Bureau d'économie théorique et appliquée (BETA)
UMR 7522

Documents de travail

« Small Worlds in Networks of Inventors and the Role of Science: An Analysis of France »

Auteurs

Francesco Lissoni, Patrick Llerena, Bulat Sanditov

Document de Travail n° 2011 - **18**

Septembre 2011

Faculté des sciences économiques et de gestion
Pôle européen de gestion et d'économie (PEGE)
61 avenue de la Forêt Noire
F-67085 Strasbourg Cedex

Secétariat du BETA
Géraldine Manderscheidt
Tél. : (33) 03 68 85 20 69
Fax : (33) 03 68 85 20 70
g.manderscheidt@unistra.fr
http://cournot2.u–strasbg.fr/beta

cnrs
Nancy-Université
Université Nancy 2
UNIVERSITÉ DE STRASBOURG

Small Worlds in Networks of Inventors and the Role of Science: An Analysis of France

Francesco Lissoni[1], Patrick Llerena[2], Bulat Sanditov[3]

[1] *DIME – Universita di Brescia & KITeS – Universita Bocconi, Italy.*

[2] *BETA – University of Strasbourg, France.*

[3] *UNU-MERIT – Maastricht University, Netherlands*

Submitted to: *Industry and Innovation* (special issue on networks)

This draft: *13 September 2011*

Abstract

Using data on patent applications at European Patent Office, we examine the structural properties of networks of inventors in France in different technologies, and how they depend from the inventive activity of scientists from universities and public research organizations (PROs). We revisit earlier findings on small world properties of social networks of inventors, and propose more rigorous tests of such hypothesis. We find that academic and PRO inventors contribute significantly to patenting in science-based fields. Such contribution is decisive for the emergence of small world properties.

Keywords: networks, inventors, academic patenting, small world

JEL codes: D35, O31, O34

Acknowledgements: Successive drafts of this paper have benefitted from comments by participants at the following conferences and workshops: EMAEE 2011 (Pisa, February 14-16, 2011), DIME final conference (Maastricht, April 6-8, 2011), 4th ZEW Conference on Economics of Innovation and Patenting (Mannheim, May 19-20, 2011), 3rd APE-INV "Name Game" workshop (Brussels, September 5-6, 2011). Koen Frenken and Valerio Sterzi discussed early versions of the paper and provided key suggestions for its further development. The following sources of financial support are acknowledged: AnCoRA, a project funded by ANR (ANR-06-APPR-003; Llerena and Sanditov); DIME, a network of excellence funded by the 7th FP (n°51 33 96- CIT3; Llerena); APE-INV, a Research Networking Programme funded by the European Science Foundation (Lissoni and Sanditov). Marc Ledoux, former director of Direction de la Politique Industrielle of CNRS, provided access to CNRS data.

1. Introduction

Over the past 10 years, an increasing number of economic and managerial studies on innovation have appeared, which make use of structural analysis of large social networks (surveys by Ter Wal and Boschma, 2009; Steen et al., 2010). All of these studies take advantage of increasing availability of large repositories of archival data which lend themselves to relational analysis, such as patents, scientific publications and firm alliances. At the same time, they all contribute to the investigation of a specific theoretical concept, that of "small world" networks, whose resurrection and systematization also owes greatly to the availability of digitalized relational data, and not just of the bibliometric type (in fact, most recent studies refer to physical networks, such as the Internet or electrical grids, or to business directories, such as those containing information on actors, directors or producers credited for movies or musical shows; see Uzzi et al., 2007; and Schnettler, 2009).

In this paper, we contribute to this literature by proposing a methodology for detecting small world in networks of inventors, as derived from bipartite patent-inventor graphs; by suggesting that inventors' networks in science-based technological fields exhibit small world properties that do not appear in other fields; and by suggesting that scientists from academia and public research organizations (PROs) contribute decisively to the emergence of such properties.

Following Watts and Strogatz (1998) we describe small world networks as intermediate configurations between ordered and random graphs, of which they share, respectively, the properties of high clustering (they appear to be composed of several tightly knitted clusters of nodes, loosely related one to another) and short average distance between nodes (ensured by those few out-of-cluster ties which provide bridges across otherwise unconnected components).

In particular, small world properties of science-based technologies derive from two related, but distinct social properties of scientific research, namely that:

(i) modern scientific research is conducted by teams, and such teams tend to be large;

(ii) scientific research is conducted not only in companies, but also in universities and public research organizations (PROs), which gives academics and PRO staff the opportunity to get directly involved into patenting.

Teamwork explains the emergence of small world properties because it increases the number of ties between actors, and the more so the larger the teams. As for the involvement of academics and PRO researchers into patenting, this contributes to reducing average distances in the network (a crucial small world property) insofar these actors' ties can more easily escape companies' boundaries. While most inventors are patentees' employees who tie up mainly with inventors from the same company, and therefore have many redundant ties, PRO staff and especially academics tend to act as freelance inventors, whose ties provide shortcuts between inventors from different companies.

We exploit a database of over 50000 French inventors, as listed on patent applications at the European Patent Offices with priority dates comprised between 1980 and 2004. Around 4% of these inventors are found to be affiliated either to a university or to CNRS (Centre national de la recherche scientifique), the leading French PRO. Information on the academic status of inventors come from the French section of the KEINS database (Lissoni et al., 2006), while information on CNRS status was collected on purpose for this paper and it allows for the first ever published estimate of this institution's contribution to patenting in France.

We proceed as follows. In section 2, we re-examine earlier findings on the structural properties of networks of inventors, and of the peculiarities of such networks that call for a careful choice of the proper random network against which small worlds properties should be tested. We also discuss the connections between this literature and the literature on teams in science and technology.

In section 3 we describe our dataset, and place special emphasis on the possibility to rank technological fields of patents according to their science intensity, as well as to the relevance of academics' and CNRS staff's contribution to patenting. We also provide key descriptive statistics on French universities' and CNRS' contribution to patenting, and on size of inventor teams.

In section 4 we examine small world properties in network of inventors as bipartite graphs, and show to what extent such properties correlate to science intensity and academic/CNRS presence in technological fields.

In section 5, we test the hypothesis that academic and CNRS inventors play a crucial structural role, especially in fields whose overall science intensity depends crucially upon their involvement in patenting.

Section 6 concludes.

2. Background literature

Data on patents and inventors, like most archival data containing relational information, lends themselves to be transformed into affiliation networks (Burt, 1983). Affiliation networks are bipartite graphs with two types of nodes: actors and events, where actors are linked to events in which they participate. Typical examples are found in corporate interlocks (Koenig and Gogel, 1981), research collaborations (Powell et al., 1996), scientific co-authorship (Newman, 2001), and underwriting syndicates (Baum et al., 2003). In the case of patents the actors are the inventors and the events correspond to patent applications. In social network analysis of inventors, affiliation networks are most often reduced to "one-mode" graphs in which inventors are directly connected, based on the assumption that, during the production of the patent, contributors to the same patent have come to know each other. By assuming further that co-inventors necessarily share some information, one can investigate networks of inventors as vehicles for knowledge diffusion, often measured with patent citations (Breschi and Lissoni, 2005). The two most important applications of this type of analysis concern the geography of innovation (Breschi and Lissoni, 2005 and 2009; Singh, 2005; Cantner and Graf, 2006; Fleming et al., 2007; Lobo and Strumsky, 2008) and the interplay between science and technology. This second line of enquiry is closely related to studies on co-authorship networks in science, from which it draws concepts and tools, as well as comparison. In particular, starting with Balconi et al. (2004), scholars have examined the structural properties of networks of inventors, and their similarities or differences with respect to those of networks of scientific authors. In addition, efforts have been made to identify who among the inventors may be academic scientists, and whether such identity correlates with their position in the network.

With respect to structural properties, much attention has been devoted to testing the existence of "small world" properties [1]. These properties can emerge only in complex graphs composed of (at least) several hundreds nodes and ties. Any such graph is said to exhibit small world properties whenever it is characterized by high *clustering* (which we can summarily define as the probability to find a direct connection between any two nodes which are also indirectly connected via a third node)[2] and short *average distances* between nodes (due to the presence of a small number of connections between otherwise unconnected nodes). One pre-requisite for the emergence of such properties is high connectedness, which can only occur when most nodes exhibit several ties (as opposed to just one or none).

In small networks a relevant share of all nodes are all connected one to another, thus giving origin to a "giant component", namely a subset of interconnected nodes which is much larger than any other

[1] Among geographical studies, Fleming et al. (2007) focus on the correspondence between small world properties of networks of inventors in metropolitan areas, and the latter's innovative performance. For a critical revision and extension, see Breschi and Lenzi (2011). For a general review of the literature on links between small worlds and innovation performance (well beyond application to network of inventors) see Steen et al., 2010.

[2] We define the clustering coefficient of a network as in Newman et al. (2001):

$$C = \frac{3 \times number\ of\ triangles\ on\ the\ graph}{number\ of\ connected\ triples\ of\ vertices}$$

By “triangle” it is meant here a set of three nodes (a “triple”), each of which is connected to both the others; while a “connected triple” is a triple in which at least one node is connected to both the others. The value of ***C*** is comprised between 0 and 1. This definition of clustering coefficient is sometimes referred as “transitivity” (or “transitivity ratio”).

set[3]. The most widely accepted theoretical treatment of small world networks is that proposed by Watts and Strogatz (1998), who describe small world networks as an intermediate configuration between ordered and random graphs, two theoretical configurations that exhibit maximal clustering and minimal average distance between nodes, respectively. In particular, small worlds exhibit average distance very close to that of a random network of equivalent size, but much higher clustering. Theoretical models by Cowan and Jonard (2003, 2004) suggest that small world networks ensure fast knowledge accumulation and high inventive rates.

Small world properties, as predicted by Watts' and Strogatz's theoretical definition, has been found to characterize several scientific networks, as derived from data on co-authorship of scientific publications (Newman, 2001 and 2004). Small words of scientific authors are compatible with classic theories in the sociology of science, which describe scientific communities as composed by a number of 'invisible colleges' (self-referential groups of scientists working in the same research area, possibly along strict disciplinary lines) which avoid mutual isolation thanks to a number of intellectual leaders who belong to several colleges (De Solla Price , 1963). They also allow to generalize early results of applications of network analysis to bibliometric data (Crane, 1972; Mullins et al., 1977; for a survey: Zuccala, 2006) and to rationalize theoretical treatment of scientific knowledge production as a collective effort (Dasgupta and David, 1994).

Networks of scientific authors exhibit small world properties to the extent that research is conducted in teams, so that most scientists are connected one to another by common work on one or more papers. Indeed, the importance of teamwork in science has grown steadily over the past half-century, as it has the average size of teams. As this is true also of teamwork and team size in invention, we may expect small world properties to be detectable also in networks of inventors (on the general importance of teams in science and technology, see Wuchty et al., 2007).

However, scientific disciplines vary in the extent and importance of teamwork, so that scientific authors' networks in some disciplines do not exhibit small world properties, witness the structural differences between sociology and economy, as found respectively by Moody (2004) and Goyal et al. (2006). Technological fields also differ with respect to their reliance on teamwork for invention, as one can see from data reported by Jones (2009; see especially table 6), which may explain why some fields appear to exhibit small world properties, while others do not (as found by Balconi et al.,

[3] Providing technical definitions of measures and concepts used by social network analysis goes beyond the purposes of this summary of the literature. Definitions of measures and concepts we use in our analysis will be provided as they become necessary, especially in section 4, along with their applications or relevant examples. For a non technical and very short introduction, see Borgatti et al. (2009). For slightly more technical treatments, with focus on small worlds, see Wang and Chen (2003) and Uzzi et al. (2007). For a complete, but accessible introduction to the important notion of centrality, see Freeman (1977).

2004). Quite strikingly, teams characterize networks of inventors especially in science-based technological fields, such as Pharmaceutics or Chemicals. This suggests that the emergence of small world properties in networks of inventors may have be associated to the importance of science as a source of innovation.

Whether such importance requires the direct involvement of universities and public research organizations in the invention (and in the patenting of inventions) may depend on whether universities and public research organizations are the dominant or only source of scientific research. This is not always the case. Breschi and Catalini (2010) study the overlapping between networks of scientific authors and networks of inventors in science-based technological fields (both of which are found to exhibit some small world network properties). They find that individuals who are both inventors and authors of scientific papers contribute to connect the two networks: but while in the US many such individuals appear to be industrial researchers working for R&D intensive companies, in Europe they come disproportionately from the academic environment.

Balconi et al. (2004) find that Italian academic inventors (university scientists whose names appear on patents, irrespectively of whether the latter are owned by universities or companies or private individuals) tend to occupy central positions in the networks of inventors, thus contributing to shorten average distances. As their presence is more widespread in science-based fields, this is suggestive of the decisiveness of their contribution in the emergence of small word properties in such fields. Balconi et al. (2004), however do not explain whether centrality of academic inventors is a result of their belonging to large-than-average teams or to their affiliation to patents from several different applicants, which would result in a more diverse set of ties to otherwise unconnected inventors. The weight carried by academic inventors in a number of technical fields is confirmed by an increasing number of studies, not all of which, however, investigate networks of inventors (Iversen et al., 2007; Lissoni et al., 2008 and 2009; for surveys of less recent contributions, see: Geuna and Nesta , 2006; and Verspagen, 2006).

One common drawback of studies of networks of inventors is the lack of adaptation of measures of clustering, centrality, and component's size to the specificities of one-mode graphs derived from affiliation networks. As suggested by Uzzi et al. (2007), standard measures of clustering tend to be inflated by the very high number of instances in which several inventors, for the simple fact of belonging to the same patent team, are all connected one to another. This inflates also the so-called Q statistic, a common indicator of the existence of small world properties (see below), which possibly lead observer to rush too quickly to the conclusion that the observed network exhibit such properties, even when this is not the case. Emphasis on centrality of inventors may also be misplaced, as it is ties between well-connected (but not necessarily super-connected) actors that really matters for shortening average distances. In the remainder of the paper we will attach great importance to the correct measurement of small world properties in networks of inventors, whose provision is indeed one of our objectives. However, for ease of exposition, we find it more convenient to set aside technical details for now and move to the description of our data (next section). Armed with examples from our data, we will come back to discussing measurement issues in section 4 (see especially section 4.2.1).

3. Data

Our data come from the EP-INV database produced at KITeS-Università Bocconi, which contains all patent applications filed at the European Patent Office (EPO) since its opening in 1978, reclassified by applicant and inventor. Originally based upon legal information produced by EPO for patent attorneys, the EP-INV database is now currently updated by reading, cleaning, and matching applicants' and inventors' names and addresses as published in PatStat, the EPO worldwide statistical database[4]. Links to PatStat allow assigning to each patent application its forward and backward citations to prior art (that is, other patents) as well as backward citations to non-patent literature (NPL), of which scientific publications represent a sizable share (Breschi and Catalini, 2010).

The EP-INV methodology of reclassification by inventor is explained in detail by Lissoni et al. (2006) and can be summarized as follows:

- First, names and addresses of inventors are standardized, in order to assign a unique code to all inventors with the same name, surname, and address;
- Second, "similarity scores" are calculated for all pairs of inventors from the same country with the same name and surname, but different addresses;
- Third, a threshold value for the similarity score is identified, over which two inventors in a pair are considered the same individual and assigned the same unique inventor code[5].

An important subset of the EP-INV database is the KEINS database, which identifies "academic inventors" and provides information on their affiliation as well as on their disciplinary field, gender, date of last promotion, and date of birth. Academic inventors are identified by matching inventors' names with names of professors on active service either on 2004 or 2005, depending on the country.

---

[4] EPO Worldwide Patent Statistical Database(http://forums.epo.org/epo-worldwide-patent-statistical-database/; last visited: July, 2011)

[5] In the case of France, the threshold was set at the median value of the similarity score distribution for all French inventors

Matches are then filtered by emailing or phoning the matched academics, asking for confirmation of their inventor status. Lissoni et al. (2006) provide more methodological details.

Accordingly, we refer to any patent application signed by at least one academic inventor as an "academic patent", no matter whether the applicant is a university, a company, a public or private research organizations, a government agency, or an individual (most typically, the inventor herself). That is to say that we are first and foremost interested into the origin of the invention (whether academic or not), rather than to its property. Indeed, over 60% of French academic patents appear to be owned by business companies, and only around 10% by universities, while PROs own around 25% (Lissoni et al., 2008).

For France, the KEINS database contains information on academic inventors from the hard sciences, medicine and engineering, who were active as *maitre de conference* or professors in any university of the country in 2004 and at least one patent signed after 1993 and before 2005. This make the data increasingly prone to return negatively biased estimates of academic patenting activity the farther back we go in time and highly unreliable for years before 1994 or after 2004.

For the purposes of this paper, the French section of the KEINS database was complemented with information on CNRS inventors, also collected following the KEINS methodology. This integration was considered necessary, due to the important role played by PROs in general, and CNRS in particular, in the French science system[6]. Data from CNRS used for name matching include all *chercheurs*, *ingénieurs de recherche*, and technical staff on duty in 2007, in all the hard sciences as well as medical and engineering disciplines. By analogy with the definitions provided above, we will refer to "CNRS inventors" and "CNRS patents", as the logical equivalents of academic inventors and patents (with CNRS-invented patents including both CNRS-owned patents, and patents over inventions by CNRS staff, but owned by other organizations or staff members themselves; more details in Guarisco, 2009; Thibaut, 2009; and Llerena 2010).

Table 1 reports the populations of French patent applications, inventors, academics and CNRS researchers, as from the database described above. Academic inventors are responsible for 2715 patent applications and CNRS researchers for 1715 applications, which amount respectively to 3.45% and 2.18% of all domestic patent applications at EPO. In total, academic and CNRS patents account for 4044 patent applications that is 5.13% of all patent applications filed at EPO by French inventors.

---

[6] Traditionally, the French public research system has assigned a greater role to PROs than to universities. Even after more than a decade of reforms aimed at correcting the imbalance, CNRS employs over 11000 full time researchers vs. around 54000 tenured staff in universities, all of which either with teaching duties only or mixed teaching-research duties (CNRS, 2010; MESR, 2009; Larédo and Mustar, 2001).

Notice that the latter figure is less than the sum of the previous two. This is because several patents are signed jointly by academics and CNRS researchers, so we counted them only once.

Table 1 EPO patents by technology: all, university, CNRS, and academic (university & CNRS)

		INVENTORS						
	TECHNOLOGICAL FIELDS	*All inv. (1)*	*Ac. inv. (2)*	*(2)/(1)*	*CNRS inv (3)*	*(3)/(1)*	*(2)+(3)*	*(2)+(3) / (1)*
1	*Electrical engineering. Electronics*	13610	217	1.59%	123	0.90%	340	2.50%
2	*Instruments (Scientific; Control)*	9714	363	3.74%	178	1.83%	541	5.57%
3	*Chemicals. Materials*	8653	336	3.88%	259	2.99%	595	6.88%
4	*Pharmaceuticals. Biotechnology*	5980	396	6.62%	280	4.68%	676	11.30%
5	*Industrial processes*	8159	153	1.88%	97	1.19%	250	3.06%
6	*Mech. Eng. Machines. Transport*	10386	64	0.62%	22	0.21%	86	0.83%
7	*Consumer goods. Civil eng.*	5158	12	0.23%	5	0.10%	17	0.33%
	ALL TECHNOLOGIES	51403	1201	2.34%	735	1.43%	1936	3.77%
		PATENTS						
		All pat. (4)	*Ac. pat. (5)*	*(5)/(4)*	*CNRS pat (6)*	*(6)/(4)*	*(5)+(6)*	*(5)+(6) / (4)*
1	*Electrical engineering. Electronics*	18237	385	2.11%	167	0.92%	504	2.76%
2	*Instruments (Scientific; Control)*	10164	513	5.05%	210	2.07%	658	6.47%
3	*Chemicals. Materials*	12157	801	6.59%	654	5.38%	1336	10.99%
4	*Pharmaceuticals. Biotechnology*	7346	713	9.71%	523	7.12%	1119	15.23%
5	*Industrial processes*	10043	195	1.94%	126	1.25%	290	2.89%
6	*Mech. Eng. Machines. Transport*	13796	91	0.66%	28	0.20%	113	0.82%
7	*Consumer goods. Civil eng.*	7057	17	0.24%	7	0.10%	24	0.34%
	ALL TECHNOLOGIES	78800	2715	3.45%	1715	2.18%	4044	5.13%

(1) All inventors who have signed at least 1 EPO patent application in 1994-2004 (EP-INV data)
(2) Academics active in 2005, 2005 in a French universities (sciences, medicine, engineering) who have signed at least 1 EPO patent application in 1994-2004 (source: KEINS database)
(3) Researcher active in 2007 in CNRS, in the hard sciences, medicine, and engineering who have signed at least 1 EPO patent application in 1994-2004 (source: elaboration on CNRS and EP-INV data)
(4) Nr of EPO patents signed by at least one inventor with French address with at least one patent in 1994-2004 (source: EP-INV database)
(5) Nr of EPO patents signed by at least one French academic with at least one patent in 1994-2004 (source: KEINS database)
(6) Nr of EPO patents signed by at least one CNRS researcher with at least one patent in 1994-2004 (source: elaboration on CNRS and EP-INV data)
* When computing (5)+(6), care was taken to avoid double counting the patents signed by both academics and CNRS researchers.

Table 1 also provides details by technological fields, which differ by science-intensity, as measured by the share of NPL citations over total backward citations (see figure 1)[7]. We notice some correspondence between the degree of science-intensity and the combined weight of academic and CNRS patents. The latter is well over 10% in the two most science-intensive fields, namely *Biotechnology & Pharmaceuticals* and *Chemicals & Materials*, while it is below 1% for the two least science-intensive ones, *Mechanics* or *Consumer Goods*. Besides, the field of *Scientific and Control Instruments* ranks similarly for science intensity and for weight of academic and CNRS patents

[7] Technological fields are computed on the basis of the original International Patent Classification (IPC) assigned by EPO officers to each patent application (see: http://www.wipo.int//classifications/ipc/en/; last visited: July 12, 2011), and regrouped by OST (2008).

(respectively, fourth and third). Partial exceptions to this correlation patterns are the fields of *Electrical engineering & Electronics* and *Industrial Processes,* the former with a relatively lower presence of academic and CNRS compared to its science intensity, the latter with a relatively high one.

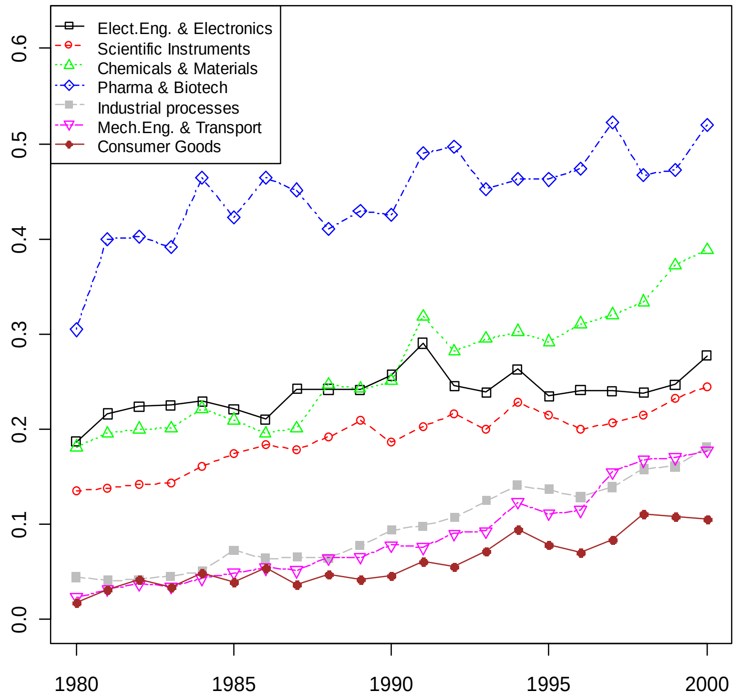

Figure 1 – Science intensity* of technological fields, per year; 1980-2000

* Total citations to non-patent literature (NPL) over total citations

As for the technological distribution of academic and CNRS patents (figure 2) this is also biased towards science-intensive fields, and in line with findings for other countries (as, for example, the US; see Mowery and Sampat, 2005). Academic patents are concentrated in a few technological fields, first and foremost in *Chemicals & Materials*, and *Biotechnology & Pharmaceuticals*, followed by *Scientific and Control Instruments*. A comparison between CNRS and academic patents reveals that the former are relatively more concentrated in *Chemicals & Materials* and *Biotechnology & Pharmaceuticals*, and less in *Electrical engineering & Electronics* and *Instruments*.

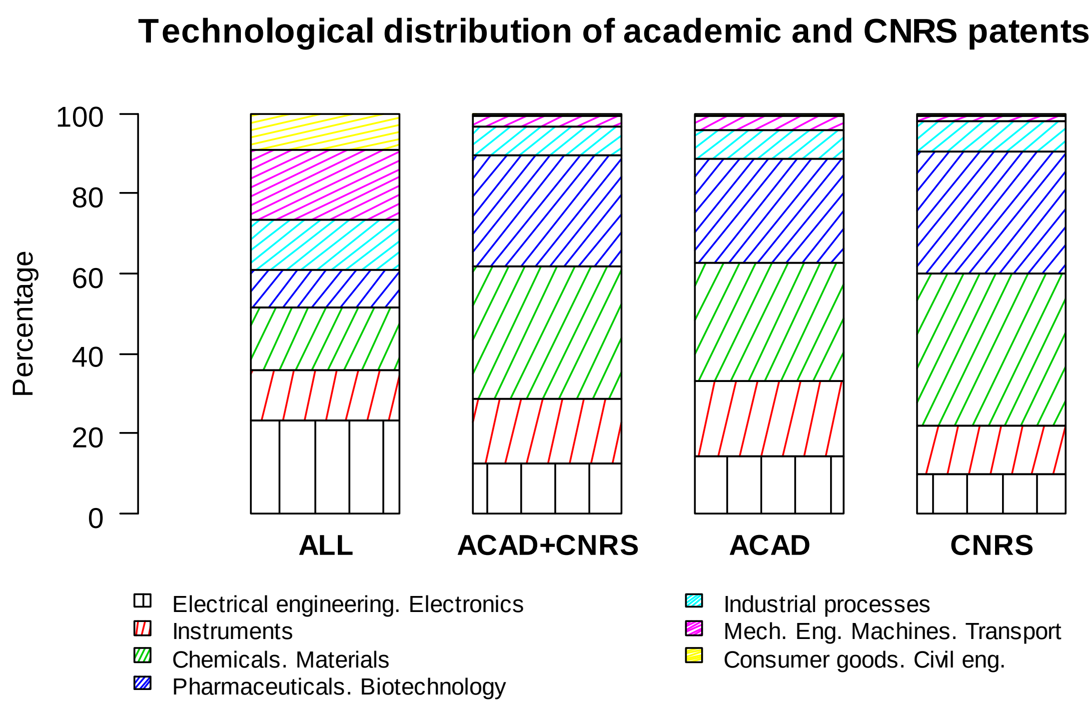

Figure 2 Distribution of patents across technological fields *

"ALL" - all inventors, "ACAD" - academic inventors, "CNRS" - CNRS inventors

In none of the science-intensive technological fields academic and CNRS inventors do not appear to be more productive than other inventors, not even in the technological fields to which they contribute most (table 2). But they appear to work in larger teams than other inventors (table 3a). However, this is in part due to a statistical artefact, which arises from our definition of "academic" or "CNRS" team as any team that includes at least one academic or CNRS inventor. In fact, the probability to have at least one academic or CNRS inventor in a team increases with the size of the team, even if team size were independent on the presence of an academic or CNRS member. This means that even if teams were assembled at random, the expected size of academic teams would be greater than average. The magnitude of the bias is larger, the larger the share of academic inventors in the technological field. In order to correct for this statistical effect we produce a "baseline" distribution of academic and CNRS teams according to the procedure outlined in Appendix A. In table 3b, we compare the average size of academic and CNRS teams derived from such baseline distribution with the observed one, for all type of inventor teams.

We first notice that in *Chemicals & Materials* and *Pharmaceuticals & Biotechnology*, the most science-intensive sectors, the difference between observed and "baseline" averages is less pronounced than in other technologies. We also notice, more importantly, that even after correcting for the statistical artefact explanation, differences in size across academic and CNRS inventors, and generic persist, which requires a substantial explanation.

Table 2 Productivity of inventors, by technological field*

TECHNOLOGICAL FIELD	*All inventors*	*Academic Invs (1)*	*CNRS inventors (2)*	*(1) + (2)*
Electrical engineering. Electronics	3.12	3.25	3.18	3.22
Instruments	3.59	3.47	3.50	3.48
Chemicals. Materials	4.72	4.32	4.36	4.34
Pharmaceuticals. Biotechnology	4.12	3.06	3.14	3.09
Industrial processes	4.00	4.60	5.07	4.78
Mech. Eng. Machines. Transport	3.16	5.89	5.14	5.70
Consumer goods. Civil eng.	3.32	5.92	6.40	6.06
ALL TECHNOLOGIES	2.79	2.84	2.95	2.88

* Only inventors who signed at least one patent between 1994 and 2004 and their patents.

We explain this evidence as follows: the more science-intensive the inventive effort, the larger the team of inventors (possibly due to higher division of the research labour, as suggested by Jones (2009)). The argument is immediately plausible when comparing R&D-based inventive efforts in sectors such as *Pharmaceuticals & Biotechnology* to the solitary trial-and-error toiling of technicians in traditional industries. But also within the same industry, irrespective of their propensity to resort to science as a source for inventions, academics may be involved in projects that require more science inputs, which goes along with having larger teams.

Table 3 Size of inventors' teams and productivity of inventors, by technological field*

	3a - Team size (observed)			
TECHNOLOGICAL FIELD	*All inventors*	*Academic Invs (1)*	*CNRS inventors (2)*	*(1) + (2)*
Electrical engineering. Electronics	2.01	3.01	3.43	3.06
Instruments	2.10	3.45	3.80	3.43
Chemicals. Materials	2.68	3.98	3.84	3.87
Pharmaceuticals. Biotechnology	2.51	3.63	3.98	3.68
Industrial processes	1.87	3.72	3.81	3.65
Mech. Eng. Machines. Transport	1.81	3.05	3.04	3.00
Consumer goods. Civil eng.	1.59	2.41	2.14	2.33
ALL TECHNOLOGIES	2.08	3.59	3.82	3.59
	3b - Team size (baseline)			
Electrical engineering, Electronics	See above	2.72	2.75	2.70
Instruments		2.90	2.96	2.87
Chemicals, Materials		3.44	3.46	3.38
Pharmaceuticals, Biotechnology		3.28	3.33	3.20
Industrial processes		2.53	2.54	2.52
Mech. Eng. Machines. Transport		2.44	2.44	2.43
Consumer goods. Civil eng.		2.08	2.08	2.08
ALL TECHNOLOGIES		2.85	2.87	2.82

* Only inventors who signed at least one patent between 1994 and 2004 and their patents.

This explanation is coherent with other information we can draw from table 3b, namely that:

(a) the average team size (no matter whether the patent is academic, CNRS or neither) is the largest in *Chemicals & Materials* and *Pharmaceuticals & Biotechnology*;

(b) the latter are also the technologies in which differences in size between generic or academic/CNRS teams are the smallest, in percentage terms (around 25%, as opposed to 35% for other technologies): this is because these are the fields in which the science input is always high, so that the inclusion of academics in a team makes less difference than elsewhere.

Our data also indicates that academic and CNRS inventors are more mobile across organizational boundaries when compared to other inventors. Table 4 reports average ratios of the count of distinct applicants (excluding individual applications) to all patents signed by inventors during their research carrier for "serial" inventors, i.e. inventors who have signed more than one patent.[8] Over the course of their research career academic and CNRS inventors appear to work with more applicants (be they business companies, universities or other organizations) than their counterparts in industry, who mostly work and patent for the same employer.

Table 4 Average number of applicants per inventor

TECHNOLOGICAL FIELDS	*All inventors*	*Acad. inventors*	*CNRS inventors*
Electrical engineering. Electronics	1.69	2.75	3.03
Instruments (Scientific; Control)	1.97	3.06	3.01
Chemicals. Materials	2.11	3.23	3.19
Pharmaceuticals. Biotechnology	2.16	3.01	3.04
Industrial processes	1.90	3.84	3.71
Mech. Eng. Machines. Transport	1.66	3.47	3.86
Consumer goods. Civil eng.	1.63	4.17	4.80
ALL TECHNOLOGIES	1.74	2.75	2.89

* Only inventors who were active between 1994 and 2004 and signed more than one patent.

** Not counting individual applicants.

4. Networks of inventors as small worlds

In this section, we first report some basic statistics on the density, cliquishness, and closeness of the French networks of inventors, by technology, which replicate similar exercises by Balconi et al. (2004) for Italy, and are indicative of small world properties, especially for science-based

[8] For an inventor who never signed more than one patent for the same applicant, the number of applicants and the number of patents are the same, so that the number per patent will be equal to 1 (or more, if one or more patents have multiple applicants).

technologies (section 4.1). We then apply two more stringent tests, both specifically adapted to bipartite graphs, which again suggest the existence of small world properties, especially for science-based technologies (section 4.2). All networks were constructed using data on patents from 1980 to 2004. As a consequence, when it comes to estimating the percentage of academic and CNRS inventors in the overall network or one of its components, they tend to underestimate it.

4.1 Basic evidence

Networks of inventors appear to be highly fragmented, as they are composed of a large number of distinct components and isolate nodes (inventors). However, as shown in columns "C1/ALL" and "C1/C2" of table 5, the more science-based the technological field, the larger its main (largest) component, with respect to both the entire network and the second largest component.

The largest components in *Electrical engineering & Electronics* and *Instruments* collect 25% and 28% of all inventors, respectively, with the second largest component being just one or two tenth of it. Largest components are even more noticeable in *Chemicals & Materials* and *Pharmaceuticals & Biotechnology*, where their share of the inventors is well over half the total size of the network, and the second largest follows at a ratio of 1:100. As for *Industrial Processes,* we observe an intermediate situation, with the weight of the principal component in line with what found for *Electrical engineering & Electronics* and *Instruments* (20%), but a first/second component size ratio closer to that of *Chemicals & Materials* and *Pharmaceuticals & Biotechnology* (2:100)[9].

Table 5 also shows that academic and CNRS inventors are more likely to be found in the main components of science-based technological fields, chiefly *Chemicals & Materials* and *Pharmaceuticals & Biotechnology*, with *Instruments, Industrial Processes,* and *Electrical Engineering & Electronics* following in decreasing order. Notice that, across technological field, the main component is the smaller (both in absolute and relative terms) the lower the presence of academic and CNRS scientists among inventors; this is the case of *Mechanical Engineering* and *Consumer Goods*.

Table 6 provides information on the main characteristics of the largest components in our networks. The first column in the table reports the number of inventors in the largest component (*N*). The following three columns report as many distinct graph centralization measures, namely

---

[9] Our values for C1/ALL and C2/C1 ratios are inferior to those found in networks of scientists, as in Newman (2001). But we should keep in mind that our networks are less connected for a number of reasons: taken two comparable populations of inventors and scientific authors active in related technologies and scientific disciplines, the number of patents per inventor is usually lower than the number of papers per author; and the average number of co-inventors in a patent is lower than the average number of co-authors per scientific paper. In addition, our networks spread through technological fields which are much larger and more heterogeneous than Newman's scientific disciplines.

“betweenness”, “degree” and “closeness” centralization index (B_{CENT}, D_{CENT}, and C_{CENT}, respectively), which vary from 0 to 1, and take maximum value in star graphs[10].

Table 5 Networks of inventors: size of the main components, and distribution of academic inventors therein

		Nr. of inventors						**C1/ALL**		**C2/C1**
		ALL [(1)]		**C1** [(2)]		**C2** [(3)]				
		(4)	(5)	(4)	(5)	(4)	(5)	(4)	(5)	
Electrical eng. Electronics	**All inventors,** of which:	**23183**	**13610**	**6459**	**3978**	**928**	**567**	**0.28**	**0.29**	**0.14**
	Academic	*226*	*217*	*99*	*94*	*1*	*1*	*0.44*	*0.43*	
	CNRS	*166*	*123*	*71*	*49*	*0*	*0*	*0.43*	*0.40*	
Instruments	**All inventors,** of which:	**18419**	**9714**	**4542**	**2870**	**128**	**70**	**0.25**	**0.30**	**0.03**
	Academic	*370*	*363*	*149*	*147*	*0*	*0*	*0.40*	*0.40*	
	CNRS	*256*	*178*	*108*	*77*	*0*	*0*	*0.42*	*0.43*	
Chemicals & Materials	**All inventors,** of which:	**15908**	**8653**	**9611**	**5723**	**85**	**48**	**0.60**	**0.66**	**0.01**
	Academic	*345*	*336*	*276*	*268*	*0*	*0*	*0.80*	*0.80*	
	CNRS	*369*	*259*	*298*	*208*	*0*	*0*	*0.81*	*0.80*	
Pharma & Biotechnology	**All inventors,** of which:	**9134**	**5980**	**5213**	**3608**	**28**	**19**	**0.57**	**0.60**	**0.01**
	Academic	*398*	*396*	*232*	*232*	*0*	*0*	*0.58*	*0.59*	
	CNRS	*362*	*280*	*242*	*183*	*0*	*0*	*0.67*	*0.65*	
Industrial processes	**All inventors,** of which:	**15677**	**8159**	**3203**	**2049**	**54**	**18**	**0.20**	**0.25**	**0.02**
	Academic	*155*	*153*	*85*	*84*	*0*	*0*	*0.55*	*0.55*	
	CNRS	*132*	*97*	*86*	*68*	*0*	*0*	*0.65*	*0.70*	
Mech. Eng., Machines, Transport	**All inventors,** of which:	**19869**	**10386**	**1005**	**647**	**881**	**611**	**0.05**	**0.06**	**0.88**
	Academic	*64*	*64*	*2*	*2*	*2*	*2*	*0.03*	*0.03*	
	CNRS	*28*	*22*	*4*	*4*	*2*	*2*	*0.14*	*0.18*	
Consumer goods. Civil eng.	**All inventors,** of which:	**10310**	**5158**	**201**	**150**	**171**	**137**	**0.02**	**0.03**	**0.85**
	Academic	*13*	*12*	*1*	*1*	*1*	*1*	*0.08*	*0.08*	
	CNRS	*9*	*5*	*0*	*0*	*0*	*0*			

(1) Overall network of inventors.
(2) Largest connected component of the inventors’ network.
(3) Second largest component of the inventors’ network.
(4) All inventors (including those who signed no patents after 1993).
(5) Only inventors with at least one patent in 1994-2004.

In general, large social networks based upon archival data, such as ours, exhibit a degree centralization index very close to zero: no single node is connected to all others, and most nodes exhibit more than one tie, which suggest a structure very far from that of a star graph. The main

[10] A star graph of *n* nodes is made of one central node that is connected to all the other (*n-1*) nodes, none of which is connected to each other. This implies that reaching a non-central node from another non-central node always requires following a path through the central node. The degree (betweenness or closeness) centralization index of the network is calculated as the ratio between the average degree (betweenness or closeness) centrality of its nodes and the average degree (betweenness or closeness) centrality for a star graph network of the same size. The degree centrality of any node *j* is simply the number of ties pointing to the node. The betweenness centrality of node *j* is calculated as the share of shortest paths linking any two other nodes, which cross node *j* (node *j* stands “in between” the two other nodes) standardized by the number of all pairs of nodes. All network measures in the paper have been calculated using *igraph* package of R (Csardi and Nepusz 2006).

component of our network is no exception: there are no inventors who worked with a very large share of the total population of inventors, but many inventors who worked with more than one partner, either because they have joined different teams throughout their career, or, more commonly, because the only team they joined comprised more than two inventors. As a result, the degree centralization $\mathbf{D_{CENT}}$ is very low (close to zero). The least science-based technologies (*Consumer Goods* and *Mechanical Engineering*) exhibit somehow higher values.

Table 6 Networks of inventors: Properties of the largest components*

TECHNOLOGICAL FIELD	N	B_{CENT}	D_{CENT}	C_{CENT}	S	C	D	L
Electrical engineering. Electronics	6459	0.194	0.009	0.101	0.07%	0.345	35	12.4
Instruments	4542	0.133	0.016	0.091	0.11%	0.546	39	12.3
Chemicals. Materials	9611	0.118	0.018	0.112	0.06%	0.319	31	8.7
Pharmaceuticals. Biotechnology	5213	0.115	0.014	0.120	0.11%	0.390	28	8.8
Industrial processes	3203	0.166	0.016	0.108	0.15%	0.350	35	9.8
Mechanical eng. Machines. Transport	1005	0.482	0.039	0.097	0.50%	0.441	34	10.4
Consumer goods. Civil engineering	201	0.390	0.076	0.181	1.95%	0.306	11	5.3

N = Nr of inventors in the largest component
B_{CENT}= Betweenness centralization index of the largest component (Avg betweenness/Avg betweenness of N-node star graph)
C_{CENT}= Closeness centralization index of the largest component (Avg closeness centrality/Avg closeness centrality of N-node star graph)
D_{CENT}= Degree centralization index of the largest component (Avg degree centrality/Avg degree centrality of N-node star graph)
S = network density (number of links/max. number of links)
C = Clustering coefficient of the largest component (definition in main text)
L = Avg path length of the largest component
D = Diameter of the main component
* All inventors (including those who signed no patents after 1993).

However, we find that the betweenness and closeness centralization indices of our networks ($\mathbf{B_{CENT}}$ and $\mathbf{C_{CENT}}$) are rather high, which suggests the existence of a few inventors who occupy key positions in between other inventors or groups of inventors, without whom the average path lengths would be higher. As we will see in section 5, a disproportionate number of these in-between inventors are indeed academics and CNRS researchers, or their team mates. Notice that the value of $\mathbf{B_{CENT}}$ for the least science-based technologies (*Consumer Goods* and *Mechanical Engineering*) is much higher than that for the other technologies, which suggest a different structure.

The fifth to eight columns of table 6 report the network density (*S*), the clustering coefficient (*C*) the average length (*L*) and the diameter (*D*) of the main components of our networks. The values for such indicators, especially in our science-based fields, have some properties typical of small worlds.

Density *S* is defined as defined as the ratio of observed links to the number of all possible links between the network's nodes. The clustering coefficient *C* can be interpreted as the probability of two randomly chosen inventors to be co-inventors conditional on that they have (at least) one co-inventor in common. The average path length *L*, and the diameter *D*, are both related to the notion of shortest path ("geodesic"), as well as to the size of the network: *L* is the length of the shortest path between two nodes on the graph, averaged over all pair of nodes in the graph; *D* is the maximum of all geodesics of the graph (Wasserman and Faust, 1994).

These indicators suggest that our networks are quite sparse, as it indeed it is the case with most large scale networks: density *S* for the entire network (not reported) is very low, lying as it does below 0.05% (the densest network is that of *Pharmaceuticals & Biotechnology*, which has a density of only 0.045%). Even largest components, which are generally denser than overall networks, have density limited to 0.15%, as shown in table 6. The only exceptions are the fields of *Consumer Goods* & *Civil Engineering*, with density close to 2% (and a very limited size), and *Mechanical Engineering, Machines & Transport*, with density over 0.5%; however, such density measures apply to very small sub-graphs

Second, clustering coefficients (*C*) appear to be very high. In particular, short-hand calculations suggest that the values reported in table 5 for *Chemicals & Materials* and for *Pharmaceuticals & Biotechnology*, are more than 1000 times larger than those we would find in a the theoretical random graph with the same number of nodes and ties.

Third, the average path length (*L*) is rather short, as it never exceeds 13 steps and it is around 8 steps for the more science-based technologies. Similarly, the diameter is also rather short.

4.2 Small world properties of inventor networks

The evidence we produced so far suggest that networks of inventors, in five out of seven technological fields (namely: *Electrical engineering & Electronics*, *Instruments*, *Industrial processes* and, especially, *Chemicals & Materials* and *Pharmaceuticals & Biotechnology*) exhibit some small world properties. These are the same fields where the presence of academic and CNRS inventors is most noticeable, even more so in the principal component. To investigate further in this direction, we proceed to a more thorough analysis of such small world properties. First, we produce a small-world ratio measure, specifically adapted to "affiliation" networks such as networks of inventors. Second, we perform a "rewiring-based" test of small-world's behaviour.

4.2.1 A inventor-network-specific small world ratio

We follow Watts and Strogatz (1998) in investigating small world properties of our networks by comparing the observed values of clustering coefficients and average path lengths to the values they would take in a benchmark random network (BRN), that is a graph with the same number of nodes and ties, where ties are distributed randomly. A synthetic way to conduct this comparison is proposed by Davis et al. (2003), who suggest calculating the following "small world ratio" Q:

$$Q = (C_{obs}/C_{BRN})/(L_{obs}/L_{BRN})$$

where C_{obs} and L_{obs} are respectively the clustering coefficient and the average path length in the observed network, and C_{BRN} and L_{BRN} the values of the same indicators in the benchmark random network. High values of Q suggest that the observed network has a higher clustering coefficient, but a similar average path length of a comparable random graph; as such they are indicative of small world qualities in the observed network.

The BRN used to calculate Q has to be generated by a stochastic process compatible with the nature of the data at hand [11]. Accordingly, we follow Molloy and Reed (1995) in building a random two-mode inventor-patent network, which preserves the same patents-per-inventor and inventors-per-patent ratios of the original network; and then project it onto the set of inventors, thus obtaining a one-mode network to be compared with the observed one. [12]

Table 7 reports the structural characteristics of the simulated BRN averaged over 100 simulation runs (second line, in italics).

First, we notice a wide difference between the sizes of the first two largest components (columns **C1** and **C2**) in the observed network and the BRN. The random matching of inventors and research teams does not respect the existing boundaries between organizations, localities and technological niches typical of the real world, so that the level of connectedness in a BRN is much higher than that observed in real networks. This is also reflected in BRN's lower values for C2 and much higher C1/C2 ratios.

Second, connectivity of networks is achieved through most productive inventors, who participate in many patents. This effect, which is also present in observed networks, is disproportionally inflated in the BRN. While in the real world a prolific inventor tends to move little in space and across firms or

---

[11] Our choice contrasts with the more common choice to found in the literature, where the "corresponding random graph" is usually the Erdos-Renyi (ER) one. In the ER model the random process is such that a tie between any pair of nodes is generated with equal probability, independently of the existence of other ties. The probability of a tie's existence is simply equal the network density. However, a difficulty arises when we wish to produce ER random graphs for networks of inventors, due to the latter's origin in a inventor-patent affiliation network (Uzzi et al., 2007). As in all affiliation network, connections are carried by events (in our case, the patents), so they come in a bunch. As a consequence, we cannot 'rewire' network ties one by one: inventors, in fact, establish contacts with whole teams of other co-inventors (as listed on a patent), hence rearranging each time a whole set of contacts with all co-inventors in the corresponding patent teams. This will necessarily result in high clustering coefficients, which in turn drive up artificially the small world ratio Q. It follows that the standard ER random graph, often used as a benchmark in studying small-world network structures, is not, the appropriate model for networks of inventors (and, in general, any other network constructed by projecting a bipartite network).

[12] In network literature Molloy-Reed algorithm in which random graphs are generated for *a given sequence of nodal degrees* is referred as "configuration model". Newman et al. (2001) study somewhat different family of random graphs, random networks from a *given distribution of degree sequences*. Both models converge to the same result for networks as large as the ones discussed in this paper. See also Robins and Alexander (2004) and Kogut and Belinky (2008)

technologies, in the BRN the probability that the same inventors will engage in repeated collaborations is negligibly small. This explains why the degree centralization of the main component (**D_{CENT}**), which we interpret as a measure of variation in the number of an individual's collaborators, is somewhat higher in BRNs than in the observed networks. BRNs' larger main components and shorter social distances also result in BRNs exhibiting a higher centralization of the main component (**C_{CENT}**).

Indicators in table 7 reinforce our earlier statement about the presence of central inventors: observed networks have significantly higher betweenness centralization (**B_{CENT}**) than BRNs, which implies that some inventors are significantly more important for network connectivity than others.

Table 7 Observed inventors' network vs. simulated random graph*

TECHNOLOGICAL FIELDS		**C1**	**C2**	**B_{cent}**	**D_{cent}**	**C_{cent}**	**C**	**D**	**L**	**Q**
Electrical engineering. Electronics	observed	6459	928	0.194	0.009	0.101	0.345	35	12.4	0.6
	simulated	*16922*	*7.9*	*0.068*	*0.011*	*0.193*	*0.262*	*14.7*	*5.5*	
Instruments	observed	4542	128	0.133	0.016	0.091	0.546	39	12.3	1.1
	simulated	*12955*	*8.3*	*0.089*	*0.015*	*0.206*	*0.216*	*14.7*	*5.4*	
Chemicals. Materials	observed	9611	85	0.118	0.018	0.112	0.319	31	8.7	1.6
	simulated	*13784*	*4.7*	*0.038*	*0.018*	*0.225*	*0.096*	*10.5*	*4.2*	
Pharmaceuticals. Biotechnology	observed	5213	28	0.115	0.014	0.120	0.390	28	8.8	1.8
	simulated	*7789*	*4.8*	*0.063*	*0.034*	*0.257*	*0.101*	*10.1*	*4.0*	
Industrial processes	observed	3203	54	0.166	0.016	0.108	0.350	35	9.8	1.4
	simulated	*10232*	*6.9*	*0.075*	*0.016*	*0.210*	*0.124*	*13.3*	*5.0*	
Mechanical eng. Machines. Transport	observed	1005	881	0.482	0.039	0.097	0.441	34	10.4	1.4
	simulated	12147	9.5	0.081	0.010	0.181	0.174	16.2	5.9	
Consumer goods. Civil engineering	observed	201	171	0.390	0.076	0.181	0.306	11	5.3	2.2
	simulated	5039	9.1	0.097	0.016	0.187	0.147	15.1	5.6	

C1 = Nr of inventors in the largest component
C2 = Nr of inventors in the second largest component
B_{CENT}= Betweenness centralization index of the largest component (Avg betweenness/Avg betweenness of N-node star graph)
C_{CENT}= Closeness centralization index of the largest component (Avg closeness centrality/Avg closeness centrality of N-node star graph)
D_{CENT}= Degree centralization index of the largest component (Avg degree centrality/Avg degree centrality of N-node star graph)
C = Clustering coefficient of the largest component (definition in main text)
L = Avg path length of the largest component
D = Diameter of the main component
Q = Small worlds ratio (definition in main text)
* All inventors (including those who signed no patents after 1993).

Turning to the last three columns of table 7 we first notice that the small world ratio (column **Q**) ranges from 0.6 in *Electrical engineering & Electronics* to 1.8 in *Pharmaceuticals & Biotechnology*. This suggests that the network of inventors in *Electrical engineering & Electronics* is not a small word at all (because for small worlds *Q* must be significantly larger than 1). Still, it does not provide any conclusive evidence on whether networks of inventors in the other technologies have small-world structures. This is because we do not know, a priori, how big Q should be "to qualify" a network as a small world one, nor we can be sure that Q , as calculated above, works well for networks, such as ours, which are projections of affiliation networks[13].

4.2.2 A "rewiring-based" test

In order to dispel ambiguities in the interpretation of Q values, we introduce an extension of the Watts-Strogatz model to one-mode projections of bipartite networks. In the same way we constructed the BRN above, we now introduce randomness into the observed networks. Starting from the original bipartite network of inventors-patents, we randomly 'rewire' existing ties between inventor and patent. We run a number of "rewiring" exercises, each one with a higher rewiring probability. As in the Watts-Strogatz model, the higher the rewiring probability the closer we move to the corresponding random graph (in which both the clustering coefficient and the average path length are very low).

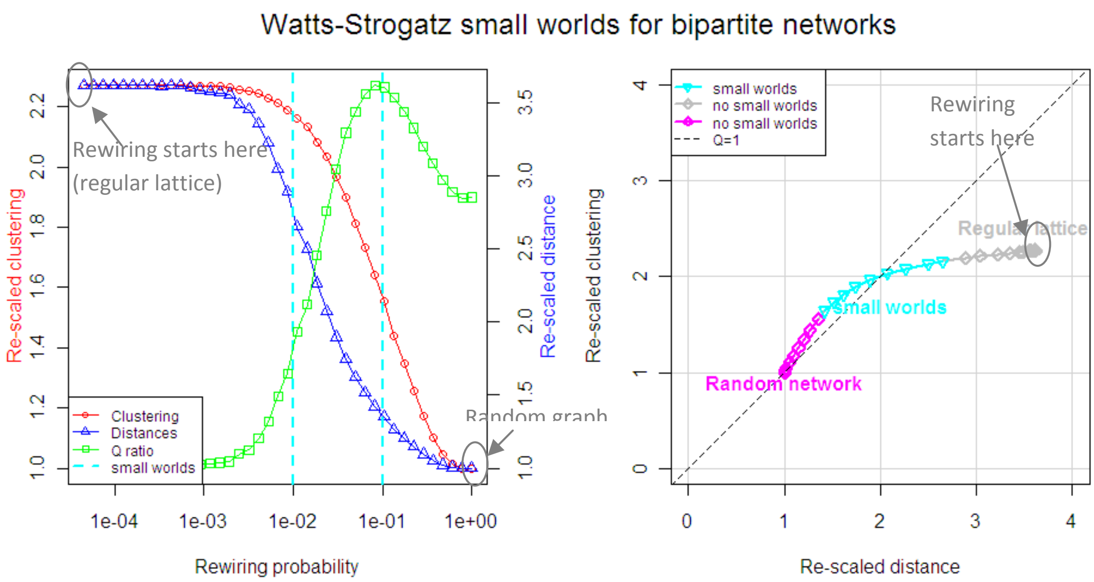

Figure 1 Small worlds networks (model).

Left: Re-scaled clustering (C/C_{BRN}), re-scaled distances (L/L_{BRN}), and small world ratio as functions of the probability of rewiring.
Right: Same results in the rescaled distance – rescaled clustering coordinates; as the probability of rewiring increases we move from the original regular lattice in the North-East to completely random BRN in (1,1).

---

[13] We also notice that the network with the highest value of Q is the one in *Consumer goods*. The primary reason for that is the extremely small size of the main component of the observed network. This is 25 times less than the one of the corresponding BRN, which makes the ratio (L_{obs}/L_{BRN}) relatively small and Q high. However, as we shall see in the following section, despite such a high Q value, networks in *Consumer goods* not are not a small-worlds.

The left panel of figure 3 shows a typical relationship between rewiring probability, clustering, and the average path length in the main component for an idealized network with 100 inventors collaborating on 100 patents, where each patent is produced by a team of 5 inventors and two consequent patents overlap by 4 inventors, so that corresponding one-mode projection is a regular lattice (ring structure) as in Watts and Strogatz (1998). The diagram to the left closely replicates the figure in the original work of Watts-Strogatz (1998), which shows that small worlds emerge when moving, through increased re-wiring probability, from a regular lattice towards a random graph. The dotted lines in the diagram represent values for average path length re-scaled by the values of the corresponding BRN (L/L_{BRN}) and re-scaled clustering coefficient (C/C_{BRN}), and the resulting small world ratio Q for networks obtained by re-wiring the initial regular lattice with an increasing re-wiring probability. We notice that, as the re-wiring probability increases, the lines for L and C initially diverge, as the former drops quickly to the level of the random network, while the latter follows with some delay. As a result "re-wired" networks exhibit small world properties when the rewiring probability is in the range between 0.01 and 0.1.

The right panel of figure 3 presents the same results by means of a diagram whose coordinates measure L and C (scaled by average path length and average clustering coefficient of BRN), and the dotted line reports the values obtained for such indicators when moving from a regular lattice for values of the rewiring probability which increase by moving from right to left. Notice that each rewiring exercise alters the structure of the network: we start from the regular lattice (ring) characterized by high clustering and large average path lengths, but small world ratio $Q<1$, in the North-East corner of the diagram; as the probability of rewiring increases, both C and L decline but at first C declines less rapidly than L. As a result, we reach the South-West corner and the structure typical of a random graph [which is a point with coordinates (1,1) on distance-clustering plain] and, most importantly, we reach it by moving first to the right and then top-down. The interval of rewiring probability over which the re-wired network exhibits small world properties (in blue) correspond to the portion of the dotted line across the bisector, with grade less than one. Notice that the bisector is the locus of points where the small world ratio Q is always equal to one, so that $Q > 1$.

This simple exercise suggests that the logic of the Watts-Strogatz model (1998) can be directly extended to our case. We can 'randomize' our observed networks of inventors, progressively and gradually, by re-wiring the original bipartite network with increased probability, as explained above; and then map the results onto a diagram such as the one in the right panel of figure 3.

Since all our networks of inventors with a significant presence of academics and CNRS staff have a small world ratio Q greater than 1 (the only exception being *Electrical Engineering and Electronics*), their corresponding dots in the diagram will appear in the region over the bisector. If we observe that rewiring reduces both C and L, with L declining faster than C, then we will also observe an alignment of the dots from right to left and top-down, i.e. a dotted line steeper than the bisector. In this case, we will conclude that our network of inventors exhibit small world properties, to the extent that, by rewiring it, we observe only minor negative effects on the average path length (which is already low enough to be close to that of a random network), but a much higher (negative) effect on clustering (which is still high enough to be close to that of a regular lattice). By contrast, if the inventor network does not exhibit any clear small world feature, and it is composed of many isolated cliques, rewiring should have either a strong effect on both distance and clustering, or possibly a strong effect on the former and a weaker one on the latter.

The results of our experiments are shown in figure 4. Networks of inventors in *Chemicals & Materials* and *Pharmaceuticals & Biotechnology* are somewhat closer to a random network, while networks in *Scientific & Control Instruments* and *Industrial Processes* are closer to a regular lattice (although not as close as to be located under the bisector). Yet all these networks show a similar pattern: as the rewiring proceeds their dotted lines initially diverge from the bisector, only to drop precipitously towards it, which suggests that they have small world structure. By contrast, the inventor network in *Electrical engineering & Electronics* exhibit a somewhat different pattern, in that rewiring at first increases both distances and clustering in the largest component. As for *Consumer Goods* and *Mechanical Engineering and Transports* their rewired paths are very erratic and far from indicating any small world property.

Notice that in all five technologies, rewiring increases the size of the largest component, as redundant ties (ties between otherwise connected nodes) within dense cliques are replaced by non-redundant ones (which end up linking formerly isolated nodes or cliques to the largest component). In principle, this may increase the average path length L, which we measure only for the largest component, because the more nodes we have in the component, the longer may be the paths connecting them. However, the substitution of redundant ties with non-redundant ones has the opposite effect, as short-cuts are created, which shorten the distance between otherwise far-apart nodes; and this is precisely what we may expect by re-wiring a small world network or a regular lattice (which has the potential – through rewiring – to become a small world). While the second effect prevails in the network for *Instruments*, *Chemicals & Materials*, *Pharma & Biotech*, and *Industrial Processes*, it is the first one that, as the rewiring proceeds, dominates *Electrical Engineering & Electronics* as well as *Consumer Goods* and *Mechanical Engineering & Transports*.

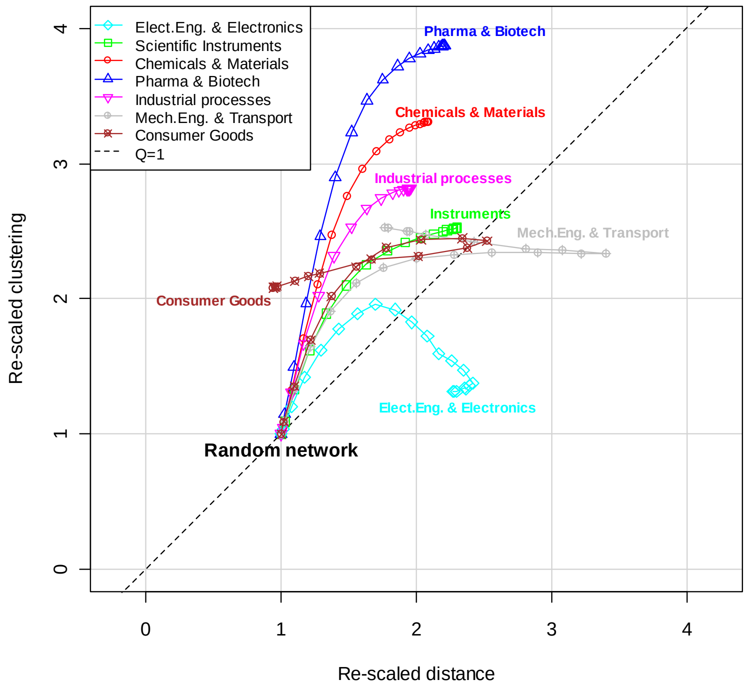

Figure 2 Small worlds in inventors' networks (networks of all inventors, including those who signed no patent after 1993).

5. Academic and CNRS inventors as small world catalysts

In this section, we prove the importance of academic and CNRS inventors in ensuring the small world properties of networks of inventors in all technologies in which their presence is non-negligible, with the exception of *Electrical engineering & Electronics.* In order to do so we first examine these inventors' properties as individual nodes, namely their centrality in the network, as well as the centrality of the teams they belong to (section 5.1).

We then proceed to a more stringent test, which consists in deleting randomly a number of nodes and examining to what extent this weakens the small world properties of the network; we show that this is particularly the case when deleted nodes represent academic or CNRS inventors (section 5.2).

5.1 Individuals' and teams' centrality

Table 8 reports the average betweenness, closeness and degree centrality scores (respectively B_{CENT}, C_{CENT} and D_{CENT}) for academic and CNRS inventors as opposed to all inventors, within the main component of each technological field[14]. Academic and CNRS inventors tend to occupy more central positions than the average inventor, both locally (D_{CENT}) and globally (B_{CENT} and C_{CENT}).

Whatever the technological field considered, D_{CENT} for both academic and CNRS inventors is higher than average which implies that academic and CNRS inventors have a higher-than-average number of co-inventors. As for B_{CENT} and C_{CENT} they are also higher for academic and CNRS inventors than for the all inventors, with the only exception of *Electrical engineering & Electronics* and *Instruments* (for C_{CENT}).

Table 8 Position of academic inventors in the main component*

TECHNOLOGICAL FIELDS		N	B_{CENT}	C_{CENT}	D_{CENT}
	All inv.	**3978**	**0.0024**	**0.0837**	**4.9**
Electrical engineering. Electronics	*Uni inv*	*94*	*0.0027*	*0.0811*	*5.5*
	CNRS inv	*49*	*0.0037*	*0.0856*	*5.5*
	All inv.	**2870**	**0.0034**	**0.0841**	**5.7**
Instruments	*Uni inv*	*147*	*0.0069*	*0.0840*	*6.5*
	CNRS inv	*77*	*0.0039*	*0.0844*	*5.4*
	All inv.	**5723**	**0.0011**	**0.1210**	**7.1**
Chemicals. Materials	*Uni inv*	*268*	*0.0019*	*0.1256*	*8.2*
	CNRS inv	*208*	*0.0019*	*0.1257*	*7.9*
	All inv.	**3608**	**0.0018**	**0.1186**	**6.4**
Pharmaceuticals. Biotechnology	*Uni inv*	*232*	*0.0034*	*0.1216*	*7.0*
	CNRS inv	*183*	*0.0026*	*0.1246*	*7.7*
	All inv.	**2049**	**0.0035**	**0.1098**	**5.6**
Industrial processes	*Uni inv*	*84*	*0.0081*	*0.1146*	*6.8*
	CNRS inv	*68*	*0.0038*	*0.1177*	*6.0*
	All inv.	**647**	**0.0116**	**0.1024**	**5.6**
Mech. engineering. Machines. Transport	*Uni inv*	*2*	*0.0020*	*0.1019*	*3.5*
	CNRS inv	*4*	*0.0049*	*0.0795*	*3.8*
	All inv.	**150**	**0.0238**	**0.1983**	**4.5**
Consumer goods. Civil engineering	*Uni inv*	*1*	*0.0000*	*0.1566*	*1.0*
	CNRS inv	*0*			

B_{CENT}= Avg betweenness centrality of inventors considered
C_{CENT}= Avg closeness centrality of inventors considered
D_{CENT}= Avg degree centrality of inventors considered
* Inventors who signed at least one patent between 1994 and 2004.

Notice than an inventor's higher-than-average centrality may have three different explanations:

- *(a)* a higher-than-average number of patents signed by the inventors (productivity effect);
- *(b)* a higher-than-average number of co-inventors, due to participation to larger-than-average teams of inventors (team effect):

[14] For a definition of D_{CENT} see footnote 4.

(c) a higher-than-average mobility across different invention teams, which implies fewer repeated interactions with the same co-inventors, but a larger number of contacts within the network (mobility effect).

Patent-per-inventor data we presented in section 3 lead us to exclude explanation (a), as we do not find any indication of a higher-than-average productivity of academic or CNRS inventors.

On the contrary, we know for sure that explanation (b) applies, as shown by data on team size, also presented in section 3. Even correcting for team size, however, academic and CNRS inventors appear to be more central than other inventors. This can be seen by calculating the *group degree centrality* scores for inventors' teams normalized by the size of the teams (table 9). Group degree centrality is defined as the number of non-group nodes that are connected to group members (Everett and Borgatti, 1999). In our case, a group is a research team as it appears in the patent document. However, larger teams (such as academic and CNRS ones) may exhibit higher centrality simply because they have many members, which altogether stand greater chances to be connected to non-group members simply as a result of their numerosity.

To correct for this bias we divide the raw centrality scores by the size of the team, so that scores reported in table 9 can be interpreted as the number of connections with inventors outside of the research team per member of the team. The result is then averaged over all academic or CNRS teams and teams which are neither. By comparing column (1) with columns (2) and (3) we notice that, irrespective of the field, academic and CNRS teams are always more central than the other teams, the only exception being *Pharmaceuticals & Biotechnology*, for which differences are barely noticeable.

Table 9 Group degree centrality of research teams normalized by the size of the team

TECHNOLOGICAL FIELD	(1)	(2)	(3)	(2)+(3)
Electrical engineering. Electronics	3.2	4.6	4.4	4.6
Instruments	2.8	3.7	3.8	3.8
Chemicals. Materials	6.2	10.3	7.9	9.4
Pharmaceuticals. Biotechnology	6.6	6.3	7.1	6.8
Industrial processes	2.7	7.2	6.9	7.2
Mechanical eng. Machines. Transport	2.3	4.3	5.1	4.3
Consumer goods. Civil engineering	1.7	3.3	3.7	3.5

(1) Group centrality score averaged over patents with no academic or CNRS inventor.
(2) Group centrality score averaged over patents with at least one academic (university) inventor.
(3) Group centrality score averaged over patents with at least one CNRS inventor.

As for explanation (c), concerning the mobility of academic and CNRS inventors we recall from section 3 that the latter indeed tend to change applicant more frequently that other inventors (see table 4). This is because most inventors are R&D employees of business companies: while working for the same employer, they cannot change team easily, unless the employer is very large and hosts many different teams; not they can change employer very often. Academic and CNRS inventors, on the contrary, behave more like "free lance" inventors, who look for different sources of funding and partnerships, and change inventors' teams accordingly. They may also end up producing some inventions as the result of involvement within large, collaborative research programmes between university or CNRS and companies, which guarantee their participants considerable freedom in picking up collaborators and partners for specific projects.

5.3 A node-deletion test

We have just seen that academic and CNRS inventors occupy central positions in the network of inventors. As such they contribute to keep average distances low, and to connectedness. We now investigate whether they also play a decisive role in the emergence of small world properties. We do so by analyzing consequences their removal from the network would have, on both distance and connectedness.

Two effects are possible. First, in weakly connected network (such as our inventor networks in *Consumer Goods* or *Mechanical Engineering & Transports*) the removal of central nodes may result in carving-out components which are connected to the rest of the network only through the removed nodes[15]. Second, even if a network remains connected, the deletion of central nodes would generally result in increasing social distances, because the removal of a node also removes of all the shortest paths passing through it.

Table 10 reports the results of our experiment, which consists comparing the effects of removing from our networks either the academic and CNRS inventors (with at least one patent after 1993; we refer to these as "case") or a "control" sample of other inventors, chosen at random (more precisely, we repeat the random removal 100 times, and calculate the average effects). The two effects we monitor are:

- the change in size of the largest component (ΔC1), which we expect to be negative; and
- the percentage in the average path length of the same component (ΔL/L), which we expect to be positive.

As for ΔC1, we notice that, in all technologies, the removal of academics and CNRS researchers as well as controls makes network more disconnected, as the reduction in size of the largest component largely exceeds the number of the removed inventors. On the contrary, for “control”

[15] In graph theory connectivity of a network is the size of minimum cut-set, a set of nodes such that removal of these nodes makes the network disconnected.

removals the total size of disconnected components is closer to the number of the removed inventors. We also notice that the difference between “case” and “controls” removals varies across technological fields: from insignificant in *Consumer Goods* and *Mechanical Engineering*, and fairly marginal in *Electrical Engineering & Electronics,* to very high in *Instruments*, where removal of 224 academics and CNRS researchers disconnects 866 additional inventors from the largest component (while the effect of 224 “controls” is 347 disconnected inventors). Nevertheless, in all technological fields (but *Instruments*) the largest components in the resulting networks retain more than 80% of their inventors.

Table 10. Effect of removal of academic and CNRS inventors from inventors' networks on the size of the largest component and average path length therein

TECHNOLOGICAL FIELD		Nr removed	C1	ΔC1	L	ΔL/L
	(1)	143	6459		12.4	
Electrical engineering. Electronics	(2)		6068	391	12.7	2.2%
	(3)		6118	341	12.6	1.7%
	(1)	224	4542		12.3	
Instruments	(2)		3452	1090	12.1	-2.0%
	(3)		3971	571	12.6	2.0%
	(1)	476	9611		8.7	
Chemicals. Materials	(2)		8538	1073	9.6	9.5%
	(3)		8827	784	9.0	2.6%
	(1)	415	5213		8.8	
Pharmaceuticals. Biotechnology	(2)		4247	966	9.5	8.6%
	(3)		4443	770	8.9	1.7%
	(1)	152	3203		9.8	
Industrial processes	(2)		2769	434	11.2	14.3%
	(3)		2882	321	9.9	0.8%
	(1)	6	1005		10.4	
Mech. engineering. Machines. Transport	(2)		994	11	10.4	-0.3%
	(3)		987.6	17	10.3	-0.9%
	(1)	1	201		5.3	
Consumer goods. Civil engineering	(2)		200	1	5.3	-0.2%
	(3)		199.4	0	5.3	0.2%

C1 = Nr of inventors in the largest component
L = Avg path length of the largest component
For each of the technological fields different rows refer to
(1) The largest component in the observed network
(2) The largest component in the network after removal of academic & CNRS inventors from the largest component of the observed network
(3) The largest component in the network after removal of control sample (average over 250 random sample draws) from the largest component of the observed network

As for the effect on ΔL/L we notice that, in *Consumer Goods*, *Mechanical Engineering* and *Electrical Engineering & Electronics*, we see again only marginal difference in the effects of removal of “case” and “controls” on the distances in the network. In *Instruments* the removal of academics and CNRS researchers decreases the average path length, but this due to the fact that, after the removal, the largest component shrinks considerably. In the other three fields the effect of the "case" removal largely exceeds the effect of removing the “controls”. In particularly, in *Industrial processes*, the removal of academics and CNRS researchers increases the average social distance by 14.3%, while the effect of random “controls” on distance is below 1%.

This evidence suggests that, in science-based fields, which we know to be characterized both by the non-negligible presence of academics and CNRS inventors and by small world properties, the academics and CNRS inventors play an important role in the emergence of two small world properties, such as the presence of a large component and short average distances.

6. Conclusions

The study of networks of inventors has attracted a great deal of attention, both theoretical and empirical. Theoretical models such as those by Cowan and Jonard (2003, 2004), which suggest that small world networks maximize knowledge diffusion and innovation rates, have led to many efforts of reconstructing networks of inventors from patent data, and to test for the existence of small world properties. Several studies in this tradition, however, do not take fully into account the methodological problems that arise when traditional measures of clustering and centralization are applied to affiliation networks (as networks of inventors), as opposed to the theoretical one-mode networks they were first thought for (and which have a counterpart in networks based on sociometric data, as in the tradition of social network analysis).

In this paper we have re-elaborated a set of specific tools for testing for the existence of small world properties in networks of inventors, and successfully detected them in science-based technological fields.. In order to do so, we have relied on an extensive data set of patents and inventors in France, which is the first to identify the contribution of both academics and researchers from CNRS (the largest French PRO) to patenting.

We have shown that reliance on science as a source of knowledge for inventive activity implies both to assign a great role to teamwork and to involve a relatively high number of academic and PRO scientists and PRO researchers in the inventive activity. Both effects contribute to the emergence of small world properties, which we find to be the stronger the more science-based the technological field. While the dominance of teamwork contributes to the emergence of small world by increasing the connectivity of inventors, the presence of academic and CNRS inventors ensures the creation of several non-redundant ties, which do not detract to cliquishness in the network, but reduces average distances. This is because academic and CNRS inventors appear to be more mobile across applicants, as they are not bound by employer-employee relationships in their choice of projects and teams. In fact, most of their patents do not belong to their employers (the universities or CNRS) but to companies with which they collaborate, very much like those of freelance inventors. The presence of academics and CNRS inventors also contribute to the emergence of small worlds, due to the larger average size of teams including such inventors. From this finding, an important policy implication may follow: academic and CNRS inventors appear to contribute to inventive activity not only in a direct way, through the patents they produce, but also through their mobility across organizations, which may lead to knowledge diffusion and further inventive activity. This contrasts with the increasing emphasis placed by European governments (including the French one) on measuring the quantity of patents produced by universities, rather than on the collaborative features of the underlying research, which we can detect by analysing the resulting social networks (Della Malva et al., 2011).

Notice that we find no difference between academic and CNRS inventors neither with respect to their characteristics (productivity, team size, mobility across applicants) nor with respect to their network status. This contrasts, at least in principle, with the different status of the groups, the former being at best part time researchers (due to their teaching duties), the latter being full time scientists. We need further investigation on the affiliation of academic inventors, who could indeed come from UMRs, the mixed units of research co-funded by CNRS in several universities and enjoy some reduction in their teaching load due to their high scientific status.

In the influential survey by Uzzi et al. (2007), the production of more accurate analyses of small world properties in affiliation networks was listed as a top priority for the small world research agenda. We have followed that recommendation and plan to follow also the other ones in the near future. In particular, also as recommended by Uzzi et al. (2007), we will investigate the dynamics of network formation, and the contribution by academic and CNRS scientists to such dynamics. We also plan to investigate the consequences of the emergence of small world properties for the speed of knowledge diffusion, as measured by patent citations, whether to other patents or the non-patent literature. Ongoing data collection efforts will also help us extending our analysis to other countries than France.

References

Balconi M., Breschi., Lissoni F. (2004), “Networks of inventors and the role of academia: an exploration of Italian patent data” , *Research Policy* 33(1), pp. 127-14

Baum, J.A.C,, Shipilov A.V., Rowley T.J. (2003), "Where do small worlds come from?" *Industrial and Corporate Change*, 12(4): 697-725.

Borgatti S.P., Mehra A., Brass D.J., Labianca G. (2009), "Network Analysis in the Social Sciences", *Science*, 323 (5916), pp. 892-895

Breschi S., Catalini C. (2010), "Tracing the links between science and technology: An exploratory analysis of scientists’ and inventors’ networks", *Research Policy* 39, pp.14-26

Breschi S., Lenzi C. (2011), " Net and the city. Co-invention networks and the inventive productivity of US cities", *mimeo*

Breschi S., Lissoni F. (2005) “Knowledge networks from patent data: Methodological issues and research targets”, in: Glänzel W., Moed H., Schmoch U. (eds), *Handbook of Quantitative S&T Research*, Kluwer Academic Publishers

Breschi S., Lissoni F. (2009) "Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows," *Journal of Economic Geography*, 9(4), pp. 439-468

Burt R. S. (1983), "Network Data from Archival Records" In: R. S. Burt and M. J. Minor (Eds.). *Applied Network Analysis*. Sage: Beverly Hills

Cantner C., Graf H. (2006), "The network of innovators in Jena: An application of social network analysis", *Research Policy*, 35(4), pp. 463-480

CNRS (2010), *Le CNRS en chiffres*, Centre national de la recherche scientifique, Paris

Cowan R., Jonard N. (2003) “The dynamics of collective invention”, *Journal of Economic Behavior & Organization*, 52(4), pp. 513-532.

Cowan R., Jonard N. (2004) “Network structure and the diffusion of knowledge, *Journal of Economic Dynamics and Control*, 28(8), pp. 1557-1575.

Crane, D., 1972. *Invisible colleges*. Chicago University Press: Chicago

Csardi G., Nepusz T. (2006), “The igraph software package for complex network research”, *InterJournal*, *Complex Systems*, p. 1695.

Dasgupta P, David PA. (1994), "Toward a new economics of science", *Research Policy*, 23, pp.487-521.

De Solla Price D.J (1963), *Little Science, Big Science*, New York: Columbia University Press

Della Malva A., Lissoni F., Llerena P. (2011), "Institutional Change and Academic Patenting: French Universities and the Innovation Act of 1999", *Journal of Evoluionary Economics* (forthcoming)

Everett M.G., Borgatti S.P. (1999), “The Centrality of Groups and Classes”, *Journal of Mathematical Sociology*, 23, pp. 181-202.

Fleming, L., King, C., & Juda, A. I. (2007). Small Worlds and Regional Innovation, *Organization Science*, *18*(6), 938-954.

Freeman L.C. (1977), " A Set of Measures of Centrality Based on Betweenness", *Sociometry* 40(1), pp. 35-41

Geuna A., Nesta L. (2006) "University Patenting and its Effects on Academic Research: The Emerging European Evidence", *Research Policy*, = 35, pp. 790-807

Goyal, S., M.J. van der Leij and J.L. Moraga-Gonzalez (2006) “Economics: an emerging small world”, *Journal of Political Economy*, 114(2), 403{412

Guarisco S. (2009), Technology *transfer at public research organisations: the case of CNRS*, unpublished MA dissertation, University of Brescia, Faculty of Engineering

Iversen E.J., Gulbrandsen M., Klitkou A. (2007) “A baseline for the impact of academic patenting legislation in Norway”, *Scientometrics* 70, pp. 393–414

Jones B.M. (2009), "The Burden of Knowledge and the ‘Death of the Renaissance Man’: Is Innovation Getting Harder?" *Review of Economic Studies* 76(1), pp. 283-317

Koenig T., Gogel R. (1981), "Interlocking Corporate Directorships as a Social Network", *American Journal of Economics and Sociology*, 40(1), pp. 37-50

Kogut,B, and Belinky,M.(2008), "Comparing Small World Statistics over Time and Across Countries", *European Management Review*, 5(1): 1-10

Larédo P., Mustar P, 2001. French Research and Innovation Policy: Two Decades of Transformation. In: Laredo P, Mustar P (Eds), *Research and innovation policies in the new global economy*. Edward Elgar: Cheltenham; 447-496.

Lissoni F., Llerena P., McKelvey M., Sanditov B., (2008), “Academic Patenting in Europe: New Evidence from the KEINS Database”, *Research Evaluation* 16(2), pp. 87-102

Lissoni F., Lotz, P., Schovsbo J., Treccani A. (2009) “Academic Patenting and the Professor’s Privilege: Evidence on Denmark from the KEINS database”, *Science and Public Policy*, 36(8), 2009, pp. 595-607

Lissoni F., Sanditov B., Tarasconi G. (2006), “The Keins Database on Academic Inventors: Methodology and Contents”, *CESPRI WP* 181 (http://www.cespri.unibocconi.it/workingpapers)

Llerena (2010)?

Lobo J., Strumsky D. (2008) "Metropolitan patenting, inventor agglomeration and social networks: A tale of two effects", *Journal of Urban Economics*, 63(3), pp. 871-884

MESR (2009), *Demographie des personnels enseignants affectés dans l’enseignement superieur au cours de l’année 2008-2009*, Ministère de l’Enseignement supérieur et de la Recherche, Paris

Molloy M., Reed, B. (1995), “A Critical Point for Random Graphs with a Given Degree Sequence”, *Random Structures and Algorithms*, 6, pp. 161-179.

Moody J. (2004) ‘‘The structure of a social science collaboration network:Disciplinary cohesion from 1963–1999’’. *American Journal of Sociology*, 69, pp.213–238.

Mowery D.C., Sampat B.N. (2005), “The Bayh-Dole Act of 1980 and University-Industry Technology Transfer: A Model for Other OECD Governments”, *Journal of Technology Transfer*, 30, pp. 115-127

Mullins NC, Lowell LH, Hecht PK, Kick EL (1977) "The group structure of co-citation clusters: A comparative study” *American Sociological Review* 42, pp.552-562.

Newman M.E.J. (2001), “The structure of scientific collaboration networks”, *Proceedings of the National Academy of Science USA*, 98, pp.404-409

Newman M. E.J (2004), ‘‘Who is the best connected scientist? A study of scientific coauthorship networks’’. In E. Ben-Naim, H. Frauenfelder and Z. Toroczkai (eds.) *Complex networks*. Berlin: Springer

OST (2008), *Indicateurs de sciences et de technologies* – Annex A, Observatoire des sciences et des techniques, Paris

Powell W.W., Koput K.W., Smith-Doerr L. (1996), ‘‘Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology’’. *Administrative Science Quarterly*, 41: 116–145.

Robins, G., and Alexander, M. (2004), "Small worlds among interlocking directors: Network structure and distance in bipartite graphs". *Computational & Mathematical Organization Theory*, 10: 69–94.

Schnettler S. (2009), "A structured overview of 50 years of small-world research", *Social Networks* 31(3), pp. 165-178

Singh, J.H. (2005) “Collaborative Networks as Determinants of Knowledge Diffusion Patterns”, *Management Science*, 51 (5): 756-770.

Steen J., MacAulay S., Kastelle T. (2010) “A review and critique of the small worlds’ hypothesis: The best network structure for innovation?”, Paper presented at the DRUID Summer Conference.

Strogatz, S. (2001), ‘‘Exploring complex networks’’. *Nature*, 410,pp. 268–276.

Ter Wal A., Boschma R. (2009), "Applying social network analysis in economic geography: framing some key analytic issues", *Annals of Regional Science*, 43(3), pp. 739-756

Thibaut E. (2009), *Transfert de Technologie : Cas du CNRS. Structure de la Pensée Créatrice Autonome*, unpublished PhD dissertation, Bureau d’Economie Théorique et Appliquée, Université de Strasbourg

Thursby J., Fuller A., Thursby M. (2006), “US Faculty Patenting: Inside and Outside the University”, mimeo

Uzzi B., Amaral L.A.N, Reed-Tsochas F. (2007) "Small-world networks and management science research: a review", *European Management Review*, 4, pp.77-91

Verspagen B. (2006), “University Research, Intellectual Property Rights and European Innovation Systems”, *Journal of Economic Surveys* 20 (4), pp. 607–632

Wang X.F., Chen G. (2003) "Complex networks: small-world, scale-free and beyond", *IEEE Circuits and Systems Magazine*, 3(1)pp.6-20

Wasserman S., Faust K. (1994), *Social Network Analysis: Methods and Applications*, Cambridge University Press.

Watts D.J. , Strogatz S.H. (1998), “Collective dynamics of 'small-world' networks”, *Nature*, 393, pp. 440-442.

Wuchty S., Jones B.F., Uzzi B. (2007), “The Increasing Dominance of Teams in Production of Knowledge”, *Science*, 316, pp. 1036-1039

Zuccala, A. (2006) “Modeling the invisible college”, *Journal of the American Society for Information Science and Technology,* 57(2): 152-168.

Appendix A: academic inventors' team size

Let q_m for $m = 1,2..$ be the observed distribution of sizes for all teams (academic and non-academic), i.e. the probability to extract at random from our data a team of size m, and p be the probability to select at random from our data an academic inventor.

If the inventors' teams were assembled at random, then probability for a team of size m to be "academic" (i.e. to include at least one university researcher) would be equal to

$$\Pr\{\text{team of size } m \text{ } is \text{ academic}\} = q_m(1-(1-p)^m).$$

It follows that the "baseline" distribution defined as the distribution of sizes of "academic" teams when research teams are assembled randomly, b_m, would be

$$b_m = \frac{q_m(1-(1-p)^m)}{1-\sum_k q_k\left(1-(1-p)^k\right)}$$

(in the limit of small p this expression reduces to $b_m = q_m m/\langle m\rangle$).

Then we can calculate it for all technological fields, compare to it our observed distribution of academic and CNRS teams, a_m, in the same field and examine whether the two differ.

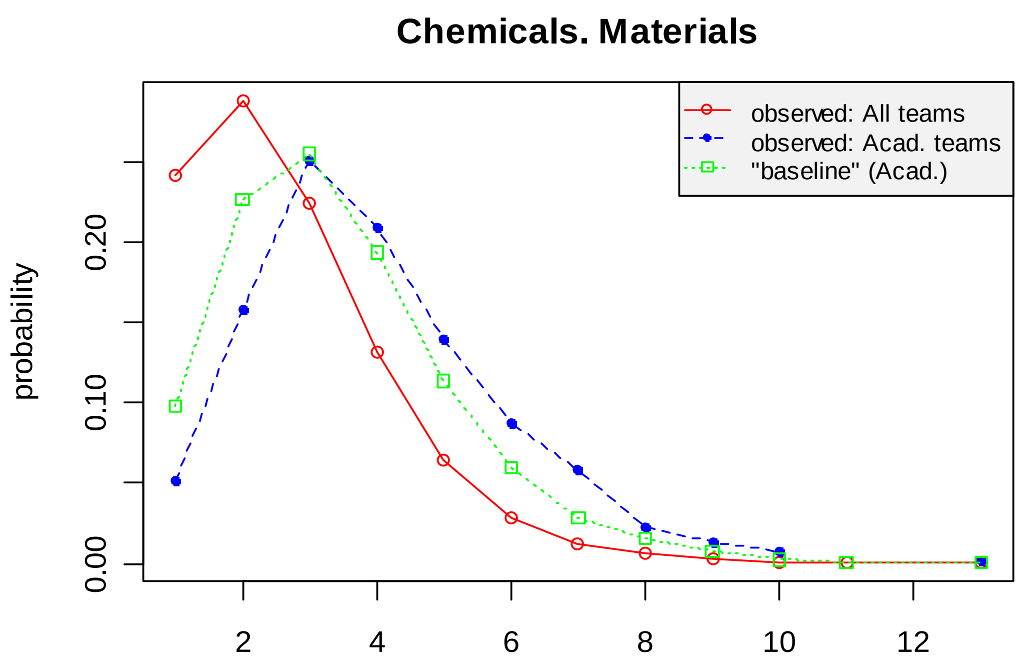

Figure A 1. Observed and "baseline" distributions of the team sizes in *Chemicals & Materials*.

As an illustration consider academic patents in technological field *Chemicals & Materials*. Figure A 1 shows the distribution of all teams (q_m), observed distribution of academic teams (a_m) and corresponding "baseline" distribution of (b_m). At the first glance the observed distribution of academic teams (blue solid circles), differ widely from the distribution for all teams (red hollow circles). However, once we make the correction for the statistical effect, and calculate the “baseline” distribution (green squares), the difference between academic teams and all teams “baseline” distribution become less apparent (although even the corrected “baseline” distribution is still more right-skewed than the observed one).

Documents de travail du BETA

2011–**01** *La création de rentes : une approche par les compétences et capacités dynamiques*
Thierry BURGER-HELMCHEN, Laurence FRANK, janvier 2011.

2011–**02** *Le Crowdsourcing : Typologie et enjeux d'une externalisation vers la foule.*
Claude GUITTARD, Eric SCHENK, janvier 2011.

2011–**03** *Allocation of fixed costs : characterization of the (dual) weighted Shapley value*
Pierre DEHEZ, janvier 2011.

2011–**04** *Data games: sharing public goods with exclusion (2nd version)*
Pierre DEHEZ, Daniela TELLONE, janvier 2011.

2011–**05** *Règle du taux d'intérêt et politique d'assouplissement quantitatif avec un rôle pour la monnaie*
Meixing DAI, janvier 2011.

2011–**06** *Ambiguity and Optimal Technological Choice: Does the Liability Regime Matter?*
Julien JACOB, février 2011.

2011–**07** *Politique budgétaire et discipline budgétaire renforcée dans une union monétaire*
Irem ZEYNELOGLU, mars 2011.

2011–**08** *L'évolution de la segmentation du marché du travail en France : 1973-2007.*
Magali JAOUL-GRAMMARE, avril 2011.

2011–**09** *The scarring effect of unemployment in ten European countries : an analysis based on the ECHP.*
Olivia EKERT-JAFFE, Isabelle TERRAZ, avril 2011.

2011–**10** *Almost common value auctions: more equilibria.*
Gisèle UMBHAUER, juin 2011.

2011–**11** *Increasing returns to scale in U.S. manufacturing industries: evidence from direct and reverse regression*
Xi CHEN, juin 2011.

2011–**12** *Cycle d'innovation dans les services, différences technologiques et similarités organisationnelles dans les entreprises de mécanique françaises et allemandes.*
Thierry BURGER-HLEMCHEN, juillet 2011.

2011–**13** *Reversibility and switching options values in the geological disposal of radioactive waste*
Oana IONESCU, Sandrine SPAETER, juillet 2011.

2011–**14** *Nuclear Waste Disposal in France : the Contribution of Economic Analysis*
Jean-Alain HERAUD, Oana IONESCU, juillet 2011.

2011–**15** *Assessing the impact of the EU ETS using firm level data*
Jan ABRELL, Anta NDOYE FAYE, Georg ZACHMANN, juillet 2011.

2011–**16** *The EU legislation game: the case of asylum law*
Jenny MONHEIM-HELSTROFFER, Marie OBIDZINSKI, août 2011.

2011–**17** *Fiscal disciplining effect of central bank opacity: Stackelberg versus Nash equilibrium*
Meixing DAI, Moïse SIDIROPOULOS, septembre 2011.

2011–**18** *Small Worlds in Networks of Inventors and the Role of Science: An Analysis of France*
Francesco LISSONI, Patrick LLERENA, Bulat SANDITOV, septembre 2011.

La présente liste ne comprend que les Documents de Travail publiés à partir du 1er janvier 2011. La liste complète peut être donnée sur demande.
This list contains the Working Papers writen after January 2011, 1rst. The complet list is available upon request.