JENA ECONOMIC RESEARCH PAPERS

2010 – 078

Structural properties of cooperation networks in Germany: From basic to applied research

by

Tom Broekel
Holger Graf

www.jenecon.de

ISSN 1864-7057

The JENA ECONOMIC RESEARCH PAPERS is a joint publication of the Friedrich Schiller University and the Max Planck Institute of Economics, Jena, Germany. For editorial correspondence please contact markus.pasche@uni-jena.de.

Impressum:

Friedrich Schiller University Jena
Carl-Zeiss-Str. 3
D-07743 Jena
www.uni-jena.de

Max Planck Institute of Economics
Kahlaische Str. 10
D-07745 Jena
www.econ.mpg.de

Structural properties of cooperation networks in Germany: From basic to applied research

Tom Broekel* Holger Graf†

12th November 2010

Abstract

Economists pay more and more attention to knowledge networks and drivers of their development. Consequently, a rich literature emerged analyzing factors explaining the emergence of intra-organizational links. Despite substantial work focusing on the dyad level, only little is known about how and why (global) network structures differ between technologies or industries. The study is based on a new data source on subsidized R&D cooperation in Germany, which is presented in detail and discussed with respect to other types of relational data. A comparison of networks within ten technologies allows us to identify systematic differences between basic and applied research networks.

JEL codes: L14, I28, O38
Keywords: R&D subsidies, network, cooperation, Foerderkatalog, Germany

1 Introduction

Driven by the need to access external knowledge firms are relying more and more on cooperation with other organizations. This has been recognized by researchers, who analyze to an increasing extent inter-organizational knowledge networks. Prominently, Powell et al. (1996) show that firms' embeddedness within knowledge networks significantly influences their economic and innovative success. However, the relation between embeddedness and economic performance does not always yield positive effects as highlighted by Uzzi (1996). In addition to studying the effects of knowledge networks, this field of study is also concerned with factors driving network development. For instance, by focusing on the dyadic (micro) level ter Wal (2009) and Balland (2010) show empirically that linkages between organizations are more likely if they are proximate in a geographical, social, or cognitive dimension.

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

† Friedrich Schiller University Jena, Department of Economics and Business Administration, Carl-Zeiss-Str. 3, 07743 Jena, Germany, *E-mail*: holger.graf@uni-jena.de.

Despite the interesting results obtained from these dyad-oriented studies, they are likely to miss factors that shape the global structures of knowledge networks. While many authors have shown that the heterogeneity of industries can explain differences in the innovation process or market structure (see, e.g., Pavitt, 1984; Malerba and Orsenigo, 1996), it is still unknown if, and especially how, these differences translate into the structure of particular knowledge networks. First steps in that direction have been taken by Broekel and Boschma (2010), who argue that differences in the underlying knowledge bases of industries influence the density of their knowledge networks. Similarly, Graf (2009) puts forward that characteristics of the search process determine the interaction structure and therefore the evolution of innovation networks.

With the present study we aim to shed some light on these issues by comparing structural properties of cooperation networks in ten different technologies. While these technologies differ in multiple manners, our primary focus is on the role of (public) science organizations within the respective networks and how their embeddedness relates to differences in overall network properties. For that, we a develop a taxonomy that gradually categorizes technologies on a scale ranging from basic to applied research. Subsequently, global network statistics as well as characteristics of the most central actors are related to this taxonomy.

The second objective of this paper is to present a detailed description of a new database on subsidized (and mostly cooperative) R&D projects in Germany. While this database has already been used for analyses within specific fields, (e.g., Fornahl et al., 2010) a study comparing different technological areas has still been missing. The paper therefore explicitly discusses issues concerning the handling and usage of this data base for network studies and relates it to other, commonly used types of relational data, especially patents.

Our findings highlight that differences exist between basic and applied research, which relate global network properties as well as the characteristics of actors holding the most central positions. Concerning the first, knowledge networks in basic research are smaller and consequently more dense, however, they involve more isolates. They are also organized in a more centralized manner, i.e., the bulk of linkages is concentrated on a few actors. The most central actors in knowledge networks of applied research are large firms, while universities are of less relevance compared to public research institutes.

Our study is organized as follows. In section 2, we provide a short review of the knowledge network literature and discuss how their structures might depend on characteristics of the underlying technology. The employed data is described in section 3 and the resulting network structures are analyzed in section 4 with the goal of providing a typology of research areas in Germany. The determinants of actor centrality in these networks are identified in section 5. Section 6 summarizes our main findings and points towards future research.

2 Innovation and social networks

The focus of economists on cooperative R&D and the importance of external knowledge sources somewhat shifted during the last decades from a bilateral perspective as in D'Aspremont and Jacquemin (1988) or in Hagedoorn (2002) to a multilateral or network view (e.g., Powell et al., 1996). Researchers approach a steadily growing variety of issues related to firms' embeddedness into knowledge networks. For example, the development of scientific communities is explored by Barabasi et al. (2002) and Moody (2004). Channels of technology transfer in general and university-industry relations in particular are investigated by Balconi et al. (2004). Fleming et al. (2006) and Fleming and Frenken (2007) use patent data to model the evolution of local clusters, while the structures and characteristics of regional networks are explored in a number of recent studies (ter Wal, 2008; Graf and Henning, 2009; Graf, 2010). Still, the majority of studies on different industries, technologies, nations, and network types focus on the individual connection (micro level) and little is known about the structural differences at the level of the complete network (macro level).

It is well known that industries differ substantially in their underlying technologies, with an influence on entrepreneurship opportunities, cumulativeness, and appropriability of knowledge (Pavitt, 1984; Malerba and Orsenigo, 1996). We argue that such heterogeneities are likely to translate into varying structural characteristics of knowledge networks. For instance, it seems straightforward to assume (but is unproven) that in a "supplier dominated industry" of Pavitt (1984) large suppliers hold central network positions. However, do universities hold more central positions in science-based industries' knowledge networks (e.g., chemicals) although the main application developers are huge multinational corporations (e.g., Dow Chemicals, BASF, Bayer)?

For the present study we focus on just one dimension, namely the degree to which the (public) science sector is involved in innovation processes. More precisely, we investigate if (and how) the location of ten technologies on a continuum from basic to applied research relates to structural characteristics of their respective knowledge networks.

There are a number of reasons why the degree to which a technology is application oriented (i.e., its position within a continuum ranging from basic to applied research) can influence the structures of the according knowledge networks. For example, fundamental differences in the way new knowledge is created exist between technologies (Asheim and Coenen, 2005). The mode of innovation generation in basic research implies trial settings, formal models, and high degrees of codification, which is similar to the 'scientific' way of knowledge creation. This eases the exchange, cooperation, and cross-fertilization between the industrial and the scientific sector making universities more likely to hold central network positions. In addition, in science based technologies new knowledge is more frequently published in public journals and magazines. This makes it easier accessible, which can reduce obstacles for collaboration (see, e.g., Kesteloot and Veugelers, 1995) eventually leading to more dense networks (Broekel and Boschma, 2010).

Similarly Malerba et al. (2000) differentiate between *generic* and *specific* knowledge, which may be of varying importance in some technologies. While the first "refers to knowledge of a very broad nature" the second describes "knowledge specialized and targeted to specific applications" (Malerba et al., 2000, p. 392). Depending on the structure of the public supply of generic or specific (scientific) knowledge, some technologies may rely to higher degrees on public research organizations.

Technologies also differ with respect to their market vicinity, which is strongly related to the intensity of firms' engagement in these technologies. We should expect that the vicinity to the market influences the ways in which network relations are formed. For example, Rowley et al. (2000) show that dense (ego-)networks are more important in contexts of knowledge exploitation whereas structural holes (Burt, 1992) are important to draw on a variety of knowledge sources in exploration contexts.

3 Data

3.1 Networks based on R&D subsidies data

In order to construct knowledge networks for different technologies, we exploit a new database on joint R&D projects that were subsidized by the German federal government. This database allows us to retrieve information on organizations that receive subsidies and to establish connections between organizations working on joint projects. This type of information documents linkages at an earlier stage compared to patent data and includes actors from the private and from the public research sector.[1]

R&D subsidies are used to stimulate private research in fields that are politically desirable. In Germany this applies to new technologies and so-called key technologies that are foremost supported (Fier, 2002). Moreover, policy aims at increasing investments in R&D because the latter is perceived to be below a desired level. A rich literature analyzes the effects of R&D subsidies, finding that subsidies impact firms' R&D efforts (e.g., Busom, 2000; Gorg and Strobl, 2007) and employment growth (e.g., Brouwer et al., 1993; Koski, 2008), and collaborating and patenting activities (see, e.g., Czarnitzki and Hussinger, 2004; Czarnitzki et al., 2007). In other words, R&D subsidies are economically highly relevant.

For the purpose of this paper it is however even more important that R&D subsidies are increasingly being used to stimulate collaboration activities. There are several reasons why R&D subsidies are more frequently been awarded to cooperative projects. First of all, some projects are just too big to be realized by a single actor making cooperation a necessity. Second, policy tries to stimulate technology transfer from public to private actors, which is why universities and research organizations are frequently participating. Thirdly, collective learning processes and knowledge spill-overs are expected from cooperative projects. For this reason actors have to sign substantive knowledge sharing agreements in order to receive the subsidies. The relevance of the cooperative element of R&D subsidies has recently been empirically confirmed by Fornahl et al. (2010).

[1] ter Wal and Boschma (2009) discuss the pros and cons of the various forms of data collection for the application of social network analysis in the context of economic geography.

Like most other advanced countries the German federal government is actively supporting public and private research and development activities with R&D subsidies programs (Czarnitzki et al., 2007). For example, in 2001 in total 7,227,838,000 Euro were spent on these measures. In 2008 this sum grew to 9,126,670,000 Euro (BMBF, 2008a). While the Federal Ministry of Education and Research (BMBF) is the primary source of this type of funding, the Federal Ministry of Economics and Technology (BMWi) and the Federal Ministry for Environment, Nature Conservation and Nuclear Safety (BMU) contribute as well. In addition to the federal ministries, the ministries of the federal states provide significant subsidies programs. Nevertheless, the federal level is still more important (Hassink, 2002).

The above mentioned ministries publish comprehensive information on the supported projects in the so-called "Foerderkatalog" (subsidies catalog), which is accessible via the website www.foerderkatalog.de. It lists detailed information on more than 110,000 individual grants that were supported between 1960 and 2009. Amongst this information are a grant's starting and ending date, a title including a very short description, the granting sum, the name and location of the receiving organization, the name and location of the executing organization, as well as a classification number. In the following some of these information are explained in more detail.

The differentiation between receiving and executing organization applies primarily to large organizations with several sub-units. In particular universities and research organizations fall into this category. While the university is the grant receiving organization, a particular chair, faculty, or institute maybe the actual executing unit. The same holds for large public research organizations like the Max-Planck-Society, which is the receiving organization and its different institutes are the executing units. This differentiation will be discussed in more detail later in the paper.

The classification number (in German "Leistungsplansystematik") is an internal classification scheme developed by the German Federal Ministry of Education and Research (BMBF) and consists of 16 main classes, which, for example, include biotechnology, energy research, sustainable development, health and medicine. These main classes are disaggregated into a varying numbers of sub-classes. These are considerably fine-grained as they allow for instance the differentiation between photonics (class: I25020), and optoelectronics (class: I25010) or plant genomics (K04210) and micro-organic genomics (K024220). At the highest level of disaggregation more than 1,100 unique activity classes have been assigned to projects between 1960-2009.[2]

In addition, the title of the project contains information on the cooperative or non-cooperative nature of projects. More precisely, cooperative projects are labeled as "Verbundprojekt" or "Verbundvorhaben" marking so-called 'joint-projects'. Organizations that participate in such joint projects agree to a number of regulations among which the following are the most important (self-translated extract of the information sheet concerning the application of subsidies for joint projects (see BMBF, 2008b)):

[2]The classification scheme as well as the assignment of projects to activity classes has been subject to some change over the years (Czarnitzki et al., 2002). For the cross-sectional approach taken in this paper this does however not causes significant biases.

1. Every partner is authorized to make unrestricted use of the project's results.
2. Intensive collaboration are the basis for finding solutions.
3. Within the scope of the project, partners grant each other a positive and free-of-charge covenant on their know-how, copy and intellectual property rights, which existed before the project's begin.
4. Amongst the project's results inventions have a special status. Extraordinary contributions to an invention have to be acknowledged.

In particular, the first three points allow for intensive knowledge exchange, while the fourth point provides incentives for innovation.

Accordingly, two organizations are argued to be cooperate if they participate in the same joint project because these involve significant knowledge exchange. We manually identify such joint projects on the basis of the title entry in the database. In general, it is a first indication of a joint project if the title contains words like "Verbundprojekt", "Verbundvorhaben", "Forschungsverbund", and "Verbund". In other cases projects have the same title but there is no indication if it is a joint project. This applies to certain special cases as, for example, the cooperation network created for the analysis of genes ("Genomforschungsnetz"). In this case, an internet search on the title is conducted to retrieve additional information. If no definite indication for a joint project is found it is treated as non-cooperative.

Large projects are frequently divided into work packages ("Teilvorhaben", "Teilprojekt"). In case it is divided into at least two work packages and each work package includes at least two partners, we only defined those organizations to be linked that participate in the same work package. An additional necessary requirement for any link to exist is that two organizations' participation periods in one project must overlap.

From the Foerderkatalog we extracted all entries that refer to ongoing projects in 2001 and 2002.[3] The resulting data set comprises 8,201 entries, which corresponds to 4,410 subsidized projects. Of these, about 41% started before 01.01.2001 and about 68% end after 31.12.2002. On average, a project lasts for about 530.7 days, which is why we consider two consecutive years for the data selection. The distribution of the length is however fairly skewed lasting from 2 days to 41 years.

3.2 Definition of network actors

Based on the data different definitions of an individual 'actor' or 'organization' can be applied. In particular, four unequal definitions are possible. The name of the granted organization is given as 'receiving organization' ("Zuwendungsempfänger"). It is differentiated from the executing unit ("Ausführende Stelle"). The latter implies that organizations are split into subunits, which is not unproblematic. For example, firms might be divided into

[3]The reason for choosing this time period is that this data has been cleaned in an earlier project as well as merged with some additional information, e.g., the types of actors. We compared the data to other periods but did not find strong variance in terms of project numbers or the share of cooperative projects.

departments ("Abteilung"). As we think of this level being to small we aggregate these entries to the next higher level, which is in most cases the firm. For very large firms however a splitting into different business fields seems appropriate. Universities and public research institutes are the most problematic cases. Sometimes they are disaggregated all the way down to working groups and chairs (individual professors). If the information is available we aggregate all universities at the level of institutes, which is a level between faculties and working groups or professor chairs. In case this information is not given the faculty level is used instead. Public research organizations are also aggregated at the institute level (if available). For both instances (receiving organization and executing unit) information on their geographic location are given, which can also be used in the definition of an appropriate actor.

We define the following four scenarios corresponding to different definitions of actors. An illustrative example is moreover provided in appendix A.

Scenario A: In the first scenario actors are identified by the receiving organization ("Zuwendungsempfänger") entry, which indicates the organization receiving the subsidies. This is the most aggregate level resulting in 2,652 actors.

Scenario B: For the second scenario we use the name of the receiving organization but disaggregated according to the location of the executing unit. In particular large organizations like Siemens AG or the Fraunhofer Society have numerous sub-units / institutes, which have different locations. Of course, applying this disaggregation implies that we don't differentiate between sub-units located in one postal code area. Using this approach we find 3,026 actors.

Scenario C: In this scenario the name of the executing unit ("Ausführende Stelle") is used to identify actors. Here we identify 4,426 actors.

Scenario D: Again, we use the name of the executing unit as defining characteristic. However, we additionally split these actors according to assigned postal code areas. This results in 4,584 individual actors.

4 Network structure

4.1 The German research network under different modes of aggregation

For an analysis, we construct networks subsuming all technologies for the four scenarios defined above and present the resulting descriptive statistics in table 1. The average project size is 1.8 mill. Euro and as expected, the number of actors is increasing with disaggregation. We can distinguish four types of actors with the majority being firms. Universities are important types of players in the domain of public research. Associations and research institutes are usually public as well but might also be public-private partnerships or non-profit organizations sponsored by private actors.

Table 1: Description of different aggregation scenarios

	Scenario A	Scenario B	Scenario C	Scenario D
Number of projects	4410	4410	4410	4410
Average size of project (TEuro)	1801	1801	1801	1801
Number of actors	2652	3026	4426	4584
Number of projects per actor	1.66	1.46	1.00	0.96
Share of firms	0.75	0.70	0.47	0.47
Share of universities	0.09	0.11	0.39	0.38
Share of institutes	0.03	0.06	0.06	0.07
Share of associations	0.13	0.13	0.08	0.08
Average subsidies firm (TEuro)	1070	1007	1067	1024
Average subsidies universities (TEuro)	6765	4890	933	922
Average subsidies institutes (TEuro)	25455	11196	7463	6800
Average subsidies association (TEuro)	5920	5306	5734	5362
Projectsum Share of firms	26.78%	26.88%	27.72%	27.72%
Projectsum Share of universities	19.85%	20.26%	20.27%	20.27%
Projectsum Share of institutes	27.57%	27.07%	26.22%	26.20%
Projectsum Share of associations	25.80%	25.79%	25.78%	25.80%
Density	0.0027	0.0023	0.0012	0.0011
Number of components	797	966	1513	1576
Size of largest component	1697	1860	2525	2587
Share in largest component	0.6399	0.6147	0.5705	0.5644
Isolates	725	871	1351	1397
Share of Isolates	0.2734	0.2878	0.3052	0.3048
Centralization (degree)	0.2283	0.0662	0.0508	0.0196
Centralization (betweenness)	0.1730	0.0453	0.0818	0.0330
Mean degree	7.2640	6.8956	5.2942	5.1684
Transitivity	0.1889	0.2752	0.4226	0.5027
Average distance	2.9867	3.3801	4.7762	5.1073
Average distance (within MC)	2.9882	3.3818	4.7790	5.1103
Diameter	7	8	12	14

Switching from the aggregation method based on the receiving organization (A and B) to scenarios based on the conducting organization (C and D) makes a substantial difference in the number of universities since they become disaggregated into faculties or institutes. Depending on the definition, universities receive average subsidies ranging between 0.9 mill. Euro to 6.8 mill. Euro. Irrespective of the scenario, we observe that institutes receive the largest funding per actor and firms receive considerably less than the typical non-profit organizations. However, this greatly underestimates firms' contributions to a project as they are required to provide own additional funding between 10% and 80% of their grant. Looking at the importance of these four types of actors as measured by the received funding, we observe roughly equal shares with firms receiving 28 per cent of the whole funding, universities receiving 20 per cent, and institutes and associations 26 per cent respectively. As such, public research is clearly dominating since institutes and associations can be either public or non-profit organizations.

Turning to the network statistics, we observe density to be decreasing with disaggregation. This is not surprising given that this measure relates the number of active links to the number of possible linkages and the latter is always increasing from scenarios A to D. Interestingly, the share of actors in the largest component (i.e., in the largest connected part of the network) seems to be rather robust against disaggregation and only decreases from 63 to 56 per cent. An indication of connectedness is also given by the mean degree, which is simply the average number relations an actor has. For scenario B, we find that on average actors are cooperating with roughly seven other actors, for scenarios C and D this measure drops to slightly more than five since the observed entities become smaller. Important differences exist in the centralization of the networks[4]. In scenario A, centralization is three to four times higher than in scenarios B and C, which are themselves more than twice as centralized as the network based on the aggregation heuristic of scenario D. The reason for the differences lies in the specificities of the large players such as the Fraunhofer-Society, which is one actor in scenario A and split up into a multitude of actors in scenarios B to D. This is documented in the increasing number of institutes from scenarios A to D and the dramatic increase in the number of universities, especially when moving to the aggregations based on the conducting organization in scenarios C and D. Contrary, the number of firms and associations remain rather stable.

We are also interested in distance measures of the network. Average distance tells us how many steps in the network have to be taken on average to contact any other actor. Since there is no connection between actors belonging to different components this only makes sense within the main component. The network diameter is the maximum distance between any two actors in the network. Unsurprisingly, both of these measures increase with higher disaggregation.[5]

In the following, we concentrate on scenario B. The reasons for this choice are based on two rationales: First, in the context of the paper we think it is not appropriate to treat very large, multi-location firms and institutes as one single actor in which information and knowledge flows freely as it is assumed in scenario A. Second, university institutes and faculties in one location should be treated as one single actor since we can assume that there is at least as much knowledge exchange across departments (or even chairs) as takes place within a firm or institute of comparable size. This would not be the case in scenarios C and D.[6]

The biggest surprise to us is the high connectedness of the network observed over such a short period and across heterogeneous technologies. This opens the question about which actors connect the network. Table 2 provides us with a first hint to answering this question. Here, the mean degree, i.e., the average number of linkages is given by actor type. Universities and research institutes are clearly the actors connecting the network with an average degree three times and almost twice as high as the mean degree of firms. The characteristics of central actors in the network will be analyzed in more detail in section 5.

[4]Centralization is a measure of concentration and can be calculated on the basis of different measures of individual level centrality.

[5]In comparison to random graphs of equal size and tie probabilities (density) leads to significantly lower centralization, a larger share of actors in the main component and larger distance measures.

[6]For other research questions, however the other actor definitions might be more appropriate.

Table 2: Mean degree by type of actor (Scenario B)

Type.Actor	Mean degree Scenario_A	Mean degree Scenario_B	Mean degree Scenario_C	Mean degree Scenario_D
association	1.70	1.63	1.73	1.67
firm	5.92	5.78	5.94	5.79
institute	18.27	10.63	7.87	7.33
university	22.96	18.09	4.84	4.80

4.2 A typology of technological systems (basic vs. applied research)

On the basis of the *Leistungsplansystematik* we can differentiate between a number of funding areas. Overall, there are 22 main fields of subsidizing (see table 3), that differ substantially in their nature. During our period of observation (2001 and 2002) we can identify projects funded in 16 research areas as indicated by an 'a' in the third column of table 3. We concentrate on the ten most important ones in terms of relevance for technological innovation and size according to funding and number of actors. Fields indicated by a 'b' in the third column of table 3 are not considered. For example, *large equipment of basic research* is excluded since this type of funding is almost per definition not cooperative and rather close to institutional funding.

While the classification scheme primarily emphasizes differences in research topics, some of the topics capture technological differences. For instance, the categories *K Biotechnology*, *D Space research and technology*, and *I ICT*, seem to have a clear technological dimension. However, similar cannot be said for categories like *H R&D improvement of working conditions*. Accordingly, interpreting these topics as technologies might not be appropriate in all instances. For this reason we refer to 'research areas' in the following. Clearly, more research is needed for example relating the 'Leistungsplansystematik' to technological classifications like the IPC codes used for patent applications and standard industry classifications like the NACE scheme.

Out of the 22, we focus on ten research areas in the following, for which we are interested if differences in the involvement of public research organizations translate into differences in the structure of the according knowledge networks.

Networks for the ten most important research areas are visualized in figure 4 and the respective network statistics are given in table 4. The networks are quite heterogeneous in terms of size but also in terms of structure. The smallest network is in *Geoscience* with 39 actors and the largest network includes 1,292 actors in *ICT*. Structural differences are for example visible in the intensity of interaction, i.e., the density of the network. Table 4 highlights that the number of links (Edges) is apparently quite correlated with the size of the networks. Therefore, density measures the share of active relations relative to all possible relations and is usually an accepted measure for the interaction intensity in a network. The problem with density is that it is also sensitive to the size of the network (Scott, 2000). Accordingly, this measure is highest in the small technologies of *Biotechnology*, *Energy*, and *Geoscience* with more than five per cent.

Table 3: Research areas according to German funding schemes (Data for 2001–2002)

Field	Research area/Technology	Short name	Excl.	Funding (Mill. EURO)	Actors
A	Institutional funding		a	–	0
B	Large equipment of basic research		b	103.2	50
C	Oceanography and polar research; offshore technology	Offshore		152.1	73
D	Space research and technology		b	25.6	4
E	Energy	Energy		802.6	49
F	Environmental research, sustainable development	Environment		552.6	529
G	Health and medical science	Medicine		766.3	172
H	R&D for the improvement of working conditions	Work		221.1	420
I	Information technology	ICT		1798.3	1292
K	Biotechnology	Biotech		262.0	82
L	Materials; phys. and chem. technologies	Materials		196.6	241
M	Aviation		a	–	0
N	Traffic		a	–	0
O	Geoscience and securing resources	Geoscience		72.3	39
P	Regional and urban planning		b	0.2	1
Q	Healthy nutrition		a	–	0
R	Sustainability		a	–	0
S	Educational research		b	325.0	401
T	Innovation, general framework conditions		a	–	0
V	Arts and humanities; social sciences		b	13.4	15
W	Miscellaneous activities	Miscellaneous		940.9	497
Y	Not R&D-relevant educational spending		b	1708.5	78

Reasons for exclusion: a – no data available, b – (almost) no cooperative funding

Cooperation intensity in our case is better measured by the mean degree which ranges between 2.26 in *Geoscience* and 10.23 in *ICT*. It is also quite high in *Medicine*, *Biotechnology*, and *Materials*. These fields are also those with the highest shares of actors in the largest component of the networks. Isolates are actors not being engaged in cooperative research at all. The share of isolates provides us with an indication how easy it is to perform research without external knowledge. In *Energy* more than 50 percent of all actors are isolated whereas in *ICT* this share lies below 10 percent. Centralization is a measure of concentration of the linkages in the network on few actors. It can be based on different individual measures of centrality, in our case on degree and on betweenness centrality.[7] *Energy* and *Biotechnology* are highly centralized networks as well as *Medicine* and *Offshore*. Networks are less concentrated in *Miscellaneous* activities, *Environment*, improvement of *Working Conditions*, and *ICT*.

The ten networks clearly show very different structures and we can now turn towards the question if this is related to the a research area's degree of application orientation. For answering this question, we first have to find a measure reflecting the degree to which an area involves basic or applied research. A natural approach is to look at the share of subsidized business firms, which we argue is proportional to the degree of application orientation. The highest shares are found in *Environment*, *Working Conditions*, *ICT*, and in *Materials*. At the other end of this scale are *Geoscience*, *Biotechnology*, and *Energy* research that accordingly can be regarded basic science. At least for *Biotechnology* this corresponds to similar classifications in the literature (Meyer-Krahmer and Schmoch, 1998). The research areas in table 4 are arranged so that the share of firms in this technology is increasing from left to right.

In this respect, it is also informative to look at the actual interaction (cooperation in joint projects) between the private (industrial) and the public science sector. Figure 1 shows the shares of linkages between different types of partners and visual inspection reveals that the intensity of interaction between science (institutes and universities) and industries increases with the share of firms. Accordingly, *Materials* shows the highest interaction intensity between the two spheres, followed by *Miscellaneous* activities, *Environment* and *Working Conditions*. *ICT* shows a somewhat lower interaction intensity, but still on a significantly higher level than the five research areas on the left side of the figure.

However, the possibilities for science-industry relations depend on the distribution of actors within the two spheres in a non-linear way. Neither without firms nor without science we could expect such relations. To account for this, we plot the relation between the share of firms and the share of science-industry relations in figure 2. The dotted curve gives the theoretical expectations if actors cooperate irrespective of partner type, i.e., the share of science-industry linkages in a complete graph. The solid curve is retained from the estimated, quadratic relationship between both variables. We can clearly observe that research areas with a higher share of firms are closer to the (unbiased) theoretical

[7] The degree centrality of actor i is the number of its ties divided by the number of possible ties $C_i = d_i/(g-1)$. *Network centralization* is then given by $C = \sum_{i=1}^{g} (\max(C_i) - C_i)/(g-2)$.

Table 4: Network description for different technologies

	Geo	Biotech	Energy	Offshore	Medicine	Misc.	Materials	Environment	Work	ICT
Nodes	39	82	49	73	172	497	241	529	420	1292
Firms	0.15	0.20	0.24	0.29	0.32	0.40	0.60	0.64	0.71	0.75
Institutes	0.33	0.40	0.35	0.27	0.35	0.42	0.21	0.20	0.15	0.12
Universities	0.53	0.40	0.41	0.44	0.33	0.18	0.19	0.16	0.14	0.13
Edges	88	470	148	218	1422	1134	1352	1982	1456	13220
Density	0.0594	0.0708	0.0629	0.0415	0.0483	0.0046	0.0234	0.0071	0.0083	0.0079
Number of components	15	33	26	32	46	282	57	184	139	142
Size of largest component	10	50	24	34	117	56	156	227	175	1111
Share in largest component	0.26	0.61	0.49	0.47	0.68	0.11	0.65	0.43	0.42	0.86
Isolates	6	32	25	29	38	243	38	156	97	122
Share of Isolates	0.15	0.39	0.51	0.40	0.22	0.49	0.16	0.29	0.23	0.09
Centralization (degree)	0.1871	0.2818	0.2819	0.2430	0.2528	0.0318	0.1445	0.0803	0.0828	0.0898
Centralization (betweenness)	0.0228	0.1024	0.0767	0.0673	0.0834	0.0046	0.0729	0.0511	0.0586	0.0817
Mean degree	2.2564	5.7317	3.0204	2.9863	8.2674	2.2817	5.6100	3.7467	3.4667	10.2322
Transitivity	0.5725	0.6741	0.5727	0.4874	0.4865	0.7082	0.7226	0.5396	0.5211	0.3421
Average distance	1.2020	2.0837	1.9101	2.0479	2.3461	2.4936	3.2904	3.8457	3.5687	3.3379

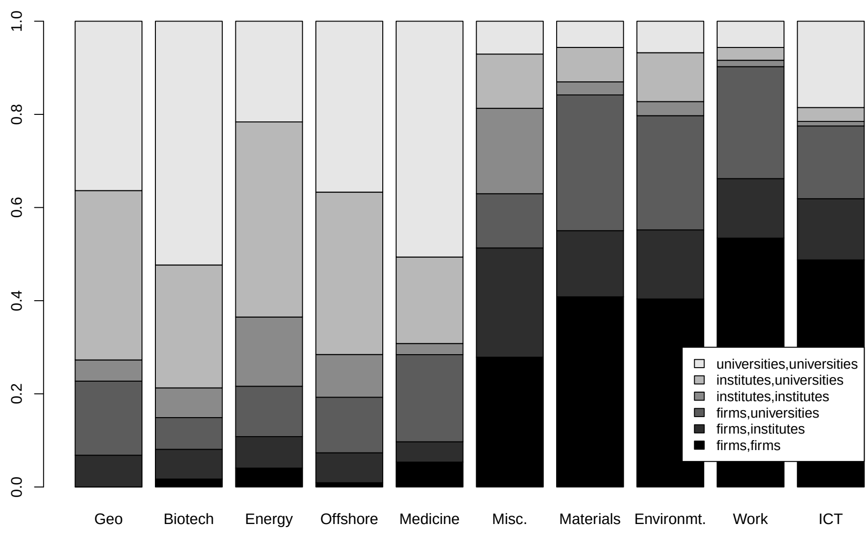

Figure 1: Shares of linkages between types of actors

expectation, indicating that the propensity for cooperation between science and industry is increasing with the orientation towards applied research. The second observation is that across all fields science-industry relations are below the unbiased expectation. Consequently our findings show a clear bias towards cooperation within the own sphere but this bias decreases within research areas that have a stronger focus on applied research, as indicated by the share of firms active in the field.

In light of these findings, the share of firms appears as a suitable measure for the application orientation of a research area. It simultaneously approximates the extent to which innovation processes are carried out by private firm and the focus of science organizations on the generation of applied knowledge.

4.3 Basic vs. applied research and network characteristics

Table 4 already reveals that certain network characteristics seem to be related to the application orientation of research. For instance, as pointed out above, the density of the network seems to fall with increasing application orientation. However, this may also be explained by the size of the network (number of nodes). In our search for the relationship between application orientation and global network structures, we correlate various properties of the networks (see table 6). The results indicate that networks within applied research are larger, less centralized, have a higher mean degree, and show a larger share of actors in the main component. To refine this observation, we perform a principal component analysis of the ten networks with respect to six network statistics. We chose density, centralization (betweenness), mean degree, share in the main component, share of isolates, and lastly the share of firms as indication of the application orientation. The results of this analysis are presented in figure 3. The two components that are plotted in the figure explain almost 85 percent of the variance between network characteristics. Those fields with a focus on application are positioned in the top left corner of the plot, while those closer to basic science are in the bottom right. *Medicine* appears to be between the

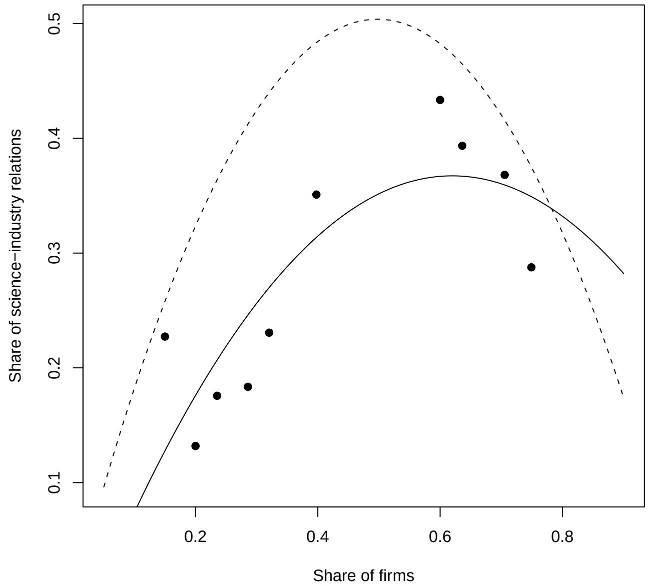

Figure 2: Linkages between firms and public research. The dotted line represents the share of linkages between firms and public research if they are formed irrespective of the type of partner. The points are the actual shares and the solid line is a quadratic fit.

typical applied and basic research networks while, unsurprisingly, *Miscellaneous* activities do not fit too well into this typology as it appears to be a conglomerate of applied and basic research.

The figure clearly reveals that the nature of the technological problems shapes the overall structure of cooperation networks. Within basic research, networks involve more isolates, are more centralized, and are more dense (due to their smaller size). The latter confirms the results of Broekel and Boschma (2010). Following their argument, a higher degree of codification and unrestricted access to (public) knowledge reduces the obstacles for cooperation. Similarly interesting is the trend for higher betweenness centralization in basic research networks. It means, that there are few crucial actors that hold the network together, i.e., connecting actors that are unconnected otherwise or only at a larger geodesic distance. Not surprisingly, this role seems to be played by universities. In contrast, the share of actors in the main component seems to be unrelated to the degree of application orientation. Similar applies to the mean degree, which seems to depend less on the basic-applied research dichotomy than we would have expected from existing work arguing for more cooperation activities in basic research (Broekel and Boschma, 2010).

As a general finding of this exercise it follows that research areas differ in their general network structure. Applied research, characterized by a large share of firms active in the field, is to a larger extent performed through cooperation as indicated by the low share of isolates. These networks are also larger, leading to a lower density, while the average number of cooperation partners does not seem to fall into this dichotomy. Basic research deals with topics tackling goals of national foresight (e.g., *Energy*, *Medicine*) in which large

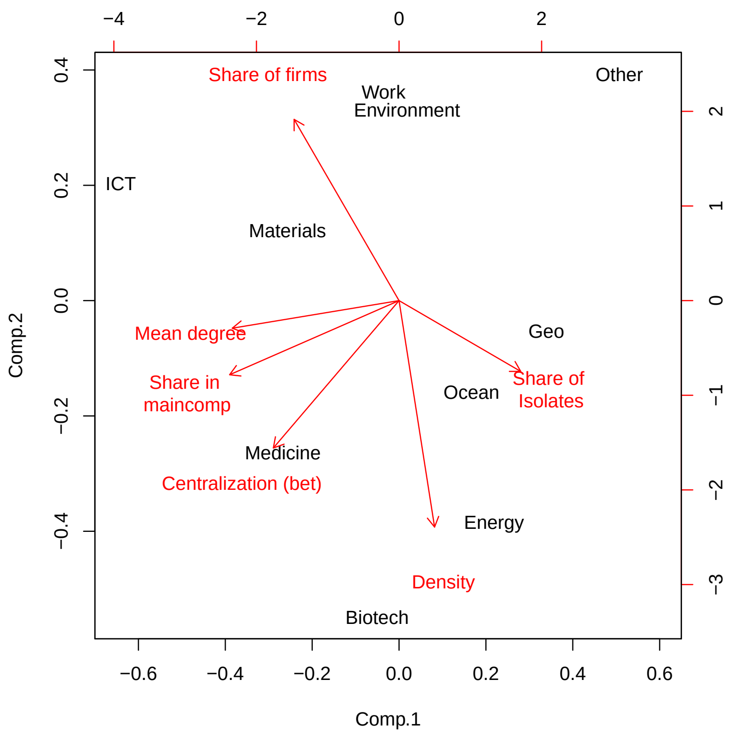

Figure 3: Biplot of principal component analysis of research networks

public research organizations are installed by the state to perform this research leading to a higher centralization.

5 Explanation of degree centrality

5.1 Network autoregression

Having explored the relationship between application orientation and the general structure of knowledge networks, we now take a more actor-oriented viewpoint. More precisely, we explore the characteristics of central actors and how they differ between basic and applied research. Given the explorative nature of our work, in the following we focus on the most simple representation of actor centrality: *degree centrality* (C_d).

In order to test the relevance of actors' characteristics on their degree centrality we use a network autocorrelation model. Accounting for network autocorrelation is necessary because network data violates a central assumption of standard regressions namely independent observations. The network characteristics of one actor are not independent of an other actors' network attributes.

The use of so-called network autocorrelation models allows circumventing this problem (see, e.g., Anselin, 1988; Leenders, 2002). Here the regression model is specified by

$$y = W1 * y + \beta * X + e, e = W2 * e + \nu \tag{1}$$

with y as the response, and X the covariance matrix. The error term ν has the usual characteristics. $W1$ and $W2$ are defined as

$$W1 = \sum_{i=1}^{p} \rho 1_i * W1_i \text{ and } W2 = \sum_{i=1}^{q} \rho 2_i * W2_i \tag{2}$$

with $W1_i$ and $W2_i$ as the elements of one or two network adjacency matrices. In this sense $W1$ and $W2$ describe the relationships between nodes in the network. $\rho 1$ can be regarded as an autoregression parameter (AR) that parameterizes the autoregression of each y value on its neighbors in the network $W1$. $\rho 2$ captures the moving average (MA) and parameterizes the autocorrelation of each disturbance in y on its neighbors in network $W2$. It accounts for an incorrect or misspecified unit of analysis (Anselin and Bera, 1988). In addition, $\rho 2$ may also take into account effects when certain events or shocks diffuse through the entire network.

In this paper we are interested in degree centrality as the dependent variable. As pointed out above, this measure is constructed by dividing the absolute number of an actor's links by the number of possible links. This transforms the count variable (degree) into a proportion (standardized degree) implying that it is bounded in the interval $[0, 1]$. We take this into account by applying a logit transformation to the variable:

$$\hat{C_d} = \frac{C_d}{1 - C_d} \quad . \tag{3}$$

To explore differences between applied and basic research the regressions are separately conducted for each of the ten research areas.

5.2 Definition of determinants

The following variables are defined to represent actors' most important characteristics, i.e., factors that might determine their centrality in knowledge networks. For these variables we test if they explanatory power of an actor's degree centrality differs between networks of basic and applied research.

We don't have any reliable information on actors' R&D activities, which are however very likely the most important determinant of an actor applying and receiving R&D subsidies. Accordingly, these activities determine the projects' sizes and the number of projects an actor is engaged in. In order to control for this effect we assume that an actors' R&D activities can be approximated by the totaled amount of received grants per day to control for the length of the projects). To be more precise, we split the variable into cooperative R&D grants (*Grant co.*) and non-cooperative R&D grants (*Grant no-co.*). This allows differentiating between actors for whom cooperative applications are relevant and actors that receive primarily non-cooperative subsidies. Such endeavor implies that we neglect all performance differences concerning the success in acquiring R&D subsidies. Based on this assumption we interpret all other factors to explain the deviation from the expected degree, which is determined by the received amount of subsidies per day. The bivariate correlation between these variables and the degree supports this approach, see table 7 in the appendix. The cooperative grants are highly correlated with degree centrality (0.9^{***}). The received non-cooperative grants are strongly negatively correlated with degree centrality (-0.6^{***}). Of course, this very high correlation is more or less induced by the construction of the variables as degree as well as received cooperative grants depend on the number of cooperation projects actors are engaged in. Again, it is not our aim to explain the intensity of cooperation but rather the network position of an actor controlling for his possibilities to cooperate. Accordingly, the bivariate correlations already reveal that those actors strongly engaged in applying for non-cooperative grants show below average degrees while the opposite is true for actors that primarily apply for cooperative grants.

In Section 3 we pointed out that the length of projects varies greatly. This hints at the existence of structurally heterogenous projects, which range from support for organizing meetings not longer than two days to the creation of new R&D jobs for many years. An actors' involvement in such structurally different projects is approximated with the average length of an actors' projects. Our primary focus is on actors' cooperation activities, wherefore we consider only cooperative projects for the construction of this variable (*Length co.*). Although it is estimated as the mean over all cooperative projects it still tends to increase with size and number of projects an actor is engaged in, which is visible in the strong correlation of this variable with *Grant co.* Since shorter projects absorb less resources, they leave more possibilities for a variety of cooperations and consequently we expect *Length co.* to influence centrality in a negative way.

While we don't have information on the size of universities and research organizations, the size of firms is approximated with their number of employees. As such, employment can be interpreted as an interaction term between employment and the firm dummy. We have to point out though that this information is far from complete (large number of empty values) and the employment numbers are only available at the most aggregate level of the receiving organization ("Zuwendungsempfänger"). Accordingly, all subsidiaries of a firm have the same employment number corresponding to the firm's total employment. Quite naturally, we expect larger firms to hold more central positions. Note furthermore that the variables accounting for the received subsidies are also likely to capture some size effects as large organizations have more resources to be simultaneously active in multiple projects.

Two variables are included that account for actors being universities (*Univ.*) or firms (*Firm*) as opposed to the heterogenous control group (e.g., research organizations and associations). In particular, universities are often shown to be central actors, at least in regional networks (Graf and Henning, 2009). In the chosen Scenario B, universities are furthermore treated as single actors aggregating all their faculties and institutes. Obviously, we expect universities to be more central in basic research while firms should be the central players in applied research (see, e.g., Niosi and Banik, 2005). Research institutes and associations cannot be as clearly assigned to one or the other orientation as they are quite heterogenous with respect to their research focus. They are therefore employed as control group.

In order to account for potential agglomeration and head-quarter effects, we include a number of regional dummies. Large multinational firms are particularly located in Munich, Frankfurt, Stuttgart, and Berlin, which is why dummies for these regions are constructed. Another dummy is created capturing if an actor is located on the territory of the former GDR, i.e., in the *Neue Länder*. These regions still suffer from the troublesome starting conditions after the German reunification, during which organizational structures have been distorted and established relationships were broken (see for a discussion Fritsch and Graf, 2010). Nevertheless, East German organizations seem to engage more actively in collaboration than their West German counterparts (Cantner and Meder, 2008) and tend to show a stronger focus on intra-regional relationships (Beise and Stahl, 1999). Moreover, East German organizations might be preferred by policy incentive schemes to accelerate the catching-up process, which may also lead to higher average centrality (Czarnitzki and Hussinger, 2004).

We also consider variables covering the funding history to account for previous success in the acquisition of R&D subsidies since the likelihood of an actor participating in multiple projects is foremost determined by its ability to produce high quality applications. This quality is not only dependent on an actor's technological superiority but also on its capability to write excellent proposals, find appropriate cooperation partners, and coordinate their contributions. A good track record and a positive image of being a trustworthy cooperation partner should also increase the chances to participate in multiple projects. *Past experience* is a dummy indicating if an actor has previously received R&D subsidies. Similarly, *past degree* refers to the degree centrality of an actor in the subsidies network existing between 01.01.1991 and 31.12.2000.[8] This variable captures three effects. Firstly, the more central an actor has been in the past, the more visible it is for future cooperation partners. The actor may hence build up a certain reputation for participating in particular funding schemes. Secondly, similar to the previous variable past degree captures past experience in subsidies programs. Thirdly, as for the case of East German firms, some policy programs may be designed to support a specific group of actors resulting in higher degrees of these actors. It is however beyond the scope of the paper to disentangle these effects.

The last variable is the number of research areas an actor is active in. This accounts for an actor's diversification and engagement in different research areas (i.e., technologies). While we expect diversification to increase with size (employment) the two variables are only weakly positive correlated (see table 7 in the appendix) This is probably because of missing employment data for universities, research institutes, and associations. Instead, universities are engaged in above average numbers of research areas. We pointed out before that this is caused by the heterogeneity of universities, which are not split into faculties.

[8]For the construction of this network we used automatic name matching procedures including fuzzy string matching.

5.3 Determinants of centrality

The results of the regressions on actors' degree centralities for each research area are shown in table 5. The model fits (adj. R^2) are satisfying. With the exceptions of *Energy* and *Medicine* they are well above 0.4 and even reach 0.7 for three research areas. The significance of the $\rho1$ and $\rho2$ parameters in most models highlights that network autocorrelation is present. This comes at no surprise given the existence of a large main component in all research areas' networks (see section 4).

The discussion of our results is organized in a way to separate more general findings that apply in all research areas from those results that allow us to identify characteristic differences with respect to the orientations towards basic or applied research. First, we take a brief look at factors that determine actors' degree centrality equally in all networks. Subsequently, we discuss factors that systematically differ between the networks.

5.3.1 General findings

Meeting our expectations, cooperative funding per day (*Grant co.*) is positively significant in all models, suggesting that actors following a cooperative strategy when applying for R&D subsidies take more central positions in the network. The negative coefficient for non-cooperative funding (*Grant no-co.*) additionally underlines this point. Actors either apply for cooperative or for non-cooperative subsidies implying that actors following the latter strategy are isolates in the R&D networks. In line with our expectations, the length of cooperative projects (*Length co.*) is negative significant in almost all research areas. The only research area with a positive coefficient is *ICT*. This indicates that with the exception of *ICT*, shorter project lengths generally lead to more collaboration and consequently higher centrality. Hence, reducing the length of projects might be an appropriate strategy to stimulate collaboration between different partners.

With respect to the regional dummies for agglomerations strong in science and high-tech industries (*Munich*, *Berlin*, *Frankfurt*, and *Stuttgart*) the results clearly contrast our expectations. In the (rare) significant cases we only observe negative coefficients, which is especially puzzling in the case of Biotechnology where Munich was one of the winners of the BioRegio contest, a policy measure directly aimed at fostering cooperation on a regional level (Dohse, 2000). Apart from that, our results imply that we can rule out critical biases arising from the suspected 'headquarter' effect. We very likely control for this effect by taking into account the location of the executing unit in the actor definition (see section 3.2).

We don't observe any indication for East German actors being more cooperative than their West German counterparts. To the contrary and against our hypothesis, actors in East Germany show significantly smaller degree centralities than their West German counterparts in six research areas. In case of *ICT* our results add to the critique on the BMBF to have failed its objectives in East Germany (see Zumpe, 2002). While our data does not allow us to assess if the subsidies added to a decrease in the technological gap between the two parts of Germany, our findings imply that cooperation was not stimulated in East Germany. Under the assumption of a positive effect of cooperation on innovation (Powell et al., 1996) this can also be interpreted as a policy failure.

The size of the technology portfolio (*# Tech.*) shows significant negative coefficients in five research areas, including basic research areas *Offshore* and *Energy* as well as rather applied research areas *Working Conditions*, *ICT*, and *Materials*. As such, no clear relation with our basic vs. applied dichotomy can be reported. This variable is however highly correlated with the university dummy and non-cooperative grants. Accordingly, diversified organizations (such as universities) do not engage in many cooperations within one research area. Most likely this is due to their cooperative engagements being spread over different fields. While this may seem logical if we assume that organizations can handle only a limited number of cooperations it has a strong implication: Organizations that bridge different research areas are not the most central actors when separately looking at each research area. It suggests that research areas (i.e., technologies) constitute separate networks that are linked by few actors. The finding explains also why universities have the highest mean degree of all types of actors (see table 2) while in the regressions being a university goes along with a below average degree.

With respect to persistency in the research networks, past degree never becomes significant and accordingly we cannot put it into relation with the degree of application orientation. The reason for the low significance levels of this variables appears to be the low number of positive values, which is an interesting result on its own. Out of 3,026 actors present in all research areas' networks, only 371 are characterized by nonzero past degree values.

5.3.2 Differences between basic and applied research

Apart from our general findings, a number of factors determining actors' degree centrality seem to differ between basic and applied research areas. The coefficient for *Universities* is significantly negative within five research areas, which implies that universities in these research fields are on average less central than our control group of research institutes and associations. These five areas can be regarded as rather applied since they are among the six research areas with the highest share of firms. Among these, only in *Materials* universities do not seem to differ from the control group with respect to centrality. Overall, this implies that the control group is more central than universities in applied but not in basic research. This concerns in particular the generally perceived prominent position held by applied research institutes, such as the Fraunhofer society, within the German national system. Unfortunately, our data does not allow us to control for size effects, which might influence this result as well. Nevertheless, this does not mean that universities do not cooperate in applied research. There are a number of universities that are very important in these networks as well. For example, the Katholische Universität Eichstätt-Ingolstadt and the University of Potsdam are highly central in the network of *ICT*.

On average, firms are characterized by relatively smaller grants and they cover less research areas than universities and the control group. This is controlled for, which is why the *Firm* dummy can be used to identify differences between the three types of

Table 5: Determinants of degree centrality

	Geo	Biotech	Energy	Offshore	Medicine	Miscellaneous	Materials	Envir.	Work	ICT
Grant co.	3.34***	2.43***	2.83***	3.25***	1.06***	1.78***	1.08***	1.58***	1.67***	0.76***
Grant no-co.	-2.78***	-2.29***	-2.33***	-2.28***	-2.09***	-2.45***	-1.75***	-2.40***	-2.49***	-1.57***
Length co.	-0.02***	-0.01***	-0.01***	-0.02***	-0.01***	-0.01***	-0.01***	-0.01***	-0.01***	0.00***
Univ.	0.41	-1.74	-2.66	0.25	-3.65**	-2.60**	0.82	-2.80***	-2.34*	-7.29***
Firm.	17.71*	-3.24	-2.01	-6.09**	-5.88***	-5.22***	-5.77***	-6.87***	-10.60***	-7.73***
Employ.	-5.29**	0.01	-0.95*	-0.09	0.02	0.11	-0.07	0.11	0.22*	0.15**
Munich	2.17	-5.84**	6.09	-3.59	-3.86**	-3.23.	-0.55	-2.43*	1.19	-0.91*
Berlin	-2.10	-3.74*	-0.64	-2.71	-1.49	-3.02***	-0.94	-1.38	-0.81	-1.28**
Frankfurt		-5.37*	-3.45	0.01	-2.88	-3.56.	-0.19	-0.39	0.72	-1.95**
Stuttgart	1.34	3.76	-0.09	-0.39	0.54	-4.21*	-0.32	-2.06.	-0.04	-0.90*
East	2.83	0.42	-4.54*	-3.50*	0.25	-2.98***	-1.33	-1.87**	-2.51**	-1.36***
Past	0.50	-1.40	-1.11	-0.22	-0.63	-1.74***	-1.22.	-1.28**	-0.07	-0.66**
Past degree		7745.00	6172.00	201.50	365.80	-155.40	-311.20	438.40	-138.50	-511.00
# Tech.	-0.21	-0.41	-0.59*	-0.75**	-0.16	-0.04	-0.77***	-0.21	-0.24.	-0.87***
$\rho 1$	-0.11	0.00	0.02	-0.01	-0.01	-0.05.	-0.06***	-0.04**	-0.04**	0.00
$\rho 2$	-0.17	0.11***	0.13***	0.04	0.04***	0.08***	0.00	0.04***	0.04*	0.03***
Adj. R^2	0.6164	0.649	0.1408	0.742	0.3423	0.6959	0.6058	0.7094	0.7447	0.4633
Log li. (Dof):	-104.3	-249.8	-134.6	-215.1	-534.1	-1548	-694.7	-1556	-1173	-3455
	(24)	(65)	(32)	(56)	(155)	(480)	(224)	(512)	(403)	(1275)
AIC:	238.7	533.6	303.1	464.2	1102	3131	1423	3146	2379	6943
Null log li. (Dof):	-131.3	-298.3	-181.5	-266.4	-586.9	-1776	-787.3	-1827	-1424	-3811
	(37)	(80)	(47)	(71)	(170)	(495)	(239)	(527)	(418)	(1290)
AIC:	266.5	600.6	367	536.8	1178	3555	1579	3659	2852	7625
AIC diff.:	27.83	66.95	63.91	72.66	75.73	424.8	155.1	513.4	473.2	682.3

Empty entrances imply zero variance in the variable.

organizations. When analyzing the firm dummy from left to right in table 5, i.e., from basic to applied research areas, it shows that firms become decreasingly central. In the most basic field, *Geoscience*, the coefficient is positive and significant, within *Biotech* and *Energy* it is insignificant, and it turns negative within the most applied fields. However, this effect has to be discussed against the background of the coefficient of *Employment*. As acknowledged above, we have employment data only for firms. Therefore, *Employment* captures some of the firm effect, especially in fields with large variances in firm size that tends to increase with the share of firms.

Firm size, in terms of employment, is significant only in four out of ten research areas. We find a negative influence on degree centrality in two basic research areas (*Energy*, *Geoscience*) and a positive influence in two applied fields (*ICT*, *Work*). Taking the effects of the *Firm* dummy and of *Employment* together leads us to the following interpretation. Despite the positive coefficient, firms are less central in basic research, since this effect is captured by employment, which is negative. In applied research, firms are not more central per se (negative firm dummy) but only if they are sufficiently large (positive effect of employment). This is not coming at any surprise as large firms benefit from having more resources and a higher absorptive capacity, which allows them to identify more opportunities to cooperate as well as manage and exploit their networks (Cohen and Levinthal, 1990). Our results are also inline with the empirical evidence on the basis of patent application data (see, e.g., Giuri and Mariani, 2005). Despite the strong focus of the BMBF on small and medium-sized firms (see BMBF, 2002) the result indicates that there is no bias towards smaller firms in the granting policy of cooperative R&D subsidies.

Another influence that differs between basic and applied research is past experience. The dummy indicating if an actor previously received a subsidy grant becomes significant in three applied research areas (*Environment*, *ICT*, *Miscellaneous*). Accordingly, we argue that in application oriented research areas past degree is negatively impacting degree centrality, which is somewhat surprising. There are a number of explanations for the negative effects of this variable. From a pessimistic viewpoint one can argue that actors either have experienced negative effects of cooperation in the past or that they have proven to be non-reliable cooperation partners and the bad reputation is now preventing them from getting access to new cooperative projects. In contrast, a more optimistic interpretation is that subsidies are not always granted to the same organizations. Be it because the goals of the subsidy programs change or that organizations fail to install effective lobbying activities giving them continuous access to subsidies. However, the variable is troubled by a low number of positive values. Out of 3.395 total actors present in all technologies' networks, only 497 are identified to have received subsidies before. As such, the low persistence might also be due to a higher turbulence characterizing fields with many firms.

6 Conclusion

With the present study, we pursued two research objectives. First, we introduced a novel data source for the analysis of cooperations in R&D and resulting network structures. Second, we apply this data to show that knowledge networks differ systematically between technologies.

The increasing interest of academe and policy in structures of interaction within the innovation process calls for high quality network data that not only covers technological and geographical dimensions but also allows for longitudinal network studies. We presented a new data source that fulfills these criteria and can be seen as complimentary to the heavily used patent and publication data. The presented data covers R&D cooperations that received subsidies from the German federal government. It is very rich in detail and freely available, however it has rarely been used for the research on cooperation networks. While all available secondary network data have their drawbacks, R&D subsidies are subject to the political decision to subsidize R&D in particular fields and according to certain granting schemes. For this reason, results based on subsidy data have to be interpreted with care and it remains to be explored if (and how) the observed network structures are 'designed' by policy. Comparing networks based on R&D subsidies with networks constructed from alternative data sources, e.g., patents, publications, seems to be a natural approach for shedding light on this issue.

However, compared to other network data, data on subsidies also has some important advantages as it covers technologies at an earlier stage of the innovation process and in different areas of research than patents. For example, technologies in which patents are useful means of appropriating returns to R&D are characterized by lower degrees of market failure since the produced good (knowledge) is made private through patent protection. Industries with intense patenting should therefore not be extensively subsidized by public policy. For fields in which patenting is less common, we might distinguish two cases. In the first case, other means of appropriation are viable and apparently more appropriate to grasp the fruits of innovation. In such cases there is also no need for subsidies. Consequently, we would only cover parts of these systems with either patent or subsidies data. In the second case, appropriability is limited and externalities are especially high calling for public intervention. These are the areas of the national innovation system that are better grasped by subsidies data than by patents. Accordingly, we argue that data on subsidies is to be viewed as a type of secondary network data, which has its flaws just like other types of publicly available data on networks (patents, publications) but additionally serves as a tool for the analysis of public policy towards research and development.

With respect to our second objective, we extended the existing body of research by comparing knowledge networks in ten technologies concerning global network properties and characteristics of central actors. We showed that knowledge networks in basic research areas are considerably different from knowledge networks concerning application oriented research. In particular, networks in basic research are smaller, more centralized, and more dense but involve more isolates. In applied research large firms and public research institutes are highly central and play an integral role, while universities are of smaller relevance.

However, our study only marks a starting point for further research as knowledge networks could be classified according to multiple dimensions with the degree of application orientation being just one. For instance, future investigations may explore if knowledge networks in related technologies are more similar than knowledge networks in unrelated fields, thereby introducing the nature of the knowledge base and the related search process as a determinant of interaction structures. Similarly interesting is the impact of varying geographical, institutional, and organizational settings. A major challenge also exists in the integration of the macro-level approach, i.e., global network properties, with the rich literature focusing on the micro level, i.e., characteristics of the dyads. Accordingly, it is not sufficient to explain network formation for a particular case (industry, region, technology), but we require more comparative studies, which add to our understanding of the driving forces behind network formation from a macro *and* from a micro perspective.

References

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Kluwer, Norwell, MA.

Anselin, L. and Bera, A. (1988). Spatial Dependence in Linear regression Models with an Introduction to Spatial Econometrics. In Ullah, A. and Giles, D. E., editors, *Handbook of Applied Economic Statistics*. Marcel Dekker, New York.

Asheim, B. T. and Coenen, L. (2005). Knowledge bases and regional innovation systems: Comparing nordic clusters. *Research Policy*, 34(8):1173–1190.

Balconi, M., Breschi, S., and Lissoni, F. (2004). Networks of inventors and the role of academia: An exploration of italian patent data. *Research Policy*, 33:127–145.

Balland, P.-A. (2010). Proximity and the evolution of collaborative networks: Evidence from r&d projects within the gnss industry. *Regional Studies*, forthcoming.

Barabasi, A., Jeong, H., Neda, Z., Ravasz, E., Schubert, A., and Vicsek, T. (2002). Evolution of the social network of scientific collaborations. *Physica*, A 311(3–4):590–614.

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

BMBF (2002). Mittelstand Innovativ: Kleine und mittlere Unternehmen im Fokus der Bildungs- und Forschungspolitik. *BMBF Publik*.

BMBF (2008a). Förderbereichen / Förderschwerpunkten und Förderarten. *Statistiken des Bundesministeriums für Bildung und Forschung*.

BMBF (2008b). Merkblatt für Antragsteller/Zuwendungsempfänger zur Zusammenarbeit der Partner von Verbundprojekten. *Bundesministerium für Bildung und Forschung, BMBF-Vordruck 0110/10.08*.

Broekel, T. and Boschma, R. A. (2010). Aviation, space or aerospace? exploring the knowledge networks of two industries in the Netherlands. *Working Papers on Innovation and Space*, 2010-05.

Brouwer, E., Kleinknecht, A., and Reijen, J. (1993). Employment growth and innovation at the firm level. *Journal of Evolutionary Economics*, 3:153–159.

Burt, R. S. (1992). *Structural Holes: The Social Structure of Competition.* Harvard University Press, Cambridge, MA.

Busom, I. (2000). An empirical evaluation of the effects of R&D subsidies. *Economics of Innovation and New Technology*, 9(2):111–148.

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

Cohen, W. and Levinthal, D. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, 35(1):128–152.

Czarnitzki, D., Doherr, T., Fier, A., Licht, G., and Rammer, C. (2002). Öffentliche Fördergung der Forschungs- und Innovationsaktivitäten von Unternehmen in Deutschland. *ZEW - Berichte*.

Czarnitzki, D., Ebersberger, B., and Fier, A. (2007). The relationship between R&D collaboration, subsidies, and R&D performance. *Journal of Applied Econometrics*, 22(7):1347–1366.

Czarnitzki, D. and Hussinger, K. (2004). The link between R&D subsidies and R&D spending and technological performance. *ZEW Discussion Paper*, 56.

D'Aspremont, C. and Jacquemin, A. (1988). Cooperative and noncooperative R&D in duopoly with spillovers. *American Economic Review*, 78(5):1133–1137.

Dohse, D. (2000). Technology policy and the regions - the case of the BioRegio contest. *Research Policy*, 29(9):1111–1133.

Fier, A. (2002). *Staatliche Förderung industrieller Forschung in Deutschland. Eine empirische Wirkungsanalyse der direkten Projektförderung des Bundes.* Nomos Verlagsgesellschaft, Baden-Baden.

Fleming, L. and Frenken, K. (2007). The evolution of inventor networks in the Silicon Valley and Boston regions. *Advances in Complex Systems*, 10(1):53–71.

Fleming, L., King, C., and Juda, A. (2006). Small worlds and regional innovation. *SSRN eLibrary*, page 892871.

Fornahl, D., Broekel, T., and Boschma, R. A. (2010). What drives patent performance of German biotech firms? The impact of R&D subsidies, knowledge networks and their location. *Papers in Evolutionary Economic Geography*, 10.09.

Fritsch, M. and Graf, H. (2010). How national conditions affect regional innovation systems – The case of the two Germanys. Jena Economic Research Papers 2010-054, Friedrich-Schiller-University Jena, Max-Planck-Institute of Economics.

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

Gorg, H. and Strobl, E. (2007). The effect of R&D subsidies on private R&D. *Economica*, 79(294):215–234.

Graf, H. (2009). Inventor networks in emerging key technologies: information technology vs. semiconductors. Jena Economic Research Papers 2009-059, FSU Jena.

Graf, H. (2010). Gatekeepers in regional networks of innovators. *Cambridge Journal of Economics*, Advance Access(beq001):1–26.

Graf, H. and Henning, T. (2009). Public research in regional networks of innovators: A comparative study of four east German regions. *Regional Studies*, 43(10):1349–1368.

Hagedoorn, J. (2002). Inter-firm R&D partnerships: An overview of major trends and patterns since 1960. *Research Policy*, 31(4):477–492.

Hassink, R. (2002). Regional innovation support systems: Recent trends in germany and east asia. *European Planning Studies A*, 10(2):153–164.

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

Koski, H. (2008). Public R&D subsidies and employment growth - microeconomic evidence from Finnish firms. *Keskusteluaiheita - Discussion Papers of the Research Institute of the Finnish Economy*, 1143.

Leenders, R. (2002). The specification of weight sturctures in network autocorrelation models of social influence. *University of Groningen, ResearchResearch Report, Institute SOM (Systems, Organisations and Management)*, 02B09.

Malerba, F. and Orsenigo, L. (1996). Schumpeterian patterns of innovation are technology-specific. *Research Policy*, 25(3):451 – 478.

Malerba, F., Orsenigo, L., and Breschi, S. (2000). Technological regimes and schumpeterian patterns of innovation. *Economic Journal*, 110(463):388 – 410.

Meyer-Krahmer, F. and Schmoch, U. (1998). Science-based technologies: university-industry interactions in four fields. *Research Policy*, 27:835–851.

Moody, J. (2004). The structure of a social science collaboration network. *American Sociological Review*, 69:213–238.

Niosi, J. and Banik, M. (2005). The evolution and performance of biotechnology regional systems of innovation. *Cambridge Journal of Economics*, 25.

Pavitt, K. (1984). Sectoral patterns of technical change: Towards a taxonomy and a theory. *Research Policy*, 13(6):343 – 373.

Powell, W. W., Koput, K. W., and Smith-Doerr, L. (1996). Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Administrative Science Quarterly*, 41(1):116–145.

Rowley, T., Behrens, D., and Krackhardt, D. (2000). Redundant governance structures: An analysis of structural and relational embeddedness in the steel and semiconductor industries. *Strategic Management Journal*, 21(3):369–386.

Scott, J. (2000). *Social Network Analysis: A Handbook, 2nd edition.* Sage Publications, London.

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

ter Wal, A. L. J. (2008). Cluster emergence and network evolution: a longitudinal analysis of the inventor network in Sophia-Antipolis. Technical Report 08.10, Papers in Evolutionary Economic Geography.

ter Wal, A. L. J. (2009). *The structure and dynamics of knowledge networks: a proximity approach.* Geowetenschapen Proefschriften.

Uzzi, B. (1996). The sources and consequences of embeddedness for the economic performance of organizations: The network effect. *American Sociological Review*, 61(4):674–698.

Zumpe, W. (2002). Forderungen an das Programm IT-Forschung 2006. *Ergebnisnachweis - Transparenz - Effizienz - Kostenersparnis.*

A Example data to illustrate scenarios A–D

In our example we observe 4 institutes (Inst A to D), all belonging to the same society (Inst), one university department (with 2 Profs), and one firm with two subsidiaries. ZE is the receiving organization, AS the organization conducting the research, AO is the location of the conducting organization and PRJ is the identification of the project.

Example data

ZE	AS	AO	PRJ
Inst	Inst A	1	A
Inst	Inst B	1	B
Inst	Inst C	2	C
Inst	Inst D	3	D
Uni	Prof A	2	A
Uni	Department A	2	E
Uni	Department A	2	F
Uni	Prof B	2	C
Firm	Firm	1	B
Firm	Firm	3	F

Scenario A

ZE	#
Inst	4
Uni	4
Firm	2

3 Actors

Scenario B

ZE	AO	#
Inst	1	2
Inst	2	1
Inst	3	1
Uni	1	4
Firm	1	1
Firm	3	1

6 Actors

Scenario C

AS	#
Inst A	1
Inst B	1
Inst C	1
Inst D	1
Prof A	1
Department A	2
Prof B	1
Firm	2

8 Actors

Scenario D

AS	AO	#
Inst A	1	1
Inst B	1	1
Inst C	2	1
Inst D	3	1
Prof A	2	1
Department A	2	2
Prof B	2	1
Firm	1	1
Firm	3	1

9 Actors

B Network visualizations

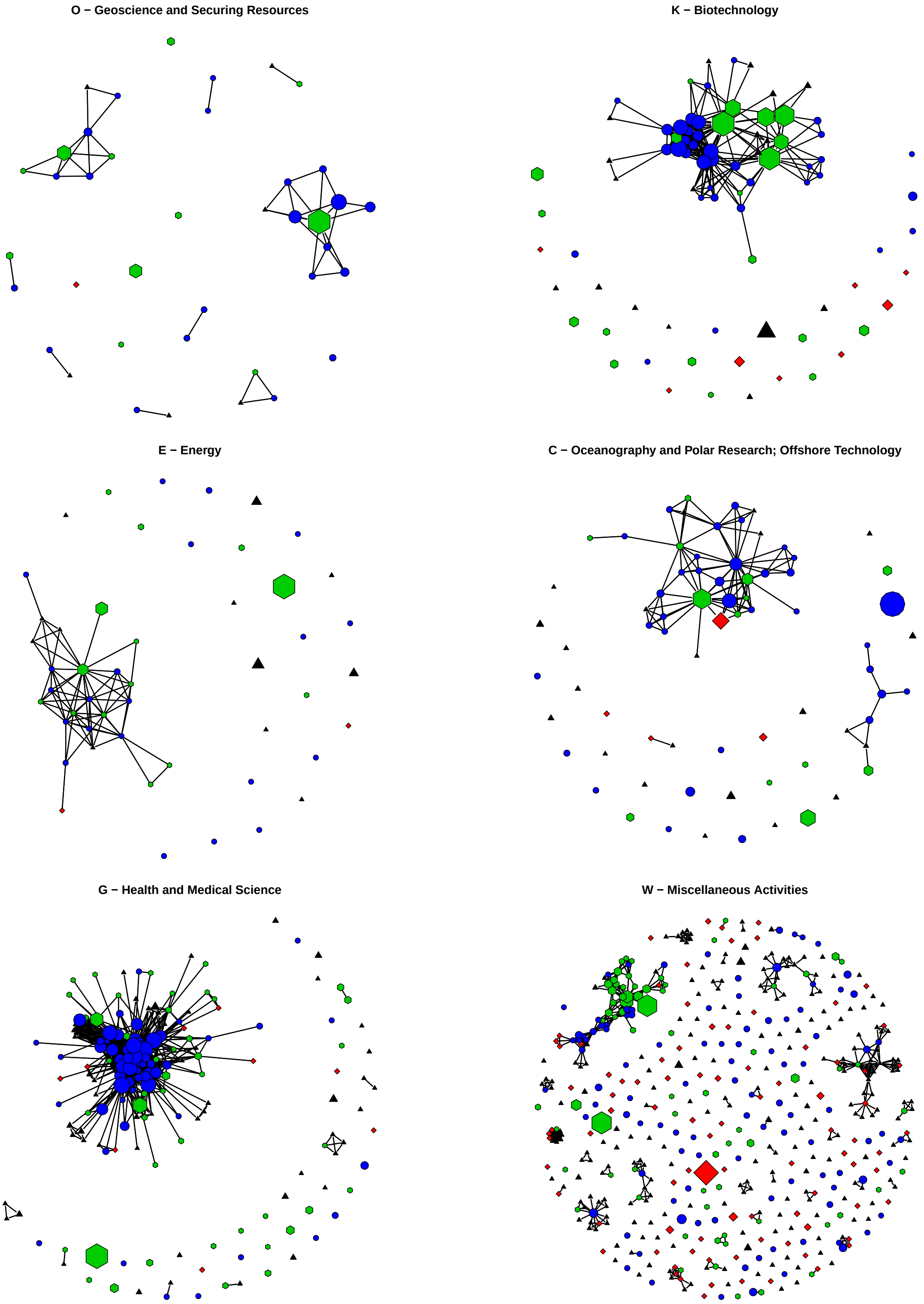

Figure 4: Network representations in different research areas
Note: Firms – black triangle, Association – red square, Institute – green hexagon, University – blue circle, node size is proportional to the amount of funding received

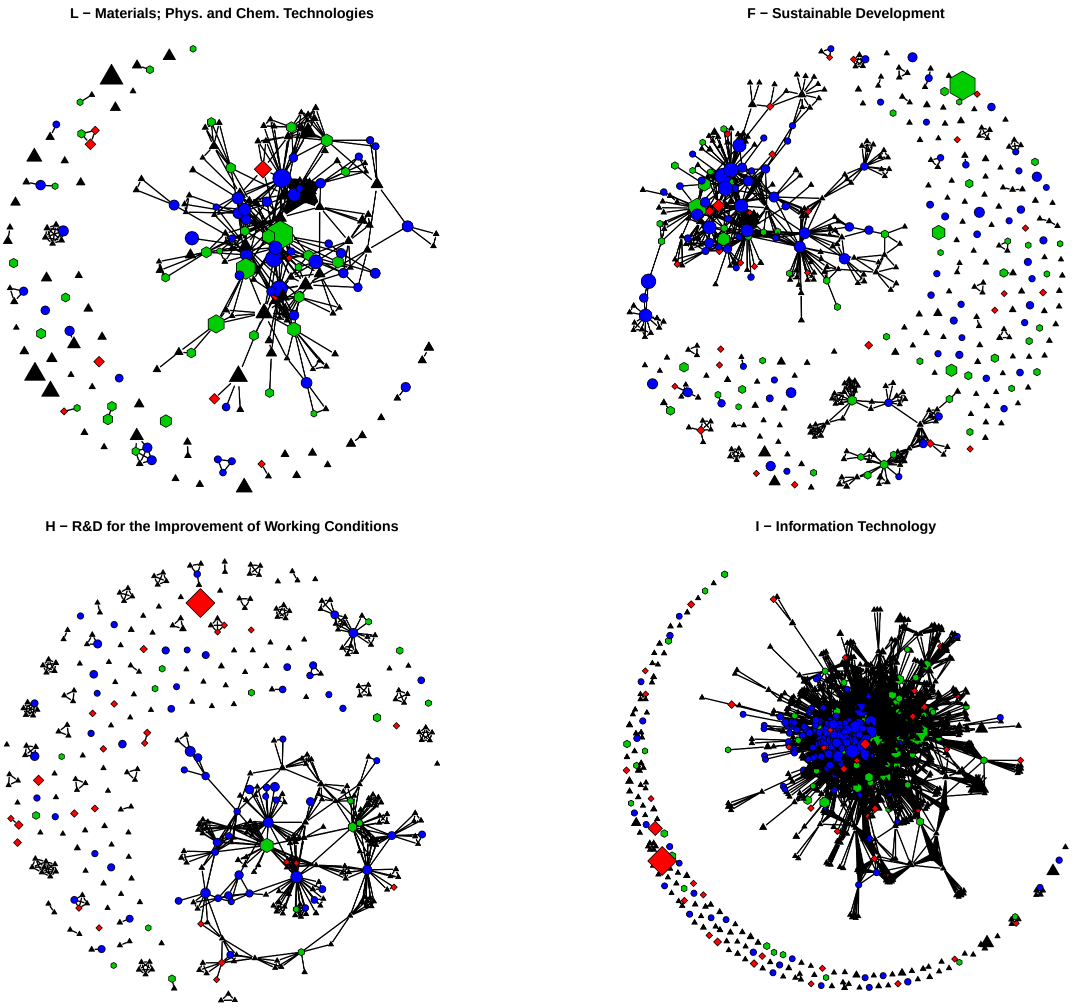

Figure 4 continued

C Correlations between network attributes

Table 6: Correlations between network attributes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
(1) Nodes	1.00															
(2) Edges	0.93	1.00														
(3) Density	-0.73	-0.46	1.00													
(4) Number of components	0.60	0.28	-0.84	1.00												
(5) Size of largest component	0.93	1.00	-0.48	0.27	1.00											
(6) Isolates	0.60	0.29	-0.80	0.99	0.27	1.00										
(7) Centralization (degree)	-0.68	-0.39	0.93	-0.88	-0.40	-0.83	1.00									
(8) Centralization (betweenness)	-0.02	0.22	0.39	-0.56	0.25	-0.54	0.60	1.00								
(9) Mean degree	0.57	0.75	-0.09	-0.09	0.75	-0.07	0.05	0.63	1.00							
(10) Transitivity	-0.53	-0.66	0.16	0.08	-0.67	0.07	0.01	-0.27	-0.50	1.00						
(11) Average distance	0.63	0.44	-0.82	0.56	0.49	0.52	-0.66	0.12	0.32	-0.17	1.00					
(12) Share in largest component	0.37	0.60	0.09	-0.38	0.62	-0.37	0.30	0.87	0.89	-0.48	0.30	1.00				
(13) Share of Isolates	-0.42	-0.52	0.25	0.15	-0.55	0.20	0.26	-0.13	-0.56	0.41	-0.32	-0.48	1.00			
(14) Firms	0.78	0.62	-0.87	0.55	0.66	0.50	-0.74	0.04	0.39	-0.32	0.95	0.33	-0.48	1.00		
(15) Institutes	-0.62	-0.60	0.59	-0.11	-0.67	-0.05	0.45	-0.19	-0.37	0.57	-0.72	-0.46	0.67	-0.84	1.00	
(16) Universities	-0.74	-0.51	0.90	-0.76	-0.54	-0.72	0.81	0.07	-0.33	0.08	-0.92	-0.18	0.25	-0.92	0.57	1.00

D Correlations between variables

Table 7: Correlation of attributes

	Degree	Grant co.	Grant no-c.	Length co.	Univ.	Firm.	Employ.	Munich	Berlin	Frankfurt	Stuttgart	East	Past grants	Past deg.
Grant co.	0.9***													
Grant no-c.	−0.6***	−0.43***												
Length co.	0.79***	0.71***	−0.72***											
Univ.	−0.01	0.07***	0.29***	−0.12***										
Firm.	0.13***	0.01	−0.44***	0.23***	−0.65***									
Employ.	0.16***	0.1***	−0.33***	0.27***	−0.46***	0.7***								
Munich	0.01	0.04**	0	0.03	−0.02	−0.01	0.04**							
Berlin	−0.05***	−0.01	0.08***	−0.03*	−0.02	−0.05***	−0.05***	−0.06***						
Frankfurt	0	0.01	−0.01	0.02	−0.03*	−0.02	0.01	−0.04**	−0.04**					
Stuttgart	0.07***	0.07***	−0.08***	0.08***	−0.06***	0.07***	0.09***	−0.05***	−0.06***	−0.04**				
East	−0.05***	−0.06***	0.03*	−0.07***	0.02	−0.04**	−0.11***	−0.09***	−0.1***	−0.06***	−0.09***			
Past grants	−0.05***	−0.02	0.11***	−0.08***	0.06***	−0.08***	−0.07***	0	0	−0.04**	0	0.02		
Past deg.	0.01	0.02	−0.02	0.02	−0.02	0.01	−0.01	0.01	0.06***	−0.01	0	−0.01	0.21***	
# Techs.	0.05***	0.14***	0.38***	−0.12***	0.66***	−0.58***	−0.39***	0.02	0.05***	−0.02	−0.04**	−0.05***	0.08***	−0.02

[1] only estimated for firms. Correlations are estimated on the pooled data for all industries.
Correlations are estimated using the pooled data of all industries.