The Dynamics of Inter-Regional Collaboration – An Analysis of Co-Patenting

06.11

Sidonia von Proff and Thomas Brenner

Philipps Universität Marburg

Impressum:

Working Papers on Innovation and Space
Philipps-Universität Marburg

Herausgeber:

Prof. Dr. Dr. Thomas Brenner
Deutschhausstraße 10
35032 Marburg
E-Mail: thomas.brenner@staff.uni-marburg.de

Erschienen: 2011

The Dynamics of Inter-Regional Collaboration – An Analysis of Co-Patenting[1]

Sidonia von Proff[2] and Thomas Brenner

both from Section Economic Geography and Location Research, Institute of Geography, Philipps-University, Marburg.

Abstract:

The paper at hand investigates how co-patenting over distance develops when aggregating inventive activities on a regional level. That means, the object of analysis is a link between two regions in contrast to other studies, where links between two individuals or firms are investigated. We analyse which regional characteristics influence the creation and continuation of such links. The main focus lies on different types of distance. The approach adds a dynamic view to the existing, often static literature about collaboration behaviour. The regressions are done for all patent-relevant industries in Germany. We find that several distance types decrease – as expected – the likelihood of link creation but also - not in all cases expected - of link continuation.

Keywords: patents, research collaboration, collaboration dynamics, inter-regional links, Germany.

JEL Classifications: R11, O34, L14

[1] Work in progress – please do not cite! October 2011.

[2] Corresponding Author: Sidonia von Proff, Philipps-University Marburg, Deutschhausstraße 10, 35032 Marburg, Germany. E-Mail: vonproff@geo.uni-marburg.de.

1 Introduction

Collaboration in innovation is largely localized (Rosenkopf and Almeida 2003; Picci 2010). The complexity of the task and the inherent tacit knowledge make it hard to collaborate over long distance. However, other forms of proximity like social, cognitive, or institutional proximity (see Boschma 2005) can substitute geographical proximity to a certain extent (Sternberg 2007). Distinguishing these types of proximity is of relevance for in-depth analyses of innovation activity in space. However, such a distinction is challenging, since they are often correlated and difficult to measure. Earlier work decomposed the observable clustering into links established with the help of social networks and those established by mere geographical proximity (e.g. Breschi and Lissoni 2005, Ponds et al. 2007). For the establishment of non-regional relationships the underlying personal links are especially important, i.e. the collaborators have to know each other either from collocation at an earlier point in time (which ended for example because of job mobility), from belonging to the same organization, or from having met at conferences/fairs or business meetings (Dettmann and Brenner 2010, Torre 2008).

The importance of several forms of proximity for innovative collaboration can be analyzed on the level of individuals, the level of the firm and on a regional level. Each of these three approaches emphasizes specific aspects and thus has advantages and disadvantages: An analysis on the level of individuals allows to study the personal relationships between people and is well suited to study spatial, social and cognitive proximity. However, it takes quite some effort to create a reliable data base for such an analysis. An analysis on the firm-level is well suited to study institutional proximity and the integration of external competencies dependent on spatial and organizational proximity. However, firms might have multiple locations which makes the analysis of spatial proximity problematic. An analysis on the regional level focuses on regional competencies and is well suited to address the relevance of spatial proximity. However, other kinds of proximity can only be indirectly approached.

A comprehensive understanding of the impact of the various kinds of proximity can only be reached by analyses on all three levels. While others have studied the individual and the firm level (e.g. Breschi and Lissoni 2005, Broekel and Boschma forthcoming, Hoekman et al. 2010), in the paper at hand, we add a view on the regional level of inventor networks. A firm-level analysis will be subject of a follow-on paper.

Regional competencies vary due to agglomeration effects of certain industries and research institutions. Hence, each collaboration over distance displays not only a combination of firm-level competencies, but also a combination of regional competencies. Hence, an analysis on the regional level is not only adequate to study spatial proximity but also offers the possibility to study the structural proximity of regions and its impact only innovative collaboration.

Furthermore, the regional level is of special interest for policy makers. In the literature it is repeatedly stated that intra-regional as well as inter-regional linkages are needed for a healthy economic development of regions (see, e.g., Broekel and Meder 2012). Therefore, policy makers have begun to support collaborative activities between specialized regions (EU INTERREG 2011). However, they have done so without knowing which regions are promising candidates for interregional collaboration regarding innovation. We will analyze how new innovation links between regions can be established and which regional characteristics make such links more likely.

In addition, it is well known that regional development is path-dependent (for a recent review see Martin and Sunley 2006). Hence, collaboration that has been established in the past can be expected to be repeated with a high probability. However, little is known about whether factors on the individual level or on the regional level influence the likelihood of continuing collaboration. In addition, it is not studied so far whether the factors that influence the continuation of collaborations are the same as the factors that influence starting collaboration. To sum up, the establishment of new links and the repetition of cooperation are two distinct things and it improves the understanding of collaboration to examine them separately. To the authors' knowledge this has not been done so far.

The existing literature on innovation collaboration links is to the largest extent of static nature. Although Ter Wal and Boschma (2009) have recommended to take a dynamic view on networks only a few papers in economic geography have tackled this task (e.g. Orsenigo et al. 1997; Ter Wal 2010). Analyzing the evolution of business or innovation networks can be expected to provide interesting insights on how stable relations and networks are; how young and/or small firms find and hold a position in their environment; how knowledge spreads in time and space; how policy measures influence the interaction among firms; and more (cf. e.g. Ernst and Kim 2002, Walker et al. 1997). The paper at hand investigates the dynamics of inter-regional collaboration behavior, i.e. how the establishment of new collaborations and the repetition of old collaborations depend on several regional characteristics and different types of distance.

Prior work methodologically closest to our approach is the paper by Hoekman et al. (2009) who use a gravity model for a static investigation. They explain the amount of co-patenting of inventors from two regions by the amount of overall patenting in the two regions, their distance, the existence of a national border between the regions, and by controlling whether the regions comprise top scientific institutions or a capital city. In all model specifications, spatial distance decreases the amount of collaboration, which is in line with earlier findings about the necessity of proximity for complex interaction.

For our analysis, data on co-inventing activities is taken from the PATSTAT database. The analysis is restricted to German regions and to all patents with inventors from at least two regions in Germany. We use the German labor market regions as spatial unit. Two regions are said to have collaboration activity with each other if there exists a patent (or more patents) listing one inventor from each of these regions (at least). The co-invented patents of a time period of three years represent the stock of collaboration activities between regions. This is necessary because a continued collaboration does not necessarily lead to a patent each year. We then calculate the probability of a continuation of the relationship or the generation of a new relationship respectively in the following year using a logit model.

The remainder of the paper is structured as follows. The following section presents the existing literature in the field of the persistence of innovation collaboration and the spatial dimension of innovation and contains our theoretical considerations about the regional characteristics influencing link creation and link continuation. The third section presents the database and the methods. The empirical results are presented and discussed in section four. Section five concludes.

2 Dynamics of collaboration

It is known from regional innovation literature that collaboration in innovative activities enhances innovation performance, especially if collaboration partners are close-by (Arndt and Sternberg 2000) and if there is some variety in the partners (Toedtling et al. 2010). In addition, having also links to actors outside the home region enhances innovativeness (Bathelt et al. 2004; Whittington et al. 2009). However, too many and unbalanced collaboration links can do damage as well (Broekel and Meder 2012). There is a trade -off for firms between the higher innovation potential by collaborating inter-regionally (larger variety of partners) and the higher risk of failure and hold-up due to a lower level of trust and higher costs of communication and coordination. In short, firms have to select carefully the right collaboration partners in order to benefit from the collaboration.

How does this selection take place? Most often, innovation projects emerge from contacts to suppliers and customers (Cohen et al. 2002). When they are mutually dependent from each other, a reasonable level of trust exists already. Earlier studies found that innovation activities of firms are rather persistent and relationships often long-term oriented (Cefis and Orsenigo 2001; Orsenigo et al. 1997). Of course, necessity-driven one-time cooperation projects do also exist which will not be pursued after the solution of the research problem.

However, we study inter-regional collaboration on the regional level. Therefore, we have to keep in mind that inter-regional collaboration links can also emerge out of links between different plants or subsidiaries of large companies. Another difference between the regional level and the firm level is that an inter-regional link does not have to be maintained by the same co-inventors. An ongoing link between two regions might be based on the same pairs of inventors in most cases, but might also be newly established between different actors again and again, whereby later links might be, but have not to be, connected to the first link. Even though the co-patenting firms or individuals may not be the same, there are environmental factors influencing inter-regional collaboration, and not only the existence, but also the dynamics. The core of our analysis is to investigate the differences between the continuation and the creation of inter-regional links and the effects of different types of distance on both, continuation and creation.

Why should there be differences between link creation and continuation? We might distinguish two kinds of factors influencing link dynamics. Firstly, some factors facilitate or hinder the interaction itself and thus have an influence on both link creation and link continuation (e.g. geographical distance). Secondly, there are factors that only influence the establishment of collaborations either because they make them looking more attractive or because they make their establishment easier. Examples are technological similarity, which makes initial communication easier, and social proximity, which causes trust between the actors and, hence, increases the willingness to cooperate. Such factors influence link creation but not link continuation, because in the latter case a link is already created and its attractiveness can be judged on the basis of past experience as well as trust is established by collaboration.

In the following, we will discuss the various environmental factors influencing inter-regional collaboration one after the other.

<u>Distance</u>

One of the main aims of this paper is to study the impact of various kinds of distances on the creation and continuation of innovative collaboration. The reason for this aim is that social interaction decreases with distance, but various different kinds of distances are involved in this relationship.

The human nature favors short-distance connections for regular and intensive communication. Empirical studies find a decrease of collaboration with increasing geographical distance between firms (Hoekman et al. 2010, Christ 2010), between academia and industry (von Proff and Dettmann 2010), and between academics (Frenken et al. 2009). Not even the rise of information and communication technologies was able to increase long-distance collaboration strongly, especially not when the joint work is complex as it is the case for innovation projects (Olson and Olson 2000). If inter-regional links exist, one can assume that they have arisen out of personal acquaintances, mediated by former collocation, by meetings on conferences and so on (Dettmann and Brenner 2010, Torre 2008). Due to a probable lack of regular meetings and profound exchange of problems and wishes, there is a high risk of either project failure or termination of cooperation over time. In addition, since a person from one region knows more potential partners in near regions than in regions far away, there is also a higher chance to start a new collaboration with someone from a near region. Hence, we expect:

Hypothesis 1a: Establishing new collaboration as well as continuing existing collaboration is less likely over larger distances.

As mentioned above, there are more dimensions of distance/proximity than purely geographical distance. On the regional level, technological proximity is of relevance. Technological proximity refers to the capabilities of research and development activities in the regions, e.g. measured by the similarity of the patenting behavior. Whether the similarity is of advantage for inter-regional co-operation depends on the purpose of collaboration: collaboration can serve the purpose (next to further aims of collaboration) of adding complementary capabilities or the purpose of increasing the capacity of existing specialization (cf. Hagedoorn 2002; Nooteboom 1999).

In the first case, dissimilarity of regions will increase the probability of a link between them. Regions with a low level of innovation in a certain industry will try to form links to regions which are successful in this industry. Firms in regions where the industry is strong can achieve competitive advantages by integrating knowledge from elsewhere. However, the technological dissimilarity must not be too large, because then mutual understanding and joint invention activities will not be possible or, at least, difficult to conduct (Nooteboom 1999). Once collaboration is conducted despite dissimilarities, further collaboration is easier to conduct because a common understanding is developed in the first collaboration. Furthermore, existing collaboration leads to some assimilation. Hence, dissimilarity especially influences link creation while there is no or much less effect on link continuation.

In the second case – resource expansion – a similar technological profile in both regions will be an advantage. However, if two regions do have a link, the degree of newness is reduced and mutual understanding increased (Nooteboom 1999). This is independent of the technology and holds for every link. Hence, we do not expect an influence of technological similarity on link continuation.

Hypothesis 1b: Increasing similarity of the technological profiles of two regions increases the probability of link creation up to a certain level before decreasing again. Hence, there is an optimal value of similarity for collaboration. Technological similarity has little influence on the probability of link continuation.

A third proximity measure is what we call here “industrial proximity”. Regions differ in their industrial structure. Industry structure refers to the target market, i.e. a similar industry structure implies that two regions compete. While this similarity increases competition and by this complicates collaboration, it enhances mutual understanding. Some clusters (e.g. Saxony Valley) have shown a positive effect of competition on co-operation, but this co-operation takes place preferably intra-regionally. In the inter-regional case (the focus of this study) the local “buzz” is missing and control cannot be exerted as good as over short spatial distance. The risk of hold-up is correspondingly high. Therefore, the barrier of competition can be assumed to have a stronger effect than the facilitated understanding. As soon as a link exists, trust is built up, so that industry structure should not have an influence on link continuation.

Hypothesis 1c: Similarity in the industry structure of two regions decreases the probability of link creation and has no influence on link continuation.

The last distance variable in our context is social distance, i.e. indirect links (“friends-of-friends”) of inventors. As mentioned in the introduction, direct and indirect links enhance link formation (Breschi and Lissoni 2005). At first sight, this concept is of limited explanatory power on an aggregated level, since an indirect link between two regions does not necessarily display an indirect link on the individual level. However, the more indirect links between two regions exist, the more likely is indeed an indirect link between two inventors, regarding the low number of inventors compared with the inhabitants of a region. Therefore the aggregated level may display the underlying social network of inventors. Regarding the case of link continuation, an influence of second degree links seems unlikely: Either the collaboration is working well or not. Whether other people are known that collaborate with the same partner does not strongly influence our judgment if we can rely on own experience.

Hypothesis 1d: The probability of link creation increases with the number of indirect links between two regions. The probability of link continuation is independent of the number of indirect links.

Universities and research institutes

Research on university-industry linkages found that academics use a variety of linkages to industry (D'Este and Patel 2007), i.e. informal relations like meeting on conferences as well as formal collaboration like contract or joint research. Since universities in Germany are usually short of (government) funding, it is essential to receive funding from industry in addition. It is not always easy to find adequate partners and a successful relationship (successful for both sides) will certainly be repeated (Lee 2000). If the link cannot be maintained over time, it is likely that the academics will search for new partners because they either need funding or are interested in the commercialization of research findings. It is known that the amount of academic patenting has increased over time (von Proff et al. 2012) and, although patents display only a part of all science-industry links, the rise in patenting indicates an orientation towards industrial application and commercialization of academic research. An increasing use of academic knowledge in industry (not only academic patents) has been found by Kim et al. (2005) as well as Mansfield (1998) empirically. According to the findings about the importance of proximity (see above), establishing new links will likely happen at places familiar to researchers, i.e. predominantly places where they have lived before or where other universities are (in case of inter-university collaboration). The creation of new links involving such places is more likely, because researchers at universities are a specifically mobile group of employees and have often more than one region, where they have lived before and where they could establish links to. Furthermore, researchers at universities often have many far-distance collaboration activities with other researchers whom they meet on conferences. Since these contacts are usually quite stable, universities should also contribute to the stability of links from regions that have large and/or important universities.

Hypothesis 2a: New collaboration is more likely to be established in regions with a high level of university research and existing links are more likely to be continued in such regions.

Next to the mere size of university research in two regions the distribution between them is of interest. A smaller region has a strong interest to find a partner in a region with larger research capacities, because this would help to overcome constraints. If a region with larger research capacity has researchers with unique skills, other regions can even be forced to connect to this strong region in order to get access to these resources. Regions with large research capacities may have a great interest to find a partner in other high potential regions in order to benefit best from the collaboration. But firstly, there are only a few regions with really high research capacities, and secondly, these regions compete. Therefore, we expect that the dissimilarity of research capacities increases the probability of link creation between two regions. As soon as a link exists already, the attractiveness caused by this dissimilarity should not influence the collaboration anymore.

Hypothesis 2b: Similarity in the amount of university research capacities decreases the probability of link creation and has no influence on link continuation.

<u>Employment and firm size</u>

The employment in the relevant industries in a region is, of course, a measure of the overall potential for collaboration. Most inventors of patents are employees in firms. Hence, regions with higher employment can be expected to be more likely involved in new collaboration as well as to have higher chances to continue existing collaboration links. Interestingly, there are hardly any studies on this topic. It is only known that larger firms collaborate more often overall (Fritsch 2001) and more often with public research institutions (Rothwell and Dodgson 1991). This could be due to larger resources available for searching collaboration partners. Larger resources are also of advantage for large projects which need more time to generate benefits and may decrease risk aversion. Using aggregated employment data on a regional level implies ignoring firm size effects. However, only if collaboration *per worker* differs with firm size there will be an effect visible on the regional level. For example, if per worker collaboration is larger in small firms, a region with many small firms can have overall more links than a region with few large firms. To the authors' knowledge, there is no study investigating this relationship and we do not know a convincing argument in favor of differing per worker collaboration. Nevertheless, we will test the following hypothesis.

Hypothesis 3: Regions with a larger average firm size continue and start collaboration more likely than those with smaller average firm size.

Not all kinds of employees have the same relevance for innovation activities. R&D staff is disproportionally often involved in innovative activities. Fritsch (2001) finds a positive influence of the share of R&D staff on collaboration on the firm level. This should hold also on the regional level: Regions with a high level of R&D employees will certainly have more collaboration links. Usually they hold a university degree and sometimes they even have stayed for some time in public research before moving to an R&D department in a firm. They can stay in contact with the professors who were their teachers and with fellow students who probably work elsewhere. Furthermore, highly-educated persons are spatially more flexible when changing jobs (van Ham et al. 2001) and existing contacts are not necessarily lost when moving to another region. If they have spent considerable time together the relationship may be strong enough to overcome distance even in the longer run.

Hypothesis 4a: Regions with a higher level of employment/R&D employment in a certain industry will create and continue collaboration links more often.

As in the case of university research, the distribution of employees between the two regions can have an effect exceeding the mere amount. The argumentation is analogue to the case of university research: complementarities enhance link creation, i.e. a difference in the employment level is of advantage. As soon as a link exists, greater trust or similar input levels do not matter anymore – the continuation of collaboration should be independent of similar levels of employment. We assume the following.

Hypothesis 4b: Similarity in the amount of employment/R&D employment in a certain industry will decrease the likelihood of link creation. There is no similar effect on link continuation.

Technology / industry

Innovation patterns of technologies and industries differ. There are industries in which patents are essential for developing and exploiting inventions (e.g. pharmaceuticals, chemicals) and others in which patents are of low relevance (e.g. rubber, office equipment, cf. Mansfield 1986). Similarly, there will be industry effects in the collaboration behaviour, e.g. due to differences in the spatial concentration of industries. We take care of this by analysing 19 industries separately, using the industries that are defined by a concordance developed by U. Schmoch and colleagues (the concordance is a current version of the concordance published in Schmoch et al. (2003) and was obtained directly from U. Schmoch).

3 Data and method

Patents are widely-used in innovation studies, e.g. for analyses of the location of knowledge production and the network through which the knowledge diffuses. For the latter means, two approaches are common: citations display how new knowledge spreads in space over time, and co-inventions show the network structure of collaboration in space. According to the focus of our study on collaboration links we will use co-invention data taken from the PATSTAT database. Full data is available for 1999 to 2007. Since the existence of a link between two regions is measured over three years and is lagged one year, the regression will be done for 2002 to 2007. The regional units are the 270 labor market regions in Germany. We focus exclusively on German inventors because we use a number of the independent variables that are available for us only for German regions.

The factors discussed in section two are empirically measured as follows. The geographical distance between two regions is the air-line distance between the geographical centers of the regions adjusted for earth curvature.[1] As a measure of technological proximity the correlation between the amounts of patenting in each technological field (19 technologies from the used concordance) is taken as well as the square of this measure. The proxy for “industrial distance” is the correlation between the numbers of employees in the individual industries (19 industries from the used concordance). “Social distance” (even though a flawed concept on

[1] As a check for robustness, we used airline distance without adjusting for earth curvature as well as the center of economic activity of each region as measured by firms in the MARKUS database in the final regressions. The coefficients do not change in significance and less than one per mill in size.

the regional level as explained above) is included by counting second degree links, i.e. the number of links to any same third region that two regions have in common.

The strength of university research is measured by the research budget. A robustness check with third-party funding and graduates did hardly change the results, since these variables are highly correlated. The assignment of research areas of study to the industries follows the assignment of Schlump and Brenner (2011). Overall employment in each industry and R&D employment are also highly correlated; we will use employment because the data are more reliable. For both the university budget and the employment variable not only the sum of the values in each two regions is taken (equivalent to a gravity approach) but also the difference (according to our considerations about the (dis)similarity of the regions).

Next to the independent variables explained above three control variables are included. Firstly, the sheer size of a region plays a role for innovation links. Therefore, the sum of population of each pair of regions is included. As for the other exogenous variables, the difference is also included to account for similarity. Secondly, the overall number of patents filed in a region is, of course, related to the number of patent links to other regions. We control for the amount of patenting by including the overall number of patents which have been filed in any pair of regions (fractional counting for multiple inventors and multiple patent classes on one patent) in the respective industry. Since the sum of the patent values of two regions is highly correlated with the difference between them (r = 95%), we use only the sum.[2]

At last, we calculate the average frequency of a link between two regions (pooled for all years) in order to separate two effects. First, links between two regions are more or less likely due to the characteristics of these regions. Second, in each year links between two regions might be established or continued due to the characteristics of these regions. The former effect is usually examined in empirical studies. We focus on the dynamics of links and, thus, on the latter effect. However, if we simply study whether a link is created or continued at a certain point in time, both effects will show up in our results. If a link between two regions is more likely, on average, it will also be more likely created or continued at each point in time. Therefore, we use the average frequency of a link between two regions as an independent variable in our analysis in order to separate the two effects.

[2] There are several regions without any patents in the respective industries while some regions display a high patenting activity. Therefore, the sum and the difference have such a high correlation.

All non-patent data are either from official databases from the Federal Statistical Office of Germany (graduates, employment and population) or from the German Institute for Labor (R&D employees and employees with a tertiary degree for robustness checks). Table 1 gives an overview of the variables.

Variable	Explanation	Hypotheses for impact on	
		link creation	link continuation
$Link_{ijt}$	Dummy variable taking the value one if there is at least one co-invention between regions i and j in year t		
$AvFreq_{ij}$	Average frequency of a link between regions i and j		
$Distance_{ij}$	Distance in km between the centers of regions i and j (adjusted for earth curvature)	- (Hyp. 1a)	- (Hyp. 1a)
$PatSim_{i,j,t-1}$	Correlation between the distributions of patent activity among the patent classes for regions i and j in year t-1	+ (Hyp. 1b)	0 (Hyp. 1b)
$SqPatSim_{i,j,t-1}$	Square of *PatSim*	- (Hyp. 1b)	0 (Hyp. 1b)
$IndSim_{i,j,t-1}$	Correlation between the industry structures of regions i and j in year t-1 (measured by employment in the various industries)	- (Hyp. 1c)	0 (Hyp. 1c)
$SecDeg_{i,j,t-1}$	Number of third regions to which both regions have a link (second-degree links) in the preceding year	+ (Hyp. 1d)	0 (Hyp. 1d)
$Population_{i,j,t-1}$	Sum of populations of regions i and j in year t-1		
$PopDiff_{i,j,t-1}$	Difference of populations of regions i and j in year t-1		
$Patents_{ijt}$	Sum of patents (fractional counting for multiple inventors and multiple patent classes on one patent) in regions i and j in year t		
$Firmsize_{i,j,t-1}$	Average firm size in the considered industry in regions i and j in year t-1	+ (Hyp. 3)	+ (Hyp. 3)
$Employm_{i,j,t-1}$	Sum of employment in the considered industry in regions i and j in year t-1	+ (Hyp. 4a)	+ (Hyp. 4a)
$EmploymDiff_{i,j,t-1}$	Difference of employment in the considered industry between regions i and j in year t-1	+ (Hyp. 4b)	0 (Hyp. 4b)
$UniBud_{i,j,t-1}$	Sum of university research budget in industry-related subjects in regions i and j in year t-1.	+ (Hyp. 2a)	+ (Hyp. 2a)
$UniBudDiff_{i,j,t-1}$	Difference of university research budget in industry-related subjects in regions i and j in year t-1.	+ (Hyp. 2b)	0 (Hyp. 2b)
$stockoflinks_{i,j,t-1}$	Dummy variable taking the value one if there has been at least one collaboration between regions i and j during the years t-3 and t-1		

Table 1: Variables and the expectations (hypotheses) for their impact on link creation and continuation.

For each industry two regressions are conducted: one for the probability that an existing link (link in the last three years) is continued and one for the probability that a link is built in case there has been none during the last three years. Even though for some variables the expected effects are the same, for many variables the expectations are different (see table 1). The dataset is divided into the respective parts by the variable *stockoflinks*. Expectedly, the number of region pairs with collaboration links is much smaller than those without. Since the dependent variable is dichotomous, a logistic regression is used:

Link continuation:

$$P\left(Link_{ijt}\middle|stockoflinks_{i,j,t-1} = 1\right) = f(independent\ Variables) \qquad (1)$$

Link creation:

$$P\left(Link_{ijt}\middle|stockoflinks_{i,j,t-1} = 0\right) = f(independent\ Variables) \qquad (2)$$

For research questions similar to ours some authors have used quadratic assignment procedure (QAP) test (e.g. Cantner and Graf 2006). The idea is that a link between two nodes is influenced by the existence of indirect links between them. However, due to the aggregation to a regional level indirect links between regions do not always display a “real” indirect link, i.e. one between two individuals/firms. Therefore, the logit model seems more adequate in our case.

The variance inflation factor (VIF) is used to test for multicollinearity and exclude problematic variables (exclusion if VIF>10). We do not expect structural changes during the years 2002-2007, so that we pool the data of all years (yearly dummies added).

The descriptive statistics support the earlier static findings that higher link probabilities are related to proximity, larger values for overall patenting, the various employment variables, the proxies for public research, and a larger population.

4 Results and Discussion

Table 2 and Table 3 give an overview of the results. There are much more variables significant for link creation than for link continuation as expected in the hypotheses. This is a hint that regional characteristics play only a small role for link continuation. Presumably, the experiences with the ongoing collaboration determine whether this collaboration will be continued or not.

Let us first discuss the control variables before we check our hypotheses. As expected, the average frequency of patenting is significant in every conducted regression. Total patent output has a positive influence on both link continuation and creation in several industries. Sometimes, the variable has to be excluded because it is correlated with the population variable. Similarly, *Population* is positive and significant for several industries. *PopDiff* is correlated with *Population* quite often. If it is significant, then it is negatively related to link creation; but it is not significant in any link continuation regression.

The following paragraphs will discuss the hypotheses in detail. The estimation details are available from the authors on request.

	audio-visual electronics	basic chemicals	computers	electric. machinery	electron. compon.	energy machinery	general machinery	machine tools	measurement	medical equipment	metal products	non-polymer materials	optics	others	pharma-ceuticals	polymers	special machinery	telecomm-unications	transport
AvFreq	+ ***	+ ***	+ ***	+ ***	+ ***	+ ***	+ ***	+ ***	+ ***	+ ***	+ ***	+ ***	+ ***	+ ***	+ ***	+ ***	+ ***	+ ***	+ ***
Distance	- **	- ***	- ***	- ***	- ***	- ***	- ***	- ***	- ***	- ***	- ***	- ***	- ***	- ***	- ***	- ***	- ***	- ***	- ***
PatSim									+ **				+ **					+ ***	
SqPatSim		+ **	+ ***			+ *						+ ***			+ **				
IndSim					- **	- **			- **	- **	- **		- **					- ***	- ***
SecDeg		+ ***	+ ***	+ ***	+ **				+ ***						+ ***		- *	+ ***	+ ***
Population	+ ***		+ ***	+ ***	+ ***		+ ***		+ ***		+ *		+ **	excl.	+ **		+ **	+ **	+ **
PopDiff		excl.	excl.	excl.	- ***	excl.	excl.	excl.	- ***	excl.	- **	excl.		excl.	excl.	excl.	excl.	excl.	excl.
Patents		+ ***		excl.	+ *		excl.					+ ***	+ ***		+ ***	+ ***	+ ***		
Firmsize		- *	+ *	+ ***	+ ***				+ *	+ **	- *	+ *					+ *		
Employm		+ ***			- *					+ **	+ ***			+ *			- ***		
EmploymDiff	excl.	excl.	excl.	excl.	excl.	excl.	excl.	excl.	excl.	excl.	excl.	excl.	excl.	excl.	excl.	excl.	excl.	excl.	excl.
UniBud								+ *						+ *	+ *			+ ***	
UniBudDiff		+ **														+ *	+ *		+ *

Yearly dummies included. Excl. stands for variables which had to be excluded due to multicollinearity.
+/- indicates the sign of the coefficient, */**/*** indicates the significance on the 0.1/1/5%-level

Table 2: Overview of results for link creation.

	audio-visual electronics	basic chemicals	computers	electric. machinery	electron. compon.	energy machinery	general machinery	machine tools	measurement	medical equipment	metal products	non-polymer materials	optics	others	pharma-ceuticals	polymers	special machinery	telecomm-unications	transport
AvFreq	+ ***	+ ***	+ ***	+ ***	+ ***	+ ***	+ ***	+ ***	+ ***	+ ***	+ ***	+ ***	+ ***	+ ***	+ ***	+ ***	+ ***	+ ***	+ ***
Distance		- ***	- **	- ***	- ***	- ***	- **		- ***	- ***	- *	- ***	- **	- **	- **	- ***	- ***	- ***	- ***
PatSim							+ *								+ *				
SqPatSim		+ *																	
IndSim																			+ *
SecDeg		+ ***	+ ***	+ *					+ **						+ ***			+ ***	
Population		+ *		+ **	+ **	excl.	+ ***						+ *	excl.			+ **	+ ***	+ *
PopDiff		excl.												excl.					excl.
Patents	+ *	+ ***	+ *	excl.			excl.	+ **			+ **	+ ***	+ **	+ *		+ ***	+ ***		+ ***
Firmsize						- **													
Employm							+ *								+ **				
EmploymDiff			excl.		excl.	excl.	excl.	- *			- **			excl.			excl.		
UniBud						+ *								+ *				+ **	+ **
UniBudDiff														- *					

Yearly dummies included. Excl. stands for variables which had to be excluded due to multicollinearity.
+/- indicates the sign of the coefficient, */**/*** its significance on the 0.1/1/5%-level

Table 3: Overview of results for link continuation.

Distance (Hypothesis 1)

A significant negative coefficient is found for geographical distance in all regressions for link creation and almost all regressions for link continuation. Hence, Hypothesis 1a is clearly confirmed. In line with our reasoning above, we find that the probability to establish new links as well as the probability to continue an existing link decreases with the geographical distance between the two involved regions. Spatial distance matters for collaboration.

Regarding technological distance, we find few significant results: link creation is fostered by technological similarity in measurement, optics, and telecommunications, link continuation in general machinery and pharmaceuticals. This is roughly in line with Hypothesis 1b and the findings of Gomes-Casseres et al. (2006) who find that knowledge flows between firms increase with technological proximity. In contradiction to Hypothesis 1b, the squared term of patent similarity is most often insignificant and if it is significant, than positively (link creation: basic chemicals, computers, energy machinery, non-polymer materials, and pharmaceuticals; link continuation: basic chemicals). Our results do not confirm an optimal value of similarity, at least not within the range that is covered by our data. Furthermore, *PatSim* and *SqPatSim* are never significant together. Hence, we find positive relationships between technological similarity and link creation as well as link continuation, sometimes rather in linear form and sometimes rather in quadratic form. In total we find such positive relationship for link creation for 8 out of 18 industries and for link continuation for 3 out of 18 industries. Hence, the expectation (Hypothesis 1b) that this relationship is much stronger in the case of link creation is confirmed.

Industrial proximity has only for the transport industry a significant effect on link continuation, confirming Hypothesis 1c, which states that industrial proximity is not relevant in the case of link continuation. As stated in Hypothesis 1c, the relationship between industrial similarity and link creation is significantly negative for eight industries (electronic components, energy machinery, measurement, medical equipment, metal products, optics, telecommunications, and transport). Indeed, collaborations with actors in other regions are less likely to happen between firms that serve similar markets, maybe due to a lack of trust caused by competition on the market.

The last distance measure is the number of second degree links. For 6 out of 19 industries a larger number of indirect links increases the probability of link continuation as it does for 8 out of 19 industries in the case of link creation. Significant coefficients for 6 industries disprove our hypothesis that second degree links do not matter for link continuation on the regional level. A possible explanation is that certain groups of regions are strongly interconnected and that this group structure supports the stability of the links. The hypothesis for link creation is supported. Interestingly, there is a significantly negative influence of second degree links in special machinery on the 5% level. We do not have an explanation for this negative relationship.

Overall, of all distance measures, geographical distance has by far the strongest influence on the collaboration dynamics. However, we stated above that the regional level of analysis is especially well suited to study geographical distance and its impact.

University research capacity (Hypothesis 2)

Hypothesis 2a is confirmed by our findings. Although significant results are only found for a few industries, all significant coefficients for the sum of universities' research budgets in the two regions are positive. These positive relationships are found for link creation (four industries) as well as link continuation (four industries). Hence, the presence of strong universities seems to support the creation and continuation of collaborations with other regions.

Hypothesis 2b is also confirmed by the above results. We find significant positive relationships between the difference of universities' research budgets and link creation for four industries, while we find only for one industry a significant relationship between this difference and link continuation. The latter relationship is negative and we have no explanation for it. However, since it occurs only in one industry, we focus on the positive relationships between the difference of universities' research budgets and link creation. These relationships support the above argument that actors in regions with low or no university research budgets will especially try to connect to regions with high university research budgets in order to get excess to this public research environment.

Firm size (Hypothesis 3)

A larger average firm size in a region has no effect on link continuation (exception: significantly negative relationship in energy machinery) but for seven out of 19 industries we find a significantly positive relationship between average firm size and link creation (for basic chemicals and metal products a negative one). The assumptions of larger resources for searching collaboration partners and the probably lower risk aversion seem to play a role for link creation. However, if a link exists these arguments seem to lose their relevance. Thus, Hypothesis 3 is partly supported. The impact of firm size seems to follow the same logic as many other factors: It is relevant for the creation of collaboration but not for the continuation of collaboration.

Employment (Hypothesis 4)

In Hypothesis 4a we have claimed that a high level of employment should positively influence the creation of collaborations between regions. Our results do not confirm this hypothesis. The results are mixed: We find significantly positive relationships for some industries, significantly negative relationships for some other industries and insignificant coefficients for most industries. This holds also for including the amount of R&D employment instead of overall employment as a robustness check (results not displayed). It seems that the variables do not capture the theoretically intuitive effect. Instead, the effects of industry-specific employment on link creation seem to depend on the industry that is studied. An explanation could be that firms differ strongly in their patent propensity and a few large firms often dominate the patent activity (like Bosch and Siemens). These are usually rather large firms that imply a high number of employees in the respective regions. Their behavior might strongly influence the empirical findings.

The second part of Hypothesis 4a states a positive relationship between employment and link continuation. Such a relationship is significantly confirmed for two industries. Thus, we find some evidence for the expectation that regions with more employees are more likely to continue collaborations.

Hypothesis 4b can only be studied partly. Due to multicollinearity we are not able to include the difference in employment numbers in any of the regressions for link creation. We are able to include this variable in some of the regressions for link continuation. In some of these cases we find significantly negative coefficients, which contradict Hypothesis 4b. Differences in employment numbers might rather cause, at least in some industries, a higher probability for collaborations to break up. Maybe differences in firm sizes make collaboration more complex and, therefore, cause less positive evaluations of the collaborations by the involved actors, so that they are not continued.

Differences between link creation and link continuation

To sum up, we find several differences between the factors related to link creation and the factors related to link continuation. Especially many regional factors have a significant relationship to link creation but no significant relationship to link continuation. We conclude from this that regional characteristics influence the choice of new partners while the decision to continue collaboration is made on the basis of experiences in this collaboration which are rarely related to regional characteristics. Most likely, the characteristics of the individuals (inventors and firms) shape the continuation of collaboration more strongly. In a new analysis on the basis of firms this should be investigated in detail.

An overview on the different findings for link creation and link continuation is given in Table 4. If we ignore all variables for which a significant relationship is found only for one industry (testing 19 industries this might be a statistical artifact), there are three kinds of variables. First, there are variables that have the same relationship to link creation and continuation: Spatial distance (*Distance*), technological distance (*PatSim* and *SqPatSim*), social distance represented by second degree links (*SecDeg*), and the sum of universities' budget (*UniBud*). Three of these four factors with a clear impact on link creation and link continuation represent some kind of distance.

Second, there are variables that only have a significant relationship to link creation, but not to link continuation: Industrial similarity (*IndSim*) and the difference of the universities' budget (*UniBudDiff*). In both cases differences in the relevant endowment (private and public activities) of the regions increases the likelihood of link creation, but not the likelihood of link continuation. Actors seem to search for partners in locations with a different structure of the industry and public research, while once such a partner is found, the evaluation of the collaboration is independent from these regional characteristics.

Third, there are variables that show no or little significant relationship with link continuation and a significant relationship with link creation that depends on the industry under consideration: firm size (*Firmsize*) and relevant employment (*Employm*). Again, once collaboration is established it seems to matter little for the evaluation, and thus continuation, of the collaboration whether firm and the regional total employment are large or small. However, in some industries more links are generated within regions with large firms or high industrial activity, while in other industries more links are generated within regions with small firms or low industrial activity.

Variable	Significant relationships for	
	link creation	**link continuation**
$Distance_{ij}$	**- (19)**	**- (17)**
$PatSim_{i,j,t-1}$ or $SqPatSim_{i,j,t-1}$	**+ (8)**	**+ (3)**
$IndSim_{i,j,t-1}$	**- (8)**	+ (1)
$SecDeg_{i,j,t-1}$	**+ (8)** / - (1)	**+ (6)**
$Firmsize_{i,j,t-1}$	**+ (7) / - (2)**	- (1)
$Employm_{i,j,t-1}$	**+ (4) / - (2)**	**+ (2)**
$EmploymDiff_{i,j,t-1}$	excl.	**- (2)**
$UniBud_{i,j,t-1}$	**+ (4)**	**+ (4)**
$UniBudDiff_{i,j,t-1}$	**+ (4)**	- (1)

Table 4: Significant relationships found for the various variables with link creation and continuation. Number of industries for which the relationship is found is given in parentheses.

5 Conclusion

In this paper we studied the relationship between a number of regional characteristics and the probability of collaboration between actors from different regions. In particular, we apply four distance measures and distinguish between link creation and link continuation. The most striking result is the strong negative effect of various kinds of distances not only on the creation but also on the continuation of collaboration links. This supports the literature claiming that distance matters. Invention is a highly complex and difficult task and distance is a barrier for communication and knowledge exchange. This seems to have an impact on link creation as well as on link continuation. We find significant negative relationship for spatial, technological, and social distance. Industrial distance, instead, seems to play a role only for link creation, but not for link continuation. Furthermore, industrial distance increases the likelihood of link creation.

Differences in university endowment between regions seem to have similar impacts as differences in the industry structure: they enhance link creation but have no significant impact on link continuation. Hence, we find that distances – geographic, technical and social – seem to have a negative impact on link creation and continuation, while differences in the actor structure – industrial structure and public research – seem to have a positive impact on link creation but no impact on link continuation.

Differences between industries are mainly found in the relationships between link creation and firm size as well as relevant regional employment. In the case of these factors more detailed, industry-specific studies would help to gather further insights.

Regarding policy measures for interregional collaboration our results imply a focus on region pairs which have low spatial, technological, and social distance. By taking into account these distances the probability of a long-term relationship can be increased. In addition, indirect measures are imaginable, i.e. programs reducing social distance by fostering employee mobility or reducing technological distance by fostering the investment in certain technologies.

For more detailed insights, it is necessary to analyze also the firm level of cooperation. However, the analysis in the paper at hand is an interesting starting point to see how link continuation and link creation differ and how distance does not only play a role for the existence of links but also for the dynamics of link creation and continuation.

6 References

Arndt, Olaf and Sternberg, Rolf, 2000. Do Manufacturing Firms Profit from Intraregional Innovation Linkages? An Empirical Based Answer. *European Planning Studies*, 8(4), 465 - 485.

Bathelt, Harald, Malmberg, Anders and Maskell, Peter, 2004. Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation. *Progress in Human Geography*, 28(1), 31-56.

Boschma, Ron, 2005. Proximity and Innovation: A Critical Assessment. *Regional Studies: The Journal of the Regional Studies Association*, 39(1), 61-74.

Breschi, Stefano and Lissoni, Francesco, 2005, Knowledge Networks from Patent Data, *Handbook of Quantitative Science and Technology Research*, pp. 613-643.

Broekel, Tom and Boschma, Ron, forthcoming. Knowledge networks in the Dutch aviation industry: the proximity paradox. *Journal of Economic Geography*.

Broekel, Tom and Meder, Andreas, 2012. Collaboration intensity and regional innovation efficiency in Germany - A conditional efficiency approach. *Industry & Innovation*, forthcoming.

Cantner, Uwe and Graf, Holger, 2006. The network of innovators in Jena: An application of social network analysis. *Research Policy*, 35(4), 463-480.

Cefis, Elena and Orsenigo, Luigi, 2001. The persistence of innovative activities: A cross-countries and cross-sectors comparative analysis. *Research Policy*, 30(7), 1139-1158.

Christ, Julian P., 2010, The geography and co-location of European technology-specific co-inventorship networks, FZID Discussion Paper 14-2010.

Cohen, Wesley M., Nelson, Richard R. and Walsh, John, 2002. Links and Impacts: The Influence of Public Research on Industry R&D. *Management Science*, 48(1), 1-23.

D'Este, P. and Patel, P., 2007. University-industry linkages in the UK: What are the factors underlying the variety of interactions with industry? *Research Policy*, 36(9), 1295-1313.
Dettmann, Anja and Brenner, Thomas, 2010, Proximity is a Social Process: A conceptual Modification, DRUID Winter Conference 2010, Aalborg.
Ernst, Dieter and Kim, Linsu, 2002. Global production networks, knowledge diffusion, and local capability formation. *Research Policy*, 31(8-9), 1417-1429.
EU INTERREG 2011. INTERREG IVC Programme. http://i4c.eu
Frenken, Koen, Hoekman, Jarno, Kok, Suzanne, Ponds, Roderik, van Oort, Frank et al., 2009, Death of Distance in Science? A Gravity Approach to Research Collaboration. In: Andreas Pyka and Andrea Scharnhorst (Eds.), *Innovation Networks - New Approaches in Modelling and Analyzing*. Springer, pp. 43-57.
Fritsch, Michael, 2001. Co-operation in Regional Innovation Systems. *Regional Studies*, 35(4), 297-307.
Gomes-Casseres, Benjamin, Hagedoorn, John and Jaffe, Adam B., 2006. Do alliances promote knowledge flows? *Journal of Financial Economics*, 80(1), 5-33.
Hagedoorn, John, 2002. Inter-firm R&D partnerships: an overview of major trends and patterns since 1960. *Research Policy*, 31(4), 477-492.
Hoekman, Jarno, Frenken, Koen and Tijssen, Robert J. W., 2010. Research collaboration at a distance: Changing spatial patterns of scientific collaboration within Europe. *Research Policy*, 39(5), 662-673.
Hoekman, Jarno, Frenken, Koen and van Oort, Frank, 2009. The geography of collaborative knowledge production in Europe. *The Annals of Regional Science*, 43(3), 721-738.
Kim, Jinyoung, Lee, Sangjoon John and Marschke, Gerald, 2005, The Influence of University Research on Industrial Innovation, NBER Working Paper.
Lee, Yong S., 2000. The Sustainability of University-Industry Research Collaboration: An Empirical Assessment. *The Journal of Technology Transfer*, 25(2), 111-133.
Mansfield, Edwin, 1986. Patents and Innovation: An Empirical Study. *Management Science*, 32(2), 173-181.
Mansfield, Edwin, 1998. Academic research and industrial innovation: An update of empirical findings. *Research Policy*, 26(7-8), 773-776.
Martin, Ron and Sunley, Peter, 2006. Path dependence and regional economic evolution. *Journal of Economic Geography*, 6(4), 395-437.
Nooteboom, Bart, 1999, *Inter-firm alliances: analysis and design*. Routledge, New York.
Olson, Gary M. and Olson, Judith S., 2000. Distance Matters. *Human-Computer Interaction*, 15(2), 139-178.
Orsenigo, Luigi, Pammolli, Fabio, Riccaboni, Massimo, Bonaccorsi, Andrea and Turchetti, Giuseppe, 1997. The Evolution of Knowledge and the Dynamics of an Industry Network. *Journal of Management and Governance*, 1(2), 147-175.
Picci, Lucio, 2010. The internationalization of inventive activity: A gravity model using patent data. *Research Policy*, 39(8), 1070-1081.
Ponds, Roderik, Oort, Frank van and Frenken, Koen, 2007. The geographical and institutional proximity of research collaboration. *Papers in Regional Science*, 86(3), 423-443.
Rosenkopf, Lori and Almeida, Paul, 2003. Overcoming Local Search Through Alliances and Mobility. *Management Science*, 49(6), 751-766.
Rothwell, Roy and Dodgson, Mark, 1991. External linkages and innovation in small and medium-sized enterprises. *R&D Management*, 21(2), 125-138.
Schlump, Charlotte and Brenner, Thomas, 2011, Universities and Regional Innovation Output: A Detailed Empirical Study of 19 Technologies in Germany, Paper for the DIME final conference, Maastricht.

Sternberg, Rolf, 2007. Entrepreneurship, Proximity, and Regional Innovation Systems. *Tijdschrift voor Economische en Sociale Geografie*, 98, 652-666.

Ter Wal, Anne L. J., 2010, The Spatial Dynamics of the Inventor Network in German Biotechnology: Geographical Proximity Versus Triadic Closure, Best Paper Proceedings of the 2010 Academy of Management Meeting.

Ter Wal, Anne L. J. and Boschma, Ron A., 2009. Co-evolution of Firms, Industries and Networks in Space. *Regional Studies*, forthcoming.

Toedtling, Franz, Grillitsch, Markus and Hoeglinger, Christoph, 2010, Knowledge Sourcing and Innovation in Austrian ICT companies - Does Geography matter?, mimeo.

Torre, André, 2008. On the Role Played by Temporary Geographical Proximity in Knowledge Transmission. *Regional Studies*, 42(6), 869-889.

van Ham, M., Mulder, C. H. and Hooimeijer, P., 2001. Spatial flexibility in job mobility: macrolevel opportunities and microlevel restrictions. *Environment and Planning A*, 33(5), 921-940.

von Proff, Sidonia, Buenstorf, Guido and Hummel, Martin, 2012. University patenting in Germany before and after 2002: what role did the professors' privilege play? *Industry & Innovation*, forthcoming.

von Proff, Sidonia and Dettmann, Anja, 2010, Inventor collaboration over distance – a comparison of academic and corporate patents, Marburg Working Papers in Innovation and Space #01.10.

Walker, Gordon, Kogut, Bruce and Shan, Weijian, 1997. Social Capital, Structural Holes and the Formation of an Industry Network. *Organization Science*, 8(2), 109-125.

Whittington, Kjersten Bunker, Owen-Smith, Jason and Powell, Walter W., 2009. Networks, Propinquity, and Innovation in Knowledge-intensive Industries. *Administrative Science Quarterly*, 54(1), 90-122.