Networks, proximities and inter-firm knowledge exchanges

Stefano Usai, Emanuela Marrocu, Raffaele Paci

University of Cagliari, CRENoS

This is a pre-copyedited, author-produced version of an article accepted for publication in INTERNATIONAL REGIONAL SCIENCE REVIEW, following peer review.

The Version of Record: USAI, STEFANO, MARROCU, EMANUELA, PACI, RAFFAELE (2017). Networks, Proximities, and Interfirm Knowledge Exchanges. INTERNATIONAL REGIONAL SCIENCE REVIEW, vol. 40, p. 377-404, ISSN: 0160-0176, is available online at https://doi.org/10.1177/0160017615576079

Abstract

Building on previous literature that provides extensive evidence concerning flows of knowledge generated by inter-firm agreements, in this paper, we aim to analyze how the occurrence of such collaborations is driven by multi-dimensional proximity among participants and by their position within firms' networks. More specifically, we assess how the likelihood that two firms set up a partnership is influenced by their bilateral geographical, technological, organizational, institutional and social proximity and by their position within networks. Our analysis is based on agreements in the form of joint ventures or strategic alliances, announced over the period 2005-2012, in which at least one partner is localized in Italy. We consider the full range of economic activities, which allows us to offer a general scenario and to investigate specifically the role of technological relatedness across different sectors. The econometric analysis, based on the logistic framework for rare events, provided three noteworthy results. First, all five dimensions of proximity jointly exert a positive and relevant effect in determining the probability of inter-firm knowledge exchanges, signaling that they complement each other rather than function as alternative channels. Second, the highest impact on probability is due to technological proximity, followed by organizational, geographical and institutional proximities, while social proximity has a limited effect. Third, we find evidence concerning the positive role played by networks, through preferential attachment effects, in enhancing the probability of inter-firm agreements.

Keywords: knowledge flows, strategic alliances, joint ventures, proximities, networks

JEL: L14, O31, O33, R12

Acknowledgments: The research leading to these results received funding from the European Union's Seventh Framework Programme FP7-SSH-2010-2.2-1 (2011-2014), under grant agreement n° 266834 SEARCH project. We would like to thank Marta Foddi and Benedetta De Magistris for the excellent work in preparing the database. We have benefited from valuable comments by participants at workshops in Barcelona, Cagliari and Nice and at the Conferences SIE in Bologna, GCW in Rotterdam, T2S in Bergamo, WRSA in San Diego and RES in Manchester.

1. Introduction

The exchange of knowledge among firms is facilitated by their geographical proximity, given that knowledge has, in part, a tacit nature that tends to bind the spatial scope of spillovers (Jaffe et al., 1993; Audretsch and Feldman, 1996). Notwithstanding the much-investigated role of geography, the most recent literature has highlighted that inter-firm exchanges can also be mediated by other dimensions of closeness, which may have an a-spatial nature, such as cognitive, institutional or organizational proximity (Torre and Gilly, 2000; Boschma, 2005). Moreover, interactions among economic agents create social links that, over time, tend to evolve into wider networks, which are likely to facilitate the future exchanges of knowledge and moderate the adverse effects of other distances (Boschma and Frenken, 2009).

A growing body of empirical research has extensively analyzed the characteristics of networks that are expected to prompt innovation diffusion by considering various forms of connections among agents. These include participation in research programs (Autant-Bernard et al., 2007; Maggioni et al., 2007; Balland 2012), co-patenting (Cantner and Meder, 2007; Maggioni et al., 2007; Cassi and Plunket, 2013), citations (Maurseth and Verspagen, 2002; Paci and Usai, 2009), co-publications (Ponds et al., 2007), applicant-inventor relationships (Maggioni et al., 2011; Picci, 2010) and human capital mobility (Miguélez and Moreno, 2013; Breschi and Lissoni, 2009).

In this paper, we intend to follow a novel route by investigating the knowledge exchanges generated by two particular modes of agreements among firms: joint ventures and strategic alliances. The management and the economic literature (Kogut, 1988; Inkpen, 2000; Oxley and Sampson, 2004, Gomes-Casseres et al. 2006) has remarked how such inter-firm agreements, regardless of their specific nature and main motivation, create the conditions for knowledge sharing, mainly based on monitoring mechanisms and alignment of firms' incentives, and thus represent an important channel of knowledge exchange among the companies involved. Indeed, firms perform several activities before, during and after the agreements that allow partners to access and share knowledge-based resources, often embedded within the organizations and thus restricted to their members (Muthusamy and White, 2005; Janowicz-Panjaitan and Noorderhaven, 2008; García-Canal et al., 2008). Beginning during the preliminary stages of the agreement, such activities involve information flows among managers and employees, which may entail access to new technologies and organizational competencies, integration, sharing or transfer of capabilities, human and organizational resources, and, finally, formal and informal inter-organizational learning processes.[1]

It is worth noting that knowledge spillovers are an unavoidable result of involvement in inter-firms agreements, notwithstanding the efforts that individual companies put to protect their knowledge capital. Such inevitable transfer of knowledge is so ingrained in joint ventures or strategic alliances that in certain cases, such as when firms are competitors in end products (Oxley and Sampson, 2004), partners tend to restrict the scope of the agreement activities up to the point where they can still accrue the potential gains of cooperation but sharing a limited and selected amount of knowledge.

The aim of this paper is to analyze how the occurrence of inter-firm collaborations and the consequent knowledge exchanges among partners are driven by different dimensions of proximity among participants and by the features of the networks in which they are involved. More specifically, we assess the likelihood that any two firms choose to activate a bilateral partnership (or take part in a multi-participant agreement) in relation to their reciprocal geographical, technological, organizational, institutional and social proximities. Moreover, on the basis of the past experience of each firm, we assess whether their position within the network has an additional effect on the occurrence of inter-firm agreements, while controlling for several other characteristics at firm, industrial and regional level.

We base our empirical analysis on announced agreements over the period from 2005 to 2012, in which at least one firm is localized in Italy, including both domestic and international collaborations. In total, we examine 631 agreements involving 1078 firms. An original feature of our study is that we consider agreements covering all economic activities, which allows us to offer a wide-ranging scenario with respect to previous contributions on the role of proximity based on individual data. To the best of our knowledge, previous studies limit their investigations to a single industry, such as footwear (Boschma and Ter Wal, 2007), nanotechnology (Autant-Bernard et al., 2007), aviation (Broekel and Boschma, 2012), biotechnology (Fornahl et al., 2011), global navigation satellite systems (Balland, 2012), wine (Giuliani, 2010) and genomics (Cassi and Plunket, 2013). Other studies give a global picture of the role of proximity with respect to the whole economy but are conducted on data aggregated at the regional level (Marrocu et al., 2013; Maggioni

[1] Inkpen (2000) reports an illuminating example on how firms learn through alliances. He comments on the case of alliances formed by Sony Corp. with computer and telecommunication firms with the aim to create new technology for its consumer electronic products. "These alliances provide Sony with access to a wealth of new knowledge, such as how to manage product development cycles that are much faster in the computer industry than in consumer electronics. In forming these alliances, Sony personnel at various organizational levels will gain access to new knowledge. The challenge for Sony and other firms involved in alliances, and for all firms seeking access to knowledge beyond their boundaries, is to incorporate disparate pieces of individual knowledge into a wider organizational knowledge base" (page 1021).

et al., 2015). Our study represents a novel contribution in investigating five dimensions of proximity within a multi-sector framework and in testing whether they exert distinctive effects or they can substitute each other to some extent. Moreover, our empirical analysis is one of the first to be carried out within the *French School of Proximity*[2] approach for the case of joint venture and strategic alliances, which so far has been under-investigated within such perspective.

The choice to restrict the sample to the set of agreements of just one country (Italy in our case) is recurrent along this research avenue for two reasons. Firstly, when analyzing firm partnerships across different countries, the network structure becomes virtually worldwide, reaching high degrees of complexity. The country-study approach, therefore, allows researchers to obtain a manageable dataset and computationally less demanding proximity and network indicators for each pair of firms. Secondly, this choice does not alter the general validity of the results because, as noted by Narula and Hagedoorn (1999), firms' propensities to initiate an agreement are much more influenced by sectoral determinants than by country features[3].

This is confirmed by examining the distribution of agreements by the number of partners: in our sample, 90% of them involve only two partners; such proportions are very similar to the one (88%) reported by García-Canal et al. (2008) for 15 countries of the European Union. Additionally, the geographical distribution of the deals does not show any relevant country-specific features: for the Italian case, only 15% of the firm pairs involved in the agreements are domestic (i.e., both firms located in Italy), which is very similar to what is recorded for France (11%) or Germany (12%).[4]

The econometric analysis, based on the logistic framework for rare events, provided three main results. First, all five dimensions of proximity jointly exert a positive and relevant effect in determining the probability of inter-firm knowledge exchanges, signaling that they complement each other rather than act as alternative channels. Second, technological proximity, exhibits the highest impact on probability, followed by the geographical, organizational and institutional proximities, whereas social proximity has a limited effect. Third, firms' network positioning, in terms of preferential attachment, significantly enhances the probability of inter-firm agreements.

[2] The French School of Proximity was founded in the early 1990s by a group of French regional scholars, whose main aim was to develop and apply a multi-faceted concept of proximity to the study of industrial and spatial dynamics (https://frenchschoolofproximity.wordpress.com).

[3] This is particularly true when one compares main advanced countries, which share similar industrial structures and where the use of strategic alliances varies mainly across industries. Their use is particularly frequent in capital and knowledge intensive sectors, that is those sectors where firms have expanded internationally fastest and where they need not only to compete in various markets simultaneously, but also to exploit and acquire assets and technologies (Narula and Duysters, 2004). Inertia may also explain some industry specific trends which may depend on emulation. According to Narula and Hagedoorn (1999, p. 292) "firms simply do whatever firms in the same industry do".

[4] Moreover, the recent empirical literature (see among others, Helpman et al. 2004) on international exchanges has emphasized the crucial role played by firms' specific characteristics in determining their cross-border strategic behavior.

The remainder of the paper is organized as follows. In the second section, a detailed description of the data on inter-firm agreements is offered. In the third section, we present the empirical model and describe how we operationalize the proximity and network measures. In the fourth section, we address some econometric issues and present our estimation strategy. The econometric results are discussed in section 5, while concluding remarks are provided in section 6.

2. Inter-firm agreements

In this section, we describe the data on inter-firm agreements, which we propose as an indicator apt to account for knowledge exchanges occurring among companies. Data on announced agreements over the period 2005-2012 are collected from the SDC Platinum database (Thomson Financial) and include all the deals involving at least one partner located in Italy. Our data on inter-firm agreements comprise both joint ventures and strategic alliances. A joint venture is defined as a cooperative business activity formed by two or more firms that creates an independent organization and allocates the ownership, operational responsibilities and financial risks and rewards to each partner while preserving their separate identities. A strategic alliance is a cooperative activity formed by two or more organizations for a wide range of strategic purposes (manufacturing, licensing, marketing, supply, technology transfer) that does not create an independent entity but establishes a contractual agreement among the partners, which remain independent organizations.

From Table 1, we see that the total number of agreements is 631, which involved 1078 different organizations, of which 511 are Italian. Agreements can be simple or complex depending on the number of potential partners involved[5]. Table 1 clearly shows that most of the partnerships (570) do not go beyond the simplest form, a single pair of firms. Only 10% of total exchanges involve more than two partners, with a maximum of seven organizations engaged. Given the presence of deals among multiple partners, the number of actual pairs – 887 – is higher than the number of agreements, as shown in Table 1. The firm dyads are formed either in joint ventures (607) or in strategic alliances (280).

Table 1 also offers interesting information on the quota of announced agreements that were completed (382, equal to 43% of the total). The agreements aggregated in the uncompleted category can take on a different status, such as pending, letter of interest or renegotiated. It is important to remark that because we are using the agreements as a proxy of knowledge exchanges among partners and given that these exchanges also take place in the preliminary and earlier stages of the contract, independently from their successive progress, we prefer to consider all the announced agreements in our analysis. Another interesting aspect concerns the localization of partners, that is,

[5] This makes our network an affiliation one (a two-mode network), even though this particular feature is not exploited in our analysis.

the place where the headquarters are located. Table 1 shows that most pairs are formed by an Italian and a foreign partner (72%), with both partners located in Italy in only 15% of cases, and both partners being foreign in 14% of cases.[6] The last information included in Table 1 is crucial to understand the nature of our network. The proportion of actual pairs on total population is just 0.15%, which makes the occurrence of an agreement a rare event, so that the implied network has a very low density.

Given our interest in spatial proximity, in Table 2, we report the geographical location of the participants, which are located in 61 different countries and in 197 Territorial Level 2 (TL2, OECD) regions all over the world. Most partners (47%) are obviously firms located in Italy, followed by those situated in other EU countries (13%); as expected, EU firms represent the most frequent partners for the Italian companies due to their closeness in terms of geography and other proximity dimensions. Widespread exchanges are also recorded, with almost 12% of the partners located in the United States. Interestingly, among the most common partners, we also find firms located in emerging countries, such as India (7%), China (4%) and Russia (4%).

Finally, in Table 3, we report the distribution of the 631 agreements (first two columns) and of the 1078 participants (last two columns) across economic sectors, according to the Standard Industrial Classification (SIC) divisions. As expected, most agreements refer to manufacturing (34%), while another large proportion (14%) refers to service sectors, such as Personal and Business, Finance Insurance and Trade and Transportation, Energy and Sanitary services. Nearly the same shares can be found for the distribution of participants, except for the fact that the manufacturing sector has a higher quota (47%).

3. The empirical model

As stated in the introduction, we focus on the case of cooperation agreements as an indicator of knowledge flows because they imply a complex and lengthy process of interactions involving two or more partners. The purpose of our analysis is to model the probability that any two firms exchange knowledge by means of taking part in an agreement as a function of the bilateral geographical, technological, organizational, institutional and social measures of proximity and of the individual firm's network positions. The general form of our empirical model is:

Prob (inter-firm agreement) = f (proximities, firm network position, firm/regional/industrial controls) (1)

[6] Given the selection criteria of our sample, the pairs with both foreign participants are necessarily part of a larger agreement where at least one Italian firm is also included.

In this section we discuss the rationale for including the five dimensions of proximity and the network indicators, and we describe in detail how they are measured. The list of variables, along with some basic statistics, is reported in the Appendix.

3.1 The dependent variable

The observational unit in our model is represented by pairs of firms, and the dependent variable is constructed as a binary variable that takes value 1 when an agreement was announced between any two companies over the period 2005 – 2012 and 0 when a pair of firms could have set up a deal but did not. We refer to the latter as "potential" pairs. To identify the potential firm dyads, we apply the approach followed, among others, by Autant-Bernard et al. (2007) and Cassi and Plunket (2013). This requires pairing the 1078 firms involved in the 631 agreements included in our sample to obtain all possible pairs, which are 580,503.[7] Of this total, 887 pairs were involved in actual agreements, while the remaining 579,616 were not; therefore, they are considered "potential" pairs and act as "controls" in the estimation because they are formed by the most similar firms with respect to those which were involved in the agreements. They have in common the propensity to take part in such a type of collaborations, a characteristic which is not observable for firms without a contract. If the latter were also included, a source of unobservable firm heterogeneity would have been introduced, and this could have biased the results. In our sample, the number of firm pairs involved in agreements is equal to 0.15% of total possible dyads: setting up partnerships is clearly a rare event (Table 1). Therefore, we apply the methodology for rare events proposed by King and Zeng (2001), discussed in detail in section 4, where we address some relevant estimation issues.

3.2 Proximity dimensions

In this section we provide a detailed description of the different proximity dimensions included in the empirical models, and on how we operationalized their measurement. It is worth noting that in most cases they are measured as binary variables taking the value of 1 when any two firms share the same characteristic, so that they can be considered proximate with respect to that specific dimension.

Geographical proximity. The ability of a firm to use ideas and technologies created and developed by other firms is a crucial mechanism for knowledge accumulation and economic growth at both the micro and macro levels (Rallet and Torre, 1999; Romer, 1986). Such diffusion can be facilitated when knowledge, especially in its tacit form, can be transmitted among agents that are physically proximate (Von Hippel, 1994). Consequently, spatial proximity has been the most

[7] Because actual agreements are set up by firms that may operate in different productive sectors, we do not impose any restriction on the potential pairs on the basis of firms' productive relatedness.

thoroughly investigated dimension in the wide literature on knowledge flows and spillovers (Jaffe, 1986; Jaffe et al., 1993; Anselin et al., 1997). We measure geographical proximity by the inverse of distance between the locations of the partners (in kilometers).[8] We also control for co-location effects at the country and regional level by considering two additional dummies, *same country* and *same region*, taking value 1 when any two firms are located in the same country or in the same TL2 region, respectively. It is worth noting that co-location may enhance the partnering probability thanks to proximity arising from a wide range of factors, which includes sharing the same institutional setting and/or cultural background.[9]

Technological proximity. It is a commonly accepted idea that knowledge transfer is not an easy, smooth or generally accessible process (Cohen and Levinthal, 1990). It may require specific and appropriate absorptive capacity, which entails a homogenous cognitive base to understand and effectively process the available knowledge (Nooteboom, 2000). Essentially, firms are technologically proximate when they are either related to each other or they are similar in what they produce and/or how they produce it.[10] Following Ellwanger and Boschma (2013) we assume that such a common base can be adequately approximated by technological relatedness between partners. Therefore, we compute a set of five mutually exclusive technological interaction dummies which measure the decreasing degree of technological similarity. These dummies are based on the primary economic activity, which is reported in the SDC database at the 4-digit SIC code for each participant.[11] The first interaction dummy (*same industry*) takes the value of 1 when the partners operate in the same 4-digit SIC Industry and the value of 0 when the two firms operate in different Industries. It is interesting to note that this strong sectoral affinity is not unusual, as it occurs in almost 28% of the firm pairs involved in the actual agreements. The dummy *same industry group* takes the value of 1 when the highest degree of industrial relatedness is at the 3-digit SIC Industry Group and 0 when the two firms operate in different industry groups or are related at a finer

[8] For the case of extra-European companies, given the difficulties of finding the exact location of the firms, we have used the location of the country capital as a proxy.

[9] In order to account for similarity at country level mainly related to cultural and institutional factors we also consider a dummy taking value 1 when the firms are located in countries which have in common the same official language and the difference in the level of perceived corruption in the public sector (source: Transparency International). The latter variable is expected to account for cultural differences, with a connotation specifically related to cultural elements directly linked to economic activities.

[10] A more general aspect of technological proximity relates to cognitive proximity, that is closeness in cognitive tools, language, theories, factual knowledge and methodologies. Boschma (2005) includes both concepts in one category, that is cognitive proximity, whereas Knoben and Oerlemans (2006) make a subtle distinction. They suggest that the main dissimilarity rests on the fact that cognitive proximity "refers to the extent to which actors can communicate efficiently, whereas technological proximity refers to the extent to which actors can actually learn from each other" (p. 78).

[11] The Standard Industrial Classification is organized in 10 Divisions (1-digit classification), 83 Major groups (2-digit), 410 Industry groups (3-digit) and 965 Industries (4-digit).

industrial disaggregation.[12] Using the same methodology, we compute the next dummies for the 2-digit SIC Major group (*same major group*), for the 1-digit SIC Division (*same division*) and, finally, for the case (*different division*) when the partners operate in different divisions (conglomerate agreements). This last dummy is not included in the regressions so that firms operating in different divisions – the least proximate ones – represent the reference group; this final case is the most recurrent one as it is observed in 316 out of 887 cases (approximately 35%).

Organizational proximity. The exchange of information and knowledge can be influenced by the membership of individuals in the same club, group or organization, which generates strategic interdependence. The common membership implies the sharing of a set of rules and practices, based on organizational arrangements, which are crucial in reducing uncertainty and opportunistic behavior (Kirat and Lung, 1999). Such arrangements can be either within or among firms and may take different forms, ranging from informal relations among companies to formally organized firms. In our empirical analysis, as in Balland et al. (2013), we measure organizational proximity with a dummy variable (*same group*) equal to 1 if the two participants involved in a partnership have the same ultimate parent company, that is, they belong to the same corporate group, and 0 otherwise.

Institutional proximity. The exchange of ideas among economic agents may be easier and more effective if such agents share the same institutional framework. Formal and informal institutions, such as laws, rules and norms, can provide a set of standard procedures and routines that are shared by firms and, therefore, taken for granted. This common institutional background is crucial in reducing uncertainty and lowering transaction costs and, thus, favors pro-cooperative attitudes. These, in turn, enhance the possibility of an agreement and the exchange of knowledge (Maskell and Malmberg, 1999; Gertler, 2003). Following previous studies (Ponds et al., 2007; Cassi and Plunket, 2013), institutional proximity at the firm level is measured by means of a dummy variable based on the status of the two partners. More specifically, the dummy (*same status*) takes value 1 if the two firms share the same institutional status (both listed in a stock exchange market, or private, or subsidiaries, or government bodies). In our data, for 38% of pairs involved in actual agreements, the two partners are institutionally similar. As an alternative measure, we also compute another dummy (*both partners independent*) taking the value of 1 if the partners are both independent entities. This is a very frequent case in our sample, as it involves almost half of the firm pairs (47.8%).

Social proximity. The existence of social ties among individuals is another important catalyst for the exchange of ideas and knowledge (Granovetter, 1985). The analysis of social networks is

[12] Note that these first two dummy variables indicate that the two participants in the agreement operate in very similar economic activities; thus, the two dummies also proxy the direct competition among partners (García-Canal et al., 2008).

therefore vital to understanding the phenomena of knowledge creation and diffusion.[13] Social proximity is generally associated with relationships between actors resulting, for example, from past collaboration (Breschi and Lissoni 2009). Previous experiences may grant a certain level of predictability, reduce the perceived risk of conflict and add to trust among economic agents (Uzzi, 1996 and Mattes, 2012). This contributes to informal knowledge flows, which in turn lead organizations with a common partner to be more likely to interact and collaborate, especially within a risky and uncertain environment, such as that of technological change and innovation. We measure social proximity by means of social network analysis using pre-sample information on agreements announced in the past, starting from the year 2000. We assume that past direct and indirect relationships provide a facilitating environment for sharing knowledge in the future. Consequently, as in Autant-Bernard et al. (2007) and Balland (2012), we assume that the degree of social proximity is the inverse of the geodesic distance, which is a measure of the shortest path between two nodes (i.e., firms). This distance ranges from one (when two nodes are directly linked because they have been partners in the past) to infinity (when two nodes are virtually infinitely distant and neither they nor any of their direct and indirect partners ever met in the past). As a result, our social proximity indicator goes from zero (maximum geodesic distance) to one (minimum geodesic distance). To extract as much information as possible from our data, we compute the inverse of the recursive measure of geodesic distance (*inverse geodesic distance*) between firm *i* and firm *j* in all available previous years.[14] A robustness test is also performed by using the inverse of the geodesic distance computed considering only the previous five years (*inverse geodesic distance – previous 5 years*).

3.3 Individual network characteristics

Social ties may be the result of an individual attitude or customary behavior and thus have to be examined from two complementary perspectives: the single node and the entire system perspective (Bramanti and Maggioni, 1997). To this aim, we introduce a set of additional measures that takes into account each firm's single social position within the network of potential ties. Such measures supplement the information on the bilateral notion of social proximity, discussed above. The theoretical literature has shown how different network architectures may impact knowledge

[13] We also consider operationalizing social proximity by relating it to the notion of social capital. However, social capital indicators are available only either at the country level or at the regional level but for a subset of European countries. In the former case the social proximity variable would be very collinear with the *same country* dummy, which proves to be an adequate proxy for co-location effects and institutional proximity. In the latter case this choice would be inconsistent with the endogenous stratified sampling design adopted to select the "controls" and would considerably alter the network structure of the partners involved in the agreements. Since both limitations are likely to bias the results we do not include the proximity indicator related to social capital.

[14] This implies that the reference period for 2005 is the five-year period from 2000 to 2004, whereas the reference period for the 2012 observations is the period from 2000 to 2011.

growth and diffusion (Cowan and Jonard, 2004, Cowan et al., 2004 and Ter Wal and Boschma, 2009). Consequently, several empirical studies (Balland et al., 2013, Autant-Bernard et al., 2007, Cassi and Plunkett, 2013, Giuliani, 2010) have investigated to what extent the position of firms within the network influences knowledge diffusion. We follow this research path by considering three measures related to the network characteristics featured by each firm over the previous years, starting from the year 2000. These indicators are expected to account for firms' past experience in partnering. Social network analysis has defined two main hypotheses to better qualify the position of a firm within a network. These are the preferential attachment hypothesis and the transitivity hypothesis.

According to the *preferential attachment hypothesis*, actors are more inclined to link to the most connected individuals. Agents with a large number of relations are more attractive because they are supposed to be more productive or more trustworthy (Barabasi and Albert, 1999). A firm's preferential attachment is usually measured in terms of the number of its previous partnerships. Therefore, for each firm, we count the relations in which it was involved in the past, and this provides the basic information to compute its degree of centrality.

The *transitivity hypothesis* states that some agents are more reachable than others because of their relative position in the network. Some nodes are relatively closer to all other nodes, and therefore, they represent a more effective route to connect to potential nodes to obtain information and acquire knowledge. It is important to note that the literature usually refers to transitivity when organizations that have a partner in common are more likely to partner themselves, thereby effectuating triadic closures (Cassi and Plunkett, 2013). In our work, we prefer to employ a more general concept and indicator, because triadic closures are very rare in our sample. We thus measure the transitivity property by referring to either the notion of closeness centrality or the notion of betweenness centrality. The former indicator is the inverse of the sum of the distances of a node to all other nodes divided by the total distance, and it is a measure of either how long it takes to spread information from one node to all other nodes sequentially or how long it takes to retrieve information from all other nodes. The latter indicator measures the extent to which a node falls on the geodesic paths between other pairs of nodes in the network and indicates its potential for intermediation among actors. It is given by the number of shortest path which cross one node divided by the total number of shortest paths.

The expected impact for both phenomena – preferential attachment and transitivity – is positive because firms are supposed to be willing to maximize the opportunity to obtain knowledge from the whole network, thus connecting to the most joined and central firms. However, a negative effect could also arise when firms are worried that linking to a central and highly connected partner may jeopardize the appropriability of their knowledge (Autant-Bernard et al., 2007). In the empirical model we include one network indicator at a time because they are highly correlated; moreover, because the degree of centrality exhibits more variability in our sample it is considered more informative and, therefore, included in the baseline specification.

3.4 Other controls

In our empirical model we also control for various characteristics at the firm, industrial and regional level. More specifically, we include information on each firm's status, organization, ownership nationality, geographical location and principal sector of activity. Regarding the status, we have computed two dummies (*listed* and *private*) for the most recurrent actual cases that account for the firm being publicly traded on a stock exchange market (40%) or being a private company (33%); in other words, a company owned either by non-governmental organizations or by a relatively small number of shareholders, often a family in Italy.[15] We have also included a dummy (*independent*) taking the value of 1 when the participant is an independent firm (i.e., when the ultimate parent company corresponds with the partner itself, 68% of the cases) and a dummy (*foreign*) taking the value of 1 for foreign-owned companies (10%).

The propensity to exchange knowledge through inter-firm agreements may also depend on the characteristics of the type of economic activity performed by firms and by the economic context where firms are localized. We, therefore, control for such differences by including ten mutually exclusive dummies for the 1-digit SIC divisions of economic activity and seven mutually exclusive regional dummies to take into account firm's spatial localization. This can be in one of the five Italian NUTS1 macro regions (North-West, North-East, Centre, South and Islands) or in one of the European Union countries or in an extra-EU country. Moreover, although we expect that firms tend to integrate more the higher is the knowledge intensity of their production, this can happen with some variability depending on the local context. To account for this feature we also include sectoral-regional interaction terms, which are obtained by multiplying the regional dummies with binary variables for the manufacturing and the service macro-sectors.

4. Estimation issues for rare event logit models

The analysis of the effects of networks and proximities on the probability that two firms exchange knowledge through an agreement is performed within the logistic framework for rare events. As stated above, this entails creating the dependent variable (Y) taking value 1 for pairs of firms (887) which actually established a cooperative link during the period 2005-2012 and 0 for

[15] In addition to the main institutional status represented by these two dummy variables, firms can also be subsidiaries, joint ventures or governmental organizations.

dyads of firms (579,616) that could have set up an agreement but did not. These are considered "potential" pairs and act as "control" pairs in the estimation, as detailed below.

By comparing the high number of potential pairs with the one related to actual deals (0.15% of all possible pairs), it is evident that setting up a cooperative agreement can be considered a rare event. In this case, given the disproportionate number of 0 observations, the logit model estimated on the total number of firm pairs would severely underestimate the probability of occurrences.[16] Following King and Zeng (2001, 2002), we apply the choice-based or endogenous stratified sampling approach, which requires selecting all the observations for which Y=1 (the "cases") and randomly (independently from the explanatory variables) selecting the observations for which Y=0 ("controls"). Thus, all the actual pairs are considered relevant and informative observations, whereas this is the case only for a very limited proportion of the controls. It is important to note that selecting on the zeroes also allows for more efficient data collection because only a small part of these observations contribute to the information content of the explanatory variables. As is well known, data selection based on Y induces bias, and it is therefore necessary to apply the appropriate statistical corrections to obtain consistent and efficient estimators. The most applied ones are based on prior correction and on the weighting method, both of which require prior knowledge of the population proportion of ones.[17]

It is worth noting that we have also to face another issue related to sample selection because the decision to set up an agreement - rather than consider other forms of collaboration - might be driven by the fact that a firm knows its proximate potential partners. To attenuate the possible selection bias, we apply the independence-in-conditional-mean approach by including in our models a wide range of firms' characteristics, which are jointly likely to affect firms' collaboration modes. Such characteristics, described in Section 3, are related to firm's status, organization, ownership nationality, operating division and geographic location. Once we control for these individual firm

[16] For a comprehensive discussion of the econometric methodology to estimate logit models for rare events, refer to King and Zeng (2001). They show that in the case of rare events the intercept coefficient is estimated with bias; the negative bias in such coefficient implies that the expected probability is also underestimated. However, it is not sufficient to correct the intercept for bias, it is also necessary to account for estimation uncertainty, otherwise the bias is magnified by the too small variance of the estimator distribution. In order to account for estimation uncertainty King and Zeng propose two methods: an analytical approximation and an approximate Bayesian estimator. The latter is proved to be superior in terms of mean square error and therefore we applied it in all our empirical models.

[17] We recall that the prior correction method is less computationally demanding because it entails only correcting the constant estimate on the basis of the population proportion of ones; the maximum likelihood estimators for the coefficients associated with the explanatory variables do not need any correction because they maintain their unbiasedness and consistency properties. The weighting method entails weighting the sample observations so that the weighted proportions of ones and zeros in the sample equal the corresponding population proportions. The weighting method is robust to potential misspecification (Manski and Lerman, 1977), but it requires further corrections because the MLE for the variance-covariance matrix is severely biased.

features, we expect that the decision to select a particular partner to carry out a specific agreement is independent of higher-level collaboration or acquisition decisions.[18]

The empirical specification for the probability of observing an agreement is formalized on the basis of the following cumulative logistic distribution:

$$Prob\left(Y_{ij} = 1 \middle| X_{ij}, N_i, N_j, W_i, W_j, Z_{ij}\right) = \frac{1}{1+e^{-\left(X_{ij}\beta_1 + N_i\beta_2 + N_j\beta_3 + W_i\beta_4 + W_j\beta_5 + Z_{ij}\beta_6\right)}} \quad (2)$$

where Y_{ij} takes the value of 1 when an agreement process involves firm i and firm j and 0 otherwise; the matrix X_{ij} includes the pairs' terms that allow us to assess to what extent the agreements are driven by inter-firm proximity, measured along the various dimensions – spatial, technological, organizational, institutional and social – described in detail in the previous section. Each of the N_i and N_j matrices includes the network indicator for firm i and j, respectively, whereas each of the matrices W_i and W_j comprise the individual firm's control variables. The matrix Z_{ij} includes the sectoral-regional controls.

We estimate model (2) by performing the sequential procedure suggested by King and Zeng (2001) for selecting the zero observations.[19] More specifically, we considered several random samples by starting with the sample for which each actual pair is matched with just one random control and stopping when we obtained no further efficiency gains, signaled by a reduction in the magnitude of standard errors. For both the prior correction and the weighting method this occurred for the sample in which 10 randomly drawn potential pairs are matched with each actual pair. Comparing the alternative correction approaches, we found that overall, the estimated coefficients did not differ substantially, thus signaling the absence of any clear misspecification problem. We interpret this result in favor of our highly parameterized specification, which simultaneously accounts for five different proximity dimensions, network features and a wide range of firm characteristics to control for possible sources of heterogeneity. For these reasons, in the next section, we focus the discussion on the evidence provided by models based on the prior correction method.

5. Empirical results

The estimated models are presented in Table 4. The first model includes only the geographical proximity, while the second one comprises the five proximity dimensions. Model 1 can be seen as a sort of benchmark, which allows us to investigate to what extent the different kinds

[18] Chakrabarti and Mitchell (2013) adopt the same approach for the case of M&A determinants.
[19] All estimations are carried out by using the ReLogit software by Tomz et al. (1999).

of inter-firm proximity complement each other or act as proxies for any of the other dimensions considered. Consequently, we assess whether the almost undisputed effect of geographical closeness is maintained when the role of other proximities is taken into account. The subsequent models (3-7) address the robustness of our results across specifications that include alternative indicators for some proximity measures and network characteristics.

5.1 The baseline model

In column 1, we report the benchmark model, according to which knowledge flows are determined by geographical proximity and by the network characteristics of each partner; controls at firm, sector and regional level are also included to account for firm heterogeneity. The results show that geographical nearness is a crucial determinant of such flows and also that the preferential attachment (degree of centrality) of each partner significantly influences the cooperation decisions among firms. Most importantly, this model estimates the probability that any two firms start an agreement process (see last row in column 1) at 2.4%, which is sixteen times the basic random probability of 0.15%. We can interpret such an increase in probability as evidence of the predictive power of our model, even in its underspecified form with just the spatial proximity.

The second model, presented in column 2, is our baseline specification where the effect of proximity is assessed with respect to all the additional dimensions - technological, organizational, institutional and social - discussed in section 3. The results show that geographical closeness remains relevant even when all other dimensions of proximity are included, although the magnitude of its coefficient is almost halved. This is mainly due to the inclusion of the co-location indicator; as expected, being proxied by the *same-country* dummy, such indicator is highly correlated with the geographical proximity. The co-location indicator, which is also apt to account for the similarity in cultural and institutional dimensions, is quite relevant in enhancing the probability of inter-firm knowledge exchanges.[20]

Interestingly, the fact that closeness in space remains highly significant confirms findings in Paci et al. (2014) that geography and the other dimensions of proximity complement each other, rather than act as substitutes. This is due to the distinctive multi-facet characteristic of proximity, which the selected indicators are indeed capturing. As a matter of fact, results in column 2 show that all dimensions of proximity exhibit a positive and significant coefficient, which implies an

[20] We estimate also some models in which similarity at the country level is proxied by common language and cultural distance Although both variables exhibit the expected sign, so that the probability of observing an agreement is positively influenced by common language and negatively by the cultural distance, when they are included together with the *same country* dummy, they are no longer significant due to multicollinearity among the variables. Given that the *same country* dummy is supposed to capture any form (observed and unobservable) of proximity between any two countries (including language and cultural closeness) we think that the model which includes the *same country* term is to be preferred.

increased probability of knowledge exchange through inter-firm agreements.[21] The estimated probability for this model rises to 4.6%, which implies that, thanks to the introduction of all proximities, cooperation becomes 90% more likely than when only geographical proximity is taken into account, indicating that our baseline model has a high predictive power. This result thus highlights the importance of simultaneously accounting for the whole set of relevant proximities within a comprehensive empirical specification, as suggested by the French School of Proximity (Torre and Gilly, 2000).[22]

Another noteworthy aspect is that the coefficients of technological proximity are not only positive and significant, but that their magnitude increases with the degree of similarity of firms' productive and knowledge bases. With respect to the reference group that comprises the most unrelated firms, the smallest coefficient (0.94) is found when the highest level of technological relatedness is the SIC1 division, whereas the largest (3.98) one is associated with the case when both firms operate in the same SIC4 industry. This result confirms recent findings by Boschma et al. (2015) on the relevant role played by industrial relatedness in favoring mergers and acquisitions partnering. Firms operating in the same economic activity are more likely to set up a cooperation agreement to exploit the potential synergies in terms of products and services and to benefit from economies of scale and scope. Moreover, information asymmetries between firms that are highly technologically related are lower, and this favors their knowledge exchanges (Hussinger, 2010).

When two firms share the same ownership status, that is they are institutionally proximate, collaborative agreements are significantly more likely to occur, as in Cassi and Plunket (2013).

We also find a positive and significant effect of organizational proximity, measured in terms of membership to the same group, confirming the results from Balland (2012) and Balland et al. (2013).

As far as social proximity is concerned, the probability that two firms announce an agreement is a positive and significant function of the inverse of the geodesic distance, as in Autant-Bernard et al. (2007). We also find evidence supporting the relevant role of network characteristics, as the indicator of preferential attachment has a positive sign and is significant for both partners. We, thus, confirm previous findings (Balland et al., 2013, Balland, 2012, Cassi and Plunket, 2013) that agents prefer to interact and exchange knowledge with those that have former agreement experiences. Such a preference induces a self-reinforcing process of collaboration around the most

[21] Basile et al. (2012) provide evidence on the positive and synergic effects of different kinds of proximities (spatial, social and relational) on the productivity growth of European NUTS2 regions.

[22] In a preliminary analysis, we introduce the proximity indicators one at a time (the complete set of controls was always included) and we found that they were all significant but, as expected, for all the models the estimated probability is always well below the one obtained from the all-proximity model. The highest estimated probability (2.9) was found for the technology proximity model, followed by country level institutional (2.7), geographical (2.4), firm level institutional (1.9), organizational (1.8) and social (1.8) model.

connected firms that may lead to an increase in the degree of concentration within the network. This process has its rationale in the belief that firms that have already experienced knowledge exchanges possess more information because of those exchanges. Previous experience is also interpreted as an indirect signal of the potential value of a firm as a partner.

Moreover, having found a positive and significant coefficient for the preferential attachment terms allows us to rule out the occurrence and prevalence of the contrasting effect related to appropriability. Firms may face a trade-off between the necessity to increase the probability of getting effective information through cooperation and the concurrent necessity to control the dissemination of their own knowledge (Antonelli et al., 2011). In our case, firms are not wary of coming into contact with firms that are in the best position, not only for collecting knowledge, but also for spreading it.

5.2 Robustness tests

In columns 3-7 of Table 4, we test the robustness of our results across specifications that include alternative indicators for proximity and network measures.

In column 3 we augment model 2 with the co-location indicator represented by the *same region* dummy variable to assess whether similarity in the regional context exert an additional effect on the probability of observing a firm agreement. The inclusion of the new variable makes the spatial proximity no longer significant (the two variables have a correlation coefficient of 0.68. In terms of predictive power model 3 does not outperforms model 2, as it emerges by comparing the estimated probabilities (bottom line of Table 4). Model 2 thus remains the preferred specification.

In column 4 we include the alternative measure for institutional proximity, in this case we consider firms to be proximate if both of them are independent entities; however, it is not significant, while leaving the coefficients of the other regressors almost unchanged.

In column 5, we also test the robustness of our results with respect to a different proxy for social proximity; more precisely, we now include the inverse of the geodesic distance, computed by limiting the time span to the previous five-year period to consider the same time span for all observations. We find that this variable is only partially significant (10%), even though the average estimated probability remains almost constant.

We investigate the robustness of the degree of centrality indicator, used to capture the network positioning of each partner, by including the alternative indicators based on the notion of betweenness and closeness centrality. It is worth reminding that these two indicators are related to the transitivity hypothesis. Although the estimated models (columns 6 and 7) indicate that, overall, the new indicators could positively affect the occurrence of inter-firm agreements, they are outperformed by the degree of centrality indicator in terms of significance and predicted probability.

Finally, it is worth mentioning that we also carried out a sub-sample analysis to investigate whether relevant differences emerged when splitting the sample according to some features of the agreements, such as completed vs. uncompleted agreements, joint ventures vs. strategic alliances and manufacturing vs. service sectors. This analysis is rather preliminary because the limited number of actual agreements prevents us from estimating all the sub-samples, and thus further research is required. In any case, no significant differences were found across sub-samples, thus confirming the main findings discussed above for the whole sample.

5.3 Effects on probability

In this final section, unlike previous contributions in the literature on proximity, we take a step forward in assessing how changes in proximity or network features affect the likelihood that any two firms exchange knowledge thanks to inter-firm collaborations. Therefore, we measure the increase in the estimated conditional probability for a given change in each explanatory variable in turn. Unless otherwise stated, such a change is considered with respect to the median value and is equal to one standard deviation.

Table 5 reports the results obtained with respect to our preferred model 2 in Table 4. We recall that model 2 yielded an estimated probability of an agreement equal to 4.6% when median values are attributed to all variables.

The first and most interesting result is that, as in Paci et al. (2014) and Montobbio and Sterzi (2013), the largest impact on probability is found when the technological proximity is measured by the highest degree of technological relatedness. A positive change of one standard deviation with respect to the median yields an increase in probability of approximately 9%, with an increase of 96% with respect to the baseline estimation. Increases of approximately 50% are also registered for all other measures of technological proximity.

The co-location indicator ranks second in terms of effectiveness in increasing the agreements probability, its effect (53%) is similar in magnitude to the ones discussed above for the intermediate degrees of technological relatedness.

Regarding geographical proximity, one standard deviation change makes the estimated probability increase by 10% (from 4.6% to 5.1%). Organizational proximity produces effects comparable in magnitude to the geographical proximity ones (10%), whereas the ones attributable to institutional proximity are slightly lower (7%). Social proximity seems to induce the smallest impact (0.8%) on the response variable. Despite the modest influence of social proximity, we find that a firm's own social relations are much more effective. Considering the partners average effects, preferential attachment raises the probability of observing an agreement by a sizeable 39%.

Overall, our findings offer further support to the composite role played by proximities and network features in driving the complex diffusion of knowledge. Although they may have reciprocal moderating effects, proximities and social links are by no means interchangeable; they complement each other by contributing to favoring the transmission of knowledge among firms.

6. Conclusions

In this paper, we analyze the determinants of knowledge exchanges among firms originating from inter-firm agreements, such as joint ventures and strategic alliances. The management and economic literature has provided extensive evidence on the existence of knowledge flows generated during the various stages of inter-firm agreement processes, when partners share knowledge-based resources, often embedded within organizations and accessible only by members. The knowledge flows occur in the form of transfers of new technologies and organizational capabilities and also in active participation in formal and informal organizational learning processes.

More specifically, we assess the effects exerted by different types of proximity and by the position of participants within the network of previous ties on the probability that any two firms engage in a cooperation agreement and therefore exchange knowledge. We analyze the case of announced agreements over the period 2005-2012 in which at least one firm is localized in Italy, considering a total of 631 agreements, which involve 1078 unique firms and give rise to 887 pairs of actual partners. The analysis is performed within a logistic framework for rare events given the large number of potential firm pairs; that is, any two firms that could have set up an agreement but did not. Our preferred model simultaneously accounts for five different proximity measures (geographical, technological, organizational, institutional and social), firm network features and a wide set of covariates to control for firm heterogeneity such as status, organization, ownership nationality, principal sector of activity, geographic location and sectoral-regional productive patterns.

Results show that all proximity dimensions exhibit a positive and significant effect, thus providing further and compelling evidence that knowledge exchanges are facilitated not only by spatial proximity, as argued traditionally, but also by other dimensions of inter-firm closeness, such as sharing a common cognitive base and the same institutional background, being a part of the same organization and belonging to the same network.

Most importantly, we find that the highest impact on the probability of generating inter-firm knowledge exchanges is found when we consider the technological proximity between firms. Being located in the same country, and thus sharing the same institutional and cultural setting, is the second most effective proximity factor, followed by geographical and organizational proximity. Firm level institutional and social proximity facilitate the exchange of knowledge to a significant but smaller degree. Despite the modest influence of social proximity, the relevance of network links is supported by significant preferential attachment and transitivity effects. There is robust evidence that firms' previous experience within existing networks positively affects the probability that companies set up cooperation agreements and thus gives rise to a knowledge exchange. The concurrent effect of different proximities and firms' network positions makes the probability of observing an inter-firm agreement as high as 4.6%, which is almost 30 times higher than the random probability.

Thus, our findings highlight the importance of analyzing inter-firm knowledge flows by simultaneously accounting for the whole set of relevant proximities and network features within a comprehensive empirical model. In particular, the relative importance of cognitive/technological proximity with respect to other dimensions of closeness may result in some tentative policy reading and some related interesting suggestions for future research. As for the former, the presence of knowledge flowing along the technological space implies that countries and regions should try to envisage balanced policies to create a common wide knowledge base and specific industrial platforms to maximize the absorptive capacity and its effective application. Moreover, the presence of a-spatial technological clusters suggests the implementation of specific industrial policies to support the functioning across countries and regions of such non-contiguous industrial districts. As for the latter, technological proximity can be further analyzed by considering more complex measures of technological relatedness, for example along the production value chain. More generally, future research should focus on the impact of these agreements on firms' creation of new knowledge and ultimately in enhancing firms' economic performance.

References

Anselin L., Acs Z.J. and Varga A. (1997) Local geographic spillovers between university research and high technology innovations, *Journal of Urban Economics*, 42, 422-448.

Antonelli C., Patrucco P.P. and Quatraro F. (2011) Productivity Growth and Pecuniary Knowledge Externalities: An Empirical Analysis of Agglomeration Economies in European Regions, *Economic Geography*, 87, 23-50.

Audretsch D.B. and Feldman M.P. (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, 495-519.

Balland P.A. (2012) Proximity and the evolution of collaboration networks - evidences from R&D projects within the GNSS industry, *Regional Studies*, 46, 6, 741-756.

Balland P.A., de Vaan M. and Boschma R. (2013) The Dynamics of Interfirm Networks along the Industry Life Cycle: The Case of the Global Video Game Industry 1987-2007, *Journal of Economic Geography*, 13, 741-765.

Barabasi A.L. and Albert R. (1999) Emergence of Scaling In Random Networks, *Science*, 86, 509-512.

Basile R., Capello R. and Caragliu A. (2012) Technological interdependence and regional growth in Europe: Proximity and synergy in knowledge spillovers, *Papers in Regional Science*, 91, 697-722.

Boschma R.A. (2005) Proximity and innovation. A critical assessment, *Regional Studies*, 39, 61-74.

Boschma R.A. and Frenken K. (2009) The Spatial Evolution of Innovation Networks: A Proximity Perspective, in R. Boschma and R. Martin (eds.) *The Handbook of Evolutionary Economic Geography*, Cheltenham: Edward Elgar, 120-135.

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 & Innovation*, 14, 177-199.

Boschma R.A., Marrocu E. and Paci R. (2015) Symmetric and asymmetric effects of proximities. The case of M&A deals in Italy, *Journal of Economic Geography*, forthcoming.

Bramanti A. and Maggioni M.A. (1997) The dynamics of mileux: the network analysis approach, in: Ratti R, Bramanti A, Gordon R (eds) *The dynamics of innovative regions*. London: Ashgate.

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, 439-468.

Broekel T. and Boschma R.A. (2012) Knowledge networks in the Dutch aviation: the proximity paradox, *Journal of Economic Geography*, 12, 409-433.

Cantner U. and Meder A. (2007) Technological proximity and the choice of cooperation partner. *Journal of Economic Interaction and Coordination*, 2, 45-65.

Cassi, L., Plunket, A. (2013) Research collaboration in co-inventor networks: combining closure, bridging and proximities. *Regional Studies*, published on line
DOI: 10.1080/00343404.2013.816412.

Chakrabarti A. and Mitchell W. (2013) The Persistent Effect of Geographic Distance in Acquisition Target Selection, *Organization Science*, 24, 1805-1826.

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

Cowan R. and Jonard N. (2004) Network structure and the diffusion of knowledge. *Journal of Economic Dynamics and Control*, 28, 1557-1575.

Cowan R., Jonard N. and Ozman M. (2004) Knowledge dynamics in a network industry, *Technological Forecasting and Social Change*, 71, 469-484.

Ellwanger N. and Boschma R.A. (2013) Who acquires whom? The role of geographical proximity and industrial relatedness in Dutch domestic M&As between 2002 and 2008, *Papers in Evolutionary Economic Geography*, no. 18, Utrecht University.

Fornahl D., Brökel T. and Boschma R.A. (2011) What drives patent performance of German biotech firms? The impact of R&D subsidies, knowledge networks and their location, *Papers in Regional Science*, 90, 395-418.

García-Canal E., Valdés-Llaneza A. and Sánchez-Lorda P. (2008) Technological flows and choice of joint ventures in technology alliances, *Research Policy*, 37, 97-114.

Gertler M.S. (2003) Tacit knowledge and the economic geography of context, or the undefinable tacitness of being (there), *Journal of Economic Geography*, 3, 75-99.

Giuliani E. (2010) Network dynamics in regional clusters: the perspective of an emerging economy. *Papers in Evolutionary Economic Geography*, no. 14, Utrecht University.

Gomes-Casseres B., Hagedoorn J. and Jaffe A. (2006) Do alliances promote knowledge flows? *Journal of Financial Economics,* 80, 5-33.

Granovetter M. (1985) Economic action and social structure: the problem of embeddedness. *American Journal of Sociology*, 91, 481-510.

Helpman, E., Melitz, M. and Yeaple, S. (2004) A simple theory of international trade with multinational corporations, *The American Economic Review*, 94, 300-316.

Hussinger, K. (2010) On the importance of technological relatedness: SMEs versus large acquisitions targets, *Technovation*, 30, 57-64.

Inkpen A.C. (2000) Learning through joint ventures: a framework of knowledge acquisition. *Journal of Management Studies*, 37, 1019-1043.

Jaffe A.B. (1986) Technological Opportunity and Spillovers of R&D: evidence from Firms' Patents, Profits and Market Value, *American Economic Review*, 76, 984-1001.

Jaffe A.B., Trajtenberg M. and Henderson R. (1993) Geographic localization of knowledge spillovers as evidenced by patient citations, *Quarterly Journal of Economics*, 108, 577-598

Janowicz-Panjaitan M. and Noorderhaven N.G. (2008) Formal and informal interorganizational learning within strategic alliances, *Research Policy*, 37, 1337-1355.

King G. and Zeng L. (2001) Logistic regression in rare events data, *Political Analysis*, 9, 137-163.

King G. and Zeng L. (2002) Estimating risk and rate levels, ratios and differences in case-control studies, *Statistics in Medicine*, 21, 1409-1427.

Kirat T. and Lung Y. (1999) Innovation and proximity - Territories as loci of collective learning processes*, European Urban and Regional Studies*, 6, 27-38.

Knoben J. and Oerlemans L.A.G. (2006) Proximity and inter-organizational collaboration: A literature review, *International Journal of Management Reviews*, 8, 71-89.

Kogut B. (1988) Joint ventures: theoretical and empirical perspectives, *Strategic Management Journal*, 9, 319-332.

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, 471-493.

Maggioni M.A., Uberti T.E. and Nosvelli M. (2015) Geography, Relations and Network Structures. Mapping the structure of Knowledge Flows within EU financed Joint Research Consortia, *this special issue.*

Maggioni M.A., Uberti T.E. and Usai S. (2011) Treating patents as relational data: knowledge transfers and spillovers across Italian provinces, *Industry & Innovation*, 18, 39-67.

Manski C.F. and S. Lerman (1977) The estimation of choice probabilities from choiced based samples, *Econometrica*, 45, 1977-1988.

Marrocu E., Paci R. and Usai S. (2013) Proximity, Networking and Knowledge Production in Europe: what lessons for innovation policy? *Technological Forecasting and Social Change*, 80, 1484-1498.

Maskell P. and Malmberg A. (1999) The competitiveness of firms and regions. 'Ubiquitification' and the importance of localized learning, *European Urban and Regional Studies*, 6, 9-25.

Maurseth P. and Verspagen B. (2002) Knowledge Spillovers in Europe: A Patent Citations Analysis, *Scandinavian Journal of Economics*, 104, 531-45.

Mattes, J. (2012) Dimensions of Proximity and Knowledge Bases: Innovation between Spatial and Non-spatial Factors, *Regional Studies*, 46, 1085-1099.

Miguélez E. and Moreno R. (2013) Research Networks and Inventors' Mobility as Drivers of Innovation: Evidence from Europe, *Regional Studies*, 47, 1668-1685.

Montobbio F. and Sterzi V. (2013) The globalization of technology in emerging markets: a gravity model on the determinants of international patent collaborations, *World Development*, 44, 281-299.

Muthusamy S.K. and White M.A. (2005) Learning and knowledge transfer in strategic alliances: a social exchange view, *Organization Studies*, 26, 415-441.

Narula R. and Hagedoorn J. (1999) Innovating through strategic alliances: moving towards international partnerships and contractual agreements, *Technovation*, 19, 283-294.

Narula R. and Duysters G. (2004) Globalisation and trends in international R&D alliances, *Journal of International Management*, 10, 199-218.

Nooteboom B. (2000) Learning by Interaction: Absorptive Capacity, Cognitive Distance and Governance, *Journal of Management and Governance*, 4, 69-92.

Oxley J.E. and Sampson R.C. (2004) The scope and governance of international R&D alliances. *Strategic Management Journal*, 25, 723-749.

Paci R. and Usai S. (2009) Knowledge flows across the European regions, *Annals of Regional Science*, 43, 669-690.

Paci R., Marrocu E. and Usai S. (2014) The complementary effects of proximity dimensions on knowledge spillovers, *Spatial Economic Analysis*, 9, 9-30.

Picci L. (2010) The internationalization of inventive activity: A gravity model using patent data, *Research Policy*, 39, 1070-1081.

Ponds R., van Oort F. and Frenken K. (2007) The geographical and institutional proximity of research collaboration, *Papers in Regional Science*, 86, 423-444.

Rallet A. and Torre A. (1999) Is geographical proximity necessary in the innovation networks in the era of the global economy?, *GeoJournal*, 49, 373-380.

Romer, P.M. (1986) Increasing Returns and Long-run Growth, *Journal of Political Economy*, 94, 1002-1037.

Ter Wal L.J. and Boschma R.A. (2009) Applying social network analysis in economic geography: framing some key analytic issues, *Annals of Regional Science*, 43, 739-756.

Tomz M., King G. and Zeng L. (1999) RELOGIT: Rare Events Logistic Regression, Version 1.1 Cambridge, MA: Harvard University, October 1, http://gking.harvard.edu.

Torre A. and Gilly J.P. (2000) On the analytical dimension of proximity dynamics, *Regional Studies*, 34, 169-180.

Uzzi, B. (1996) The sources and consequences of embeddedness for the economic performance of organizations: The network effect, *American Sociological Review*, 61, 674-698.

Von Hippel E. (1994) “Sticky Information" and the Locus of Problem Solving: Implications for Innovation, *Management Science*, 40, 429-439.

Appendix. Variable definitions

Proximity indicators for partener pairs		***Mean/Frequency****
Spatial proximity		
inverse geo distance	inverse of the distance in km between partners cities	0.00025
co-location - same country	dummy = 1 if partners are located in the same country	0.18828
co-location - same region	dummy = 1 if partners are located in the same TL2 region	0.07215
Cognitive proximity - Technological relatedness		
same division	dummy = 1 if the highest degree of industrial relatedness is at SIC1	0.18828
same major group	dummy = 1 if the highest degree of industrial relatedness is at SIC2	0.10259
same industry group	dummy = 1 if the highest degree of industrial relatedness is at SIC3	0.07666
same industry	dummy = 1 if the highest degree of industrial relatedness is at SIC4	0.27621
Organisational proximity		
same group	dummy =1 if partners belong to the same group	0.01353
Institutional proximity		
same status	dummy = 1 if partners have the same insitutional status	0.38219
both partner independent	dummy = 1 if both partners are independent companies	0.47802
Social proximity		
inverse geo distance	inverse of geodesic distance with recursive time window	0.00000
inverse geo distance (5 years)	inverse of geodesic distance with 5-year window	0.00000
Network characteristics for each partner		
preferential attachment	degree centrality, number of links incident upon a node	1.97774
transitivity	betweenness, the inverse of the sum of the distances of a node to all other nodes divided by the total distance	0.00013
transitivity	closeness centrality, number of shortest path which cross one node divided by the total number of shortest paths	0.00409
Controls for individual characteristics of each partner		
listed	dummy = 1 if partner is publicly traded on a stock exchange	0.40074
private	dummy = 1 if partner status is private	0.32746
independent	dummy = 1 if partner is independent, it is not part of a group	0.67904
foreign	dummy = 1 if partner is owned by a foreign ultimate parent company	0.09833
Italy - North-west	dummy = 1 if partner location is in the Italian North-western macroregion	0.24861
Italy - North-east	dummy = 1 if partner location is in the Italian North-eastern macroregion	0.09647
Italy - Centre	dummy = 1 if partner location is in the Italian Centre macroregion	0.10482
Italy - South	dummy = 1 if partner location is in the Italian Southern macroregion	0.01763
Italy - Islands	dummy = 1 if partner location is in one of the two main Italian islands	0.00649
European country (other than Italy)	dummy = 1 if partner location in another EU countries	0.12894
SIC division	dummy = 1 if partner's economic activities is division SIC 1 (or 2-10, ten dummies)	

* *Note*: the proximity indicators values are calculated on the sample of firm pairs involved in actual agreements (N=887), whereas the values for partner's characteristics are calculated on the observed sample of firms (N=1078)

Table 1. Inter-firm agreements with at least an Italian participant, 2005-2012

Announced agreements	**631**
with 2 participants	*570*
with 3 participants	*43*
with 4 participants	*6*
with 5 participants	*8*
with 6 participants	*2*
with 7 participants	*2*
Participants	**1078**
Italian	*511*
foreign	*567*
Actual participant pairs	**887**
joint ventures	607
strategic alliances	280
completed	382
uncompleted	505
with both partners in Italy	*130*
with one partner in Italy	*636*
with both partners not in Italy	*121*
Total possibile pairs	580503
Proportion of actual pairs on population (%)	0.15

Table 2. Participants per country of origin, 2005-2011

	Number	%
Italy	511	47.4
EU countries	141	13.1
United States	127	11.8
India	72	6.7
China	44	4.1
Russian Fed.	40	3.7
Utd Arab Em.	16	1.5
Canada	13	1.2
Turkey	13	1.2
Japan	11	1.0
Rest of the World	90	8.3
Total	1078	100.0

Table 3. Agreements and participants per SIC division, 2005-2012

	Agreements		**Participants**	
	Number	%	Number	%
A Agriculture	1	0.2	2	0.2
B Mining	20	3.2	34	3.2
C Construction	12	1.9	17	1.6
D Manufacturing	213	33.8	504	46.8
E Transp., Comm., Energy, Sanitary Serv.	91	14.4	160	14.8
F Wholesale Trade	58	9.2	26	2.4
G Retail Trade	30	4.8	26	2.4
H Finance, Insurance, Real Estate	90	14.3	163	15.1
I Services (personal and business)	114	18.1	136	12.6
J Public Administration	2	0.3	10	0.9
Total	631	100.0	1078	100.0

Table 4. Logit models for the probability of inter-firm agreements

Prior correction bayesian method for rare events

	1	2	3	4	5	6	7
Spatial proximity							
inverse geographic distance	0.194 ***	0.105 ***	0.011	0.104 ***	0.100 ***	0.096 ***	0.092 ***
	(0.021)	*(0.037)*	(0.044)	*(0.037)*	*(0.037)*	*(0.037)*	*(0.037)*
same country		0.990 ***	1.138 ***	1.000 ***	1.009 ***	1.141 ***	1.126 ***
		(0.213)	*(0.219)*	*(0.213)*	*(0.212)*	*(0.214)*	*(0.212)*
same region			0.847 ***				
			(0.256)				
Technological proximity							
same division (SIC1)		0.935 ***	0.934 ***	0.934 ***	0.941 ***	0.917 ***	0.932 ***
		(0.120)	*(0.120)*	*(0.120)*	*(0.120)*	*(0.120)*	*(0.120)*
same major group (SIC2)		2.758 ***	2.759 ***	2.759 ***	2.757 ***	2.729 ***	2.744 ***
		(0.168)	*(0.168)*	*(0.167)*	*(0.167)*	*(0.167)*	*(0.168)*
same industry group (SIC3)		3.220 ***	3.208 ***	3.215 ***	3.248 ***	3.263 ***	3.295 ***
		(0.203)	*(0.205)*	*(0.204)*	*(0.202)*	*(0.201)*	*(0.199)*
same industry (SIC4)		3.978 ***	3.986 ***	3.976 ***	3.972 ***	4.009 ***	4.029 ***
		(0.164)	*(0.164)*	*(0.163)*	*(0.164)*	*(0.160)*	*(0.162)*
Organisational proximity							
same group		2.665 ***	2.716 ***	2.703 ***	2.614 ***	3.046 ***	2.718 ***
		(0.754)	*(0.775)*	*(0.756)*	*(0.749)*	*(0.840)*	*(0.778)*
Institutional proximity							
same status		0.154 *	0.153 *		0.178 **	0.187 **	0.190 **
		(0.094)	*(0.094)*		*(0.093)*	*(0.092)*	*(0.092)*
both partners independent				0.152			
				(0.192)			
Social proximity							
inverse geodesic distance		0.178 ***	0.176 ***	0.181 ***		0.194 ***	0.175 ***
		(0.061)	*(0.061)*	*(0.062)*		*(0.058)*	*(0.057)*
inverse geodesic distance (previous 5 years)					0.122 *		
					(0.065)		
Network characteristics							
preferential attachment - partner 1	0.084 ***	0.072 ***	0.073 ***	0.072 ***	0.073 ***		
	(0.007)	*(0.009)*	*(0.009)*	*(0.009)*	*(0.009)*		
preferential attachment - partner 2	0.066 ***	0.060 ***	0.061 ***	0.061 ***	0.064 ***		
	(0.009)	*(0.010)*	*(0.010)*	*(0.010)*	*(0.010)*		
Transitivity - partner 1, betweenness						120.356 ***	
						(36.553)	
Transitivity - partner 2, betweenness						3.409	
						(36.222)	
Transitivity - partner 1, closeness centrality							23.859 ***
							(5.382)
Transitivity - partner 2, closeness centrality							18.613 ***
							(5.933)
Estimated probability Y=1|X at median values (%)	2.4	4.6	4.6	4.0	4.8	3.8	3.0

See Appendix for definitions of variables

Numbers of observations: 9757. Proportion of ones:zeros observations equal to 1:10

All models include individual firm controls for status (listed, private), organization (independent, subsidiary), ownership nationality (Italian, foreign), SIC1 division, geographic location (Italy:North-west, North-east, Centre, South, Islands; another EU country; rest of the world) and interaction terms for geographic location-macrosectors (macrosectors: manufacturing, services)

Geodesic and geographic distance are log-transformed

Robust standard errors in parenthesis. Significance level *** 1%, ** 5%, *10%

Table 5. Effects of proximities and networks on the probability of inter-firm agreements

All changes are equal to one standard deviation and are measured with respect to the median values

From Model 2 Table 4: Prob (Y=1|X)=0.04617	**Standard deviation**	**Absolute difference**	**Percentage Increase**
Spatial proximity			
geographic distance	3321.7	0.0046	10.0
same country	0.458	0.0246	53.2
Cognitive proximity - Technological relatedness			
same division (SIC1)	0.424	0.0208	44.9
same major group (SIC2)	0.168	0.0248	53.7
same industry group (SIC3)	0.122	0.0205	44.3
same industry (SIC4)	0.184	0.0444	96.2
Organisational proximity			
same group	0.039	0.0048	10.3
Institutional proximity			
same status	0.470	0.0033	7.0
Social proximity			
geodesic distance	0.036	0.0004	0.8
Network characteristics			
preferential attachment (partners average)	4.353	0.0146	38.7

All effects are calculated by the Bayesian method and are significant at the 5% significance level