THE GEOGRAPHY OF SCIENCE- INDUSTRY COLLABORATIONS: SIZE AND PROXIMITY EFFECTS

Draft Version

BOUBA-OLGA O., FERRU M. AND PEPIN D.

CRIEF - University of Poitiers

Abstract: The aim of this paper is to provide explanatory elements of the geography of collaboration for innovation, highlighting the respective roles of the geography of scientific and economic resources and various forms of distance (geographic, scientific, sectoral), in a proximist perspective. The original database used is the one that collects Science-Industry agreements signed in France between 1997 and 2006, amounting to around 8,000 contracts. The econometric exploitation of the database in the form of a gravity model allows us to test the importance of size-effects and different forms of proximity effects. Given the distribution of the data used, we realize a generalized Tobit specification. Three important results are revealed. 1) The greater the scientific activity and the economic size of a region, the greater will be the number of collaborations it is involved in; 2) the spatial proximity between actors positively affects the number of collaborations; 3) neither the scientific proximity, nor the sectoral proximity explain significantly the spatial distribution of partnerships.

Introduction

Many empirical studies emphasize the importance of collaboration for the innovativeness of firms, whether these be global (Hagedoorn, 2002), European (Giuri et al., 2006) or French (Sessi, 2006). In 2006, nearly 40% of French companies reported collaboration dedicated to innovation, especially with customers, suppliers (40.3% of them), as well as with universities and public R & D (21.3%)[1]. Therefore, a question arises as to the spatial dimension of these collaborations.

A burgeoning literature concerned with the geography of these collaborations has developed since the 1990s. Initially, a variety of authors addressing the current geography of innovation (Jaffe, 1989, Audretsch and Feldman, 1996) and approaches on the theme of clusters (Porter, 1990; Beccatini, 1992, Saxenian, 1994) emphasized the local nature of innovation due to the essentially tacit knowledge exchanged. The transfer of tacit knowledge requires face-to-face interactions between the players and would be facilitated by spatial proximity: “knowledge traverses corridors and streets more easily than continents and

[1] The remaining 38.4% corresponds to relations with other companies in the same group (17.2%), with other companies (8.1%), and with consultants or private research organizations (13.1%).

oceans "(Feldman, 1994, p. 2). Based on this work, many governments and international institutions have implemented policies to support the emergence of clusters.

In recent years, new works challenged these initial models, both with respect to the forthcoming statistical results, but also in their use of explanatory terms. These different studies (Paci and Usai, 2000; Kenney and Florida, 1994; Gay and Picard, 2007; Giuri et al., 2006) qualify the role of spatial distance. Audretsch and Stephan (1996) in the case of the United States, Autant-Bernard (2000) in the case of France and Hagedoorn (2002) globally, point out the coexistence of local and non-local collaborations. As part of a major study based on patents filed with the European office, Giuri et al. (2006) obtained similar results and, referring to the weakness of local inter-organizational collaborations, explain that this result "is puzzling given the emphasis in the literature on the importance of the geographical proximity for collaboration and knowledge transfer" (p.16).

This ongoing debate about the geography of collaboration remains focused on the operating conditions of collaboration and, particularly, on the difficulties of sharing tacit knowledge between those involved in innovation.

The aim of this article is to provide additional explanatory material to enrich the analysis of the geography of collaboration by given new answers to this debate. Contrary to existing analysis, we consider the problem of partner choice has to be distinguished from location choice concerning science-industry partnerships notably: innovation projects are more and more multi-partners and medium-term period, involving an ubiquity constraint for project leaders. At the same time, thanks to the improvement of ICT and means of transport, actors are able to be here and elsewhere in the same time, they can coordinate with each other from a distance and/or to use a temporary proximity. Taking into account these elements, we assume actors involved in science-industry collaborations do not change their location for every partnership and consequently the geography of collaborations could be explained by resources endowments of regions (endowment in terms of economic resources and scientific resources) and spatial proximity between regions. We have to precise that actors can be mobile but this mobility is costly (costs in terms of transport and more generally transaction costs). We test these two explanations thanks to a gravity model.

Concerning the role of spatial proximity, we propose to distinguish the effect of the co-location of partners, and the effect of distance between regions. The analysis of interregional distance is given depth by measuring it in two complementary ways. The existence of possible effects such as spatial proximity can be explained by recourse to the nature of tacit knowledge. This is the traditionally posed hypothesis in the literature on the geography of innovation[2]. Moreover, following the proximity approach, we propose to test

[2] It can also be explained by other constraints prior to the collaborative process, including constraints on the relationship between those involved in development. The local nature of collaboration is sometimes explained less by the problems attached to the exchange of tacit knowledge than by the fact that those involved may mobilize their social relationships to find their partners (Grossetti and Bes (2003), Ferru (2010)), which have a strong local character. The data used here does not permit discrimination between these different factors.

the role of other forms of proximity which are less pointed out in the literature but which could be also determining in the spatial distribution of partnerships. More precisely, we could assess the relative importance of the sectoral and scientific proximities in addition to the spatial proximity, given that our hypothesis is the closer the specialization structures of regions are, the less they collaborate.

In addition to the analysis of proximity effects, the focus is on determinants which intervene before collaborations take place. The specific aim is to test the existence of structural effects related to the uneven spatial distribution of resources (scientific and economic) that may significantly influence the geography of collaboration. Indeed, it appears that for some low density areas with respect to scientific and economic resources, those involved in innovation have difficulty finding potential partners locally. This may explain, at least in part, the low weight put on “being local” in these regions. Instead, regions boasting many laboratories and academic institutions in R & D, as is the case for the Ile de France in particular, will necessarily raise the number of local collaborations above the average.

Empirically, to evaluate these determinants a database collecting science-industry contracts covering the 1997-2006 period is used. This comprises 7,972 contracts and is supplemented by basic data on economic and scientific resources, as well as interregional distances.

We introduce first our theoretical framework based on the School of proximity (Rallet & Torre, 2005; Boschma, 2005) to underline the potential role of proximity effects. More precisely, we use the typology developed by Bouba-Olga & Grossetti (2008) which sheds light on the role of socio-economic proximity for the geography of knowledge collaborations. After presenting in a second point the data and the methodology used, a third point relating to the results of the gravity model is addressed. This demonstrates the important role played by the geography of resources and notably economic resources, as well as the role of proximity effects, including the advantage inherent in the co-location of partners.

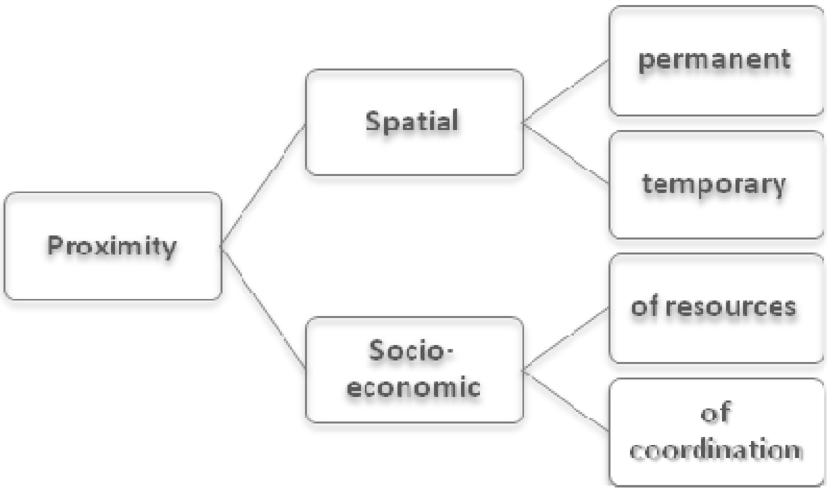
1. A Proximist framework

In the first paragraph it is shown how approaches from the "School of proximity" (Benko, 1998) can surpass the dominant concepts in the geography of innovation. It is possible to consider not only the operation of collaborations but also their construction. The proximist typology developed by Bouba-Olga and Grossetti (2008) seems quite relevant in this regard, in that it takes account of the elements of proximity that particularly link partners.

In general, approaches based on proximity constitute an analytical framework aimed at accounting for the spatial dimension of economic activities through the thorough study of the actors’ needs in co-ordinating. The authors constituting this research group mobilize various theoretical inputs, leading them to propose different analytical formulations of proximity.

Following the work of Perroux (1955), all authors refer to a physical proximity that « deals with the separation in space and the links in terms of distance », with a non-physical proximity which addresses « the economic separation in space and the links in terms of the organization of production » (Gilly and Torre, 2000, p. 12-13). There is a general consensus about the notion of geographical proximity, even though it is acknowledged that the exact of the terms geographical proximity, physical proximity and spatial proximity remain open to discussion. What really matters about spatial proximity is to distinguish the temporary spatial proximity and the permanent one (Rallet and Torre, 2005). This distinction is crucial since the assumption of an efficient coordination thanks to a temporary proximity allows to not assimilate collaborations choices and location choices. In other words, actors mobility reduces the disadvantages of spatial distance between actors involved in innovation projects.

Concerning the non spatial proximity, several typologies have been developed for some years (Boschma (2005) notably). We propose in this paper to use the one developed by Bouba-Olga and Grossetti (2008), which has the advantage of showing that certain forms of proximity also facilitate the construction of collaborations and are key to the geography of collaboration. The authors show the presence, close to physical proximity, of a socio-economic proximity, revealing the importance of embedding the actors in their economic and social environment. They distinguish, within this socio-economic proximity, a proximity of resources and a proximity of coordination (see Figure 1).



Bouba-Olga and Grossetti (2008) reflect on the constraints affecting coordination throughout the collaborative process using the concept of coordination proximity. This allows consideration of the interplay of social networks and the mechanisms of mediation in the coordination of actors, employing the notions of relational proximity and mediation of proximity, respectively. Specifically, the authors consider that there is a close relationship between two actors when social relations led to the rapprochement of the two actors. The proximity of mediation is applicable, in turn, when some mechanism of mediation allows coordination between two actors, a mediation being "anything that allows the exchange without depending on the links of personal relationships" (Grossetti et al., 2006)[3].

Besides the proximity of coordination, the authors underline the importance of **a proximity of resources** which refers to the idea of similarity or complementarity of resources (material or cognitive) that allows the bringing together of actors and fosters collaboration. Partners need to be able to more easily understand and advance their innovation project, with the minimum resources of coordination, a common language, a relatively close knowledge base, etc. In other words, proximity of similar resources is required during collaborations for the exchange of tacit knowledge. Most authors interested in the geography of innovation do so through an approach in terms of proximity (see including Boschma, 2005, Nooteboom et al. 2005, etc.). For the purposes of their innovation projects, the players are also forced to use additional productive resources. Thus, actors look for a proximity of resources to find complementary resources: in science-industry collaborations, laboratories look for complementary companies and *vice versa* companies look for complementary laboratories. This latter assertion is crucial since we assume actors do not change their location for every partnership. The spatial distribution of economic and scientific resources could highly structure the geography of partnerships. All else being equal, large regions should have a higher probability to be involved in collaborations.

Finally, the explaining framework of the geography of partnerships we want to test can be sum up as followed : i) taking into account the capacity of actors to use a temporary proximity, we assume companies and laboratories do not change their location, ii) the geography of collaborations should thus depend on resources endowment of regions (size effects), iii) it should also depend on different forms of proximity: spatial proximity between regions, on the one hand, and scientific and sectoral proximity (specialization effects), on the other, iv) to test the heuristic power of the determinants proposed, we realized a gravity model.

2. Data and Method

The empirical work presented relies on a database relating to science-industry research contracts under the auspices of the *Convention Industrielle de Formation par la Recherche* (Cifre)[4]. This is the mechanism operating in France since 1981 whereby a subsidy is due to any business, working within the French law, employing a PhD student at the heart of research collaboration with a state laboratory. Cifre grants would appear to be a relevant

[3] These two forms of coordination proximity lead to the question of the functioning of collaborations: the existence of social relations between actors or mediation devices promots the exchange of knowledge between partners by allowing, in particular, a certain degree of trust between actors. Proximity of mediation and relational proximity also appear necessary when building collaborations, particularly when this involves linking partners. Using these two forms of proximity, Grossetti and Bès (2003) and Ferru (2010) showed the local nature of collaborations is sometimes explained less by the problem attached to the exchange of tacit knowledge than by the fact that those involved may mobilize their social relationships to find their partners.

[4] This database is provided by the National Association of Research and Technology (ANRT), which manages this kind of agreements in France.

indicator in the analysis of collaborations for innovation, even if they only refer to a portion of them, since they correspond to a subset of contracts concerning only those relationships between science and industry.

The database collects all Cifre contracts established over the period 1997-2006 for the whole of France, constituting a total of 7,972 contracts. Diverse information relating to these contracts is available including i) the industrial sector of the firm (subdivided into 18 sectors), ii) the type of scientific field of the laboratory (26 fields), iii) the location of the laboratory (by municipality, county and region), iv) the location of the plant (ditto). The analysis is confined to regions of mainland France (21 regions) and excludes Corsica.

To test the existence of size and proximity effects, a gravity model is relied on. This model, commonly used in the analysis of international trade has been applied in recent years to provide some explanations for the geography of collaborations. Different authors, including Ponds et al. (2007) and Maggioni and Uberti (2008) have indeed used this model to test the role of physical distance. Size-effects and proximity effects related to the geography of resources have rarely, if ever, been included to reflect the geography of collaboration by authors using gravity models. This can be explained, as said above, by the fact that the authors have remained hitherto focused on conditions for the exchange of tacit knowledge, as well as on the role of spatial distance. A gravity model is proposed here to relate both the existence and the intensity of interregional cooperation in the realms of size and the different forms of proximity/distance between regions. In other words, in this model, the dependent variable is either the probability or the intensity to collaborate and the explanatory variables are the economic and scientific sizes of the regions, on the one hand, and the various forms of proximity (spatial, scientific and sectoral) on the other.

The model we test here is a generalized tobit (second type of tobit regarding the classification of Amemiya (1985)). It assesses in a first equation the probability that potential partners collaborate and in a second equation the intensity (number of contracts) of such collaboration (when it comes true).

In a first step, we evaluate the first equation of the model, a probit specification, prior to test the complete model in a second step.

The probit, which defines the probability to collaborate, could be written as a latent variable model :

$$d^*_{r_1,r_2} = a_0 + a_1 \ln X_{r1} + a_2 \ln Y_{r2} - a_3 D_{r_1,r_2} + a_4 \,\mathrm{intra}_{r_1,r_2} + a_5 QI_{r_1,r_2} + a_6 QS_{r_1,r_2} + u_{1t} \quad , \qquad (1)$$

$$\begin{cases} d_{r_1,r_2} = 1 \text{ si } d^*_{r_1,r_2} > 0 \\ d_{r_1,r_2} = 0 \text{ si } d^*_{r_1,r_2} \le 0 \end{cases}$$

where $u_{1t} \sim N\left(0, \sigma_1^2\right)$.

The variable $d^*_{r_1,r_2}$ is latent (partially unobservable). We just know the sign of this variable, giving if there is a collaboration between the potential partners $(d_{r_1,r_2}=1)$ or not $(d_{r_1,r_2}=0)$. $d^*_{r_1,r_2}$ could be read as a measure of the tendency/disposition/enactment/aptitude of partners (regions ?) to collaborate. The more they are inclined to collaborate, the more the probability the collaboration comes true.

The parameters a_1 to a_6 are assumed to be positive. A negative value for one of these parameters would have no sens.

The variable $\ln X_{r_1}$ is the relative economic size of region r_1.

To measure the economic size of regions, three indicators are available: the fraction of GDP, the number of establishments and the number of private researchers were taken into account. These three measures are heavily correlated and seem to be placeholder in the model. Choice indicators having very few differences in the model, we choose arbitrarily the PIB to measure economic size. More precisely, the measure is the percentage of GDP (INSEE data), averaged over 1997 and 2006.

The variable $\ln Y_{r_2}$ is the relative scientific size of region r_2. The scientific size is measured by the percentage of public researchers of the region, averaged over 1997 and 2006.

$D_{r1,r2}$ is the geographical distance between r_1 and r_2. This distance could be measured by different means. We considered notably two types of physical distance: 1) a geographic distance using a matrix of contiguities (number of border crossings to get from one region to another) and 2) a distance in terms of the duration of the train journey between each of the regional capitals. The model works with anyone of these distances but the geographic one seems to give better explanation. The results we will present in the next paragraph integrate such measure of the distance.

A dummy $intra_{r1,r2}$ is incorporated to account for the specificity of intra-regional cooperation / to capture the effect of co-location of those involved within the same region,: it equals 1 when $r_1 = r_2$ and equals 0 when $r_1 \neq r_2$. It measures the specificity of such collaboration which would not be take into account with the physical distance (which equals 0 when $r_1 = r_2$), or thanks to the dissimilarity QI_{r_1,r_2} and QS_{r_1,r_2} (which also equals 0 when $r_1 = r_2$).

$QI^{\delta}_{r1,r2}$ and $QS^{\theta}_{r1,r2}$ correspond to the sectoral proximity and scientific proximity respectively. If two regions are really close in terms of productive activities/scientific domains, companies/laboratories have no reason to look for a partner outside the region, all else being equal. On the contrary, if two regions are really different in terms of productive activities/scientific domains, a laboratory/a company have to find a complementary resource (which does not exist in the region) outside the region. Consequently, we expect a negative correlation between the number of science-industry partnerships and the scientific/sectoral proximity. Regarding the indicators used to assess the sectoral/scientific proximity and the expected correlation, we apply these indicators on the inter-regional subset of partnerships (cf. infra).

To evaluate these proximities, we build a matrix based on the Aquino Index (1978). This latter can be defined as followed:

$$Q_{ij} = \frac{\sum_k |V_{ik} - V_{jk}|}{2}$$

with

$V_{ik} = \frac{X_{ik}}{X_i}$ and $V_{jk} = \frac{X_{jk}}{X_j}$,

The proportion of the k class (activity sectors, scientific domains, etc.) for the regions i and j. An Aquino Index close to 1 reveals a great difference of the regions features. On the contrary, an index close to 0 indicates an important proximity of the regions features.

Let us indicate in the equation (1) that certain variables have logarithmic measures. This caracterization of the model appears to be theoretically necessary: if the economic size of the region r_1 or the scientific size of the region r_2 is too low, no collaboration can exist regardless of the others variables value. For instance, considering an extreme case where the region r_1 is totally deserted, no production can exist in this region, the GDP is close to 0 leading endlessly to a negative logarythm. The parameter a_1 being positive, the value of the latent variable is also endlessly negative, showing a collaboration between such region and any other region is not possible. The same reasonning can be used for the scientific size. To evaluate scientific proximity, we use a matrix calculated with data in 1999 (DIRD by scientific domain) by Mangematin and al. (2003). To assess sectoral proximity, we use the value added by sector (Insee, Nes 14) in 1999.

Table 1 : Presentation of the data

Variables	Average	std deviation	min.	max.
dependent variable (number)				
Contracts	17,94	101,22	0	1888
Intra-regional contracts	198	416	15	1888
Interregional contracts	8,94	27,46	0	388
economic size (% France)				
GDP	5,36	6,42	0,98	28,89
Number of plants	0,06	0,11	1,16	22,31

number of private researchers	0,05	0,05	0,57	45,56
scientific size (% France)				
number of state researchers	5,71	8,86	0,47	38,26
spatial proximity				
geographical distance (in terms of the borders crossed)	2,3	1,24	0	6
time of the train journey(in number of hours)	4,38	2,07	0	9
sectoral proximity				
AQI	0,12	0,05	0.04	0,31
scientific proximity				
AQS	0,54	0,20	0,07	0,99

Note : N=21. [a]N=441

The table above shows some indicators of the statistical data used. The matrix of interregional collaborations (for the period 1997-2006) used is attached (see the appendix). Its analysis particularly underscores the importance of intra-regional agreements, which represent 53% of the total, with a minimum of 22% for Picardie and a maximum of 65% for the Midi-Pyrénées region. There is also a bias towards the Ile de France, where 46% of the contracting establishments and 30% of the laboratories are concentrated. The next highest are Rhone-Alpes (respectively 15% and 17%) and Midi-Pyrénées (5% and 8%).

3. Results using the gravity model

The following results[5] come from the probit model tested thanks to the maximum likelihood method:

$$\hat{d}^*_{r_1,r_2} = \underset{(0,72)}{-0{,}14} + \underset{(0,00)}{0{,}98}\ln X_{r1} + \underset{(0,00)}{0{,}75}\ln Y_{r2} - \underset{(0,00)}{0{,}30} D_{r_1,r_2} + \underset{(0,97)}{5{,}38}\,\mathrm{intra}_{r_1,r_2} + \underset{(0,57)}{0{,}21} QI_{r_1,r_2} - \underset{(0,65)}{1{,}08}\, QS_{r_1,r_2}$$

Regarding the Wald test for the simultaneous nullity of the parameters a_1 to a_6, we have to reject the hypothesis that no variable does not impact the endeogeneous variable (the critical probability of the test being negligible). The probit regression is thus significant. The Hosmer and Lemeshow tests confirm the quality adjustment of the model (p-value = 0,67).

The analysis of individual effects, resulting from the estimated equation, highlights that the variables $\ln X_{r_1}$, $\ln Y_{r_2}$ and D_{r_1,r_2} are heavily significant whereas the variables $Intra_{r1,r2}$, QI_{r_1,r_2} and QS_{r_1,r_2} have a really low significant effect. The probability for potential partners to collaborate with each other apparently depends on the economic and scientific sizes of the regions involved and on the distance that separates them. The significant parameters

[5] Numbers into brackets correspond to the *p-value* or to the critical probability of the test for the null hypothesis of the parameter (Wald test).

have the expected signs. The economic and scientific sizes have a positive effect on the probability to collaborate whereas the distance between two regions have a negative impact on it.

In addition to the fact that the estimated parameters are significant, we interestingly observed the economic relevance of these results as we can illustrate thanks to various examples. Removing the non significant variables of the model, the probability for a collaboration between the regions r_1 et r_2 equals:

$$P\left(d_{r_1,r_2}=1\right)=\Phi\left(0{,}98\ln X_{r_1}+0{,}75\ln Y_{r_2}-0{,}30D_{r_1,r_2}\right), \quad (2)$$

where $\Phi(.)$ corresponds to the normal distribution function. In the following table, the minimum, maximum and average values of the three exogenous variables in (2) are given.

Minimum, maximum and average values of the significant variables

	Mimimum value	Average value	Maximum value
$\ln X_{r_1}$	- 0,022	1,236	3,363
$\ln Y_{r_2}$	-0,764	0,906	3,644
D_{r_1,r_2}	0	2,405	6

To assess the effect of each variable on the probability to collaborate, we measure the three probabilities in (2) that correspond to the three values given in the table above (the values of the two other variables being considered as equal to the average value). The following table sums up such estimated probabilities.

Probabilies according to each variables

	Mimimum value	Average value	Maximum value
$\ln X_{r_1}$	0,474	0,878	0,999
$\ln Y_{r_2}$	0,466	0,878	0,999
D_{r_1,r_2}	0,970	0,878	0,536

The effects, for each variable, appear to be hight : changing from the minimum value to the maximum one leads the probability to collaborate being doubled or so. For the Limousin region (that has the lowest GDP), the probability to collaborate with a region which has an average scientific size and is located at an average distance, has a probability to collaborate two times lower than the Ile de France region (that has the highest GDP). Moreover, the Limousin region (that also features the lowest scientific size) has a probability to collaborate with a region characterized by an average economic size and located at an average distance two times lower than the Ile de France region (that has also the highest scientific size). Lastly, taking into account the spatial proximity, a collaboration within the region (D_{r_1,r_2} = 0) would have a probability two times higher to exist compared to collaborations between two distant regions like the Nord-Pas-de-Calais and the Aquitaine regions (D_{r_1,r_2} =6).

In addition, the table above underlines that only one disadvantageous variable is not enough to prevent the partners from collaborating. The lower probability, which equals to 0.466, is far away from insignificant. To have really low probabilities, regions have to cumulate disadvantageous concerning at least two exogenous variables simultaneously. The following table sums up the probabilities, based on the calculations of two disadvantageous situations and one average situation regarding exogenous variables. We thus are able to consider situations with low probabilities to collaborate.

Probabilities to collaborate with two disadvantageous and one average situations

	Probability
$\ln X_{r_1}$ and $\ln Y_{r_2}$ minimum	0,094
$\ln X_{r_1}$ minimum and D_{r_1,r_2} maximum	0,126
$\ln Y_{r_2}$ minimum and D_{r_1,r_2} maximum	0,122

Even if the model is statistically significant, its quality in terms of predictions is quite low. The sample counts 336 existing collaborations out of 441, 105 cases corresponding to an absence of collaboration. Let us assume that when the probability is higher than 0.5, the existence of collaboration is predicted and, on the contrary, when the probability is lower than 0.5, the absence of collaboration is predicted. Then, comparing the predicted values with the real ones, it appears to be easier to predict when collaboration comes true rather than when it does not: 92% of the existing collaborations are well predicted whereas only 34.3% of the absence of collaborations is. Consequently, the model does not explain two third of the non collaborations cases. The endogenous variables integrated in the model may not be the more relevant ones to explain the existence or the absence of collaboration. However, such variables appear to be more predictive to explain the intensity of existing collaborations. The probability and the intensity of collaborations could be assessed jointly thanks to a generalized tobit:

$$d^*_{r_1,r_2} = a_0 + a_1 \ln X_{r1} + a_2 \ln Y_{r2} - a_3 D_{r_1,r_2} + u_{1t} \quad ,$$

$$\begin{cases} d_{r_1,r_2} = 1 \text{ si } d^*_{r_1,r_2} > 0 \\ d_{r_1,r_2} = 0 \text{ si } d^*_{r_1,r_2} \leq 0 \end{cases},$$

$$\ln C^*_{r_1,r_2} = b_0 + b_1 \ln X_{r1} + b_2 \ln Y_{r2} - b_3 D_{r_1,r_2} + b_4 \,\mathrm{intra}_{r_1,r_2} + b_5 QI_{r_1,r_2} + b_6 QS_{r_1,r_2} + u_{2t},$$

$$\begin{cases} C_{r_1,r_2} = C^*_{r_1,r_2} & \text{si } d^*_{r_1,r_2} > 0 \\ C_{r_1,r_2} = 0 & \text{si } d^*_{r_1,r_2} \leq 0 \end{cases},$$

where ${}^t(u_{1t}\ u_{2t})$ follows a two-dimensional normal law.

The first part of the model defines the probit, without the non-significant variables. As said above, it measures the probability to collaborate. The second part of the model determines the intensity of the collaboration. By definition, the effective number of contracts C_{r_1,r_2}

equals to zero when there is no collaboration ($d^*_{r_1,r_2} \leq 0$). When there is a collaboration, C_{r_1,r_2} is defined by an underlying model that links the number of contracts to our explicative variables ($\ln X_{r_1}$, $\ln Y_{r_2}$, D_{r_1,r_2}, intra_{r_1,r_2}, QI_{r_1,r_2} and QS_{r_1,r_2}). As previously explained, we use a logarithmic specification for certain variables. Such a specification is used for the variable C_{r_1,r_2} in order to ensure that its estimation is positive.

The estimation of the generalized tobit (maximum likelihood method) gives the following results :

$$\hat{d}^*_{r_1,r_2} = \underset{(0,00)}{-0{,}94} + \underset{(0,00)}{1{,}73}\ln X_{r1} + \underset{(0,00)}{0{,}98}\ln Y_{r2} - \underset{(0,00)}{0{,}37}D_{r_1,r_2}$$

$$\begin{cases} d_{r_1,r_2} = 1 \ if \ d^*_{r_1,r_2} > 0 \\ d_{r_1,r_2} = 0 \ if \ d^*_{r_1,r_2} \leq 0 \end{cases},$$

$$\ln \hat{C}^*_{r_1,r_2} = \underset{(0,00)}{-0{,}72} + \underset{(0,00)}{1{,}32}\ln X_{r1} + \underset{(0,00)}{0{,}75}\ln Y_{r2} - \underset{(0,00)}{0{,}28}D_{r_1,r_2} + \underset{(0,00)}{2{,}70}\,\text{intra}_{r_1,r_2} + \underset{(0,91)}{0{,}001}QI_{r_1,r_2} - \underset{(0,99)}{0{,}001}QS_{r_1,r_2},$$

$$\begin{cases} C_{r_1,r_2} = C^*_{r_1,r_2} \quad if \ d^*_{r_1,r_2} > 0 \\ C_{r_1,r_2} = 0 \qquad if \ d^*_{r_1,r_2} \leq 0 \end{cases}.$$

The comparison between these results and the previous ones reveals that the generalized tobit estimation is rather sensitive, since the coefficients value change. Nevertheless, they remains significant and have the expected signs.

The analysis of the estimation of the number of contracts, when the collaboration comes true, reveals that the dissimilarity measures are not significant and whereas the variables $\ln X_{r_1}$, $\ln Y_{r_2}$ and D_{r_1,r_2} are. Contrary to the probit model, the variable $INTRA_{r1,r2}$ has a significant effect. Thus, the economic and scientific sizes of partners and the distance between regions have an impact on the number of collaborations. The intensity of collaborations increases when the collaboration exists within the region for a reason that is not linked to the distance (taking into account by the variable D_{r_1,r_2}) or to the dissimilarity measures (that are not significant).

Moreover, the quality of the model is rather good, since the residual variance is around one fourth out of the total variance , thus the R^2 equals 0.75. The significant variables give a good explanation of the collaboration intensity.

CONCLUSION

The aim of this paper was to complement the empirical work on the spatial dimension of collaboration for innovation, drawing on an original database to measure possible effects of various parameters. The empirical investigations conducted have shown that the geography of collaboration for innovation is essentially explained by the geography of resources and by distance effects. The influence of the sectoral and scientific proximities seems to be limited. Although this last determinant is generally indicated by a great part of the authors in the literature, it has been rarely tested with relational data on direct science-industry collaborations. Concerning the second determinant tested, the geography of resources, it is generally neglected even if its effects seem to be obvious.

This model could be enriched by addressing the role of other determinants involved earlier in the process of collaboration. Beyond the role of the geography of resources, the relationships between those involved could be set out formally. Such studies would be necessary to discriminate between these effects and measure more precisely the respective influence of such prior constraints, compared with the influence of others constraints involved in the actual collaborative process.

REFERENCES

Amemiya T., 1985, *Advanced econometrics*, Basil Blackwell, Oxford.

Audretsch D. and Feldman M., 1996, “Knowledge Spillovers and the Geography of Innovation and Production”, *American Economic Review*, 86 (3), p. 630-640.

Audretsch B. and Stephan P., 1996 Company-scientists locational links: the case of biotechnology, *American Economic Review*, Vol. 86, pp. 641-652.

Autant-Bernard C., 2000, *Géographie de l'innovation et externalités locales de connaissances. Une étude sur données françaises,* Thèse de Doctorat ès Sciences Economiques, Université Jean Monnet St-Etienne.

Becattini G. 1992 Le district marshallien: une notion socio-économique, In Benko G., Lipietz A. Ed., *Les régions qui gagnent - Districts et réseaux: les nouveaux paradigmes de la géographie économique*, Presses Universitaires de France, Paris, pp. 35-55.

Benko G., 1998, *La science régionale*, Presses Universitaires de France

Boschma R., 2005, « Proximity and innovation: a critical assessment », *Regional studies*, Vol. 39(1), pp.61-74.

Bouba-Olga O. et Grossetti M., 2008, « Socio-économie de proximité », *Revue d’Economie Régionale et Urbaine*, n°3.

Feldman M., 1994, “The geography of innovation”, Kluwer Academics Publishers, Dordrecht.

Ferru M., 2010, « Formation process and geography of science-industry partnerships : the case of the University of Poitiers », Industry and Innovation, Forthcoming.

Gay C., Picard F., 2007, « Individus, organisation, réseaux : triptyque d'une politique de création de connaissances », In Capron H. (Ed.), Politique régionale européenne : convergence et dynamique d'innovation, p. 215-243.

Gilly JP. et Torre A., 2000, “Dynamique de proximité”, Ed. L’Harmattan, Paris.

Giuliani E. and Arza V. (2009) What drives the formation of `valuable' university-industry linkages?: Insights from the wine industry. *Research Policy*, 38 (6), 906-921.

Giuri P., Mariani M., Brusoni S., Crespi G., Francoz D., Ganbardella A., Garcia-Fontes RG., Harhoff D., Hoisl K., Le Bas C., Luzzi A., Magazzini L., Nomaler O., Palomeras N., Patel P., Romanelli M.and Verspagen B., 2006, "Everything you always wanted to know about inventors (but never asked): evidence from the PatVal-EU survey", Discussion paper series, n°5752, Center for economic policy research.

Grossetti M. and Bès MP., 2003, « Dynamique de réseau et des cercles. Encastrements et découplages », *Revue d'Economie Industrielle*, numéro spécial sur la morphogenèse des réseaux, n°103, 2ème et 3ème trimestres, pp. 43-58.

Grossetti M., Barthe JF, Beslay C., 2006, «La dimension spatiale des processus de création d'entreprises innovantes». «colloque «Les systèmes d'innovation: multiplicité des échelles, diversité des espaces», Bordeaux, 16 et 17 novembre 2006.

Hagedoorn J. 2002, « Inter-firm R&D partnerships: an overview of major trends and patterns since 1960 », *Research Policy*, Elsevier, Vol. 314, p. 477-492, Mai

Jaffe A., 1989, "Real effects of academic research", *The American Economic Review*, Vol. 79, issue 5, p. 957-970.

Joly P.B. and Mangematin V. (1996) Profile of public laboratories, industrial partnerships and organisation of R&D, *Research Policy,* 25, 901-922

Kenney M. and Florida R., 1994, "The organisation and geography of japanese R&D: results from a survey of japanese electronics and biotechnology firms", *Research Policy*, n°23, p. 305-323.

Mangematin V., Massard N., Autant-Bernard C., Carrere M., Coronini R., Nesta L., Riou S., 2003, « Etude sur les profils scientifiques et techniques régionaux », Rapport pour le ministère de l'éducation nationale

Nooteboom, B., Wuyst, S., Colombo, M.G., et Duta, S., 2005, "Empirical test of optimal cognitive distance", Journal *of economic behaviour and organization*, Vol. 58, pp. 277-302.

Paci R.and Usai S., 2000, "Externalities, knowledge spillovers and the spatial Distribution of innovation", European Regional Science Association conference papers.

Perroux F., 1955, « Note sur la notion de pôle de croissance », Economie Appliquée, 1-2, pp. 307-320.

Porter*, M.* 1990, *The Competitive Advantage of Nations,* Macmillan, London.

Rallet A. et Torre A, 2005, « Proximity and localization », *Regional studies*, Vol. 391, pp.47-59.

Rallet A. and Torre A. (2007) *Quelle proximité pour innover ?.* L'Harmattan.

Saxenian A., 1994, *Regional advantage: culture and competition in Silicon Valley and Route 128*, Cambridge, Harvard University Press.