



**HAL**  
open science

# EXPLORING SPATIAL FEATURES OF SCIENCE-INDUSTRY PARTNERSHIPS: A STUDY ON FRENCH DATA

Marie Ferru, Olivier Bouba-Olga, Dominique Pepin

► **To cite this version:**

Marie Ferru, Olivier Bouba-Olga, Dominique Pepin. EXPLORING SPATIAL FEATURES OF SCIENCE-INDUSTRY PARTNERSHIPS: A STUDY ON FRENCH DATA. *Papers in Regional Science*, 2011, pp.10.1111/j.1435-5957.2011.00385.x. 10.1111/j.1435-5957.2011.00385.x . hal-00566109

**HAL Id: hal-00566109**

**<https://hal.science/hal-00566109>**

Submitted on 15 Feb 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# EXPLORING SPATIAL FEATURES OF SCIENCE-INDUSTRY PARTNERSHIPS: A STUDY ON FRENCH DATA

*Olivier Bouba-Olga, Marie Ferru, Dominique Pépin*

*Laboratoire CRIEF-TEIR EA 2249 – Université de Poitiers*

## **Abstract**

The aim of this paper is to provide explanatory elements of the geography of collaboration by testing various potential factors related to the partners' features (sector of activity, location, affiliation to a parent firm, ...) and to the pair of regions involved in the partnership (economic and scientific endowment of regions and proximities that separate them). Based on a database collecting around 15 000 Science-Industry agreements signed in France between 1981 and 2006, we first realized a probit model that focus on the local dimension of science-industry collaborations on an inter-individual level. We then test a gravity model (sample selection model) in order to explain the probability and the intensity of interregional collaborations.

**Keywords:** geography, partnerships, endowment, proximities

**Code JEL:** O31, R12, O39

## 1. Introduction

For twenty years now, there has been an increasing amount of literature on the subject of science-industry collaborations whose major part focuses on the geographical dimensions of these partnerships. In the 1990s, the very first works on the subject of geographical innovation highlighted the presence of local technological externalities (Jaffe, 1989; Audretsch and Feldman, 1996). Later on, the results obtained were questioned especially regarding the limits linked to the data gathered (co-location data rather than collaboration data). Then, the very first researchers on the subject encountered « difficulties in finding localised data on innovation and knowledge processes and in addressing methodological issues due to the spatial aggregation of the available data » (Autant-Bernard et al., 2007a). This led to a second series of works involving genuine relational data that revealed that these so-called externalities do not simply diffuse in the air because « the geographical dimension is in complex ways related to other mechanisms that are still barely identified and that take place at different geographical levels » (cf. Autant-Bernard et al., 2007a).

Our paper is in line with these latest studies and aims to go a step further through the analysis of these different mechanisms thanks to a new relational database concerning the science-industry collaborations: it gathers various information about the Cifre contracts developed in France between 1981 and 2006, averaging 15 000 contracts. Its novelty lies in the fact that it provides genuine long-term relational data.

Thanks to this database, we propose to investigate the spatial patterns of science-industry partnerships through a twofold analysis applied to two complementary levels:

1) First, we focus on the local or non local dimension of science-industry collaborations. This indicator of the geography of innovation is computed at an inter-individual level. This first analysis rests on a binary probit model that tests the influence of the partners' various features (sector of activity, scientific field of study, the partners' location and their affiliation to a parent firm) and the influential role played by the period during which the contract was carried out.

2) Then, we propose a complementary analysis by concentrating on second indicators of this geography: the probability and the intensity of collaborations between pairs of regions. To be more exact, a sample selection model over the end of the period studied (1997-2006) is tested. It allows us to measure the effect of inter-regional proximities (spatial and non-spatial) and of regional endowments (industrial and scientific resources) on both the probability and the intensity of collaborations.

Our paper is organised as follows: in the second section we will review the existing literature to highlight the existence of various potential determinants in the geography of collaborations. In the third section, we will present the data and the first results of descriptive statistics. The main results obtained from the econometrics models will be exposed in sections four and five and we will finally present our conclusions and discuss the results obtained in the sixth section.

## 2. Literature background

From an empirical and theoretical point of view, the university seems to be an essential partner in innovation by providing new sources of knowledge (Gibbons et al., 1994; Feldman, 1994; Saxenian, 1994 ; Anselin et al., 1997). The CIS (Community Innovation Survey) inquiry reveals that more than 20

percent of the French organizations collaborate with a university laboratory or with a public research organization to complete their innovation project, a number that has kept increasing (Godin and Gingras, 1999 ; Wagner and Leydesdorff, 2005 ; Hagedoorn and Roijakkers, 2006). However in the early 1980s, the conception of collaboration to innovation was hard to imagine for scholars considering the specificity of that activity which was both difficult to initiate technically speaking and whose result was doubtful (Hagedoorn, 2002).

For more than 15 years now, the growing importance of these collaborations has led to numerous studies (Gibbons et *al.*, 1994 ; Joly and Mangematin, 1996 ; Etzkowitz et *al.*, 1998 ; Carayol, 2003 ; *etc.*). Some try to understand how they are structured, their objective; others test the link between the different features of the partners (their size, sector of activity, technological intensity and so on) as well as their propensity to collaborate (see Bergman (2010) for an overview of the existing literature on the subject of science-industry collaborations). Given the strong polarity of innovation activities (Lung, 1997 ; Puga, 1999 ; Lallement et *al.*, 2007, *etc.*), other writers have focused on the essential question of the spatial dimension of these collaborations. A group of authors relied on the theoretical models of the New Geographical Economy carried out several empirical studies (Jaffe, 1989 ; Feldman, 1994 ; Feldman and Florida, 1994 ; Audretsch and Feldman, 1996) to evidence the existence of local technological externalities. These results have largely contributed to the emergence of the idea that collaborations require geographical proximity though recently several authors have questioned these results: the empirical inquiries at the origin of that thesis were essentially based on a limited spatial scale, regarding local proximity as the most favourable place for innovation through collaborations. A new series of empirical studies showed that the local factor does exist but is far from being the only one to take into account (See Ferru (2009) for a recapitulative table).

On a theoretical level, several works (Rallet and Torre, 2005 ; Boschma, 2005 ; Boschma and Frenken, 2009 ; Bouba-Olga and Grossetti, 2008) justify the diversity of spatial scales by the plural aspect of proximity, then showing that “the proper impact of the geographical dimension must be more precisely assessed in relation to other types of proximity” (Autant-Bernard et al., 2007a, p.343), such as social technological, organizational or institutional proximity. On the empirical level, the authors today rely on genuine relational database to test the role played by the different theoretical determinants and the effects of proximities. Table 1 provides a non exhaustive overview of these empirical studies.

**Table 1 : Empirical studies about the determinant factors of the geography of science-industry partnerships**

Authors/Title	Database	Main results
Grossetti & Nguyen, 2001, « La structure spatiale des relations science-industrie en France: l'exemple des contrats entre les entreprises et les laboratoires du CNRS »	Data about 13827 French research contracts (CNRS) from 1986 to 1998	Statistical analyses reveal <b>regional effects</b> . Collaborations with firms located in <b>Ile de France region</b> are dominating
Singh, 2005, "Collaborative networks as determinants of knowledge diffusion patterns"	More than 2 540 000 patents citations recorded by UPSTO from 1986 to 1995	Choice based sampling confirms the <b>positive influence of spatial proximity</b> but this effect decreases when taking into account former ties
Levy & al., 2009, "A study of science-industry collaborative patterns in a large european university"	Data about 1000 firms having collaborated with ULP from 1990 to 2002	A multinomial logit model reveals the <b>role of proximity is even more important in the case of intense collaborations</b>
Ponds & al., 2007, "The geographical and institutional proximity of research collaborations"	Data about more than 240 000 co-publications in Netherland from 1988 to 2004	A censored regression and a gravity model reveal <b>geographical proximity is more relevant for "mixed" collaborations</b> and is a way to overcome institutional proximity
Maggioni & al., 2007, "Space versus networks in the geography of innovation"	Co-patent between regions from Germany, Italy, Spain, France and the UK	A gravity model and a spatial error model underline the importance of public R&D expenditure, previous collaborations and <b>proximities (spatial and technological)</b>
Autant-Bernard & al., 2007, « Social distance versus spatial distance in R&D cooperation »	Individual data about 139 european firms and 75 french firms in micro and nano-technology	A binary logit model highlights the significant role of R&D potential and <b>social distance and the insignificant role of spatial distance</b>
Abramovsky & Simpson, 2008, "Geographical proximity and firm-university innovation linkages: evidence from Great Britain"	Data from the UK office for national statistics about firms with intra-mura R&D (2000-2003)	Pharmaceutical firms locate R&D close to frontier chemistry research departments and are <b>more likely to engage with universities</b>
Ferru, 2010, "Formation Process and Geography of Science-Industry Partnerships: the case of the University of Poitiers"	Individual data about more than 100 research contracts between laboratories of Poitiers and firms	A qualitative analysis reveals that the geography of collaborations is structured by the specificity of the resources needed and the <b>uneven spatial distribution of these resources</b> . Vectors of connection with partners are also decisive
D'Este & Iammarino, 2010, "The spatial profile of university-business research partnerships collaborations from 1999 to 2003"	4 525 joint research partnerships in UK (from EPSRC)	Negative binomial and OLS estimates reveal that research quality and <b>geographical proximity</b> are positively associated with the frequency of partnerships but with differences between scientific domains
Laursen & al., 2010, "Exploring the effect of geographical proximity and university quality on University-industry collaboration in the United Kingdom"	26 172 OCDE and CIS data from 2002 to 2004	logistic regressions reveal that collaborations are positively influenced by <b>geographical proximity</b> and quality of university but this latter is a preferential factor for firms compared to the former
Hoekman & al., 2010, "Research collaboration at a distance : changing spatial patterns of scientific collaboration within Europe"	47 000 data on co-publications in 33 European countries from 2000 to 2007	A gravity model underlines <b>changing effects of physical distance and territorial borders</b>

Our paper is in line with these works. Considering the data gathered, we especially test the existence of the following effects:

#### 1) Regional and size effects

Regions are heterogeneous regarding size, specialization, reputation and scientific quality which leads us to assume that regional effects do exist (some regions collaborate more than others on the local level). Grossetti and Nguyen (2001) have shown the existence of such effects in France (the importance of local collaborations differ depending on the regions the partners belong to) and the hypertrophy of the region Ile-de-France regarding the spatial distribution of the contracts between CNRS and industries. Hoekman et al. (2010) have noticed the over representation of regional capitals in the spatial distribution of collaborations for innovation on the European level. They have emphasized the existence of regional differences and suggested that “researchers based in (these) city-agglomerations are attractive partners, possibly reflecting their access to dense local or regional research networks”.

These regional effects would by and large result from more or less numerous opportunities to find a potential local partner in accordance with the economic and scientific resources allocated to the region the partner belongs to. As shown by Ferru (2010), the decision to collaborate with one partner rather than another « depends on the resources needed that are located in a particular place. This underlines the relevance of the geography of resources which generally favours territories concentrating numerous resources like capital regions and major cities. On the contrary, regions with only few resources are disadvantaged (...) and must necessarily develop collaborations outside the region: they have fewer opportunities to find a partner locally since their resources are limited in their total number and specialized in a restricted number of industrial sectors or scientific fields ». That assumption can be sustained through a qualitative analysis of relational data which show that the inequalities in the economic and scientific resources to be found on the French territory accounts for the over representation in the number of research contracts between the University of Poitiers and organizations located in Ile-de-France - a region counting most of the R&D centers - and a low representation of local contracts considering that Poitou-Charentes count very few R&D centers. That is why, considering these different results, we not only expect the amount of local collaborations versus non-local ones to be different depending on the regions (with positive or negative regional effects), but we also expect a size effect (depending on the economic and scientific resources allocated to regions).

#### 2) Physical proximity effects and border effects

The role played by the physical proximity effects has widely been documented to explain the geography of collaborations: though it was commonly accepted that proximity favors the development of collaborations (through a reduction of research and coordination costs), some scholars consider that today its scope is limited (due to the development of the ICT and means of transportation) and even assert that it is “the end to the tyranny of distance” (Castells, 1996; Cairncross, 1997). This trend has been confirmed by Frenken et al. (2009) in a survey revealing an overall increase in the number of long-distance partnerships. However, Maggioni et al. (2007), Abramosky and Simpson (2008) as well as Laursen et al. (2010) have all noticed the existence of physical proximity effects on the development of science-industry collaborations. Ponds et al. (2007) have shown that these effects have a significant impact in the case of collaborations involving

partners from different fields of activity (science versus industry) though it is not the case for partnerships between universities or between industries. Levy *et al.* (2009) have also shown that the more tighter the collaborations, the more significant the role played by proximity. Hoekman *et al.* (2010) have evidenced the continuity of these proximity effects from 2000 to 2007 thanks to the introduction of a dynamic perspective whereas Grossetti and Nguyen (2001) showed an increase in local collaborations from 1986 to 1998. Finally, for D'Este and Iammarino (2010), spatial proximity has a secondary effect compared to other factors. In the same line, Singh (2005) has noted that the impact of these effects gets weaker when taking into account the anteriority of collaborations. That trend had previously been evidenced by Almeida and Kogut (1998) and Grossetti and Bes (2003) who showed that the effects of physical proximity only result from previous social relations between local partners, which Autant-Bernard *et al.*'s works also confirmed (2007b).

Some writers also mentioned the potential role that borders can play. Hoekman *et al.* (2010) noted that « a systematic comparison between the effect of (spatial) distance and territorial borders is required to analyse the changing spatial patterns in research collaborations ». Indeed, beyond the effects linked to physical distance separating the partners and considering the existence of regional systems of innovation (Cooke, 1994), we can reasonably expect a concentration of collaborations within regions and the partners' difficulties to cross the borders. Hoekman *et al.* (2010) confirmed the significant influence of the border effects but also showed they gradually decreased with time.

### 3) Sectoral effects and sectoral proximity

A great number of writers insist on the necessity to take into account the scientific and technological profile of regions, consequently implying the existence of effects linked to the sectoral specialization (whether on the industrial or scientific level) of the partners. It seems reasonable to think that the actors, depending on their sectoral specialization, show narrow or wide possibilities to integrate knowledge from a distance, these possibilities varying with the technological intensity of the sectors of activity (Cohen and Levinthal (1990)). Ponds *et al.* (2007) have assessed the role played by proximity for each scientific field and revealed that it ostensibly changed between the scientific fields linked to biology and those linked to physics. Hoekman *et al.* (2010) evidenced the positive effect of sectoral proximity: "the larger the differences in research portfolios, the less co-publication activity occurs". Maggioni *et al.* (2007) also drew the same conclusion for inter-regional collaborations showing that « co-patenting activity of two regions is positively correlated to the degree of technological similarity of their innovation systems ». According to these works, the sectoral proximity would significantly facilitate inter-individual as well as inter-regional collaborations. However, if in the first study, the sectoral proximity was assessed for inter-university collaborations and referred to the similarity existing between the partners' scientific profiles, in the second study, the sectoral proximity was assessed in the case of mixed collaborations and referred to the existence of complementary in the scientific and technological profiles of the partners. In this latter analysis, the authors refer to what Frenken *et al.* (2007) and Boschma and Iammarino (2009) called "related variety"<sup>1</sup>.

---

<sup>1</sup> « Knowledge will spill over effectively between sectors when complementarities exist among sectors in terms of shared competences. Such competencies are captured by the notion of related variety » (Boschma and Iammarino, 2009).

The data we have do not permit to assess these complementarity effects, they only permit to assess the effects correlated to the proximity of the profiles between industries on one side and the profiles between science on the other. However, the causality relation between the proximity of these profiles and the geography of collaborations remains ambiguous: one can reasonably think that if two regions have quite similar industrial profiles, the laboratories in the first region have no interest in signing contracts with firms located in the second region since they are geographically close to the partners matching the profile required. By signing local contracts, firms would make savings on their transaction costs and/or could benefit from face-to-face contacts facilitating the exchange of tacit knowledge. But we can also think that these laboratories may have developed useful skills through local contacts and they could use them to develop contracts with similar firms operating in other regions. The proximity of the industrial profiles between the two regions would positively lead to potential collaborations. The same reasoning can be applied to the scientific profiles of regions.

### 3. Data, methodology and first descriptive statistics

#### *Data and global methodology*

The empirical work presented relies on a database about science-industry research contracts under the auspices of the Convention Industrielle de Formation par la Recherche (Cifre)<sup>2</sup>. This is the mechanism operating in France since 1981 whereby a state subsidy is due to any firm, working within the French law, to hire a PhD student in order to develop research in collaboration with a public laboratory. Such diadic relations between firms and laboratories appear to be a relevant 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 1981-2006 for the whole of France, constituting a total of 14 669 contracts. Diverse information relating to these contracts is available (see table 2). Instead of choosing a geographical breakdown by commune, we preferred to focus on a regional breakdown (21 regions) for the location of laboratories and institutions. It is on the regional scale that we possess essential complementary data (especially regarding economic and scientific resources of regions).

**Table 2: Presentation of the database used**

Variables	N	Features
YEAR	14669	1981 – 2006
FIELD OF STUDY (DOMRECH)	11809	Main field of study. ANRT terminology - All scientific domains
GROUP	13213	Affiliation of the firm to a parent firm or independant firm
SECTOR OF ACTIVITY	13613	Sector of activity (ANRT terminology)- All activities
ADRETAB_CITY	14637	Location of the firm
ADRETAB_COUNTRY	14668	Country of the firm
ADRELAB_CITY	14301	Location of the laboratory
ADRLAB_COUNTRY	14584	Country of the laboratory

<sup>2</sup> This database is provided by the National Association of Research and Technology (ANRT), which manages this kind of agreements in France.



This database provides a whole range of information; however we lack information concerning some variables, which prevents us from using them for our analysis. The classification used for the sector variable does not refer to any usual classification, which creates a constraint for the control of sectoral bias. Moreover, the lack of data for some contracts limits our sample of study: the probit is based on a sample of 10 263 contracts. Considering these different constraints and our subject of study, we focused on a subset of variables: i) the industrial sector of the firm (subdivided into 16 sectors), ii) the type of scientific field of the laboratory (21 fields), iii) the location of the laboratory (by NUTS 2 classification), iv) the location of the plant (ditto), iv) the year of the realized contract. The analysis is confined to NUTS 2 regions of mainland France, the number of contracts involving a foreign partner being really low (105 contracts only).

We will first of all concentrate on the local vs. non local dimension of the contracts by testing a binary probit model over the 1981 and 2006 period. This model tests the influence of the partners' various features (sector of activity, scientific field of study, the partners' location and their affiliation to a parent firm) and the influential role played by the period during which the contract was carried out.

The data available and the results obtained during that first step (weak time effects and strong regional effects), led us to test a sample selection model over the end of the period studied (1997-2006). It allows us to measure the effect of inter-regional proximities (spatial and non-spatial) and of regional endowments (industrial and scientific resources) on both the probability and the intensity of collaborations. In this perspective, we collected additional data thoroughly presented in section 5. Unfortunately, given the available data, we cannot test the model all over the period. In addition, it becomes impossible to carry out tests by sector of activity although they are often relevant (see Hoekman *et al.*, 2010).

This two-step methodology and the econometric strategy used are sum up in the table 3.

**Table 3 : The econometric strategy**

		<b>Probit model</b>	<b>Sample selection model</b>
<b>Explained variable</b>		Local vs. Non local relationship (10 262 obs.)	Number of contract between regions (441 obs.)
<b>Explanatory variables</b>	Regional effects	Yes	Yes
	Specialisation effects	Yes	No
	Dynamic effects	Yes	No
	Group affiliation	Yes	No
	Regional ressources endowment	No	Yes
	Proximity effects	No	Yes

#### *First descriptive statistics*

The firms signing CIFRE contracts with university laboratories essentially belong to groups (more than 41%) and independent firms employing less than 500 people (more than 35%). These firms operate in different sectors of activity and the four main sectors (electronics and electric equipment, services, para-chemistry and mechanical construction) account for more than 55% of the total number of contracts, a trend which is relatively constant. As for laboratories, a few sectors dominate in the collaborations between science and industry: computer science, physics, humanities & social

sciences and chemistry laboratories represent around 40% of the total number of CIFRE contracts. Once again this trend keeps constant through time save a slight drop in the contracts with the main scientific fields of study. It results that the actors in Ile de France, Rhone Alpes, Midi-Pyrénées and Provence Alpes Côte d'Azur regions are the most active in terms of science-industry collaborations with more than 65% of the total number of contracts. Once again this trend is quite stable over the period studied.

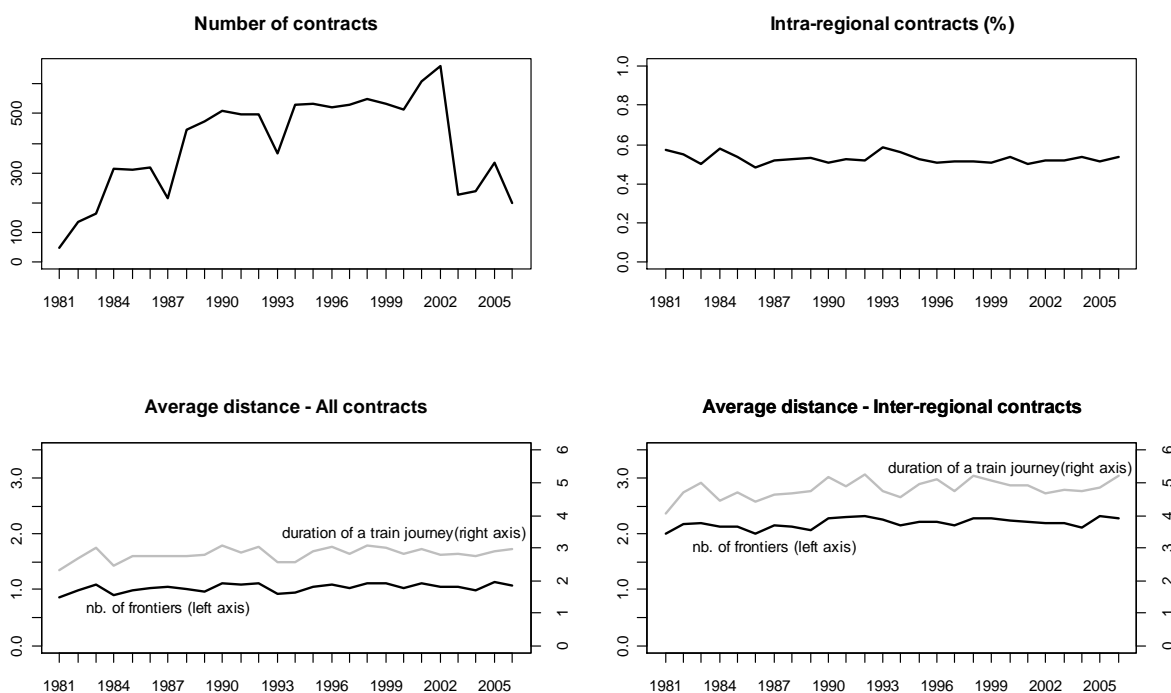
The data provided by the CIFRE database make it possible to assess the multi-scale patterns of collaborations dedicated to innovation, the contracts being developed with both local and non local partners. However, intra-regional collaborations still prevail: for the period and all the regions, intra-regional contracts represent 52.4%. Concerning inter-regional contracts, the Ile-de-France region appears to be involved in a great part (41%) of the total number of Cifre contracts.

The largest regions keep a strong importance, especially Ile de France, despite the decentralization policies started in the 1960s. The laboratories and above all the organizations located in Ile de France are privileged actors of the CIFRE contracts. One of the main reasons explaining that situation lies in the inequalities of economic and scientific resources allocated to regions and their strong concentration in the capital region: on average, Ile-de-France accounts for 29% of the GDP over the last decade and counts 38% of the researchers in the public sector contrary to the Limousin region with the lowest allocation of economic (1% of the GDP) and scientific (0.4% of the total number of researchers) resources. We thus observe a spatial structure based on 2 main levels: the national level characterized by a hypertrophy of the capital region and the regional level.

Despite the evolution of the number of CIFRE contracts (following a significant upward trend from 1981 to 2002), the number of inter-regional contracts has remained particularly stable over the period studied (See figure 1). The results obtained do not sustain the idea of an agglomeration process, as expected by Boschma and Frenken (2009) (« the emerging innovation network is most likely to cluster spatially »), neither do they show a dispersion process, expected by Menzel (2008) for example. These results may be explained by the reproduction of former relations established with former partners but also by the impact of the geography of resources provided that this geography of resources is also characterized by a strong inertia.

In order to give an opinion on the evolution of physical distances between partners, we have used two indicators: the number of borders to cross or the time it takes to go by train from one regional capital to another. It then becomes possible to calculate the average distance separating the partners over the period of study. We then noted that this average distance was apparently stable, whatever the indicator chosen. To confirm that observation, we once again tested the existence of a trend. It results that the trend is relatively increasing for all the average distances but the effect remains moderate, not to say very low: only the value linked to the average distance covered by train and the time it takes to cover it for all the contracts is relevant at a threshold of 5%.

**Figure 1 : Evolution of the CIFRE contracts (total number, regional contracts) and distances (total number of contracts and interregional contracts)**



These very first empirical elements reveal a crystallization trend of the geography of collaborations: inertia prevails over agglomeration or dispersion.

#### 4. The local dimension of science-industry collaborations: inertia and regional effects

##### *Model*

In order to assess the impact of the structural determinants on the spatial dimension of collaboration, we have tested a probit model,  $LOC_k^*$  being the latent variable and  $LOC_k$  the associated binary variable which takes on the value of 1 if the  $k^{ième}$  contract is signed between two partners in the same region, if it is not the case, it takes on the value of 0.

$$LOC_k = \begin{cases} 1 & \text{if } LOC_k^* \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

With  $LOC_k^* = \alpha + Z_k\beta + \mu_k$

$Z_k$  represents the vector of the explicative variables and  $\beta$  the vector of the associated parameters.

The explicative variables correspond to the dates of the contracts divided into five periods having the same amplitude (*period*), the affiliation to a group (*group*), the sector of activity of the firm (*industry*), the field of study of the laboratory (*scientific\_domain*) and the laboratory's location (*laboratory\_location*), the firm's location (*firm\_location*). For the regions whether dealing with laboratories or firms, the referential modality is the Aquitaine region. As for the sectors of activity and the scientific fields of study, the referential modalities are respectively energy and computer science.

To measure the quality of our model, we will measure (in addition to classical tests such as AIC criteria, Likelihood ratio, etc.) a pseudo-R<sup>2</sup> that will compare our model's forecasts to that of a null model. Concerning the model's forecasts, we assume when the probability is superior to 0.5 that the existence of the collaboration can be forecast and vice versa for a probability inferior to 0.5. The comparison between the forecasts and the values really observed allows building up a confusion matrix and calculating the number of "true positive" and "true negative". However, the analysis of these numbers cannot be carried out *ex-nihilo*; it is necessary to compare them to the values obtained from a trivial model made up of a single constant.

In the case of the trivial model, the rule is to attribute the modality of the most frequent explicative variable to all individuals; it then becomes possible to define an Adjusted Count R<sup>2</sup> defined as follows:

$$R^2_{AC} = \frac{\#correct - \max_k(n_k)}{n - \max_k(n_k)}$$

$n$  corresponds to the total number of observations,  $\#correct$  to the number of correct predictions obtained with our model and  $\max_k(n_k)$  the most frequent explicative variable. If the model tested does not do any better than the trivial model, we obtain a pseudo-R<sup>2</sup> equals to 0; if we have a perfect model, able to classify all the individuals correctly, we then have a R<sup>2</sup> equals to 1.

### Results

In order to assess the existence and the impact of the different effects tested, we first of all led a selection process of the variables by optimization according to Akaike's AIC and Schwartz's BIC criterion which over penalizes the addition of new variables as soon as the population number increases. Through the Akaike criterion all the explicative variables have been retained in the following order: *laboratory\_location*, *firm\_location*, *industry*, *scientific\_domain*, *group* and *period*. With the BIC criterion, the last two variables have not been retained. So, the regional effects seem to prevail, followed by sectoral specialization effects (industrial or scientific). The rejection of the *period* variable confirms the inertia of the geography of collaborations.

The detailed results obtained with the model optimized according to the AIC criterion feature in Table 4.

First of all, the Likelihood Ratio test allows us to reject the null hypothesis of all the coefficients. From the confusion matrix, we have also calculated the rate of "true positive" equal to 76% and the rate of "true negative" equal to 79%. The comparison between the predictions resulting from our model and those of the null model leads to a Pseudo-R<sup>2</sup> of 0.530, which can be considered as a satisfactory result.

**Table 4 : The influence of partners characteristics on the geography of collaboration (probit model)**

	INTRA-REGIONAL/INTER-REGIONAL COLLABORATION				
	ESTIMATE	STD. ERROR	Z VALUE	Pr(> z )	SIGNIFICANCE
(Intercept)	0.235	0.116	2.028	0.043	*
		<b>Period</b>			
1981-1985	Ref.				
1986-1990	-0.082	0.054	-1.524	0.127	
1991-1995	-0.069	0.053	-1.313	0.189	
1995-2000	-0.168	0.053	-3.196	0.001	**
2000-2006	-0.125	0.054	-2.311	0.021	*
		<b>Group</b>			
Group	-0.149	0.032	-4.588	0.000	***
		<b>Industry</b>			
Agriculture	-0.175	0.106	-1.641	0.101	
Energy	Ref.				
Metallurgy	-0.234	0.084	-2.788	0.005	**
Minerals Production	-0.082	0.094	-0.878	0.380	
Chemistry	-0.167	0.096	-1.747	0.081	.
Para-chemistry	0.118	0.080	1.479	0.139	
Mechanical construction	-0.024	0.070	-0.343	0.732	
Electric/Electronic Material	0.121	0.064	1.881	0.060	.
Armament	0.129	0.077	1.664	0.096	.
Manufacture of food products	-0.228	0.095	-2.412	0.016	*
Various Industries	0.056	0.115	0.488	0.626	
Paper/Plastic	-0.186	0.124	-1.496	0.135	
Construction	-0.025	0.113	-0.222	0.824	
Transport/Telecommunication	0.147	0.092	1.596	0.110	
Services	0.272	0.066	4.147	0.000	***
Others	0.227	0.161	1.408	0.159	
		<b>Scientific Domain</b>			
Mathematical	0.037	0.066	0.567	0.571	
Computer Science	Ref.				
Electronics	-0.021	0.061	-0.339	0.735	
Instrumentation	-0.261	0.103	-2.531	0.011	*
Automation	-0.001	0.083	-0.017	0.986	
Engineering	-0.336	0.095	-3.538	0.000	***
Physics	-0.111	0.055	-2.025	0.043	*
Metallurgy	-0.312	0.085	-3.666	0.000	***
Mechanical	-0.187	0.082	-2.291	0.022	*
Fluid mechanics	-0.226	0.086	-2.621	0.009	**
Energy	-0.626	0.122	-5.136	0.000	***
Chemistry	-0.221	0.060	-3.697	0.000	***
Environment	-0.147	0.095	-1.557	0.119	
Biomedical engineering.	-0.413	0.113	-3.661	0.000	***
Agrobusiness	-0.180	0.092	-1.949	0.051	.
Pharmaceutical	-0.154	0.100	-1.539	0.124	
Civil Engineering	-0.135	0.138	-0.976	0.329	
Humanities & Social Sciences	0.205	0.058	3.500	0.000	***
Biotechnology	-0.219	0.090	-2.426	0.015	*
Agricultural	0.118	0.144	0.815	0.415	
Others	0.008	0.090	0.088	0.930	

<b>Laboratory location</b>					
Alsace	-0.326	0.120	-2.709	0.007	**
Aquitaine	Ref.				
Auvergne	-0.393	0.148	-2.655	0.008	**
Basse-Normandie	0.558	0.194	2.875	0.004	**
Bourgogne	-0.184	0.152	-1.214	0.225	
Bretagne	-0.152	0.116	-1.307	0.191	
Centre	0.406	0.147	2.767	0.006	**
Champagne-Ardenne	0.247	0.203	1.218	0.223	
Franche-Comté	0.046	0.152	0.301	0.764	
Haute-Normandie	0.075	0.152	0.496	0.620	
Ile-de-France	1.116	0.088	12.688	< 2e-16	***
Languedoc-Roussillon	-0.803	0.117	-6.870	0.000	***
Limousin	-0.549	0.168	-3.260	0.001	**
Lorraine	-0.259	0.110	-2.369	0.018	*
Midi-Pyrénées	-0.138	0.098	-1.410	0.159	
Nord-Pas-de-Calais	0.043	0.110	0.389	0.697	
Pays De La Loire	-0.138	0.120	-1.151	0.250	
Picardie	0.031	0.133	0.234	0.815	
Poitou-Charentes	-0.328	0.146	-2.240	0.025	*
Provence-Alpes-Côte D'azur	0.024	0.101	0.242	0.808	
Rhône-Alpes	0.015	0.091	0.163	0.871	
<b>Firm location</b>					
Alsace	0.287	0.141	2.042	0.041	*
Aquitaine	Ref.				
Auvergne	-0.359	0.170	-2.110	0.035	*
Basse-Normandie	-0.384	0.182	-2.114	0.034	*
Bourgogne	-0.788	0.156	-5.053	0.000	***
Bretagne	0.147	0.135	1.090	0.276	
Centre	-0.889	0.141	-6.309	0.000	***
Champagne-Ardenne	-1.094	0.190	-5.747	0.000	***
Franche-Comté	-0.431	0.156	-2.764	0.006	**
Haute-Normandie	-0.982	0.158	-6.225	0.000	***
Ile-de-France	-0.540	0.093	-5.776	0.000	***
Languedoc-Roussillon	0.636	0.146	4.365	0.000	***
Limousin	0.358	0.200	1.792	0.073	.
Lorraine	0.323	0.127	2.537	0.011	*
Midi-Pyrénées	0.512	0.109	4.672	0.000	***
Nord-Pas-de-Calais	0.005	0.121	0.038	0.970	
Pays De La Loire	-0.250	0.135	-1.853	0.064	.
Picardie	-0.657	0.140	-4.699	0.000	***
Poitou-Charentes	-0.119	0.168	-0.709	0.478	
Provence-Alpes-Côte D'azur	-0.012	0.110	-0.108	0.914	
Rhône-Alpes	0.145	0.100	1.450	0.147	
Chi-2 : 2544.13	True positive : 76%	AIC : 11820	df : 80		
Pseudo-R <sup>2</sup> : 0.51	True negative : 79%	p(>Chi-2) : 0.000			

Notes : \*P<0.1, \*\*P<0.05, \*\*\*P<0.01

The referential modalities are given in the table (modalité « Ref. »)

Several results deserve to be emphasized:

- i) Regarding the temporal dimension, the last two periods negatively influence the probability of local collaborations compared to the first period but the effect is quite moderate. An additional test on the equality of the coefficients of the different modalities of the *period* variable confirms the weakness of that influence since the hypothesis of the equality of the coefficients cannot be rejected (p-value = 0.1464). The periods of time have consequently a weak influence over local collaborations; this seems to confirm the relative stability of the geography of collaborations that was previously observed and infirm the spatial agglomeration trend that some observed (Grossetti and Nguyen, 2001).
- ii) The affiliation to a group negatively impacts the probability of local collaborations. That influence can result from the fact that the affiliation to a group offers opportunities of collaborations outside the region where the firm involved is located since this firm can benefit from the relation network of the whole set of entities controlled by the parent firm.
- iii) As for the sectoral specialization effects (*industry, scientific\_domain*) they are stronger on the scientific field of study (12 out of 20 fields of study have significant coefficients) than on the industry (6 sectors out of 15). The sector of services on the one hand and the social and human sciences fields on the other hand are the only ones to positively impact the probability of local collaborations. No significant differences have been noticed between biology and physics contrary to what was observed by Ponds et al. (2007).
- iv) The regional effects (*laboratory\_location, firm\_location*) are quite pronounced for laboratories as well as for firms. It is worth noting the very significant positive coefficient of the laboratories located in Ile de France, which confirms the “capital effect” noted by Hoekman et al. (2010) and Grossetti and Nguyen (2001) and the very significant negative values obtained for the firms located in the neighbouring regions of Ile de France (Haute-Normandie, Champagne-Ardennes, Bourgogne, Picardie, Centre).

This last result seems particularly interesting; indeed, two complementary explanations can account for that situation: either the neighbouring regions benefit from a proximity effect with Ile de France or they are overshadowed by the region Ile de France. In order to check the relevance of these explanations we have first tried to determine if the weak localism of the neighboring regions could be explained by a greater number of relations with Ile de France since the administrative division of regions can conceal the effects of geographical proximity. The data available confirm that assumption: though the average number of contracts between Ile de France and the province represents 12%, it rises to 23% for Haute-Normandie, 22% for Champagnes-Ardennes and the Centre region, 19% for Picardie and 16% for Bourgogne; Auvergne is the only region with a higher rate of 25%. The weak localism of these regions is also due to a shadow effect (Brouillat and Lung, 2010) by benefiting from low investments in public research, which leads back to the question of the geography of resources and implies the existence of size-effects. When the relative scientific allocation is calculated for these regions (by comparing their importance in the total number of researchers to their importance in the GDP (2006 data)), we indeed note that Champagne-Ardennes,

Picardie and Haute Normandie are the three regions having the lowest allocation with respective indices of 32, 33 and 38 against an average of 80 for the Province regions.

## 5. Interregional collaborations: Importance of size and proximity effects

### *Model*

In the second part of our analysis, we will test a gravity model to evaluate the existence of potential spatial and non spatial proximity effects as well as the existence of size effects involved in the collaborations between each pair of regions. 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. It is this type of model that was used by Hoekman *et al.* (2010) to explain the intensity of co-publications in Europe and by Ponds *et al.* (2007) for the Netherlands and finally by Liang and Zhu (2002) and Scherngell and Hu (2010) for China. Maggioni and Uberti (2007) used a similar model to deal with the geography of research collaborations (within the EU Fifth Framework Program) and EPO co-patent applications.

The model we are testing here is more precisely a sample selection model (two-step Heckman model). We first suppose that the existence of collaborations for each pair ( $i$ ) of regions  $r_1$  and  $r_2$  (with  $i = 1, \dots, r_1 * r_2$ ) depends on a very first latent variable  $d_i^*$ , both unobservable and continuous, the linear combination of the exogenous variables  $Z_i$  representing here 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:

$$d_i^* = Z_i\beta^{(1)} + \mu_{i1}$$

When  $d_i^*$  is superior to a certain limit, we note that the two regions collaborate and when it remains inferior to that limit, they do not collaborate. Thus:

$$d_i = \begin{cases} 1 & \text{if } d_i^* \geq 0 \\ 0 & \text{if } d_i^* < 0 \end{cases}$$

The  $d_i$  indicator allows to see whether it is possible to observe the number of collaborations between two regions, represented by a second latent variable ( $f_i^*$ ) explained by the same exogenous variables since we would like to determine whether the two variables (probability and intensity of interregional collaborations) depend on the same factors:

$$f_i^* = Z_i\beta^{(2)} + \mu_{i2}$$

The intensity of collaborations  $f_i = \ln C_i$ ,  $C_i$  corresponding to the number of contracts between the  $r_1$  and  $r_2$  regions, is obtained as follows:

$$f_i = \begin{cases} f_i^* & \text{if } d_i^* \geq 0 \\ 0 & \text{if } d_i^* < 0 \end{cases}$$

We assume that the errors  $\mu_{i1}$  and  $\mu_{i2}$  are normal.

The explanatory variables are measured as follow:



1) Size effects: as explained before, the inequalities in the economic and scientific resources allocated to regions may influence the interregional breakdown of collaborations. Then, we integrate two explanatory variables linked to the size of regions,  $\ln X_{r_1}$  and  $\ln Y_{r_2}$ . The variable  $\ln X_{r_1}$  is the relative economic size of region  $r_1$ . To measure it, three indicators are available: the fraction of GDP, the number of establishments and the number of private researchers. Since these three measures are heavily correlated they are placeholder in the model; we thus arbitrarily choose the GDP to measure economic size. More precisely, the measure is the percentage of GDP (INSEE data), averaged over the period between 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 the period between 1997 and 2006.

2) Spatial proximity:  $D_{r_1, r_2}$  is the geographical distance between  $r_1$  and  $r_2$ . There were different complementary ways of measuring this distance. As said above, we considered 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 a train journey between each of the regional capitals. We will test the role of each kind of distance to check whether or not their respective effects differ.

A dummy  $INTRA_{r_1, r_2}$  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 that is not taken into account with the physical distance (that equals 0 when  $r_1 = r_2$ ), or with the scientific and sectoral proximities (see below).

To precise the effect of spatial separation between partners, it is important to control for non spatial separation measures that may influence the spatial distribution of collaborations. Scientific disciplines are not equally distributed across regions and certain research partnerships profiles are likely to be geographic dependant considering their specialization. We thus integrate a measure of distance between any two regions in terms of sectoral profile by constructing a sectoral proximity index.

3) Sectoral proximity :  $QI_{r_1, r_2}$  and  $QS_{r_1, r_2}$  correspond to the industrial proximity (sector of activity of firms) and scientific proximity (scientific domain of laboratories) respectively. To evaluate these proximities, we designed a matrix based on the Aquino Index (1978). This can be defined as follows:

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

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 between the regions in terms of sectoral activity and scientific domain. On the contrary, an index close to 0 indicates an important proximity between the regions. In order to use a relevant index that increases when the proximity increases, we simply write:

$$Q_{ij} = 1 - A_{ij}$$

In other words, two regions that have exactly the same industrial structure have an index equal to 1. The more the industrial weights differ, the less the proximity index. To evaluate scientific proximity, we use a matrix calculated by Mangematin *et al.* (2003) with data from 1999 (GERD by scientific domain). To assess industrial proximity, we use the value added by sector (INSEE, 14 sectors) in 1999.

4) Control variable: a second dummy for each region is integrated to control the potential effect of regions the partners belong to.

It must be noted that some variables have logarithmic measures. This characterization of the model appears to be theoretically necessary: if the economic size of the region  $r_1$  or the scientific size of the 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 to an endlessly negative logarithm. The parameter associated being positive, the value of the latent variable is also endlessly negative, showing a collaboration between such region and any other region to be impossible. The same reasoning can be used for the scientific size.

### Results

Concerning the different forms of proximity, we first note that on average, French regions are more distant from one another regarding scientific specialization than regarding industrial specialization, with the Aquino index being respectively 0.488 against 0.887 for the total number of French regions. The dispersion around this average is stronger regarding the scientific profile of regions (0.229 against 0.054). From a scientific point of view, The PACA and Aquitaine region are the closest with an index of 0.930); The Limousin and Auvergne regions being the most distant (with an index of 0.010). From the industrial point of view, the most distant are Ile de France and Champagne-Ardennes (with an index of 0.694) and the closest being Midi-Pyrénées and Aquitaine (with an index of 0.955), followed by Nord Pas de Calais and Lorraine (with an index of 0.952).

The correlation matrix between the different forms of proximity (Table 5) reveals that there is a good correlation between the two measures of spatial proximity. However, the other correlations are rather weak. It is especially the case for scientific and industrial proximity being weakly correlated ( $r=0.140$ ). Nevertheless, even low, the correlation coefficients between spatial proximity and sectoral proximity are always negative.

**Table 5 : spatial and sectoral proximities (correlation matrix)**

	D1	D2	QI	QS
D1 : Spatial proximity (frontiers)	1.000			
D2 : Spatial proximity (train)	0.627	1.000		
QI : Industrial proximity	-0.207	-0.065	1.000	
QS : Scientific proximity	-0.279	-0.356	0.140	1.000

The estimation of the model gives the results presented in table 6. The quality of the model is rather good: the confusion matrix designed from the predictions of the selection equation leads to an 85% rate of true positives and a 74% rate of true negatives<sup>3</sup>. As for the output equation, the  $R^2$  equals 0.82, we can consider that the significant variables provide a good explanation for the collaboration

<sup>3</sup> The pseudo- $R^2$  of the selection equation (presented in the previous section) equals 0.385.

intensity. Finally, the inverse of Mills ratio is significantly positive at 1% level. It indicates the presence of a selection bias that was controlled by our model.

**Table 6 : The determinants of interregional collaborations (sample selection model)**

PROBABILITY/INTENSITY OF INTERREGIONAL COLLABORATIONS					
Selection equation :					
	Estimate	Std. Error	t value	Pr(> t )	Significance
(Intercept)	-0.879	1.120	-0.785	0.433	
lnX	1.925	0.373	5.167	0.000	***
lnY	1.156	0.238	4.858	0.000	***
D	-0.198	0.091	-2.189	0.029	**
INTRA	5.412	1166.178	0.005	0.996	
QI	13.094	3.775	3.468	0.001	***
QS	0.625	0.523	1.196	0.233	
Alsace	-0.101	0.360	-0.279	0.780	
Aquitaine	Ref.				
Auvergne	0.465	0.450	1.032	0.303	
BasseNormandie	0.026	0.449	0.058	0.954	
Bourgogne	0.101	0.436	0.231	0.818	
Bretagne	0.453	0.335	1.353	0.177	
Centre	-0.056	0.382	-0.147	0.883	
ChampagneArdenne	1.192	0.544	2.191	0.029	**
FrancheComte	1.345	0.532	2.526	0.012	**
HauteNormandie	0.452	0.444	1.017	0.310	
IDF	5.198	1248.017	0.004	0.997	
LanguedocRoussillon	0.550	0.376	1.462	0.144	
Limousin	2.074	0.653	3.177	0.002	***
Lorraine	0.360	0.381	0.945	0.345	
MidiPyrenees	1.249	0.462	2.703	0.007	***
NordPasdeCalais	0.401	0.432	0.929	0.353	
PaysdeLaLoire	0.379	0.351	1.079	0.281	
Picardie	0.748	0.445	1.683	0.093	*
PoitouCharentes	0.392	0.420	0.933	0.351	
PACA	0.812	0.631	1.287	0.199	
RhoneAlpes	0.252	0.581	0.434	0.664	
Outcome equation:					
	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	-0.550	0.380	-1.448	0.148	
lnX	1.322	0.105	12.590	< 2e-16	***
lnY	0.839	0.084	10.028	< 2e-16	***
D	-0.301	0.040	-7.588	0.000	***
INTRA	2.512	0.255	9.870	< 2e-16	***
QI	-1.166	1.186	-0.983	0.326	
QS	0.433	0.228	1.903	0.058	*
Alsace	0.069	0.159	0.431	0.666	
Aquitaine	Ref.				
Auvergne	-0.006	0.189	-0.032	0.975	
BasseNormandie	0.128	0.211	0.604	0.546	
Bourgogne	-0.053	0.189	-0.282	0.778	
Bretagne	0.128	0.141	0.908	0.364	
Centre	-0.231	0.158	-1.464	0.144	
ChampagneArdenne	0.027	0.226	0.120	0.904	

FrancheComte	0.314	0.213	1.474	0.141	
HauteNormandie	0.087	0.195	0.444	0.657	
IDF	-0.688	0.255	-2.694	0.007	***
LanguedocRoussillon	-0.189	0.151	-1.249	0.212	
Limousin	0.875	0.243	3.601	0.000	***
Lorraine	0.384	0.166	2.313	0.021	**
MidiPyrenees	0.341	0.145	2.356	0.019	**
NordPasdeCalais	0.189	0.178	1.060	0.290	
PaysdeLaLoire	-0.185	0.141	-1.314	0.190	
Picardie	0.265	0.187	1.419	0.157	
PoitouCharentes	-0.072	0.178	-0.403	0.687	
PACA	-0.318	0.182	-1.745	0.082	*
RhoneAlpes	0.257	0.141	1.820	0.069	*

Multiple R-Squared:0.8257

Adjusted R-Squared:0.8095

Error terms:					
	Estimate	Std. Error	t value	Pr(> t )	
invMillsRatio	0.4584	0.175	2.619	0.00916	***

Notes : \*P<0.1, \*\*P<0.05, \*\*\*P<0.01

The analysis of individual effects highlights that the variables are heavily significant both in the selection and output equation. It results that the economic and scientific size-effects and spatial distances are decisive factors for the probability and intensity of collaborations. We thus confirm the existence of size-effects suggested in the literature by Grossetti and Nguyen (2001) notably and the important role of any kind of spatial distances as already observed by Maggioni *et al.* (2007), Abramovsky and Simpson (2008) or Laursen *et al.* (2010).

The industrial proximity has a significant and positive effect in the selection equation but not in the output equation. In other words, the probability to collaborate increases with the industrial proximity but the intensity does not. This result seems to confirm the previous hypothesis: by contracting locally laboratories of a certain region are endowed with relevant resources they can use in the future to collaborate with firms belonging to the same industry, located outside the region. Conversely, the scientific proximity does not impact the probability to collaborate but impact the intensity of partnerships. In addition, it plays a negative role on the intensity of collaborations even if this impact is reduced (the variable is significant at a 10% level). These opposite results between the scientific proximity and the industrial one are difficult to interpret without any complementary researches. At least, behaviors seem different according to the type of partner.

In the selection equation, industrial proximity has a very significant negative impact but has no effect in the output equation. Thus, the proximity of industrial profiles lowers the probability of collaborations but does not impact the intensity of collaborations. It is the reverse for scientific proximity: it does not impact the probability of collaborations but it has a negative impact on the intensity of collaborations, even if the effect is relatively moderate (the variable being significant at a 10% level). In both cases, we must note that the proximity of sectoral profiles negatively affects collaborations for innovation, since they take place between different regions rather than similar ones.

As for the regional effects, they are relatively low (5 significant coefficients in the selection equation, 6 in the output equation), the explanatory variables mostly providing answers to the probability and intensity of collaborations. In the case of Midi-Pyrénées, we observe the existence of positive and significant regional effects in the two equations. These two regions seem to outperform regarding collaborations for innovation. In the selection equation, all the regional fixed effects are positive (for Limousin, Midi-Pyrénées, Franche-Comté, Picardie and Champagne-Ardenne). In the output equation, if 4 regions obtain positive and significant coefficients, two others obtain negative and significant ones, namely Ile de France and the PACA region. Thus the Capital region seems to underperform regarding collaborations for innovation.

We finally noted that the very significantly positive impact of the INTRA variable on the intensity of collaborations but not on the probability of these collaborations. The co-location of the actors in the innovation process would lead to a relatively great number of collaborations for other reasons than those evidenced by the physical distance variable.

## **6. DISCUSSION AND CONCLUSION**

The aim of this paper was to complement empirical studies on the spatial dimension of collaboration for innovation, drawing on a relational and long term database to measure factors already known and effects sometimes neglected in the literature.

The statistical and econometric investigations conducted confirm first the multiscale dimension of the geography of science-industry collaborations and the important and positive effect of spatial proximity, this latter being measured in terms of time travel or thanks to regional borders.

We add also interesting results taking into account a dynamic perspective. We reveal the inertia of this geography and highlight the stability of the local dimension (the weight of intra-regional contracts does not increase during the studied period in spite of the evolution of the number of contracts) and of the average distance that separates partners.

Last but not least, we give new explanations of the geography of science-industry collaborations testing regional and sectoral effects. Despite what could be expected considering the existing literature, the first econometric model carried out evidence the weakness of sectoral effects. Only the service sector (for the industry) and the humanities and social science field of study (for scientific domains) seem to be atypical influencing positively and significantly the geography of collaborations. In the second econometric model (sample selection model), we highlight the existence of sectoral effects: the industrial proximity plays a positive role in the probability to collaborate but plays any part in the intensity of collaborations; the scientific proximity plays a negative role (but less strongly) in the intensity but not in the probability. Complementary investigations are necessary to better interpret these results.

On the contrary, our paper reveals the particularly structuring role played by the regional effects. Though these results may seem obvious, it still deserves to be underlined since the geography of collaborations is particularly influenced by these effects. Moreover, we have provided detailed information regarding this result and showed the existence of a capital effect: the data collected not only evidence the hypertrophy of Ile De France in the spatial breakdown of science-industry contracts but also the presence of the capital region's overshadowing effects on the neighboring regions. These last two effects especially raise the question of the geography of resources (i.e. inequality of the regional resources allocated) and implies the existence of size effects, as confirmed

by the gravity model. Both the probability and intensity of the collaborations are mainly influenced by the size of the regions and especially their economic size.

This last set of results involves strong implications in terms of public policies. Indeed, if the tendency to collaborate essentially depends on the resources allocated to regions, the measures taken to prompt researchers and firms to collaborate together are likely to have a limited efficiency: the fact that some regions tend to collaborate much less than other regions would be linked to the structural features of territories rather than to the under optimal behavior of the actors. If policy-makers' aim is to raise the number of collaborations, then it will be necessary to act on these structural features by tackling the question of the economic development of regions (how to strengthen the creation process of resources in the neighboring regions?) and posing the problem of the geography of scientific resources on which public policies have a strong freedom of action.

## REFERENCES

- Abramovsky L. and Simpson H., 2008, Geographical proximity and firm-university innovation linkages: evidence from Great Britain, *Working paper CMPO*.
- Anselin L., Varga A. and Acs Z., 1997, « Local geographic spillovers between university research and technology innovation », *Journal of Urban Economics*, Vol. 42 (3), pp. 422-448.
- Audretsch D. and Feldman M., 1996, "Knowledge Spillovers and the Geography of Innovation and Production", *American Economic Review*, 86 (3), pp. 630-640.
- Autant-Bernard C., Mairesse J. and Massard N., 2007a, "Spatial knowledge diffusion through collaborative networks", Introduction to the special issue of *Papers in Regional Science*, Vol. 86 (3).
- Autant-Bernard C., Billand P. and Massard N., 2007b, « Social distance versus spatial distance In R & D cooperation: Empirical evidence from European collaboration choices In micro and nanotechnologies », *Papers In Regional Science*, Vol. 86 (3), pp. 495-519.
- Bergman E., 2010, « Knowledge links between European universities and firms: A review », *Papers in Regional Science*, Vol. 89 (2), pp. 311-333.
- Boschma R. and Frenken K., 2009, « The spatial evolution of innovation networks. A proximity perspective », In R. Boschma and R. Martin (eds) *Handbook of Evolutionary Economic Geography*, Cheltenham.
- Boschma R., 2005, « Proximity and innovation: a critical assessment », *Regional studies*, Vol. 39 (1), pp. 61-74.
- Bouba-Olga O. and Grossetti M., 2008, « Socio-économie de proximité », *Revue d'Economie Régionale et Urbaine*, n°3.
- Brouillat E. and Lung Y., 2010, « Spatial distribution of innovative activities and economic performances : a geographical-friendly model », *Druid Summer conference*, London, 16-18 June.
- Carayol N., 2003, « Objectives, agreements and matching in science-industry collaborations: Reassembling the pieces of the puzzle », *Research Policy*, Vol. 32, pp. 887-908.
- Cairncross F., 1997, *The Death of Distance: How the Communications Revolution will Change Our Lives*, Harvard Business School Publishing, Cambridge.
- Castells M., 1996, *The Rise of the Network Society*, Blackwell Publishers, Oxford.
- Cohen W. and Levinthal D., 1990, « Absorptive capacity: A new perspective on learning and innovation », *Administrative Science Quarterly*, Vol. 35, pp. 128-152.
- Cooke, N., 1994, « Varieties of knowledge elicitation techniques », *International Journal of Human-Computer Studies*, Vol. 41 (6), pp. 801-849.
- D'Este P. and Iammarino S., 2010, « The spatial profile of university-business research partnerships », *Papers in Regional Science*, Vol. 86, pp. 423-443.
- Etzkowitz H., Webster A. and Healey P., 1998, « Capitalizing knowledge. New intersections of industry and academia », SUNY.
- Feldman M., 1994, *The geography of innovation*, Kluwer Academics Publishers, Dordrecht.

- Feldman MP. and Florida R., 1994, « The Geographic Sources of Innovation: Technological Infrastructure and Product Innovation In the United States », *Annals of the Association of American Geographers*, Vol. 84, pp. 210-229.
- Ferru M., 2010, « Formation process and geography of science-industry partnerships : the case of the University of Poitiers », *Industry and Innovation*, Vol. 17 (5).
- Ferru M., 2009, “La géographie des collaborations pour l’innovation”, Thèse ès sciences économiques, Université de Poitiers.
- Frenken K., Hardeman S., Hoekman J., 2009, “Spatial scientometrics: towards a cumulative research program”, *Journal of Informetrics*, Vol. 3 (3), pp. 222–232.
- Gibbons M., Limoges C., Nowotny H., Schwartzman S., Scott P. and Trow M., 1994, *The new production of knowledge, The dynamics of science and research In contemporary societies*, Sage, Londres.
- Godin B. and Gingras Y., 1999, « L’impact de la recherche en partenariat sur la production scientifique », *Dossier de recherche de l’AUCC*, Vol. 3 (3).
- Grossetti M. and Bès MP., 2003, « Dynamique de réseau et des cercles. Encastremets et découplages », *Revue d’Economie Industrielle*, Numéro spécial sur la morphogenèse des réseaux, n°103, 2<sup>ème</sup> et 3<sup>ème</sup> trimestres, pp. 43-58.
- Grossetti M. and Nguyen D., 2001, « La structure spatiale des relations science-industrie en France : l’exemple des contrats entre les entreprises et les laboratoires CNRS », *Revue d’Economie Régionale et Urbaine*, Vol. 2, pp. 311-328.
- Hagedoorn J. 2002, « Inter-firm R&D partnerships: an overview of major trends and patterns since 1960 », *Research Policy*, Vol. 314, pp. 477-492.
- Hagedoorn J. and Roijackers N., 2006, « Inter-firm R&D partnering in pharmaceutical biotechnology since 1975: trends, patterns, and networks », *Research Policy*, Vol. 35 (3), pp. 431-446.
- Hoekman J., Frenken K. and Tijssen R., 2010, « Research collaboration at a distance : changing spatial patterns of scientific collaboration within Europe », *Research Policy*, Vol. 39 (5), pp. 662-673.
- Jaffe A., 1989, “Real effects of academic research”, *The American Economic Review*, Vol. 79 (5), pp. 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.
- Lallement R., Mouhoud EM. and Paillard S., 2007, « Polarisation et internationalisation des activités d’innovation : incidences sur la spécialisation technologique des nations », In A. Rallet et A. Torre A. (dir), *Quelle proximité pour innover ?*, Ed. L’Harmattan.
- Laursen, K., Reichstein, T., Salter A., 2010, « Exploring the effect of geographical proximity and university quality on university-industry collaboration in the UK », *Regional Studies*, pp. 1-17.
- Levy R., Roux P. and Wolff S., 2009, « Study of science-industry collaborative patterns in a large European university », *Journal of Technology Transfer*, Vol. 34 (1), pp. 1-23.
- Liang L. and Zhu L., 2002, « Major factors affecting China’s inter-regional research collaboration : Regional scientific productivity and geographical proximity », *Scientometrics*, Vol. 31, pp. 31-43.



- Lung Y., 1997, « Organisation spatiale et coordination des activités d'innovation des entreprises », *Rapport pour le Commissariat Général du Plan*, octobre.
- Maggioni M., Nosvelli M. and Uberti T., 2007, « Space versus networks in the geography of innovation », *Papers in Regional Science*, Vol. 86(3), pp. 471-493.
- Maggioni M. and Uberti T., 2007, « Inter-regional knowledge flows in Europe : an econometric analysis », In Frenken K. (eds) *Applied evolutionary, economics and economic geography*, Edward Elgar, Cheltenham, UK.
- Mangematin V., Massard N., Autant-Bernard C., Carrere M., Coronini R., Nesta L. and Riou S., 2003, « Etude sur les profils scientifiques et techniques régionaux », *Rapport pour le ministère de l'éducation nationale*.
- Menzel MP., 2008, « Dynamic proximities Changing relations by creating and bridging distances », *Papers In Evolutionary Economic Geography*, n°08.16, Utrecht University.
- Ponds R., Van Oort F. and Frenken K., 2007, « The geographical and institutional proximity of research collaboration », *Regional Science*, Vol. 86 (3), pp. 423-443.
- Puga D., 1999, « The rise and fall of regional inequalities », *European Economic Review*, Vol. 43 (2), pp. 303-334.
- Rallet A. and Torre A. 2005, « Proximity and localization », *Regional Studies*, Vol. 391, pp. 47-59.
- Saxenian A., 1994, *Regional advantage: culture and competition in Silicon Valley and Route 128*, Cambridge, Harvard University Press.
- Scherngell T. and Hu Y., 2010, « Collaborative Knowledge Production in China: Regional Evidence from a Gravity Model Approach », *Regional Studies*, In press.
- Singh, 2005, « Collaborative networks as determinants of knowledge diffusion patterns », *Management Science*, Vol. 51 (5), pp. 756-770.
- Wagner C. and Leydesdorff L., 2005? "Network structure, self-organization, and the growth of international collaboration in science", *Research Policy*, Vol. 34 (10), pp. 1608-1618.