Paper to be presented at the DRUID 2012

on

June 19 to June 21

at

CBS, Copenhagen, Denmark,

Co-patenting and inventive performance: in search of the proximity paradox

Lorenzo Cassi
Centre d'Economie de la Sorbonne - Paris1 University
Observatoire des Sciences et des Techniques - OST - Paris
lorenzo.cassi@univ-paris1.fr

Anne Plunket
Paris Sud
ADIS
anne.plunket@u-psud.fr

Abstract

Understanding how networks and proximity issues interact and whether their interaction contributes to explain the levels of innovation performance is a key and challenging issue. The aim of our study is to contribute to this vivid discussion by investigating how networks and proximity factors interact to influence the performance of inventors, measured by the value of their patents measured by citations. In particular, we study how geographical and technological proximity on the one hand, and the actors network position on the other hand influence inventive performance. In particular, we consider how network strategies, that is, the fact that inventors choose to collaborate within or outside their local network and their organization or technological boundaries impact the value of patents. This will enable us to address the so called proximity paradox, that is the fact that high degree of proximity is a prerequisite for collaboration and learning processes, but too high degree of proximity may harm individuals and hamper innovative performance.

Jelcodes:D85,O31

Co-patenting and inventive performance:

In Search of the Proximity Paradox

Lorenzo Cassi
lorenzo.cassi@univ-paris1.fr
Université Paris 1, CES - Centre d'Economie de la Sorbonne &
OST - Observatoire des Sciences et des Techniques, Paris

Anne Plunket
anne.plunket@u-psud.fr
Université Paris Sud 11, ADIS

Abstract

This paper investigates how network relations, proximity and their interplay affect collaboration and their inventive performance. Using patent citations as a proxy for patent quality, we investigate how the network and proximity characteristics of co-inventors enable them to access different sources of knowledge in different geographical and organizational contexts, and finally affect the quality of inventive collaboration. Our findings enable to address the proximity paradox, which states that proximity facilitates collaboration and knowledge sharing, but it does not necessarily increase innovative performance, too much proximity may even harm innovation (Boschma and Frenken, 2009; Broekel and Boschma, 2011).

Keywords: Social networks, geographical proximity, technological proximity, co-patenting, network formation.

JEL codes: O31, O33, R11, D85, L65

1. Introduction

The predominance of geographical proximity for knowledge creation and interactive learning is largely explained by the characteristics of knowledge underlying the processes of innovation (Audrestch and Feldman, 1996; Bathelt et al., 2004). First, spatial propinquity facilitates information and knowledge sharing through frequent interactions, especially when knowledge is tacit, complex and sticky. Second it contributes to solving coordination problems through trust building and inter-organizational learning.

However, geographical proximity *per se* only explains part of the story. Since knowledge is not in the air, actors need to be embedded in local networks and occupy central positions in order to gain access to information and resources that influence innovation (Whittington et al., 2009). And indeed, social proximity within networks is increasingly acknowledged as a key mechanism to understand knowledge flows underlying inter-active learning and innovation (Agrawal et al. 2008; Breschi and Lissoni, 2009).

Besides geography and networks, two other factors mediate knowledge flows and interactive-learning, one technological and one organizational. Because of absorptive capacity, some level of technological proximity is needed between actors in the network for interactive learning and knowledge exchange to take place (Jaffe, 1989; Cohen and Levinthal, 1990; Nooteboom, et al. 2007). Organisational proximity also plays a role in the innovation process since it reduces uncertainty, limits the risk of opportunism and facilitates communication between actors and increases performance.

While a high degree of proximity and network embeddedness are main drivers of network formation and knowledge diffusion, the impact on innovative performance is rather

ambiguous, “since proximity between actors does not necessarily translate into higher innovative performance” (Boschma and Frenken, 2009). The so-called “proximity paradox” argues that the drivers of network formation should be distinguished from the determinants of innovative performance. If proximity and network embeddedness clearly explain the formation of network relations (Autant-Bernard, et al. 2007; Cassi and Plunket, 2010) interactive learning and knowledge flows (Agrawal et al. 2006; Breschi and Lissoni, 2009), they may not necessarily benefit innovative performance, they may even be harmful for interactive learning (Boschma, and Frenken, 2009).

This paper aims to provide further evidence addressing the so-called proximity paradox. In order to do so, we contrast how proximity and network relations, separately or in combination, affect the formation of collaborations on the one hand, and their inventive performance, on the other hand. The formation of collaborations is studied through co-inventor dyads and the inventive performance is measured by the patent forward citations, as a proxy for the value of inventions or patent quality (Trajtenberg, 1990; Harhoff, et al. 2003). We analyse the networks and proximity characteristics of co-inventor ties raising two questions. First, is there an optimal level of proximity, in the sense that too much or too little proximity may harm the invention process? Second, how the network position combined with proximity enable actors to increase the quality of their inventions.

We address these issues through the analysis of co-inventor networks in the field of genomics in Europe between 1990 and 2006. Our results provide partial support for the proximity paradox as geographical and organizational proximity play little role for explaining innovation performance. In contrast, the actors network position and technological

proximity are prominent.

The paper is organized as follows: Section 2 reviews the recent literature on the topic positing our contribution. Section 3 provides a description of data and an explanation of how networks have been built up. Section 4 describes the estimation design; results are discussed in section 5. Section 6 concludes.

2. Networks, proximity and the performance of innovation: the proximity paradox

In this section, we consider how network and proximity determinants, as well as their interplay may differ in explaining innovative activities and performance.

2.1. Geographical proximity and networks

Knowledge diffusion and innovation are known to be highly localised and embedded in industrial clusters, and a large literature has investigated localized knowledge externalities and their impact on knowledge creation. Since a decade or so, an intense debate has been going on to understand under what conditions individuals and firms benefit from these knowledge externalities. Besides the death of distance, a number of arguments have been advanced to show that, 1) knowledge is far from being in the air and freely accessible to all, and 2) permanent proximity may not be necessary (Torre and Rallet, 2005).

The first argument emphasizes the various channels through which knowledge diffusion occurs. A strong result is that individuals and firms need to be embedded in local networks

in order to benefit from knowledge diffusion, especially when knowledge is sticky, complex or tacit, because then, individuals need frequent interactions that facilitate communications, interactive learning and trust building. Although networks are non-geographical mechanisms, they strongly interact with it (Ter Wal, 2011). Two reasons for that: first, knowledge diffusion follows inter-personal channels such as labour mobility, professional acquaintances, co-ethnicity, friendship or kinship (Agrawal et al. 2006). Second, these knowledge flows are localized to the extent that individuals and inter-personal networks are also localized, essentially because individuals are not very mobile in space (Breschi and Lissoni, 2009). This contributes to explain why geography end-up playing little or even no role, once social networks are accounted for (Autant-Bernard et al. 2007; Maggioni et al., 2007).

Concerning the second argument of "non-permanent proximity", the role of networks is once again put forward. First, other forms of proximity, such as belonging to the same organization, may play similar roles to geographical proximity in sharing tacit knowledge and solving coordination problems. In a recent paper, Cassi and Plunket (2010) have indeed shown that organizational, social and geographical proximity endorse similar roles and act as substitutes in explaining the formation of co-inventor collaborations. In other words, organizational and social proximity compensate geographical distance when explaining technological collaborations. Second, individuals and firms need access to external and non-local sources of knowledge. In other words, the innovative performance also depends on the extent to which they are also embedded in a global network comprising more distant partners (Bathelt et al., 2004; Witthington et al. 2009).

Regarding the proximity paradox, geographical and organizational proximity play a major

role for the formation of collaboration and networks. Once these networks exist, they may play little or even no role for subsequent collaborations. Regarding the performance of innovation, the need to access external knowledge may limit the role of proximity and rather favour networks.

2.2. Network positions, knowledge flows and innovation

Once the role of networks has been accepted as being prominent, it may be that some positions within these networks are more favourable for accessing knowledge and for innovating. The relationship between network positions and performance is usually considered through closure and bridging positions (Coleman, 1988; Burt, 1992; Fleming et al., 2007; Baum, et al., 2012).

Closure positions refer to a situation in which actors are at least indirectly connected within the same sub-network component. When actors are embedded in networks in which they are densely connected, collaborations and knowledge flows are facilitated for the following reasons. First, individuals belong to the same community; they know each other, at least indirectly. This social proximity contributes to generate common language codes, behavioural routines and group norms that promote trust and collaboration (Coleman, 1988). Second, cohesive networks also facilitate communication between individuals, and exchanges of tacit and complex knowledge (Reagans and McEvily, 2003). For both of these reasons, increased cohesion solves in part coordination problems by reducing uncertainty as well as the risks of opportunism. Thus, it reduces the cost of coordination and increases the likelihood of collaboration.

As closure ties occur within a close and dense network of actors already indirectly linked,

they use redundant information paths within a close community of inventors with similar knowledge bases and technological skills. The proponents of cohesion argue that closure ties provide organisations with two main advantages (Fleming et al. 2007). First, the redundancy facilitates exchanges and mobilization of knowledge. Second, the cost of coordination is largely reduced since actors share a high level of trust. These two reasons explain why closure ties contribute to strengthen the innovation process and promote performance. However, being embedded in very dense and strongly cohesive networks may also harm individuals in their search of new knowledge and their learning processes. In fact, Burt (1992) argues that knowledge accessing is more efficient when individuals occupy structural holes that enable the link of unconnected actors. Individuals positioned in structural holes are able to broker knowledge flows across unconnected groups (e.g. Gargiulo and Benassi, 2000).

Bridging ties provide a brokerage position by connecting actors that have no network link in common since they belong to separate network components. These ties allow establishing a channel across clusters, which facilitates access to non-redundant and even novel sources of information and knowledge (Burt, 1992). This is even reinforced when bridging ties enable actors to access distant clusters endowed with diverse sources of knowledge and resources (Bathelt, et al. 2005).

The impact of bridging ties on innovation performance is also ambiguous. On the one hand, bridging positions are inherently “weak” and fragile ties, since they link together unconnected actors that are more difficult to mobilize and coordinate (Fleming et al. 2007). However, this disadvantage may be compensated through some form of organizational proximity, that is, a governance structure that reduces uncertainty and favours coordination

in the innovation process (Boschma, 2005). On the other hand, bridging ties promote creativity and provide opportunities for novel combination and recombination of ideas (Obstfeld, 2005; Fleming et al. 2007), which might increase the performance of innovation.

In sum, although closure and bridging ties play different roles, they may both promote innovative performance depending on the type of knowledge created and, as will be discussed now, on the optimal level of technological proximity (Nooteboom, et al. 2007, Rowley et al. 2000, Sorenson, et al. 2006).

2.3. Technological proximity and innovative performance

When it comes to knowledge diffusion and especially creation, technological or cognitive proximity plays a key role compared to other forms of proximity, such as social, organizational or geographical boundaries, which aim is primarily that of coordination and learning "facilitator" (Broekel and Boschma, 2011).

Technological proximity means that actors share the same knowledge base or technology. Actors are more likely to collaborate when they have very similar knowledge bases, since it makes communication, learning processes and knowledge sharing easier. However, too much cognitive proximity may yield diminishing and even negative returns since the learning process may not be very rich if actors have a very high degree of technological overlap. As a result, there is an inverted-U relationship between technological proximity and the formation of collaborations (Mowery et al. 1998). A similar relationship is also expected when considering the innovative performance. Some degree of dissimilarity in the knowledge base is needed since resource heterogeneity provides an opportunity for learning and innovation. This is true up to a certain "optimal" level of cognitive distance since after

this point, uncertainty and complexity become too important for collaboration to be coordinated and managed (Nooteboom et al. 2007). In the specific case of the Dutch aviation industry, Broekel and Boschma (2011) do not find evidence of an inverted u-shape relationship between technological proximity and innovation performance. Instead, they find a true negative impact supporting the proximity paradox arguments.

However the relationship between technological proximity and innovation depends on the type of knowledge, and the underlying learning process, as well as the level of proximity between actors. When knowledge is complex, its diffusion is more complicated when actors are distant from a geographical, organizational or social point of view. Diffusion and recombination of knowledge is easier when knowledge is of moderate complexity and actors are close to some extent, organizationally or socially (Sorenson et al. 2006). Thus, when knowledge is complex, closure ties may play a more prominent role for promoting innovative performance as compared to bridging ties.

The industrial and technological environments and the underlying learning regimes do also moderate the relationship between technological proximity and innovative performance. If competition is driven by innovations based on exploitative learning, it is expected that innovative performance will be higher when actors are technologically close, and form closure ties (Rowley et al. 2000). In both of these situations, actors share similar knowledge bases, they may benefit from specialization effects and cumulative knowledge. However, technological proximity should not be too close either as already explained. In contrast, if innovation is rather based on exploration, actors may have higher innovative performances if they use different and non-redundant knowledge bases, which means technological proximity should not be too close and actors could benefit from forming bridging ties.

However as claimed by Rowley et al. (2000), closure and bridging ties may not necessarily be contradictory. It may indeed be the case that in some environments, actors may benefit from relying on both situations, closure and bridging ties at the same time, in order to be able to benefit and combine the advantages of exploitation and exploration.

We explore the role of various forms of proximity and network positions and their interaction in determining network tie formation (bridging or closure) and innovation performance, measured in terms of technological value of the patents (i.e. forward citations). We intend to test if these two sets of variables play similar or opposite roles in this context. Moreover, we try to provide some empirical support to the proximity paradox: proximity makes knowledge sharing easier but too much proximity could harm innovation performance.

3. Data and network formation method

The dataset under investigation is composed of all the genomic patents published at the European Patent Office between 1990 and 2006, with at least one inventor reporting a European postal address (EU15, Switzerland and Norway, see Appendix for number of patents and inventors distinguished by country of residence) and the citations received by those patents over their period 1994-2010 by other patents independently of the technological content.

The database was built during a recent research project carried out by ADIS-Paris Sud, LERECO-INRA and the OST – Observatoire des Sciences et des Techniques - supported by the French national research agency (ANR – Agence National pour la Recherche). The EPO

Worldwide Patent Statistical Database (PATSTAT)[1] was searched using a specific strategy involving genetics and genomics keywords in order to define the genomic field (Laurens, Zitt and Bassecoulard, 2010). "Genetics *stricto sensu* is the science of gene heredity and variation of organisms by looking at single genes […] in contrast, genomics typically looks at all the genes or at least at large fractions of a genome as a dynamic system, over time, to determine how they interact and influence biological pathways, networks and physiology, in a much more global sense" (ibid, p.649). A number of experts were asked to validate the lexical query for filtering genomics out of genetic and finally the field delineation and the border areas.

Our final database is a sub-sample of 12,968 patents filed by 4,406 distinct applicants and 24,708 inventors[2]. The data include all patents with at least one inventor reporting a European postal address (EU15, Switzerland and Norway) and their co-inventors. Every patent provides information on the inventors, their name and postal address, which enables to define their geographical location at the NUTS 3 level for European inventors and the geographical distance between them. The patent offers also information on applicants, for which we have determined whether they are private companies, research institutes and universities, non for-profit organizations or individuals. For each patent, we also know their IPC – International Patent Classification – codes, which identify their technological fields. We use all these information in order to define the inventor's individual characteristics such as geographical location, technological specialization and affiliation. The affiliation is in this case the organization for which the patent is filed and not necessarily the employer. For

[1] The original project used PATSTAT October 2007 version. The citation data refer to a more recent version of PATSTAT, April 2011.

[2] The disambiguation of inventors' homonymies has been dealt following the methodology proposed in Carayol and Cassi (2009).

instance, it may happen in a number of cases that academic inventors file patents for a private company instead of their own university, but we are not able to identify them.

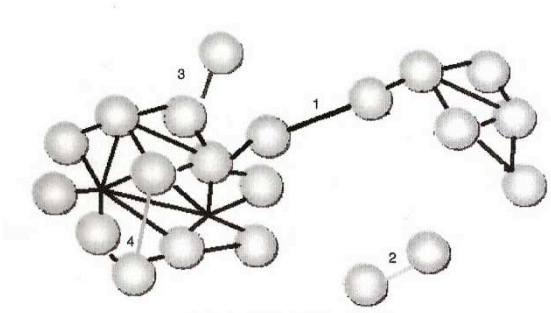
Concerning the citation data, we consolidate (received) citations using equivalent (Martinez, 2010): we consider only citations coming from other EPO patents and from patents of other offices with an equivalent European patent. Moreover, we identified and eliminated self-citations defined at the level of inventors, i.e. a citation has been considered as self-citation if the two patents have at least one inventor in common. For each patent belonging to the dataset, we have considered the number of citations received in a five-year window.

3.1. Network building and tie definitions

In order to build the network,[3] we assign a link (edge) between any two inventors (nodes) who file a patent together. The actors that co-patent form small components that increase over time and eventually connect to other components through new co-patenting activities. Networks may thus be described as bundles of actors that are connected but all the actors within a network are not necessarily linked. Networks evolve over time through the formation of dyads between co-inventors. Following Amburgey et al. (2008), we classify each new link according to the connectivity to the overall network. Four categories of links may be distinguished, as represented in Figure 1: (1) a link bridging two components; (2) a link determining the creation of a new component; (3) a pendant to an existing component; or (4) an intra-component link.

[3] Social Network Analysis computation has been programmed by the authors themselves with SAS. The SPAM modules developed by James Moody (2000) have been extremely helpful.

Figure 1. Type of network ties



Source: Fig. 10, Amburgey et al. (2008)

These new links are explained by the network structure and the inventors' individual characteristics. In order to avoid simultaneity biases, we consider all determinants with a lag of one period. For this reason, we may only investigate links among already active actors, that is, bridging and intra-component ties. Another reason for investigating these links comes from the specificity of patents as compared to publications (Fafchamps et al., 2010, Ponds et al., 2007); co-inventors of a given patent have, by definition, the same affiliation[4] and technological field (IPC codes). For this reason, this information cannot be used to highlight organizational or technological determinants.

Bridging and intra-component ties have very different consequences on network structure. Bridging allows for the linking of separate groups of inventors and establishing channels that facilitate the access to resources or other assets. Intra-component ties allow for the establishment of a direct link between actors already (indirectly) connected and the increase of cohesion. In order to consider the impact of cohesion on network formation and

[4] Even for industry-university collaborations, most of the time there is only one affiliation for a given patent, for this reason inventors of a given patent have the same affiliation even if the applicant designated in the patent does not employ them.

innovation, we focus on closure ties that occur within a close social proximity when geodesic distance is equal to 2 or 3. These ties represent 83% of all intra-component ties[5].

Finally, since ties may die out after a certain period of time, we use a five-year moving window to get a more realistic picture of the network for any given year. So, for instance, the network in 1994 is built up considering all the patents published between 1990 and 1994. Accordingly, an inventor is considered as active (e.g. in 1994), if she has at least one patent over the 1990-1994 period.

3.2. Proximity and networks: descriptive statistics

In this section, we briefly describe the characteristics of the co-inventor dyads in terms of proximity regarding geography, organizational arrangements (Table 1) and technology by the types of ties (Figure 1).

More than half of all inventors are located in the three largest European countries, that is, Germany, United Kingdom and France. Regarding geography, 86% of all ties occur within the same country and 34% within the same NUTS3 region. Closure ties occur more frequently within regions. When dyads occur across countries, they happen through bridging ties in 65%. Regarding organizational arrangements, 42% of all ties occur within the same organization. But, these are mainly closure ties in 70% of those cases. When collaboration happens outside the organization, they occur mainly between organizations with similar types, specifically between firms and through bridging ties.

Figure 1 illustrates the link between forward citations and geographical as well as technological distance by types of ties. We see that closure ties receive more citations than

[5] Among all intra-component ties, 66% occur within a geodesic distance of 2 (i.e., with the partner of one's partner), and 17% within a geodesic distance of 3.

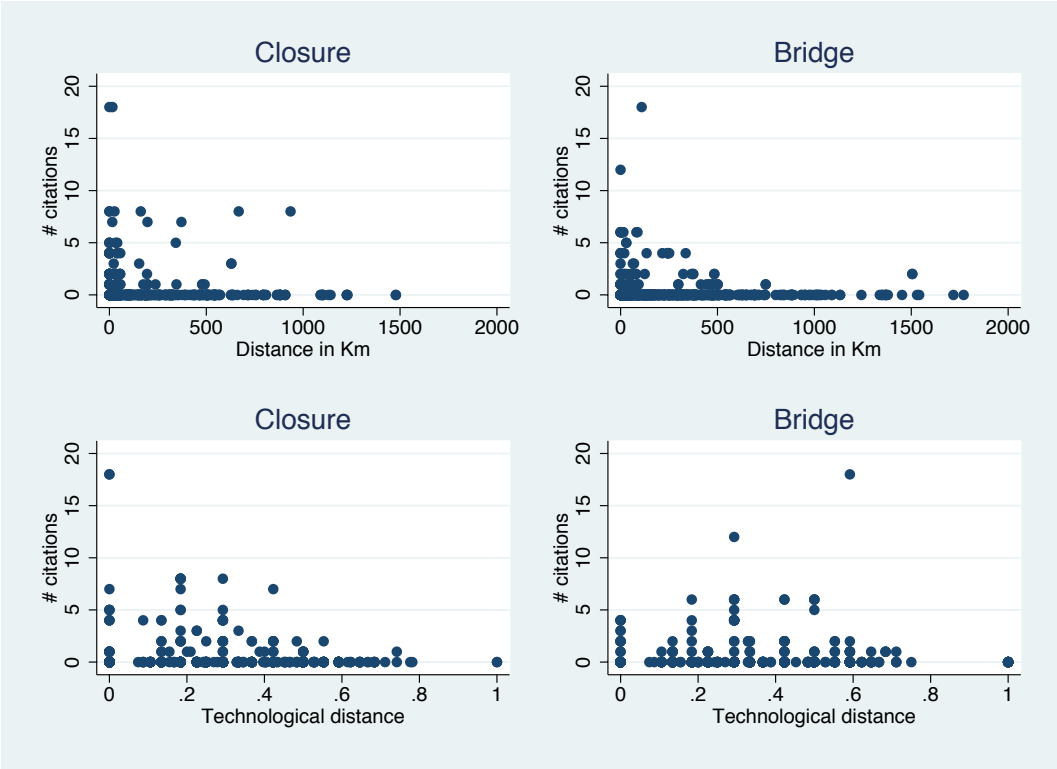
bridging ties, and that bridging ties are more geographically spread. Citations occur for a smaller technological distance for closure ties and larger technological distance for bridging ties.

Table 1. Geographical proximity and organizational arrangements

	Total	(1)	Bridge	(2)	Intra-component	(2)	Closure	(2)
Total	2133	1,00	1065	0,50	1068	0,50	885	0,41
Geographical proximity								
Same country	1825	0,86	864	0,47	961	0,53	796	0,44
Same region	725	0,34	311	0,43	414	0,57	348	0,48
Different country	308	0,14	201	0,65	107	0,35	89	0,29
Organizational arrangements								
Same organization	902	0,42	272	0,30	630	0,70	351	0,39
Same type	751	0,35	445	0,59	306	0,41	242	0,32
Between firms	679	0,32	413	0,61	266	0,39	202	0,30

Note: Number of realized dyads – distribution among types of ties (Intra-component + Bridge = all ties)
(1) Fraction of all dyads ; (2) Fraction of the total in rows by categories (geography and organization)

Figure 1. Citations given geographical and technological distance



In summary, and as expected bridging ties occur at higher geographical, technological and organizational distance than closure ties. Does it make a difference for tie formation and inventive performance? This is estimated in the following sections.

4. Estimation and variables

In order to explore the proximity paradox, we contrast the impact of proximity and network determinants, first, on the decision to collaborate (section 5.1.) and second, on the performance of this collaboration (section 5.2.). We use similar variables in both analyses to ease the comparison.

4.1. Explaining tie formation

In this first analysis, we estimate how prior network relations and various forms of proximity drive network tie formation. For two inventors i and j, we estimate the probability of forming a tie p_{ij}. If the tie is observed, the dependent variable takes the value of 1 and it is 0 otherwise. Three cases are considered whether we estimate the probability of forming a closure, a bridging or any type of tie as compared to not forming any tie. All realized and possible dyads between any two pairs of inventors represent more than ten million observations, that raise important difficulties of estimation. In order to handle the problem, we adopt a case-control approach: for any realized tie and its related co-inventors, we randomly select five possible but not realized co-inventors that have filed a patent in the same year as the observed tie, which provide five controls for each co-inventor. In summary, for each realized tie, we have ten controls. Since the proportion of ties in the sample (11%) is

much higher than the proportion of ties in the population (around 0.0022%), logistic regressions may be biased. For this reason, rare event logistic models may be more appropriate to estimate models based on a case-control design (King and Zeng, 2001; Sorenson et al., 2006)[6]. The strategy is to select all the "cases" for which the event is realized (p_{ij}=1, we observe a realized tie in the population as well as in the sample) and consider a random selection of controls (p_{ij} =0, the tie is potential but not realized). Using this sampling method, we know the fractions of ones in the population; in our case, we have 1065 bridging ties and 885 closure ties. To estimate the rare event logit, we implement the prior correction procedure, which involves computing the usual logistic regression and correcting the estimates using prior information about the fraction of ones in the population. In doing so it is possible to correct the estimation, taking in account the difference between the probability of a positive case observed in the sample and the *rarity* of the event actually observed in the population. In our case, we compute the fraction of ones in the population by dividing the number of realized ties by the number of potential ties.

For each regression, we include a full set of year dummies and we use a cluster robust procedure to adjust standard errors for intra-group correlation between realized ties and their controls.

4.2. Explaining inventive performance

The inventive performance is assessed through patent forward citations. The citations, received by other patents in subsequent years following application, provide a measure of the technological quality and economic value of a patent (Trajtenberg, 1990; Harhoff et al. 2003, Gambardella et al. 2008). We compute citations based on patent families without self-

[6] Rare event logit has been implemented through the ReLogit Stata routine proposed by Tomz (1999).

citations (Martinez, 2010). Since the dependent variable is a count measure, a Poisson process is used to estimate the model (Hausman et al., 1984). The basic Poisson process is written as:

$$\Pr\left[p_{it}\right] = f(p_{it}) = \exp(-\lambda_{it}(x_{it}))\left[\frac{\lambda_{it}^{p_{it}}(x_{it})}{p_{it}!}\right]$$

The underlying assumption of the Poisson process is that the probability as well as the variance of the number of events during a time period is equal to the rate λ_{it}, thus making the assumption that there is no heterogeneity in the sample. Since the variance often exceeds the mean and since such over-dispersion causes the standard errors of the parameters to be underestimated, the statistical significance levels are overstated. For this reason, the negative binomial model is often preferred (Cameron and Trivedi, 2005) as it is the case here since the dispersion parameter is highly significant. The model is fitted using a conservative Huber/White/sandwich robust variance estimator and controls for year fixed effects to capture the possible correlation of the dependent variable with omitted time-invariant variables. For the negative binomial we use only the realized patents, which leaves the sample with 885 observations (dyads) for closure ties and 1065 bridging ties.

4.3. Independent Variables

The network variables are tested through the impact of prior network ties. They are assessed through closure and bridging ties. The variable "*Closure"* estimates the impact of social proximity, when the geodesic distance prior to collaboration is equal to 2 or 3. It should increase the likelihood of collaboration since inventors collaborate more easily with their partners' partners because "knowing" them facilitates trust and collaboration, and

favour inventiveness. It should also ease innovation output since closure facilitates the use of complex and tacit knowledge. The 1,950 dyads are part of inventor teams that have generated 922 patents. Among those, 44 patents (and 223 dyads) are composed of teams that combine both closure and bridging ties. We introduce a dummy "*Closure and Bridging" patent* to estimate the fact that combining cohesion and brokerage may yield higher returns (Rowley et al. 2000).

The proximity variables are tested using geographical, technological and organizational proximity. In order to calculate the "*geographical proximity"* in kilometres, we locate each inventor at the NUTS 3 level based on its postal address. All European inventors are identified this way; the non-European inventors have been dropped from the regressions. The distance is calculated using the latitude and longitude coordinates of each NUTS 3 centroïd[7]. Geographical proximity is thought to have a positive impact on the likelihood of forming a tie and the innovative performance since frequent interaction decreases transaction costs, and enables to invest in more complex and productive learning processes.

Collaboration is easier among inventors that share similar technological interests and specializations. For this reason, we suppose that "*technological distance"* decreases the likelihood of collaboration. It is computed as the complement of the Jaffe's (1989) index$^{t}_{ij}$, which is a proximity measure ranging between zero and one, depending on the degree of overlap between the co-inventors' prior patent IPC codes.

[7] We adjust the latitude and longitude coordinates for the earth curvature; thus the distance in km between two points A and B is computed as:

d(A,B) = 6371 × arccos[sin(latitude(A)) × sin(latitude(B)) + cos(latitude(A)) × cos(latitude(B)) × cos(|longitude(A) – longitude(B) |)]

$$t_{ij} = \frac{\sum_{k=1}^{K} f_{ik} f_{jk}}{\sqrt{\sum_{k=1}^{K} f_{ik}^2 \sum_{k=1}^{K} f_{jk}^2}}$$

f_{ik} and f_{jk} represent each inventor *i* and *j* technological position. Innovative performance is an inverted u-shape of technological distance since some technological distance is needed to go beyond 'local search' but too much proximity may limit communication, knowledge exchange and recombination (Gilsing et al. 2008).

We then consider the impact of organizational proximity. Organizational proximity occurs when two inventors file a patent for the same applicant. When inventors file a patent for different organizations, two inventors may work for similar types of organizations, either among academia and public research centres or among private companies (Ponds et al. 2007). We suppose that inventors are more likely to form ties within their own organizational boundaries or with inventors belonging to similar organizational types. In order to account for different organizational settings, we consider three occurrences: "Same applicant" takes the value of 1 when inventors have patented for the same organization prior to tie formation and 0 otherwise; "Same type" takes the value of 1 when inventors have patented for the same organizational type (firms or companies) and 0 otherwise; and "Different type, different applicant" as the last occurrence, in our case university – industry relationships.

Network and proximity interaction variables. We interact closure with geographical proximity and technological distance in order to test if they may have substitutable or complementary impacts on network tie formation and performance. Our hypothesis is that inventors will choose closure ties when they require similar competences that may be found

in a close neighbourhood. They will choose bridging links when they need distinct skills that may not be found in their own environments. Closure and geography should be substitutes since they endorse similar roles in knowledge communication (Cassi and Plunket, 2010). The sign should therefore be negative. Closure should also moderate the negative impact of technological distance.

Control variables

A number of control variables are introduced in the regressions. First, we account for preferential attachment based on the *degree centrality measure*. Since the study considers the likelihood of two inventors in forming a tie, we must examine this measure for both inventors and consider the average $\bar{n}_{ij}$ and the difference Δn_{ij} of both inventors' degrees (Fafchamps et al. 2010).

$$\bar{n}_{ij} = \frac{(n_i + n_j)}{2}$$

$$\Delta n_{ij} = \left| n_i - n_j \right|$$

For each type of tie, we expect a different sign. In particular, we expect the average measure to be positive and the difference to be negative for closure ties and *vice versa* for bridging ties. When actors belong to the same sub-network, individuals tend to link to partners similar to themselves in terms of degree: thus the difference in the number of partners should tend to zero. This is even more important for individuals with a greater number of collaborations since they are more visible within the network. When individuals are searching for an effective collaboration that enables them to access new and different resources, it is likely that similarity is less important or even plays a negative role. In this

case, a greater difference would have a positive effect on tie formation and, consequently, we should expect a negative effect of the average degree as well.

We introduce other types of controls. First, we control for the distinction between domestic and foreign inventors' collaboration. Since being a foreigner is strongly correlated with geographical distance, we prefer to consider the specific case of foreigners located in *border* countries by introducing a dummy for inventors located in two countries with a common frontier.

Second, we also consider the team size through the *number of inventors* involved in research determining the new link. We expect a positive impact on the performance, since the number of inventors could be interpreted a proxy of resources invested.

Finally, we also consider the number of years since the first tie in order to control for experience with the patent process. Again, in order to account for the symmetric relation, we introduce the difference and average value of both inventors' experiences, namely *Experience – absolute difference* and *Experience – average difference.*

5. Estimation results and discussion

Table 2 displays the results for the rare event logit where the dependent variable is the tie formation. Table 3 reports the results explaining the performance measured by the number of forward citations (in a five-years window) estimated with a negative binomial.

Collaboration ties formation (Table 2).

Models 1-3 demonstrate the impact of relational and proximity variables on the likelihood of forming closure ties, models 4-6 test the same variables on bridging ties. The last three models (7-9) consider the two types of ties jointly and introduce a dummy for closure ties for allowing comparison. Across models, variables and controls remain overall consistent in sign and magnitude, suggesting that they are rather robust to the introduction of additional variables.

Regarding the *proximity mechanisms*, all the sources of similarity impact collaborations, as expected. The likelihood of forming a tie is larger when co-inventors share similar technological fields and work within close spatial distance. Organizational proximity is also strongly significant and positive; the likelihood of forming a tie increases when inventors patent for the same applicant, even in the case of bridging ties. This confirms the fact that inventors patent first of all with individuals that belong to their own organization (Singh, 2005).

In summary, collaborations mainly occur when inventors are located in close geographical distance to each other, work in similar technological areas and presumably patent for the same organization. This is true for both types of ties. However some difference should be emphasized. First technological distance plays a less important role for bridging ties: these collaborations enable to connect individuals that are more far away in terms of knowledge features as will be seen in the next section. Also the organisational proximity effect is different: it plays a greater role for closure ties (i.e. *same type* variable). Moreover, the model 9 shows that forming a closure tie is extremely easier relative to bridging ties,

especially because social distance is complementary to geographical proximity.

Inventive performance: the patent quality impact (Table 3).

As in the previous table, models 1-3 demonstrate the impact of relational and a proximity variable on the quality of patents when closure ties are involved. Models 4-6 test the same variables on bridging ties. The last three models (7-9) consider the two types of ties jointly and introduce a dummy for closure ties for allowing comparison. Across models, variables and controls remain consistent overall in sign and magnitude, suggesting that they are rather robust to the introduction of additional variables.

Comparing closure (model 1-3) to bridging ties (model4-6) show clear differences. For closure ties, the innovative performance is larger when the patent collaboration also involves bridging ties (i.e. variable *Closure and Bridging patent)*. The opposite is no true in the case of bridging ties (the variable is not significant across models 4-6). Otherwise, the only other variable with positive effect is the average number of partners of each individual especially if the difference between the two is not huge. Technological distance is negative: if two individuals, already indirectly connected, decide to start this collaboration, the value of their output decreases with increasing technological distance.

We have a completely different picture for bridging ties. Organisational proximity plays a very important role: when trust and confidence is established institutionally, the collaboration is able to get better results. Moreover, concerning technological distance we get an inverted U-shaped relationship. It is important to identify the optimal distance in terms of knowledge between the individuals involved.

Tablec2. Rare event logit – Tie formation

VARIABLES	(1) Closure	(2) Closure	(3) Closure	(4) Bridge	(5) Bridge	(6) Bridge	(7) All	(8) All	(9) All
Closure							3.098***	0.386	0.474
							[22.32]	[1.55]	[1.30]
Closure x geographical proximity								-0.580***	-0.548***
								[-13.49]	[-7.13]
Closure x technological distance								1.084+	0.729
								[1.90]	[0.69]
Closure x technological distance x geographical proximity									-0.119
									[-0.51]
Geographical proximity	0.590***	0.590***	0.627***	0.657***	0.656***	0.641***	0.562***	0.645***	0.652***
	[22.68]	[22.67]	[11.46]	[25.68]	[25.67]	[13.58]	[30.63]	[32.61]	[17.56]
Technological distance	-2.742***	-2.855***	-3.455**	-1.707***	-2.094***	-1.847*	-2.126***	-2.390***	-2.506***
	[-9.54]	[-3.95]	[-3.19]	[-8.63]	[-4.02]	[-2.10]	[-4.74]	[-5.08]	[-3.57]
Technological distance sq		0.191	0.309		0.547	0.525	0.645	0.814	0.824
		[0.18]	[0.30]		[0.82]	[0.77]	[1.09]	[1.31]	[1.33]
Geographical x technological distance			-0.119			0.046			-0.023
			[-0.84]			[0.40]			[-0.27]
Same applicant	3.321***	3.319***	3.327***	2.352***	2.345***	2.344***	1.566***	1.483***	1.487***
	[18.21]	[18.20]	[18.00]	[11.12]	[11.10]	[11.11]	[10.43]	[9.69]	[9.68]
Same type	0.020	0.020	0.019	-0.117	-0.118	-0.118	-0.179**	-0.141*	-0.142*
	[0.15]	[0.15]	[0.14]	[-1.52]	[-1.54]	[-1.54]	[-2.68]	[-2.01]	[-2.02]
Degrees - Avrg	0.765***	0.765***	0.767***	-0.475***	-0.472***	-0.471***	-0.613***	-0.637***	-0.636***
	[5.61]	[5.61]	[5.60]	[-4.32]	[-4.29]	[-4.30]	[-6.87]	[-6.95]	[-6.94]
Degrees - Abs.diff.	-0.191*	-0.191*	-0.191*	0.286***	0.285***	0.285***	0.192**	0.199**	0.199**
	[-2.35]	[-2.34]	[-2.34]	[3.84]	[3.83]	[3.82]	[3.16]	[3.25]	[3.25]
Border	-1.388***	-1.389***	-1.383***	-1.121***	-1.121***	-1.121***	-1.094***	-1.099***	-1.097***
	[-8.15]	[-8.15]	[-8.13]	[-9.95]	[-9.95]	[-9.97]	[-10.94]	[-10.17]	[-10.17]
Experience - Abs.diff	-0.243+	-0.243+	-0.243+	0.269**	0.271**	0.272**	0.335***	0.342***	0.341***
	[-1.78]	[-1.77]	[-1.76]	[2.91]	[2.94]	[2.95]	[3.98]	[3.95]	[3.93]
Experience - Avrg	-0.119	-0.119	-0.118	-0.152*	-0.151*	-0.150*	-0.126*	-0.124*	-0.124*
	[-1.27]	[-1.27]	[-1.25]	[-2.18]	[-2.17]	[-2.16]	[-2.09]	[-2.01]	[-2.00]
# inventors per patent	-0.030	-0.032	-0.029	-1.577***	-1.580***	-1.579***	-1.190***	-1.127***	-1.127***
	[-0.07]	[-0.07]	[-0.07]	[-3.37]	[-3.38]	[-3.38]	[-3.33]	[-3.51]	[-3.51]
Constant	-8.624***	-8.612***	-8.443***	-3.849***	-3.810***	-3.896***	-4.581***	-4.253***	-4.218***
	[-9.60]	[-9.53]	[-8.86]	[-4.20]	[-4.15]	[-4.17]	[-6.56]	[-6.68]	[-6.41]
Observations	9,569	9,569	9,569	11,656	11,656	11,656	21,225	21,225	21,225
D.F.	20	21	22	20	21	22	22	24	26

Robust z-statistics in brackets *** p<0.001, ** p<0.01, * p<0.05, + p<0.1

Table 3. Negative binomial – Forward citations

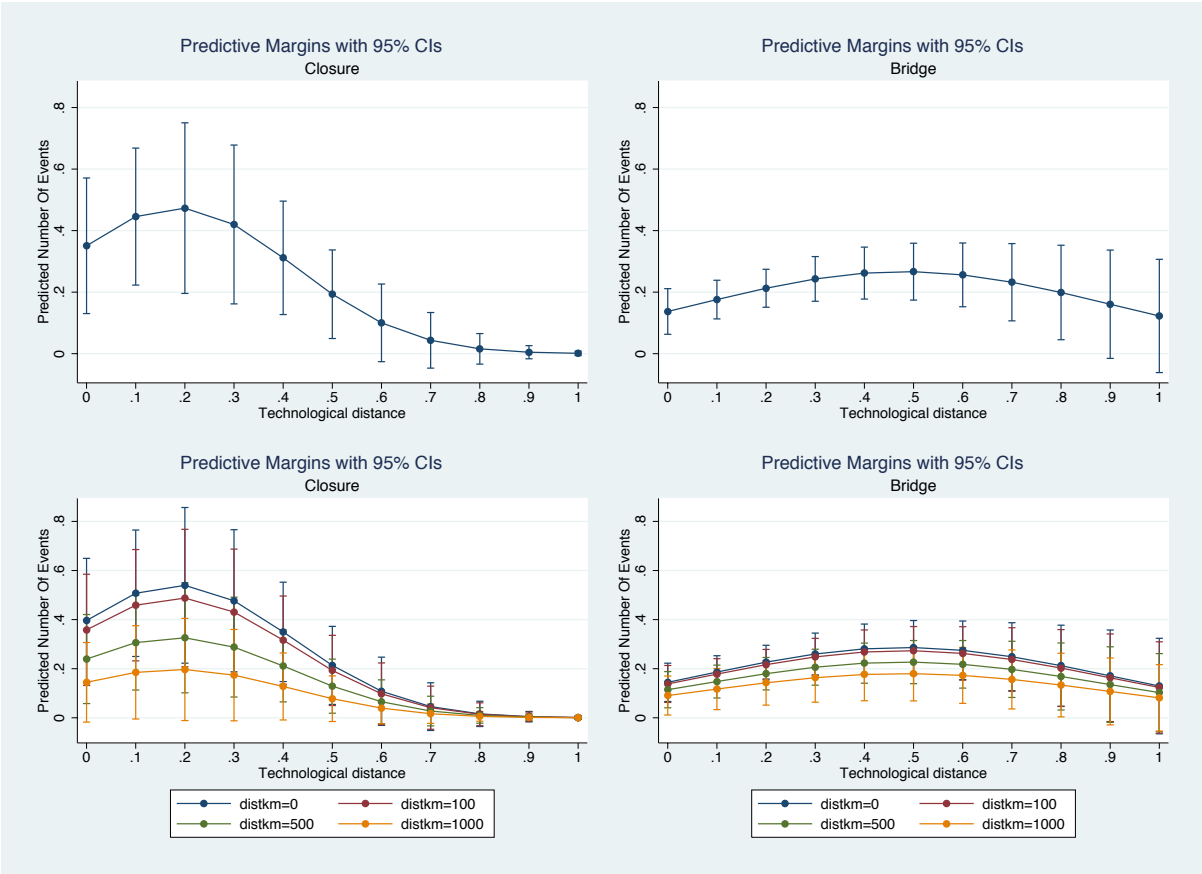
VARIABLES	(1) Closure	(2) Closure	(3) Closure	(4) Bridge	(5) Bridge	(6) Bridge	(7) All	(8) All	(9) All
Closure and Bridging patent	1.874***	1.794***	1.748***	0.269	0.231	0.257	0.688*	0.721*	0.743**
	[4.22]	[4.16]	[4.10]	[0.78]	[0.68]	[0.75]	[2.33]	[2.51]	[2.59]
Closure							-0.041	0.743+	1.575**
							[-0.19]	[1.88]	[3.23]
Closure x geographical proximity								-0.011	0.265*
								[-0.15]	[2.22]
Closure x technological distance								-3.069**	-6.344***
								[-3.08]	[-4.23]
Closure x technological distance x geographical proximity									-1.119**
									[-3.04]
Geographical proximity	0.036	0.035	0.190*	-0.050	-0.050	-0.176+	0.016	0.011	-0.064
	[0.69]	[0.68]	[2.13]	[-0.91]	[-0.91]	[-1.81]	[0.42]	[0.20]	[-0.69]
Technological distance	-1.531*	3.288	2.364	0.943+	2.770+	4.602**	1.961	3.666*	5.066**
	[-2.19]	[1.12]	[0.79]	[1.70]	[1.86]	[2.96]	[1.60]	[2.55]	[3.02]
Technological distance sq		-8.961	-10.760+		-2.882	-3.620*	-3.833*	-4.647*	-5.605**
		[-1.58]	[-1.75]		[-1.53]	[-2.02]	[-2.23]	[-2.49]	[-3.03]
Geographical x technological distance			-0.703*			0.394			0.232
			[-2.37]			[1.64]			[0.98]
Same applicant	0.047	0.129	0.148	0.913*	0.921*	0.967**	0.528*	0.558*	0.582*
	[0.13]	[0.36]	[0.42]	[2.50]	[2.54]	[2.73]	[2.00]	[2.10]	[2.23]
Same type	0.455	0.438	0.413	0.897**	0.942**	0.919**	0.958***	0.998***	0.975***
	[1.14]	[1.10]	[1.08]	[2.93]	[3.03]	[2.96]	[3.93]	[4.08]	[4.02]
Degrees - Avrg	1.551***	1.610***	1.548***	0.155	0.112	-0.002	0.873***	0.818***	0.706**
	[3.75]	[3.68]	[3.57]	[0.41]	[0.29]	[-0.01]	[3.53]	[3.36]	[2.88]
Degrees - Abs.diff.	-0.548**	-0.569**	-0.569**	-0.201	-0.193	-0.167	-0.404**	-0.374**	-0.368**
	[-3.01]	[-3.08]	[-3.16]	[-0.97]	[-0.93]	[-0.81]	[-3.19]	[-2.96]	[-2.95]
Border	0.543	0.588	0.490	-0.858+	-0.880+	-0.951*	-0.806*	-0.812*	-0.841*
	[0.79]	[0.82]	[0.79]	[-1.84]	[-1.91]	[-2.06]	[-2.26]	[-2.21]	[-2.43]
Experience - Abs.diff	-0.437	-0.551	-0.550	-0.177	-0.167	-0.130	-0.232	-0.103	-0.082
	[-1.24]	[-1.49]	[-1.45]	[-0.56]	[-0.53]	[-0.40]	[-0.96]	[-0.43]	[-0.33]
Experience - Avrg	-0.130	-0.084	-0.104	-0.031	-0.075	-0.081	-0.144	-0.175	-0.192
	[-0.73]	[-0.48]	[-0.60]	[-0.18]	[-0.44]	[-0.48]	[-1.08]	[-1.33]	[-1.47]
# inventors per patent	-1.245**	-1.261**	-1.230**	-0.673**	-0.628*	-0.663**	-0.957***	-0.958***	-0.930***
	[-2.78]	[-2.76]	[-2.68]	[-2.63]	[-2.50]	[-2.60]	[-3.90]	[-4.01]	[-3.90]
Constant	-2.023*	-2.329*	-2.014+	-2.882**	-3.013**	-3.353***	-2.503***	-2.946***	-3.225***
	[-1.97]	[-2.24]	[-1.94]	[-2.97]	[-3.07]	[-3.51]	[-3.30]	[-3.77]	[-4.08]
Dispersion parameter	2.409***	2.352***	2.317***	2.525***	2.517***	2.525***	2.640***	2.613***	2.598***
	[15.70]	[14.49]	[13.74]	[17.56]	[17.39]	[17.57]	[25.45]	[25.32]	[25.01]
Observations	885	885	885	1,065	1,065	1,065	1,950	1,950	1,950
Log Likelihood	-380.3	-378.5	-377.1	-427.5	-427.0	-426.4	-824.4	-821.4	-819.0
D.F.	21	22	23	21	22	23	23	25	27
Chi2	126.9	118.7	115.2	77.84	78.96	104.6	137.6	161.8	180.1

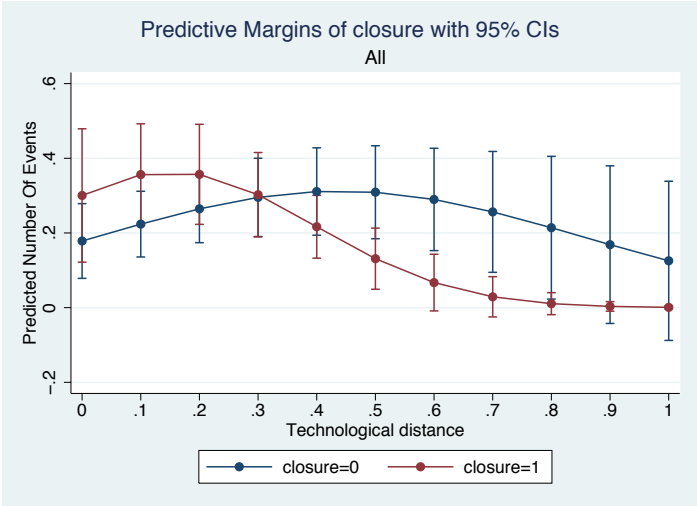
Robust z-statistics in brackets *** p<0.001, ** p<0.01, * p<0.05, + p<0.1

Thus bridging ties allow effective collaboration at some degree of distance. Last but not least, geographical proximity does not play any role. Models 7-9 confirm the result of this comparison. Figure 2 enable to visualize the impact of technological distance in different settings.

Figure2: Predicting patent forward citations

Impact of technological distance given geography and networks





These figures show the predicted number of forward citations estimated from the regressions with confidence intervals for Closure, Bridging regressions and All (closure and bridging ties combined). First, we see that the optimal level of technological distance is lower for closure (around .2) than for bridging ties (around .5) (first two panels). This confirms our understanding of bridging ties that enable collaborations to combine different knowledge bases as opposed to closure ties that occur within a high technological overlap. Although geographical proximity is not significant as a variable in the regressions, we see that geographical proximity enables nevertheless to yield higher levels of performance at least for closure ties; the difference is less striking for bridging ties.

Finally, we see with the last panel, that closure ties lead to higher performance when technological distance is small (.3). When the need for technological distance increases, bridging ties lead to higher performance in terms of forward citations. This confirms the important difference between both types of ties.

6. Conclusions

The aim of the paper is to investigate how the network and proximity characteristics of co-inventors enable them to access different sources of knowledge in different geographical and organizational contexts, and finally affect the quality of inventive collaboration. Doing so, we are able compare the effect of the proximity and network variables in establishing collaboration between individuals and the quality of this collaboration. The main conclusion is that both kinds of variables play a different role in the two contexts: what facilitates collaboration does not imply a better collaboration.

First, proximity variables, in every declination tested - geographical, technological and organisational – have a positive impact on establishing collaboration. This is true, even if with some difference, for both kinds of ties. Second, when the quality of the collaboration output is considered, geographical distance is not significant.

Third, there are main differences if a closure tie or a bridging tie is involved. In the former case, only technological proximity is playing a positive role. That is not the case for bridging ties that are able to manage effective collaborations at an optimal technological distance and if the institutional set is clearly defined, as in intra-organisation collaborations or between individuals sharing the same set of incentives and rules.

A number of limitations must be raised. Considering the impact of collaboration on innovative performance through co-inventor dyads reduces the scope of the study, since it does not enable to easily consider the network position and characteristics of the organizations in which they invent. The performance of an invention is not only due to the inventor characteristics and abilities, it also depends on their respective organisations.

Finally, the dyad approach does not enable us to consider the innovative performance at the team level rather than at the dyad level. This limits also the scope of the results since we are left with a dichotomous approach that does not consider all the complexity of the collaboration.

References

Agrawal, A. & Kapur, D. & McHale, J., (2008), How do spatial and social proximity influence knowledge flows? Evidence from patent data, *Journal of Urban Economics*, 64(2), 258-269.

Audretsch, D. and Feldman, M. (1996), R&D Spillovers and the Geography of Innovation and Production, *American Economic Review,* **86**, 630-640.

Autant-Bernard C., Billand P., Frachisse D. and Massard N. (2007), Social distance versus spatial distance in R&D cooperation: Empirical evidence from European collaboration choices in micro and nanotechnologies, *Papers in Regional Science*, 86(3), pages 495-519

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

Baum J.A.C, McEvily B., and Rowley T.J. (2012), Better with Age? Tie Longevity and the Performance Implications of Bridging and Closure, *Organization Science,* 23:529-546

Boschma, R. A. (2005): Proximity and Innovation: A Critical Assessment. Regional Studies, 39(1), 61–74.

Boschma, R. A. and Broekel, T. (2009), Knowledge networks in the Dutch aviation industry: the proximity paradox. Papers in Evolutionary Economic Geography# 09.15.

Broekel T. and Boscham R. (2011), The cognitive and geographical composition of ego-networks of firms – and how they impact on their innovation performance, Papers in Evolutionary Economic Geography # 11.18

Boschma, R. A. and ter Wal, A. L. J. (2007): Knowledge Networks and Innovative Performance in an Industrial District: The Case of a Footwear District in the South of Italy. Industry and Innovation, 14(2), 177–199.

Boschma, R.A. and Frenken, K. (2009): The Spatial Evolution of Innovation Networks. A Proximity Perspective. In: Boschma, R.A., Martin, R. (Eds.), Handbook of Evolutionary Economic Geography, Cheltenham: Edward Elgar, chapter 5.

Breschi, S. and Lissoni, F. (2009), Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows, *Journal of Economic Geography* **9**(4), 439-468.

Broekel, T. and Boschma, R. (2011), Knowledge Networks in the Dutch Aviation Industry - The Proximity Paradox, *Journal of Economic Geography*, (forthcoming)

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

Carayol N. and Cassi L. (2009), Who's Who in Patents. A Bayesian approach, Université paris 1 Panthéon-Sorbonne, hal-00631750, HAL.

Cassi L. and Plunket A. (2010), The determinants of co-inventor tie formation: proximity and network dynamics, Papers in Evolutionary Economic Geography (PEEG), 1015

Cameron, A. C. and Trivedi, P. K. (2005), *Microeconometrics: Methods and Applications*, Cambridge University Press, N. Y

Czarnitzki D., Hussinger K. Schneider C. (2011), Commercializing academic research: the quality of faculty patenting *Industrial and Corporate Change*, 20(5): 1403-1

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

Coleman, J. S. (1988), Social Capital in the Creation of Human Capital, *The American Journal of Sociology* **94**, S95--S120.

Fafchamps, M., Goyal, S. and van der Leij, M. J. (2010), Matching and Network Effects, *Journal of the European Economic Association* **8**(1), 203—231

Fleming, L., Mingo S. and Chen D. (2007), Collaborative brokerage, generative creativity, and creative success. *Administrative Science Quarterly* 52(3): 443-475.

Fornahl, D., Broekel, T. & Boschma, R. (2011), What drives patent performance of German biotech firms? The impact of R&D subsidies, knowledge networks and their location, Papers in Regional Science, 395-418.

Gambardella, A., Harhoff, D., Verspagen, B. (2008). The Value of European Patents, European Management Review

Gargiulo, M. and Benassi, M. (2000), Trapped in Your Own Net? Network Cohesion, Structural Holes, and the Adaptation of Social Capital, *Organizational Science* **11**(2), 183-196.

Jaffe, A. B. (1989), Real Effects of Academic Research, *The American Economic Review* **79**(5), 957-970.

King, G., Zeng, L., 2001. Logistic regression in rare events data. *Political Analysis* 9, 137–163

Harhoff, D., Scherer, F.M., Vopel, K. (2003). Citations, Family Size, Opposition and the Value of Patent Rights - Evidence from Germany. Research Policy 32: 1343-1363.

Laurens, P., Zitt, M. and Bassecoulard, E. (2010), Delineation of the genomics field by hybrid citation-lexical methods: interaction with experts and validation process, *Scientometrics* **82**(3), 647-662.

Maggioni, M. A. Nosvelli, M. and Uberti, T.E. (2007), Space versus networks in the geography of innovation: A European analysis, *Papers in Regional Science*, 86(3), 471-493

Martinez, C. (2010), "Insight into Different Types of Patent Families", *OECD Science, Technology and Industry Working Papers*, 2010/02, OECD Publishing

Mowery D.C., Oxley J.E. and Silverman B.S. (1998), Technological overlap and interfirm cooperation: implications for the resource-based view of the firm, *Research Policy*, 27 (5) 507-523

Moody, J. (2000), *SPAN: SAS Programs for Analyzing Networks.* University of North Carolina at Chapel Hill

Nooteboom B., Van Haverbeke W., Duysters G., Gilsing V. van den Oord A (2007), Optimal cognitive distance and absorptive capacity, *Research Policy*, 36 (7),

Obstfeld D. (2005), Social Networks, the Tertius Iungens Orientation, and Involvement in Innovation, *Administrative Science Quarterly*, 50, 100-130

Ponds, R., Van Oort, F. and Frenken, K. (2007), The geographical and institutional proximity of research collaboration, *Papers in Regional Science* **86**(3), 423-443.

Reagans R. and McEvily B . (2003) Network Structure and Knowledge Transfer: The Effects of

Cohesion and Range, *Administrative Science Quarterly*, 48, 240-267

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

Singh, J. (2005), Collaborative Networks as Determinants of Knowledge Diffusion Patterns, *Management Science* **51**(5), 756-770

Sorenson O., Rivkin J.W. and Fleming L. (2006), Complexity, networks and knowledge flows, *Research Policy*, 35, 994-1017

Ter Wal A. (2011), The Dynamics of the Inventor Network in German Biotechnology: Geographical proximity and Triadic Closure, Papers in Evolutionary Economic Geography, 11.02

Tomz, M., 1999. Relogit (Stata ado file). Available at http://www.stanford.edu/~tomz/software/software.shtml

Torre, A. and Rallet, A. (2005), Proximity and localization, *Regional Studies* **39(1)**, 47-59.

Trajtenberg, M. (1990). A Penny for Your Quotes: Patent Citations and the Value of Innovations. RAND Journal of Economics 21(1): 172-187.

Wittington K.B., Owen-Smith J. and Powell W.W. (2009), Networks, Propinquity, and Innovation in Knowledge-intensive industries, Administrative Science Quarterly, 54 :90-122

Appendix

A1. Inventors and patent by country (fractional and cumulative counting)

Country	Patent		Inventors	
	fractional	*cumulative*	*fractional*	*cumulative*
AT	213,26	844	419,83	427
BE	612,17	1983	897	908
CH	467,13	1554	871,33	896
DE	3630,55	13879	5717,75	5776
DK	510,83	2193	865,75	874
ES	185,92	818	603	606
FI	168,36	662	402	405
FR	1816,37	6407	3456,83	3491
GR	17,83	45	35	35
IE	57,40	215	118	119
IT	358,17	1193	767,67	772
LU	0,25	1	1	1
NL	845,38	2797	1341,42	1357
NO	105,83	299	200	201
PT	9,19	38	32	32
SE	426,50	1303	758,33	767
UK	2241,41	6773	3639,50	3687
others	1301,46	*7321*	4252,58	*4354*
Total	12968	48325	24379	24708

A2. Inventors: total, active and size of giant and 2nd largest components

Year	Total	Active Inventor	Size of GC		Size of 2nd largest component		
			Inventor	As % of active inventor	Inventor	As % of active inventor	2nd comp/GC
1994	2944	2853	284	9,95	229	8,03	80,63
1995	3840	3294	181	5,49	140	4,25	77,35
1996	4740	3766	168	4,46	124	3,29	73,81
1997	5867	4488	189	4,21	115	2,56	60,85
1998	7081	5254	194	3,69	135	2,57	69,59
1999	8400	6139	274	4,46	210	3,42	76,64
2000	10580	7607	388	5,10	257	3,38	66,24
2001	12780	9147	968	10,58	687	7,51	70,97
2002	15202	10629	1065	10,02	1013	9,53	95,12
2003	17607	12096	1590	13,14	667	5,51	41,95
2004	20083	13541	1465	10,82	515	3,80	35,15
2005	22499	14172	1682	11,87	376	2,65	22,35
2006	24379	14190	1101	7,76	480	3,38	43,60

A3. Descriptive statistics

Variables	Mean	SD	Min	Max
# citations	0.24	1.14	0.00	18.00
Geographical proximity (log)	-3.01	2.48	-7.99	0.00
Distance in Km	154.62	271.52	0.00	2944.23
Technological distance	0.28	0.21	0.00	1.00
Technological distance sq	0.12	0.15	0.00	1.00
Same applicant	0.42	0.49	0.00	1.00
Same type	0.35	0.48	0.00	1.00
Degrees – Avrg (log)	1.97	0.53	0.69	3.60
Degrees - Abs.diff. (log)	1.62	0.85	0.00	3.95
Border	0.06	0.23	0.00	1.00
Experience - Abs.diff (log)	1.57	0.48	0.69	2.74
Experience – Avrg (log)	1.18	0.75	0.00	2.77
# inventors per patent (log)	1.93	0.42	1.10	3.85

Note : 2133 observations for the realized dyads

	Bridging ties				Closure ties			
	Mean	SD	Min	Max	Mean	SD	Min	Max
# citations	0.22	1.03	0.00	18.00	0.29	1.33	0.00	18.00
Geographical proximity	-3.37	2.47	-7.99	0.00	-2.67	2.46	-7.30	0.00
Distance in Km	185.47	301.41	0.00	2944.23	127.18	237.82	0.00	1478.30
Technological distance	0.29	0.21	0.00	1.00	0.26	0.20	0.00	1.00
Technological distance sq	0.13	0.16	0.00	1.00	0.11	0.13	0.00	1.00
Same applicant	0.26	0.44	0.00	1.00	0.60	0.49	0.00	1.00
Same type	0.42	0.49	0.00	1.00	0.27	0.45	0.00	1.00
Degrees - Avrg	1.82	0.49	0.69	3.60	2.16	0.52	0.69	3.51
Degrees - Abs.diff.	1.53	0.82	0.00	3.93	1.72	0.87	0.00	3.95
Border	0.09	0.28	0.00	1.00	0.04	0.19	0.00	1.00
Experience - Abs.diff	1.61	0.45	0.69	2.74	1.53	0.49	0.69	2.71
Experience - Avrg	1.22	0.73	0.00	2.77	1.14	0.77	0.00	2.71
# inventors per patent	1.91	0.46	1.10	3.85	1.95	0.38	1.10	2.94

Note: 1065 bridging ties and 885 closure ties