Titze, Mirko; Brachert, Matthias

Conference Paper

Determinants for the partners' choice within the 7th EU Framework Programmec

54th Congress of the European Regional Science Association: "Regional development & globalisation: Best practices", 26-29 August 2014, St. Petersburg, Russia

Provided in Cooperation with:
European Regional Science Association (ERSA)

Determinants for the partners' choice within the 7th EU Framework Programme[c]

by Mirko Titze[ab] and Matthias Brachert[a]

1st draft, please do not quote without the authors' permission.

Abstract:

The EU Framework Programme (FP) belongs to the most important instruments promoting transnational collaborative R&D projects in Europe. Its main objective is to initiate cross-border complementarities in order to exploit knowledge resources and to conduct large scale research. Within the EU FPs the applicants are free to choose partners from all over Europe. The key research task of our paper is to identify the determinants of the formation of dyadic intra- and interregional collaborations within EU Framework Programmes. To explain partner selection, we adopt a cluster as well as proximity perspective including the cognitive, spatial and institutional dimensions. The analysis is conducted on EU NUTS 2 level. We focus on two technology fields, biotechnology and aerospace. In doing so, we are able to capture general and technology specific characteristics. We apply a gravity type spatial interaction modelling framework. The empirical analysis is carried out using a negative binomial specification. We find evidence that geographical factors matter – but that cognitive proximity is more importantly for link formation. Institutional proximity is especially important from a common language or language family perspective. Moreover, industrial clusters positively affect the setting up of links. However, we also prove that the mere size in terms of employment or establishments is not necessarily required to establish cross-region collaborations. When linked to industrial clusters, also small actors can form links within the framework programmes.

Keywords: European Cluster Networks, R&D Collaborations, EU Framework Programme, Spatial Interaction Modelling

JEL classifications: D85, L14, R12, R15

[a] Halle Institute for Economic Research, Halle, Germany

[b] Corresponding author: mirko.titze@iwh-halle.de

[c] This research is part of the GRINCOH project that is funded within the 7th EU Framework Programme. The acronym GRINCOH stands for Growth-Innovation-Competitiveness Fostering Cohesion in Central and Eastern Europe (http://www.grincoh.eu/).

1 Introduction

The EU Framework Programme (FP) belongs to the most important instruments promoting transnational collaborative R&D projects in Europe. Its main objective is to initiate cross-border complementarities in order to exploit knowledge resources and to conduct large-scale research. The idea of this programme is in line with recent insights of knowledge production. Several scholars highlight that knowledge production is increasingly performed in a collaborative way (Meyer and Bhattacharya 2004; Hoekman et al. 2009). Furthermore, collaborative knowledge production is seen to be more efficient in terms of the number of patent citations as well as publications (Frenken et al. 2005, Hoekman et al. 2009).

This knowledge production can be analysed in the light of a regional, national and European perspective. Our study adopts a European inter-regional perspective. Using data from the 7th EU framework programme (FP), we try to identify the role of proximity and of industrial clusters on link formation in EU collaborations. Proximity influences the probability of interaction through bringing down perceived costs of collaboration (Boschma 2005). Regarding the importance of proximity for partner selection, an inter-regional EU perspective allows us to draw attention on the relevance of space for partner choice. In doing so, we are able to distinguish whether spatial proximity as well as institutional proximity promotes R&D collaboration. Spatial proximity is thought to affect collaboration intensity though increasing collaboration costs with distance (Hoekman et al. 2009). Institutional proximity exerts an effect on collaboration intensity via increasing collaboration costs when being confronted with different institutional frameworks (Edquist and Johnson 1997). Furthermore, we make use of Pre-Framework information on regional technology profiles measured by patents to test to what extent cognitive proximity affects link formation. Some, but not too much cognitive distance is seen as a necessary condition to ensure common understanding but to enable learning within collaboration (Boschma and Frenken 2011). The same argument holds true for industrial clusters that are furthermore expected to hold special benefits for collaboration in terms of research quality both from a private (centres of excellence) as public perspective (specialized public infrastructure). The key research task of our paper is to identify the determinants of the formation of dyadic intra- and interregional collaborations within EU FPs.

The paper is organized as follows. The next chapter presents some more in depth discussion on determinants of partner choice in R&D collaborations. Chapter 3 introduces the data and methods used in the empirical framework. Chapter 4 discusses the regression results. Chapter 5 concludes the paper and gives an outlook on further fields of research.

2 The role of proximities and clusters on partner choice in R&D collaborations

Today, considerable parts of knowledge generation and application base upon interactive learning processes. Reasons for this can be dedicated to the increasing technological complexity of modern products and services. This comes in line with increasing difficulties for organizations to hold all resources needed to sustain their competitive advantages. In this context R&D collaboration allows enhancing organization's performance (Cohen and Levinthal 1990). They give access to partner resources and allow benefits from collective and organizational learning.

Malmberg and Maskell (2006) point out that the exchange of tacit knowledge plays a crucial role for the adaption and development of new technologies. However, collaboration configurations differ with respected to perceived risk and costs. Boschma (2005) therefore argues that different types of proximities are of special relevance in this context and are able to give explanations for differences in the probability of interaction.

Boschma (2005) for example claims that geographical proximity is neither a necessary nor a sufficient condition for interactive learning processes. Furthermore, geographical proximity strengthens other forms of proximity, e. g. cognitive, organizational, social, institutional proximity. Though, the different forms of proximity do not necessarily show simple linear correlation. Not only too little, but also too much proximity may be harmful for interactive learning processes. With respect to Boschma's (2005) proximity concept we consider specific issues of his approach. We focus on specific forms of proximity in order to capture their impact on interactive learning processes in isolation. We place emphasis on geographical, institutional and technological proximity as they have been detected as crucial in the Scherngell and Barber (2009) study.

Industrial clusters embrace various forms of proximity simultaneously (e. g. cognitive, organizational, social, institutional issues). They provide a critical mass of activities in a certain knowledge domain and are therefore seen as conducive to innovation activities (Porter 1998). A bulk of literature has emerged regarding the nature and the impact of different forms of knowledge interaction with respect to the presence of industrial clusters. Bathelt et al. (2004) placed emphasis on the co-existence of two forms of knowledge interactions being relevant for clusters: local and global communication. Both modes are possible channels for the exchange of tacit knowledge. Local interactions (buzz) concern learning processes between locally embedded actors. Furthermore, knowledge-based flows (pipelines) also occur from and to selected actors outside the region. The authors argue that the interplay of local buzz and global pipelines provide advantages for embedded actors that are not available to outsiders.

Bathelt and Turi (2011) point out that the exchange of tacit knowledge does not necessarily require co-location and real face-to-face interaction. Moreover, the use of novel information and communication technologies enables communication over large distances. Temporary and virtual interaction, which is given in collaborative R&D projects, is recognized as the basis for establishing worldwide production and innovation linkages. Dependent on the circumstances a specific combination of modes of interactions may be advantageous in some production and innovation contexts but not in others. Bathelt and Li (2013) furthermore argue that clusters also shape the emergence of trans-local (interregional) connections. Against this backdrop clusters are more likely to build connections to other industrial clusters to keep up with wider industrial developments and to keep connections to other centres of excellence in production and knowledge generation. Although they prove this argument in the context of FDI data, the same arguments are likely to hold for R&D collaborations.

In our paper we focus on R&D collaboration networks. Our work relies on an approach developed by Scherngell and Barber (2009). They studied determinants for the number of cross-region collaborations using data on collaborative R&D projects granted within the 5th EU Framework Programme (FP).

We tie in with this research and provide an extended approach that considers cluster-specific characteristics. In line with Scherngell and Barber (2009) we use data on EU FPs in order to investigate cross-region knowledge interactions. The unit under analysis is the NUTS 2-level. In order to capture technology specific characteristics we differentiate between two technology fields, biotechnology and aerospace. Against this backdrop we formulate our hypotheses as follows:

H1: Clusters are more likely to generate a higher number of cross-region collaborations.

H2: Geographical effects matter.

H3: Technological proximity shows a strong effect on the number of cross-region collaborations.

H4: The effect of cognitive proximity differs across technology fields, because modes of innovations are likely to differ.

3 Data and methods: variable definition and some descriptive statistics

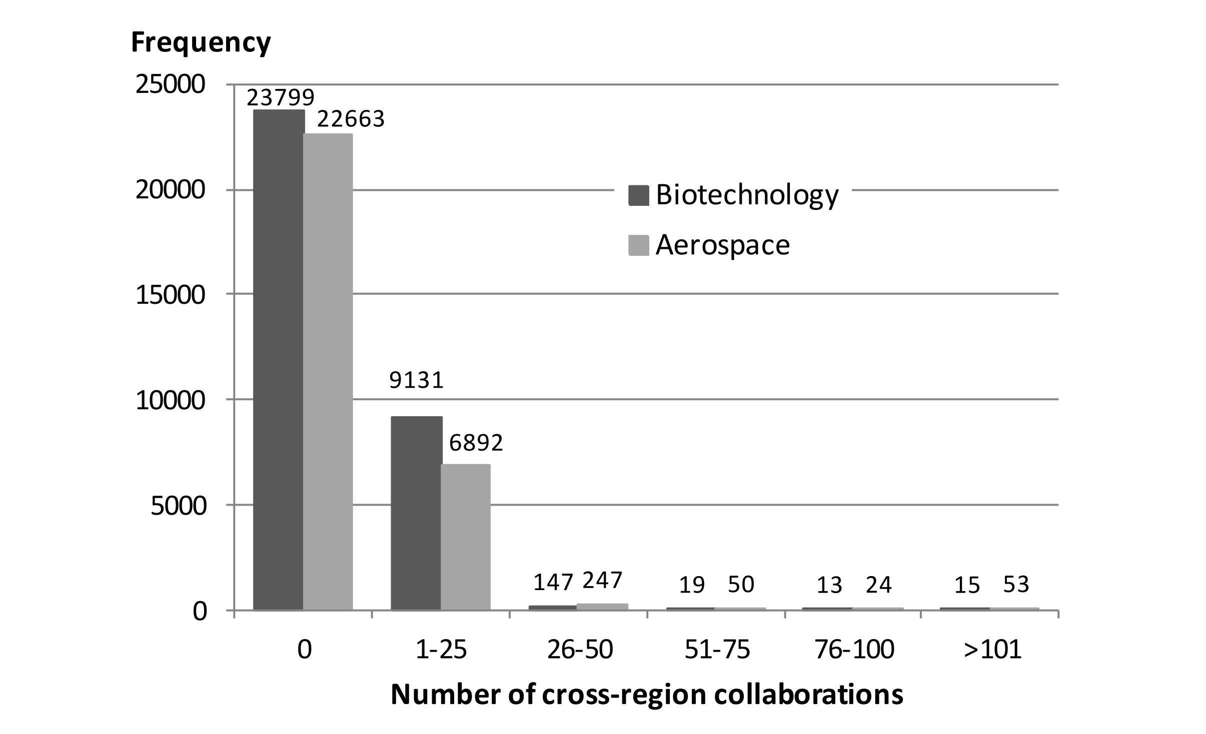
Dependent variable: number of cross-region collaborations

As we made clear in the introduction our paper focuses on the question which determinants propel the emergence of cross-region collaborations. Data on granted collaborative projects are regarded as an appropriate tool to depict cross-region R&D networks (e. g. Scherngell and Barber 2009). We carry out our study using data on collaborative projects that have been funded within the 7th EU FP (as of March 1, 2013). The respective data are provided in the ECORDA and CORDIS databases. Regarding the spatial entities we have chosen the NUTS 2-regions of the EU 27 countries. The French DOM (Guadeloupe, Martinique, Guyane and Réunion), the Canary Islands, the Azores, Madeira and the Spanish exclaves in northern Africa have been excluded. Finland is not included in the dataset due to changes in the NUTS 2 definition. Moreover, we did not consider non EU members, such as Norway and Switzerland. In total, the dataset contains 253 NUTS 2 regions yielding to 253x253= 64009 possible pairs of regions.

With respect to the direction of knowledge interactions we assume bilateral flows. We claim that each partner in a collaborative project participates in the process of knowledge generation and knowledge transfer within the project team. If three partners are involved in one collaborative project (A, B, C) we receive six pairs of actors representing knowledge interactions: AB, BC, AC, BA, CB and CA. Using this basic table we develop a (symmetric) cross-region collaboration matrix by aggregating the project information on regional level.

In order to capture technology specific effects we focus on two fields: biotechnology and aerospace. Figure 1 illustrates the distribution of the number of cross-region collaborations in the two technology fields under analysis.

Figure 1: Frequency of cross-region collaborations



Sources: Raw data, ECORDA and CORDIS; authors' own calculations.

The graph indicates that cross-region collaboration obviously is a rare event. A large number of pairs of regions was not capable to develop at least one collaboration. Only a small number of regions shows a high intensity of collaboration. The graph does not indicate significant differences between the two technology fields. The final biotechnology dataset includes 33,124 observations. The maximum number of collaborations is 458. Regarding the technology field aerospace the number of observations is 29,929, and the maximum number of collaborations in the symmetric cross-region matrix is 1,286. The number of collaborations forms our dependent variable.

Spatial interaction model and estimation procedure

Our empirical model bases on a spatial interaction modeling framework and relies on the approach developed in LeSage et al. (2007), Scherngell and Barber (2009) and Scherngell and Barber (2011). Within this framework they applied a gravity type specification. It combines three types of functions: i) a function describing the origin of the collaboration, ii) a function describing the destination of the collaboration, and iii) a function describing the distance (respectively proximity) between the pair of regions. The origin and destination function (i) and ii)) include weighted origin and destination variables. The separation function (iii) contains a set of variables that are described by an exponential functional form. The estimation procedure follows a count data specification. Particularly in the case of rare events the Poisson distribution is generally accepted to describe discrete processes. Another advantage of this distribution is the consideration of "0" as a natural outcome. As we have shown in the previous section a large number of pairs of regions does not collaborate in EU FP. The application of the Poisson model specification assumes that the variance equals the expected value (equidispersion). Since the model includes a multiregional setting it is very likely that this assumption does not hold true. In this case overdispersion occurs, and unobserved heterogeneity exists which leads to biased estimates. The problem of unobserved heterogeneity can be captured by introducing a stochastic heterogeneity parameter. Then overdispersion is allowed. The implementation of this parameter assumes a binominial distribution of the dependent variable. We take into account this issue and apply the test for overdispersion in our regressions.

Independent variables

Since the paper applies the global cluster networks conception developed by Bathelt and Li (2013) we have to consider regional cluster specific characteristics in the mentioned modeling framework. Cluster characteristics are implemented twofold in our approach. At first, we include the absolute number of *employees* and *establishments* as variables describing the origin respectively destination region (Table 1). Typically, localization economies are measured by using specific indices (e. g. localization coefficient, concentration rate, cluster index etc.; e. g. Aiginger et al. 1999 or Sternberg and Litzenberger 2004). Though, the variety of different measures prevents the accumulation of empirical evidence. Since the focus in cluster research is on the identification of externalities caused by co-location of workers and/or firms the simple measure of the absolute number might be regarded as an appropriate procedure (Frenken et al. 2013).[1] Moreover, we might assume that European NUTS 2 regions have been defined in such a manner that they represent (roughly) comparable spatial entities. Thus, the absolute number of workers/establishments in NUTS 2 regions could represent a rough relative measure. The data stem from the cluster observatory homepage (www.clusterobservatory.eu). Table 1 indicates that in the technology field biotechnology a relatively small number of workers is employed in a relatively large number of establishments. Conversely, in the field of aerospace many employees work in a relatively small number of establishments. Obviously the two technology fields are characterized by different (optimal) firm sizes.

The cluster concept is also reflected through the separation variables *diff_employees* and *diff_establishments*. These variables are derived from *employees* respectively *establishments* and capture size effects between the origin and destination region. We included these measures in order to prove the findings of Bathelt and Li (2013). They found that clusters are more likely to set up connections to other clusters. Conversely, non-clustered organizations are less likely to get connected with clusters.

In line with the theoretical considerations of Boschma (2005) we include further separation variables in our model measuring different forms of proximity (respectively distance). A first category of variables captures geographical effects, such as geographical *distance* and some dummy variables that control for language, country, neighbor and border effects. Geographical distance is measured as physical distance in kilometers between the centroids of two regions. Boschma (2005) pointed out that geographical proximity alone is neither a necessary nor a sufficient condition for interactive learning processes. Nevertheless, geographical proximity facilitates other forms of proximity. In general, geographical proximity brings people together. A large distance is detrimental for intensive face-to-face

[1] Frenken et al. (2013) refer to two studies supporting their hypothesis. For a region in the U. S. a study by Stough et al. (1998) proved evidence that the concentration of technically skilled workers is strongly related to higher growth rates of new firms. Another study developed by Raspe and Van Oort (2008) show for all Dutch establishments that the (absolute) endowment with knowledge resources positively affects firm growth. Frenken et al. (2013) consider these findings as indirect evidence for agglomeration economies.

contacts. However, these contacts enable the exchange of tacit knowledge and therefore interactive learning.

The second group concerns cognitive proximity which is measured through the variable *tech_dist*. Knowledge generation capabilities are not distributed equally in space. In other words, knowledge is allocated between different actors (Antonelli 2000). Cohen and Levinthal (1990) stress, that knowledge transfer is effective if the involved actors possess absorptive capacity. So, these actors are capable to identify, interpret and exploit new knowledge. Boschma (2005, p. 63) regards cognitive proximity as ability that “people sharing the same knowledge base and expertise may learn from each other.” In our approach we use technological distance as proxy for cognitive proximity. According to Scherngell and Barber (2009) technological distance is measured as *1-r*, where *r* represents the Pearson correlation coefficient between two vectors t_i and t_j. The vector t_i is calculated as $x_{ir}/\Sigma_r x_{ir}$. The variable x denotes the number of number of patent applications in the IPC subclass (7 digit level) r in region i. In sum, the vector t represents the share of patent applications a region has in a specific technology field. It reveals information about the region’s technological specialization. The vector t_j is determined similarly. Patent applications are taken from the RegPat database. The assignment of patent subclasses to the technology fields aerospace and biotechnology follows a Eurostat classification scheme.[2] In total, we used 51 subcategories for aerospace and 50 for biotechnology. The *tech_dist* variable ranges from 0 to 2. In the first case regional technological characteristics are identical in the latter they are completely different.

Table 1: Descriptive statistics

Independent variables	Technology field	Observations	Mean	Std. Dev.	Max
Weighted origin/destination variables					
employees	Biotechnology	33,124	274.6	464.9	2,934
	Aerospace	29,929	1,513.8	3,584.6	26,200
establishments	Biotechnology	33,124	27.2	45.3	250
	Aerospace	29,929	7.8	9.6	56
Separation variables					
distance	Biotechnology	33,124	1,351.4	808.7	4,958.5
	Aerospace	29,929	1,268.9	754.5	4,388.7
tech_dist	Biotechnology	33,124	0.6	0.4	1.1
	Aerospace	29,929	1.0	0.2	1.2
diff_employees	Biotechnology	33,124	375.0	540.0	2,933
	Aerospace	29,929	2,390.3	4,470.5	26,199
diff_establishments	Biotechnology	33,124	36.7	52.4	249
	Aerospace	29,929	8.5	10.6	55
dummy variables[a]					

[a] language_identical, neighbour, border_region, country_identical

Sources: Raw data, ECORDA and CORDIS; authors’ own calculation.

[2] http://epp.eurostat.ec.europa.eu/cache/ITY_SDDS/Annexes/pat_esms_an8.pdf.

Boschma (2005) pointed out that non-linearity might occur regarding proximity: not only too little, but also too much proximity might be detrimental for effective interactive learning processes. Noteboom (2000) emphasizes that too little distance lead to a lack of novelty and too much distance means that actors do not understand each other. We considered this issue by introducing a quadratic term for technological and geographical distance.

4 Estimation results

In the previous section we have mentioned that the Poisson modeling framework is an appropriate tool to cover discrete processes, particularly in the case of rare events. Though, we have to test for overdispersion. The respective test is highly significant so that we can reject equidispersion. To sum up, the negative binominal specification fits best in our case (Winkelmann and Boes 2009, pp. 288-290).

In order to avoid multicollinearity we examined the variance inflation factors (VIF) and pairwise correlations between all dependent variables (Appendix 1). All VIF values remain below 10 which is regarded as a critical value (e. g. Ross et al. 2013, p. 18). Though, some variables are highly correlated. This particularly applies to the variables *country_identical* and *language_identical*. This result is not unexpected, but due to theoretical reasons we included both dummies in order to capture effects that occur between different countries speaking the same official language (e. g. Germany and Austria, France and Belgium, Great Britain and Ireland).

Table 2 reports the results of our estimations. Because we are interested in technology specific effects we carried out identical regressions for both technology fields. At first glance, almost all coefficients are highly significant. Regarding the effects of clusters on the number of collaborations we found a positive effect for both technology fields. The coefficients for the number of employees and establishments are significant positive. Nevertheless, the variables measuring size differences (*diff_employees* and *diff_establishments*) - albeit the coefficients are very small. These findings support hypothesis H1 that clusters are more likely to generate a higher number of cross-region collaborations. Though, small actors have been chosen as partners in European R&D networks, too.

Regarding hypothesis H2 that geographical effects matter we found significant coefficients for the physical distance (*distance, distance^2*). The analysis of the marginal effects has shown only a small impact on the number of cross-region collaborations. Moreover, table 2 reports that the number of cross-region collaborations significantly decreases if a partner belongs to the same country. This effect is probably caused by the funding schemes of the EU FPs: the overall collaboration project must include a sufficient number of international actors. Nevertheless, we can observe a significant positive language effect that can be interpreted as institutional proximity. The effects of neighbouring regions and border regions are only significant for the biotechnology field.

The influence of technological proximity (variables *tech_dist* and *tech_dist^2*) is highly significant. Furthermore, we checked marginal effects and proved a strong effect on the number of cross-region collaborations. Insofar, we can confirm hypothesis H3. However, we could observe a considerable difference between the two technology fields: technological distance follows a U-shaped trend in biotechnology indicating that either very little or very much proximity lead to a large number of cross-region collaborations. Conversely, the field of aerospace tells a different story. The inverted U-shaped trend implies that there is a specific distance that corresponds with a maximum number of cross-region collaborations. The latter finding is in line with the proximity paradox described in Boschma (2005): not only too little, but also too much proximity may hamper the number of interactions. We guess that differences in the innovation processes between biotechnology and aerospace are responsible for the different trends regarding technological proximity. Against this backdrop also hypothesis H4 can be confirmed.

Table 2: Estimation results

Dependent variable Number of cross-region collaborations	Biotechnology		Aerospace	
Weighted origin/destination variables				
from_employees (ln)	0.380***	(0.010)	0.142***	(0.011)
to_employees (ln)	0.380***	(0.010)	0.142***	(0.011)
from_establishments (ln)	0.197***	(0.012)	0.279***	(0.024)
to_establishments (ln)	0.197***	(0.012)	0.279***	(0.024)
Separation variables				
1.language_identical	0.350***	(0.059)	0.303***	(0.070)
distance	-0.001***	(0.000)	-0.000**	(0.000)
distance^2	0.000***	(0.000)	0.000***	(0.000)
1.neighbour	0.279**	(0.098)	-0.098	(0.135)
1.border_region	-0.288*	(0.168)	0.581	(0.423)
1.country_identical	-0.584***	(0.077)	-0.325***	(0.095)
tech_dist	-4.097***	(0.165)	2.243***	(0.480)
tech_dist^2	2.246***	(0.135)	-3.177***	(0.283)
diff_employees	0.000***	(0.000)	0.000***	(0.000)
diff_establishments	0.002***	(0.000)	0.014***	(0.002)
_cons	-3.592***	(0.105)	-1.929***	(0.218)
/lnalpha	0.600	(0.024)	1.545	(0.021)
Alpha	1.823	(0.043)	4.688	(0.100)
Number of obs	33124		29929	
Log pseudolikelihood	-33229.951		-31619.33	
Wald chi2	11983.61		5590.48	
Prob > chi2	0.0000		0.0000	

Notes: Values in parantheses = robust standard errors. *** significant at the 0.001 level, ** significant at the 0.05 level, * significant at the 0.10 level.

Sources: Raw data, ECORDA and CORDIS; authors' own calculation.

5 Conclusions and outlook for further research

The main purpose of the paper was the identification of determinants which affect the number of cross-region collaborations. The starting point for our analysis was the observation that considerable parts of economic activity take place on regional level despite dynamic globalisation processes. Industry concentration respectively clusters are caused by co-location of workers and/or firms. A rich body of literature explains co-location through the argument that the transfer of tacit knowledge, which is regarded as a key driver for the development of innovative capacities in a region, is spatially bounded.

Otherwise, there are strands in the literature emphasizing that geographical proximity is neither a necessary nor a sufficient condition for interactive learning processes to take place (Boschma 2005). Moreover, physical proximity may facilitate other forms of proximity, such as cognitive, social, institutional and organizational. Bathelt and Li (2013) proved for FDI as a specific kind of knowledge transfer that multinational firms are more likely to set up new affiliates in other, similarly specialized clusters. In doing so, these clusters are capable to keep up to wider industrial development.

We picked up this idea and investigated the partner's choice in collaborative R&D projects that have been funded within the 7^{th} EU FP. In line with Scherngell and Barber (2009) we apply spatial interaction modelling framework that bases on a gravity model. The estimation has been carried out by using a negative binominial model. We applied an extended version of the Scherngell and Barber (2009) approach since we considered cluster specific characteristics. Moreover we differentiated between two technology fields: biotechnology and aerospace in order to cover technology specific characteristics.

We found that cluster actors are indeed more likely to set up linkages to other actors also located in clusters. Nevertheless, non-cluster actors are involved in interregional R&D networks, too. Moreover, geographical effects matter: while the effects of physical distance are relatively low, identical languages affect the number of cross-region collaborations positively. A relative strong influence can be observed for technological proximity. We found differences regarding the effect technological proximity between the technology fields biotechnology and aerospace. In biotechnology either very little or very much proximity yields to a large number of cross-region collaborations. The contrary holds true for aerospace. Here, an optimal value for technological proximity exists. Too much but also too little proximity are detrimental for cross-region interactions.

Research efforts should address dynamic issues and actor based models. So far, we carried out our analysis on the level of regions. However, the data underlying our analysis in principle allow investigation on the level of actors. In doing so, specific characteristics of the involved actors may be taken into account. An important issue that is pointed out in the proximity debate concerns social proximity. This form of proximity may be captured by the actor's network position (e. g. Maggioni et al. 2011). Also organizational proximity may be analyzed advantageously at the level of actors. However, a bulk of interesting research questions remains and calls for further explication.

Literature

Aiginger, K.; Boeheim, M.; Gugler, K.; Pfaffermayr, M. and Wolfmayr-Schnitzer, Y. (1999): Specialisation and (geographic) concentration of European manufacturing. Background paper for The Competitiveness of European Industry: 1999 Report. http://karl.aiginger.wifo.ac.at/fileadmin/files_aiginger/publications/1999/cr99_back.pdf (accessed July 2010).

Antonelli, C. (2000): Collective knowledge communication and innovation: the evidence of technological districts. In: Regional Studies, vol. 34, pp. 535–547.

Asheim, B.; Gertler, M. (2006): Regional innovation systems and the geographical foundations of innovation. In: Fagerberg, J.; Mowery, D. C.; Nelson, R. R. (2004) (eds): The Oxford handbook of innovation. Oxford University Press, Oxford.

Bathelt, H. and Li, P.-F. (2013): Global cluster networks—foreign direct investment flows from Canada to China. Forthcoming in: Journal of Economic Geography.

Bathelt, H. and Turi, P. (2011) Local, global and virtual buzz: the importance of face-to-face contact and possibilities to go beyond. In: Geoforum, vol. 42, pp. 520–529.

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

Boschma, R. and Frenken, K. (2011): Technological relatedness and regional branching. In: Bathelt, H. M.P. Feldman and D.F. Kogler (eds.): Beyond Territory. Dynamic Geographies of Knowledge Creation, Diffusion and Innovation. Routledge, London and New York.

Boschma, R. A. (2005): Proximity and Innovation: A Critical Assessment. In: Regional Studies, vol. 39, pp. 61–74.

Cohen, W. M. and Levinthal, D. A. (1990): Absorptive capacity: a new perspective on learning an innovation. In: Administrative Science Quarterly, vol. 35, pp. 128–152.

Cooke, P. (2004): Regional innovation systems: An evolutionary approach. In: Cooke, P.; Heidenreich, M.; Braczyk, H. J. (2004) (eds): Regional innovation systems: The role of governances in a globalized world. Routledge, London.

Edquist, C. and Johnson, B. (1997): Institutions and organisations in systems of innovation. In: Edquistm C. (ed): Systems of innovation. Technologies, Institutions and Organizations. Pinter, London.

Frenken, C.; Hölzl, W. and Vor, F. (2005) The citation impact of research collaborations: the case of European biotechnology & applied microbiology (1988-2002). In: Journal of Engineering and Technology Management, vol. 22, pp. 9-30.

Frenken, K.; Cefis, E. and Stam, E. (2013): Industrial dynamics and clusters: a survey. Utrecht School of Economics. Tjalling C. Koopmans Research Institute. Discussion Paper Series 13-11.

Hoekman, J.; Frenken, K. and van Oort, F. (2009): The geography of collaborative knowledge production in Europe. In: Annals of Regional Science, vol. 43, pp. 721-738.

Jaffe, A. B.; Trajtenberg, M. and Henderson, R. (1993): Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. In: The Quarterly Journal of Economics, MIT Press, vol. 108, pp. 577-598.

LeSage, J. P.; Fischer, M. M.; Scherngell, T. (2007): Knowledge spillovers across Europe: Evidence from a Poisson spatial interaction model with spatial effects. In: Papers in Regional Science, vol. 86, pp. 393–422.

Maggioni, M.; Uberti, T. E. and Usai, S. (2011): Treating Patents as Relational Data: Knowledge Transfers and Spillovers across Italian Provinces. In: Industry and Innovation, vol. 18, pp. 39-67.

Malmberg, A. and Maskell, P. (2006): Localized learning revisited. In: Growth and Change, vol. 37, pp. 1–18.

Meyer, M. and Bhattacharya, S. (2004): Commonalities and differences between scholarly and technical collaboration. An exploration of co-invention and co-authorship analyses. In: Scientometrics, vol. 61, pp. 443-456.

Noteboom, B. (2000): Learning and Innovation in Organizations and Economies. Oxford University Press, Oxford.

Raspe, O. and Van Oort, F. G. (2008): Firm growth and localized knowledge externalities. In: Journal of Regional Analysis and Policy, vol. 38, pp. 100–116.

Ross, J.-M.; Perkmann, M. and Fini, R. (2013): Translational research: When do public science projects result in real world impact? Paper to be presented at the 35th DRUID Celebration Conference 2013, Barcelona, Spain, June 17-19.

Scherngell, T. and Barber, M. (2011): Distinct spatial characteristics of industrial and public research collaborations: evidence from the fifth EU Framework Programme. In: The Annals of Regional Science, vol. 46, pp. 247–266.

Scherngell, T. and Barber, M. J. (2009): Spatial interaction modelling of cross-region R&D collaborations: empirical evidence from the 5th EU framework programme. In: Papers in Regional Science, vol. 88, pp. 531-547.

Scott, A. and Storper, M. (2007): Regions, globalization, development. In: Regional Studies, vol. 41, pp. 191–205.

Sternberg; R., and Litzenberger, T. (2004): Regional clusters in Germany—their geography and their relevance for entrepreneurial activities. In: European Planning Studies, vol. 12, pp. 767–791.

Stough, R. R.; Haynes, K. E. and Campbell, H. S. (1998): Small business entrepreneurship in the high technology services sector: An assessment for the edge cities of the U.S. national capital region. In: Small Business Economics, Vol. 10, pp. 61–74.

Winkelmann, R. and Boes, S. (2009): Analysis of Microdata. Berlin: Springer, 2nd edition.

Appendix

Appendix 1: Pairwise correlations and variance inflation factors (VIF)

		1	2	3	4	5	6	7	8	9	10	VIF
	Biotechnology											
1	employees	1,00										0,00
2	diff_employees	0,63	1,00									1,76
3	establishments	0,22	0,10	1,00								1,69
4	diff_establishments	0,10	0,10	0,63	1,00							1,77
5	language_identical	0,05	0,04	-0,03	-0,09	1,00						1,70
6	distance	-0,12	-0,13	0,05	0,08	-0,45	1,00					2,98
7	neighbour	0,00	0,00	0,00	-0,03	0,32	-0,23	1,00				1,38
8	country_identical	0,05	0,02	0,02	-0,05	0,80	-0,38	0,37	1,00			1,62
9	border_region	0,00	0,00	-0,01	-0,02	0,03	-0,11	0,48	-0,02	1,00		3,03
10	tech_dist	-0,19	-0,17	-0,21	-0,18	-0,15	0,28	-0,03	-0,15	0,00	1,00	1,41
	Aerospace											
1	employees	1,00										0,00
2	diff_employees	0,66	1,00									3,53
3	establishments	0,73	0,46	1,00								3,03
4	diff_establishments	0,45	0,67	0,59	1,00							3,10
5	language_identical	0,04	0,04	0,00	-0,02	1,00						2,59
6	distance	-0,02	-0,01	-0,02	-0,01	-0,45	1,00					2,83
7	neighbour	0,01	0,00	0,01	-0,01	0,32	-0,23	1,00				1,29
8	country_identical	0,05	0,04	0,03	0,00	0,78	-0,38	0,37	1,00			1,63
9	border_region	0,00	0,00	0,00	0,00	0,04	-0,11	0,49	-0,02	1,00		2,89
10	tech_dist	-0,15	-0,14	-0,18	-0,13	-0,18	0,18	-0,06	-0,19	-0,01	1,00	1,42

Sources: Raw data, ECORDA and CORDIS; authors' own calculation.