



Munich Personal RePEc Archive

# **The cognitive and geographical structure of knowledge links and how they influence firms' innovation performance**

Broekel, Tom and Boschma, Ron

Department of Economic Geography, Faculty of Geosciences,  
Utrecht University, The Netherlands

February 2017

Online at <https://mpra.ub.uni-muenchen.de/76805/>

MPRA Paper No. 76805, posted 15 Feb 2017 11:57 UTC



## **The cognitive and geographical structure of knowledge links and how they influence firms' innovation performance**

**Tom Broekel** Firms' embeddedness in knowledge networks has received much attention in literature. However, little is known about the structure of firms' knowledge exchange with respect to different types of proximities. Based on survey data of 295 firms in 8 European regions, we show that firms' knowledge exchange systematically differs in their geographical and cognitive dimensions. We find that firms' innovation performance is enhanced if the firm primarily links to technologically related as well as technologically similar organizations. Connecting with organizations at different geographical levels yields positive effects as well.

Department of Economic Geography,  
Faculty of Geosciences,  
Utrecht University,  
The Netherlands  
E-mail:  
t.broekel@uu.nl

**Ron Boschma**

Department of Economic Geography,  
Faculty of Geosciences,  
Utrecht University,  
The Netherlands  
E-mail:  
r.boschma@geo.uu.nl

**Keywords:**  
geographical proximity,  
knowledge networks,  
technological relatedness,  
innovation performance

### **Introduction**

Access to external knowledge is crucial for firms' research and development activities as it allows complementing internal resources (Powell et al. 1996). This implies that firms' embeddedness in knowledge networks is crucial for their economic success (cf. Uzzi 1996, Cantner–Graf 2004). This raises the question about the type of factors that influence the structure of knowledge networks. Boschma (2005) proposed five

types of proximity that influence the probability of organizations to establish a knowledge link. Researchers have analysed the role these proximities play for the development of knowledge networks (cf. Breschi–Lissoni 2003, Balland, 2011) and firms' performance (cf. Gluckler 2007, Broekel–Boschma 2012).

To assess the relevance of proximities for firms' innovative success, studies usually relate the innovative output of a firm to the *average* proximity (based on one of the proximity dimensions) with its knowledge exchange partners. This means that these dimensions are often condensed into a single measure. However, little is known about the *structure* of the links of firms for knowledge exchange, concerning the types of proximity. This also applies to the impact of such structures on their innovative success. In addition, studies tend to focus only on one type of proximity, and when they do consider multiple proximity types, these are treated as being independent of each other. However, proximity in one dimension can substitute for proximity in another dimension. For instance, when two partners are cognitively distant, being geographically close allows for better communication because of the better possibilities to interact face-to-face. The proximity types may also work together when it comes to their effects on innovation. In this respect, cognitive proximity has an exceptional position among the proximity types (Boschma 2005). Similar to the other proximities, it not only facilitates knowledge exchange but also defines the learning potential for the creation of novel ideas and solutions. Therefore, we argue that the effects of the other types of proximities on innovative success have to be identified and empirically investigated in relation to that of cognitive proximity.

This paper analyses the cognitive and geographical structures of 295 firms' links for knowledge exchange in 8 European regions. In particular, we focus on the relationship between the geographical and cognitive structure of links and analyze whether firms interact more intensively with technologically similar, related, or dissimilar organizations at various geographical levels (regional and non-regional). By means of a cluster analysis technique, we establish an interaction between the cognitive with the geographical proximity dimension and systematically identify groups of firms with particular *combinations* of links with specific geographical and cognitive characteristics. We use that information to test whether particular geographical or cognitive structures of links or combinations of these are more conducive for innovative performance than other structures.

The paper is structured as follows. Section 2 briefly introduces the proximity framework and presents arguments on how the different proximities matter for innovative success. Section 3 unveils the employed data, and Section 4 sets out the empirical approach. Sections 5 and 6 present the results. Section 7 concludes the study.

### **Proximity, networks, and innovation**

Firms' embeddedness into knowledge networks is increasingly being recognized as a factor influencing their economic and innovative success (Powell et al. 1996). A rich

body of literature has emerged that analyses factors driving the formation of knowledge networks. Inspired by the French school of proximity (cf. Torre–Rallet 2005), Boschma (2005) proposes five types of proximity (i.e. cognitive, social, organizational, institutional, and geographical) that may affect firms' likelihood to engage in knowledge exchange with other organizations. A flourishing body of empirical literature has taken up these ideas and has investigated the driving forces behind network formation (cf. Breschi–Lissoni 2003, Cantner–Graf 2004, Ter Wal–Boschma 2009, Balland 2011, Broekel–Boschma 2012).

In addition, the structure of firms' links has received attention, with studies focusing on geographically proximate and distant partners (cf. Meyer–Kraemer 1985, Arndt–Sternberg 2000, Bathelt et al. 2004) as well as cognitively proximate and distant partners (cf. Fritsch 2003, Nooteboom et al. 2007, Balland et al. 2013). Especially, the economic relevance of the type of partner (in terms of the proximity dimensions) has been the focus of researchers. This focus is more relevant as high degrees of proximity may ease communication and make the establishment of links between firms more likely, but high levels of proximity may not lead to higher innovation performance *per se*, and may even possibly reduce it. This has been referred to as the 'proximity paradox' (Boschma–Frenken 2010). Accordingly, the benefits a firm can gain from particular links may be related to optimal levels of proximity in various dimensions. For instance, Sternberg & Arndt (2001) showed that balancing intra- and inter-regional links is most conducive for firms' innovation performance. Phelps (2010) found that access to greater variety improves innovation activities. Nooteboom et al. (2007) and Broekel & Boschma (2012) provided evidence that, in terms of innovative success, an optimal level of cognitive proximity might exist. That is, both higher cognitive proximity (greater technological similarity) and lower cognitive proximity (lower technological similarity) tend to reduce the innovation performance of firms.

In this context, Boschma & Frenken (2010) made a distinction between social, organizational, institutional, and geographical proximities on the one hand and cognitive proximity on the other hand. For the first four types of proximity, an 'optimal level' refers to a *combination* of links with low and high proximity. For instance, with respect to geographical proximity (and similarly for the other three proximity types), they state that an optimal level of geographical proximity does not equal an optimal geographical distance between two agents. Instead, agents should focus on balancing or combining local and non-local linkages. Contrarily, to be beneficial, they propose that an optimal level of cognitive proximity needs to exist for each link. This is not only relevant from a theoretical perspective but it also matters empirically. For instance, many studies that test for the effect of cognitive proximity estimate the *average* technological similarity (cognitive distance) of a focal firm to its direct knowledge exchange partners (Nooteboom et al. 2007, Broekel–Boschma 2012). However, the estimated *average* can be identical for very different distributions. In a very simplified example, two firms might have two partners each. For firm **A**,

one link has a maximum technological similarity index of 1 and the other one is at a minimum with value 0. The average technological similarity of this firm to its partners is 0.5, which is the same value obtained for firm **B**, which has two partners with a similarity of 0.5 each (mean similarity). Accordingly, both firms obtain the average similarity index value of 0.5. However, in the proximity framework, firm **A**'s links are expected to lower the innovation performance of the firm as a great cognitive overlap (similarity of 1) implies a lack of variety for innovation creation and a low cognitive overlap (similarity of 0) goes hand in hand with a lack of effective communication. Firm **B** is expected to have a high innovation performance, since effective communication and sufficient levels of new variety characterise all the links of firm **B**. Analyses using the averaged similarity index treat both firms **A** and **B** identically, and therefore fail to discriminate between the two very different cases.

Boschma (2005) points out that proximity types are related to each other in two ways. First, proximities can be substitutes rather than complements when it comes to the establishment of links and ensure their success. It can be sufficient to be proximate in just one dimension to link with an organization: being proximate in another dimension does not yield further effects (Boschma–Frenken 2010). This argument is backed by some empirical studies. For instance, Singh (2005) shows that geographical proximity fosters the establishment of links among researchers who are distant in the cognitive dimension. Ponds et al. (2007) find that geographical proximity helps to overcome institutional distance. As revealed in their study, industry-university relationships are more likely to be established if firms and universities are geographically proximate, while university-university relationships span greater geographical distances.

Second, proximities can be related in terms of their effects on innovation. It is still empirically unclear if proximity in one dimension can substitute for proximity in another dimension in terms of firms' innovation performance. For instance, it seems plausible that geographical proximity allows effective communication because of the possibilities of face-to-face interaction, even when two partners might be cognitively distant. However, in many cases geographical proximity might just capture the effects of other types of proximity (social, institutional, organizational), questioning its direct and independent impact.

However, cognitive proximity has an exceptional position among the other proximity types. Similar to the other types of proximity, cognitive proximity not only facilitates knowledge exchange, but also defines the learning potential, i.e. the potential for the creation of novel ideas and solutions. In this context, some cognitive distance is necessary for facilitating interactive learning and innovation (Nooteboom 2000). The other proximity types (i.e. geographical, social, institutional, and organizational) are not directly related to the learning potential and primarily impact communication efficiency and the success of knowledge exchange (Boschma 2005). Therefore, we propose that the effects of these proximities on innovative success

must be seen in relation to cognitive proximity. Our study takes up this issue by analyzing the relationship between geographical proximity (as an enabling factor of knowledge exchange) and cognitive proximity for innovation performance of firms. To be precise, we perform tests to determine if particular geographical or cognitive structures (or a combination of both) of knowledge exchange links are more conducive for firms' innovation performance.

## Data

Our empirical study is based on a database collected as part of a European project on '*Constructing regional advantage: Towards state-of-the-art regional innovation policies in Europe?*' funded by the European Science Foundation. Research teams in a number of European countries (the Czech Republic, Germany, the Netherlands, Sweden, and Turkey) interviewed firms in different European regions and industries with a harmonized questionnaire. The questions were concerned with the R&D activities of firms, their innovative success, and their engagement in knowledge exchange activities. The firms provided detailed information on all kinds of organizations (other firms, associations, universities, etc.) that they were interacting with during the last three years and with whom they exchanged technological knowledge relevant for their innovation activities. The interviews were conducted by the word of mouth, and the interviewer adapted the questions according to the context of corresponding industry. The information facilitates an analysis on the sources of technological knowledge of the firms and the manner in which it affects their innovation performance.

Table 1

### Overview of the cases

Industry	Region	Country	Firms
Automotive	Southwest Saxony	Germany	58
Aerospace	Netherlands	Netherlands	71
Biotechnology	North Rhine-Westphalia	Germany	23
Biotechnology	Prague	Czech Republic	16
Biotechnology	Scania	Sweden	30
ICT	Prague	Czech Republic	29
Video Game	Hamburg	Germany	20
Moving Media	Scania	Sweden	37
Textiles	Denizli	Turkey	31
Food	Scania	Sweden	28
Electronics	South Moravia	Czech Republic	29
<i>Total</i>			<i>372</i>

As shown in Table 1, data was collected on 372 firms in eight European regions and nine industries. A total of 295 firms reported at least one link with which they exchanged technological knowledge. On average, 5.31 links were mentioned (see Table A1). The distribution of the number of links is strongly skewed, with 50% of the firms reporting four or more links. Only 8% of the 295 firms mentioned just one organization with which they exchanged technological knowledge. The maximum number of links was 36.

In this study, we grouped firms that have similar links concerning their cognitive and geographical proximity. Our data restricted us to these two proximity types. In a similar manner as Cassiman et al. (2005), the technological similarity between an interviewed firm and its knowledge exchange partner is reported on a scale of 1 to 5 for each link. Accordingly, we measured technological similarity in a relative manner allowing for comparison across industries. For all 2,110 links, on average, a technological similarity of 2.83 was reported. For all firms and contacts, the geographical location (i.e. the latitude and longitude) was recorded, allowing the estimation of the geographical distance between the firm and the contact organization. The mean geographical distance is large with 886 km, which is however caused by a number of long distance links (>10,000). The median distance is just about 177 km.

## **Empirical Approach**

### **Defining the structure of links**

As pointed out earlier, studies analysing the effects of proximities on innovation often rely on the average geographical and cognitive distance between a firm and its direct knowledge exchange partners. However, the average distance gives only limited insights since the mean can be identical for very different empirical distributions.

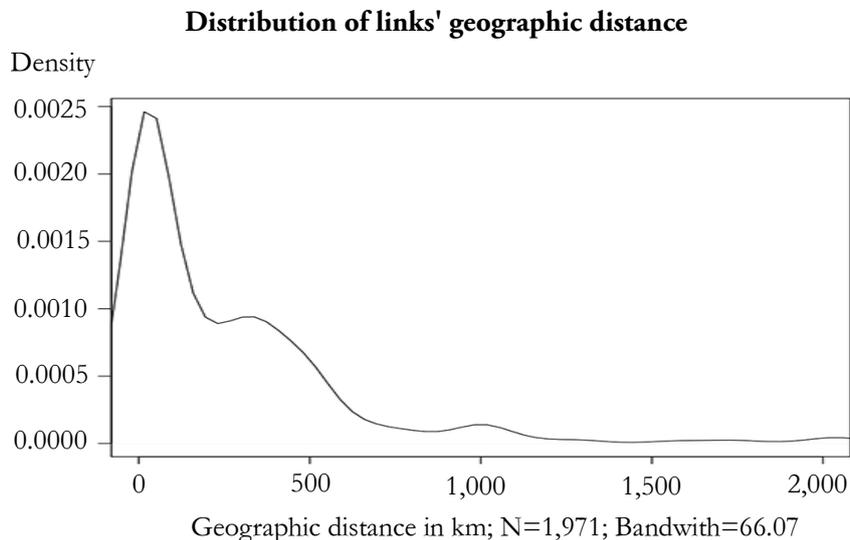
To obtain a picture of the cognitive structure of firms' links, we constructed three variables: (1) the share of links that show low technological similarity; (2) the share of links with medium technological similarity; and (3) the share of links with organizations whose knowledge base is similar to that of the focal firm. These shares are derived from the 1-5 similarity variable and are as follows: values equal to 1 are defined as low technological similarity (SHARE-VARIETY), values from 2 to 3 are considered as medium similarity relationships (SHARE-RELATED), and values of 4 and 5 refer to high similarity (SHARE-SIMILAR). This classification is based on own experiences during interviews and results in a uniform distribution of observations across these classes, as shown in Table 2. We then grouped the firms according to the similarity in their profiles by performing a cluster analysis on the three shares.

Table 2

<b>Mean values basic variables</b>	
Variables	Mean
SHARE-VARIETY	0.21
SHARE-RELATED	0.47
SHARE-SIMILAR	0.32
SHARE-REGIONAL	0.36
SHARE-MEDIUM-DIST	0.45
SHARE-LARGE-DIST	0.19
SHARE-REGIONAL&VARIETY	0.07
SHARE-REGIONAL&RELATED	0.17
SHARE-REGIONAL&SIMILAR	0.12
SHARE-NO-REGIONAL&VARIETY	0.13
SHARE-NO-REGIONAL&RELATED	0.30
SHARE-NO-REGIONAL&SIMILAR	0.20

We took an approach similar to that taken for cognitive dimension to analyse geographical dimension. We are particularly interested in the embeddedness of firms in regional knowledge networks, which refer to links with organizations at low geographical distances. It implied the need to differentiate between regional and non-regional linkages. The data covers observations from different countries and regions. For this reason, 'regional' may refer to quite different distances in each case. For example, distances between Dutch cities rarely exceed 200 km, while this is different in Sweden. Accordingly, the size of regions is likely to vary substantially, which makes the definition of a universal threshold distance problematic. Figure 1 shows the distribution of the geographical distances of all links (excluding those larger than 2,000 km), smoothed with a standard Gaussian kernel density. A clear peak of link distances at around 40 km and a flattening of the curve from 600 km onwards are visible. Based on this, three variables are created: (1) the share of links with distances equal or less than 40 km (SHARE-REGIONAL), (2) the share of linkages with distances between 40 and 600 km (SHARE-MEDIUM-DIST), and (3) the share of knowledge relations exceeding a distance of 600 km (SHARE-LARGE-DIST). Clusters for these three variables are identified with a cluster analysis.

Figure 1



As shown above, we treated the cognitive and geographical structure of firms' partners independently of each other. However, this does not take into account the relationship between geographical and cognitive proximity, which refers to the idea that the two might have a complementary or substitutive relationship. To consider this empirically, the two dimensions are made to interacted. In order to keep the complexity at an acceptable level, we dropped the differentiation between the medium-range and large distance linkages, as most of the literature regarding the effects of geographical proximity emphasize the difference between regional and non-regional knowledge links. Based on this, the following variables were created: (1) the share of regional links with low technological similarity (SHARE-REGIONAL&VARIETY), (2) the share of regional links with medium technological similarity (SHARE-REGIONAL&RELATED), and (3) the share of regional links with high technological similarity (SHARE-REGIONAL&SIMILAR). The same variables are defined for non-regional links: SHARE-NO-REGIONAL&VARIETY, SHARE-NO-REGIONAL&RELATED, and SHARE-NO-REGIONAL&SIMILAR. As stated before, a cluster analysis is performed on the six variables to group firms with similar geographical and cognitive links.

### **The cognitive and geographical structure of firms' knowledge links**

While there are several approaches to performing cluster analysis (Kaufman–Rousseeuw 1990), we used the *pdfCluster* method (Bin–Risso 2011). It employs nonparametric model-based clustering techniques as proposed by Azzalini & Torelli (2007). The advantages of this method are that the 'correct' number of clusters is

derived from the data and that it works well in situations of high dimensionality, i.e. situations in which few observations but many variables are available (for more details see Bin–Risso 2011).

### Cognitive structure of firms' knowledge links

The application of the *pdfCluster* method to the three shares representing the cognitive dimension (SHARE-VARIETY, SHARE-RELATED, and SHARE-SIMILAR) results in the identification of six clusters, whose characteristics are presented in Table 3. Each of the six clusters summarizes at least 30 observations. Cluster 1, 2, and 4 clearly represent firms with primary links to technologically similar (2: SIMILAR), technologically related (1: RELATED), and technologically unrelated (4: VARIETY) organizations. The other three clusters summarize firms with many related and some highly similar links (3: RELATED&SIMILAR), firms with many highly similar and a few related links (5: SIMILAR&RELATED), and firms with equal shares of links with low and medium similarity (6: VARIETY&RELATED).

Table 3

#### Cluster: Cognitive dimension

Cluster	1	2	3	4	5	6
Size	70	36	71	36	50	32
SHARE-VARIETY	0.02	0.01	0.07	0.93	0.11	0.46
SHARE-RELATED	0.95	0.01	0.61	0.03	0.27	0.47
SHARE-SIMILAR	0.02	0.99	0.33	0.04	0.62	0.07
Cluster Names:	<b>RELATED</b>	<b>SIMILAR</b>	<b>RELATED&amp; SIMILAR</b>	<b>VARIETY</b>	<b>SIMILAR&amp; RELATED</b>	<b>VARIETY&amp; RELATED</b>
Share of AUTO	0.09	0.12	0.30	0.04	0.32	0.14
Share of BIO	0.36	0.14	0.22	0.09	0.14	0.05
Share of AERO	0.34	0.06	0.25	0.09	0.06	0.21
Share of TEXT	0.12	0.24	0.12	0.32	0.08	0.12
Share of MEDIA	0.23	0.13	0.19	0.23	0.13	0.10
Share of ICT	0.00	0.26	0.26	0.16	0.32	0.00
Share of VIDEO	0.10	0.05	0.35	0.10	0.35	0.05
Share of FOOD	0.46	0.04	0.23	0.12	0.04	0.12
Mean LINKS_TEC	5.24	3.56	8.18	5.14	10.10	6.50
Mean AGE	37.39	21.00	30.42	26.53	20.86	30.47
Mean EMPL	170.27	82.69	706.17	78.06	84.40	118.63
Mean SKILL	49.13	43.22	41.82	46.31	36.34	39.69
Mean RD_INT	0.27	0.28	0.30	0.37	0.36	0.31
Mean SIM	2.52	4.45	3.02	1.07	3.59	1.91

In our example, average similarity (Mean SIM) seems to be a reliable approximation of the distribution as it clearly allows differentiating between firms that are dominated by links of low technological similarity (VARIETY), those dominated by links of medium similarity (RELATED), and firms that primarily link with very similar organizations (SIMILAR) links. In case of the two clusters (i.e. RELATED&SIMILAR and SIMILAR&RELATED), the cluster analysis offers richer information on the cognitive structure of firms' links than the simple mean value of technological similarity (SIM).

Some more insights into the structure are derived when relating the cluster classification to a number of firm characteristics collected during the interviews. The first characteristic is the absorptive capacity of firms. It is approximated by four variables: the share of R&D employees in total employment (RD\_INT), the share of employees with at least a bachelor's degree (SKILL), firm's age (AGE), and the number of employees (EMPL). Last, it might be worthwhile to look at the absolute number of links for exchanging technological knowledge (LINKS\_TEC). This variable represents a firms' general propensity to engage in technological knowledge exchange. Moreover, based on the eight industry dummies (AUTO, BIO, AERO, TEXT, MEDIA, ICT, VIDEO, and FOOD), we estimate the share of firms in a cluster with respect to each of these industries.

Table 3 highlights that there are no ICT-producing firms in Cluster RELATED and firms based on the textile and video gaming industry are barely present in clusters SIMILAR and VARIETY. It is important to note that no cluster is dominated by a single industry and firms of each industry are included in almost all the clusters. Therefore, the clusters do not represent industry characteristics. Cluster RELATED&SIMILAR tends to contain larger organizations (Mean EMPL) with more links (Mean LINK\_TEC). Firms mainly connected with very similar organizations (SIMILAR) tend to have relatively small number of links.

### **Geographical structure of firms' links**

The results of the cluster analysis for the geographical dimension suggest the existence of five clearly defined groups of firms, as shown in Table 4. The smallest cluster summarizes 18 firms and the largest summarizes 88 firms. Similar to the cognitive dimension, clusters 1, 2 and 4 represent the 'specialized' cases that summarize: firms which primarily link at medium distances (1: MEDIUM-DIST); firms that mainly have regional linkages (2: REGIONAL); and firms that extensively engage in long distance knowledge exchange (4: LARGE-DIST). Cluster 5 (5: MIX-DIST) includes firms that maintain an equal share of linkages to organizations at all three spatial levels. Contrarily, Cluster 3 (3: REGIONAL&MEDIUM-DIST) represents firms that lack large-distance linkages.

Table 4

**Clusters: geographical dimension**

Clusters	1	2	3	4	5
Size	69	53	88	18	67
SHARE_REGIONAL	0.04	0.94	0.44	0.03	0.24
SHARE_MED-DIST	0.91	0.04	0.52	0.04	0.32
SHARE_LARGE-DIST	0.05	0.02	0.05	0.94	0.44
Cluster Names:	<b>MEDIUM-DIST</b>	<b>REGIONAL</b>	<b>REGIONAL&amp;MEDIUM-DIST</b>	<b>LARGE-DIST</b>	<b>MIX-DIST</b>
Share of AUTO	0.35	0.02	0.53	0.02	0.09
Share of BIO	0.17	0.14	0.28	0.05	0.36
Share of AERO	0.49	0.09	0.23	0.02	0.17
Share of TEXT	0.12	0.44	0.04	0.12	0.28
Share of MEDIA	0.10	0.39	0.19	0.16	0.16
Share of ICT	0.00	0.16	0.47	0.11	0.26
Share of VIDEO	0.05	0.30	0.30	0.10	0.25
Share of FOOD	0.19	0.23	0.23	0.04	0.31
Mean LINKS_TEC	7.52	4.06	8.05	4.11	6.84
Mean AGE	37.09	27.15	27.96	17.72	25.81
Mean EMPL	151.58	817.17	91.57	235.00	146.31
Mean SKILL	27.57	47.019	45.36	44.50	52.70
Mean RD_INT	0.25	0.36	0.32	0.31	0.32
Mean DIST	437	107	322	4,447	1,887

The average geographical distance measure (Mean DIST) differentiates well between the 'specialized' cases: REGIONAL, MEDIUM-DIST, and LARGE-DIST. However, the measure works less well for differentiating between these three specialized clusters and those cluster summarizing firms whose links fall in multiple distance categories (clusters 3: REGIONAL&MEDIUM-DIST and cluster 5: MIX-DIST). In particular, firms with equal shares of links at all three geographical levels (MIX-DIST) will be classified as firms with rather long-distance relations when using the average geographical distance of their links, which obtains the second highest value of 1,887 km. However, the majority of these firms' links (56%) is still shorter than 600 km.

As in the case of the clusters based on the technological characteristics of links, we relate the cluster classification to firm characteristics. Table 4 reveals that all the clusters in this analysis contain firms of different industries. Therefore, the identified geographical patterns do not 'simply' reflect industry specifics. Moreover, firms in the

clusters REGIONAL and LARGE-DIST have, on average, fewer linkages for exchanging technological knowledge. On average, the cluster REGIONAL has larger firms (EMPL), while the cluster LARGE has relatively younger firms (AGE). Cluster MED consists of smaller (EMPL) and older firms (AGE) and firms with lower shares of highly qualified employees (SKILL).

Table 5

### Clusters: cognitive and geographical dimension

Clusters	1	2	3	4	5	6	7
Size	97	22	46	15	58	40	17
SHARE-REGIONAL&SHARE_VARIETY	0.03	0.00	0.10	0.01	0.02	0.03	0.70
SHARE-REGIONAL&SHARE-RELATED	0.04	0.02	0.02	0.96	0.42	0.10	0.15
SHARE-REGIONAL&SHARE_SIMILAR	0.03	0.02	0.22	0.00	0.09	0.39	0.06
SHARE-NO-REGIONAL&SHARE-VARIETY	0.06	0.02	0.55	0.01	0.06	0.08	0.06
SHARE-NO-REGIONAL&SHARE-RELATED	0.66	0.05	0.07	0.02	0.28	0.09	0.01
SHARE-NO-REGIONAL&SHARE-SIMILAR	0.18	0.89	0.03	0.00	0.14	0.32	0.03
Cluster names:	NO-REGIONAL	NOREGIONAL&SIMILAR	REGIONAL&SIMILAR/ NO-REGIONAL&VARIETY	REGIONAL&RELATED	RELATED	SIMILAR	REGIONAL&VARIETY
Share of AUTO	0.42	0.12	0.07	0.00	0.23	0.16	0.00
Share of BIO	0.36	0.06	0.13	0.09	0.16	0.17	0.03
Share of AERO	0.51	0.04	0.15	0.06	0.15	0.08	0.02
Share of TEXT	0.08	0.00	0.48	0.04	0.16	0.08	0.16
Share of MEDIA	0.16	0.13	0.19	0.07	0.23	0.07	0.16
Share of ICT	0.11	0.21	0.16	0.00	0.16	0.32	0.05
Share of VIDEO	0.20	0.05	0.10	0.05	0.25	0.25	0.10
Share of FOOD	0.39	0.00	0.12	0.08	0.31	0.04	0.08
Mean LINKS_TEC	8.51	5.14	4.85	2.40	7.33	6.58	5.24
Mean AGE	34.04	13.82	25.33	53.33	31.17	20.00	19.24
Mean EMPL	159.20	102.46	92.65	65.87	863.95	57.00	30.88
Mean SKILL	38.61	45.50	39.70	53.80	46.95	44.28	49.71
Mean RD_INT	0.29	0.33	0.29	0.22	0.32	0.31	0.51

### Combining cognitive and geographical dimensions of knowledge links

In this section, we explore whether firms interact more intensively with technologically similar or dissimilar organizations at particular geographical distances. For the sake of simplicity, we reduced the geographical dimension to two variables: share of regional (SHARE-REGIONAL) and non-regional links (SHARE-NO-REGIONAL). These variables are interacted with the three variables on technological similarity (SHARE-VARIETY, SHARE-RELATED, SHARE-SIMILAR). This results in six variables that serve as inputs for the cluster analysis.

As shown in Table 5, the analysis identifies seven distinct clusters, which vary in size between 15 and 97 firms. Cluster 5 (RELATED) and cluster 6 (SIMILAR) resemble mainly the cognitive structure of links. Two clusters represent firms with primarily links to organizations outside their region that are either technologically related (Cluster 1, NO-REGIONAL&RELATED) or highly similar (cluster 2, NO-REGIONAL&SIMILAR). Two other clusters consist of firms interacting primarily with regional organizations that are either technologically dissimilar (cluster 6, REGIONAL&VARIETY) or technologically related (cluster 4, REGIONAL&RELATED). Last, there is one 'mixed' cluster (cluster 3, REGIONAL&SIMILAR/NO-REGIONAL&VARIETY) that summarizes firms with complex partner structures: firms have large shares of regional relationships with technologically similar organizations and large shares of non-regional links with dissimilar organizations.

#### Substitution versus complementarity

The identification of these clusters is interesting insofar as they allow inference to either a substitutive or a complementary relationship between the cognitive and geographical dimension. According to the concept of a substitutive relationship, geographical proximity allows for easier communication and overcoming cognitive distance between firms. The existence of the REGIONAL&VARIETY cluster corresponds precisely to this concept as it summarizes firms interacting with technologically dissimilar knowledge sources at close geographical distance. Cluster NO-REGIONAL&SIMILAR describes the opposite case in which large geographical distances are overcome when the knowledge bases of two organizations are technologically similar (low cognitive distance). However, we did not find a particular geographical pattern for the exchange of related knowledge: organizations interact with organizations having related knowledge at close geographical distance (REGIONAL&RELATED), over larger distances (NO-REGIONAL&RELATED), or at a combination of geographical distances (RELATED).

However, two clusters confute this idea of substitution. They rather suggest that geographical proximity enhances the interaction with technologically similar organizations. The first cluster, SIMILAR, refers to firms linking with technologically

similar organizations at the regional as well as at the non-regional level. The cluster is orthogonal to the existence of a substitutive relationship between the two dimensions. To an even larger extent, the same applies to the cluster REGIONAL&SIMILAR/NO-REGIONAL&VARIETY. Here, firms interact with technologically similar organizations at close geographical distance and simultaneously with dissimilar organizations at high geographical distance. Therefore, these two clusters rather hint at a complementary relationship between the two proximity dimensions.

Out of the total number of firms assigned to each cluster, it is observed that... 39 firms fully support the substitution hypothesis (part of the NO-REGIONAL&SIMILAR and REGIONAL&VARIETY clusters), while 46 firms (part of the REGIONAL&SIMILAR/NO-REGIONAL&VARIETY clusters) suggest a complementary relationship. Therefore, the cluster analysis indicates that both complementary and substitutive relationships exist between the geographical and cognitive proximity types. These relationships between the two proximity dimensions was explored further by identifying the characteristics of the firms. This was achieved through a multinomial regression analysis. In this analysis, we selected all the clusters that did not indicate either of the relationships as reference groups, i.e. all the clusters that represented relationships between organizations with medium levels of technological similarity (NO-REGIONAL&RELATED, REGIONAL&RELATED, RELATED). We also aggregated the two clusters that represented the substitution hypothesis (REGIONAL&VARIETY, NO-REGIONAL&SIMILAR), owing to the presence of a small number of firms in the REGIONAL&VARIETY cluster. The two clusters that support the complementarity hypothesis (REGIONAL&SIMILAR/NO-REGIONAL&VARIETY and SIMILAR) are kept separate, as the SIMILAR cluster does not clearly indicate complementarity.

The variables that were introduced as firm characteristics include RD\_INT and SKILL, AGE, EMPL, LINKS\_TEC, and eight industry dummies. However, to obtain estimates that are more precise we included additional control variables. We control for the technological orientation of firms by considering the share of engineers (ENG) in total employment. Second, organizations may also differ with respect to their openness towards external knowledge. The variable (OPEN) describes if a firm perceives external knowledge as being highly relevant for its innovation activities.<sup>1</sup> Third, the total number of links regarding the exchange of market knowledge (LINKS\_MA) is included to control for the differences in firms' general collaboration behaviour and the potential innovation stimulating effects resulting from superior market information. As we have data on biotech firms in three

---

<sup>1</sup> This information was collected by the following question we posted during the interviews: 'Please indicate in terms of percentage the relative importance of: a) knowledge acquired inside the company; b) knowledge acquired outside the company (adding up to 100%)'.

different countries, we took into account two sector-country dummies (SWED\_BIO, CZECH\_BIO) to control for potential country effects in this sector.

The results of the analysis are presented in Table 6. The first interesting finding is concerned with the coefficients of the sector dummies. While we have already pointed out that all the clusters contain firms that belong to different industries, the regression analysis adds further evidence that none of the detected geographical and cognitive structures is restricted to just one industry. For instance, the cluster NO-REGIONAL&SIMILAR/REGIONAL&VARIETY seems to correspond primarily to the characteristics of firms in the ICT sector. However, firms of the same sector are also significant members in the cluster SIMILAR (positive significant coefficient of ICT). Similar findings are observed for other sector dummies and clusters.

Table 6

### Multinomial regression clusters

Clusters	NO- REGIONAL&SI MILAR/ REGIONAL&V ARIETY	REGIONAL&SI MILAR/ NO_REGIONA L&VARIETY	SIMILAR	NO- REGIONAL&R ELATED REGIONAL&R ELATED RELATED
Obs:	39	46	40	170
CONST	0.903	0.143	-0.178	Reference group
Log(AGE)	-0.569*	0.101	-0.103	
Log(EMPL)	-0.215	-0.561**	-0.412	
SKILL	-0.002	0.001	-0.009	
ENG	-0.005	0.006	-0.007	
RD_INT	0.295	-1.140	-1.256	
AUTO	1.478	0.185	2.336*	
BIO	-0.231	1.656*	2.968**	
AERO	-0.103	0.176	1.243	
TEXT	1.553	3.143***	1.947	
MEDIA	0.960	0.939	1.347	
VIDEO	0.498	0.221	2.822*	
ICT	2.273*	1.285	4.016***	
FOOD	(omitted)	(omitted)	(omitted)	
CZECH_BIO	0.429	-2.041	-1.154	
SWED_BIO	-0.290	-2.119*	-1.331	
EXTERN	0.007	0.000	-0.006	
LINKS_TEC	-0.132*	-0.110	-0.029	
LINKS_MA	0.005	0.044	-0.007	
Obs:	295.000			
LR chi2(51)	96.28			
Log Likelihood	-289.879			
Pseudo R2	0.1424			

Another interesting finding is that firms in cluster NO-REGIONAL&SIMILAR and REGIONAL&VARIETY tend to be younger (AGE), which suggests that the substitutive relationship between cognitive and geographical proximity holds primarily for younger firms. Given their limited experiences, geographical proximity seems to be more important for these firms for maintaining successful collaborations with technologically dissimilar organizations. Cognitive proximity eases communication and as a result enables younger firms to interact with more geographically distant partners. Moreover, a substitutive relationship between the cognitive and geographical dimension tends to be more relevant for firms with low numbers of links concerning technological knowledge (LINKS\_TEC). It might be attributed to the reluctance of firms to collaborate (having fewer links) and the caution they exercise while selecting partners. They may choose geographically proximate organizations because they can monitor them easily. Moreover, social relationships, which are more likely to exist between geographically proximate partners, might imply higher levels of trust. These relationships, in turn, may reduce the perceived obstacles to exchange knowledge between cognitively distant partners.

In contrast, the negative coefficient of EMPL for REGIONAL&SIMILAR/NO-REGIONAL&VARIETY suggests a complementary relationship between the two types of proximities that holds particularly for smaller firms. Accordingly, smaller firms (note that EMPL is not correlated with AGE) are well embedded in regional clusters of technologically similar firms. At the same time, they maintain significant relationships with dissimilar firms at larger distances. Apparently, this should make them gatekeepers that provide access to geographically distant and diverse knowledge sources. This finding clearly contradicts our expectation according to which this role should have been played by large firms (see also Lazerson–Lorenzoni 1999 and Graf 2011). We did not observe statistically significant differences between firms of cluster SIMILAR and our control group.

So far, our analysis demonstrated that systematic differences exist in the structures of firms' links along the geographical and cognitive dimension. In the subsequent sections, we focus on whether these structural differences can impact the innovation performance of firms.

## **Impact of link structures on firms' innovation performance**

### **Empirical approach**

As is common in innovation studies, we approximated the innovation performance of firms by means of the share of significantly improved products/processes on a

firm's turnover (INN).<sup>2</sup> In addition to the geographical and cognitive composition of firms' partners, we considered all previously presented firm characteristics, which might influence the innovative success of firms.

Our dependent variable (INN) is a share within the interval [0,1]. The OLS regressions can be applied to such data when the variable is 'logit' transformed ( $j = \log(y / (1 - y))$ ). However, Ferrari & Cribari-Neto (2004) point out that such an approach has several shortcomings, which include the problematic interpretation of coefficients and tendency towards asymmetric distribution of proportions (shares). Moreover, OLS regressions based on such data tend to be heteroskedastic. To overcome these drawbacks, they proposed *beta regression*, which is based on the assumption that the dependent variable is beta-distributed. Depending on the specification of some parameter values, beta's density can assume different shapes and it can be used to model left and right skewed distributions. We use beta regression with maximum-likelihood estimation and a 'logit' specification, as proposed by Cribari-Neto & Zeileis (2010), which implies that the obtained coefficients can be interpreted as change in the logit.<sup>3</sup> In some instances, our dependent variable is equal to, 0, or 1. To apply a beta regression, as suggested by Cribari-Neto & Zeileis (2010) and Smithson & Verkuilen (2006), we transformed the variable ( $y$ ) using the following equation:  $(y^{*(n-1)} + 0.5) / n$ , whereby  $n$  represents the number of observations (372).

### Factors influencing firms' innovation performance

The cluster dummies created previously are employed to approximate differences in firms' choice of knowledge exchange partners. Using the beta regression approach, these cluster dummies and other firm characteristics are regressed on the share of a firm's turnover attributable to new or significantly improved products or processes that are not older than 3 years (INN). The results are shown in Table 7. We did not report the model with all cluster dummies because the overlap between the dummies (firms are classified into multiple clusters) causes linear dependency and multicollinearity issues. We tested each proximity dimension (cognitive and geographical) separately before considering the combined clusters (cognitive & geographical).

<sup>2</sup> This information was collected with the following question: 'Please indicate how much percentage of the turnover of your firm is attributed to new, dramatically improved products/processes introduced in the last three years?'

<sup>3</sup> We also ran models specifying a 'log' and a 'probit' link function, but the results remained stable.

Table 7

**Impact on innovation performance of firms**

Model	I	p-value	II	p-value	III	p-value	IV	p-value	V	p-value	VI	p-value	VII	p-value
CONST	-0.258	0.6186	-0.238	0.653	-0.182	0.729	-0.170	0.744	-0.124	0.813	-0.149	0.774	-0.163	0.752
Log(AGE)	-0.146	0.119	-0.142	0.136	-0.151	0.108	-0.152	0.108	-0.163*	0.084	-0.159*	0.092	-0.161*	0.083
Log(EMPL)	-0.072	0.2644	-0.090	0.178	-0.092	0.166	-0.111*	0.097	-0.112*	0.094	-0.097	0.139	-0.099	0.128
SKILL	0.005**	0.0493	0.005**	0.037	0.005*	0.052	0.005**	0.047	0.005**	0.044	0.005*	0.057	0.005**	0.040
ENG	0.003	0.2887	0.003	0.260	0.003	0.242	0.003	0.184	0.003	0.197	0.003	0.231	0.003	0.229
RD_INT	-0.596*	0.0887	-0.705*	0.057	-0.678*	0.068	-0.732**	0.048	-0.709*	0.055	-0.586*	0.099	-0.664*	0.056
AUTO	-0.109	0.7408	-0.335	0.360	-0.245	0.508	-0.305	0.403	-0.256	0.483	-0.070	0.836	-0.217	0.512
BIO	0.765*	0.0871	0.637	0.177	0.588	0.214	0.633	0.179	0.562	0.231	0.742	0.101	0.562	0.209
AERO	-0.155	0.5855	-0.304	0.337	-0.297	0.351	-0.342	0.277	-0.315	0.316	-0.169	0.563	-0.255	0.370
TEXT	-0.199	0.6105	-0.292	0.488	-0.281	0.504	-0.199	0.638	-0.232	0.581	-0.104	0.799	-0.262	0.501
MEDIA	-0.326	0.4025	-0.453	0.276	-0.417	0.318	-0.398	0.335	-0.411	0.318	-0.266	0.503	-0.420	0.277
SOFT	0.071	0.8365	-0.082	0.824	-0.054	0.884	-0.044	0.904	-0.096	0.794	0.063	0.857	-0.112	0.747
VIDEO	0.886**	0.0311	0.641	0.154	0.745*	0.097	0.685	0.123	0.608	0.171	0.822**	0.049	0.611	0.141
SWED_BIO	-0.852**	0.0185	-0.890**	0.016	-0.934**	0.011	-0.893**	0.014	-0.923**	0.011	-0.942***	0.009	-0.909**	0.012
CZECH_BIO	0.048	0.9144	0.131	0.772	0.115	0.800	0.163	0.718	0.215	0.633	0.183	0.684	0.170	0.702
EXTERN	-0.004*	0.0989	-0.004*	0.099	-0.005*	0.075	-0.004	0.104	-0.004*	0.084	-0.005**	0.069	-0.004*	0.092
LINKS_MA	0.011	0.6057	0.022	0.255	0.005	0.811	-0.001	0.945	-0.001	0.945	0.002	0.935	0.005	0.816
LINKS_TEC	0.022	0.2506	0.003	0.884	0.022	0.256	0.022	0.244	0.022	0.249	0.022	0.256	0.022	0.257
REL			0.058	0.824										
HIGH			0.186	0.515										
REL_HIGH			0.646**	0.019					0.376*	0.062	0.370**	0.041	0.519***	0.003
LOW			-0.004	0.989										
HIGH_REL			0.297	0.329										
LOW_REL			0.319	0.301										
MED					0.231	0.388								
REG					0.043	0.876								
REG_MED					0.210	0.426								
LARGE					0.069	0.847								
GEO_MIX					0.501*	0.065				0.269	0.140	0.306*	0.089	0.306*
NONREG_REL							0.483*	0.069	0.275	0.323	0.026	0.883		
NONREG_HIGH							-0.004	0.991	-0.090	0.792	-0.324	0.262		
REG_HIGH_NONREG_LOW							0.032	0.907	0.004	0.989	-0.210	0.330		
REG_REL							-0.213	0.570	-0.205	0.584	-0.391	0.251		
COM_REL							0.512*	0.068	0.234	0.440	-0.020	0.936		
COM_HIGH							0.494*	0.091	0.429	0.145			0.339	0.120
REG_LOW							0.006	0.987	-0.026	0.945	-0.256	0.431		
(phi)	1.056***	0.000	1.044***	0.000	1.027***	0.000	1.092***	0.000	1.109***	0.000	1.102***	0.000	1.100***	0.000
Log-Lik.				348 on 26 Df		345.3 on 25 Df		323.5 on 27 Df		326.4 on 29 Df		325.3 on 28 Df		324.9 on 23 Df
Pseudo R <sup>2</sup>				0.196		0.183		0.197		0.212		0.207		0.205

Before discussing the results for the cluster dummies, we analyzed firm characteristics that gained significance in all model specifications (I–VII). The variable SKILL is positive and significant in all the models, implying that larger shares of highly qualified employees increase the innovation performance of firms. Surprisingly, the share of R&D employees (RD\_INT) is negative and significant in all models. However, the variable is strongly positively correlated with SKILL (+0.33\*\*) and INNO (+0.13\*). It is therefore likely that SKILL captures the positive effect of RD\_INT. As a result, the remaining variance of RD\_INT that is not captured by SKILL might reflect some sector specifics (e.g. the negative correlation between TEXT and INNO).<sup>4</sup> Moreover, EXTERN is negative and significant in almost all models. It suggests that firms that attribute a high importance to external knowledge are less innovative. While the negative significances of EMPL in models IV and V as well as that of AGE in models V, VI, and VII are in line with literature (cf. Frenkel–Schefer 1998), they appear only in a limited number of models. Although there are no signs of multicollinearity, it was shown above that EMPL and AGE are significantly related to particular structures in the combined cognitive and geographic dimensions. As these variables become significant in those models including all cluster dummies of the combined cognitive and geographic dimensions, they probably capture some effects of the corresponding clusters. However, given the weakness of their effects, we refrained from disentangling these variables any further.

Here, we focus on the effects of the cognitive and geographical proximity dimensions. In model II, we included only the dummies for clusters representing differences in the cognitive dimension of the firms' partners. We expected firms in clusters SIMILAR, VARIETY, VARIETY/RELATED, and SIMILAR/RELATED to be outperformed by the firms in clusters RELATED and RELATED/SIMILAR because only these latter two have access to related knowledge, i.e. contacts that offer sufficient variety of knowledge but still allow for effective communication. Only one dummy (RELATED/SIMILAR) was positive and significant. The results meet our expectations as they show that having a mixture of links to related and technologically similar firms is conducive for innovation performance of firms. However, we also expected RELATED to be positive and significant, which is not the case.

In model III, only the dummies of the geographical proximity dimension are included. We expected firms with a mix of linkages at different spatial levels (MIX-DIST and REGIONAL/MEDIUM-DIST) to outperform firms that focus mainly on only one type of spatial level for interaction (REGIONAL, MEDIUM-DIST, LARGE-DIST). Only the MIX-DIST cluster was positive and significant. This confirms our hypothesis that a balance of links at all three spatial levels facilitates the innovation performance of firms.

---

<sup>4</sup> It might however also account for a potential negative size effect because, when excluding EMPL from the regression, RD\_INT loses its significance.

In Model IV, the cluster dummies for the combined cognitive and geographical proximity dimensions are tested. Three of these have positive and significant effects (NO-REGIONAL&SIMILAR, RELATED, and SIMILAR). It suggests that innovation performance is particularly enhanced when firms have significant links with technologically similar organizations outside their region (NO-REGIONAL&SIMILAR), links with technologically similar organizations located at different spatial levels (SIMILAR), and linkages with technologically related organizations at various spatial scales (RELATED). A prominent observation is that no cluster dummies are significant with primarily regional linkages.

It is apparent that the geographical dimension plays only a minor role in defining these clusters. Therefore, we confront the dummies of the combined cognitive and geographical proximity dimensions with the two dummies that were found significant in the previous models (RELATED/SIMILAR, MIX-DIST). As a result, as shown in model V, NO-REGIONAL&SIMILAR, RELATED, and SIMILAR lose their significance. Their effect is captured by RELATED/SIMILAR. However, the previously significant MIX-DIST also loses its significance and regains significance when SIMILAR is removed (model VI). Although the overlap between the two clusters SIMILAR and MIX-DIST is small, their positive impact on innovation seems to be interrelated. To disentangle their effects, we excluded all insignificant dummy variables and estimated model VII that includes only RELATED-SIMILAR, SIMILAR, and MIX-DIST. The RELATED/SIMILAR, SIMILAR, and MIX-DIST clusters remained highly significant, insignificant, and positive and significant, respectively. Therefore, the innovation performance of firms is enhanced when they have access to knowledge sources at different spatial scales or link with a combination of organizations that offer related and similar technological knowledge. This implies that the best practices facilitated by the geographical structure of knowledge exchange partners (with a combination of regional and non-regional linkages) is an alternative to the best practice facilitated by the cognitive structure (with a combination of cognitively similar and related organizations).

## Conclusions

The paper investigated firms' embeddedness in knowledge networks. Taking a proximity approach, we focused on firms' knowledge exchange partners and the roles of cognitive and geographical proximities. Our investigation contributed to the literature in two ways.

First, we showed that the structure of firms' links differs systematically along the cognitive and geographical dimensions. With respect to the cognitive dimension, we found a large number of firms having relationships primarily with one type of organization (technologically dissimilar, related, or similar). Nevertheless, somewhat more than 50% of firms exchange knowledge with organizations possessing varying

degrees of knowledge similarity. For the geographical dimension, similar findings were obtained with about 53% of firms having links to organizations at various spatial scales. We also found that the two proximity dimensions share a relationship. For younger firms and those with fewer technological knowledge links, geographical proximity helps to overcome cognitive distance. In contrast, geographical and cognitive proximities are complementary for smaller firms (but not younger firms), which implies that these firms tend to be well embedded in regional clusters of technologically similar firms. At the same time, they maintain significant relationships with dissimilar firms at larger geographical distances.

Second, we tested the effects of these different partner structures on innovation. The results clearly confirmed that firms increase their innovation performance by primarily linking with technologically related firms as well as technologically similar organizations. This is in line with Boschma and Frenken's findings (2010) and confirms the findings of other studies on this issue (cf. Broekel–Boschma 2012, Fornahl et al. 2011). In accordance with existing research, it also helps firms to access knowledge sources at various geographical scales (cf. Arndt–Sternberg 2000). However, in our study, we did not find an indication for the interrelated effects of cognitive and geographical proximity (in either a substitutive or a complementary sense) on firms' innovation performance.

This study has several shortcomings. We only measured links that are realized by the firms, not the interaction possibilities that firms may possess. For instance, the absence of technologically dissimilar organizations in a region fails to facilitate communication between such organizations at the regional level. Accordingly, they might be driven to search for such collaboration partners at greater geographical distances. The study does not consider that the existing potential for collaborating with a particular organization at a particular spatial level might influence the results on the relationship between cognitive and geographical dimensions for the establishment of knowledge links.

Our measure of geographical proximity might also have captured the effect of other types of proximity, like social proximity (Boschma 2005). In this study, we cannot entirely rule out this potential of spurious correlation. Therefore, we clearly see the need for further research to include data on all proximity types that were not considered in this paper, namely social, institutional, and organizational proximity. This would certainly increase our understanding of what drives the structure of the knowledge exchange partners of firms, whether proximity dimensions act as substitutes or complements, and the manner in which different structures of knowledge exchange partners influence the innovation performance of firms.

## REFERENCES

- ARNDT, O.–STERNBERG, R. (2000): Do manufacturing firms profit from intraregional innovation linkages? An empirical based answer *European Planning Studies* 8 (4): 465–485.
- AZZALINI, A.–TORELLI, N. (2007): Clustering via nonparametric density estimation *Statistics and Computing* 17 (1): 71–80.
- BALLAND, P.-A. (2011): Proximity and the evolution of collaboration networks: Evidence from Research and Development projects within the Global Navigation Satellite System (GNSS) industry *Regional Studies* 46 (6): 741–756.
- BALLAND, P.-A.–SUIRE, R.–VICENTE, J. (2013): Structural and geographical patterns of knowledge networks in emerging technological standards: Evidence from the European GNSS industry *Economics of Innovation and New Technology* 22 (1): 47–72.
- BATHELT, H.–MALMBERG, A.–MASKELL, P. (2004): Clusters and knowledge: Local buzz, global pipelines and the process of knowledge creation *Progress in Human Geography* 28 (1): 31–56.
- BIN, R. D.–RISSO, D. (2011): A novel approach to the clustering of microarray data via nonparametric density estimation *BMC Bioinformatics* 12:49.
- BOSCHMA, R. A. (2005): Proximity and innovation: A critical assessment *Regional Studies* 39 (1): 61–74.
- BOSCHMA, R.–FRENKEN, K. (2010): The spatial evolution of innovation networks. A proximity perspective In: BOSCHMA, R.–MARTIN, R. (eds.) *The Handbook of Evolutionary Economic Geography* pp. 120–135., Edward Elgar, Cheltenham.
- BRESCHI, S.–LISSONI, F. (2003): *Mobility and social networks: Localised knowledge spillovers revisited* CESPRI Working Paper, No. 142., Università Commerciale “Luigi Bocconi”, Milano.
- BROEKEL, T.–BOSCHMA, R. (2012): Knowledge networks in the Dutch aviation industry: the proximity paradox *Journal of Economic Geography* 12 (2): 409–433.
- CANTNER, U.–GRAF, H. (2004): The network of innovators in Jena: An application of social network analysis *Research Policy* 35 (4): 463–480.
- CASSIMAN, B.–COLOMBO, M. G.–GARRONE, P.–VEUGELERS, R. (2005): The impact of M&A on the R&D process: An empirical analysis of the role of technological- and market-relatedness *Research Policy* 34 (2): 195–220.
- CRIBARI-NETO, F.–ZEILEIS, A. (2010): Beta regression in R *Journal of Statistical Software* 34 (2): 1–24.
- FERRARI, S. L. P.–CRIBARI-NETO, F. (2004): Beta Regression for modeling rates and proportions *Journal of Applied Statistics* 31 (7): 799–815.
- FORNAHL, D.–BROEKEL, T.–BOSCHMA, R. A. (2011): What drives patent performance of German biotech firms? The impact of R&D subsidies, knowledge networks and their location *Papers in Regional Science* 90 (2): 395–418.
- FRENKEL, A.–SCHEFER, D. (1998): Local milieu and innovations: Some empirical results *The Annals of Regional Science* 32 (1): 185–200.
- GLUCKLER, J. (2007): Economic geography and the evolution of networks *Journal of Economic Geography* 7 (5): 619–634.

- GRAF, H. (2011): Gatekeepers in regional networks of innovation *Cambridge Journal of Economics* 35 (1): 173–198.
- KAUFMAN, L.–ROUSSEEUW, P. J. (1990): *Finding groups in data* John Wiley and Sons Ltd. West Sussex, England.
- LAZERSON, M. H.–LORENZONI, G. (1999): The firm that feed industrial districts: A return to the Italian source *Industrial and Corporate Change* 8 (2): 235–266.
- Martin, R.–Moodysson, J. (2011): Innovation in Symbolic Industries: The Geography and Organization of Knowledge Sourcing *European Planning Studies* 19 (7): 1183–1203
- MEYER-KRAHMER, F. (1985): Innovation behaviour and regional indigenous potential *Regional Studies* 19 (6): 523–534.
- NOOTEBOOM, B. (2000): *Learning and innovation in organizations and economics* Oxford University Press, Oxford.
- NOOTEBOOM, B.–VAN HAVERBEKE, W.–DUYSTERS, G.–GILSING, V.–VAN DEN OORD, A. (2007): Optimal cognitive distance and absorptive capacity *Research Policy* 36 (7): 1016–1034.
- PHELPS, C. C. (2010): A longitudinal study of the influence of alliance network structure and composition of firm exploratory innovation *Academy of Management Journal* 53 (4): 890–913.
- PONDS, R.–VAN OORT, F.–FRENKEN, K. (2007): The geographical and institutional proximity of research collaboration *Papers in Regional Science* 86 (3): 423–443.
- POWELL, W. W.–WALTER, W.–KOPUT, K. W.–SMITH-DOERR, L. (1996): Interorganizational collaboration and the locus of innovation: Networks of learning in Biotechnology *Administrative Science Quarterly* 41 (1): 116–145.
- SINGH, J. (2005): Collaborative networks as determinants of knowledge diffusion patterns *Management Science* 51 (5): 756–770.
- SMITHSON, M.–VERKUILEN, J. (2006): A better lemon squeezer? Maximum-Likelihood Regression with Beta-Distributed Dependent Variables *Psychological Methods* 11 (1): 54–71.
- STERNBERG, R.–ARNDT, O. (2001): The firm or the region: What determines the innovation behavior of European firms? *Economic Geography* 77 (4): 364–382.
- TER WAL, A.–BOSCHMA, R. (2009): Applying social network analysis in Economic Geography: Framing some key analytical issues *Annals of Regional Science* 43 (3): 739–756.
- TORRE, A.–RALLET, A. (2005): Proximity and localization *Regional Studies* 39 (1): 47–59.
- 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.

**APPENDIX**

Table 1

**Descriptives**

Variables	Mean	Sd	Median	Min	Max	Skew	Kurtosis
INNO	0.29	0.31	0.20	0.00	1.00	1.04	-0.01
AGE	27.12	41.62	16.00	1.00	502.00	5.83	52.75
EMPLOY	220	2129	30.00	1.00	40877	18.72	361
HIGH	43.12	37.60	30.00	0.00	100.00	0.39	-1.46
ENG	48.19	40.16	50.00	0.00	100.00	-0.02	-1.67
RD_INT	0.25	0.31	0.08	0.00	1.00	1.11	-0.18
EXTERN	37.06	25.83	30.00	0.00	100.00	0.66	-0.11
CAR	0.16	0.36	0.00	0.00	1.00	1.89	1.64
BIO	0.19	0.39	0.00	0.00	1.00	1.61	0.64
AERO	0.19	0.39	0.00	0.00	1.00	1.57	0.50
TEXT	0.08	0.28	0.00	0.00	1.00	3.00	7.20
MEDIA	0.10	0.30	0.00	0.00	1.00	2.67	5,25
SOFT	0.08	0.27	0.00	0.00	1.00	3.14	8.04
VIDEO	0.05	0.23	0.00	0.00	1.00	3.94	13.86
FOOD	0.08	0.26	0.00	0.00	1.00	3.21	8.50
CZECH_BIO	0.04	0.20	0.00	0.00	1.00	4.49	18.56
SWED_BIO	0.08	0.27	0.00	0.00	1.00	3.07	7.61
LINKS_TEC	5.31	5.57	4.00	0.00	36.00	1.80	4.78
LINKS_MA	5.37	5.13	4.00	0.00	35.00	2.10	7.04