IWH

Halle Institute for Economic Research

Member of the Leibniz Association

Discussion Papers

No. 1

January 2019

Identifying Cooperation for Innovation – A Comparison of Data Sources

Michael Fritsch, Matthias Piontek, Mirko Titze

Authors

Michael Fritsch
Friedrich Schiller University Jena and
Halle Institute for Economic Research (IWH) –
Member of the Leibniz Association
E-mail: m.fritsch@uni-jena.de

Matthias Piontek
Friedrich Schiller University Jena
E-mail: matthias.piontek@uni-jena.de

Mirko Titze
Corresponding author
Halle Institute for Economic Research (IWH) –
Member of the Leibniz Association, Centre for
Evidence-based Policy Consulting (IWH-CEP)
E-mail: mirko.titze@iwh-halle.de
Tel +49 345 7753 861

The responsibility for discussion papers lies solely with the individual authors. The views expressed herein do not necessarily represent those of IWH. The papers represent preliminary work and are circulated to encourage discussion with the authors. Citation of the discussion papers should account for their provisional character; a revised version may be available directly from the authors.

Comments and suggestions on the methods and results presented are welcome.

IWH Discussion Papers are indexed in RePEc-EconPapers and in ECONIS.

Editor

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

Address: Kleine Maerkerstrasse 8
D-06108 Halle (Saale), Germany
Postal Address: P.O. Box 11 03 61
D-06017 Halle (Saale), Germany

Tel +49 345 7753 60
Fax +49 345 7753 820

www.iwh-halle.de

ISSN 2194-2188

Identifying Cooperation for Innovation – A Comparison of Data Sources

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 subsidised 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. The key role of universities that connect many regional actors is identified in all three types of data.

Keywords: knowledge interactions, social network analysis, regional innovation systems, data sources

JEL classification: O30, R12, R30

1. Introduction[1]

There is strong indication that knowledge exchange and division of labor play an increasingly important role for innovaton processes (Wuchty, Jones and Uzzi 2007). 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.[2] 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 interactions. We describe the comprehensiveness and selectivity of these three types of data. We combine the three types of data 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.[3] This empirical

[1] 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. 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).

[2] For example, patent data disregard cooperation for inventions that are not patented (e.g. Arundel and Kabla 1998).

[3] 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 et al. (2015) combines three different datasets (granted projects supported within the EU framework

exercise covers the period 2000–2010 and is performed for six regions in Germany with varying levels of innovation activity and population density.

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. 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 respectively network development programs.

The remainder of this paper is organized as follows. In section 2 we discuss the related literature. Section 3 describes the data sources and introduces the case study regions. 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; Broekel, Fornahl and Morrison 2015), the empirical evidence on the actual extent of these differences is rare. Most empirical analyses of innovative interactions are based on co-patents (e.g. Graf and Henning 2009; Hoekman et al. 2009)[4], co-

programs, co-patents, and co-publications). Although, these datasets are merged at the regional level and not at the level of actors.

[4] Fischer and Griffith (2008) as well as Fischer et al. (2006) use patent citations to depict innovative interactions with the help of the patent indicator.

publications (e.g. Ponds et al. 2007; Hoekman et al. 2009; Hoekman et al. 2010), or data on publicly supported collaborative R&D projects[5].. Few studies such as Lata et al. (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[6] 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). However, because inventions are just ideas many of which are never commercially applied, these analyses can hardly say anything about later stages of the innovation process. Moreover, patenting requires a certain level of newness, so that 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 in these large firms are more likely to file patents than small and medium-sized

[5] E.g. Maggioni, Nosvelli and Uberti (2007), Broekel, Fornahl and Morrison (2015), Scherngell and Barber (2011), Scherngell and Lata (2013), Barber and Scherngell (2013).

[6] Regional level means that the region is the unit of observation. Knowledge transfer is then modelled as knowledge transfer between regions.

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 underreported as the public institutions have only comparatively weak incentives to patent.

Based on data of subsidized collaboration projects in Germany Broekel and Graf (2010) distinguish between joint projects in basic and in applied research. In projects basic research they find smaller and more centralized networks which results in a relatively pronounced concentration of interactions between 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 et al. (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.

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 be 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 Data sources and matching procedures

Our analysis builds on Titze et al. (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, which has a crucial role in the management 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 the 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, the public grants are addressed to institutions (universities, external research institutes, firms, etc.) but not individuals. Consequently, this database does not include the names of the people involved in a project. Three key variables from the Subsidies Catalogue are relevant to our investigation: primary keys for sub-projects and for the entire collaboration roject[7], the name and location of the executing organizations,[8,9] and the

[7] In case of several subprojects these subprojects summarized to one main collaboration project.

[8] 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.

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. We account for those projects that involve at least two collaboration partners.

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 information. We consider patent applications with at least two applicants[10]. Compared to the OECD RegPat data this source has several severe advantages. First, since the patent identification number do 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.[11] 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.

---

[9] 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.

[10] 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.

[11] 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.

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 et al. 2000; Mansfield et al. 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.

Using bibliometric data presents certain well-known and -discussed difficulties in the literature (e.g. Abramo et al. 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 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 precleaning 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 b 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.

Table 1: Features of the raw data in the sources applied

	Co-patents	Co-publications	Collaborative R&D projects
	Form of interaction		
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
	Information at the level of individuals		
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	Postal codes	No	No
	Information about the organization/affiliation		
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	Postal codes	Sometimes, but different spellings possible (no unique identifier)	Administrative regional codes (*Amtlicher Gemeindeschluessel*)
Industry	No	No	2-digit NACE codes (universities and research institutes at a more disaggregated level)

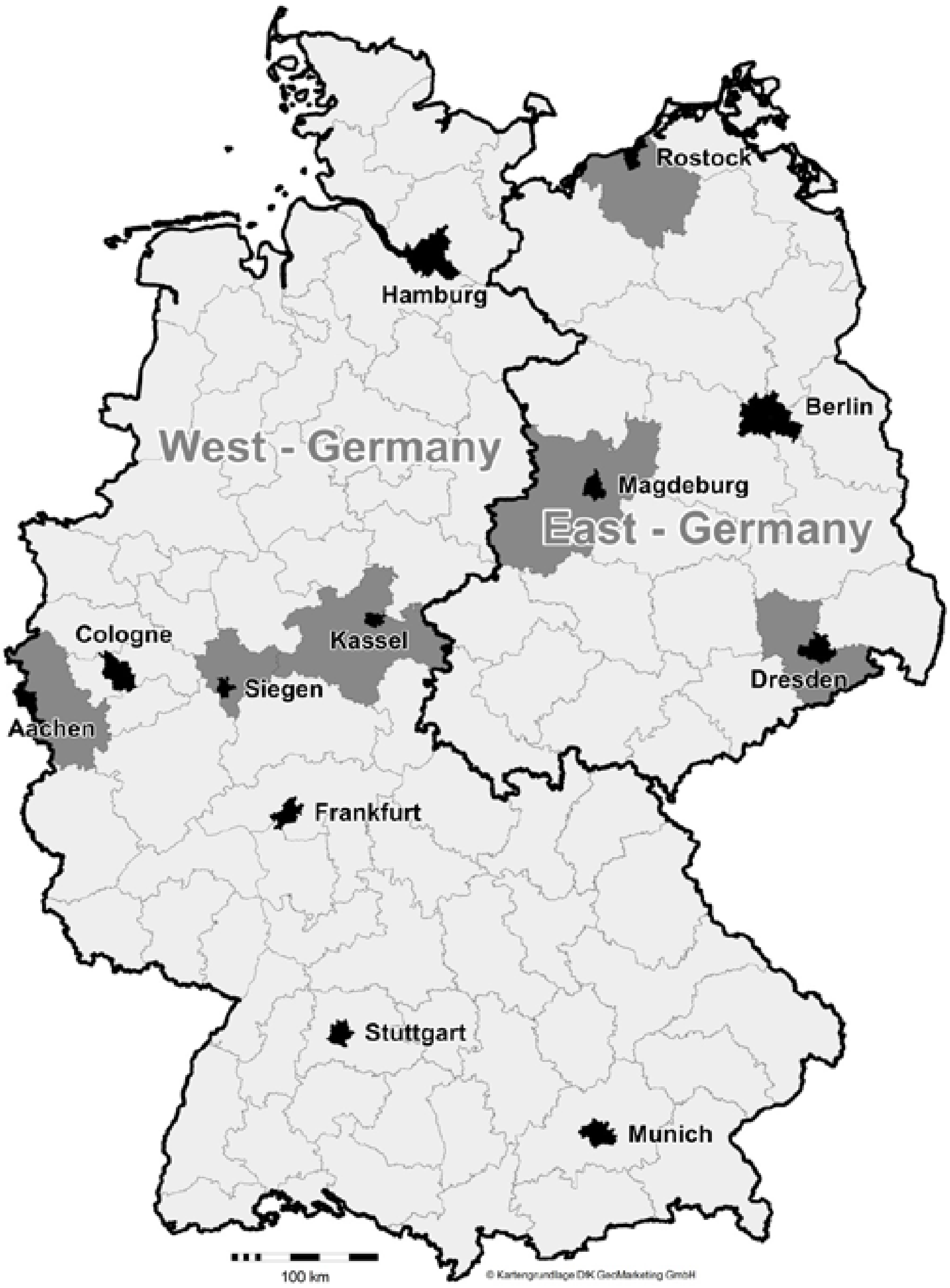
Source: Own illustration.

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. The patent data and data on publicly funded R&D collaboration also provide the opportunity to include inter-regional relationships.

3.2 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 the district as their workplace.



Source: Own illustration.

Figure 1: Case study regions in Germany

Our sample of regions comprises three types of settlement structures: agglomerations, moderately congested regions, and rural areas.[12] Large innovation centers such as Munich or Stuttgart are not included. All the case study regions host at least one university. Figure 1

[12] 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).

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 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.

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. 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; 839 (43.3%) actors were identified either in the Amadeus firm database or the Research Explorer database.

Table 2: Actors in overlapping channels of knowledge transfer[a]

Channels of knowledge transfer (pooled 2000-2010)	Types of actors: All actors	Firms	Universities	Research institutes	Other[b]
	Number of actors by type				
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)
	Sum shares[c] (in %)				
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.

The lion's share of the 1,940 unique actors recorded in our dataset (Table 2) are involved in one type of innovative link, either co-patenting (50.5%)[13], 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 2 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

[13] 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.

source. In the co-publications data and the information on publicly funded collaborative R&D activities, the shares are 29.5% and 46.1%, 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 2), 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.

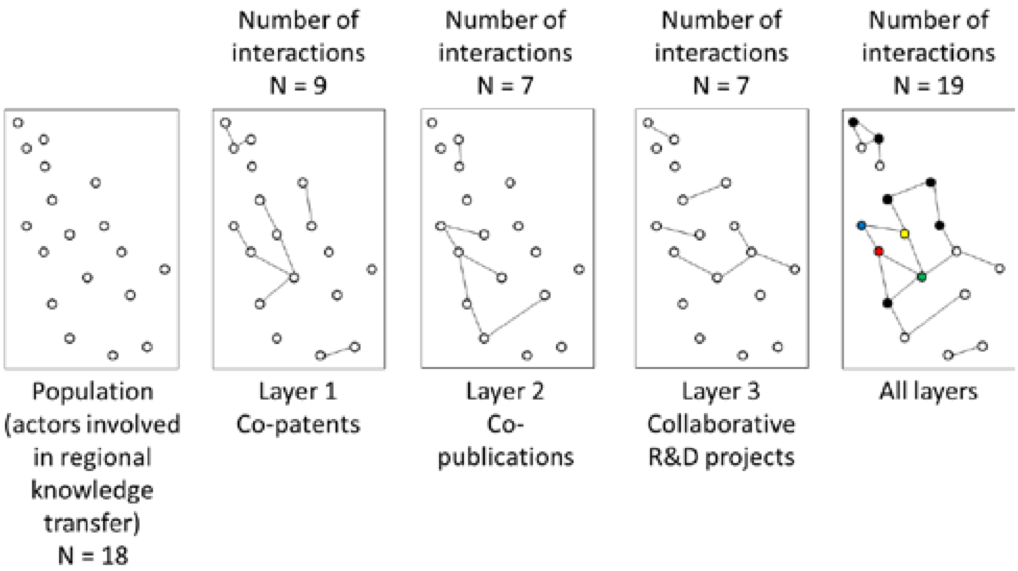
4.2 Overlapping knowledge transfer channels

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 3 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 3). If we distinguish these interactions by type, we find 15,542 (56.7%) interactions between two institutions simultaneously appear as co-patents, co-publications, and publicly funded collaborative R&D projects; 3,729 (13.6%) relationships are mere co-publications; 3,233 (11.8%) represent co-publications and joint R&D projects; and 2,875 (10.5%) are pure co-patents.

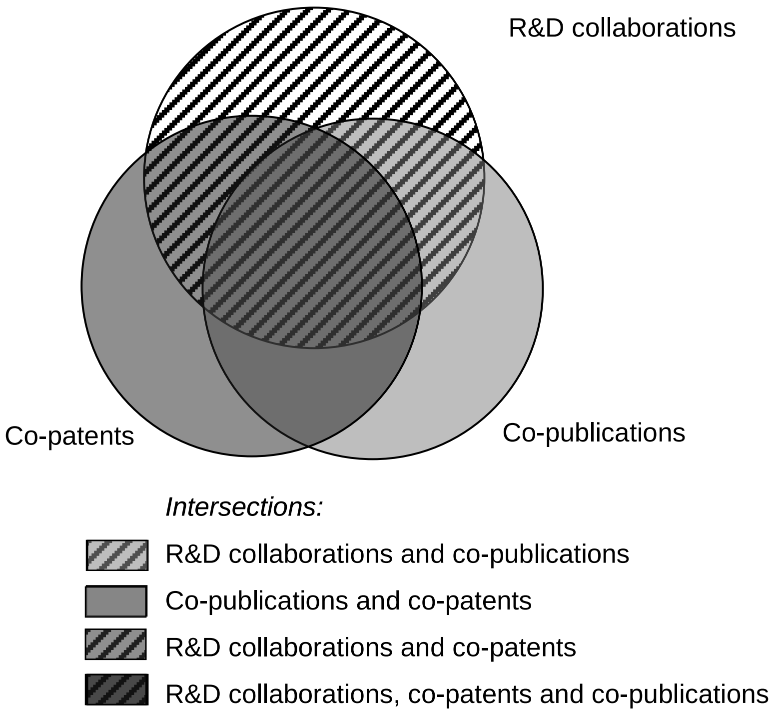
The bottom of Table 3 presents the total share of regional links captured by each type of interaction. The data reveals that more than two-thirds (69.4%) of all recorded regional relationships can be identified in the patent statistics, whereas 30.6% of the relationships are not recorded in the patent data. The shares of the other two data sources are higher. Publicly funded R&D collaborations provide information representing approximately 74.8% of the total links, and co-publications cover 83.1%. Hence, all three data sources disregard considerable shares of all identified links.

These results change dramatically if we distinguish among the types of actors. Private sector firms tend not to use simultaneous channels of knowledge transfer; instead, they are involved in only one form, namely, co-publications (24.5%), co-patents (21.9%), or collaborative R&D projects (20.3%). An exception is the combination of collaborative R&D projects and co-publications. Co-patents display only 43.9% of all intraregional links of firms. 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 approximately half of the intraregional innovative relationships of private sector firms (Table 3).



Source: Own illustration.

Figure 2: Combinations of the forms of cooperation



Source: Own illustration.

Figure 3: Combinations of different forms of knowledge transfer

Table 3: Overlapping channels of knowledge transfer[a]

Channels of knowledge transfer (pooled 2000-2010)	Types of actors				
	All actors	Firms	Universities	Research institutes	Other[b]
	Number of interactions by type				
Co-patents only	2,875 (10.5)	1,312 (21.9)	2 (0.0)	252 (2.4)	1,309 (72.4)
Co-publications only	3,729 (13.6)	1,465 (24.5)	5 (0.1)	1,900 (18.3)	359 (19.9)
Collaborative R&D only	1,442 (5.3)	1,214 (20.3)	52 (0.6)	168 (1.6)	8 (0.4)
Co-publications and co-patents	304 (1.1)	192 (3.2)	0 (0.0)	79 (0.8)	33 (1.8)
Collaborative R&D and co-patents	309 (1.1)	262 (4.4)	14 (0.2)	33 (0.3)	0 (0.0)
Collaborative R&D and co-publications	3,233 (11.8)	681 (11.4)	16 (0.2)	2,438 (23.5)	98 (5.4)
All layers	15,542 (56.7)	860 (14.4)	9,157 (99.0)	5,525 (53,2)	0 (0.0)
Total	27,434 (100.0)	5,986 (100.0)	9,246 (100.0)	10,395 (100.0)	1,807 (100.0)
	Sum of shares[c] (in %)				
Co-patents	69.4	43.9	99.2	56.7	74.3
Co-publications	83.1	53.4	99.3	95.6	27.1
R&D collaborations	74.8	50.4	99.9	78.5	5.9

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 which share of regional knowledge transfer is captured by co-patents, co-publications, and (granted) R&D collaboration projects. The sum of the overlapping transfer channels the shares is more than 100%.

Source: Own calculations.

These findings are completely different from the results obtained for universities. About 99% of universities' knowledge links occur in all three layers, indicating that universities are reliably represented in each of the three databases. The figures for the extra-university public research institutes are in between those for firms and universities. Table 3 presents another remarkable finding, namely, the importance of public research for regional knowledge transfer: 71.6% of all the identified interactions include either universities or extra-university public research institutes. This clearly highlights the central role of these actors in regional knowledge transfer.

4.3 Network descriptives

The previous two sections have demonstrated that the numbers and shares of regional interactions differ considerably across the three types of data. Table 4 depicts main network descriptives for the seven 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

3c and B1c to B5c 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 4: Network descriptives for the data sources and regions

	Average	Aachen	Dresden	Kassel	Magde-burg	Rostock	Siegen
	All layers (baseline scenario)						
Number of nodes[b]	323.3 (100)	581 (100)	588 (100)	145 (100)	278 (100)	245 (100)	103 (100)
Share of actors in the largest component [a,c] (%)	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
	Collaborative R&D projects						
Number of nodes[b]	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[d]	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 [d]	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
	Co-publications						
Number of nodes[b]	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 [d]	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[d]	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
	Co-patents						
Number of nodes[b]	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 [d]	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 [d]	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 base line scenario).

Source: Own calculations.

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.

4.4 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 largest main component.[14]

Figures 3a–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.[15] The most central actors in the entire multilayer network (Figure 3a) 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

[14] See the network graphs for all case study regions in Appendix B.

[15] 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.

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 is a crucial broker role for the network. 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 3a). 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 3a): this share is 83% of the network based on collaborative R&D projects (Figure 3b), 100% of the co-publication network (Figure 3c), and 47.2% of the patent network (Figure 3d). 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 Tables4). Only twelve of these "other" actors (7.6%) are connected to the main component of the patent network in Dresden (Figure 3d). 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.[16]

[16] The reason for this is that patenting of all Fraunhofer institutes is centrally managed at the headquarters of its society in Munich.

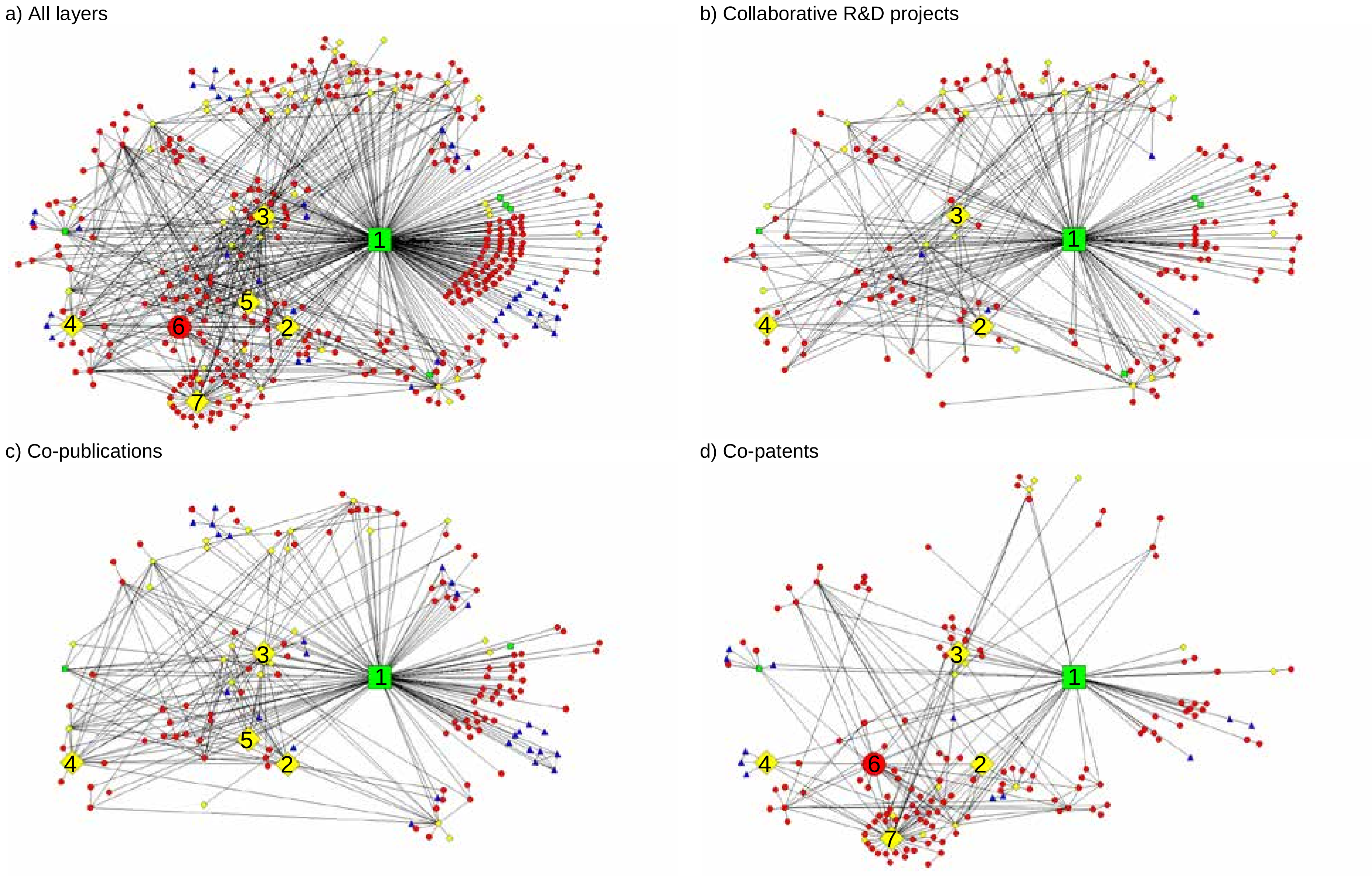


Figure 3: Networks of the Dresden region (largest component, period 2000–2010)

Source: Authors' own illustration.

Legend:

Private Sector (firms) University Research Institute Other actors

Actors with central positions:

1: TU Dresden 2: Leibniz Institute for Solid State and Materials Research 3: Helmholtz-Zentrum Dresden-Rossendorf
4: Leibniz Institute for Polymer Research 5: Max Planck Institute for Chemical Physics of Solids 6: Infineon 7: Fraunhofer Society

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 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 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 et al. 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 secondary 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.[17]

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 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.

[17] For a detailed characterization of the three types of knowledge bases see Asheim et al. (2007).

References

Abramo, G.; D'Angelo, C.A.; Caprasecca, A. (2009): Allocative efficiency in public research funding: can bibliometrics help? *Research Policy*, 38(1), 206–215.

Arundel, A.; Kabla, I. (1998): What percentage of innovations are patented? Empirical estimates for European firms. *Research Policy*, 27, 127–141.

Asheim, B.; Coenen, L.; Moodysson, J.; Vang, J. (2007): Constructing knowledge-based regional advantage: implications for regional innovation policy. *International Journal of Entrepreneurship and Innovation Management*, *7*(2-5), 140-155.

Barber, M. J.; Scherngell, T. (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.

Bathelt, H.; Malmberg, A.; Maskell, P. (2004): Cluster and knowledge: local buzz, global pipelines and the process of knowledge creation. *Progress in Human Geography*, 28, 31–56.

BBSR (2015): http://www.bbsr.bund.de/BBSR/DE/Raumbeobachtung/Raumabgrenzungen/Regionstypen/regionstypen.html?nn=443048. Accessed 6 June 2017.

Blind, K.; Edler, J.; Frietsch, R.; Schmoch, U. (2006): Motives to patent: Empirical evidence from Germany. *Research Policy*, *35*(5), 655-672.

Brandes, U. (2001) Drawing on Physical Analogies. In: Kaufmann, M.; Wagner, D. (eds): *Drawing Graphs: Methods and Models*. Berlin: Springer.

Broekel, T.; Graf, H. (2010): Structural properties of cooperation networks in Germany: From basic to applied research. *Jena Economic Research Papers* 2010 – 078.

Broekel, T.; Graf. H. (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.

Broekel, T.; Fornahl, D.; Morrison, A. (2015): Another cluster premium: Innovation subsidies and R&D collaboration networks. *Research Policy,* 44, 1431–1444.

Cohen, W. M.; Nelson, R. R.; Walsh, J. P. (2000): Protecting their intellectual assets: Appropriability conditions and why US manufacturing firms patent (or not). National Bureau of Economic Research (No. w7552).

Cowan, R.; Jonard, N. (2004): Network structure and the diffusion of knowledge, *Journal of Economic Dynamics and Control*, 28(8), 1557-1575.

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.; Griffith, D. A. (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.

Fischer, M. M.; Scherngell, T.; Jansenberger, E. (2006). The geography of knowledge spillovers between high-technology firms in Europe: Evidence from a spatial interaction modelling perspective. *Geographical Analysis*, 38, 288–309.

Fleming, L.; King, C.; Juda, A. I. (2007): Small Worlds and Regional Innovation. *Organization Science*, 18(6), 938-954.

Fritsch, M.; Aamoucke, R. (2017): Fields of Knowledge in Higher Education Institutions, and Innovative Start-Ups—An Empirical Investigation. *Papers in Regional Science*, 96, 1-27.

Fritsch, M.; Graf, H. (2011): How sub-national conditions affect regional innovation systems: The case of the two Germanys. *Papers in Regional Science*, 90(2), 331-353.

Fritsch, M.; Slavtchev, V. (2011): Determinants of the efficiency of regional innovation systems. *Regional Studies*, 45, 905–918.

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.

Graf, H.; Henning, T. (2009): Public Research in Regional Networks of Innovators: A Comparative Study of Four East German Regions. *Regional Studies*, 43, 1349-1368.

Griliches, Z. (1990): Patent statistics as economic indicators: a survey. *National Bureau of Economic Research* (No. w3301).

Hanneman, R. A.; Riddle, M. (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

Hoekman, J.; Frenken, K.; Tijssen, R. J. (2010): Research collaboration at a distance: Changing spatial patterns of scientific collaboration within Europe. *Research Policy*, 39(5), 662-673.

Hoekman, J.; Frenken, K.; Van Oort, F. (2009): The geography of collaborative knowledge production in Europe. *The Annals of Regional Science*, 43(3), 721-738.

Kauffeld-Monz, M.; Fritsch, M. (2013): Who Are the Knowledge Brokers in Regional Systems of Innovation? A Multi-Actor Network Analysis. *Regional Studies*, 47, 669–685.

Lata, R.; Scherngell, T.; Brenner, T. (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.

Mansfield, E.; Schwartz, M.; Wagner, S. (1981): Imitation costs and patents: an empirical study. *The Economic Journal*, 91(364), 907-918.

Maggioni, M.A.; Nosvelli, M.; Uberti, T.E. (2007): Space versus networks in the geography of innovation: a European analysis. *Papers in Regional Science*, 86 (3), 471–493.

Meyer, M. (2002): Tracing knowledge flows in innovation systems. *Scientometrics*, 54(2), 193–212.

Mowery, D.C.;Sampat, B. N. (2006): Universities in national innovation systems. In: Jan Fagerberg, David C. Mowery and Richard R. Nelson (eds): *The Oxford Handbook of Innovation*. Oxford: Oxford University Press, 209-239.

Ponds, R.; Van Oort, F.; Frenken, K. (2007): The geographical and institutional proximity of research collaboration. *Papers in regional science*, 86(3), 423-443.

Research Explorer (2017): http://www.research-explorer.de/research_explorer.de.html. Accessed 6 June 2017.

Scherngell, T.; Barber, M. J. (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.

Scherngell, T.; Lata, R. (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.

Schmoch, U. (1999): Interaction of universities and industrial enterprises in Germany and the United States. A comparison. *Industry and Innovation*, 6(1), 51–68.

Storper M.; Veneables, A.J. (2004): Buzz: face-to-face contact and the urban economy. *Journal of Economic Geography*, 4, 351-370.

Ter Wal, A. ; Boschma, R. (2009): Applying social network analysis in economic geography: framing some key analytic issues. *Annals of Regional Science*, 43(3), 739-756.

Titze, M.;Schwartz, M.;Brachert, M. (2012): A systemic view on knowledge-based development metrics. *International Journal of Knowledge-Based Development*, 3(1), 35-57.

Wuchty, S; Jones B.F.; Uzzi, B. (2007): The Increasing Dominance of Teams in Production of Knowledge. *Science*, 316, 1036-1039.

Youtie, J.; Shapira, P. (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.

Appendix A

Table A1: Case study regions at a glance

Planning Region	*Aachen*	*Dresden*	*Siegen*	*Rostock*	*Kassel*	*Magdeburg*
Macro-region in Germany	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
Private sector 2000						
Number of establishments[b]	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 (%)[c]	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
Research sector (2000)						
Number of research institutes[a]	21	38	0	14	7	19
Number of universities[ad]	3	10	1	2	3	4
Total research teaching staff[d]	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 (%)[de]	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[d]	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 (%)[d]	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

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.

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	Entire network	149	588	25.3
(all layers)	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

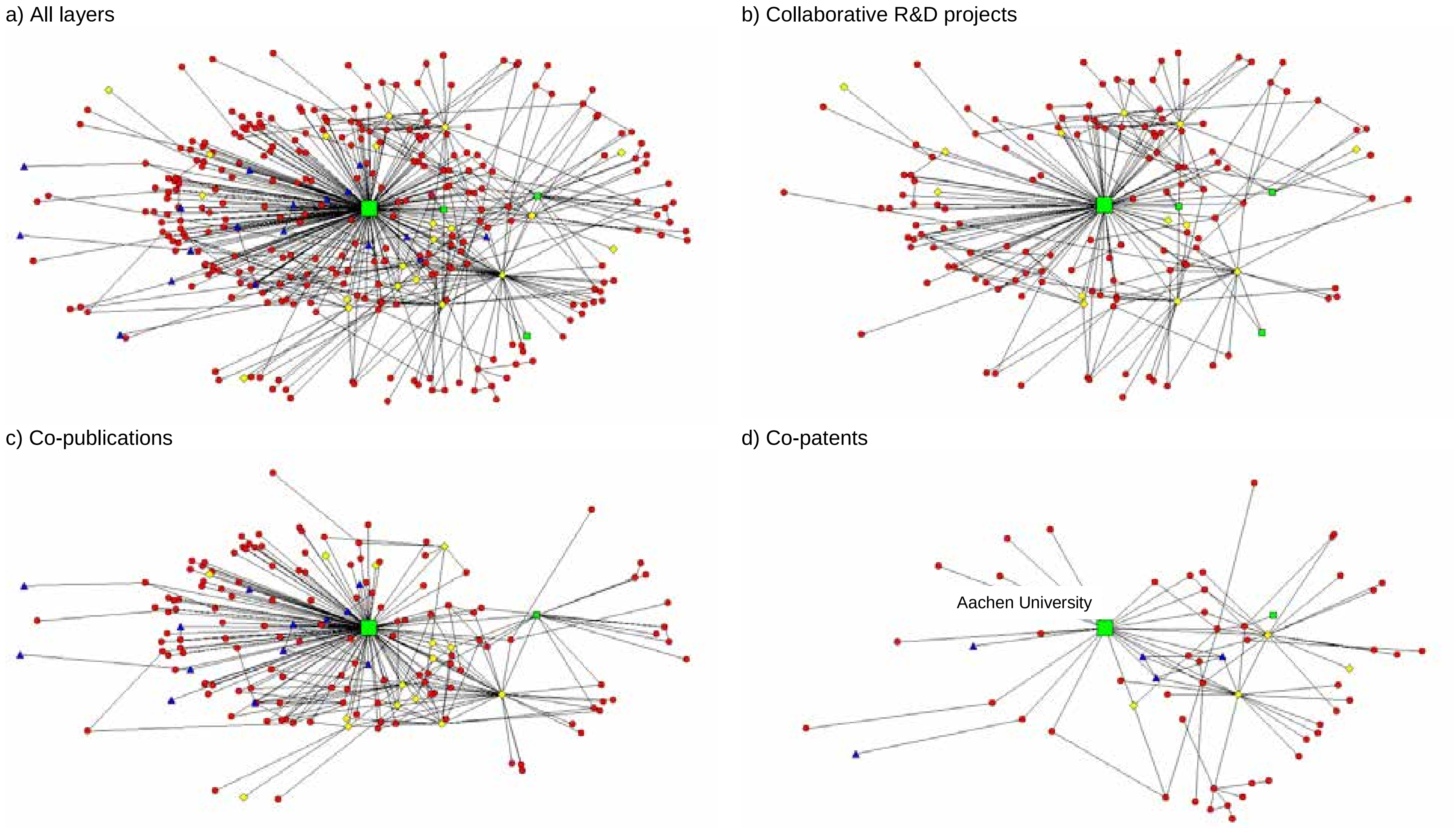


Figure B1: Networks of the Aachen region (largest component, period 2000–2010)

Source: Authors’ own illustration.

Legend:
Private Sector (firms) University Research Institute Other actors

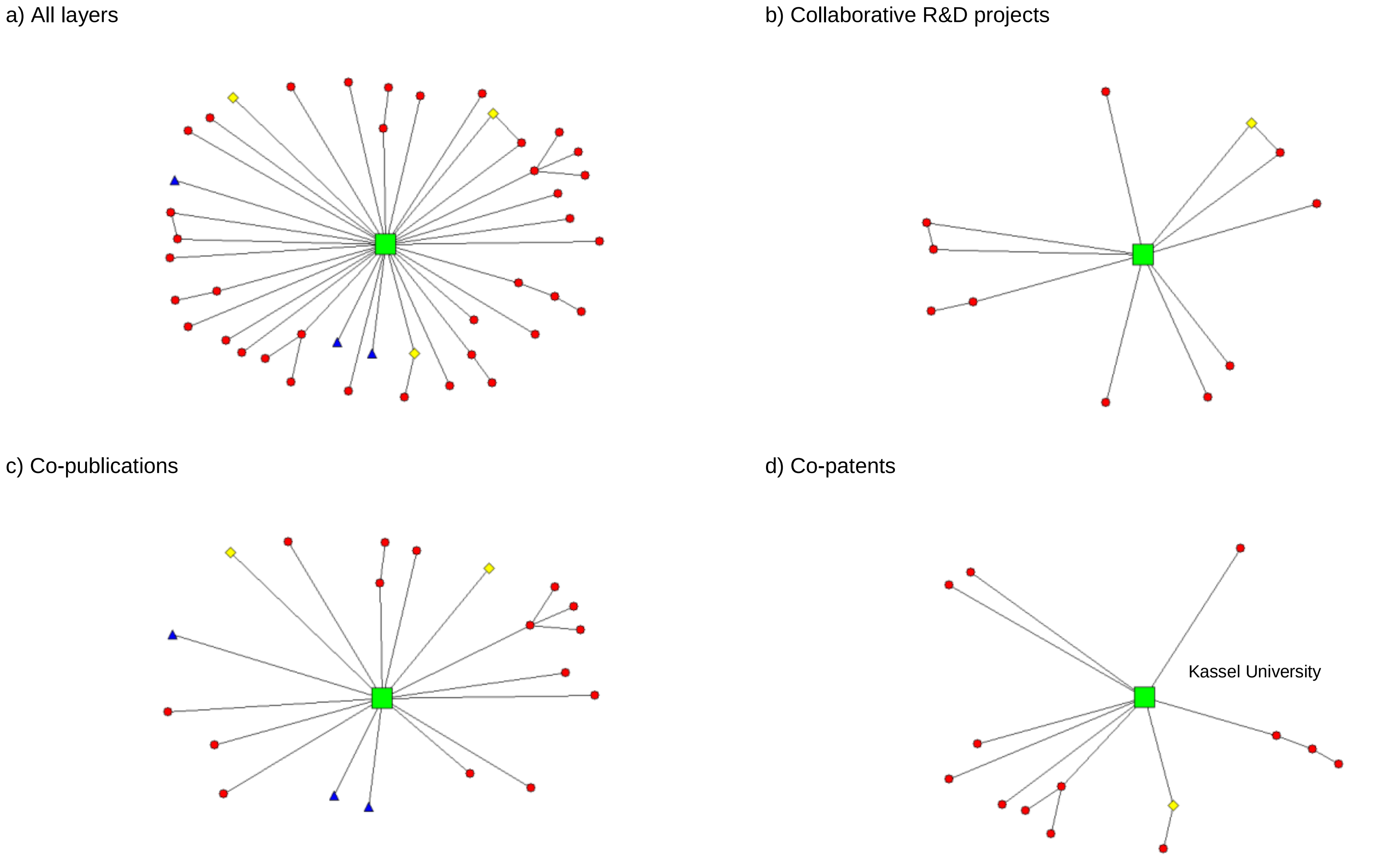
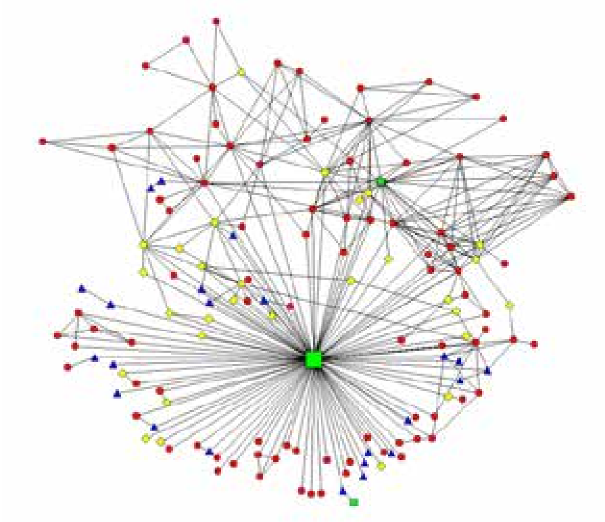


Figure B2: Networks of the Kassel region (largest component, period 2000–2010)

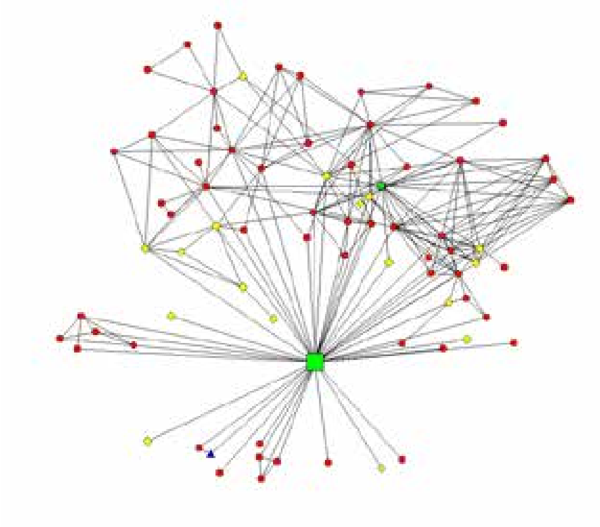
Source: Authors' own illustration.

Legend:
Private Sector (firms) University Research Institute Other actors

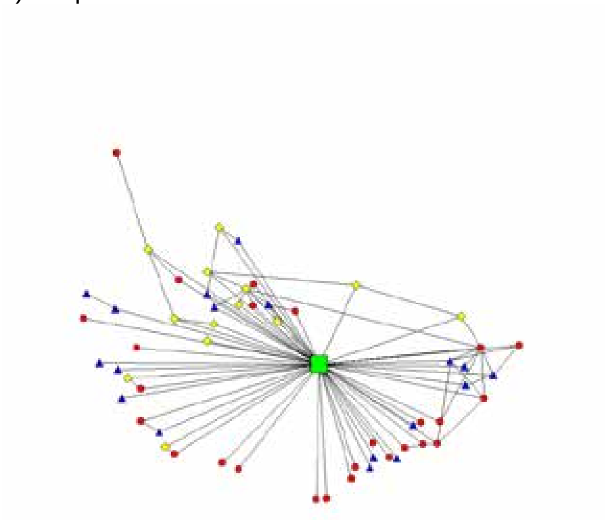
a) All layers



b) Collaborative R&D projects



c) Co-publications



d) Co-patents

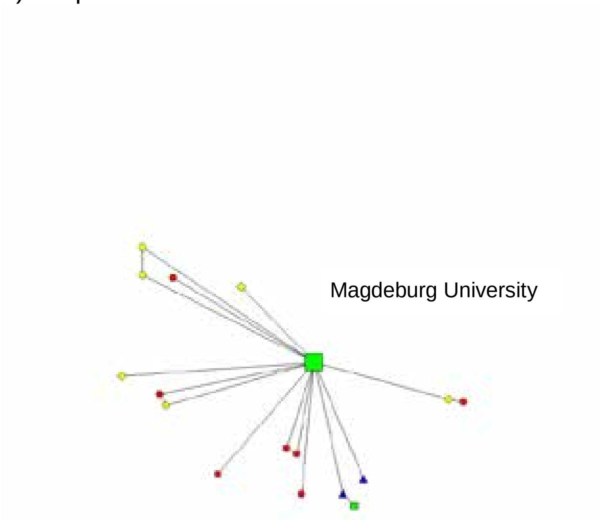


Figure B3: Networks of the Magdeburg region (largest component, period 2000–2010)

Source: Authors' own illustration.

Legend:
Private Sector (firms) University Research Institute Other actors

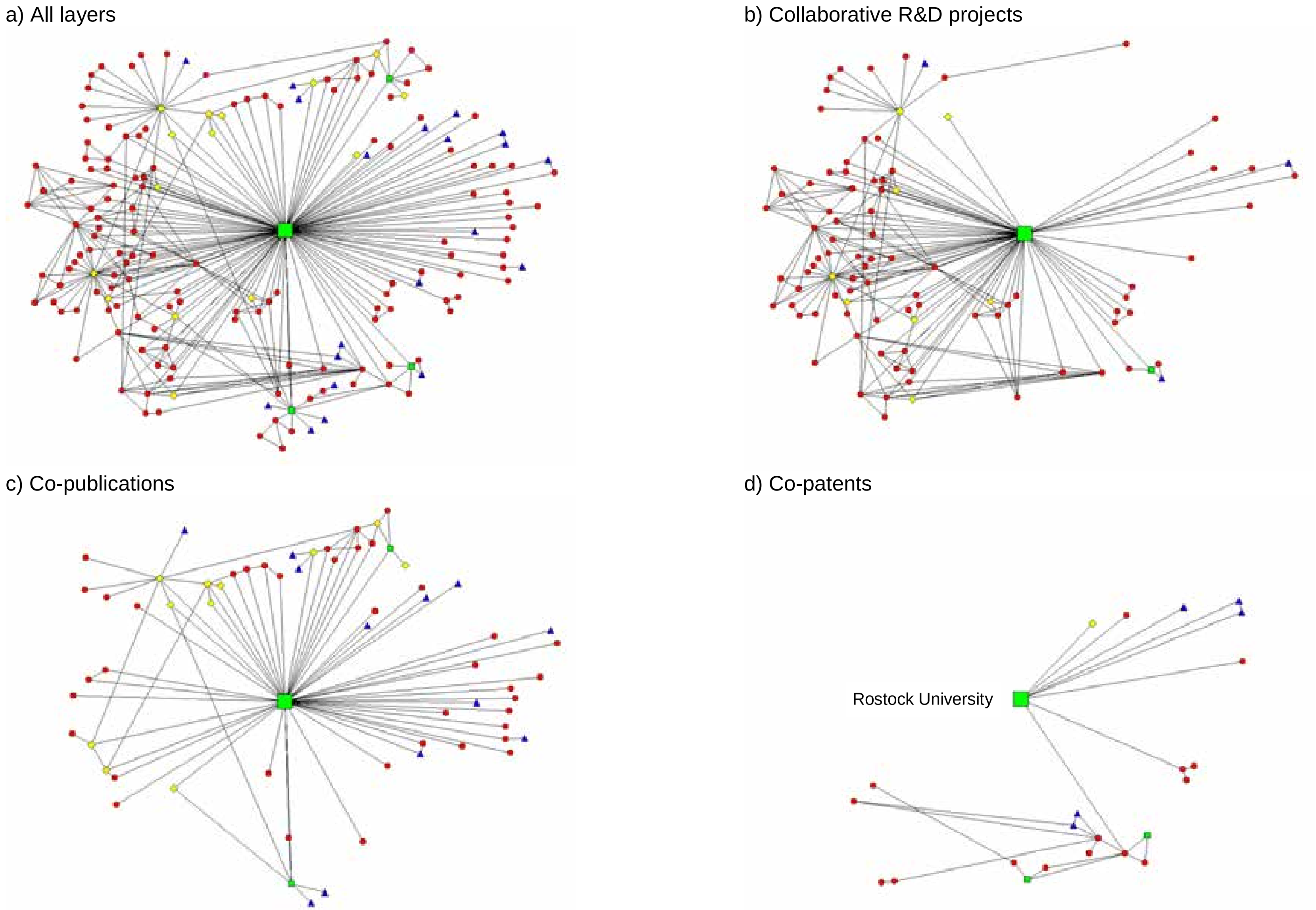
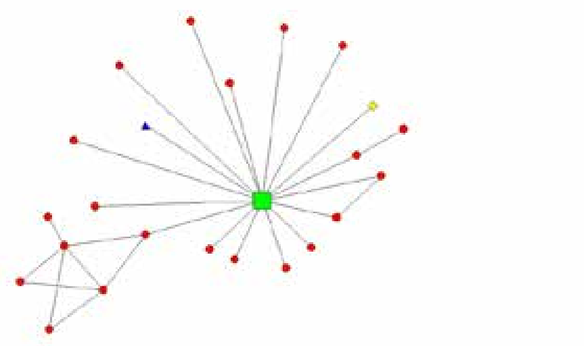


Figure B4: Networks of the Rostock region (largest component, period 2000–2010)

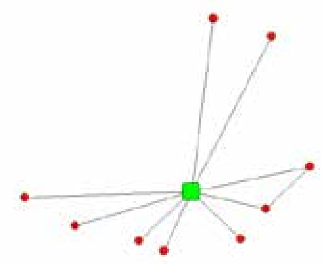
Source: Authors' own illustration.

Legend:
Private Sector (firms) University Research Institute Other actors

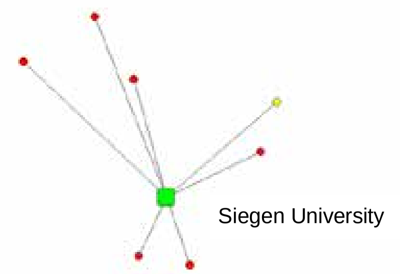
a) All layers



b) Collaborative R&D projects



c) Co-publications



d) Co-patents

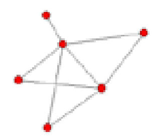


Figure B5: Networks of the Siegen region (largest component, period 2000-2010)

Source: Authors' own illustration.

Legend:
Private Sector (firms) University Research Institute Other actors

Halle Institute for Economic Research –
Member of the Leibniz Association

Kleine Maerkerstrasse 8
D-06108 Halle (Saale), Germany

Postal Adress: P.O. Box 11 03 61
D-06017 Halle (Saale), Germany

Tel +49 345 7753 60
Fax +49 345 7753 820

www.iwh-halle.de

ISSN 2194-2188