Papers in Evolutionary Economic Geography

11.18

The cognitive and geographical composition of ego-networks of firms – and how they impact on their innovation performance

Tom Broekel and Ron Boschma

Utrecht University
Urban & Regional research centre Utrecht

The cognitive and geographical composition of ego-networks of firms – and how they impact on their innovation performance

Tom Broekel*,# and Ron Boschma**

* Institute of Economic and Cultural Geography, Leibniz University of Hanover
**Urban and Regional research centre Utrecht (URU), Faculty of Geosciences, Utrecht University, The Netherlands

1 November 2011

Abstract
Firms' embeddedness into knowledge networks has received much attention in the literature. However, little is known about the composition of firms' ego-networks with respect to different types of proximities. Based on survey data of 295 firms in eight European regions, we show that the ego-networks of firms systematically differ in their geographical and cognitive embeddedness. We find that firms' innovation performance is stimulated if the firm primarily links to technologically related firms as well as technologically similar organizations. Connecting with organizations at different geographical levels yields positive effects as well.

Keywords: geographical proximity, knowledge networks, technological relatedness, innovation performance, ego-networks
JEL Codes: R11, R12, O18, O33

 # Corresponding author: Tom Broekel, Insitute of Economic and Cultural Geography, Leibniz University of Hanover, Schneiderberg 50, 30167 Hanover, Germany, Tel: +49 511 762 2765, broekel@wigeo.uni-hannover.de

1. Introduction

Access to external knowledge is crucial for firms' research and development activities as it allows to complement internal resources (Powell et al., 1996). This implies that firms' embeddedness into knowledge networks is crucial for their economic success (e.g. Uzzi, 1996; Cantner & Graf, 2004). This raises the question of what factors influence the structure of knowledge networks. Boschma (2005) proposed five proximity types that influence the probability of organizations to establish a knowledge link. Researchers have analyzed the role these proximities play for the development of knowledge networks (e.g., Breschi & Lissoni, 2003; Balland, 2011) and firms' performance (e.g. Gluckler, 2007; Broekel and Boschma, 2011).

To assess the relevance of proximities for firms' innovative success, studies usually relate the innovative output of a firm to the *average* proximity (on one of these dimensions) with the knowledge exchange partners of the firm. This means that firms' ego-networks are often condensed into a single measure. However, little is known about the *composition* of firms' ego-networks concerning the types of proximity. This also applies to the impact of particular proximity structures in firms' ego-networks on firms' innovative success. In addition, studies tend to focus only on one type of proximity, or 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, being geographically close allows for better communication when two partners are cognitively distant 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. Cognitive proximity has an exceptional position among the proximity types (Boschma (2005). Similar to the other proximities, it facilitates knowledge exchange, but it 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 seen and empirically investigated in relation to that of cognitive proximity.

This paper analyses the cognitive and geographical composition of the ego-networks of 295 firms in 8 European regions and 9 different industries. In particular, we focus on the relationship between the geographical and the cognitive structure of firms' ego-networks and analyze whether firms interact more intensively with technologically similar, related or dissimilar organizations at particular geographical levels (regional and non-regional). By means of a cluster analysis technique, we interact the cognitive with the geographical proximity dimension and systematically identify groups of firms whose ego-networks show particular *combinations* of links with specific geographical and cognitive characteristics. We use that information to test whether particular geographical or cognitive ego-network structures or combinations of these are more conducive for innovative performance than others.

The paper is structured as follows. In Section 2, the proximity framework is briefly introduced and arguments are made for how the different proximities matter for innovative success. The employed data is unveiled in Section 3, while Section 4 sets out the approach. Section 5 presents the results. Section 6 concludes.

2. Proximity, networks and innovation

Firms' embeddedness into knowledge networks is increasingly being recognized to influence their economic and innovative success (Powell et al., 1996). A rich literature has emerged that analyzes factors driving the formation of knowledge networks. Inspired by the French school of proximity (see, e.g., Torre & Rallet, 2005), Boschma (2005) proposes five types of proximity (i.e. cognitive, social, organizational, institutional, and geographical) that may impact firms' likelihood to engage into knowledge exchange with other organizations. A flourishing empirical literature has taken up these ideas and has investigated the driving forces behind network formation (see, e.g., Breschi & Lissoni, 2003; Cantner & Graf, 2004; Ter Wal & Boschma, 2009; Balland, 2011; Broekel & Boschma, 2011).

Also, the composition of firms' links has received attention in terms of geographically proximate and distant partners (see, e.g., Meyer-Krahmer, 1985; Arndt & Sternberg, 2000; Bathelt et al., 2004) as well as cognitively proximate and distant partners (see, e.g., Fritsch, 2003; Nooteboom et al., 2007; Balland et al., 2010). Especially the economic relevance of the type of partner (in terms of the proximity dimensions) has been the focus of researchers. For instance, (Sternberg & Arndt, 2001) showed that a balance between intra- and interregional 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 (2011) found evidence that in terms of innovative success, an optimal level of cognitive proximity might exist. That is, both higher cognitive proximity (i.e. greater technological similarity) as well as lower cognitive proximity (lower technological similarity) tends to reduce the innovation performance of firms. And Uzzi (1996) demonstrated that a mixture of low (social) proximity and high (social) proximity, that is, a combination of arm's length ties and embedded ties, is most conducive for the competitiveness of firms.

Boschma & Frenken (2009) proposed the concept of "proximity paradox". While high degrees of proximity ease communication and exchange of knowledge and make the establishment of links between firms more likely, too high levels of proximity may not lead to higher innovation performance *per se*, and may possibly even reduce it. Accordingly, the benefits a firm can gain from particular links may be related to optimal levels of proximity in its various dimensions. In this respect, Boschma & Frenken (2010) made an important 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 *mixture* of links with low and high proximity. For instance, they argue for geographical proximity (and similarly for the other three proximity types): "[*w*]*hen thinking about an optimal level of geographical proximity, this does not mean determining an optimal geographical distance between two agents. Instead, one should think of a balance of local and non-local linkages*" (Boschma and Frenken, 2010, p. 6-7). In contrast, 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, 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 and Boschma, 2011). 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 ego-network 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 same 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) implies a lack of effective communication. In contrast, firm **B** is expected to have a high innovation performance since the links with both partners are characterized by effective communication and sufficient levels of new variety. Any analysis using the averaged similarity index treats both firms **A** and **B** identically and is therefore unable to discriminate between the two very different cases.

Boschma (2005) also pointed out that the proximity types are related to each other in two ways. Firstly, proximities can be substitutes rather then complements when it comes to the establishment of links and their success. It can be sufficient to be proximate in just one dimension to link to an organization: being proximate in another dimension does not yield further effects (Boschma and Frenken, 2010). This argument is backed by empirical findings. For instance, Singh (2005) showed that geographical proximity fosters the establishment of links among researchers that are distant in the cognitive dimension. Ponds et al. (2007) found that geographical proximity helps to overcome institutional distance. In their study, industry-university relations are more likely if firms and universities are geographically proximate, while university-university relations span greater geographical distances.

Secondly, the proximities can be related when it comes to 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 being geographically close allows for more efficient communication even when the two partners might be cognitively distant because of the better possibilities to interact face-to-face. 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.

Cognitive proximity though has an exceptional position among the other proximity types. Similar to the other types, cognitive proximity facilitates knowledge exchange, but it also defines the learning potential, i.e. the potential for the creation of novel ideas and solutions. In that respect, some cognitive distance is needed for 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 on communication efficiency and the success of knowledge exchange (Boschma, 2005). We argue therefore that the effects of these proximities on innovative success have to be seen in relation to that of 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. More precisely, we test if particular geographical or cognitive structures in firms' ego-networks are more conducive for their innovation performance than others. Subsequently, these structures are confronted with information on combined cognitive and geographical structures and we investigate which of these structures enhance firms' innovation performance.

3. Data

Our empirical study is based on a database collected for an 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. Questions concerned R&D activities of firms, their innovative success, and their engagement in knowledge exchange activities. The latter included detailed information on organizations they have been in contact with during the last three years, and with which they exchanged technological knowledge relevant for their innovation activities. This information is used to construct their ego-networks concerning the sourcing of technological knowledge.

Industry	Region	Country	Firms
Automotive	Southwest Saxony	Germany	58
Aerospace	Netherlands	Netherlands	71
Biotechnology	North Rhine-Wesphalia	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
Total			372

Table 1: Overview included cases

As shown in Table 1, data was collected on 372 firms in 8 European regions and 9 industries. 295 firms reported at least one link concerning the exchange of 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 4 links or more. Only 8 percent of the 295 firms mentioned just one organization with which they exchanged technological knowledge. The maximum number of links is 36.

In this study, for two proximity types (cognitive and geographical), we will group firms that have similar structures in ego-networks concerning the respective proximity. Our data restricts us to these two 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 between 1 and 5 for each link. For all 2,110 links, on average, a 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 firm and contact organization. The mean geographical distance is fairly 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.

4. Empirical Approach

4.1 Defining the structure of ego-networks

As pointed out before, studies analyzing the effects of proximities on innovation often rely on the average geographical and cognitive distance of a focal firm to its ego-network. However, the average gives only limited insight into the structure of ego-networks as the mean can be identical for very different empirical distributions.

To obtain a picture of the cognitive composition of the ego-network of a firm, 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 to organizations whose knowledge base is similar to that of the focal firm. These shares are derived from the 1 to 5 scale similarity variable as follows: values equal to 1 are defined as low technological similarity (SIM_LOW), values from 2 to 3 are characterized as medium similarity relations (SIM_MEDIUM), and values of 4 and 5 refer to high similarity (SIM_HIGH). This classification is based on own experiences during interviews and re-samples the distribution of the original values quite well, as shown in Table 2. We then grouped the firms according to the similarity in their profiles by applying a cluster analysis to the three shares.

In a similar fashion, we analyze the geographical dimension. We are particularly interested in firms' embeddedness in regional knowledge networks, which refer to links to firms at low geographical distances. This implies we need to differentiate between regional and non-regional linkages. The data covers firms 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 kilometer, while this is different in Sweden. Accordingly, the size of regions is likely to vary substantially, which makes the definition of an 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 is visual at around 40 km and a flattening of the curve from 600 km onward. Based on this, three variables were created: (1) the share of links with distances equal or less than 40 km (REGIONAL); (2) the share of linkages with distances between 40 and 600 km (MEDIUM); and (3) the share of knowledge relations exceeding a distance of 600 km (LARGE). As shown in Table 2, these shares also resemble the distribution of geographical distances quite well. Clusters for these three variables are identified with a cluster analysis.

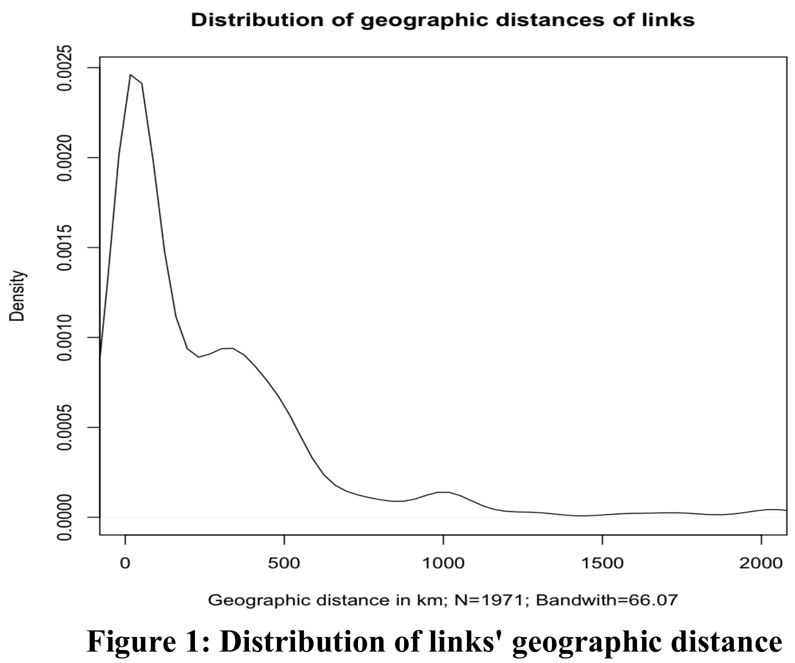

Figure 1: Distribution of links' geographic distance

Above, we treated the cognitive and geographical structure of firms' ego-networks independently of each other. However, this does not take into account the case of geographical and cognitive proximity being related to each other, which refers to the idea that the two might have a complementary or substitutive relationship. To consider this empirically, the two dimensions need to be interacted. In order to keep the complexity at an acceptable level, we decided to drop the differentiation between 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 (REG_LOW); (2) the share of regional links with medium technological similarity (REG_MED), (3) the share of regional links with high technological similarity (REG_HIGH). The same variables were defined for non-regional links: NONREG_LOW, NONREG_MED and NONREG_HIGH. As before, the six variables are entered into a cluster analysis to group firms with similar geographical and cognitive ego-network structures.

Variables	**Mean**
SIM_LOW	0.21
SIM_MEDIUM	0.47
SIM_HIGH	0.32
REGIONAL	0.36
MEDIUM	0.45
LARGE	0.19
REG_SIM_LOW	0.07
REG_SIM_MED	0.17
REG_SIM_HIGH	0.12
NON_REG_SIM_LOW	0.13
NON_REG_SIM_MED	0.30
NON_REG_SIM_HIGH	0.20

Table 2: Mean values basic variables

4.2 Cluster analysis

First, we grouped the firms according to the similarity in their profiles by applying a cluster analysis to their shares on the two proximity dimensions. Many approaches exist to cluster observations (Kaufman & Rousseeuw, 1990). We used the *pdfCluster* method (Bin & Risso, 2011), which employs nonparametric model-based clustering techniques, proposed by Azzalini & Torelli (2007). Advantages of this method are that the "correct" number of clusters is derived from the data and that it also works well in situations of high dimensionality (i.e. few observations but many variables).

4.3 Testing the effects on innovation performance of firms

As 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).[1] In addition to the geographical and cognitive composition of the ego-networks of firms, the following firm characteristics are measured that might also impact on firms' innovative success.

First, we accounted for the absorptive capacity of firms, which might positively impact on their innovation performance. This has been approximated by the share of R&D employees in total employment (RD_INT) and the share of employees with at least a university bachelor degree (SKILL). Since our dependent variable is measured as a share of turnover, we have ensured that these independent variables are also independent of firm size. Second, we have included firm's age (AGE) and the number of employees to account for any firm size effects (EMPL). Third, we account for some industry-specifies by including the share of engineers (ENG) on total employment. Fourth, 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.[2] Fifth, the total number of links regarding the exchange of market knowledge (LINKS_MA) was included to control for differences in firms' general collaboration behavior and for potential innovation stimulating effects resulting from superior market information. Sixth, we included the absolute number of links for exchanging technological knowledge (LINKS_TEC). This variable controls for any effects related to the size of firms' ego-networks, i.e. their propensity to engage in technological knowledge exchange. We also included 8 industry dummies (AUTO, BIO, AERO, TEXT, MEDIA, SOFT, VIDEO, FOOD) to control for industry specifies. As we have data on biotech firms in three different countries, we also included 2 sector-country dummies (SWED_BIO, CZECH_BIO) to control for potential country effects in this sector.

---

[1] This information was collected with the following question: "Please indicate how much (in percent) of the turnover of your firm is attributed to new, dramatically improved products / services introduced in the last three years?"

[2] 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%)".

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

5. Results

5.1 Clustering the composition of firms ego-networks

5.1.1 Cognitive composition

Applying the *pdfCluster* method to the three shares representing the cognitive dimension (SIM_LOW, SIM_MEDIUM, SIM_HIGH) resulted in the identification of 6 clusters, whose characteristics are presented in Table 3. All 6 clusters summarize at least 30 observations. Cluster 1, 2 and 4 clearly represent firms with primarily links to technologically similar (2: HIGH), technologically related (1: REL), and technologically unrelated (4: LOW) organizations. The other three clusters rather represent firms with many related and some highly similar links (3: REL_HIGH), firms with many highly similar and some related links (5: HIGH_REL) and firms with equal shares of links with low and medium similarity (6: LOW_REL).

In our example, average similarity (Mean SIM) seems to be a fairly reliable approximation of the distribution as it clearly allows differentiating between firms with primarily low similarity links (LOW), medium similarity (REL), and high

[3] We also ran models specifying a “log“ and a “probit“ link function but the results remained stable.

similarity (HIGH) links. In the case of the two clusters REL_HIGH and HIGH_REL, the cluster analysis offers richer information on the differences in the cognitive composition of the ego-networks than the mean value of SIM. Table 3 highlights that there are no software-producing firms in Cluster REL and firms of the textile and video gaming industry are barely present in clusters HIGH and LOW. Cluster REL_HIGH tends to contain larger organizations (Mean EMPL) with more links (Mean LINK_TEC). Firms with mainly connections to very similar organizations (HIGH) tend to have relatively small number of links.

Cluster	**1**	**2**	**3**	**4**	**5**	**6**
Size	70	36	71	36	50	32
SIM_LOW	0.02	0.01	0.07	0.93	0.11	0.46
SIM_MEDIUM	0.95	0.01	0.61	0.03	0.27	0.47
SIM_HIGH	0.02	0.99	0.33	0.04	0.62	0.07
NAME	**REL**	**HIGH**	**REL_HIGH**	**LOW**	**HIGH_REL**	**LOW_REL**
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 SOFT	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

Table 3: Cluster: Cognitive dimension

5.1.2 Geographical composition

The results of the cluster analysis for the geographical dimension suggest the existence of 5 clearly defined geographical compositions of ego-networks, as shown in Table 4. The smallest cluster summarizes 18 firms and the largest 88. Similar to the cognitive dimension, clusters 1, 2 and 4 represent the “extremes” summarizing firms that have primarily medium-range linkages (1: MED), regional linkages (2: REG) or large-distance linkages (4: LARGE). Cluster 5 (5: GEO_MIX) includes firms that maintain a fairly equal share of linkages to organizations at all three spatial levels. In contrast, Cluster (3: REG_MED) represents firms that miss large-distance linkages.

Clusters	**1**	**2**	**3**	**4**	**5**
Size	69	53	88	18	67
REG	0.04	0.94	0.44	0.03	0.24
MED	0.91	0.04	0.52	0.04	0.32
LARGE	0.05	0.02	0.05	0.94	0.44
	MED	**REG**	**REG_MED**	**LARGE**	**GEO_MIX**
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 SOFT	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	4447	1887

Table 4: Clusters: geographical dimension

The average geographical distance measure (Mean DIST) differentiates well between the "extreme" cases REG, MED and LARGE. However, this measure is less reliable for discriminating between these "extreme" cases and those of REG_MED and GEO_MIX. In particular, the latter one is likely to be wrongly classified when just looking at the average geographical distance. While it obtains the second highest average distance (1,887 km), the largest shares of links of its firms (56 percent) is still shorter than 600 km.

The clusters REG and LARGE have on average fewer linkages concerning technological knowledge. The cluster REG has on average 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).

5.1.3 Combining cognitive and geographical composition

Now we turn to the question 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

Clusters	1	2	3	4	5	6	7
Size	97	22	46	15	58	40	17
REG_LOW	0.03	0.00	0.10	0.01	0.02	0.03	0.70
REG_MED	0.04	0.02	0.02	0.96	0.42	0.10	0.15
REG_HIGH	0.03	0.02	0.22	0.00	0.09	0.39	0.06
NONREG_LOW	0.06	0.02	0.55	0.01	0.06	0.08	0.06
NONREG_MED	0.66	0.05	0.07	0.02	0.28	0.09	0.01
NONREG_HIGH	0.18	0.89	0.03	0.00	0.14	0.32	0.03
	NONREG_REL	**NONREG_HIGH**	**REG_HIGH_NONREG_LOW**	**REG_REL**	**COM_REL**	**COM_HIGH**	**REG_LOW**
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 SOFT	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

Table 5: Clusters: cognitive and geographical dimension

variables: share of regional (REG) and non-regional links (NONREG). These are interacted with the three variables on technological similarity (SIM_LOW, SIM_MEDIUM, SIM_HIGH). This results in six variables that serve as inputs for the cluster analysis.

As shown in Table 5, the analysis identifies 7 distinct clusters, which vary in size between 15 and 97 firms. Cluster 5 (COM_REL) and cluster 6 (COM_HIGH) resemble primarily cognitive structures in firms' ego-networks. Two clusters represent firms with primarily non-regional linkages to organizations with either medium similarity (Cluster 1, NONREG_REL) or high similarity (Cluster 2, NONREG_HIGH). Two other clusters consist of firms with mostly regional relations to organizations that are either technologically dissimilar (Cluster 6, REG_LOW) or technologically related (Cluster 4, REG_REL). Lastly, there is one "mixed" cluster (cluster 3, REG_HIGH_NONREG_LOW) that summarizes firms with complex ego-networks: firms have large shares of regional relations with technological similar firms and large shares of non-regional links to dissimilar organizations.

5.1.4 Substitution versus complementarity

The identification of these clusters is interesting insofar as they allow inference about a substitutive or complementary relationship between the cognitive and geographical dimension. According to the idea of a substitutive relationship, geographical proximity allows for easier communication and overcoming cognitive distance between firms. The existence of cluster REG_LOW corresponds precisely to this idea as low technological similarity is accessed at close geographical distance. Cluster NONREG_HIGH describes the opposite case in which large geographical distances are overcome when the knowledge bases of two organizations are technologically similar (low cognitive distance). We did not find a particular geographical pattern for the exchange of related knowledge though: organizations interact with organizations having related knowledge at close geographical distance (REG_REL), over larger distances (NONREG_REL) or at a mixture of geographical distances (COM_REL).

However, there are two clusters that confute this idea of substitution. They rather suggest that geographical proximity stimulates the interaction with technologically similar organizations. Cluster COM_HIGH describes firms that link with technologically similar organizations at the regional as well as at the non-regional level. While the second is in line with the above interpretation, the first is clearly orthogonal to the existence of a substitutive relationship between the two dimensions. This is even more so for cluster REG_HIGH_NONREG_LOW. Here firms interact at close geographical distance with technologically similar organizations and simultaneously at high geographical distance with dissimilar organizations. Accordingly, these two clusters rather hint at a complementary relationship between the two proximity dimensions.

When looking at the number of firms assigned to each cluster, we found that 39 firms fully back the substitution hypothesis (part of clusters NONREG_HIGH and REG_LOW), while 46 firms (cluster REG_HIGH_NONREG_LOW) rather suggest a complementary relationship. Accordingly, the cluster analysis suggests that both types of relationships between geographical and cognitive proximity seem to exist.

We explored this further by identifying firms' characteristics that might explain this difference in the relationship between the two proximity dimensions. This is achieved by a multinomial regression analysis. We chose all clusters that are unrelated to this issue as reference group, i.e. all clusters that concern relations between organizations with medium levels of technological similarity (NONREG_REL, REG_REL, COM_REL). We also aggregated the two clusters that represent the substitution hypothesis (REG_LOW, NONREG_HIGH) because of the low number of firms in REG_LOW. The two clusters that rather support the complementarity hypothesis (REG_HIGH_NONREG_LOW and COM_HIGH) are kept separate, as COM_HIGH is not a fully clear case of complementarity.

The results of the analysis are represented in Table 6. We refrain from discussing the coefficients for the sector dummies. The findings indicate that firms in cluster NONREG_HIGH & REG_LOW 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 to maintain successful collaborations with technologically dissimilar organizations. And when cognitive proximity eases the communication, younger firms also interact with more geographically distant partners. Moreover, a substitutive relationship between the cognitive and geographical dimension also tends to be more relevant for firms with low numbers of links concerning technological knowledge (LINKS_TEC). A reason for this might be that firms that are more reluctant to collaboration (having fewer links) are more careful in selecting partners. They may choose geographically proximate organizations because they can monitor them more easily. Moreover, existing social relations, which are more likely to exist between geographically proximate partners, might imply higher levels of trust. This in turn may reduce the perceived obstacles to exchange knowledge between cognitively distant partners.

Clusters	NONREG_HIGH & REG_LOW	REG_HIGH_ NON_REG_LOW	COM_HIGH	NONREG_REL & REG_REL & COM_REL
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*	
SOFT	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 Likeli	-289.879			
Pseudo R2	0.1424			

Table 6: Multinomial regression clusters

In contrast, the negative coefficient of EMPL for REG_HIGH_NON_REG_LOW 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 into regional clusters of technologically similar firms. At the same time, they maintain a large share of relations to dissimilar firms at larger distances. Apparently, this should make them gatekeepers that provide access to knowledge variety in distant places. 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 (2010). We do not observe statistically significant differences between firms of cluster COM_HIGH and our control group.

So far, our analysis has demonstrated that systematic differences exist in the structures of firms' ego-networks along the geographical and cognitive dimension. Next, we turn towards the question whether these differences in the structures of the ego-network of firms do really matter for the innovation performance of firms.

5.2 Impact on innovation performance of firms

We used the previously created cluster dummies to approximate differences in firms' ego-networks. Using the beta-regression approach, these cluster dummies and other firm characteristics are regressed on the share of turnover of a firm that is attributed to new or significantly improved products not older than 3 years (INN). The results are shown in Table 7. We do 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 test each proximity dimension (cognitive and geographical) separately before the combined clusters (cognitive & geographical) are considered.

Before discussing the results for the cluster dummies we take a look at the firm characteristics that gain significance in all model specifications (I – VII). In all models, the variable SKILL is positive and significant 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 positive correlated to SKILL ($+0.33^{***}$) and INNO ($+0.13^{**}$). It is therefore likely the case that SKILL captures the positive effect of RD_INT. The remaining variance of RD_INT that is not captured by SKILL might therefore reflect some sector specifics (e.g. the negative correlation between TEXT and INNO).[4] 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 the literature (see, e.g., 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 & geographic dimension. As they become significant in those models including all cluster dummies

[4] It might however also account for a potential negative size effect because when excluding EMPL from the regression, RD_INT loses its significance.

of the combined cognitive & geographic dimension, they probably capture some effects of the according clusters. However, given the weakness of their effects we refrain from disentangling this any further.

- Table 7 here -

Now, we can turn towards 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 in firms' ego-networks. We expected firms in clusters HIGH, LOW, LOW_REL and HIGH_REL to be outperformed by firms in clusters REL and REL_HIGH because only these latter two have access to related knowledge, i.e. contacts that offer sufficient knowledge variety but still allow for effective communication. Only one dummy (REL_HIGH) is positive and significant. Meeting our expectations, having a mixture of links to related and technologically similar firms is conducive for innovation performance of firms. However, we also expected REL 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 that firms with a mix of linkages at different spatial levels (GEO_MIX and REG_MED) to outperform firms that focus mainly on only one type of spatial level for interaction (REG, MED, LARGE). Only GEO_MIX is positive and significant. This confirms our hypothesis that a balance of links to all three spatial levels facilitates the innovation performance of firms.

In Model IV, the cluster dummies for the combined cognitive & geographical proximity dimensions are tested. For these our expectations are a combination of the above. Three of these have positive and significant effects (NONREG_HIGH, COM_REL, and COM_HIGH). It suggests that innovation performance is particularly stimulated when firms have large shares of linkages to technologically similar organizations outside their region (NONREG_HIGH) and at different spatial levels (COM_HIGH) and when firms have a large share of linkages to technologically related organizations at various spatial scales (COM_REL). What is striking is that no cluster dummies are significant with primarily regional linkages.

Obviously, the geographical dimension plays only a minor role for the definition of these clusters. For this reason, we confront the dummies of the combined cognitive & geographical proximity dimensions with the two dummies that were found significant in the previous models (REL_HIGH, GEO_MIX). The results are shown in model V: NONREG_HIGH, COM_REL and COM_HIGH loose their significance. Their effect is captured by REL_HIGH. However, the previously significant GEO_MIX also looses its significance. It regains significance when COM_HIGH is removed (model VI). Although the overlap between the two clusters COM_HIGH and GEO_MIX is small they seem to relate to the same positive impact on innovation. To disentangle their effects, we excluded all insignificant dummy variables and estimate model VII that includes only REL_HIGH, COM_HIGH and GEO_MIX. REL_HIGH remains highly significant, COM_HIGH is insignificant and GEO_MIX is positive and significant. Accordingly, 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-practice geographical ego-network structure (with a mixture of regional and non-regional linkages) is an independent alternative to the best-practice cognitive ego-network structure (with a mixture of cognitively similar and related organizations).

6. Conclusions

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

Firstly, we showed that the structure of firms' ego-networks differs systematically along the cognitive and geographical dimension. With respect to the cognitive dimension, we found large numbers of firms with relations primarily to one type of organization (technologically dissimilar, related or similar). Nevertheless, somewhat more than 50 percent of firms that exchange knowledge link to organizations with varying degrees of knowledge similarity. For the geographical dimension, similar findings were obtained with about 53 percent of firms having links to organizations at various spatial scales. We also found the two proximity dimensions to be related. For younger firms and those with fewer technological knowledge links, geographical proximity helps to overcome cognitive distance. In contrast, geographical and cognitive proximity are complementary for smaller firms (but not younger firms), meaning that these firms tend to be well embedded into regional clusters of technologically similar firms. At the same time, they maintain a large share of relations to dissimilar firms at larger geographical distances.

Secondly, we tested the effects of these different ego-network structures on innovation. The results clearly showed that firms can increase their innovation performance by primarily linking to technologically related firms as well as technologically similar organizations. It also helped firms to access knowledge sources at various geographical scales. In our study, we did not find an indication for cognitive and geographical proximity being related (either in a substitutive and complementary sense) in their effects on firms' innovation performance.

This study has a number of shortcomings. We measured only links that firms actually realized, not the interaction possibilities firms had. If, for instance, no technologically dissimilar firms are present in a region, firms cannot interact with such firms at this spatial level. Accordingly, they might be driven to search for such collaboration partners at larger geographical distances. Not considering 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 dimension 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. We therefore see clearly 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 formation of ego-networks of firms, whether the proximity dimensions in ego-networks of firms act as substitutes or complements, and how the different structures of ego-networks affect the innovation performance of firms.

Appendix

	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

Table A1: Descriptives

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
INNO	**(1)**												
AGE	**(2)**	-0.18***											
EMPL	**(3)**	-0.05	0										
SKILL	**(4)**	0.27***	-0.29***	0.01									
ENG	**(5)**	0.05	0.04	-0.06	-0.21***								
RD_INT	**(6)**	0.13**	-0.24***	-0.07	0.33***	-0.25***							
EXTERN	**(7)**	-0.12**	-0.03	0.03	-0.01	-0.04	0.06						
AUTO	**(8)**	-0.05	0.15***	-0.01	-0.38***	0.33***	-0.06	-0.06					
BIO	**(9)**	0.17***	-0.18***	-0.04	0.48***	-0.39***	0.35***	0.05	-0.21***				
AERO	**(10)**	-0.07	0.1*	-0.02	-0.13***	0.31***	-0.19***	0.03	-0.21***	-0.23***			
TEXT	**(11)**	-0.18***	-0.04	0.01	-0.28***	-0.33***	-0.12**	0.09	-0.13**	-0.14***	-0.15***		
MEDIA	**(12)**	-0.03	-0.13**	0.14***	0.2***	-0.27***	0.22***	-0.04	-0.14***	-0.16***	-0.16***	-0.1*	
SOFT	**(13)**	0.06	-0.09*	-0.02	0.19***	0.19***	-0.07	0.01	-0.12**	-0.14***	-0.14***	-0.09*	-0.1*
VIDEO	**(14)**	0.23***	-0.12**	-0.02	0.1*	0.09*	0.16***	-0.11**	-0.1**	-0.11**	-0.12**	-0.07	-0.08
FOOD	**(15)**	-0.11**	0.36***	0	-0.12**	-0.23***	-0.11**	-0.04	-0.12**	-0.14***	-0.14***	-0.09*	-0.09*
CZECH_BIO	**(16)**	0.1*	-0.07	-0.02	0.17***	-0.24***	-0.09*	0.03	-0.09*	0.44***	-0.1**	-0.06	-0.07
SWED_BIO	**(17)**	0.06	-0.12**	-0.03	0.39***	-0.23***	0.32***	0.14***	-0.13**	0.62***	-0.14***	-0.09*	-0.1*
LINKS_TEC	**(18)**	0.06	0.03	-0.02	-0.05	0.07	0.35***	0.02	0.43***	0.12**	-0.16***	-0.1*	-0.07
LINKS_MA	**(19)**	0.09*	0	0.18***	0.06	0.05	0.2***	-0.07	0.25***	-0.08	-0.19***	-0.11**	0.21***

		(13)	(14)	(15)	(16)	(17)	(18)
VIDEO	**(14)**	-0.07					
FOOD	**(15)**	-0.08	-0.07				
CZECH_BIO	**(16)**	-0.06	-0.05	-0.06			
SWED_BIO	**(17)**	-0.09*	-0.07	-0.08	-0.06		
LINKS_TEC	**(18)**	-0.07	0.09*	-0.07	-0.15***	0.11**	
LINKS_MA	**(19)**	0.02	0.14***	-0.05	-0.16***	0.02	0.67***

Table A2: Correlations

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
FOOD	-0.014	0.9724	-0.169	0.692	-0.205	0.631	-0.213	0.616	-0.216	0.609	-0.052	0.895	-0.156	0.690
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*	0.083
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^2			0.196		0.183		0.197		0.212		0.207		0.205	

Table 7: Impact on 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 collaborative networks: Evidence from R&D projects within the GNSS industry. Regional Studies, IFirst, 27.1.2011, DOI: 10.1080/00343404.2010.529121.

Balland, P.-A., Suire, R., & Vicente, J. (2010). How do clusters/pipelines and core/periphery structures work together in knowledge processes? Evidence from the European GNSS technological field. Papers in Evolutionary Economic Geography, 10.08.

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.

Breschi, S. & Lissoni, F. (2003). Mobility and social networks: Localised knowledge spillovers revisited. CESPRI Working Paper, No. 142.

Broekel, T. & Boschma, R. (2011). Knowledge networks in the Dutch aviation industry - The proximity paradox. Journal of Economic Geography, doi:10.1093/jeg/lbr010.

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.

Frenkel, A. & Schefer, D. (1998). Local milieu and innovations: Some empirical results. The Annals of Regional Science, 32, 185-200.

Gluckler, J. (2007). Economic geography and the evolution of networks. Journal of Economic Geography, 7(5), 619-634.

Graf, H. (2010). Gatekeepers in regional networks of innovation. Cambridge Journal of Economics, online access (doi:10.1093/cje/beq001).

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.

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, 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).

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. and Boschma, R. (2009). Applying social network analysis in Economic Geography: Framing some key analytical issues. Annals of Regional Science, 43:739–756.

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.