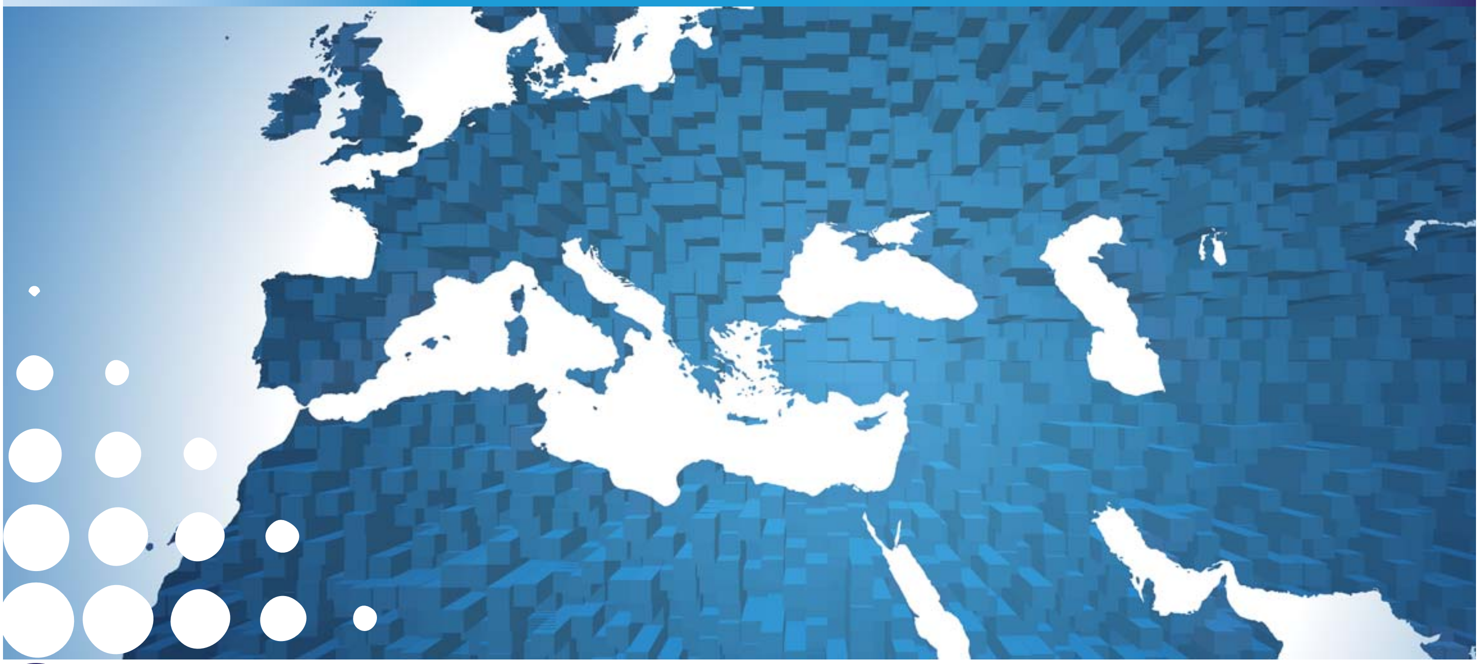
WP4/21 SEARCH WORKING PAPER

Networks, proximities and inter-firm knowledge exchanges

Stefano Usai, Emanuela Marrocu, Raffaele Paci

September 2013



SEARCH
SHARING KNOWLEDGE ASSETS: INTERREGIONALLY COHESIVE NEIGHBORHOODS

EUROPEAN COMMISSION
European Research Area
SEVENTH FRAMEWORK PROGRAMME
Funded under Socio-economic Sciences & Humanities

The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2010-2.2-1) under grant agreement n° 266834

Networks, proximities and inter-firm knowledge exchanges

Stefano Usai, Emanuela Marrocu, Raffaele Paci

University of Cagliari, CRENoS

Abstract

Building on previous literature providing extensive evidence on flows of knowledge generated by inter-firm agreements, in this paper we aim to analyse how the occurrence of such collaborations is driven by the multi-dimensional proximity among participants and by their position within firms' network. 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 in terms of centrality and closeness. 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 localised in Italy. We consider the full range of economic activities and this allow us to offer a general scenario and to specifically investigate the role of technological relatedness across different sectors. The econometric analysis, based on the logistic framework for rare events, yielded three noteworthy results. First, all the five dimensions of proximity jointly exert a positive and relevant effect in determining the probability of inter-firm knowledge exchanges, signalling that they are complementary rather than substitute channels. Second, the higher impact on probability is due to the technological proximity, followed by the geographical one, while the other proximities (social, institutional and organizational) have a limited effect. Third, we find evidence on the positive role played by networks, through preferential attachment and transitivity 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 has 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 seminar in Nizza.

June 2013

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 bound 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 be also mediated by other dimensions of closeness, which may have an a-spatial nature, such as technological, institutional or organisational proximity (Torre and Gilly, 2000; Boschma, 2005). Moreover, interactions among economic agents create social links that, over time, tend to evolve in wider networks, which are likely to facilitate the exchanges of knowledge and moderate the adverse effects of other distances (Boschma and Frenken, 2009).

A growing body of empirical research has extensively analysed the characteristics of networks that are expected to prompt innovation diffusion by considering various forms of connections among agents. These include participation in research programmes (Autant-Bernard et al., 2007; Maggioni et al., 2007; Balland 2012), co-patenting (Cantner and Meder, 2007; Maggioni et al., 2007; Cassi and Plunket, 2012), citations (Maurseth and Verspagen, 2002; Paci and Usai, 2009), co-publications (Ponds et al., 2007), applicant-inventors relationships (Maggioni et al., 2011; Picci, 2010) and human capital mobility (Miguelez and Moreno, 2011; 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 agreement among firms, such as joint ventures and strategic alliances. The management literature (Kogut, 1988; Inkpen, 2000; Oxley and Sampson, 2004, Phelps et al., 2012) has remarked how such inter-firm agreements, regardless their specific nature and motivation, create the conditions for knowledge sharing and thus represent an important channel of knowledge exchanges 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 organisations and thus restricted to their members (Muthusamy and White, 2005; Janowicz-Panjaitan and Noorderhaven, 2008; García-Canal et al., 2008). Since 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.

The aim of this paper is to analyse 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 they are involved in. 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 proximity. Moreover, on the basis of the past experience of each firm within networks, we assess whether the preferential attachment and the transitivity characteristics have an additional effect on the occurrence of inter-firm agreements.

We base our empirical analysis on announced agreements over the period 2005-2012 in which at least one firm is localised in Italy, including both domestic and international collaborations. In total we examine 631 agreements that involved 1078 firms. An original feature of our study is that we consider agreements covering all economic activities and this allow 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 limited the investigation 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 system (Balland, 2012), wine (Giuliani, 2010) and genomics (Cassi and Plunket, 2012). Other studies give a global picture of the role of proximity with respect to the whole economy, but were conducted on data aggregated at a regional level (Marrocu et al., 2013; Maggioni et al., 2013). Our study represents a novel contribution in investigating five dimensions of proximity within a multi-sector framework and in testing whether they act as substitutes or complements in nowadays complex economic systems.

Given the large sample dimensionality difficulties that arise when analysing firm partnerships across different countries – because the network structure becomes virtually worldwide, reaching high degrees of complexity – we chose to restrict our sample to the set of agreements with at least one firm located in Italy. This allows us to end up with a manageable dataset and computationally less demanding proximity and network indicators for each pair of firms. It is worth noting that our choice is unlikely to alter the general validity of the results because, as noticed by Narula and Hagedoorn (1999), the firms' propensity to start an agreement is much more influenced by sectoral heterogeneity than by country differences. This is confirmed by looking at the distribution of agreements by number of partners: in our sample, comprising agreements with at least one Italian participant, 90% of them involve only two partners; such proportion is very similar to the one (88%) reported by García-Canal et al. (2008) for 15 countries of the European Union. Also the geographical distribution of the deals does not show any relevant country-specific feature: 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 happens in France (11%) or in Germany (12%).

The econometric analysis, based on the logistic framework for rare events, yielded three main results. First, all the five dimensions of proximity jointly exert a positive and relevant effect in determining the probability of inter-firm knowledge exchanges, signalling that they are complementary rather than substitute channels. Second, technological proximity exhibits the higher impact on probability, followed by the geographical one, while the other proximities (social, institutional and organizational) have a limited effect. Third, firms network positioning, in terms of both preferential attachment and transitivity, significantly enhances the probability of inter-firm agreements.

The remainder of the paper is organised 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 deal with some estimation issues and present our estimation strategy. The econometric results are discussed in section 5, while some 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 across companies. Data on announced agreements over the period 2005-2012 are collected from the SDC Platinum database (Thomson Financial) and they include all the deals that comprise at least a partner located in Italy. Our data on inter-firm agreements contain both joint ventures and strategic alliances. A joint venture is defined as a cooperative business activity, formed by two or more firms, which creates an independent organisation and allocates the ownership, the operational responsibilities and the financial risks and rewards to each partner, while preserving their separate identity. 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, etc.) which 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 organisations, of which 511 are Italian. Agreements can be simple or complex depending on the number of potential partners involved. Table 1 clearly shows that most partnerships (570) do not go beyond the simplest form, which is a pair of firms involved. Only 10% of total exchanges involve more than two partners with a maximum of seven organizations engaged. Given the presence of deals with multiple partners the number of actual pairs – 887 – is higher than the number of agreements, as it is shown in Table 1. The firm dyads are formed either in joint ventures (607) or in strategic alliances (280).

Table 1 offers interesting information also 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 different status, like pending, letter of interest or renegotiated. It is important to remark that as we are using the agreements as a proxy of knowledge exchanges among partners and given that these exchanges take place also in the preliminary and earlier stage of the contract, independently from their successive progress, we prefer to consider all the announced agreements in our analysis. In the robustness analysis we test whether there are significant differences between completed and uncompleted deals. Another interesting aspect concerns the localization of partners, that is the place where the headquarter is located. Table 1 shows that most pairs are formed by an Italian and a foreign partner (72%), while in only 15% of cases both partners are located in Italy and in 14% they are both foreigners.[1]

Given our interest in spatial proximity, in Table 2 we report the geographical location of the participants; these are located in 61 different countries all over the world. Most partners (47%) are obviously firms located in Italy, followed by those situated in other EU countries (13%), which, as expected, represent the most common partners for the Italian companies, due to their closeness in terms of space and other proximity dimensions. Widespread exchanges are also recorded with partners located in the United States with a share of almost 12%. Interestingly, among most common partners we also find firms located in the 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 the economic sectors, according to the Standard Industrial Classification (SIC) divisions. As expected, most agreements refer to manufacturing (almost 50%), whereas another big fraction (42.5%) refers to service sectors, such as Personal and Business ones, Finance Insurance and Trade and Transportation, Energy and Sanitary services. Almost the same shares can be found for the distribution of participants, but for the fact that the manufacturing sector has a lower quota (around one third).

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, organisational,

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

institutional and social measures of proximity and of the individual firms' network positions. The general form of our empirical model is:

Prob (agreement between any two firms) = f (proximities, firms' network positions, firms' controls) (1)

In this section we discuss the rationale for including the five dimensions of proximity and the two network indicators, as well as describing in detail how they are measured. The list of variables is reported in the Appendix 1.

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 which 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. In order to identify the potential firm dyads, we apply the approach followed by Autant-Bernard et al. (2007) and Cassi and Plunket (2012), among others. This requires pairing the 1078 firms, involved in the 631 agreements included in our sample, to obtain the number of all possible pairs, which is 580,503.[2] Of this total, 887 pairs were involved in actual agreements, while the remaining 579,616 ones were not and therefore they are considered as potential pairs. Thus in our sample the number of firm pairs involved in agreements is equal to 0.15% of total potential 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 deal with the estimation issues.

3.2 Proximity dimensions

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 both at the micro and the macro level (Rallet and Torre, 1999; Romer, 1986). Such diffusion can be facilitated when knowledge, especially in its tacit component, can be transmitted among agents which are physically proximate (Von Hippel, 1994). Consequently, the spatial proximity has been the most thoroughly investigated dimension by the wide literature on knowledge spillovers and flows (Jaffe, 1986; Jaffe et al., 1993; Anselin et al., 1997). We measure geographical proximity by the inverse of distance (*Inv_dist*) between the locations of the partners (in kilometers).[3] As an alternative, spatial closeness between partners is also measured by means of a binary variable (*ID_intra_reg*) which takes value 1 when both partners are located in the same Italian region.

Technological or cognitive proximity. It is a commonly accepted idea that knowledge transfer is not an easy and smooth process open to everybody (Cohen and Levinthal, 1990). It may require specific and appropriate absorptive capacity, which entails a homogenous cognitive base in order to understand and effectively process the available knowledge (Nooteboom, 2000). In practical terms, we expect that firms having a similar knowledge base exchange knowledge more easily and efficiently. We account for the technological relatedness between partners with a set of five mutually exclusive technological interaction dummies ordered by increasing technological similarity

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

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

(Ellwanger and Boschma, 2012). These dummies are based on the primary economic activity, which is reported in the SDC database at the 4digit SIC code for each participant.[4] The interaction dummy (*ID_intra_SIC4*) takes value 1 when the partners operate in the same 4digit SIC Industry and value 0 when the two firms operate in different Industries. It is interesting to note that this strong sectoral affinity is not unusual, since it occurs in almost 28% of the firm pairs involved in the agreements. The dummy (*ID_intra_SIC3*) takes value 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 industrial disaggregation.[5] With the same methodology we compute the dummies (*ID_intra_SIC2*) for the 2 digit SIC Major group, (*ID_intra_SIC1*) for the 1 digit SIC Division and finally the dummy (*ID_inter_SIC1*) 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 happens in 316 out of 887 cases (around 35%).

Organisational proximity. The exchange of information and knowledge can be influenced by the membership of individuals to the same club, group or organisation, 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 behaviour (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 organised firms. In our empirical analysis, as in Balland et al. (2013), we measure organizational proximity with a dummy variable (*ID_intra_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, firms in our case, 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, favours 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, 2012), the institutional similarity is measured by means of a dummy variable based on the status of the two partners. More specifically, the dummy (*ID_status*) takes value 1 if the two firms share the same institutional status (both listed on a stock exchange, or private, or subsidiaries, or government bodies). In our data, for 38% of pairs involved in the agreements the two partners are institutionally similar. As an alternative measure, we also compute a dummy (*ID_indep*) taking value 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 presence of social ties among individuals is another important catalyst factor for the exchange of ideas and knowledge (Granovetter, 1985). The analysis of social networks is therefore vital to understand the phenomena of knowledge creation

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

[5] 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).

and diffusion. Social proximity refers mainly to reputation and trust effects created by the experience of past collaborations and repeated contacts between partners. Previous experience, which breeds reputation and trust, 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 by using also pre-sample information on agreements announced in the past, starting from the year 2000. We assume that direct and indirect relationships in the past 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 decreases with geodesic distance, which measures the shortest path between two nodes (i.e. firms). Therefore, our social proximity indicator is the inverse of the geodesic distance and goes from zero (when two nodes are virtually infinitely distant, they never met in the past, nor did any of their direct and indirect partners) to one (when two nodes are directly linked because they have been partners in the past). In order to extract as much information as possible from our datas we compute the inverse of the recursive measure of geodesic distance (*Inv_geod_rec*) between firm i and firm j in all available previous years.[6] A robustness test is also performed by using the inverse of the geodesic distance computed considering only the previous five years (*Inv_geod_5y*).

3.3 Individual network characteristics

Social ties may be the result of an individual attitude or customary behaviour and must be thus examined from two complementary perspectives: the single node's one together with the whole system's perspective (Bramanti and Maggioni, 1997). This is why we introduce two additional measures, which take into account the single social position of firms within the network of potential ties. Such measures supplement the information on the bilateral notion of social proximity, discussed above. Theoretical literature has shown how networks architecture may impact on knowledge growth and diffusion (Cowan and Jonard, 2004, Cowan et al., 2004 and Ter Wal and Boschma, 2009). Consequently several empirical studies (such as Balland et al., 2013, Autant-Bernard et al., 2007, Cassi and Plunkett, 2012, Giuliani, 2010) have investigated to what extent the position of firms within the network influences knowledge diffusion. We follow this research path by introducing two measures on the network characteristics of individual firms over previous years starting from the year 2000, which are expected to account for their previous experience in partnering.

Preferential attachment hypothesis. According to this hypothesis actors are more inclined to link to the most connected individuals. Agents with a large number of relations are more attractive, as they are supposed to be more productive or more trustworthy (Barabasi and Albert, 1999). 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 its degree of centrality (*P_deg*).

Transitivity hypothesis. This 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 in order to get 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. In our work we prefer to employ a more general concept and indicator, since triadic closures

[6] 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 going from 2000 to 2011.

in our sample are very rare. We thus measure the transitivity property by referring to the notion of closeness centrality (*P_clo*), that is the inverse of the sum of the distances of a node to all other nodes. This indicator can be regarded as 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 expected sign for both indicators – preferential attachment and transitivity – is positive since firms are supposed to be willing to maximise the opportunity to get knowledge from the whole network thus connecting to the more joined and central firms. However, it may also possible to observe a negative effect if firms are worried that linking to a central and highly connected partner may jeopardize the appropriability of their knowledge (Autant-Bernard et al., 2007).

3.4 Individual firms' characteristics

In our empirical model we also control for various characteristics at the firm level. More specifically, for each firm we include information on its status, organization, ownership nationality, geographical location and principal sector of activity. Regarding the status we have computed two dummies (*PD_listed* and *PD_private*) to account for the firm being publicly traded on a stock exchange market or a private company, that is a company owned either by non-governmental organizations or by a relatively small number of shareholders, often a family in Italy.[7] We have also included a dummy (*PD_indep*) taking value one when the participant is an independent firm (i.e. when the ultimate parent company corresponds with the partner itself) and a dummy (*PD_fo*) taking value one for foreign-owned companies. The categorical variables, like economic activity and spatial location, have been transformed in dummy variables defined for each of their categories. In this way we have computed ten mutually exclusive dummies for the 1 digit SIC Divisions of economic activity. Five mutually exclusive dummies have been created for the firm's spatial localisation, which can be in one of the three Italian macro regions - North, Centre and South - or in one of the European Union countries or in an extra-EU country.

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 dyads of firms (579,616) that could have set up an agreement but did not.

Comparing the high number of potential pairs with the one related to actual deals (0.15%), 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. 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"). It is important to note that selecting on the zeros allows also for a more efficient data collection because only a small part of such observations contribute to the information content of the explanatory variables. As it is well known, data selection based on Y induces bias, therefore it is necessary to apply the

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

appropriate statistical corrections in order 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 one observations.
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. In order to attenuate the possible selection bias we apply the independence in conditional-mean approach by including in our models a wide range of firm's characteristics, which are jointly likely to affect firm's 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 features, we expect that the decision to select a particular partner in order to carry out a specific agreement is independent of higher-level of collaboration or acquisition decisions.[8] Therefore, the empirical specification for the probability of observing an agreement is formalized on the basis of the cumulative logistic distribution as follows:

$$Prob \;\; Y_{ij} = 1 \; X_{ij}, N_i, N_j, W_i, W_j \;\; = \frac{1}{1+e^{-\; X_{ij}\beta_1 + N_i\beta_2 + N_j\beta_3 + W_i\beta_4 + W_j\beta_5}} \qquad (2)$$

where the matrix X_{ij} includes the interaction 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 two network indicators for firm *i* and *j*, respectively, whereas the matrices W_i and W_j comprise the individual control variables.
We estimate model (2) by performing the sequential procedure suggested in King and Zeng (2001) for selecting the zero observations. More specifically, we considered several random samples by starting with the proportion of ones/zeros observations equal to 0.5 (each actual pair matched with just a random control) and stopping when we get no further efficiency gains, signalled by a reduction in the standard errors magnitude. This occurred for the 0.1 proportion (1 actual pair matched with 10 other randomly drawn potential pairs) sample for both the prior correction and the weighting method.[9] Comparing the alternative correction approaches we found that overall the estimated coefficients did not differ substantially, thus signalling the absence of any clear misspecification problem. We interpret this result in favour of our highly parameterized specification, which simultaneously accounts for five different proximity dimensions, network features and for 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. Results on model comparisons across correction methods and different sample sizes are reported in Appendix 2.[10]

5. Empirical results

[8] The same approach is adopted by Chakrabarti and Mitchell (2013) for the case of M&A determinants.
[9] For a thorough discussion on the correction methods refer to King and Zeng (2001). Some technical aspects are also reported in Appendix 2.
[10] All estimations are carried out by the ReLogit software by Tomz et al. (1999).

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 of inter-firm proximity act as substitutes or complements and to assess whether the almost undisputed effect of geographical closeness is maintained when the role of other proximities is taken into account. The subsequent models 3-5 address the robustness of our results across specifications that include alternative indicators for some proximity measures.

5.1 The baseline model

In column 1 we report the benchmark model according to which knowledge flows are affected by geographical proximity and by the network characteristics of each partner; controls at individual firm level are included to account for firm heterogeneity. Results show that geographical nearness is a crucial determinant of such flows and also that preferential attachment (degree of centrality) significantly influences the cooperation decisions among firms. At the same time, closeness centrality has the right sign but it results only marginally significant. Most importantly, this model estimates the probability that any two firms start an agreement process (see last row in column 1) at 2.7%, which is eighteen 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, organisational, institutional and social, discussed in section 3. Results show that, first of all, geographical closeness remains relevant even when all other dimensions are controlled for; its estimated coefficient does not change if not negligible (from 0.254 to 0.256). This evidence is in line with the conclusions by Paci et al. (2013) that geography and the other dimensions of proximity are not substitutes but rather complements. As a matter of fact, results in column 2 show that all dimensions of proximity exhibit a positive and significant coefficient, which is expected to imply an increased probability of knowledge exchange through inter-firm agreements.[11] As a matter of fact, the estimated probability for this model rises up to 3.7%, which implies that, thanks to the introduction of all proximities, cooperation becomes 40% 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 (Kirat and Lung, 1999; Torre and Gilly, 2000).

Another noteworthy aspect is that the coefficients of technological proximity are not only positive and significant, but that their values are increasing with the degree of similarity of firms' productive and knowledge base. 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 (3.97) one is associated to the case when both firms operate in the same SIC4 industry. This result confirms recent findings by Ellwanger and Boschma (2012) on the relevant role played by industrial relatedness in favouring 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

[11] Basile et al (2012) provide evidence on the positive and synergic effects of different kinds of proximities – such as spatial, social and relational – as important channels of knowledge spillovers.

scale and scope. Moreover, information asymmetries between firms which are highly technologically related are lower and this favours their exchange of knowledge (Hussinger, 2010).

As far as social proximity is concerned, the inverse of the geodesic distance is always positively and significantly linked to the probability that two firms announce an agreement, as in Autant-Bernard et al. (2007). Moreover, the fact that two firms share the same ownership status, that is that they are institutional proximate, is also a favouring factor for collaborative agreements, as in Cassi and Plunket (2012). Finally, we find a positive and significant effect of organizational proximity, measured in terms of membership to the same group, confirming the results by Balland (2012) and Balland et al. (2013).

We also find evidence supporting the relevant role of network characteristics as the two indicators of centrality have a positive sign and are significant for both partners. Regarding the preferential attachment hypothesis (degree of centrality), we confirm previous findings (Balland et al., 2013, Balland, 2012, Cassi and Plunket, 2012) that agents prefer to interact and exchange knowledge with those which have former agreement experiences. Such a preference induces a self-reinforcing process of collaboration around the most 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 having already experienced knowledge exchanges, possess more information as a result of those exchanges. Previous experience is also interpreted as an indirect signal of the potential value of a firm as a partner. Moreover, we find that the closeness centrality of firms within the network is an additional facilitating factor for knowledge exchanges providing evidence in favour of the transitivity hypothesis. Our findings thus confirm that firms are chosen for cooperation because their position within the network makes them potentially more able to connect to all other nodes in order to get external knowledge.

It is important to note that the positive and significant signs exclude that another contrasting effect occurs and prevails, the one 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 getting in contact with firms which are in the best position, not only for collecting knowledge, but also for spreading it.

5.2 Robustness tests

In the last three columns of Table 4, we test the robustness of our results across specifications that include alternative indicators for the proximity measures.

In column 3, we find that spatial proximity affects knowledge exchanges also when we refer to the sharing of the same regional location by partners. All other regressors keep their signs and significance. However, the average estimated probability declines to 2.8%, which is definitely lower than the one obtained from model 2, which remains the preferred one.

In column 4 we include an alternative measure for institutional proximity, i.e. the fact that both partners are independent entities; however, it turns out to be not significant, while leaving almost unchanged the coefficients of the other regressors. In column 5 we also test the robustness of our result with respect to a different proxy for social proximity; more precisely, we now calculate the inverse of the geodesic distance limiting the time span to the previous five years in order to consider the same time span for all observations. We find that this variable is only partially significant (10%), even though the average probability remains almost constant.

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 prevent us to estimate all the sub-samples and thus further research is required. In any case, no significant differences were found across subsamples, thus confirming the main findings discussed above for the full sample.

5.3 Effects on probability

In this final section, unlike previous contributions in the literature on proximity, we take a step further in assessing how changes in proximity or network features affects 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 it 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, when median values are attributed to all variables, equal to 3.7%.

The first most interesting result is that, as in Paci et al. (2013) and Montobbio and Sterzi (2013), the highest impact on probability is found when the technological proximity measured by the sharing of the same industry increases by one standard deviation with respect to the median. The probability goes up to around 7.4%, with an increase of 98% with respect to the baseline estimation. Increases of around 50% are also registered for all other measures of technological proximity.

Regarding the geographical proximity, one standard deviation change makes the estimated probability increase by 23% (from 3.7% to 4.7%), which is around one quarter of the effect produced by a change in the same-industry dummy. As a matter of fact, the same effect induced by an increase in the highest degree of technological relatedness would be obtained with a reduction in the geographical distance as remarkable as moving from the median value (1728 km) to just 100 kilometers.

As for the other proximities, the effect on the estimated probability is always positive but smaller: a change in the organizational proximity induces a change of 12.2%, while 9.3% and 0.8% are the increases in probability due to institutional proximity and social proximity, respectively. Despite the modest influence of social proximity, we find that firm’s own social relations are much more effective. Considering the partners average effects, the preferential attachment raises the probability of observing an agreement by 31.1%, while 8.1% is the increase due to the transitivity property.

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 a reciprocal moderate effects, proximities and social links are by no means interchangeable, they supplement each other in contributing to favor 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 literature has provided extensive evidence on the knowledge flows generated by inter-firm agreements as partners during the various stages of the agreement process share knowledge-based resources, often embedded within the organizations. These flows are derived by the integration and transfers of new technologies and organizational capabilities and also by the access to formal and informal organizational learning processes.

More specifically we assess the role played by five different types of proximity (geographical, technological, organisational, institutional and social) and by the position of participants within the network of previous ties. We also control for firms heterogeneity by introducing individual characteristics like status, organization, ownership nationality, principal sector of activity and geographic location.

We apply our empirical analysis to the case of announced agreements over the period 2005-2012, in which at least one firm is localized in Italy. We consider 631 agreements, which involve 1078 unique firms and give rise to 887 pairs of partners. The analysis of the effects of networks and proximities on the probability that two firms set up an agreement, and therefore exchange knowledge, 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.

The results of our preferred model − which include five different proximity measures, the two network effects and a full set of individual controls − show that all dimensions of proximity exhibit a positive and significant coefficient. Thanks to the concurrent effect of the five proximity dimensions the probability of knowledge exchanges through inter-firm agreements raises up to 3.7%, which is 25 time higher than the random probability. Our results underline that knowledge exchanges are facilitated not only by spatial proximity, as argued by the traditional approach, but also by other dimensions of inter-firm closeness, like sharing a common cognitive base, have the same institutional background, being a part of the same organisation, belong 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, rather than geographical proximity. The latter, however, remains much more effective than the other proximities in driving the agreement process. Organizational, institutional and social proximities facilitate the exchange of knowledge with a significant but smaller contribution. Despite the modest influence of social proximity, the relative importance of network links is evidenced by the presence of preferential attachment and transitivity effects. There is robust evidence that the degree of centrality, that is firm's previous experience, within the existing networks positively affects the probability that companies set up cooperation agreements and thus give rise to a knowledge exchange. At the same time firms are attracted by those partners which are on average closer to all other firms, as they are probably more able to obtain and process information across the network.

This highlights the importance of analysing inter-firm knowledge flows simultaneously accounting for the whole set of relevant proximities and network features within a comprehensive empirical specification.

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*, doi:10.1093/jeg/lbs023

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

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.

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.

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. and Plunket A. (2012) Research collaboration in co-inventor networks: combining closure, bridging and proximities, *MPRA working paper* n. 39481

Chakrabarti A. and W. Mitchell (2013) The Persistent Effect of Geographic Distance in Acquisition Target Selection, Organization Science, online before print, DOI: 10.1287/orsc.1120.0811.

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. (2012) Who acquires whom? The role of geographical proximity and industrial relatedness in Dutch domestic M&As between 2002 and 2008, mimeo.

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* 10–14.

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

Inkpen A.C. (2000) Learning through joint ventures: a framework of knowledge acquisition. *Journal of Management Studies*, 37, 1019–43.

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

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, 8, 1337-1355

King G. and Zeng L. (2001) Logistic regression in rare events data, *Political Analysis*, 9(2), 137-163.

King G. and Zeng L. (2002) Estimating risk and rate levels, ratios and differences in case-control studies, *Statistics in Medicine*, 21(10), 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.

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. (2013) 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(8), 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.* Doi: 10.1016/j.techfore.2013.03.004.

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.

Miguélez E. and Moreno R. (2011) Research Networks and Inventors' Mobility as Drivers of Innovation: Evidence from Europe, *Regional Studies*, DOI: 10.1080/00343404.2011.618803

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

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.

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

Oxley, J.E., 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. (2013) The complementary effects of proximity dimensions on knowledge spillovers, forthcoming in *Spatial Economic Analysis*.

Phelps C., Heidl R. and Wadhwa A. (2012) Knowledge, Networks, and Knowledge Networks : A Review and Research Agenda, *Journal of Management*, published online, DOI: 10.1177/0149206311432640.

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, *The 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.

Von Hippel E. (1994) “Sticky Information" and the Locus of Problem Solving: Implications for Innovation, *Management Science*, 40(4), 429-439

Appendix 1. Variable definitions

Dependent variable		
	Y	dummy: = 1 if the two participants have an agreement; = 0 otherwise
Interaction variables for proximity dimensions between each partner in a pair		
Spatial proximity		
	Inv_dist	inverse of the distance in km between partners cities (log)
or	ID_intra_reg	dummy: = 1 if partners are located in the same Italian region
Technological proximity		
	ID_intra_SIC4	dummy: = 1 if the highest degree of industrial relatedness is at SIC4
	ID_intra_SIC3	dummy: = 1 if the highest degree of industrial relatedness is at SIC3
	ID_intra_SIC2	dummy: = 1 if the highest degree of industrial relatedness is at SIC2
	ID_intra_SIC1	dummy: = 1 if the highest degree of industrial relatedness is at SIC1
Organisational proximity		
	ID_intra_group	dummy: =1 if partners belong to the same group
Institutional proximity		
	ID_status	dummy: = 1 if partners have the same insitutional status
or	ID_indep	dummy: = 1 if both partners are independent companies
Social proximity		
	Inv_geod_rec	inverse of geodesic distance with recursive window (log)
or	Inv_geod_5y	inverse of geodesic distance with 5-year window (log)
Network characteristics for each partner		
Preferential attachment		
	P_deg	degree centrality, number of links incident upon a node
Transitivity		
	P_clo	closeness centrality, inverse of the sum of its distances to all other nodes
Control dummies for individual characteristics of each partner		
	PD_listed	partner is publicly traded on a stock exchange
	PD_private	partner status is private
	PD_indep	partner is independent, it is not part of a group
	PD_fo	partner is owned by a foreign ultimate parent company
	PD_north	partner location in northern Italy
	PD_centre	partner location in central Italy
	PD_south	partner location in southern Italy
	PD-EU	partner location in another EU countries
	PD_div1 - 10	partner economic activity in 10 SIC divisions (ten dummies)

Appendix 2. Methodological issues

For a comprehensive discussion on the correction methods for logistic models for rare events refer to King and Zeng (2001a). The two mostly applied correction methods are the prior correction method and the weighting method. The first one is less computationally demanding because it entails only correcting the estimate constant on the basis of population proportion of ones. Assuming β_0 is the estimated constant coefficient, the prior correction emends it by means of the following expression: $\beta_0 - ln \; \frac{1-\tau}{\tau} \; \frac{y}{1-y}$, where τ and y are the population and the sample fraction of ones, respectively. The Maximum likelihood estimators for the coefficients associated with the explanatory variables do not need any correction since they maintain their unbiasedness are consistency property.

By applying the weighting method the sample observations are weighted so that the weighted proportion of ones and zeros in the sample equals the true proportion in the population. The weighting method is robust to potential misspecification (Manski and Lerman, 1977), but it requires further corrections since the MLE for the variance-covariance matrix is severely biased.

The Table A2 below compares the baseline specification (model 2, Table 4) across different sample sizes (obtained by increasing the number of zero observations) and correction methods.

Table A2. Alternative sample sizes and correction methods for the baseline model

Logit models for the probability of inter-firms agreements

	1	2	3	4
		baseline		
Correction for rare events	**Prior correction**	**Prior correction**	**Weighting**	**Weighting**
Proportion of ones : zeros observations	1:5	1:10	1:5	1:10
Spatial proximity				
inverse geographic distance	0.218 ***	0.256 ***	0.235	0.319 ***
	(0.029)	*(0.026)*	*(0.029)*	*(0.027)*
Technological proximity				
same division (SIC1)	0.884 ***	0.939 ***	0.672 ***	0.861 ***
	(0.125)	*(0.112)*	*(0.142)*	*(0.128)*
same major group (SIC2)	2.787 ***	2.733 ***	2.329 ***	2.446 ***
	(0.190)	*(0.167)*	*(0.206)*	*(0.180)*
same industry group (SIC3)	3.392 ***	3.174 ***	3.240 ***	3.104 ***
	(0.248)	*(0.205)*	*(0.262)*	*(0.207)*
same industry (SIC4)	3.920 ***	3.972 ***	3.481 ***	3.759 ***
	(0.204)	*(0.166)*	*(0.195)*	*(0.162)*
Organisational proximity				
same group	3.158 ***	3.073 ***	1.994 **	3.040 ***
	(0.788)	*(0.719)*	*(0.844)*	*(0.753)*
Institutional proximity				
same status	0.195 *	0.199 **	0.126	0.191 *
	(0.104)	*(0.094)*	*(0.124)*	*(0.105)*
Social proximity				
inverse geodesic distance	0.131 **	0.175 ***	0.086 *	0.123 **
	(0.059)	*(0.059)*	*(0.048)*	*(0.050)*
Network characteristics				
preferential attachment - partner 1	0.070 ***	0.068 ***	0.068 ***	0.070 ***
	(0.010)	*(0.008)*	*(0.010)*	*(0.008)*
preferential attachment - partner 2	0.084 ***	0.063 ***	0.055 ***	0.056 ***
	(0.018)	*(0.009)*	*(0.011)*	*(0.009)*
transititvity - partner 1	14.441 *	13.453 **	9.029	10.909
	(7.864)	*(6.889)*	*(11.899)*	*(9.627)*
transititvity - partner 2	11.081	17.127 **	14.117	32.039 ***
	(8.890)	*(8.198)*	*(11.520)*	*(9.693)*
Observations	5322	9757	5322	9757

See Appendix 1 for variables' definitions

All models include individual firm controls for status (listed, private), organization (independent, subsidiary), ownership nationality (Italian, foreign), SIC division, geographic location (North, Centre, South Italy, another EU country, rest of the world)

Geodesic and geographic distance are log-transformed

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

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	1.6	17	1.9
D Manufacturing	213	46.8	504	33.8
E Transp., Comm., Energy, Sanitary Serv.	91	14.8	160	14.4
F Wholesale Trade	58	2.4	26	9.2
G Retail Trade	30	2.4	26	4.8
H Finance, Insurance, Real Estate	90	15.1	163	14.3
I Services (personal and business)	114	12.6	136	18.1
J Public Administration	2	0.9	10	0.3
Total	631	100.0	1078	100.0

Table 4. Logit models for the probability of inter-firm agreements
Prior correction model for rare events

	1	2	3	4	5
Spatial proximity					
inverse geographic distance	0.254 ***	0.256 ***		0.257 ***	0.254 ***
	(0.024)	*(0.026)*		*(0.026)*	*(0.026)*
same region			1.495 ***		
			(0.261)		
Technological proximity					
same division (SIC1)		0.939 ***	0.931 ***	0.939 ***	0.945 ***
		(0.112)	*(0.118)*	*(0.119)*	*(0.119)*
same major group (SIC2)		2.733 ***	2.720 ***	2.732 ***	2.731 ***
		(0.167)	*(0.166)*	*(0.166)*	*(0.166)*
same industry group (SIC3)		3.174 ***	3.123 ***	3.167 ***	3.209 ***
		(0.205)	*(0.207)*	*(0.205)*	*(0.205)*
same industry (SIC4)		3.972 ***	3.984 ***	3.969 ***	3.966 ***
		(0.166)	*(0.167)*	*(0.165)*	*(0.166)*
Organisational proximity					
same group		3.073 ***	3.279 ***	3.128 ***	3.046 ***
		(0.719)	*(0.775)*	*(0.719)*	*(0.720)*
Institutional proximity					
same status		0.199 **	0.208 **		0.223 **
		(0.094)	*(0.093)*		*(0.093)*
both partners independent				0.205	
				(0.193)	
Social proximity					
inverse geodesic distance		0.175 ***	0.169 ***	0.178 ***	
		(0.059)	*(0.058)*	*(0.060)*	
inverse geodesic distance (previous 5 years)					0.120 *
					(0.064)
Network characteristics					
preferential attachment - partner 1	0.076 ***	0.068 ***	0.069 ***	0.068 ***	0.069 ***
	(0.007)	*(0.008)*	*(0.008)*	*(0.008)*	*(0.008)*
preferential attachment - partner 2	0.065 ***	0.063 ***	0.064 ***	0.063 ***	0.065 ***
	(0.009)	*(0.009)*	*(0.010)*	*(0.009)*	*(0.010)*
transititvity - partner 1	9.117	13.453 **	14.362 **	13.330 **	13.661 **
	(5.957)	*(6.889)*	*(6.817)*	*(6.859)*	*(6.894)*
transititvity - partner 2	12.705 *	17.127 **	17.695 **	16.928 **	17.732 **
	(6.761)	*(8.198)*	*(8.055)*	*(8.184)*	*(8.272)*
Estimated probability Y=1|X at median values (%)	2.69	3.77	2.78	3.13	3.88

See Appendix 1 for variables' definitions

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 (North, Centre, South Italy, another EU country, rest of the world)

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)=0.038	**Standard deviation**	**Absolute difference**	**Percentage Increase**
Spatial proximity			
geographic distance	3321.7	0.0087	23.1
Technological proximity			
same division (SIC1)	0.424	0.0173	45.9
same major group (SIC2)	0.168	0.0204	54.1
same industry group (SIC3)	0.122	0.0167	44.4
same industry (SIC4)	0.184	0.0371	98.4
Organisational proximity			
same group	0.039	0.0046	12.2
Institutional proximity			
same status	0.470	0.0035	9.3
Social proximity			
geodesic distance	0.036	0.0003	0.8
Network characteristics			
preferential attachment (partners average)	4.353	0.0117	31.1
transititvity (partners average)	0.005	0.0031	8.1

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