NBER WORKING PAPER SERIES

MEASUREMENT AND EXPLANATION OF THE INTENSITY OF CO-PUBLICATION IN SCIENTIFIC RESEARCH: AN ANALYSIS AT THE LABORATORY LEVEL

Jacques Mairesse
Laure Turner

Working Paper 11172
http://www.nber.org/papers/w11172

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2005

We wish to thank Michèle Crance, Serge Bauin and the members of the Unité des Indicateurs de la Politique Scientifique (UNIPS, CNRS), for their help in compiling the data base and their advice. We have also benefited from remarks by Ajay Agrawal, Marie-Laure Allain, Céline Allard-Prigent, Anne Crubelier, Dominique Foray, Claudine Hermann, and are particularly grateful for detailed comments to Bronwyn Hall and Edward Steinmueller.

Measurement and Explanation of the Intensity of Co-publication in Scientific Research: An Analysis at the Laboratory Level
Jacques Mairesse and Laure Turner
NBER Working Paper No. 11172
March 2005, Revised July 2010
JEL No. D29,O39

ABSTRACT

In order to study networks of collaboration between researchers, we propose a simple measure of the intensity of collaboration, which can be easily interpreted in terms of relative probability and aggregated at the laboratory level. We first use this measure to characterize the relations of collaboration, as defined in terms of co-publication between the scientists of the French “Centre National de la Recherche Scientifique” (CNRS) in the field of condensed-matter physic, during the six-year period 1992-1997. We then use it to investigate the importance of various factors of collaboration: mainly the geographical distance between laboratories, but also their specialization and size, their productivity and the quality of their publications, and their international openness. We find that the average intensity of co-publication of researchers within laboratories is about 40 times higher than the average intensity between laboratories if they are located in the same towns, and that it is 100 times higher than the intensity between laboratories which are not located in the same towns. Yet, geographical distance does not have a significant impact, or a very weak one, on the existence and intensity of co-publication between laboratories located in different towns. What matters is immediate proximity. We also find that the productivity of laboratories, their size and specialization profiles are significant determinants of collaboration.

Jacques Mairesse
CREST-INSEE
15, Boulevard Gabriel PERI
92245 MALAKOFF CEDEX
FRANCE
and NBER
mairesse@ensae.fr

Laure Turner
CREST - Labratoire d'Economie Industrielle
28 rue des Saints-Pere
75007 Paris
France
laure.turner@ensae.fr

1 Introduction

Since the scientific research system has become an essential sector in our modern knowledge-based economies, an important new research field has opened up. The challenge is to illuminate the role of scientific institutions in the production, diffusion and transfer of knowledge and that of science in economic development and social welfare. The "new economics of science" therefore is interested in a variety of issues concerning the functioning of scientific institutions, the labor market, training and careers of scientists, their productiveness, the allocation of public funds to basic research, the design of intellectual property rights, etc…. It thus contributes to the understanding of the organization of science and of ways it can be improved (Dasgupta and David, 1994; Gibbons and al., 1994; Diamond, 1996; Stephan, 1996; Callon et Foray, 1997; Shi, 2001; Foray, 2004).

The analysis of co-publications between scientists presented in this contribution is in keeping with the main focus of the economics of science on knowledge-production, and is part of a broader study of the determinants of scientific research productivity. We believe that membership in a dynamic and productive laboratory favors collaboration between researchers and improve their own individual productivity, and that it may be part of a process of cumulative advantage by which these researchers enhance their productivity and reputation.[1]. Given the substantial increase in the proportion of co-authored articles, it also seems that the relevant units of knowledge production tend to be more and more specific networks of researchers, whether they belong or not to the same institutions and/or countries (Gibbons and al., 1994).

In the economics of science, until recently, the literature on the interactions that favors knowledge production and diffusion primarily concerns geographic externalities. Authors have mainly studied such externalities within industries or from universities and other public research institutions to firms and industries, relying on the analysis of patent data.[2] Our work moves upstream to study knowledge externalities within the scientific research system using co-publication data. We wish to look beyond the observation of the spatial dimensions of research activity to investigate the determinants of the occurrence and intensity of collaborative relations between researchers. Audretsch and Stephan (1996) have done similar work but again concerning the relations between public research and industry. Based on data on academic scientists collaborating with US biotechnology firms, they show that such collaboration between firms and researchers is more likely when the researchers have a good academic reputation, when they belong to a geographically extensive network, and when they are involved in practice in the transfer of knowledge towards the firm (as participants in the creation of the firm or as members of the Scientific Advisory Board). Regional and local characteristics also seem to influence the strength of relations between scientists and firms.

In the sociology of science and in bibliometry, a number of studies have already highlighted some of the factors facilitating collaboration within academic research (see Beaver and Rosen, 1978 and 1979, and Katz, 1994, for a summary presentation). They include, above all, researchers' reputation and visibility, the need to access or to share the use of specific research instruments and facilities, the increasing specialization in science and geographic proximity. Two types of analyses can be found, however, in this literature, depending on their explicit or implicit conception of a network (Shrum and Mullins, 1988). In one line of analysis the actors in networks are identified through their interrelations, being mainly differentiated by their different positions in the structural configuration of their networks (e.g. whether they occupy a central position or not), not by their individual characteristics such as age, gender or skills.[3] By contrast, the second line of research takes into explicit account the status, capacities and strategies of actors, and it is these individual characteristics that mainly determine the position of agents in networks and the nature of interactions between them.[4] Yet it would be desirable to be able to include in the same analysis structural and individual elements as determinants of network interactions, and particularly for collaboration in research. Knowledge production and diffusion are based on the interactions of multiple agents and institutions with diverse interests: scientists in public and private laboratories, firms, financiers, public authorities, and so on (Callon, 1999). Investigating the existence and intensity of collaboration between researchers in relation with their specific characteristics should afford insight into the various mechanisms at play.

In this chapter we present the first results of such an attempt. We propose an intensity measure of collaboration between researchers, which have an intuitive interpretation and can be simply aggregated to the laboratory level or higher levels of aggregation. Our unit of analysis in this contribution is the laboratory and the group of laboratories at the geographic level of a town (which to be short we will call "town"). Our purpose is to explain measured differences of intensity of collaboration as revealed by co-publications by various potential determinants: precisely the geographic distance between laboratories, their thematic specialization, their size, their productivity in terms of average number of publications per researcher, their quality in terms of average citation impact factor per publication, and also their international openness.[5] In particular to what extent does the geographic distance between researchers and their laboratories strongly impede, or not, their scientific collaboration?

Our approach is basically descriptive. We measure the intensity of co-publications among the researchers of the French Centre National de la Recherche Scientifique (CNRS) in the field of condensed matter physics, during the six-year period 1992-97.[6] We first estimate the intensity of co-publication among these researchers, within their laboratories and between them, and also within and between the towns in which these laboratories are located. Next we consider by means of simple correlations the possible influence of geographical distance and other determinants on the occurrence and intensity of co-publication. We then try to better assess the specific impacts of these different factors by estimating their relative weight in a regression analysis.

The chapter is organized as follows. In section 2, we give necessary information on the scope of our study, the construction of our sample, and some descriptive characteristics of co-publication. In section 3, we define our measure of intensity of collaboration, giving a detailed example of its computation. In section 4, we present our correlation and regression results and comment on what they tell us of the respective importance of the various determinants of co-publication we have been able to consider. We briefly conclude in section 5.

2 Scope of the study and general characteristics of co-publication

2.1 Scope of the study: the collaboration between CNRS researchers in condensed matter physics

In this chapter we study the determinants of collaboration among a group of 493 physicists belonging to the condensed matter section at the CNRS, over the six-year period 1992-1997. This sample consists of practically all the CNRS physicists in this field who were born between 1936 and 1960 and were still working at the CNRS in 1997.[7] Condensed matter physics investigates, at various scales (atom, molecules, colloids, particles or cells), all states of matter from liquids to solids in which molecules are relatively close to each other. It is based on a heritage of traditions, both experimental (crystallography, diffusion of neutrons and electrons, magnetic resonance imagery, microscopy, etc.) and theoretical (solid state physics). It has recently developed a closer relation with industry, contributing to the development of materials used in electronics, plastics, food or cosmetic gels, and so forth. We chose condensed matter physics for three main reasons. First, the characteristics of this field are particularly well suited to our study: it is a domain of basic research, which is clearly defined and where the journals with a sound reputation are easily identifiable. Second, condensed matter is a fast-growing field, honored by the Nobel Prize for Physics awarded to Pierre-Gilles de Gennes in 1991, and currently accounting for close to half of all French research in physics. Third, there is relatively very little mobility among CNRS researchers outside of the field in other fields of research in CNRS, or out of CNRS towards academia or industry.

The sample of 493 physicists studied here represents a majority of all CNRS researchers in the field. The CNRS and higher education institutions are the only public research institutions in this field in France. In 1996, there were a total of 654 condensed matter physicists in CNRS, as against 1475 in universities and "Grandes Ecoles" (Barré et al., 1999).

The fact that our study is limited to researchers belonging to the same institution, the CNRS, comes in fact as an advantage. It implies a strong organizational proximity between the researchers, characterized by the sharing of common knowledge and implicit or explicit rules of organization that favor interaction and coordination (Rallet and Torre, 2000; Foray, 2004). Because they all belong to the same scientific community within the same institution, they work in a context directly conducive to cooperation that does not involve prior agreement on rules of behavior. The existence of such strong organizational proximity thus makes it possible to isolate more clearly the effects on collaboration of geographic distance proper and other factors.

The indicator of collaboration that we use in this study is co-publication. It seems to be a reliable indicator of collaboration without being an exhaustive measurement, in so far as collaboration can have results other than publications. Our data base has been compiled on the basis of all the publications drawn from the Science Citation Index (SCI), for 518 CNRS condensed matter physicists over the period 1992-1997, of whom 493 published at least one co-authored article during this six years.[8] Of the remaining 25 physicists, 21 in fact published no articles in this period, and the other four published only a total of five non-co-authored articles. Collaboration appears to be the main mode of publication for the 493 researchers. Only 132 of them also wrote articles without co-authors over the period (for a total of 252 articles), and from the total corpus of 7,784 articles they wrote over the period, 7,532 (97% !) are in fact co-authored.

In order to measure better the intensity of collaboration in our analysis, we thought appropriate to weight co-authored article in proportion to the number of couples or pairs of co-authors they involve. In other words we simply chose to study the network of collaboration "link by link," that is by pairs of co-authors. In practice, this means that an article appears in the data-base we constructed as many times as the number of different pairs of its CNRS co-authors.[9].

We also chose to center our study at the level of the laboratory, and even at the more aggregate level of groups of laboratories in the same towns or localities ("towns"). We thus consider networks of collaboration between laboratories and towns rather than directly between individual researchers. When two researchers belonging to different laboratories (towns) collaborate, we consider that these laboratories (towns) collaborate, and on this basis we can measure the intensity of collaboration between laboratories (towns). When two researchers belonging to the same laboratory (town) collaborate, we also simply consider it as a case of collaboration "within" this laboratory (town), and likewise we compute the intensity of collaboration within laboratories (within towns). We can also similarly compute intensity of collaboration between laboratories-within towns. Carrying out our study at the aggregate level of laboratories and towns simplifies somewhat the analysis and makes the use of our measure of collaboration intensity perhaps more convincing, since networks of collaboration are of course much denser at these levels than at the individual researcher level. But, as we shall see, it also has the advantage that it allows for a direct characterization of the influence on collaboration of working in the same laboratory or town, and thus of the importance of spatial proximity and easy face-to-face relations.

2.2 Two configurations of co-publication

FIGURE 10.1 and TABLE 10.1 about here or below in this sub-section

The co-authors of the articles of our group of 493 CNRS researchers, which we will simply call "CNRS researchers" from now on, can be (these) CNRS researchers themselves, or other researchers, mainly belonging to universities or other institutions, either French or foreign, whom we will call "external researchers", who are mainly from. In our analysis, we are led to distinguish between two configurations of co-publications, depending on whether a publication involves *at least two CNRS researchers* and possibly other researchers (CNRS or external), or whether it concerns *at most one CNRS researcher* and one or more external researchers. An important reason for this distinction is a practical one. We not only preferred *a priori* to focus our analysis on the collaboration among CNRS researchers in the same field, but also we could not extend it in practice to the external researchers in this field. The CNRS researchers were the only ones for whom we could have access to the name, location and some characteristics of their laboratories, in addition to their individual characteristics (age, gender, seniority, etc …).[10] This was not possible for the external researchers since we could not even retrieve the name and location of their laboratories with sufficient reliability from the SCI.[11] We were thus left for them with much more limited information than for the CNRS researchers and their laboratories.

Our group of 493 CNRS researchers generally co-publish both with the other CNRS researchers of the group and with external researchers. Precisely, as indicated in Figure 10.1, 38 of them collaborating only with CNRS researchers (never with external researchers), and 69 only with external researchers (not with the other CNRS researchers), and thus 386 (= 493 – 38 – 69) collaborating in both ways. The first configuration of co-publication (involving at least two CNRS researchers and possibly other researchers) corresponds to „Group 1' with a total of 1,823 articles (= 1741 + 82), while the second (involving only one CNRS researcher with external researchers) corresponds to „Group 2' with a total of 5709 articles (= 5012 + 697). Group 1 thus concerns 424 of our CNRS researchers (= 493 – 69), while Group 2 concerns 455 of them (= 493 – 38).

Table 10.1 shows the two-way distribution of the number of articles in Group 1 and Group 2 with respect to the number of their CNRS authors and that of their external authors (see also the related distribution shown in Figure 10.2 below). We observe immediately that, for both Group 1 and Group 2 articles, collaboration generally involves several "external" researchers, (82 articles are written by CNRS researchers only!). We can also note that most articles of Group 1, which have at least two CNRS co-authors, do not involve a third (or more) CNRS co-author (1498 out of 1823). Thus, the average number of authors per article for Group 1 is of 5.9, of which 2.2 are CNRS researchers and 3.7 external researchers, and for Group 2 it is of 4.9 (i.e., 1 CNRS researcher and 3.9 external researchers).

2.3 The selected sample of co-publications and some characteristics

Four main reasons determine our choice of limiting our analysis to the first configuration of co-publications and Group 1 of articles. The first reason, which we already stressed, is analytical. By studying co-publication between couples of CNRS researchers, we control for institutional and organizational proximity resulting in "common knowledge" of rules and practices and strongly favoring collaboration. Organizational proximity and geographical proximity being usually confounded, this has the great advantage of allowing us to unravel clearly the impact of the latter on collaboration. The second reason, which we also mentioned, is simply that we cannot identify precisely enough the laboratories of the "external" researchers, and thus cannot locate them nor characterize them, as we can do for the laboratories of the CNRS researchers.

But there is a third important reason of empirical nature for focusing our investigation on the collaboration between CNRS researchers. The occurrence of co-publication between a CNRS and an external researcher is extremely low, while it is much higher, as we would expect, between couples of CNRS researchers. The 1,823 articles in Group 1, written by 424 CNRS authors and about 3500 external co-authors, actually involve only 880 different couples of CNRS researchers out of the 89676, (=423x424/2) number of potential couples (that is one out of a hundred). The 5709 articles in Group 2, written by 455 CNRS authors with close to 10000 external co-authors, involve by contrast as much as 17500 couples of a CNRS researcher with an external researcher, out of the 4550000 potential couples (that is only four out of a thousand). Thus, the average number of articles per effective couple of co-authors is of 2.1 in Group 1 and only about 0.3 in Group 2. Likewise the probability (frequency) for a CNRS author to have another CNRS co-author (in Group 1) is much higher than to have an external co-author (in Group 2): 0.021 as against 0.001 (!).

FIGURES 10.2, 10.3 and 10.4 about here or below in this sub-section

A last consideration arises from the fact that some characteristics of co-publication in Group 1 and Group 2 are nonetheless close enough. This suggests that hopefully a number of the results we find in the analysis of co-publication between CNRS researchers might not be too different from those we would have obtained if we had been able to extend the analysis to the co-publication with non-CNRS researchers. This is clear for the three following characteristics that we can compute for the sample of 6,753 articles published by the 386 CNRS researchers involved in both types of publication (see Figure 10.1). The first of them is the frequency distribution of the number of articles per number of external co-authors. As shown in Figure 10.2 the probability (frequency) that an article is co-authored by a given number of external researchers is nearly the same in the two Groups of articles. The second very close characteristic concerns the degree of concentration of the number of articles published in the two Groups of articles by the CNRS researchers. As shown in Figure 10.3 the concentration curves practically coincide in both cases, with nearly 40% of the articles being co-authored by 10% of the most productive CNRS researchers, and about 80% by the more productive half of them.

Yet, as can be seen on Figure 10.4, the distribution of the number of articles written per CNRS researcher (our third characteristic) differs somewhat for the two Groups of articles. During the six-year period 1992-1997, the cumulative probability that a CNRS researcher publishes less than six articles in Group 1 is 50%, while it of about 35% in Group 2. Likewise, during this period, a CNRS researcher published an average of 9.9 articles in Group 1, as against 13 in Group 2.[12]

2.4 Other restrictions on the selected sample

In practice, in order to avoid having laboratories and towns with too few CNRS researchers we thought better to put two further restrictions on our sample. We imposed that laboratories had in our sample at least 5 CNRS researchers, and towns at least nine CNRS researchers. Our final sample thus consists of 470 CNRS researchers in condensed matter physics (out of the initial group of 493), located in 34 laboratories and 17 towns. Likewise, in our analysis, we thought better to avoid characterizing collaboration between two laboratories, or collaboration between two towns, on the basis of too few co-publications between their CNRS researchers. We thus defined collaboration between a couple of laboratories as involving more than 4 co-publications over the six year study period, and between a couple of towns as involving more than 6 co-publications. These two types of restrictions had the consequence of limiting also the number of Group 1 articles (with at least two CNRS co-authors), on which our analysis concentrates, to 1634 articles (out of 1741). To summarize, our investigation is thus mainly based on a sample of 470 CNRS condensed matter physicists (located in 17 towns and 34 laboratories) and a sample of 1634 articles they have co-published over the period 1992-1997.

3 Measurement of intensity of collaboration

The behaviors of agents in networks is determined by "intrinsic" individual characteristics such as age, gender, skills, motivations and objectives, and by more "structural" variables such as the density of their networks, their more or less central or peripheral situation, geographic distance, etc. As a result, the form and functioning of networks differ. If the actors were not differentiated and if they collaborated with all the others with equal probability, we would expect to observe a uniform structure of relations between all the individuals. We take this extreme case of "homogeneity" as a reference. At the aggregate level of entities such as the laboratories and groups of laboratories (towns) on which we center our analysis, the case of homogeneity corresponds to a configuration in which the frequency of collaboration of agents, the CNRS researchers, is the same, irrespective of the entities to which they belong, their geographic localization and other characteristics. Our simple measure of (relative) intensity of collaboration between two entities is simply based on the comparison between the real network as portrayed by the data and the network that would be observed in the hypothetical case of homogeneity. We generally define this measure in sub-section 3.1, and comment on its aggregation properties and on the weighting issues in sub-sections 3.2 and 3.3. In sub-section 3.4 we then provide a detailed example of its calculation.

3.1 Definition

In this sub-section we assume for simplicity that collaboration always involves at the most two (CNRS) researchers (this assumption is discussed in the next sub-section). The network of collaboration studied has a finite number of entities (laboratories or towns) consisting in total of N researchers who can form C collaboration pairs, or couples, where by definition $C = N(N-1)/2,$ the total number of possible pairs. Let n be the total number of articles produced in collaboration between the N researchers, then p the frequency of the number of co-publications per pair in the complete network is the ratio between the total number of articles n, and the number of possible pairs C, that is $p = n/C$.

Using similar notations at the level of the network's entities, consider now two entities X and Y, where N_X and N_Y are the numbers of researchers working in them respectively. The numbers of possible pairs of researchers within X and within Y are respectively $C_X = N_X(N_X-1)/2$ and $C_Y = N_Y(N_Y-1)/2$, and the number of possible pairs that can be formed between researchers from X and Y is $C_{XY} = N_X N_y$. If the total numbers of articles written jointly within X and Y are respectively n_X and n_Y , and the total number of articles written in common by researchers in X and Y is n_{XY} , the frequencies of collaboration within the entity X and Y, noted as p_X and p_Y, are the corresponding ratios between the total number of articles n_X and n_Y written together by researchers from X or Y, and the number of possible pairs C_X and C_Y of researchers in entity X and Y, that is $p_X = n_X/C_X$ and $p_Y = n_Y/C_Y$. Similarly the frequency of collaboration p_{XY} between the two entities X and Y is the ratio between the total number of articles n_{XY} written in common by researchers in X with researchers in Y, and the number of possible pairs C_{XY} of researchers from the two entities, that is $p_{XY} = n_{XY}/C_{XY}$.

The intensity of collaboration relates the frequencies obtained at the entities' level to the frequency p obtained for the complete network. We thus define the two intra- or within-entity intensities and the inter- or between-intensity as:

$$i_X = \frac{p_X}{p} = \frac{n_X / C_X}{n / C} \qquad i_Y = \frac{p_Y}{p} = \frac{n_Y / C_Y}{n / C} \qquad i_{XY} = \frac{p_{XY}}{p} = \frac{n_{XY} / C_{XY}}{n / C}$$

Note that in what follows we will be using indifferently the expression intra- or within- intensity, and inter- or between intensity.

In the reference case of homogeneity of the network we have $p_X = p_Y = p_{XY}$ for all X and Y, and. consequently we can see that $p_X = p_Y = p_{XY} = p$, or in terms of the intensity measure: $i_X = i_Y = i_{XY} = 1$. In the case of homogeneity, the frequencies of collaboration intra- and inter-entities are all equal to the overall frequency p for the network, and the intra- and inter- intensities of collaboration are all equal to unity. Otherwise, in the case of a real network, as the one we are considering, various factors influence intensities of collaboration; we can expect them of course to be very different from unity, which can be viewed as an average benchmark value.

Note that another way of looking at our measure of intensity of collaboration of an entity is to interpret it as its contribution of co-authored articles n_X to the total number n of co-authored articles in the network, normalized by its size relative to that of the complete network measured in terms of possible pairs of co-authors, i.e. . $i_X = p_X /p= (n_X/n)/(C_X/C)$. Note also that the structure of intensity of a network of E entities can be represented by means of a symmetrical matrix E by E with positive or zero coefficients where the diagonal terms are equal to the intra-entity intensities, and the off-diagonal terms are equal to the inter-entity intensities.[13] Appendix 1 gives this matrix for the 17 towns in our sample.

3.2 Aggregation properties

The (relative) intensity of collaboration as defined above has the advantage of being easy to aggregate at different levels of analysis. In order to see this, suppose that V is a town with two laboratories X and Y. The total number of co-authored articles written in V is the sum of co-authored articles by researchers from X and Y separately, and from X and Y jointly. Likewise, the number of possible pairs of researchers in V is the sum of the possible pairs of researchers in X and in Y separately, and between X and Y. We thus can write:

$$\frac{n_V}{C_V} = \frac{n_X + n_Y + n_{XY}}{C_X + C_Y + C_{XY}} = \frac{n_X}{C_X} \text{ x } \frac{C_X}{C_V} + \frac{n_Y}{C_Y} \text{ x } \frac{C_Y}{C_V} + \frac{n_{XY}}{C_{XY}} \text{ x } \frac{C_{XY}}{C_V}$$

or in terms of frequencies and intensities of collaboration:

$$p_V = w_X p_X + w_Y p_Y + w_{XY} p_{XY} \quad \text{or} \quad i_V = w_X i_X + w_Y i_Y + w_{XY} i_{XY}$$

$$\text{where} \quad w_X = \frac{C_X}{C_V} \quad w_Y = \frac{C_Y}{C_V} \quad w_{XY} = \frac{C_{XY}}{C_V} \quad \text{with } w_X + w_Y + w_{XY} = 1$$

This formula can easily be extended to groups of more than two laboratories. Aggregating over the entire network, we have

$$\sum_{\mathrm{I}} w_I i_I + \sum_{\mathrm{I,J \neq I}} w_{IJ} i_{IJ} = 1 \quad \text{with} \quad \sum_{\mathrm{I}} w_I + \sum_{\mathrm{I,J \neq I}} w_{IJ} i_{IJ} = 1$$

3.3 Remark on the weighting

Until now, we have considered for simplicity that the articles were co-authored by two (CNRS) researchers. In reality, they can also be written by threesomes or foursomes of (CNRS) researchers, etc. But, as already indicated (in sub-section 2.1), we thought appropriate to study the network of collaboration "link by link," that is, by couples or pairs of co-authors. In practice, this means that an article is repeated in our data base (and thus counted) as many times as there are pairs of different (CNRS) co-authors. For example, for an article published by three CNRS researchers, one belonging to a laboratory *X* and the two others to a laboratory *Y*, we count three co-publications – two between *X* and *Y* and one within *Y*.[14] Note that, if we follow this procedure, the aggregation formula (as we simply write it in the previous sub-section) applies more generally in the case where there are more than two (CNRS) co-authors for an article. Note also, that in practice in our case, since only 20% of the articles in Group 1 are co-authored by more than two CNRS researchers, the choice in the weighting assumption should not make really an important difference.

3.4 Practical calculation: an example

TABLE 10.2 about here or below in this sub-section or next one 4.1

Let us take the concrete example of the town of Marseille to describe in detail the calculation of our measure of the intensity of collaboration, using the information displayed in Table 10.2, which also gives the results of this calculation for the other towns. Marseille (as indicated in column 1) is a town with 18 CNRS researchers (among the 470). These researchers are involved in 34 co-publications among themselves (column 3) and in 18 co-publications with the CNRS researchers from two other towns (column 4), 10 of them with Grenoble and 8 with Strasbourg.[15] The number of possible couples of researchers working in Marseille is 18*17/2, or 153. The frequency of collaborations per couple of researchers in Marseille is therefore 34/153 or 0.22. Given that the numbers of researchers in Grenoble and Strasbourg are of 105 and 14, the number of possible couples of researchers linking Marseille and Grenoble and Marseille and Strasbourg are respectively 1890 (= 105*18) and 252 (= 14*18). The corresponding frequencies of collaborations per couple are therefore 0.0053 (= 10/1890) and 0.0317 (= 8/252).

In order to compute the intensities of collaboration, we have also to calculate *p,* the overall frequency of collaboration per couple of researchers for the complete set of the 17 towns. It is the ratio between the total weighted number of articles, 2480 (=1,715+765), and the number of possible couples that can be formed by the 470 CNRS researchers, i.e. 110215 pairs (= 470*469/2). We thus have p = 0.0225. This overall frequency p is also the intra- (or within-) and inter- (or between-) frequency of collaboration that would have been obtained for Marseille and all the other towns in the hypothetical case of homogeneity. In fact, the intra-frequency for Marseille is much higher (0.22) than this reference value, the inter-frequency of collaboration with Grenoble much lower (0.0053), and that with Strasbourg relatively closer (0.0317). Finally, the intra-town intensity for Marseille is of 0.22/*p* or 9.88 (column 5). Likewise, the Marseille-Grenoble and Marseille-Strasbourg inter- intensities are of 0.24 and 1.41 respectively, yielding a mean inter-intensity of collaboration of Marseille with all the other 16 towns of (0.24+1.41)/16 or 0.1 (column 6), and a mean inter-intensity of Marseille with only its two effective partners of (0.24+1.41)/2 or 0.82 (column 7).

4 Results: the importance of geographical proximity and quality of scientific environment

We look first at the estimated intensities of co-publication between the CNRS researchers at the town level (Table.10.2, and Appendices 1 and 2). Next, we consider in detail the statistical evidence on the potential determinants of co-publication we have been able to measure, which is mainly provided by simple correlations computed both at the town and laboratory levels (Tables 10.3 to 10.5, and Appendix 3). Lastly, we assess the robustness of these results by examining the multivariate regressions of the occurrence and intensity of co-publication on these various determinants (Table 10.6).

4.1 Intensity of co-publication at the town level

The estimated inter-town intensities of co-publication among all the different couples of towns, as we can see from the matrix of co-publication intensity in Appendix 1 and from the graph of the co-publication network in Appendix 2 (and also from their averages by towns computed in Table 10.2), are extremely dispersed. Of the 136 possible couples of towns, only 34 are effectively collaborating.[16]. Grenoble, Orsay and Paris are the main nodes in the network of collaboration, being respectively linked to 12, 9 and 7 other towns, whereas Poitiers, Orléans and Talence appear to be isolated.[17] Among all effectively linked couples of towns (or partner towns), the intensity of co-publication ranges from a lowest value of 0.24 for Grenoble-Marseille and Grenoble-Meudon to a highest value of 7.40 for Bagneux-Villeuneuve d'Ascq. As could be expected (and checked by looking at the number of CNRS researchers in our sample per town, given in column 2 of Table 10.2) the towns with the largest number of CNRS researchers are the ones which tend to have the more links with other towns but also the lower inter-town intensity estimates.

The estimated intra-town intensities of co-publication (given in column 6 of Table 10.2) are much higher than the inter-town intensities, with very few exceptions. They are always greater than one, and on average equal to 18.9, as compared to an average inter-intensity of 0.3 when computed over all couples of towns and of 1.0 when computed only over the partner towns. This strongly points to a major influence of geographical proximity on the intensity of collaboration. Note also that intra-town intensity tends to be high in towns with few partners like Meudon, Poitiers, Strasbourg, Villeneuve-d'Ascq or Villeurbanne compared to towns with many partners like Grenoble, Orsay and Paris, which have among the lowest intra-town intensities (5.4, 3.6 and 3.0 respectively). This result could mainly be explained by the larger size of the CNRS research community in Grenoble, Orsay and Paris and the fact that these towns host several laboratories, both characteristics entailing numerous potential links among which relatively many do not occur. As a matter of fact, the co-publication intensities estimated at the laboratory level for Grenoble, Orsay and Paris are much higher, being in average respectively equal to 19.4, 33.4, and 34.5, and quite comparable to that of the laboratories of the other towns (see column 3 in Table 10.5).

4.2 Determinants of the occurrence and intensity of co-publication: correlation evidence

We consider six a priori influential determinants of collaboration at the laboratory or town levels, which we have been able to approximately measure or proxy: geographic distance, specialization, size, productivity, quality of publications, and international openness. Apart from geographic distance and distance in specialization which are directly defined for a couple of laboratories or towns, it is somewhat problematic to adopt a priori a single measure for our four other variables, such as their average over the two laboratories or towns concerned, say for example $(S_I + S_J)/2$ where S is a measure of size of the laboratories or towns I and J. Thus, in addition to the average, we used the maximum and minimum values, say S_I and S_J for the couple of laboratories or towns I and J. Note that for size we have also three different possibilities: the number of CNRS researchers, say N_I for laboratory or town I; the number of possible pairs of CNRS co-authors, say (N_I x N_J) for the couple (I, J) of laboratories or towns –or (N_I x N_I-1)/2) within the laboratory or town I – and the number of publications over the six-year period, say S_I for laboratory or town I.

TABLES 10.3 and 10.4 about here or below in this sub-section

We will examine all these variables in turn. Their means at the town level are given in Table 10.3; and their correlations with the binary indicator of occurrence of co-publication and with our measures of both intra- and inter-intensity, also at the town level, are displayed in Tables 10.4. These correlations are also recorded in Appendix 3 at the laboratory level, both overall (for the 34 laboratories) and within towns (for the 7 towns out of 17, which have more than one laboratory). The statistical evidence is quite consistent at both the town and laboratory levels, with the notable exception of the correlations of the occurrence of co-publication with specialization at the laboratory level within towns. It is also in general qualitatively comparable for the occurrence and intensity of co-publication, the one major exception being the size variable positively correlated with occurrence and negatively with intensity.

4.2.1 Geographic distance

The average distance of a town from its partners can vary widely (see Table 10.3). At the two extreme, Montpellier collaborates with 7 other towns, situated at an average distance of 550 km (kilometers), while Gif-sur-Yvette is related to 3 towns much closer, at an average distance of 170 km, two of them being also located in the Parisian region. Four of the five towns situated at less than 300 km in average from their partners are in the Parisian region (besides Paris *intra-muros*, Bagneux, Gif sur Yvette, Meudon, Orsay, and Palaiseau). The geographic distance apparently plays a negligible role, or only a slightly negative one, in the occurrence of collaboration, as well as on its intensity. This is shown by the correlations computed at the town level but also at the laboratory level. The relevant two correlations at the town level (-0.09 and -0.16) are negative but both statistically non significant, and the two ones at the laboratory level (-0.09 and -0.06) are also negative, with only the first moderately significant (at a 5% confidence level).

In fact, as we already noted in comparing the values of the intra- and inter- town intensities of co-publication presented in Table 10.2, proximity has a major influence on collaboration. This is confirmed by the comparison of the intra- and inter-laboratory intensities shown in Table 10.5. But more interestingly, this can also be qualified, since by comparing the inter-laboratory-intra-town intensity to the inter-laboratory-inter-town intensity we distinguish between what we shall call immediate proximity and local proximity. Immediate proximity, which is that of researchers working in the same laboratory, usually located in a common building, favors frequent face-to-face interactions and can be expected to induce and facilitate collaboration. Local proximity, which is that of researchers working in different laboratories, but still relatively close as when in the same town, can be expected to be less conducive to collaboration than immediate proximity. Nonetheless local proximity should be more favorable to collaboration than when researchers are working in really distant laboratories, as when located in different towns. This is indeed very clearly what we find. For example for Grenoble and its six laboratories, the average intra-laboratory intensity is 19.4, which is about seven times higher than the inter-laboratory-intra-town intensity (2.7), itself about 5 times higher than the average inter-laboratory- inter-town intensity (0.5). In average for all 17 towns, the pattern is the same, the three average intensities being respectively 30.7, 0.8 and 0.3.

TABLES 10.5 about here in this sub-section

One should thus distinguish three scales of geographic distance, which influence very differently collaboration. Immediate proximity has a considerable impact and local proximity is also favorable but much less. By contrast, beyond proximity, geographic distance strongly hinders collaboration, but *per se* and only slightly, if at all, in proportion to real distance (say in kilometers). Such findings are well corroborated by prior studies on knowledge flows between public laboratories and industrial firms, which show that proximity allows face-to-face interactions and exchanges of tacit knowledge between actors inducing them to build a common understanding rather than referring only to a common „text' (see for example Zucker and al., 1998 a and b, and Leamer and Storper, 2000). New communication technologies have certainly contributed to the "death of distance", by helping researchers to collaborate much more easily and faster; however they did not do away with the crucial importance of proximity.

4.2.2 Specialization

Proximity in specialization, not only geographic proximity, should also strongly influence collaboration in a field as diverse and large as condensed matter physics. The network of co-publication presented in Appendix 2 is by itself suggestive of such an influence. Orsay and Grenoble, which appear to be two central nodes of the network, are indeed the location of the two French storage rings, which are very large facilities used by physicists of condensed matter. [18]

We have tried to take into account the specialization of laboratories (or towns), although there is no easy and good way to do so. We have defined a profile of specialization of a laboratory (or town), based on the classification by the main theoretical and/or experimental "sub-domains" of the journals in which the CNRS researchers of this laboratory (or town) are publishing. Such classification is difficult but seems carried out relatively well by the Science Citation Index (SCI). We found that the most frequently listed sub-domains of the journals in which the articles of the CNRS condensed matter physicists are published were physics-chemistry, general physics, solid-state physics, applied physics, materials science, and crystallography. We then characterized an "overall specialization profile" of each laboratory (or town) by the [7, 1] column vector defined by the proportions of publications of its CNRS researchers in these six main sub-domains and the group of other sub-domains. We also considered the seven "specific specialization profiles" corresponding to the seven [2, 1] column vectors defined by the proportions of publications in each of the seven sub-domains and all the six others.

Next, to measure the distance in specialization between all couples of laboratories (or towns) we adopt the Chi-squared distance between their specialization profiles. To facilitate interpretation we also normalized this measure in such a way that the average distance between any one laboratory (or town) and all others will at most be equal to one if the laboratories (or towns) had specialization profiles which were not statistically different at the 1% confidence level.[19] The average specialization distances, (in terms of the overall profiles) between towns, given in Table 10.3, show that we are in fact far from this hypothetical situation. Specialization profiles are quite diverse, and the specialization distances as we measure them vary widely (the lowest town average specialization distance being of about 3 for Bagneux and Villeneuve d'Ascq and the highest one of about 11 for Marseille and Talence).

With the major exception of the puzzling positive and significant between-laboratory-within town correlations with the occurrence of co-publication (and some minor ones concerning particularly the specialization indicator in crystallography and again the occurrence of co-publication), all other correlations of our overall and specific specialization distance measures with the occurrence and intensity of co-publication are consistently negative (see Table 10. 4 and Appendix 3). The correlations with the existence of co-publication, however, are mostly small and not significant, while the correlations with the intensity tend to be sizeable and statistically significant. The two between laboratory-between town and between-laboratory-within town correlations of intensity of co-publication and overall specialization distance are, for example, as large as -0.3 and -0.4 respectively. Note also that the puzzling exception of the between laboratory-within town correlations with the occurrence of co-publication may largely reflect the correlations of our measure of specialization with the other determinants of co-publication, since the corresponding coefficient in our regression analysis is not statistically significant (see sub-section 4.3 and Table 10.6).

On the whole, our evidence thus tend to show that proximity in specialization favors strongly the intensity of collaboration, but much more weakly so, or not, its existence. This latter result is not what we expected. It is clear, however, that our attempt here at measuring specialization is crude, and that much remains to be done to better characterize what it is and to assess its impact on collaboration.

4.2.3 Size

We considered three different measures of size. The first relies on the number of CNRS researchers in the laboratories (or towns) concerned, and the second on the total number (or stock) of their publications over the six year 1992-1997.[20] The third, which is particularly well suited to our definition of intensity of co-publication, is the number of possible couples of CNRS researchers for each couple of laboratories (or towns). For the first two measures, as already explained, we experimented with the average, the maximum and minimum of the number of researchers and of the stock of their publications for each couple of laboratories (or towns). Not surprisingly, these three types of size measures are overall quite consistent, as shown by the descriptive statistics in Table 10.3. Grenoble comes first of all towns with 22% of the total number of CNRS researchers involved, 20% of the total number of possible couples of co-authors among them, and 27% of their total publications. Paris comes second and Orsay third (with respectively about 18% and 14% of the number of researchers and of all possible couples of co-authors, and for each of them roughly 14% of all their publications).

A priori one would expect that the size of laboratory (or town) would impact positively the chance of collaboration, but not its intensity, since by construction our measure of intensity takes already into account such a size effect. One would even think likely that the larger the laboratories (or towns) involved in collaboration, the smaller its intensity. This is indeed what we see clearly when looking at the correlations in Table 10.4 and in Appendix 3 for the different indicators of size we used. The correlations of nearly all of them, both at the town level and the laboratory level (within- and between-town), are thus very significantly positive and substantial (ranging from 0.2 to 0.6) with the occurrence of co-publications, while very significantly negative in a comparable range (from -0.2 to -0.6) with its intensity. Note that it is also the case that the correlations of the intensity of co-publication of the researchers within their own laboratories (or towns) with the size of their laboratories (or towns) tend to be significantly negative.

4.2.4 Productivity

The productivity of the laboratories (or towns) is simply measured as the stock of publications of their CNRS researchers in the period 1992-1997 per researcher (that is as the ratio of our two first measures of size just defined in the sub-section above).[21] As can be seen in Table 10.3, productivity varies widely from one town to another, from a minimum of 6.3 articles per researcher over six years for Orleans to a maximum of 36.4 for Bagneux, the overall mean being of 15.7 articles per researcher (that is 2.6 articles per year). In contrast with size, it seemed a priori likely that both the correlations of productivity with the occurrence of co-publication and its intensity should be positive. This is definitely what we find. Nearly all these correlations, including the two ones of the within-town and within-laboratory intensities with productivity, are very significantly positive and of a sizeable order of magnitude, from 0.2 up to 0.7 (see Table 10.4 and Appendix 3).

4.2.5 Quality of publications

Our measure of the quality of publications of laboratories (or towns) is consistent with that of their productivity. It is the average impact factor (or impact score) per publication of their stock of publications over the period 1992-1997. Precisely, it is the weighted mean of the impact factors of the journals in which these publications have appeared (the weights being the numbers of publications in the different journals). The impact factors of the journals are provided by the SCI; they are defined and computed as the average number of citations per article received by the articles published in the journals over a period of two and five years. We used here the two years impact factors, but using the five years impact factors did not make a difference in our results. Our measure of the quality of publications of a laboratory (or town) is thus an estimate of the expected number of citations that the publications of its researchers will in average receive over two years. In Table 10.3, we see that this average number is overall of 3.4 citations per article over two years, and that it can differ by a factor of 2 at the town level, being lowest for Poitiers, with a citation rate of 2.3, and highest for Palaiseau, with a citation rate of 4.8.

Although we expected that the quality of publications, like productivity, would be positively correlated with both the occurrence and intensity of collaboration, the evidence (recorded in Table 10.4 and Appendix 3) is mixed. There are no statistically significant correlations with the intensity of co-publication at the town or laboratory levels. We find a significantly positive correlation with the occurrence of co-publication only when we use as our quality indicator the minimum value for the couples of laboratories or towns involved (of 0.23 at the town level and 0.06 at the laboratory level). This suggests, interestingly but tentatively, that what matters in establishing a collaboration is a minimum quality requirement on the two partners involved. To confirm such proposition would of course need a more detailed analysis and which should be performed at the researcher level and not only at the aggregate level of the laboratory. Note, however, that the between-laboratory-within-town correlations with the occurrence of co-publication are all very significantly negative, raising a similar puzzle as the one we have with our specialization indicators.

4.2.6 International openness

As we already indicated (in sub-section 2.2), we cannot precisely locate the laboratories of the very many "external" researchers who are co-publishing with our sample of 470 CNRS physicists. However, it is possible to identify the foreign (non French) addresses among all those listed in the SCI electronic records for all their articles (both in Group 1 and Group 2). In spite of the imprecision of such information, we can thus build an indicator of international openness of the laboratory as the proportion of articles of their CNRS researchers (over the six year 1992-1997), involving a least one foreign co-author. In Table 10.3, we see that this proportion is overall about 30%, and that it is the highest, about 50%, for Grenoble and Paris, and the lowest, about 15%, for Bagneux, Gif sur Yvette and Strasbourg.

Our a priori thought was that international openness would also go together with greater occurrence and intensity of collaboration between the CNRS researchers themselves and their laboratories (or towns). This is what we observe, although the evidence is not strong. Many of the correlations given in Table 10.4 and in Appendix 3 are not significantly different from zero, but those that are significant are all positive.

4.3 Regression confirmatory evidence

TABLES 10.6 and 10.7 about here or below in this sub-section

To assess the robustness of the evidence provided by our analysis of simple correlations, we did a number of regressions of both the occurrence and the intensity of co-publication on the six a priori influential variables we have been able to consider. All of them mainly told the same story, confirming all those of our observations, which were already strongly supported by the correlation evidence.[22] We present in Tables 10.6 and 10.7 the regressions we did at the laboratory level, which include all six variables measured in the simplest way (that is for specialization in terms of overall profile, for size as the average of number of researchers in all couples of laboratories, and similarly for productivity, quality of publications and international openness also as the average of the corresponding indicators on all couples of laboratories).

The geographic distance between laboratories does not influence the intensity of collaboration and has only a small significantly negative impact on its occurrence -- an increase of 100 km in the distance between two laboratories corresponding to a decrease on the frequency of co-publication of less than 1% (0.8%). As expected, the distance in specialization has a negative impact on the intensity of co-publication, which seems sizeable although statistically not very significant. An increase of one standard deviation in the distance of specialization, as we characterize it, will thus imply a fall of nearly 30% in the intensity of co-publication between-laboratory-between-town.

Laboratory size has a very significant and large impact on collaboration: positive on its occurrence, while negative on its intensity. A 10% increase of the average size of each laboratory will thus entail an increase of the frequency of collaboration within-town and between-town of about 15% and 25% respectively, while it will correspond to a decline of the intensity of co-publication within-laboratory of nearly 10%, and between-laboratory-between-town of about 10% also. Laboratory productivity has positive and mostly significant effects on both the occurrence and intensity of collaboration, which are of the same or even larger orders of magnitude than the size effects. A 10% increase of productivity will thus involve an increase of 15% and 30% of the frequency of collaboration within-town and between-town respectively, and will result in a rise of about 20% the intensity of co-publication within-laboratory.

The quality of laboratory publications does not seem to have a significant impact on collaboration, except one which is negative, contrary to our a priori expectation, on the frequency of co-publication between-laboratory-within town (confirming the puzzling simple correlations we already noted). Clearly the international openness of laboratories, at least in the way we can proxy for it, has also apparently no significant influence on collaboration.

5 Conclusion

In order to study networks of collaboration between researchers, we proposed a simple measure of the intensity of collaboration, which can be intuitively interpreted in terms of relative probability and easily aggregated at the laboratory level. We first used this measure to characterize the relations of collaboration between the scientists of the French “Centre National de la Recherche Scientifique” (CNRS) in the field of condensed-matter physic, during the six-year period 1992-1997. We then used it to investigate the importance of various factors of collaboration: mainly the geographical distance between laboratories, but also their specialization and size, their productivity, the quality of their publications and their international openness.

We find that the average intensity of co-publication of researchers within laboratories is about 40 times higher than the average intensity between laboratories if they are located in the same towns, and about 100 times higher than the intensity between laboratories if they are not located in the same town. Yet, geographical distance does not have a significant impact, or a very weak one, on the existence and intensity of co-publication between laboratories located in different towns. There is basically three scales of geographic distance. Immediate proximity, which allows easy face-to-face interactions, has a considerable impact on collaboration, while local proximity is also relatively favorable but much less. Geographic distance *per se*, that is beyond proximity, remains by contrast a strong obstacle to collaboration, but only slightly, if at all, in proportion to real distance.

Although our measure of specialization between laboratories remains crude, we find that proximity in specialization has also a large positive influence on the intensity collaboration. The size of laboratories and their productivity in terms of number publications per researcher appear to be influential determinants of collaboration, having both a positive impact on the occurrence of co-publication, but a negative impact for size and a positive one for productivity on the intensity of co-publication. Contrary to our expectations, we do not really observe significant effects of the average quality of publications and of international openness of laboratories. However, this may be due, at least in part, to the fact that these two indicators, as we have been able to construct them, are at best imperfect proxies.

In future work, it will thus be necessary to improve by and large the measurement of the potential determinants of collaboration we have been able to consider, as well as to extend the list of these determinants. Clearly it will also be important to broaden the scope of our study, which remains mainly illustrative. In particular, although we think it is appropriate and interesting to analyze collaboration at the level of the laboratory, as we did here, it will be both enlightening and challenging to carry out this analysis together with an investigation at the individual researcher level. This will undoubtedly lead to an assessment of the role of "star scientists" in the scientific performance of their own laboratories and in the formation and development of networks of collaboration. By focusing on the co-publication between researchers working in the same institutional setting, that of the French CNRS, we have been able to control for organizational proximity. Comparing similar studies in different research environments could be very instructive by itself. But of course, trying more generally to integrate institutional and organizational characteristics in the analysis and to understand how they can enhance or hinder collaboration should be a central objective in the research agenda --- one that will keep up with the high standards of Paul David.

Bibliography

AUDRETSCH D. and STEPHAN P., 1996, "Company-Scientist Locational Links: The case of biotechnology", *American Economic Review*, 86(3), 641-652.

BARRE R., CRANCE M. and, SIGOGNEAU A., 1999, "La recherche scientifique française : situation démographique", *Etudes et dossiers de l'OST*, 1.

BLAU J., 1973, "Patterns of Communication among Theoretical High Energy Physicists", *Sociometry*, 37, 391-406.

BEAVER D. and ROSEN R., 1978, "Studies in Scientific Collaboration. Part 1. The Professional Origins of Scientific Co-authorship", *Scientometrics*, 1, 65-84.

BEAVER D. and ROSEN R., 1979, "Studies in Scientific Collaboration. Part 2. Scientific Co-authorship, Research Productivity and Visibility in the French Scientific Elite", *Scientometrics*, 1, 133-149.

CALLON M., 1999, "Le Réseau comme forme émergente et comme modalité de coordination : le cas des interactions stratégiques entre firmes industrielles et laboratoires académiques", in Callon M. et al., *Réseau et coordination*, Paris: Economica.

CALLON M. and FORAY D., 1997, "Introduction : nouvelle économie de la science ou socio-économie de la recherche scientifique ?", *Revue d'Economie Industrielle*, 79, 13-37.

COLE J. and COLE S., 1973, *Social Stratification in Science*, Chicago: U. of Chicago Press.

CRANE D., 1969, "Social Structure in a Group of Scientists: A Test of the Invisible College Hypothesis", *American Sociological Review*, 34, 335-352.

CRANE D., 1972, *Invisible Colleges: Diffusion of Knowledge in Scientific Communities*, Chicago: University of Chicago Press.

CRAWFORD S., 1971, "Informal Communication among Scientists in Sleep Research", *Journal of the American Society for Information Science*, 22, 301-310.

DASGUPTA. P. and DAVID P. A., 1994, "Toward a New Economics of Science", *Research Policy*, 23(5), 487-521.

DAVID P., 1994, "A Science Simulation for Studying the US and similar Institutional Setups (SCISIMUS)", Mimeo.

DIAMOND A., 1996, "The Economics of Science", *Knowledge and Policy*, 9 (2-3).

FORAY D., 2004, *The Economics of Knowledge*, Cambridge, Mass. The MIT Press.

GIBBONS M., LIMOGES C., NOVOTNY H., SCHWARTZMAN S., SCOTT P. and TROW M., 1994, *The New Production of Knowledge*, London: Sage.

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

JAFFE A. and TRAJTENBERG M., 1998, “International Knowledge Flows: Evidence from Patent Citations”, NBER Working Paper 6507.

JAFFE A., TRAJTENBERG M. and HENDERSON R., 1993, “Geographic Localization of Knowledge Spillovers as Evidenced from Patent Citations”, *Quarterly Journal of Economics*, 108(3), 557-598.

KATZ J.S, 1994, “Geographical Proximity and Scientific Collaboration”, *Scientometrics*, 31(1), p. 31-43.

LEAMER E. and STORPER M., 2000, “The Economics of Geography at the Internet Age”, mimeo, University of California at Los Angeles and Management and Public Policy, November.

RALLET A. and TORRE A., 2000, “Is Geographical Proximity Necessary in the Innovation Networks in the Era of Global Economy?”, mimeo, Université Paris Dauphine et Institut National de la Recherche Agronomique.

SHI Y., 2001, *The Economics of Scientific Knowledge,* Chetelham, UK: Edward Elgar.

STEPHAN P., 1996, “The Economics of Science”, *Journal of Economic Literature*, 34, 1199-1235.

SHRUM W. and MULLINS N., 1988, “Network Analysis in the Study of Science and Technology”, in *Handbook of Quantitative Studies of Science and Technology*, ed. A.F.J. Van Raan, North Holland.

TURNER L. and MAIRESSE J., 2002, “Explaining Individual Productivity Differences in Public Research: How Important Are Non-Individual Determinants? An Econometric Analysis of French Physicists (1986-1997)”, Working Paper- Cahiers de la MSE 2002-66, Université Paris I- Panthéon-Sorbonne.

Unité des Indicateurs de la Politique Scientifique (UNIPS), 1999, “Les Publications des laboratoires du CNRS et leur impact”, éd. CNRS, mars.

ZUCKER L, DARBY M. and ARMSTRONG J., 1998a, “Intellectual Human Capital and the Birth of U.S. Biotechnology Enterprises”, *American Economic Review*, 88, 290-306.

ZUCKER L, DARBY M. and BREWER M. 1998b, “Geographically Localized Knowledge: Spillovers or Markets?”, *Economic Inquiry*, 36, 65-86.

Table 10.1. ***Number of articles in Group 1 and Group 2 by number of CNRS and external co- authors***

Number of "external" co-authors:	0	1	2	3	4	5	6	7	8	9	Total
Group 1 (at least 2 CNRS co-authors) Of which:	82	230	324	375	260	209	127	81	56	161	1823 (424*)
2 CNRS co-authors	64	196	268	300	218	172	106	61	47	66	1498
3 CNRS co-authors	15	31	45	60	31	33	15	15	6	7	257
4 CNRS co-authors	3	2	9	12	8	4	6	4	3	4	55
5 or more CNRS co-authors	0	1	2	3	3	0	0	1	0	3	13
Group 2 (only 1 CNRS co-author)	0	726	1087	1114	976	708	441	241	128	288	5709 (455**)
Total (group 1 and 2)	**82**	**956**	**1411**	**1489**	**1236**	**917**	**568**	**322**	**184**	**367**	**7532 (493)**

( ) The three numbers in parentheses below the numbers of articles are the numbers of CNRS researchers co-authoring these articles. * Including 38 CNRS researchers who never published with external researchers and who account for 82 publications of group 1 articles. ** Including 69 researchers who never published with other CNRS researchers and who account for 697 publications of group 2 articles.

Note that 132 CNRS researchers who also published co-authored papers have written alone 252 articles (not included in the first group or second group of articles).

Table 10.2. ***Descriptive statistics and within– and between-town intensity of co-publication at the town level #***

Towns	Number of CNRS researchers (*)	Number of laboratories per town (*)	Number of partner towns (**)	Number of articles "within"	Number of articles "between"	Intensit y within-town	Intensity between-town (average computed on all other 16 towns)	Intensity between-town (average computed on partner towns only)
Bagneux	9	1	6	51	171	63.0	1,3	3,5
Gif sur Yvette	16	1	3	11	40	4.1	0.2	0.9
Grenoble	105	6	12	666	449	5.4	0.7	0.9
Marseille	18	1	2	34	18	9.9	0.1	0.8
Meudon	9	1	2	27	19	33.3	0.1	0.5
Montpellier	20	3	7	47	83	11.0	0.3	0.8
Orléans	10	1	0	7	0	6.9	0.0	0.0
Orsay	66	3	9	174	192	3.6	0.2	0.4
Palaiseau	18	2	4	15	45	4.4	0.2	0.9
Paris	86	6	7	249	148	3.0	0.3	0.6
Poitiers	11	1	0	31	0	25.1	0.0	0.0
Saint Martin d'Hères	31	2	5	161	193	15.4	0.3	0.9
Strasbourg	14	1	2	72	20	35.2	0.1	0.9
Talence	9	1	0	8	0	9.9	0.0	0.0
Toulouse	29	2	4	88	63	9.6	0.4	1.6
Villeneuve d'Ascq	10	1	3	39	31	38.5	0.6	3.0
Villeurbanne	9	1	2	35	58	43.2	0.2	1.4
Total	**470**	**34**	**68**	**1715** [a]	**765** [b]			
Mean	**27.6**	**2.0**	**4.0**			**18.9**	**0.3**	**1.0**

The overall frequency of co-publications for the sample of 470 CNRS researchers is P = 0.0225
* Towns with less than 9 CNRS researchers and laboratories than less than 5 CNRS researchers are not considered.
** Partner towns are defined as having more than 6 articles co-published by their CNRS researchers over the six year period 1992-1997 (that is at least an average of one co-publication per year).
a Each article is weighted by the number of pairs of authors that contribute to its publication, otherwise the number of articles would be 1222.
b Each article is weighted by the number of pairs of authors that contribute to its publication, otherwise the number of articles would be 412.

Table 10.3. ***Descriptive statistics for the main determinants of co-publication at the town level***

Town	Number of scientists	Number of possible couples "between"	Number of possible couples "within"	Stock of publications between 1992 and 1997	Mean geographic distance to partners	Mean distance of specialization	Mean productivity	Mean quality of publications	Mean Proportion of articles co-authored with foreigners
Bagneux	9	4149	36	328	344	3.02	36.44	3.68	0.12
Gif sur Yvette	16	7264	120	246	171	3.93	15.38	3.07	0.14
Grenoble	105	38325	5460	1870	421	10.19	17.81	3.39	0.50
Marseille	18	8136	153	235	361	11.17	13.06	2.84	0.44
Meudon	9	4149	36	99	208	5.97	11.00	2.63	0.20
Montpellier	20	9000	190	365	548	5.63	18.25	3.47	0.21
Orléans	10	4600	45	63	0	8.24	6.30	3.54	0.32
Orsay	66	26664	2145	922	334	4.97	13.97	3.69	0.25
Palaiseau	18	8136	153	274	297	4.32	15.22	4.77	0.33
Paris	86	33024	3655	985	291	6.72	11.45	3.75	0.48
Poitiers	11	5049	55	88	0	5.17	8.00	2.34	0.26
Saint Martin d'Hères	31	13609	465	438	418	4.44	14.13	3.78	0.27
Strasbourg	14	6384	91	248	449	6.45	17.71	3.69	0.16
Talence	9	4149	36	193	0	11.07	21.44	3.94	0.43
Toulouse	29	12789	406	379	410	5.30	13.07	2.71	0.24
Villeneuve d'Ascq	10	4600	45	184	305	3.12	18.40	4.13	0.28
Villeurbanne	9	4149	36	139	188	4.55	15.44	3.02	0.23
Total	**470**	**194176**	**13127**	**7056**					
Mean	**27.6**				**279**	**6.13**	**15.71**	**3.44**	**0.29**

Table10.4: ***Correlations at the town level with the occurrence and intensity of co-publication***

	Intensity within-town (N=17)	Occurrence between-town (N=136)	Intensity between-town (N=34)
Geographic Distance	-	-0.09	-0.16
Distance in Specialization			
Overall Profile	-	-0.02	-0.21
Physics-Chemistry		-0.18**	-0.14
General Physics		-0.20***	-0.20
Solid-state Physics		-0.14	-0.24
Applied Physics		0.06	-0.08
Materials Science		-0.02	-0.06
Crystallography		0.16*	-0.17
Other		-0.00	-0.16
Size			
Number of researchers			
N_I	- 0.48 *	-	-
Maximum (N_I, N_J)	-	0.52***	- 0.41**
Minimum (N_I, N_J)		0.42***	- 0.31*
Average $(N_I + N_J)/2$		0.56***	- 0.45***
Stock of publications between 1992 and 1997			
S_I	- 0.29	-	-
Maximum (S_I, S_J)	-	0.52***	-0.29*
Minimum (S_I, S_J)		0.54***	-0.21
Average $(S_I + S_J)/2$		0.60***	-0.33*
Number of couples of researchers $C_{IJ} = N_I * N_J$	- 0.39	0.49***	- 0.32*
Productivity			
P_I	0.62***	-	-
Maximum (P_I, P_J)	-	0.11	0.67***
Minimum (P_I, P_J)		0.26***	0.47**
Average $(P_I + P_J)/2$		0.19***	0.69***
Quality of Publications			
Q_I	- 0.11	-	-
Maximum (Q_I, Q_J)	-	0.03	0.09
Minimum (Q_I, Q_J)		0.23***	0.03
Average $(Q_I + Q_J)/2$		0.16*	0.08
International Openness			
pe_I	-0.37	-	-
Maximum (pe_I, pe_J)	-	0.14*	-0.26
Minimum (pe_I, pe_J)	-	0.34***	-0.12
Average $(pe_I + pe_J)/2$	-	0.26***	-0.22

The stars ***, ** and * indicate that the correlations are statistically significant at a confidence level of 1%, 5% and 10%, respectively.

Table 10.5. ***Town averages of within- and between- town intensity of co-publication at the laboratory level***

Town	Number of laboratories per town	Intensity of collaboration within-laboratory	Intensity of collaboration between-laboratory-within-town*	Intensity of collaboration between-laboratory-between-town*
Bagneux	1	58.4	-	2.0
Gif sur Yvette	1	9.2	-	0.4
Grenoble	6	19.4	2.7	0.5
Marseille	1	11.6	-	0.1
Meudon	1	30.9	-	0.1
Montpellier	3	28.1	0.0	0.3
Orléans	1	6.4	-	0.0
Orsay	3	33.4	1.4	0.2
Palaiseau	2	11.9	0.0	0.2
Paris	6	34.5	0.2	0.2
Poitiers	1	23.2	-	0.0
Saint Martin d'Hères	2	37.1	0.0	0.4
Strasbourg	1	38.0	-	0.1
Talence	1	15.7	-	0.0
Toulouse	2	25.7	1.5	0.2
Villeneuve d'Ascq	1	98.9	-	0.5
Villeurbanne	1	40.0	-	0.4
Mean*	**2**	**30.7**	**0.8**	**0.3**
Mean**	**2**	**30.7**	**3.2**	**4.0**

* Mean computed on all the laboratories. **Mean computed on partner laboratories only.

Table 10.6 ***Regression results at the laboratory level on the occurrence of co-publication***

Variables	Occurrence of co-publication between laboratories: Within-town (N=39)	Between-town (N=522)
Geographic Distance	-	-0.008** (0.004)
Distance in Specialization	0.02 (0.04)	-0.009* (0.005)
Size	0.04*** (0.01)	0.015*** (0.001)
Productivity	0.03* (0.02)	0.013*** (0.003)
Quality of Publications	-0.18*** (0.02)	-0.017 (0.029)
International Openness	-0.08 (1.5)	0.17 (0.22)
Adjusted R2	0.374	0.107

The standard errors of the estimated coefficients are given in parentheses. The stars ***, ** and * indicate that they are statistically significant at a confidence level of 1%, 5% and 10%, respectively.

Table 10.7 ***Regression results at the laboratory level on the intensity of co-publication***

Variables	Intensity within-laboratory (N=34)	Intensity of co-publication between-laboratory: Within-town (N=15)	Between-town (N=41)
Geographic Distance	-	-	-0.02 (0.19)
Distance in Specialization	-	-0.78* (0.43)	-0.41* (0.23)
Size	-1.51*** (0.50)	-0.31 (0.34)	-0.24** (0.09)
Productivity	3.31*** (1.09)	0.32 (0.19)	0.19 (0.12)
Quality of Publications	6.88 (5.41)	-0.83 (2.42)	0.48 (1.82)
International Openness	40.1 (58.5)	33.4 (29.8)	-0.93 (12.1)
Adjusted R2	0.312	0.162	0.369

The standard errors of the estimated coefficients are given in parentheses. The stars ***, ** and * indicate that they are statistically significant at a confidence level of 1%, 5% and 10%, respectively.

Figure 10.1. ***Choosing the sample***

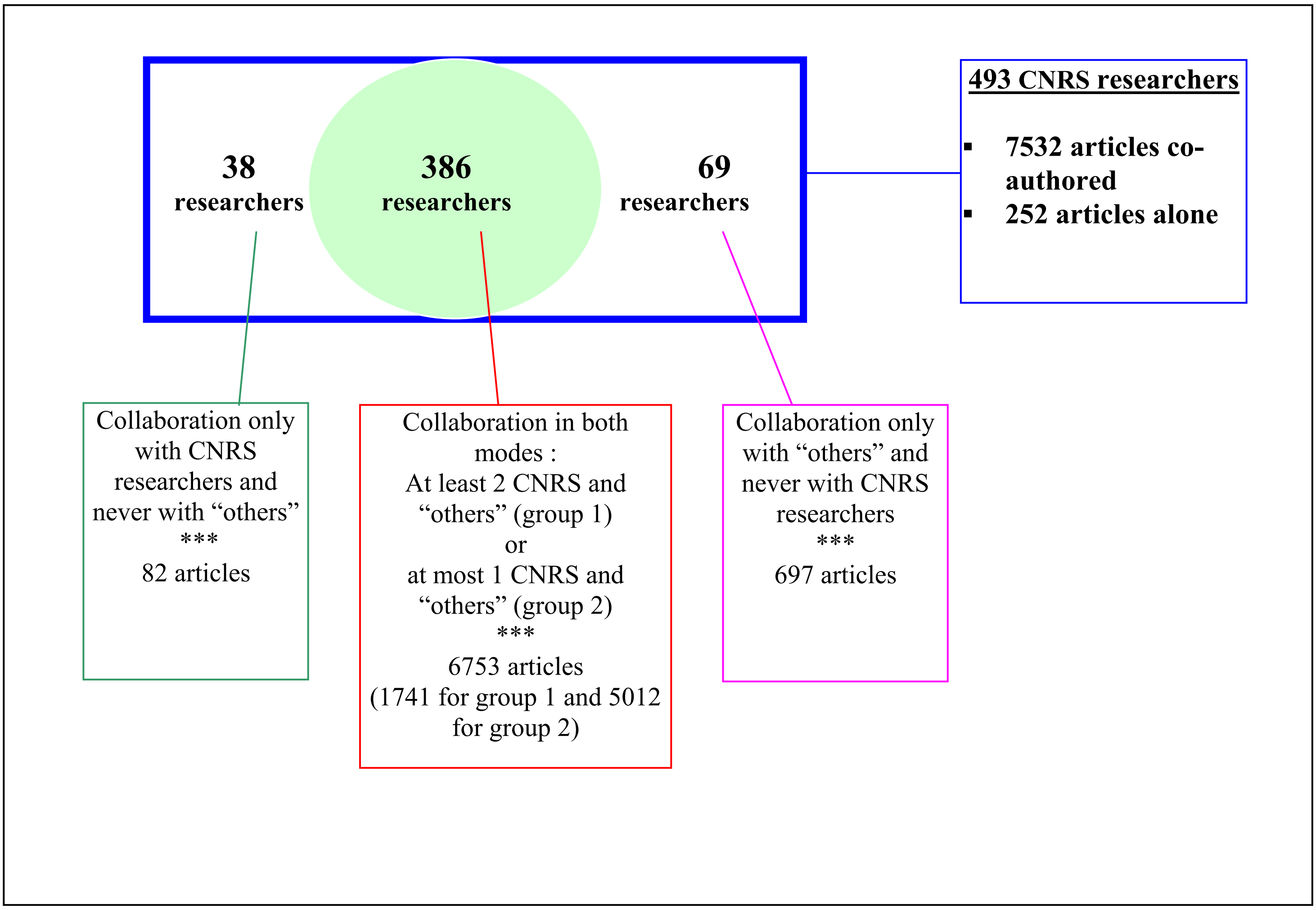


Figure10.2. ***Frequency of the number of articles written by CNRS researchers with external co-authors in Groups 1 and 2 of articles***

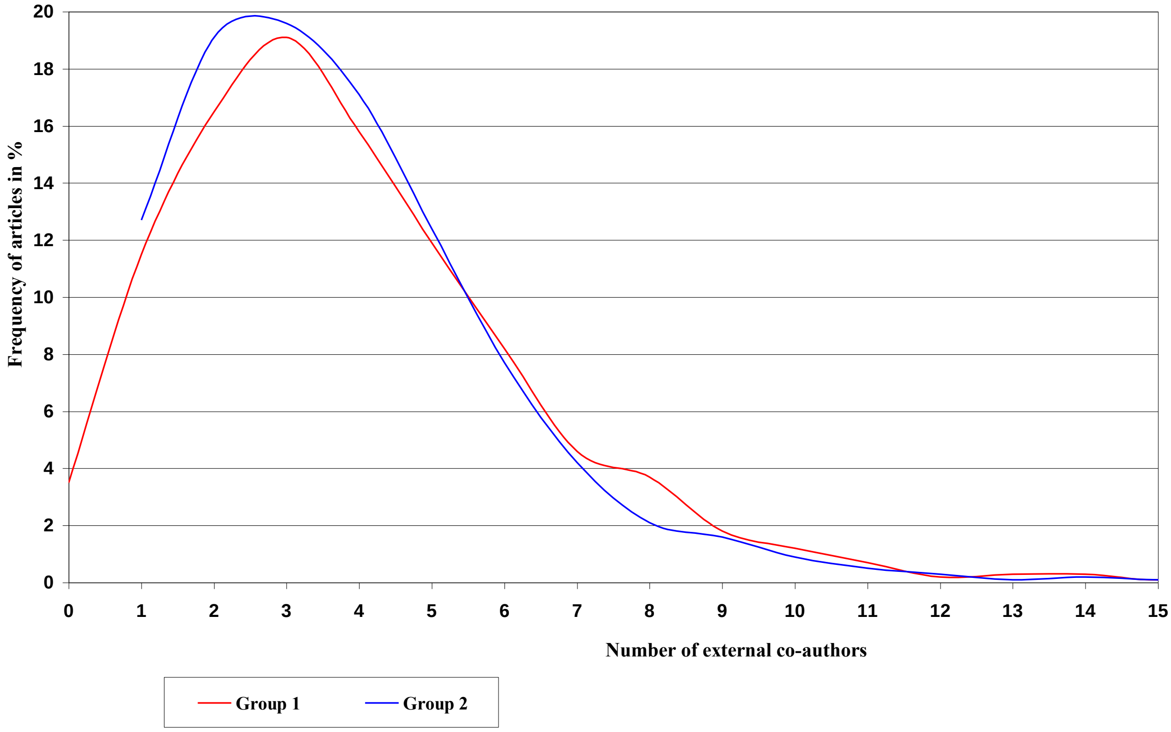


Figure 10.3. ***Concentration curves of the number of articles written by CNRS researchers in Groups 1 and 2 of articles***

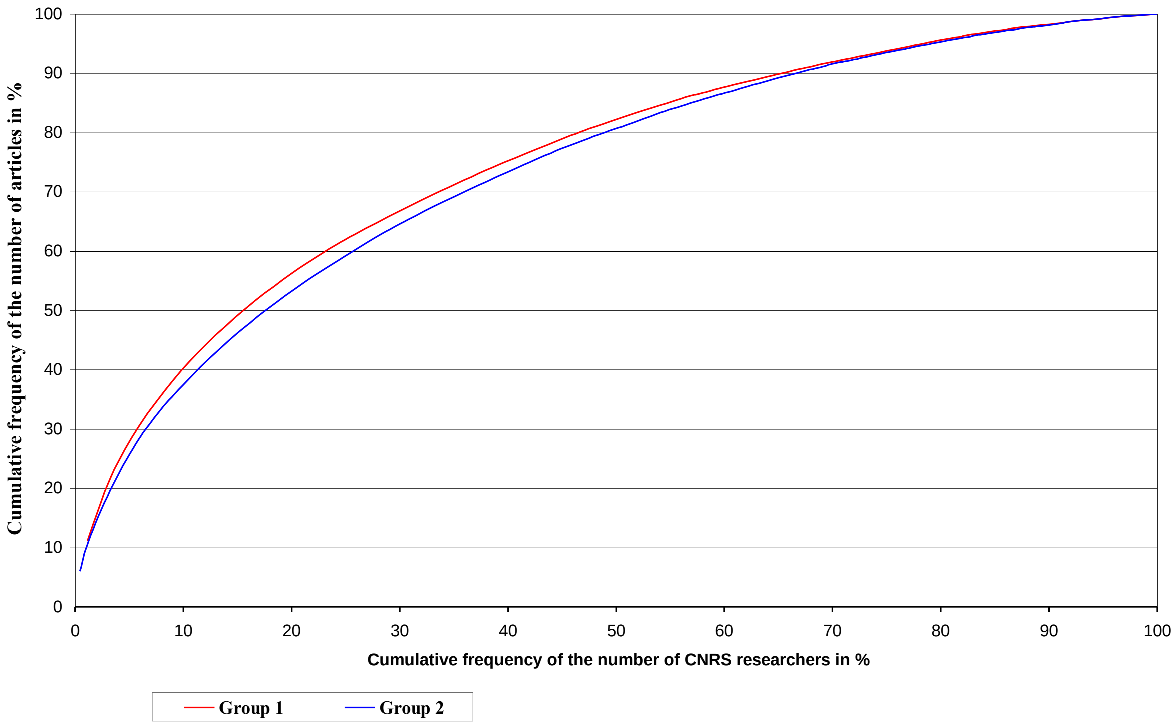
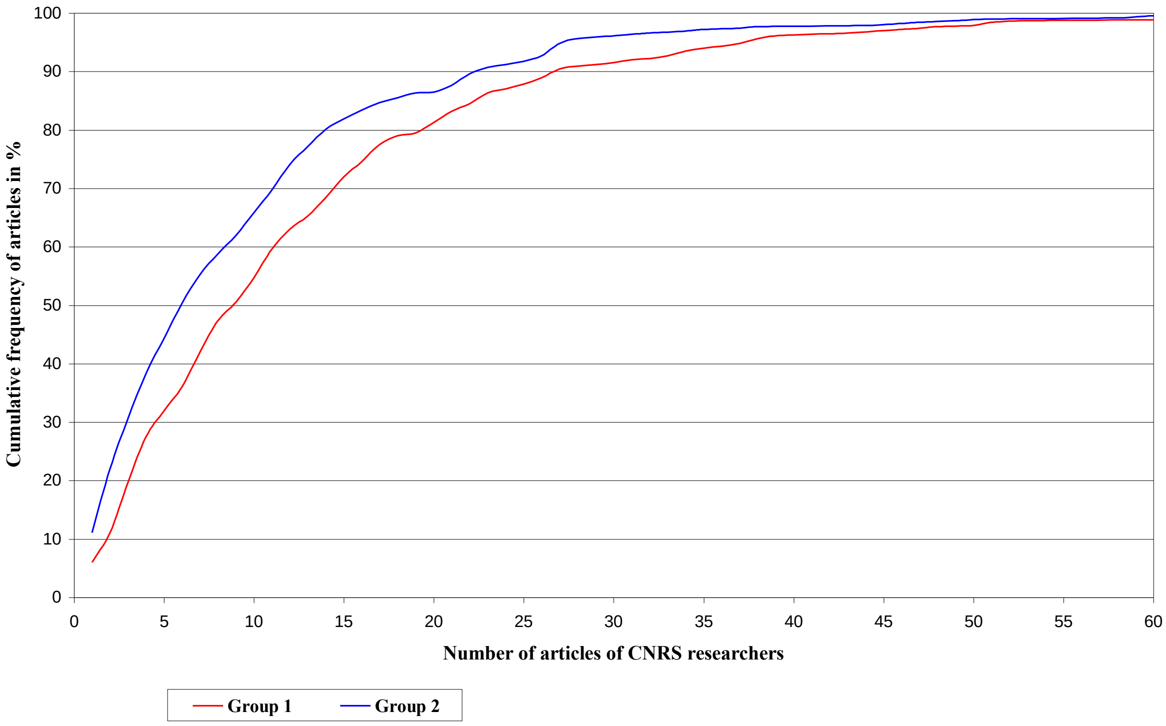


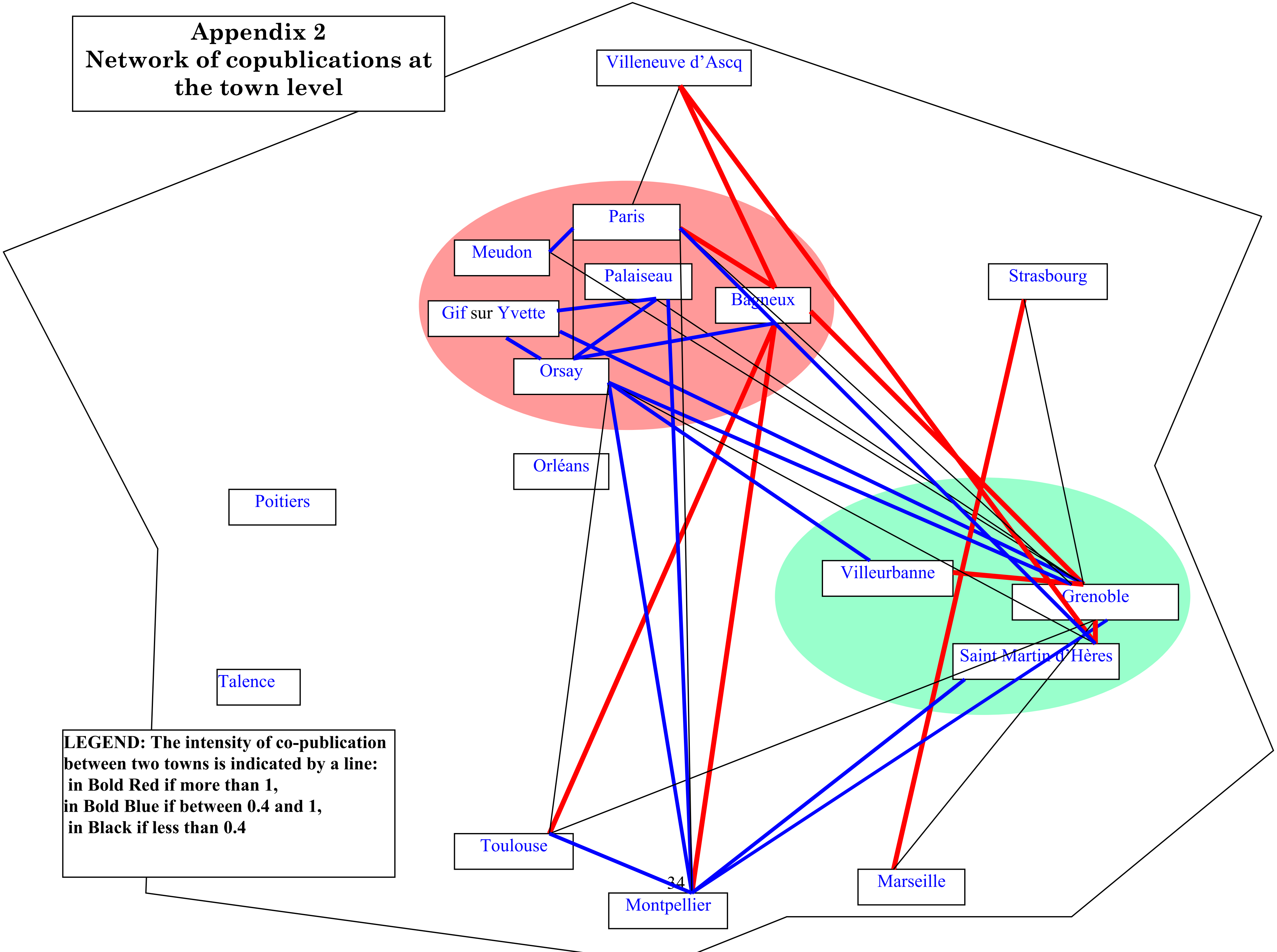
Figure 10.4. ***Distribution of the number of articles written by CNRS researchers in Groups 1 and 2 of articles***



Appendix 1 ***Matrix of intensities of co-publication between CNRS condensed matter physicists at the town level***

	Bagneux	Gif sur Yvette	Grenoble	Marseille	Meudon	Mont-pellier	Orléans	Orsay	Palaiseau	Paris	Poitiers	Saint Martin d'Hères	Stras-bourg	Talence	Toulouse	Ville-neuve d'Ascq	Villeur-banne
Bagneux	62.96		3.53			1.73		0.37		2.57					5.62	7.4	
Gif sur Yvette		4.07	0.53					0.38	1.70								
Grenoble			5.42	0.24	0.24	0.49		0.44	0.28	0.13		1.84	0.36		0.15		2.50
Marseille				9.88									1.41				
Meudon					33.33					0.80							
Montpellier						11.00		0.81	0.99	0.13		0.79			0.38		
Orléans							6.91										
Orsay								3.61	0.52	0.29		0.33			0.35		0.37
Palaiseau									4.36								
Paris										3.03		0.38				0.36	
Poitiers											25.05						
Saint Martin d'Hères												15.39				1.29	
Strasbourg													35.16				
Talence														9.88			
Toulouse															9.63		
Villeneuve d'Ascq																38.52	
Villeurbanne																	43.21

Appendix 2
Network of copublications at the town level



APPENDIX 3- Table1a. ***Correlations at the laboratory level with the occurrence of co-publication***

	Occurrence of co-publications between-laboratory-	
	between-town (N=522)	**within-town (N=39)**
Geographic Distance	-0.09**	-
Distance in Specialization		
Overall Profile	-0.02	0.40***
Physics-Chemistry	-0.05	0.30***
General Physics	-0.09***	0.35***
Solid-state Physics	-0.04	0.24**
Applied Physics	-0.01	0.05
Materials Science	0.05*	0.41***
Crystallography	0.06**	0.10
Other	-0.03	-0.11
Size		
Number of researchers		
N_I	-	-
Maximum (N_I, N_J)	0.21***	0.45***
Minimum (N_I, N_J)	0.22***	0.57***
Average $(N_I + N_J)/2$	0.24***	0.57***
Stock of publications between 1992 and 1997		
S_I	-	-
Maximum (S_I,S_J)	0.23***	0.58***
Minimum (S_I,S_J)	0.35***	0.63***
Average $(S_I + S_J)/2$	0.31***	0.63***
Number of couples of researchers $C_{IJ} = N_I * N_J$	0.27***	0.61***
Productivity		
P_I	-	-
Maximum (P_I,P_J)	0.19***	0.35***
Minimum (P_I,P_J)	0.30***	0.40***
Average $(P_I + P_J)/2$	0.26***	0.39***
Quality of Publications		
Q_I	-	-
Maximum (Q_I,Q_J)	-0.03	-0.46***
Minimum (Q_I,Q_J)	0.06**	-0.30***
Average $(Q_I + Q_J)/2$	0.02	-0.44***
International Openness		
pe_I	-	-
Maximum (pe_I, pe_J)	0.02	0.07
Minimum (pe_I, pe_J)	0.03	-0.06
Average $(pe_I + pe_J)/2$	0.03	0.01

The stars ***, ** and * indicate that the correlations are statistically significant at a confidence level of 1%, 5% and 10%, respectively.
Correlations shown are significant at the 10% level (*), 5% level (**), and 1% level (***) respectively.

APPENDIX 3- Table1b. ***Correlations at the laboratory level with the intensity of co-publication***

	Intensity within-laboratory (N=34)	Intensity of co-publication between-laboratory	
		between towns (N=41)	within towns (N=15)
Geographic Distance	-	-0.06	-
Distance in Specialization			
Overall Profile		-0.32***	-0.42**
Physics-Chemistry		-0.18	-0.10
General Physics		-0.29***	-0.30
Solid-state Physics		-0.25**	-0.05
Applied Physics		-0.05	-0.26
Materials Science		-0.21*	-0.35*
Crystallography		-0.26**	-0.16
Other		-0.11	-0.26
Size			
Number of researchers			
N_I	- 0.42**	-	-
Maximum (N_I, N_J)		-0.52***	-0.34*
Minimum (N_I, N_J)		-0.51***	-0.02
Average $(N_I + N_J)/2$		-0.60***	-0.27
Stock of publications between 1992 and 1997			
S_I	- 0.06	-	-
Maximum (S_I,S_J)	-	-0.08	0.18
Minimum (S_I,S_J)		-0.14	0.13
Average $(S_I + S_J)/2$		-0.12	0.17
Number of couples of researchers $C_{IJ} = N_I * N_J$	- 0.34**	-0.49***	-0.23
Productivity			
P_I	0.36**	-	-
Maximum (P_I,P_J)	-	0.51***	0.33*
Minimum (P_I,P_J)		0.42***	0.36**
Average $(P_I + P_J)/2$		0.54***	0.36**
Quality of Publications			
Q_I	0.20	-	-
Maximum (Q_I,Q_J)	-	-0.03	-0.25
Minimum (Q_I,Q_J)		0.12	0.11
Average $(Q_I + Q_J)/2$		0.05	-0.01
International Openness			
pe_I	0.14	-	-
Maximum (pe_I, pe_J)	-	0.34***	-0.005
Minimum (pe_I, pe_J)	-	0.17	0.35*
Average $(pe_I + pe_J)/2$	-	0.29***	0.22

The stars ***, ** and * indicate that the correlations are statistically significant at a confidence level of 1%, 5% and 10%, respectively.

[1] For a simulation analysis of this process in the institutional context of the US, see David (1994), and for a first attempt of an econometric analysis on the same data as one used in the present work, see Turner and Mairesse (2002).

[2] Three of these studies can be mentioned here. Jaffe (1989) shows that there is in the U. S. a close relationship at the state level between the number of patents and the importance of university research, which he interprets as evidence of geographic externalities. Jaffe, Trajtenberg and Henderson (1993) investigate the localization of knowledge externalities using patent citation data. The authors show that citing and cited patents belong to the same geographic region with a very high probability. Jaffe and Trajtenberg (1998), also on the basis of patent citation data, study the localization of flows of knowledge at an international scale. They find that patents cite much more frequently patents whose inventors live in the same country than patents whose inventors live in different countries.

[3] For example, by adopting a definition of a network as a "clique" in the sense of the theory of graphs (i.e.; a set of points which are connected or such that the intensity of their interconnections exceeds a certain threshold), Blau (1973) makes the following observations for a group of 411 physicists. Members of large networks are often young, work in new and innovative specialties, have a teaching post and are relatively well-known; by contrast, members of small networks are older, work in established specialties, in prestigious university departments, and are involved in administration. These findings seem to reflect the existence of a cycle in research careers, leading the most productive scientists to be also part of the administrative elite.

[4] The analysis by Cole and Cole (1973) on stratification in science is typical of this approach. The authors classify physicists in terms of different criteria such as age, prestige within university departments, productivity and scientific awards. They then measure the impact of these characteristics on the researchers' ranking in terms of scientific reputation and visibility. They finally use their results in an attempt to assess the existence and the extent of discrimination possibly arising from differences of race, gender and religion.

[5] In future work, a possibility will be to carry on this research also at the level of individual researchers. In addition to what we already can observe at the laboratory level, this would allow the analysis of the role of productive and well-known "star" scientists in shaping research networks. See for example the work by Crane 1969 and 1972, Crawford1971, Zucker and al. 1998a and 1998b.

[6] The Centre National de la Recherche Scientifique (CNRS) is the main French public organization for basic research. With 25,000 employees (11,000 researchers and 14,000 engineers, technicians and administrative staff) and over 1,200 research and service units (laboratories) throughout the country, the CNRS covers all fields of science. Directly administered by the Ministry responsible for research which is also usually responsible for higher education, the CNRS has very close links with academic research, researchers from the CNRS and from universities often working in the same laboratories.

[7] These criterions are mainly based on two practical considerations: researchers had to be "not too young" so that we had a history of their publications (the youngest researchers born in 1960 had already been publishing for a few years in 1992, when they were 32 years old); and 1997 was the year for which we could know precisely the laboratories in which the researchers were working, when we first started compiling our data base.

[8] The *Science Citation Index* (SCI) is produced by the Institute for Scientific Information (ISI). It encompasses all the ("hard") scientific disciplines and is constructed on the base of a compilation of over 3,200 of the most cited international periodicals. The quality of the data is remarkable and, in particular, the coverage of scientific publications by CNRS units is very satisfactory (UNIPS, 1999). Ninety-five percent of the scientific articles written by the CNRS researchers are in English and these are fully covered by the SCI.

[9] Another solution would be to count each article only one time, by simply weighting them by the inverse of the number of pairs of authors concerned. This point is discussed in sub-section 3.3. It seems that the main results of our analysis would have been qualitatively unchanged.

[10] This information on the individual researchers and their laboratories was provided to us by the Unité des Indicateurs de la Politique Scientifique (UNIPS) of CNRS.

[11] There is no strict rule for correspondence between authors and addresses in the SCI, since the number of authors recorded for a scientific article often differs from the number of addresses listed for them. It is possible that several co-authors have the same address, in which case their address may be listed only once. Or when the collaboration involves different laboratories, the correspondence between authors and addresses is not always clear. Another possibility is that of multiple affiliations, with one co-author mentioning his or her affiliation to two or more laboratories, which may result again in a problem of attribution.

[12] Note that for Group 1 articles this is a weighted average in which each article is counted as many times as the number of pairs of CNRS co-authors. The simple average for Group 1 is of 4.5 (= 1741/386). See sub-section 2.1 and 3.3 and footnote 9.

[13] Note that this matrix is similar to the adjacency matrix used in the graph theory. The coefficients of the adjacency matrix are equal to 1 when there is a link between the entities corresponding to rows and those

corresponding to columns; otherwise it is 0. The adjacency matrix thus characterizes only the occurrence of collaboration between entities but not their intensities.

[14] It is possible of course to proceed otherwise; that is in this example we could have counted the article for one article giving rise to "two- third" of a co-publication between X and Y and "one-third" within Y. But, as we said, since we are interested in the analysis of collaboration relations, we deemed better to consider that the more co-authors, the greater the weight of an article.

[15] Marseille in fact has also relations with Poitiers, Gif-sur-Yvette, Orsay, Toulouse and Villeurbanne, but these are not taken into account because they all involve less than six co-publications (see sub-section 2.4).

[16] As explained in sub-section 2.4, we only estimated the intensity of co-publication between any two towns (or two laboratories) when the actual number of co-publications between the CNRS researchers in these two towns (or two laboratories) was not too small, that it is less than six (or 4) over the six-year period, and set it to zero otherwise.

[17] The average number of links per town with other towns is only of four. Although we do not know of such a result in another study to which we could compare this estimate, it may seem somewhat on the low side for towns with at least 9 CNRS researchers in our sample (and given our adoption of a rather small threshold of at least 6 co-publications over six years for the definition of an actual link between two towns).

[18] Storage rings have become of great importance throughout the world. They are used to curve or oscillate the trajectory of light charged particles (electrons or positrons) that emit "synchrotron radiations". They thus constitute an extraordinary source of radiations of varying wavelengths, especially X-rays. The European ring of the ESRF (European Synchrotron Radiation Facility) is situated at Grenoble and employs as much as 500 persons on a permanent basis. France has two other rings situated at Orsay at the LURE. About thirty outside laboratories collaborate on a permanent basis with the LURE, as do twenty industrial partners, in the field of physics but also chemistry, biology and environmental science, micro-production, lithography and astrophysics. The LURE rings should soon be replaced by the "SOLEIL" ring, which will constitute a source of "super" synchrotron radiation (several thousand times brighter).

[19] The Chi-squared distance between the specialization profile column vectors Π_1 and Π_2 of two laboratories 1 and 2 is thus defined as: $\chi^2 = \sum_i \left[n_1 \frac{(\Pi_{1i} - \hat{\Pi}_i)^2}{\hat{\Pi}_i} + n_2 \frac{(\Pi_{2i} - \hat{\Pi}_i)^2}{\hat{\Pi}_i} \right]$, where Π_{1i} and Π_{1i} denote the coefficients of vectors Π_1 and Π_2 with i varying from 1 to 7 (or 1 to 2) for the overall (specific) profiles, and where n_1 and n_2 are the numbers of publications of the two laboratories, and $\hat{\Pi}$ is the specialization profile column vector of the two laboratories taken together (or weighted average profile). This Chi-squared distance (as shown in Table 10.3) is normalized by dividing it by the 99 percentile value of the Chi-squared statistic of 6 degrees of liberty for the test of equality of the overall specialization profile vectors (of dimension 7) for any two given laboratories or towns.

[20] Note that the stock of publications not only includes the co-publications of Group 1 of our final sample of 470 CNRS researchers, but also their co-publications of Group 2 and their (few) publications alone. Note also that each publication is counted only once, irrespective of the number of co-authors.

[21] Our measure of productivity thus corresponds to all the publications of our 470 CNRS researchers in the period 1992-1997 (see previous footnote).

[22] Note that estimating a generalized Tobit regression model of both the occurrence and intensity of co-publication provide also practically the same picture that the two corresponding separate linear regressions (the estimated correlation between the probit occurrence equation and the linear intensity equation being not statistically different from zero).