. Moreover, public research institutes are more important in these networks than universities.

Regarding networks based on co-publications Ponds, van Oort, and Frenken (2007) argue that there is a considerable mismatch in the incentive structure between private sector firms and public research institutions. The goal of academia is to create new knowledge, to broaden the knowledge base and to diffuse knowledge as widely as possible. In contrast, actors in the private sector are primarily interested in minimizing the diffusion of their knowledge in order to preserve an advantage over competitors. Co-publications of authors from different research fields may be an important means for knowledge transfer across different academic disciplines (see, for example, Peng and Haoxiang 2015).

To sum up, each of the three types of data under consideration—patents, subsidized research collaborations, and co-publications—captures different modes of interactions that are likely to exist in different stages of the innovation process. Against this backdrop, it is quite likely that the networks and the results of the respective analyses based on these types of data differ considerably. Hence, there is good reason to expect that an analysis based on several types of data will provide a clearer and particularly more reliable picture of regional interaction and knowledge transfer.

### 3. Empirical approach

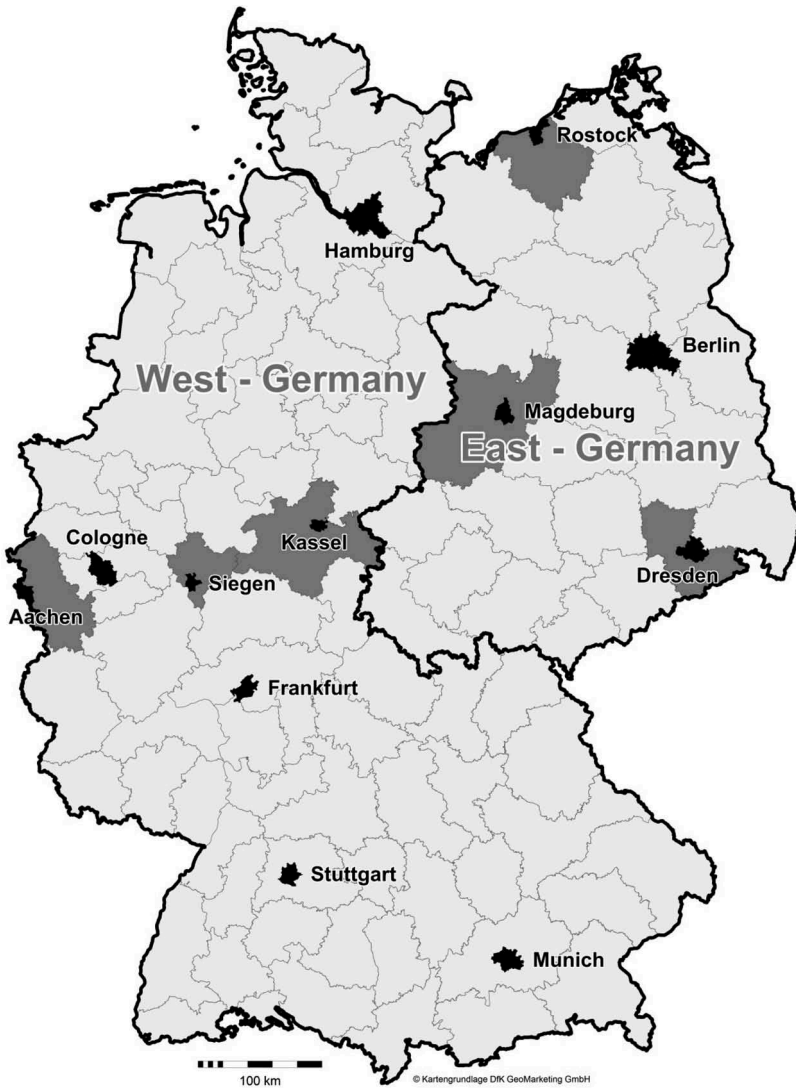
#### 3.1. Spatial framework

We choose the level of planning regions (*Raumordnungsregionen*) as the geographical unit of analysis. German planning regions typically comprise a core city (*kreisfreie Stadt*) and its neighboring districts (*Kreise*). This regional level of aggregation is considered appropriate for regional network analyses for two reasons (Graf and Henning 2009). *First*, it considers that regional channels of knowledge transfer do not necessarily end at the boundaries of a district or district-free city. *Second*, planning regions consider commuter flows. This aspect is particularly important for the analysis of patent applications because patents are assigned to the inventor's place of residence, which might not be the same district as their workplace.

Our sample of regions comprises three types of settlement structures: agglomerations, moderately congested regions, and rural areas.<sup>7</sup> Large innovation centers such as Munich or Stuttgart are not included. All the case study regions host at least one university. [Figure 1](#) shows the spatial distribution of the case study regions (areas in dark grey). Two of the regions—Aachen and Dresden—represent smaller agglomerations with comparable numbers

---

<sup>7</sup>This definition is in line with the classification of the Federal Institute for Research on Building, Urban Affairs and Spatial Development (BBSR). For details see BBSR (2015).



**Figure 1.** Case study regions in Germany.

Source: Own illustration.

of inhabitants (about 1 Mio.), establishments (28,000), and employees (approximately 235,000) (Table A1 in Appendix A). Both regions host a large university that focuses on engineering and natural science.

These regions also match with respect to the qualification structure of the workforce (the share of natural scientists and engineers in the total number of employees is approximately 4.5%) and the size of the universities (number of professors: 600–800; total research and teaching staff: 4,700–4,900). The regions of Rostock and Siegen represent smaller cities with a rural surrounding. Each has a population of approximately 430,000 and has been shrinking over the last decade. Kassel and Magdeburg are moderately congested regions with a population of approximately 1 Mio., which has slightly declined from 2000 to 2010. Regarding ‘openness’ (share of inter-regional collaborations over total number of

interactions<sup>8</sup>), we find that small regions (in particular Kassel and Siegen) are more connected to partners outside the region than agglomerations like Aachen and Dresden.

### 3.2. Data sources and matching procedures

Our analysis builds on Titze, Schwartz, and Brachert (2012), who presented a conceptual approach for analyzing networks that feature several dimensions of interactions. We develop a multilayer framework that allows us to investigate the overlapping channels of knowledge transfer at the level of institutions. Because the information on interactions in all three databases relies on officially documented interaction processes, they reveal actual collaborations more credibly than self-reported responses in interviews or questionnaires.

Data on publicly funded *R&D collaboration projects* are provided in the Subsidies Catalogue (*Foerderkatalog*) prepared by the German Federal Ministry for Education and Research and the National Aeronautics and Space Research Centre<sup>9</sup>, which has a crucial role in the administration of these projects (for a detailed description see Broekel and Graf 2012). The data comprises more than 100 thousand completed and ongoing research projects. This database may only have a limited scope. *First*, it does not contain information on collaborative R&D projects conducted without public funding. *Second*, some support schemes from the Federal States or the European Union (EU) are not included in this database. *Third*, public grants are awarded to institutions (universities, external research institutes, firms, etc.) but not individuals. Consequently, this database does not include the names of the persons involved in a project. Three key variables from the Subsidies Catalogue are relevant for our investigation: primary keys for sub-projects and for the entire collaboration project<sup>10</sup>, the names and locations of the executing organizations,<sup>11,12</sup> and the funding period. Small and medium-sized enterprises, universities, and extra-university public research institutes are generally eligible for the publicly funded projects recorded in this database. The share of large firms is relatively small because they are not eligible for some of these programs. We account for those projects that involve at least two collaboration partners in the region under analysis.

The German Patent and Trade Mark Office (DPMA) provides data on (*co-*) *patents* with at least one German organization involved. Each record includes a unique patent identification number, the title of the patent, the patent classes (IPC) and the names and locations of the inventor(s) and applicant(s). As we are interested in the actual knowledge flows, we use the applicant's name and regional

<sup>8</sup>This information relies on the Foerderkatalog (version 2017) and the EPO database.

<sup>9</sup>For our network analysis, we used the version provided in 2014.

<sup>10</sup>In case of several subprojects these subprojects summarized to one main collaboration project.

<sup>11</sup>The database distinguishes between the recipient of the grant(s) and the organization that actually works on the project (executing organization). In most cases both actors are identical. Exceptions are typically large enterprises consisting of numerous subsidiaries and large publicly funded research organizations like the Fraunhofer Society. In case of the Fraunhofer Society the recipient of the grant is the headquarter in Munich, but the actual project is conducted in a specific Fraunhofer Institute that may be located elsewhere.

<sup>12</sup>The database also contains a variable indicating the type of the actor (private firm, university, extra-university research institute and 'others'). In principle, this variable could be an appropriate indicator for measuring organizational proximity. Unfortunately, however, the raw data contains many incorrect assignments. Moreover, the spelling of the names has not been harmonized, and a unique identifier for organizations does not exist.

information. We consider patent applications with at least two applicants<sup>13</sup> in the respective region. Compared to the OECD RegPat data this source has several severe advantages. *First*, since the patent identification number does not change over the different versions of the statistics, it avoids multiple counting of the same patent. *Second*, it is considerably more comprehensive since it also contains the complete set of patents that has only been filed at the German Patent Office and that is not included in the RegPat data.<sup>14</sup> *Third*, we spent particular manual effort on the correction of typing errors and different spelling of inventor's names in order to maximize the reliability of the identification of inventors, an issue that is of key importance for the topic of our analysis.

The use of patent data in empirical analyses of innovation processes is not free from (well-known) methodological problems (e.g. Griliches 1990; Schmoch 1999; Cohen, Nelson, and Walsh 2000; Mansfield, Schwartz, and Wagner 1981; Blind et al. 2006). *First*, certain inventions are not patented because of problems such as secrecy, application cost, the effort required to demonstrate novelty, and the time span between patent filing and granting. *Second*, large companies, like Siemens and extra-university public research organizations such as the Fraunhofer Society, have centralized patent offices at their headquarters that administer all the patent applications for their organization. Thus, we follow the approach of Graf (2011) and solve the problem of headquarter applications by considering only patents where the majority of inventors have residences in one of our case study regions. These patents are then assigned to the local subsidiary of the respective company or to the local research institute of the public research organization. *Third*, patent activities differ considerably across scientific fields. *Forth*, inventions with a low degree of novelty and inventions in non-technological fields such as new methods of organization of management cannot be patented.

Finally, we rely on bibliometric data provided by the Clarivate Web of Science (formerly Thomson Reuters ISI Web of Knowledge) database for the analysis of co-publications. The packages available for the analysis were the Social Sciences Citation Index, the Science Citation Index Expanded, and the Arts & Humanities Citation Index. We use the following information from this database: the primary key of the publication (WOS number), name of the authors' affiliation, and geographical locations recorded in the authors' information. We consider those co-publications that report at least two authors from different affiliations in the region under consideration.

Using bibliometric data presents certain well-known and -discussed difficulties in the literature (e.g. Abramo, D'Angelo, and Caprasecca 2009). *First*, the Web of Science database is incomplete because it mainly contains articles published in peer-reviewed journals. *Second*, publication activities and strategies differ considerably between scientific disciplines. *Third*, there is not necessarily complete correspondence of authorship

---

<sup>13</sup>Some studies also consider 'mobility' relations. A mobility link occurs if an inventor is named on two patent applications of different applicants. The idea behind is that knowledge flows if the inventor moves from institution A to institution B (Graf and Henning 2009). We include this specific form of knowledge transfer in the patent layer, but not in the remaining two layers (co-publications, collaborative R&D collaborations). The main reason is that the data on publicly funded collaborative R&D projects contains no information about the individual researchers involved. Hence, it is not possible to analyze whether a researcher moved from institution A to B.

<sup>14</sup>The number of patents that is recorded in RegPat (version March 2018) for the same regions and period of time is only about 53 % percent of the number of patents that we find in our data base. Quite remarkably, this share varies considerably across the regions of our sample.



**Table 1.** Features of the raw data in the sources applied.

	Co-patents	Co-publications	Collaborative R&D projects
<i>Form of interaction</i>			
Database provided by	German Patent Office	Clarivate Analytics (formerly Thomson Reuters Web of Knowledge)	German Federal Ministry of Education and Research, BMBF (Subsidies Catalogue)
Identification of interaction by	Applicants/inventors with the same patent number	Authors with the same journal/book article	Institutions with the same identifier of collaboration project
Knowledge area of the interaction	International Patent Classification (IPC)	Publication classes according to the Web of Knowledge database	Technological fields according to the BMBF classification scheme
<i>Information at the level of individuals</i>			
Name	Yes, differentiated by inventor and applicant; different spellings possible (no unique identifier)	Yes, but different spelling possible (no unique identifier)	No
Surname	Yes, sometimes only initials available; academic degrees are often incomplete	Yes, but frequently only initials	No
Name of the region	Yes, but different spellings possible	No	No
Regional codes	ZIP codes	No	No
<i>Information about the organization/affiliation</i>			
Name	Inventors: no; Applicants: yes	Yes, but different spellings possible (no unique identifier)	Yes, but different spellings possible (no unique identifier)
Name of the region	Yes, but different spellings possible	Yes, but different spellings possible	Yes, but different spellings possible
Regional codes	ZIP codes	Sometimes, but different spellings possible (no unique identifier)	Administrative regional codes ( <i>Amtlicher Gemeindeschlüssel</i> )
Industry	No	No	2-digit NACE codes (universities and research institutes at a more disaggregated level)

Source: Own illustration.

of a publication and actual collaboration in the respective research. Furthermore, identifying inter-regional linkages (co-publications, scientist mobility) in the Web of Science database is problematic because the names of the affiliations are not standardized in this dataset. Table 1 summarizes the main features of the datasets.

The information about the actors, that is, their names and geographical code, from the three data sources was subject to a harmonization procedure that consisted of two steps: a pre-cleaning routine (change of the spelling to uppercase, replacement of German umlauts, removal of double spaces, etc.) and the record linkage in a narrow sense. For this purpose, we used the software Fuzzy Dupes, which provides a probability for the match of two records (see Ehrenfeld 2015a and 2015b for details). To receive further actor-specific information (e.g. type of institution, number of employees, industry code, and age), we merged the resulting dataset with the Amadeus data and the Research Explorer database (see Research Explorer 2017).

The Amadeus database comprises information on companies in Europe. For Germany, this database includes approximately 3 million companies. Every company in this data holds a unique identification number used to identify and link actors. The Research Explorer dataset comprises approximately 23,000 German universities and publicly funded external research institutes. This database complements the Amadeus

enterprise database regarding cooperation actors, because the Amadeus data usually does not include universities and extra-university research institutes.

According to the limitations of each dataset (Table 1) and for harmonization purposes, we investigate a subsample of the entire network that relies on intraregional interactions between institutions. We restrict the analysis to relationships between institutions because the data on publicly funded research collaborations does not allow the identification of the individuals involved in a project.<sup>15</sup> We developed the network by assuming that all partners involved in a collaboration are linked to all other partners in the same way so that all connections between partners are weighted equally.<sup>16</sup>

Table 2 presents some descriptives on the harmonized individual datasets. After harmonization, we can identify about 1,096 regional actors in the co-patenting, 488 in the co-publication and 605 in the collaborative R&D network. The dominant event in

**Table 2.** Descriptives on individual datasets.

	Co-patents	Co-publications	Collaborative R&D projects
Number of collaboration events <sup>a</sup>	1,528	8,923	639
<i>Number of regional<sup>d</sup> actors involved</i>			
Minimum	2	2	2
Maximum	5	5	10
Average	2.1	2.1	2.5
<i>Number of actors by type</i>			
Total number	1,096	488	605
<i>Characteristics of private sector firms</i>			
<i>Share of industry (%)<sup>e, c</sup></i>			
Agriculture, forestry, fishing (A)	0.0	0.0	1.2
Manufacturing (B-E)	52.2	23.8	37.1
Construction (F)	2.8	0.9	2.4
Trade, transportation, tourism, communication (G-J)	9.6	10.6	17.1
Banking, insurance, real estate services (K-L)	3.2	1.3	1.2
Other services (M-S)	32.3	63.4	41.1
<i>Firm age (years, base year 2010)</i>			
Mean	23.3	15.4	15.8
Standard deviation	30.3	15.0	12.8
Minimum	0	0	1
Maximum	228	111	171
<i>Share of firm size (%)<sup>f, c</sup></i>			
Small	33.3	41.4	38.3
Medium sized	35.3	26.6	37.0
Large	24.2	21.2	19.4
Very large	7.3	10.8	5.3

Notes: a) Co-patents, co-publications, collaborative R&D projects. b) This category represents actors (mainly individual inventors) who could not be assigned to an institution because the patent document does not contain information about the inventor's affiliation. c) Differences due to rounding may occur. d) The term 'regional' refers to collaboration partners located in the same spatial planning region. Actors are harmonized at the level of institutions. e) Statistical Classification of Economic Activities in the European Community (Rev. 2, 2008). f) According to the classification provided in the Bureau van Dijk databases.

Sources: Own calculations; Private sector firm characteristics: Bureau van Dijk.

<sup>15</sup>As a consequence, links between co-inventors named in a patent document are disregarded if the patent does not have several applicants.

<sup>16</sup>We are aware of the fact that interactions do not necessarily have the same intensity in one co-operation. In particular, this might hold true in collaborations with many partners, e.g. R&D collaborations in EU Framework Programmes. Here, an interaction is expected to occur only between the coordinator and the partners but not between partners (Breschi and Cusmano 2004). However, as none of our databases used in the paper provides information on the coordinator we assume identical weights for each link.

our analysis are co-publication activities (8,923). By contrast, the number of co-patents and collaborative R&D projects is considerably smaller (1,528 and 693, respectively). On average approximately two actors from the same planning region (which forms our geographical unit of analysis) are involved in co-patenting and co-publication activities. The average number of partners in collaborative R&D projects is slightly higher (2.5).

We then studied characteristics of actors involved in innovative intraregional knowledge transfer. This analysis is restricted to private sector firms because information on industry<sup>17</sup>, size and age of universities and other public research institutions is not available or not comparable. The largest share of private sector firms in co-patent networks belongs to the manufacturing sector (52.2%); the share of actors in other services amounts to 32.3%. By contrast, 23.8% (37.1%) of private sector firms in co-publication (collaborative R&D) networks can be assigned to the manufacturing sector; 63.4% (41.1%) are part of other services. All other sectors are negligible. Private sector firms in co-patent networks are on average older (23.3 years) than those in co-publication (15.4 years) and in collaborative R&D networks (15.8 years).<sup>18</sup> More than two thirds of private sector firms involved in co-operation networks is small and medium-sized. This holds across all three types of networks analyzed.

## 4. Comparing the types of activities of regional innovation networks

### 4.1. Actors involved in the different types of innovative interactions

Based on the three data sources, we identified 1,940 unique actors in the six case study regions during the period 2000–2010 (Table 3). Private sector firms represent 1,111 of these actors (57.2%), 20 actors (1.0%) are universities, and 115 actors (5.9%) are extra-university public research institutes. The remaining 694 actors (35.8%) mainly consist of individual inventors and authors who could not be assigned to an institution (see Table A2); 839 (43.3%) actors were identified either in the Amadeus firm database or the Research Explorer database (see Research Explorer 2017).

The lion's share of the 1,940 unique actors recorded in our dataset (Table 3) is involved in only one type of innovative link, either co-patenting (50.5%)<sup>19</sup>, co-publication (17.1%), or publicly funded collaborative R&D projects (22.0%). That only a small share of actors is recorded in more than one data source again supports our assertion that the use of only one type of data considerably underestimates regional innovative activity and knowledge transfer.

The bottom of Table 3 demonstrates that only 39.4% of all firms in our data are captured by co-patents. In other words, 60.6% of firms involved in regional knowledge transfer are neglected by this data source. In the co-publications data and the information on publicly funded collaborative R&D activities, the shares are 29.5% and 46.1%,

<sup>17</sup>Universities are grouped into only one industry (NACE 85.4). The same holds true for extra-university public research institutes (NACE 72). Activities of these publicly funded research institutions could be differentiated according to scientific disciplines. However, an overall classification scheme is not provided for German institutions.

<sup>18</sup>The reference year for determining age is 2010.

<sup>19</sup>The high share of actors that is only recorded in the patent statistics is particularly driven by the large number of 'other' actors representing patent applicants that could not be assigned to an institution.

**Table 3.** Actors in overlapping channels of knowledge transfer<sup>a</sup>.

Channels of knowledge transfer (pooled 2000–2010)	Types of actors				
	All actors	Firms	Universities	Research institutes	Other <sup>b</sup>
<i>Number of actors by type</i>					
Co-patents only	979 (50.5)	350 (31.5)	2 (10.0)	18 (15.7)	609 (87.8)
Co-publications only	331 (17.1)	230 (20.7)	1 (5.0)	28 (24.3)	72 (10.4)
Collaborative R&D only	427 (22.0)	391 (35.2)	6 (30.0)	24 (20.9)	6 (0.9)
Co-publications and co-patents	25 (1.3)	19 (1.7)	0 (0.0)	4 (3.5)	2 (0.03)
Collaborative R&D and co-patents	46 (2.4)	42 (3.8)	1 (5.0)	3 (2.6)	0 (0.0)
Collaborative R&D and co-publications	86 (4.4)	52 (4.7)	1 (5.0)	28 (24.3)	5 (0.7)
All layers	46 (2.4)	27 (2.4)	9 (45.0)	10 (8.7)	0 (0.0)
Total	1,940 (100.0)	1,111 (100.0)	20 (100.0)	115 (100.0)	694 (100.0)
<i>Sum of shares<sup>c</sup> (in %)</i>					
Co-patents	56.5	39.4	60.0	30.4	88.0
Co-publications	25.2	29.5	55.0	60.9	11.4
R&D collaborations	31.2	46.1	85.0	56.5	1.6

Notes: a) Numbers in parentheses represent the shares in percent. b) This category represents actors (mainly individual inventors) who could not be assigned to an institution because patent statistics do not list inventors' affiliations in some cases. c) The numbers indicate that the share of actors is captured by co-patents, co-publications, and (granted) R&D collaboration projects. Because of overlapping transfer channels, the sum of the shares is more than 100%.

Source: Own calculations.

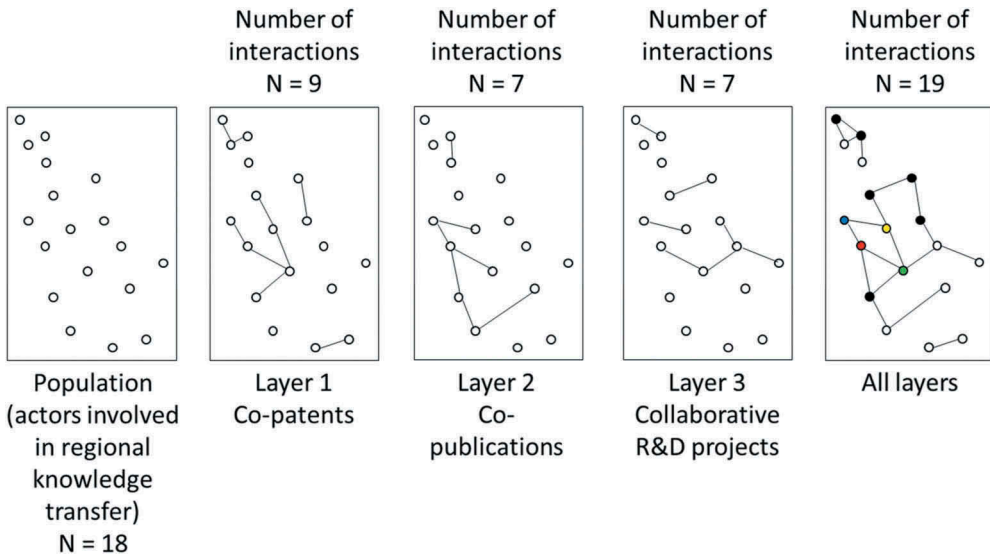
respectively. These figures strongly emphasize the necessity for an integrated and comprehensive approach in the study of regional innovation activity.

The picture changes considerably if we examine universities and extra-university public research institutes. A large share of universities is involved in multiple channels—nine out of twenty universities (45.0%) are part of all three types of collaborations. Although 85.0% of the universities recorded in the data have publicly funded R&D collaborations (bottom of Table 3), this share is much smaller for the other types of actors (56.5% of extra-university research institutes, 46.1% of private sector firms, and 1.6% of 'other' actors). A considerable share of the extra-university research institutes is involved in co-publication and publicly funded collaborative R&D activities; 60.9% of all extra-university public research institutes are covered by co-publications; and 56.5% are identified in the data on publicly funded collaborative R&D projects.

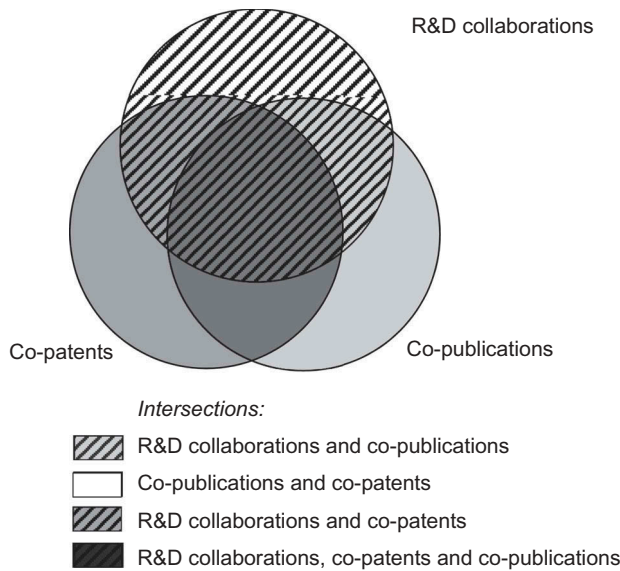
#### 4.2. Overlapping knowledge transfer channels

Each of the three data sources identified a certain type of relationship – co-patent, co-publication, or publicly funded collaborative R&D project. The network that can be constructed for a certain type of relationship forms a specific layer. Figure 2 illustrates how the separate analyses of single channels of knowledge transfer might conceal interactions that occur in another layer. The figure also demonstrates that the total main component based on all three data sources or channels of knowledge transfer is larger than those of each single layer.

To analyze how actors are involved in different forms of knowledge transfer, we form seven groups representing diverse forms of transfer (Figure 3). Table 4 presents the actors' involvement in these different transfer channels. In total, we find 27,434 interactions in the six case study regions in all three layers (column 'All actors' in Table 4). If we distinguish these interactions by type, we find that networks are dominated by co-publications. 75.4% of all



**Figure 2.** Combinations of the forms of cooperation.  
 Source: Own illustration.



**Figure 3.** Combinations of different forms of knowledge transfer.  
 Source: Own illustration.

interactions represent this type of knowledge exchange. The numbers of interactions in co-patents (3,802; 13.9%) and collaborative R&D projects (2,936; 10.7%) are considerably lower.

These results change dramatically if we distinguish by types of actors. While co-publications are the by far dominant form of interaction for universities and research institutes, we find that private firms are relatively more involved in co-patents (30.5%) and

**Table 4.** Channels of knowledge transfer<sup>a</sup>.

Channels of knowledge transfer (pooled 2000–2010)	Types of actors				
	All actors	Firms	Universities	Research institutes	Other <sup>b</sup>
<i>Number of interactions by type</i>					
Co-patents	3,802 (13.9)	1,824 (30.5)	256 (2.8)	409 (3.9)	1,313 (72.7)
Co-publications	20,696 (75.4)	2,469 (41.2)	8,354 (90.4)	9,398 (90.4)	475 (26.3)
Collaborative R&D	2,936 (10.7)	1,693 (28.3)	636 (6.9)	588 (5.7)	19 (1.1)
Total	27,434 (100.0)	5,986 (100.0)	9,246 (100.0)	10,395 (100.0)	1,807 (100.0)

Notes: a) Numbers in parentheses represent the shares in percent. b) This category represents actors (mainly individual inventors) who could not be assigned to an institution because patent statistics do not list inventors' affiliations in some cases.

Source: Own calculations.

collaborative R&D projects (28.3%). For 'other' actors who could not be assigned to an institution the dominant form of interaction is co-patenting (72.7% of all interactions).

These figures clearly show that investigating innovative relationships based only on co-patents neglects a large share of actual links that firms have. Co-publications and collaborative R&D projects capture more than half of the intraregional innovative relationships of private sector firms.

These findings are completely different from the results obtained for universities. About 90% of universities' knowledge links occur in co-publications. By contrast, the figures for the extra-university public research institutes point reveal a high share of co-patent applications.

### 4.3. Network dynamics

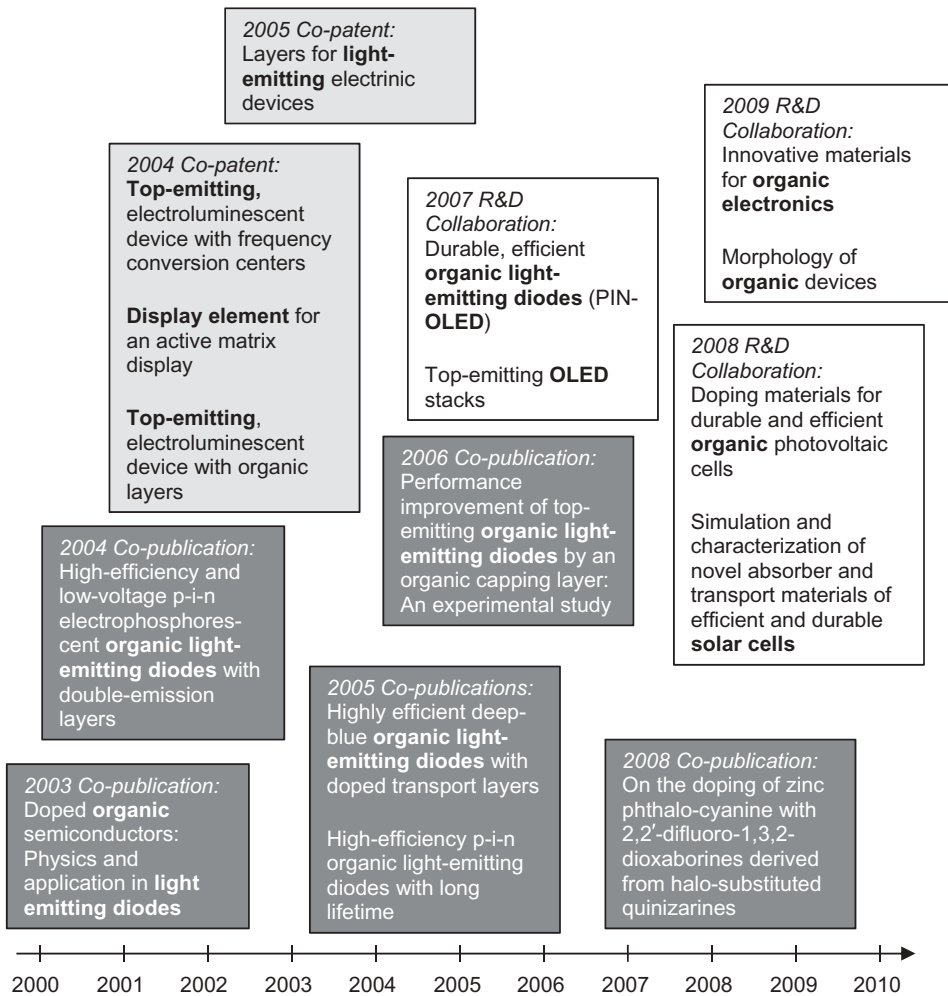
So far, we analyzed activities in different channels of knowledge transfers from a static perspective. Table 5 turns to a dynamic one and shows transitions from and into different types of innovative links. We observe how actors participate in different channels of knowledge transfer in two sub-periods (2000–2005 and 2006–2010). Similarly to Table 3 we find that a considerable share of actors is involved in only one way of knowledge transfer and use either co-patenting, co-publication or publicly funded collaborative R&D projects. A relatively small fraction uses different channels of knowledge transfer simultaneously. Table 5 also reveals that transitions between the different types of innovative links are relatively few. There are, however, high shares of actors who enter or exit a certain mode of collaboration. For example, 79.4% of actors who had a co-patent with another institution in the 2000–2005 period (85.4% who had co-patents only) use none of the three transfer channels in the 2006–2010 period. 56.9% of those with a publicly funded R&D collaboration in the first period (69.6% of actors with R&D collaboration only) have no collaboration at all in the second period.

These shares are lowest for co-publications: 39.8% of actors with a co-publication in the first period and 46.6 with co-publication only have no cooperation in the second period. The figures clearly show considerable levels of turbulence or 'fluidity' (Fritsch and Zoellner 2019) of links between actors. The figures at the diagonal line in Table 5 also show that there are also larger shares of actors that continue with a certain mode of collaboration or mix of modes. Beyond that, we also find relative large numbers at the

**Table 5.** Different channels of knowledge transfer<sup>ab</sup> – dynamics.

	2006–2010	No collaboration	Co-patents only	Co-publications only	Collaborative R&D only	Co-publications and co-patents	Collaborative R&D and co-patents	Collaborative R&D and co-publications	All Layers	Total
No collaboration			336 (49.6)	147 (21.7)	159 (23.5)	7 (1.0)	11 (1.6)	13 (1.9)	4 (0.6)	677 (100.0)
Co-patents only	556 (85.4)		87 (13.4)	1 (0.2)	3 (0.5)	1 (0.2)	1 (0.2)	1 (0.2)	1 (0.2)	651 (100.0)
Co-publications only	95 (46.6)		3 (1.5)	89 (43.6)	3 (1.5)	3 (1.5)	1 (0.5)	9 (4.4)	1 (0.5)	204 (100.0)
Collaborative R&D only	208 (69.6)		8 (2.7)	9 (3.0)	60 (20.1)	0 (0.0)	3 (1.0)	11 (3.7)	0 (0.0)	299 (100.0)
Co-publications and co-patents	3 (18.8)		0 (0.0)	6 (3.5)	3 (18.8)	1 (6.3)	0 (0.0)	0 (0.0)	3 (18.8)	16 (100.0)
Collaborative R&D and co-patents	8 (36.4)		5 (22.7)	0 (0.0)	2 (9.1)	0 (0.0)	5 (22.7)	1 (4.6)	1 (4.6)	22 (100.0)
Collaborative R&D and co-publications	7 (14.3)		1 (2.0)	13 (26.5)	2 (4.1)	0 (0.0)	1 (2.0)	19 (38.8)	6 (12.2)	49 (100.0)
All Layers	0 (0.0)		0 (0.0)	3 (13.6)	0 (0.0)	1 (4.6)	0 (0.0)	5 (22.7)	13 (59.1)	22 (100.0)
Total	877 (45.2)		440 (22.7)	268 (13.8)	232 (12.0)	13 (0.7)	22 (1.1)	59 (3.0)	29 (1.5)	1,940 (100.0)

Notes: a) Numbers in parentheses represent the shares in percent. b) Highlighted figures represent shares > 10%.  
Source: Own calculations.



**Figure 4.** Sequences in the participation in different channels of knowledge transfer – collaborations between Technical University Dresden and Novaled.

Notes: German titles of co-patent applications and subsidized collaborative R&D projects were translated into English. Source: Own illustration.

main diagonal indicating that actors obviously prefer specific types of knowledge transfer over time. Transitions between different modes of collaboration are, however, rare.

Our database contains selected numbers of examples of collaborations with different types of innovative links. As we stated above the use of overlapping channels of knowledge transfer is a rare event. Against this backdrop it is hard to identify any pattern of sequences for the use of specific forms of knowledge transfer. As an example Figure 4 presents cooperation projects between Technical University Dresden and Novaled. The firm (<http://www.novaled.com/homepage/>) is a semiconductor producer with a focus on organic light-emitting diodes. Novaled started its business in 2001. The headquarter is located in Dresden. In 2014 the firm had 135 employees, and the operating revenue amounted to approximately 42 Mio. Euro with total assets of about 32 Mio. Euro. Figure 4 shows the different types of cooperation between Novaled and the Technical University Dresden during the period of analysis. Based



on [Figure 4](#) there is no clear pattern of subsequent types of cooperation such as first collaborative R&D, then patenting, and finally publication. On the contrary, the cooperative relationships in the first years are publishing and patenting followed by publicly subsidized R&D collaboration. The firm is, however, one of the rather few examples that uses all three modes of cooperation in both sub-periods.

#### **4.4. Network descriptives**

The previous three sections have demonstrated that the numbers and shares of regional interactions differ considerably across the three types of data. [Table 6](#) depicts main network descriptives for the six regions under study. The first row (all layers, nodes) contains the number of actors involved in regional knowledge transfer as documented in at least one of the three data sources.

There is huge variation across regions with regard to the number of nodes involved that indicates rather different levels of innovation activity. In each region, the largest number of actors involved in regional knowledge transfer is found in the co-patent layers. This result is certainly shaped by the large number of ‘other’ actors, for example, inventors who could not be assigned to an institution, which is the unit of analysis in our study. The shares of all regional actors recorded in the patent statistics is particularly high in the regions of Kassel and Siegen, which have relatively low numbers of actors involved in the other two forms of intraregional knowledge transfer. The share of actors participating in regional collaborative R&D projects ranges between 11.7% in Siegen to 43.9% in Magdeburg. For co-publications, these figures vary between 9.7% in Siegen and 30.3% in Aachen.

A large share of actors in the largest component of the comprehensive network (between 33.3% and 51.7%) is active in the co-publication layer. The other two layers demonstrate smaller shares of actors in the largest component of the comprehensive network and larger deviations of these shares across regions. In comparison to the baseline scenario, the share of actors in the main component is highest for the co-publications and publicly funded collaborative R&D projects and lowest for the interactions based on co-patents. In other words, we observe a high concentration of regional actors in co-publication and R&D collaboration networks, whereas the co-patent networks tend to be more dispersed. The graphical representations of the regional co-publication networks ([Figures 5\(c\)](#) and [B1\(c\)–B5\(c\)](#) in [Appendix B](#)) strongly indicate a key role of universities in co-publication networks.

Density in all regions increases when the three layers are put together. The intensity of interaction is considerably underestimated when only one type of innovative connection is considered. However, comparing networks based on their density might not be appropriate if there are significant differences in the numbers of network nodes. Because the number of potential links increases more than proportional with the number of actors, larger networks may have lower density values.

Magdeburg and Rostock have higher densities than Aachen and Dresden, although they have fewer nodes. Remarkably, this finding does not hold for Kassel and Siegen, who are also characterized by relatively fewer nodes. This density size bias is apparent when the density and fragmented density (that only considers active actors within a certain layer) measures are compared. The differences between the two measures can be particularly pronounced in networks with smaller numbers of nodes.

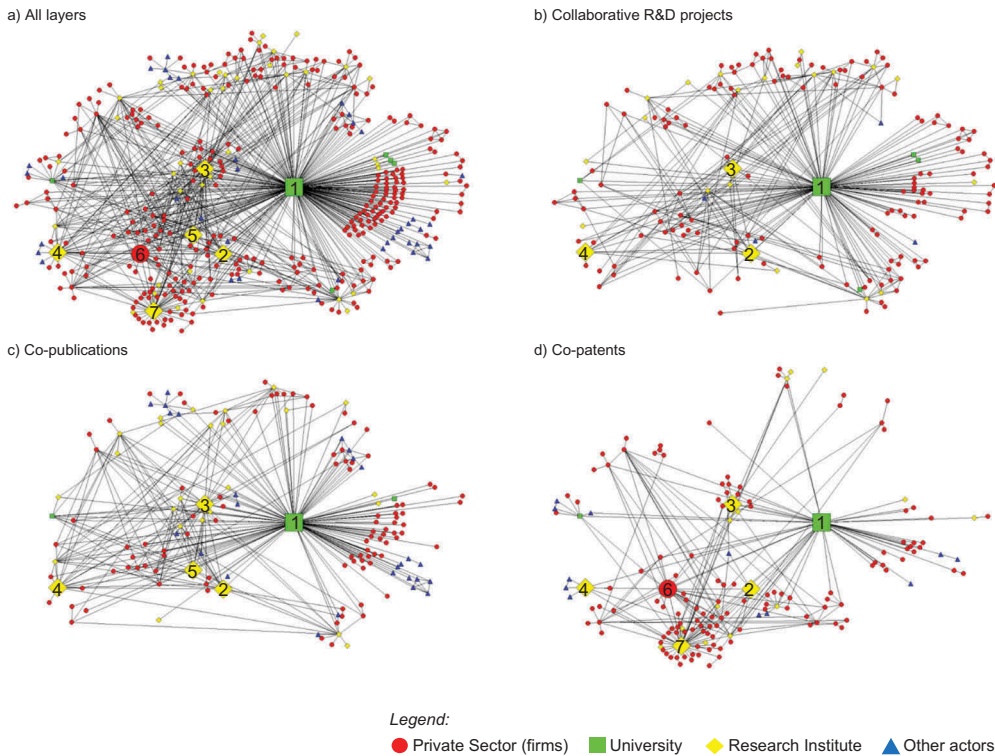


**Table 6.** Network descriptives for the data sources and regions.

	Average	Aachen	Dresden	Kassel	Magdeburg	Rostock	Siegen
<i>All layers (baseline scenario)</i>							
Number of nodes <sup>b</sup>	323.3 (100)	581 (100)	588 (100)	145 (100)	278 (100)	245 (100)	103 (100)
Share of actors in the largest component <sup>a,c</sup> (%)	50.0	54.9	68.9	30.3	63.3	59.2	23.3
Density	0.195	0.221	0.225	0.033	0.288	0.366	0.038
Mean degree	10.71	17.24	17.76	2.43	11.81	12.55	2.47
Mean degree (binary)	2.53	2.43	3.20	1.53	2.98	3.59	1.42
<i>Collaborative R&amp;D projects</i>							
Number of nodes <sup>b</sup>	100.5 (31.1)	158 (27.2)	206 (35.0)	18 (12.4)	122 (43.9)	87 (35.5)	12 (11.7)
Share of actors in the largest component (%)	82.1	91.1	83.0	66.7	78.7	89.7	83.3
Density	0.008	0.020	0.003	0.002	0.008	0.012	0.002
Density fragmented <sup>d</sup>	0.081	0.027	0.023	0.144	0.040	0.086	0.167
Mean degree	1.35	1.14	1.62	0.30	2.11	2.74	0.21
Mean degree fragmented <sup>d</sup>	4.24	4.19	4.61	2.44	4.80	7.55	1.83
Mean degree (binary)	1.03	0.84	1.23	0.28	1.68	1.91	0.21
Mean degree (binary) fragmented	3.29	3.09	3.52	2.22	3.82	5.26	1.83
<i>Co-publications</i>							
Number of nodes <sup>b</sup>	81.5 (25.2)	176 (30.3)	154 (26.4)	21 (14.5)	67 (24.1)	60 (24.5)	10 (9.7)
Share of actors in the largest component (%)	94.6	93.8	100	100	97.0	96.7	80.0
Density	0.020	0.025	0.022	0.006	0.030	0.035	0.004
Density fragmented <sup>d</sup>	0.408	0.277	0.322	0.281	0.527	0.595	0.444
Mean degree	7.62	14.66	12.90	0.81	8.38	8.60	0.39
Mean degree fragmented <sup>d</sup>	29.53	48.40	49.25	5.62	34.78	35.10	4.00
Mean degree (binary)	0.63	0.88	1.03	0.28	0.66	0.76	0.16
Mean degree (binary) fragmented	2.70	2.90	3.95	1.90	2.72	3.10	1.60
<i>Co-patents</i>							
Number of nodes <sup>b</sup>	184.7 (57.1)	320 (55.1)	335 (57.0)	113 (77.9)	134 (48.2)	121 (49.4)	85 (82.5)
Share of actors in the largest component (%)	19.7	18.8	47.2	13.3	17.9	14.0	7.1
Density	0.008	0.003	0.006	0.009	0.005	0.005	0.018
Density fragmented <sup>d</sup>	0.019	0.008	0.018	0.015	0.022	0.022	0.026
Mean degree	1.74	1.44	3.25	1.31	1.32	1.26	1.86
Mean degree fragmented <sup>d</sup>	2.96	2.63	5.76	1.70	2.85	2.57	2.23
Mean degree (binary)	1.02	0.99	1.20	1.02	0.79	1.05	1.07
Mean degree (binary) fragmented	1.71	1.64	2.14	1.32	1.71	2.15	1.28

Notes: a) In each case study region, universities belong to the largest component. An exception is Siegen, where the university is not part of the regional network of co-patents. b) Numbers in parentheses represent the relation to the number of nodes in the base line scenario in percent. Please, note that the numbers' sum does not equal 100% because some institutions are involved in several layers. c) Numbers in parentheses represent the relation to the number of nodes in the main component in the base line scenario in percent. The numbers' sum does not equal 100% because some institutions are involved in several layers. d) Fragmented network measures only include actors that are active within this layer, whereas non-fragmented network measures include all identified actors in the respective region (= nodes of the baseline scenario).

Source: Own calculations.



**Figure 5.** Networks of the Dresden region (largest component, period 2000–2010). (a) All layers (b) Collaborative R&D projects (c) Co-publications (d) Co-patents.

Source: Authors' own illustration.

Similar to the density measure result, the mean degree and binary mean degree are the highest in the combined layer for all the regions. An analysis of single layers reveals that the mean degrees in the co-publication layer tend to be considerably higher than for collaborative R&D projects and co-patents. The six regions have considerable differences regarding the number of nodes and characteristics of their networks.

The two regions with the lowest numbers of innovative actors, Kassel and Siegen, have the lowest fragmented mean degrees in each of the network types. Generally, smaller networks tend to have lower shares of actors in the largest component and a lower mean degree. However, we have to consider here, that particularly small regions tend to have a relatively large share of inter-regional collaborations, which could not be included in our analysis due to data restrictions.

#### 4.5. Illustration: Dresden

As an illustration of the scope of the data sources for identifying R&D cooperation, we provide graphical representations of the networks in Dresden. We focus on Dresden because it has the largest number of actors and the largest main component.<sup>20</sup>

<sup>20</sup>See the network graphs for all case study regions in [Appendix B](#).

Figure 5(a–d) depict the main components of the network layers for the Dresden region. The circles, rectangles, and diamonds represent private sector firms, universities, and extra-university public research institutes, respectively. The remaining category of actors (triangles) captures individuals that could not be assigned to an organization in the record-linkage procedure. Cooperative relationships between actors (linkages) are represented by straight lines.<sup>21</sup> The most central actors in the entire multilayer network (Figure 5(a)) are identified as actors 1 to 5. The first four are in all three data sources.

We observe tremendous differences between the networks based on only one of the three data sources. These differences clearly demonstrate that each of these datasets covers only a specific part of the overall regional knowledge transfer. The most comprehensive picture of the relationships is provided by a combination of all three sources of information. All four graphs also clearly demonstrate that the Technical University of Dresden assumes a central position and a crucial broker role for the network. However, this result requires some qualification. One should be aware that universities are large multi-department units where there is great diversity of knowledge fields across academic disciplines. Hence, knowledge flows between the different departments will be far from perfect. This means that if the Technical University of Dresden is connected to two firms through different departments we cannot necessarily assume that there is an indirect link between the two firms. One may, however, argue in this case that there is at least the opportunity for an indirect link between two departments as they face the same institutional framework of the university.

The extra-university public research institutes are connected to universities and each other, and most of these institutes maintain many links to private firms (Figure 5(a)). This information indicates that the knowledge in these institutes is valuable to regional firms and transferred into the regional economy.

Approximately two-thirds of all actors in Dresden (405 out of 588) are present in the main component of the comprehensive network (Figure 5(a)): this share is 83% of the network based on collaborative R&D projects (Figure 5(b)), 100% of the co-publication network (Figure 5(c)), and 47.2% of the patent network (Figure 5(d)). The patent network's relatively high level of fragmentation is, to a considerable extent, because of the 'other' actors that could not be assigned to an institution (see Table 3). Only twelve of these 'other' actors (7.6%) are connected to the main component of the patent network in Dresden (Figure 5(d)). Another reason for the differences among the patent network and networks based on co-publications and collaborative R&D projects is that it was impossible to assign the patents of the twelve Fraunhofer Institutes located in Dresden to the single institutes.<sup>22</sup>

## 5. Discussion and conclusions

### 5.1. Research contribution

We have constructed regional innovation networks based on the following types of data: patents, publications, and publicly subsidized R&D collaborations. By applying

<sup>21</sup>The position of nodes was produced using the spring embedding method (see Brandes 2001). For clarity, we do not attempt to represent the strength of a link or the number of patents, publications, and R&D projects of an actor by the thickness of an edge or the size of a node.

<sup>22</sup>The reason for this is that patenting of all Fraunhofer institutes is centrally managed at the headquarters of its society in Munich.

comprehensive record-linking techniques, we merged the three databases at the level of institutions. We observe that this combined network provides a much more comprehensive picture of regional innovative interactions than networks constructed by using only one or two data sources. Our comparisons make clear that the results of social network analyses can be considerably shaped by the characteristics of the respective database and that one should be well aware of such biases when interpreting the respective results.

A comparison of the networks based on the sources of data also allows us to assess the bias of each data source in capturing cooperative relationships. We observe that universities tend to be well-represented in all three types of data, whereas private sector firms are particularly included in publicly subsidized R&D collaboration. Our analyses suggest that patent statistics—the most frequently used database for constructing innovation networks—tend to underestimate the links of private sector firms. An obvious reason for this pattern is that patents tend to represent activities in the field of knowledge exploration, which is the domain of universities, whereas the R&D collaboration of private firms represents additional activities that are mainly knowledge exploitation. The data on co-publications add many links not identified in the patent statistics and in the data on publicly subsidized R&D collaborations. The main reason for this observation is probably that patents and publicly subsidized R&D collaborations primarily represent links that focus on the development of technologies, whereas co-publications cover a much wider spectrum of knowledge fields.

Despite such biases and incomplete representations, our analyses demonstrate the importance of R&D cooperation and division of innovative labor for innovation processes. In particular, the key role of universities and other public research organizations as brokers who link many actors and ‘organize’ regional innovation networks is obvious. As universities act as broker in the co-patent, the co-publication, and the research network they also link these three layers. Organizations that are active in only one layer are linked via universities who should be able to work as a translator for different forms of knowledge.

Moreover, our analyses reveal immense differences across the sample’s regions, regarding the intensity of networking. Such differences in the levels of cooperative relationships reflect divergent intensities of division of innovative labor that can have critical consequences for the efficiency of innovation processes at the level of individual actors and the respective regional innovation system as a whole.

The pronounced role of public research institutions, particularly of universities in regional innovation networks, qualifies them as crucial starting points for policy measures that aim to stimulate knowledge transfer and division of innovative labor in regional innovation systems (RIS). Hence, our analyses corroborate that policies aiming at stimulating the links between public research and private sector firms to improve knowledge transfer in RIS are highly appropriated.

## ***5.2. Limitations and suggestions for further research***

Although we provided new empirical evidence on the measurement of cooperation in RIS, the analyses have shortcomings that could represent starting points for further research. The main limitation of our analyses is that we only considered formal links and did not capture informal relationships. Although it is plausible to assume that many formal links are embedded in informal relationships, it would be desirable to identify these informal links directly. Moreover, we identified only intraregional links,

the ‘local buzz’ (Bathelt, Malmberg, and Maskell 2004; Storper and Venables 2004). To complement this picture, further work should include and analyze the differences among the databases in capturing inter-regional links, the ‘global pipelines.’ This inclusion would facilitate the identification and analyses of the role of gatekeepers in a RIS that is well-connected to other actors inside and outside a region (Graf 2011).

Because our data did not permit the identification of actors within private firms involved in an R&D project, we were unable to merge the three databases at the level of individuals. Hence, we had to choose the level of institutions—firms, universities, other public research institutions—as the smallest unit of observation. A main advantage of data at the level of individuals would be the possibility of including mobility across institutions as a link (Graf 2006).

The considerable differences we observed among the levels of R&D cooperation and structures of the innovation networks deserves an explanation. Given the strong role of universities in regional innovation networks, the number and size of the regional universities and their fields of knowledge may provide such an explanation. The fields of knowledge should play a role when included in a certain type of database. For example, there is good reason to expect that university researchers in the natural sciences and engineering have a much higher propensity to apply for a patent than researchers in the social and administrative sciences (Arundel and Kabla 1998; Fritsch and Aamoucke 2017). Moreover, private sector firms may find more interesting opportunities for R&D cooperation with the technologically oriented departments of a university than with, for example, humanities. Another crucial factor may be the correspondence of the knowledge fields in public and private research, in that high levels of correspondence lead to high levels of cooperation (Fritsch and Slavtchev 2011).

A further limitation of our study is that the data is limited to six regions. Because of this small number, we cannot apply statistical methods to investigate the relationship between network structure and performance of the respective RIS across regions. Hence, it would be desirable to have comparable information on a larger set of regions to have sufficient numbers of observations to perform an econometric analysis. A further shortcoming of our data is that our sample of regions does not include large, high-density centers of innovation activity, such as Munich and Stuttgart. Having such information available would allow for interesting comparisons of RIS with very different numbers of actors and degrees of density.

Because two of the three data sources (i.e. patents and collaborative R&D projects) are more or less entirely limited to analytical and synthetic types of knowledge, we cannot exclude the possibility that the links identified primarily represent the transfer of such kinds of knowledge, whereas the transfer of other types of knowledge (e.g. symbolic knowledge) may only be included in links identified by co-publications.<sup>23</sup>

Furthermore, our approach may contribute to theory development because it enables the identification of the different forms of knowledge transfer during different stages of the innovation process. Authors have claimed that certain types of actors have a particularly pronounced role in the different stages of the innovation process. A common assumption is that universities are primarily involved in knowledge exploration, whereas the activities of private firms tend to be in the field of knowledge exploitation that is, transferring knowledge to commercial applications (Mowery and Sampat 2006). Hence, a promising step for

---

<sup>23</sup>For a detailed characterization of the three types of knowledge bases see Asheim et al. (2007).

future research would be to create a longer time-series and assess the roles of the different forms of knowledge transfer along the stages of the innovation process.

## Acknowledgments

We are indebted to Holger Graf, Matthias Brachert, Stefan Luethi, and Moritz Zoellner for their helpful comments on earlier versions of this paper. Particular thanks go to Wilfried Ehrenfeld and Alexander Giebler for technical support with the data preparation. Finally, we would like to thank the two anonymous referees for their helpful comments and suggestions. The underlying data was generated in the framework of a research project that investigated the future role of universities in regions with a declining population, financed by the German Federal Ministry of Education and Research (grant 01PW11011C).

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Funding

The underlying data was generated in the framework of a research project that investigated the future role of universities in regions with a declining population, financed by the German Federal Ministry of Education and Research.

## ORCID

Michael Fritsch  <http://orcid.org/0000-0003-0337-4182>

Mirko Titze  <http://orcid.org/0000-0002-4479-9534>

## References

- Abramo, G., C. A. D'Angelo, and A. Caprasecca. 2009. "Allocative Efficiency in Public Research Funding: Can Bibliometrics Help?" *Research Policy* 38 (1): 206–215. doi:10.1016/j.respol.2008.11.001.
- Adams, J. D., and A. N. Link. 2018. "The Structure and Performance of U.S. Research Joint Ventures: Inferences and Implications from the Advanced Technology Program." *Economics of Innovation and New Technology* 27: 551–575. doi:10.1080/10438599.2017.1376169.
- Arundel, A., and I. Kabla. 1998. "What Percentage of Innovations are Patented? Empirical Estimates for European Firms." *Research Policy* 27: 127–141. doi:10.1016/S0048-7333(98)00033-X.
- Asheim, B., L. Coenen, J. Moodysson, and J. Vang. 2007. "Constructing Knowledge-based Regional Advantage: Implications for Regional Innovation Policy." *International Journal of Entrepreneurship and Innovation Management* 7 (2–5): 140–155. doi:10.1504/IJEIM.2007.012879.
- Barber, M. J., and T. Scherngell. 2013. "Is the European R&D Network Homogeneous? Distinguishing Relevant Network Communities Using Graph Theoretic and Spatial Interaction Modelling Approaches." *Regional Studies* 47 (8): 1283–1298. doi:10.1080/00343404.2011.622745.
- Bathelt, H., A. Malmberg, and P. Maskell. 2004. "Cluster and Knowledge: Local Buzz, Global Pipelines and the Process of Knowledge Creation." *Progress in Human Geography* 28: 31–56. doi:10.1191/0309132504ph469oa.
- BBSR, Bundesinstitut fuer Bau-, Stadt und Raumforschung. 2015. "Types of Regions." Accessed 6 June 2017. <http://www.bbsr.bund.de/BBSR/DE/Raumbeobachtung/Raumabgrenzungen/Regionstypen/regionstypen.html?nn=443048>

- Blind, K., J. Edler, R. Frietsch, and U. Schmoch. 2006. "Motives to Patent: Empirical Evidence from Germany." *Research Policy* 35 (5): 655–672. doi:10.1016/j.respol.2006.03.002.
- Brandes, U. 2001. "Drawing on Physical Analogies." In *Drawing Graphs: Methods and Models*, edited by M. Kaufmann and D. Wagner, 71–86. Berlin: Springer Verlag Berlin.
- Breschi, S., and L. Cusmano. 2004. "Unveiling the Texture of a European Research Area: Emergence of Oligarchic Networks under EU Framework Programmes." *International Journal of Technology Management* 27 (8): 747–772. doi:10.1504/IJTM.2004.004992.
- Broekel, T., D. Fornahl, and A. Morrison. 2015. "Another Cluster Premium: Innovation Subsidies and R&D Collaboration Networks." *Research Policy* 44: 1431–1444. doi:10.1016/j.respol.2015.05.002.
- Broekel, T., and H. Graf. 2010. "Structural Properties of Cooperation Networks in Germany: From Basic to Applied Research." *Jena Economic Research Papers*. 2010–2078.
- Broekel, T., and H. Graf. 2012. "Public Research Intensity and the Structure of German R&D Networks: A Comparison of 10 Technologies." *Economics of Innovation and New Technology* 21 (4): 345–372. doi:10.1080/10438599.2011.582704.
- Cohen, W. M., R. R. Nelson, and J. P. Walsh. 2000. "Protecting Their Intellectual Assets: Appropriability Conditions and Why US Manufacturing Firms Patent (or Not)." *National Bureau of Economic Research* (No. w7552).
- Contractor, F. J., and J. J. Reuer, eds. 2018. *Frontiers of Strategic Alliance Research: Negotiating, Structuring and Governing Partnerships*. Cambridge: Cambridge University Press. doi:10.1017/9781108236188.
- Ehrenfeld, W. 2015a. *RegDemo: Preparation and Merger of Actor Data – Technical Documentation of Routines and Datasets*. IWH Technical Reports 01/2015. Halle (Saale).
- Ehrenfeld, W. 2015b. *RLPC: Record Linkage Pre-Cleaning – Technical Documentation of Routines*. IWH Technical Reports 02/2015. Halle (Saale).
- Fischer, M. M., and D. A. Griffith. 2008. "Modeling Spatial Autocorrelation in Spatial Interaction Data: An Application to Patent Citation Data in the European Union." *Journal of Regional Science* 48 (5): 969–989. doi:10.1111/j.1467-9787.2008.00572.x.
- Fischer, M. M., T. Scherngell, and E. Jansenberger. 2006. "The Geography of Knowledge Spillovers between High-technology Firms in Europe: Evidence from a Spatial Interaction Modelling Perspective." *Geographical Analysis* 38: 288–309. doi:10.1111/j.1538-4632.2006.00687.x.
- Fritsch, M., and H. Graf. 2011. "How Sub-national Conditions Affect Regional Innovation Systems: The Case of the Two Germanys." *Papers in Regional Science* 90 (2): 331–353. doi:10.1111/j.1435-5957.2011.00364.x.
- Fritsch, M., and M. Zoellner. 2019. "The Fluidity of Inventor Networks." *Journal of Technology Transfer*. doi:10.1007/s10961-019-09726-z.
- Fritsch, M., and R. Aamoucke. 2017. "Fields of Knowledge in Higher Education Institutions, and Innovative Start-Ups—An Empirical Investigation." *Papers in Regional Science* 96: 1–27. doi:10.1111/pirs.12175.
- Fritsch, M., and V. Slavtchev. 2011. "Determinants of the Efficiency of Regional Innovation Systems." *Regional Studies* 45: 905–918. doi:10.1080/00343400802251494.
- Graf, H. 2006. *Networks in the Innovation Process: Local and Regional Interactions*. Cheltenham: Elgar.
- Graf, H. 2011. "Gatekeepers in Regional Networks of Innovators." *Cambridge Journal of Economics* 35 (1): 173–198. doi:10.1093/cje/beq001.
- Graf, H., and T. Henning. 2009. "Public Research in Regional Networks of Innovators: A Comparative Study of Four East German Regions." *Regional Studies* 43: 1349–1368. doi:10.1080/00343400802251460.
- Griliches, Z. 1990. "Patent Statistics as Economic Indicators: A Survey." *National Bureau of Economic Research* (No. w3301). doi:10.1099/00221287-136-2-327.
- Hanneman, R. A., and M. Riddle. 2005. *Introduction to Social Network Methods*. Riverside (CA): University of California at Riverside. [http://faculty.ucr.edu/~hanneman/nettext/Introduction\\_to\\_Social\\_Network\\_Methods.pdf](http://faculty.ucr.edu/~hanneman/nettext/Introduction_to_Social_Network_Methods.pdf)
- Hoekman, J., K. Frenken, and F. van Oort. 2009. "The Geography of Collaborative Knowledge Production in Europe." *The Annals of Regional Science* 43 (3): 721–738. doi:10.1007/s00168-008-0252-9.



- Hoekman, J., K. Frenken, and R. J. Tijssen. 2010. "Research Collaboration at a Distance: Changing Spatial Patterns of Scientific Collaboration within Europe." *Research Policy* 39 (5): 662–673. doi:10.1016/j.respol.2010.01.012.
- Kauffeld-Monz, M., and M. Fritsch. 2013. "Who are the Knowledge Brokers in Regional Systems of Innovation? A Multi-Actor Network Analysis." *Regional Studies* 47: 669–685. doi:10.1080/00343401003713365.
- Lata, R., T. Scherngell, and T. Brenner. 2015. "Integration Processes in European Research and Development: A Comparative Spatial Interaction Approach Using Project Based Research and Development Networks, Co-Patent Networks and Co-Publication Networks." *Geographical Analysis* 47 (4): 349–375. doi:10.1111/gean.12079.
- Link, A. N., and C. Antonelli, eds. 2018. *Strategic Alliances: Leveraging Economic Growth and Development*. London: Routledge.
- Maggioli, M. A., M. Nosvelli, and T. E. Uberti. 2007. "Space versus Networks in the Geography of Innovation: A European Analysis." *Papers in Regional Science* 86 (3): 471–493. doi:10.1111/j.1435-5957.2007.00130.x.
- Mansfield, E., M. Schwartz, and S. Wagner. 1981. "Imitation Costs and Patents: An Empirical Study." *The Economic Journal* 91 (364): 907–918. doi:10.2307/2232499.
- Meyer, M. 2002. "Tracing Knowledge Flows in Innovation Systems." *Scientometrics* 54 (2): 193–212. doi:10.1023/A:1016057727209.
- Mowery, D. C., and B. N. Sampat. 2006. "Universities in National Innovation Systems." In *The Oxford Handbook of Innovation*, edited by J. Fagerberg, D. C. Mowery, and R. R. Nelson, 209–239. Oxford: Oxford University Press.
- Peng, L., and X. Haoxiang. 2015. "Structure and Evolution of Co-authorship Network in an Interdisciplinary Research field." *Scientometrics* 103: 101–134. doi:10.1007/s11192-014-1525-y.
- Ponds, R., F. van Oort, and K. Frenken. 2007. "The Geographical and Institutional Proximity of Research Collaboration." *Papers in Regional Science* 86 (3): 423–443. doi:10.1111/j.1435-5957.2007.00126.x.
- Research Explorer. 2017. German Research Institutions. Accessed 6 June 2017. [http://www.research-explorer.de/research\\_explorer.de.html](http://www.research-explorer.de/research_explorer.de.html)
- Scherngell, T., and M. J. Barber. 2011. "Distinct Spatial Characteristics of Industrial and Public Research Collaborations: Evidence from the Fifth EU Framework Programme." *The Annals of Regional Science* 46 (2): 247–266. doi:10.1007/s00168-009-0334-3.
- Scherngell, T., and R. Lata. 2013. "Towards an Integrated European Research Area? Findings from Eigenvector Spatially Filtered Spatial Interaction Models Using European Framework Programme Data." *Papers in Regional Science* 92 (3): 555–577. doi:10.1111/j.1435-5957.2012.00419.x.
- Schmoch, U. 1999. "Interaction of Universities and Industrial Enterprises in Germany and the United States. A Comparison." *Industry and Innovation* 6 (1): 51–68. doi:10.1080/13662719900000004.
- Storper, M., and A. J. Venables. 2004. "Buzz: Face-to-face Contact and the Urban Economy." *Journal of Economic Geography* 4: 351–370. doi:10.1093/jnlecg/lbh027.
- Ter Wal, A., and R. Boschma. 2009. "Applying Social Network Analysis in Economic Geography: Framing Some Key Analytic Issues." *Annals of Regional Science* 43 (3): 739–756. doi:10.1007/s00168-008-0258-3.
- Titze, M., M. Schwartz, and M. Brachert. 2012. "A Systemic View on Knowledge-based Development Metrics." *International Journal of Knowledge-Based Development* 3 (1): 35–57. doi:10.1504/IJKBD.2012.045559.
- Tomasello, M. V., M. Napoletano, A. Garas, and F. Schweitzer. 2017. "The Rise and Fall of R&D Networks." *Industrial and Corporate Change* 26: 617–646. doi:10.1093/icc/dtw041.
- Wuchty, S., B. F. Jones, and B. Uzzi. 2007. "The Increasing Dominance of Teams in Production of Knowledge." *Science* 316: 1036–1039. doi:10.1126/science.1136099.
- Youtie, J., and P. Shapira. 2008. "Mapping the Nanotechnology Enterprise: A Multi-indicator Analysis of Emerging Nanodistricts in the US South." *The Journal of Technology Transfer* 33 (2): 209–223. doi:10.1007/s10961-007-9076-5.

## Appendix A

**Table A1.** Case study regions at a glance.

Planning Region	Aachen	Dresden	Siegen	Rostock	Kassel	Magdeburg
<i>Macro-region in Germany</i>	West	East	West	East	West	East
Population 2000	1,282,164	1,022,527	431,845	424,191	902,491	993,891
Annual change 2000–2010 (%)	0.2	–0.0	–0.4	–0.3	–0.4	–0.9
<i>Private sector 2000</i>						
Number of establishments <sup>b</sup>	28,753	27,868	9,952	11,386	21,213	24,714
Annual change 2000–2008 (%)	–0.6	–1.0	–0.9	–1.3	–0.7	–1.4
Number of employees 2000	239,343	231,352	113,680	83,781	185,882	194,111
Annual growth 2000–2010 (%)	–0.6	–0.1	–0.3	–0.6	–0.6	–0.6
Share of R&D employees 2000 (%) <sup>c</sup>	4.7	4.5	2.0	2.6	2.4	2.2
Annual change 2000–2008 (%)	0.0	0.1	0.0	–0.0	0.1	–0.0
<i>Research sector (2000)</i>						
Number of research institutes <sup>a</sup>	21	38	0	14	7	19
Number of universities <sup>ad</sup>	3	10	1	2	3	4
Total research teaching staff <sup>d</sup>	4,898	4,715	837	1,958	1,389	1,988
Annual change 2000–2010 (%)	4.1	4.3	4.7	2.2	8.4	0.8
Share of research and teaching staff in natural sciences & engineering (%) <sup>de</sup>	61.7	53.0	50.6	38.5	50.4	37.6
Annual change 2000–2010 (%)	–0.1	0.3	–0.7	0.2	–0.6	0.3
Number of professors <sup>d</sup>	649	820	231	299	318	392
Annual change 2000–2010 (%)	0.6	–0.4	0.4	–0.5	2.9	0.4
Share of professors in natural sciences and engineering (%) <sup>d</sup>	64.9	54.6	48.3	43.3	47.1	42.5
Annual change 2000–2010 (%)	–0.4	–0.1	–0.8	–0.9	–0.8	0.2
<i>Inter-regional collaboration intensity<sup>g</sup> (pooled 2000–2010)</i>						
Granted R&D collaboration projects <sup>h</sup>	82.8	79.3	95.8	60.8	86.5	66.9
Co-patent applications (inventors) <sup>i</sup>	32.7	37.0	52.9	56.7	50.9	45.9
Co-patent applications (applicants) <sup>i</sup>	59.6	59.3	74.3	60.6	57.0	56.0

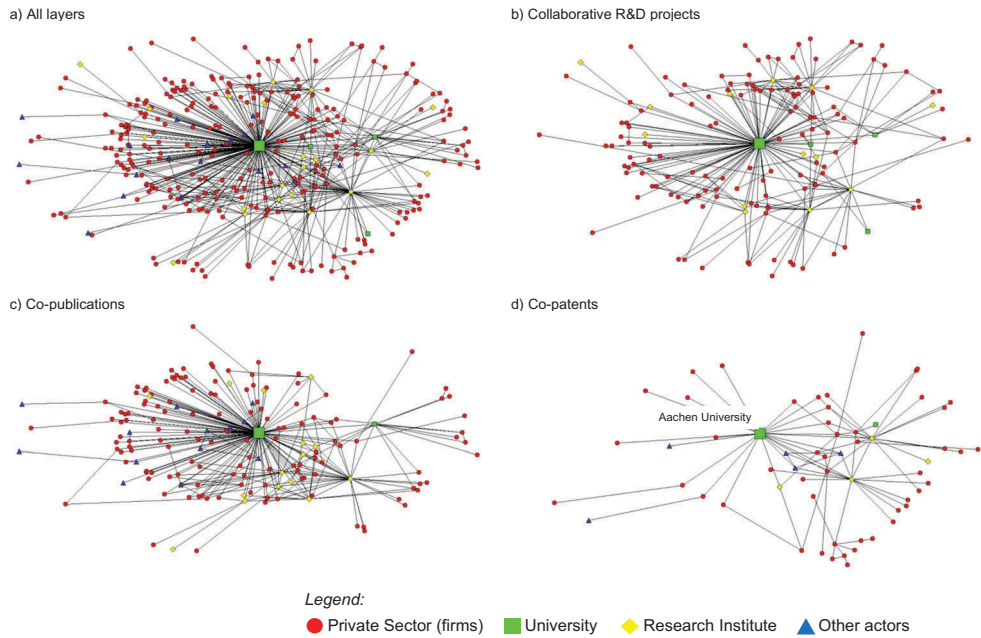
Notes: a) These figures are reported for the year 2013. b) Includes all establishments with at least one employee. c) Employees with a tertiary education in natural science or engineering. d) Includes research universities and technical colleges ('Fachhochschulen'). e) Includes three groups of scientific disciplines: natural sciences, agricultural and nutritional sciences, and engineering. Excludes medical sciences, cultural and social sciences, law and economics, and arts. f) Total of private and public sector. g) Number of inter-regional linkages as a percentage of total number of interactions. h) Calculated at the basis of the most recent version of the Foerderkatalog available for us (2017). i) Calculated at the basis of the EPO data.

Sources: German Statistical Office (population, university staff), establishment file of the German Social Insurance Statistics (establishments, employees), DFG Research Explorer (number of universities and research institutes).

**Table A2.** Proportion of the 'other actors' in the Dresden region.

Type of network	Part of the network	Number of actors classified as 'other'	Number of all actors	Share (%)
Total network (all layers)	Entire network	149	588	25.3
	Main component	38	405	9.4
R&D collaborations	Entire network	5	206	2.4
	Main component	4	171	2.3
Co-publications	Entire network	26	154	16.9
	Main component	26	154	16.9
Co-patents	Entire network	121	335	36.1
	Main component	12	158	7.6

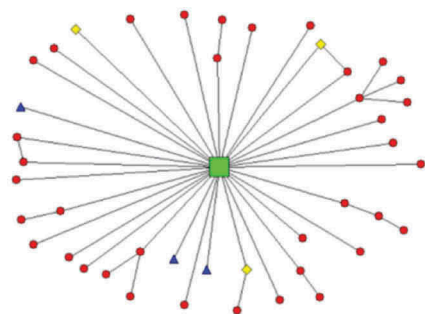
## Appendix B. Network graphs for all other case study regions



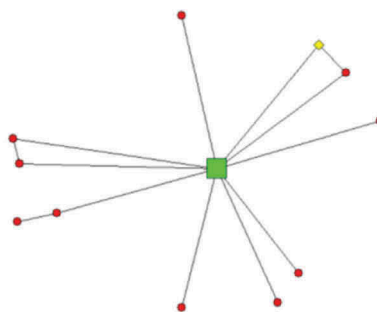
**Figure B1.** Networks of the Aachen region (largest component, period 2000–2010). (a) All layers. (b) Collaborative R&D projects. (c) Co-publications. (d) Co-patents.

*Source:* Authors' own illustration.

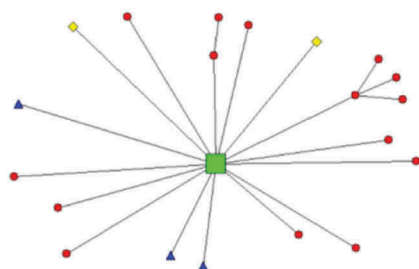
a) All layers



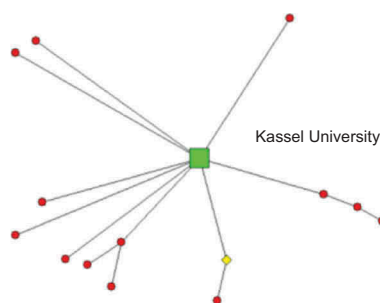
b) Collaborative R&D projects



c) Co-publications



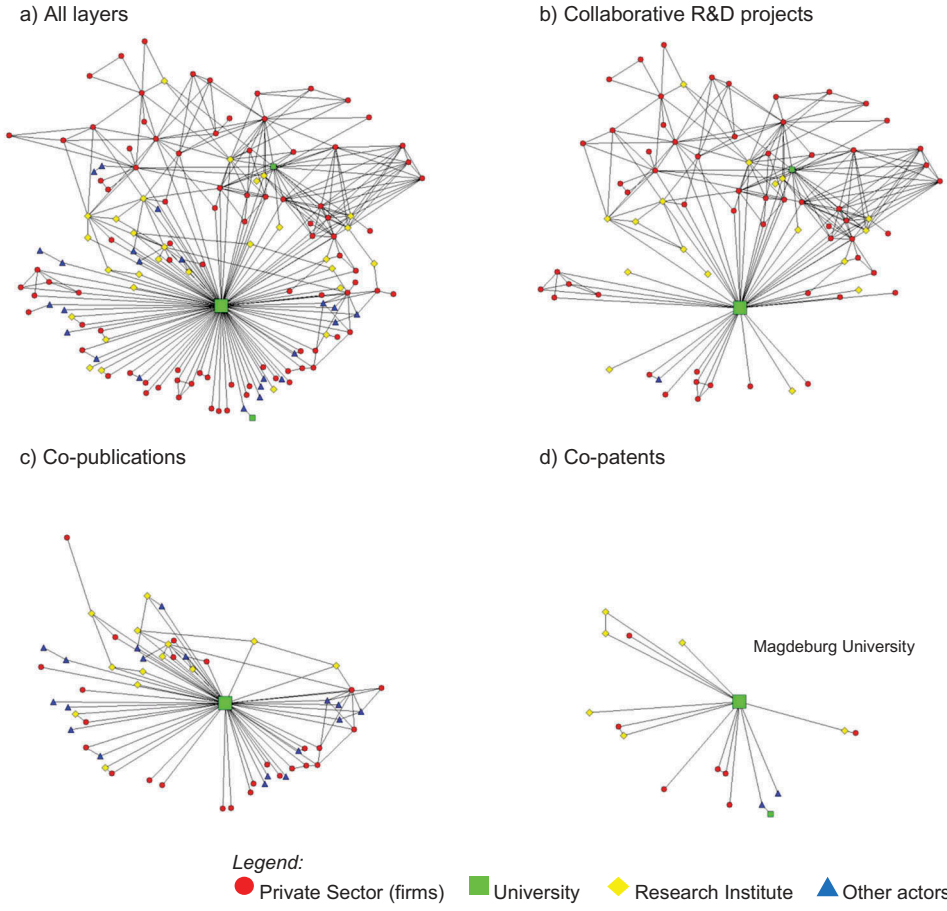
d) Co-patents



Legend:  
● Private Sector (firms)    ■ University    ◆ Research Institute    ▲ Other actors

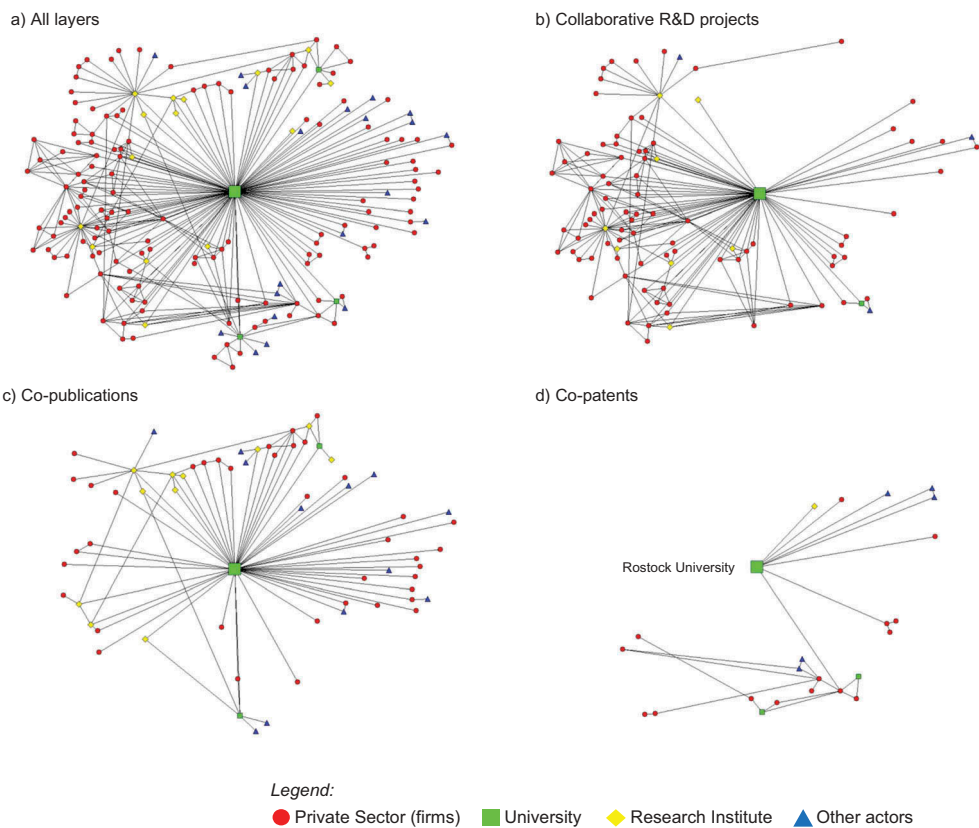
**Figure B2.** Networks of the Kassel region (largest component, period 2000–2010). (a) All layers. (b) Collaborative R&D projects. (c) Co-publications. (d) Co-patents.

Source: Authors' own illustration.



**Figure B3.** Networks of the Magdeburg region (largest component, period 2000–2010). (a) All layers. (b) Collaborative R&D projects. (c) Co-publications. (d) Co-patents.

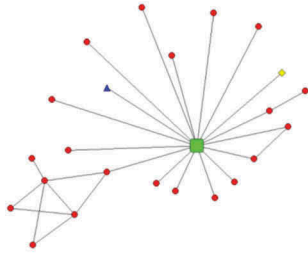
Source: Authors' own illustration.



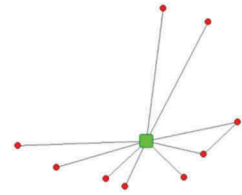
**Figure B4.** Networks of the Rostock region (largest component, period 2000–2010). (a) All layers. (b) Collaborative R&D projects. (c) Co-publications. (d) Co-patents.

Source: Authors' own illustration.

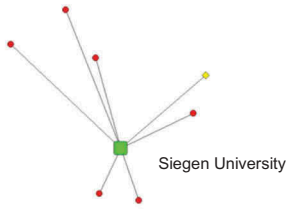
a) All layers



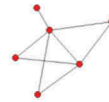
b) Collaborative R&D projects



c) Co-publications



d) Co-patents



Legend:

- Private Sector (firms)
- University
- ◆ Research Institute
- ▲ Other actors

**Figure B5.** Networks of the Siegen region (largest component, period 2000–2010). (a) All layers. (b) Collaborative R&D projects. (c) Co-publications. (d) Co-patents.

Source: Authors' own illustration.