ANNALES D'ÉCONOMIE ET DE STATISTIQUE. – N° 79/80 – 2005

"Cross-Firm" Inventors and Social Networks: Localized Knowledge Spillovers Revisited

Stefano BRESCHI* and Francesco LISSONI**

ABSTRACT. – The paper explores the role of inventors' mobility and social networks in generating localized knowledge flows. Using a sample of Italian inventors, we replicate JAFFE's, TRAJTENBERG's, and HENDERSON's [1993] test on patent citations and find similar results. We then control for the role of "cross-firm inventors" (inventors who move across, or do research for different companies), who generate personal self-citations and help creating social links across companies by entering various teams of inventors, which in turn will cite each others' patents. When controlling for personal self-citations, no localization of knowledge flows remains to be seen at the city or province level. What remains of localization effects at the regional level diminishes sensibly after controlling also for the social ties between inventors from cited, citing, and control patents. Knowledge flows thus appear to be localized to the extent that cross-firm activity of inventors and the resulting social networks are also localized. The weight of personal self-citations suggests that frequent interpretations of localized knowledge flows as spillovers, that is externalities, may be misplaced.

« Inventeurs multi-firmes » et réseaux sociaux : un ré-examen des externalités de connaissance localisées?

RÉSUMÉ. – Cet article explore le rôle de la mobilité des inventeurs et de leurs réseaux sociaux en tant que producteurs de flux localisés de connaissance. À partir d'un échantillon d'inventeurs italiens, nous répliquons le test de JAFFE, TRAJTENBERG et HENDERSON's [1993] sur les citations de brevet et nous obtenons des résultats similaires. Nous contrôlons ensuite pour le rôle des « inventeurs multi-firmes » (inventeurs changeant de firmes ou faisant de la recherche dans plusieurs firmes). Ils génèrent des auto-citations et permettent de créer des réseaux sociaux entre entreprises de part leur présence dans plusieurs équipes d'inventeurs, qui chacune citeront les brevets des autres équipes. En contrôlant pour les auto-citations, il ne reste plus d'effet de la localisation des flux de connaissance au niveau de la ville ou de la province. Les effets de localisation qui persistent au niveau régional diminuent sensiblement en tenant compte des liens sociaux entre inventeurs des brevets. Les flux de connaissance apparaissent ainsi comme localisés à l'étendue des activités « multi-firmes » des inventeurs et des réseaux sociaux qui en découlent. L'importance des auto-citations suggèrent que les interprétations fréquentes des flux de connaissances localisés comme *spillovers* peuvent être erronées.

Research for this paper has been supported by a grant provided by the Italian Ministry for Research and University (MIUR). We also gratefully acknowledge financial support from Bocconi University, Milan (Breschi) and the Italian Fulbright Commission (Lissoni); and the kind hospitality of the Sloan School of Management at MIT (Lissoni). The insightful comments of two anonymous referees and the editors have contributed to substantially improve an earlier version of this paper presented at the *International Conference in Memory of Zvi Griliches*, Paris, 25-27 August 2003. We also wish to thank participants at seminars and workshops held at Cespri (Bocconi University, Milan), UQAM (Montreal), Max-Planck Gesellschaft (Jena), ZEW (Mannheim), and Sloan-MIT (Cambridge, MA). We are of course responsible for any remaining error.
* S. BRESCHI: Associate Professor of Applied Economics and Senior Fellow of CESPRI (Centre for Research on Innovation and Internationalization) at Bocconi university, Milan (Italy).
** Fr. LISSONI: Associate Professor of Applied Economics at the university of Brescia (Italy),and Senior Fellow of CESPRI (Centre for Research on Innovation and Internationalization), Bocconi university, Milan (Italy).

1 Introduction

The search for knowledge spillovers has permeated much of the recent literature addressing the spatial dimension of innovative activities. The idea that scientific and technological knowledge may escape its producer's control, and yet diffuse only locally, has become the cornerstone of most explanations of the tendency for innovation activities to be spatially clustered. Yet, this idea has not gone entirely undisputed, the main counter-arguments being that knowledge spillovers are hardly measurable, as they are invisible and do not leave any paper trail, and that there is no compelling reason to conceive them as spatially bounded, as transport and communication technologies have dramatically reduced the costs of transmitting information (KRUGMAN, [1991]).

Although a few indirect tests had already been produced within the knowledge production function approach (GRILICHES, [1992]; for a survey, FELDMAN, [1999]), the task of proposing a direct test for measuring the localization of knowledge spillovers was first taken up ADAM JAFFE, MANUEL TRAJTENBERG and REBECCA HENDERSON (1993, hereafter JTH). In a path-breaking article, the three authors argued that knowledge spillovers may indeed leave a paper trail in the form of citations to prior art contained in patent documents. Accordingly, JTH set up an original experiment to show that citations to prior art tend to come disproportionately from the same geographical area as the cited patents, therefore supporting the notion that knowledge spillovers are spatially localized.

The basic JTH methodology has become a classical reference for any empirical work on knowledge spillovers. When it comes to interpreting its results, however, the various authors have pointed at different knowledge diffusion mechanisms, not all of them compatible with the original definition of "spillover" as a pure externality. The mobility of R&D scientists and engineers within localized clusters of firms, the existence of advice exchange networks within communities of scientists and practitioners, or even localized markets for technologies and scientific/technical consultancy, have all been mentioned as possible explanations of the JTH findings.

In this paper we propose a refinement to the JTH methodology that may help telling these different mechanisms apart. We reclassify patents according to their inventors, and find that inventors who patent across different companies ("cross-firm" inventors) contribute extensively to the observed citation patterns, both directly (through personal self-citations) and indirectly, by linking the various companies *via* a social network of inventors conducive to more citations.

To the extent that the geographical mobility of these "cross-firm" inventors is very limited, the resulting social networks and citations patterns are found to be bounded in space.

These results cast some doubts on the "spillover interpretation" of citation-based evidence on localized knowledge flows, as the cross-firm inventors' contribution to knowledge diffusion may not entirely be seen as a pure externality.

The paper is organized as follows. In Section 2, we briefly summarize the original experiment proposed by JTH. Section 3 discusses the role that social networks and the mobility of inventors across firms may play as drivers of knowledge diffusion, and outlines our original methodology to measure their effects. Section 4 describes the data used and Section 5 illustrates our results. Section 6 concludes with a discussion of our main findings and some suggestions for future research.

2 The JTH's experiment: methods and discussion of results

2.1 Methodology

The JTH's experiment started with the selection of a sample of *originating* patents[1]. For each originating patent, the authors collected all subsequent patents citing it as prior art, with the exclusion of company self-citations, i.e. pairs of citing-cited patents assigned to the same company[2]. The address of inventors recorded in patent documents was then used to assign patents to a geographical area and to match the locations of citing and cited patents[3]. Although one might be tempted to stop here and look at the frequency with which citing and cited patents match geographically, such a simple test would tell us little about the localization of knowledge spillovers, unless one controls for the pre-existing geographic concentration of innovative activities. In other words, one might find that a very large share of citations come from the same area as the cited patents simply because the *production* of technological innovations (i.e. patents) happens to be agglomerated in that area. The *production* of innovations, in turn, may be spatially agglomerated for a number of reasons, which have nothing to do with the access to the local pool of knowledge (e.g. availability of skilled labour and specialized inputs, the infrastructure endowment of cities and regions, etc.).

The important contribution of the JTH paper was to develop a methodology, which allows one to separate the effects of pure knowledge spillovers from the impact of other agglomeration forces. Specifically, JTH built a *control* sample of patents in the following way. Each citing patent was matched to a randomly drawn patent, which had the same technology class and application date as its matched citing patent, *but did not cite* the same originating patent[4]. JTH's test consisted then

1 In particular, JTH selected two cohorts of patents, one consisting of 1975 patent applications and the other of 1985 applications. The choice of two cohorts allowed to control for possible changes over time in the geographical reach of spillovers.

2 The exclusion of company self-citations was motivated by the fact that these do not represent spillovers. Originating patents with no citations received were similarly excluded from the analysis, since they do not provide any information on the localization of knowledge spillovers.

3 The rules followed by JTH to locate patents in space are indeed to complex to be summed up here. Two full paragraphs of their article are devoted to explain them (p. 585). JTH locate patents at three different levels: country (US *vs.* non-US), states and Metropolitan Areas.

4 Patent offices classify applications according to very detailed technological codes, which should reflect the technological contents of the inventions. JTH used patent data from the United States Patent Office (USPTO), which classifies patents according to the Unites States Classification System (USCS). JTH matched patents according to 3-digit patent classes of the USCS. Thompson and Fox-Kean [2005] have criticized the selection process proposed by JTH. Their main argument is that the level of technology aggregation adopted by JTH to match citing and control patents is likely to induce spurious localization effects. Although their results are surely important, we will stick to the original JTH methodology, in order to allow easier comparisons of results.

in comparing the frequency with which citing and cited patents match geographically, with the frequency with which control and cited patents match geographically. If the former turns out to be significantly greater than the latter, this should be interpreted as evidence of localization effects (i.e. spillovers) *over and above* the agglomeration effects arising from other sources. More specifically, the JTH exercise consisted in comparing

"the probability of a patent matching the originating patent by geographic area, *conditional* on its citing the originating patent, with the probability of a match *not conditioned on the existence of a citation link*. This noncitation-conditioned probability gives a baseline or reference value against which to compare the proportions of citations that match" (JTH, 1993, p. 581).

The evidence reported by JTH shows indeed that citations are highly localized. Citing patents are up to two times more likely than the control patents to come from the same state, and up to six times more likely to come from the same metropolitan area.

2.2 Interpretation issues and developments

The mobility of R&D scientists and engineers within a localized labour market, the free exchange of information within the cross-company social networks created by that mobility, and the existence of localized markets for technologies have all been reported by various authors as potential explanations of JTH results. Not all of these explanations are compatible with the "spillover" interpretation of localized knowledge flows as measured by patent citations; others have so far proved to be quite elusive to measurement efforts.

As for labour mobility, ALMEIDA and KOGUT [1999] have replicated the JTH exercise for each US state. They find evidence of localized knowledge flows only in those few regions (most notably, the Silicon Valley) where the intra-regional mobility of inventors across companies is high. It is hard to tell to what extent the "labour mobility" explanation is compatible with the "spillover" interpretation of localized patent citations: there seems to be no way to tell how much knowledge moves from company to company along with the inventor (and remains the inventor's private asset) and how much it is spread across all the companies the inventor works for (in which case externalities may be generated).

Regarding the role of social networks as drivers of localized knowledge diffusion, the typical argument is that agents who are co-located in the same region or city are more likely to be embedded in a thick web of social ties through which tacit knowledge may easily flow (AUDRETSCH, [1998]). Geographical proximity is said to reduce the cost for scientists and engineers to meet personally, and to increase the chances of serendipitous encounters at workshops, conferences and various social gatherings. Unfortunately, geography is a very broad proxy for social cohesion: within the same city or region, a scientist or an engineer may belong to many different networks, only a few of which provide him with the knowledge necessary to his research (such as professional contacts, friendship ties with former fellow students, professors or colleagues). Besides, the geographical reach of the various networks may vary, as some contacts may be retained and used over long distances, while others need to be nurtured by frequent contacts, which are only possible if they are geographically proximate. Any step toward a direct measurement of the social networks of inventors may be of great help in opening the "black box" of geography as proxy of social cohesion.

Markets for technologies may explain the JTH results to the extent that technology users need to consult frequently with suppliers. Research contracts signed by the same independent inventor with different companies may produce patents that appear to be unrelated in terms of ownership, but very close in terms of technological contents and geographical distance; to the extent that the inventor builds on her previous works, the patents will also exhibit a citation link. Some evidence in this direction exists for the case of technology licensing (Arora, [1996]; Zucker, Darby and Armstrong, [1998]; Mowery and Ziedonis, [2001]).

In the remainder of the paper, we build on these developments and insights. We propose that it is possible to make further use of patent and citation data to assess the role of inventors moving across firms (or consulting for different firms), and of social networks of inventors, as derived from co-invention data. The next section illustrates in detail our methodology.

3 Methodology and data

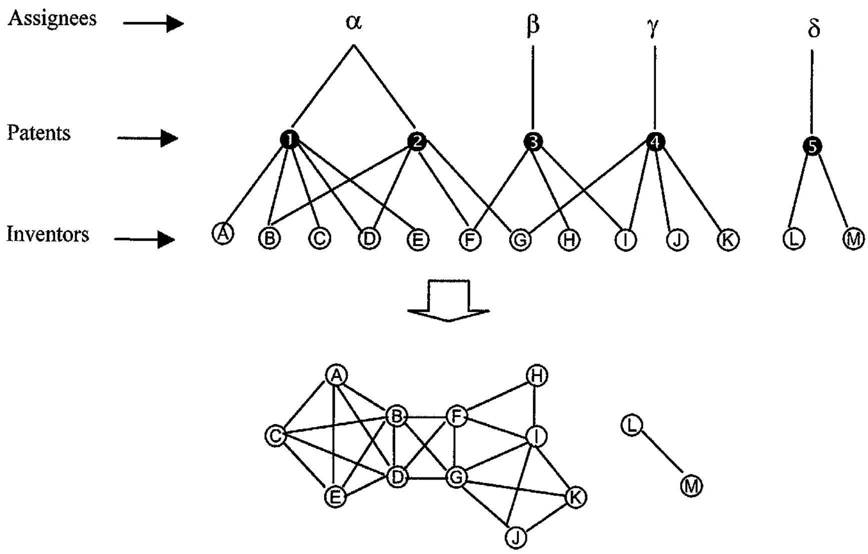
3.1 Social networks: methodology and definitions

Our methodology exploits information recorded in patent documents regarding the names, surnames, address and company affiliation of each inventor. This information can be used to track the existence of a social linkage between the citing and the cited patents and to identify instances of mobility of inventors across companies. The following hypothetical example illustrates the main idea (see Figure 1). Let's consider five patents (1 to 5) and four assignees (α, β, γ, δ). Assignee α owns two patents (1 and 2), while assignees β, γ and δ own one each. Patents have been produced by thirteen distinct inventors (A to M). For example, patent 1 assigned to company α has been produced by a team comprising inventors A, B, C, D and E. It is reasonable to assume that, due to the collaboration in a common research project, these five inventors are *socially linked* by some kind of knowledge sharing. The existence of such a linkage can be graphically represented by drawing an undirected edge between each pair of inventors, as in the bottom part of Figure 1.

Repeating the same exercise for each team of inventors, we end up with a map representing the network of all inventors[5]. Using the graph just described, we can

5 In the language of graph theory, the top part of the figure reports the affiliation network of patents, applicants and inventors. An affiliation network is a network in which actors (e.g. inventors) are joined together by common membership to groups of some kind (e.g. patents). Affiliation networks can be represented as a graph consisting of two kinds of vertices, one representing the actors (e.g. inventors) and the other the groups (e.g. patents). In order to analyze the patterns of relations among actors, however, affiliation networks are often represented simply as unipartite (or one-mode) graphs of actors joined by undirected edges – two inventors who participated in the same patent, in our case, being connected by an edge (see bottom part of Figure 1). Please note that the position of nodes and the length of lines in the graph have no specific meaning.

FIGURE 1
Bipartite Graph of Patents and Inventors



Top: Bipartite graph of assignees (α, β, γ, δ), patents (1 to 5) and inventors (A to M), with lines linking each patent to the respective inventors and assignees.
Bottom: the one-mode projection of the same network onto just inventors

measure how connected two patents are. In order to see how, we first give a few definitions:

i) For any pair of inventors, one can measure the distance between the two by calculating the so-called *geodesic distance*. The geodesic distance is defined as the minimum number of edges that separate two distinct inventors in the network[6]. In Figure 1, for example, the geodesic distance between inventors A and C is equal to 1, whereas the same distance for inventors A and H is 3. While A and C shared directly their knowledge while working on patent 1, A and H are more likely to have exchanged some word-of-mouth technical information through the mediation of other actors (such as B and F).

ii) Inventors may belong to the same *social component* or they may be located in *socially disconnected components*. A component of a graph can be defined as a subset of the entire graph, such that all nodes included in the subset are connected through some path. In Figure 1, for example, inventors A to K belong to the same component, whereas inventors L and M belong to a different component. A pair of inventors belonging to two distinct components have distance equal to infinity (i.e. there is no path connecting them).

iii) We define a *cross-firm* inventor any inventor whose name has been reported in patent documents assigned to different organizations. This kind of inventors play a fundamental role in connecting teams of inventors belonging to different organizations. For example, in Figure 1, inventor F worked for both company α and

6 For technical terms from social network analysis, see WASSERMAN and FAUST [1994].

β, thus connecting the team of inventors (B, D) with the team of inventors (H, I). Similarly, inventor G worked both for company α and γ, thus connecting the team (B, D, F) with the team (I, J, K).

In addition, cross-firm inventors are of great interest for their own merits. They may be either mobile R&D employees who move across companies, or independent inventors (such as academic scientists) selling their output to different companies[7]. At this stage of our research, we can not distinguish between the two kinds of inventors. (Besides, we are possibly underestimating the role of employee mobility since a potentially important part of mobility is unobserved[8].)

Using these definitions, we may now turn to illustrate how the existence of a linkage between patents can be ascertained. Three possible relations exist between any pair of patents from different firms:

1) The two patents exhibit *no social connection*, such as when the inventors behind them belong to socially disconnected components[9].

2) The two patents are linked by a *social connection*, such as when their inventors belong to the same social component. We also calculate the social distance *between patents* as the geodesic distance between the two closest individuals from the two teams of inventors (*minimum* geodesic distance)[10]. As such, the social distance between two socially connected patents may vary from 1 to any positive discrete value.

3) The two patents are linked by a *personal connection*, such as when at least one inventor belongs to both patents' teams. The social distance between two personally connected patents is zero[11].

3.2 **Implementation on EPO patents by italian inventors**

To implement the methodology just described, we have relied on EP-INV, a biographical dataset on 21,526 Italian inventors and their 26,898 patent applications at the European Patent Office (EPO), filed between 1978 and 1999. EP-INV is a

7 In Balconi, Breschi and Lissoni [2004] we estimate that roughly 3% of patents by Italian inventors come from academic scientists on active duty in 2000. Most of these patents belong to business companies.

8 This is the case when an inventor leaves a firm, but he does not patent in his new company, or conversely when an inventor who did not patent with the former employer starts patenting in his new company. We wish to thank an anonymous referee for drawing our attention on this point.

9 With reference to Figure 1, this is the case, for example, of patent 5 and patent 1.

10 When two patents are socially connected, all of their inventors belong necessarily to the same social component, but not all them are at the same geodesic distance. For example, in Figure 1 patents 4 and 1 are socially connected, but inventor K (from patent 4) and inventor A (from patent 1) exhibit a geodesic distance of 3, while inventors G (patent 4) and B (patent 1) have a geodesic distance of just 1. G and B are the closest inventors, and it is the geodesic distance between them that we pick up as the social distance between patents 4 and 1. In other words, the social distance between the two patents is the minimum geodesic distance between their inventors. See Breschi and Lissoni [2004] for further details, and a discussion of this choice.

11 For example, inventors G, K and J (from patent 4) and H and F (from patent 3) belong to the same social component; in addition, inventor I appears both in patent 4 and 3. In the absence of I, patents 4 and 3 would be socially connect at distance 1 (the geodesic distance between G and F). The presence of I reduces this distance; we capture this reduction by setting the distance to zero.

subset of EP-CESPRI, a much larger database that classifies *all* EPO patent applications from 1978 (EPO's first year of activity) to the current year, by company[12].

The EP-INV database allowed us to track the overall affiliation network of patents, applicants and inventors, as well as the one-mode projection of the same network onto just inventors (see Figure 1 above). The resulting network of inventors has then been used to derive measures of social linkage between citing (control) and cited patents. Specifically, for each citing (control) patent with application year T we have constructed the network of inventors taking all patent applications from 1978 to year T-1[13]. For example, in the case of a patent issued in 1995 citing a patent issued in 1987, we constructed the network of inventors on the basis of all patent applications from 1978 to 1994. On the basis of the network of inventors at time T-1, we have then categorized each pair of citing(control)-cited patents at time T as non-connected, socially connected, and personally connected, as from the definitions in the previous section.

3.3 Geographic assignment of patents

A major problem in measuring the frequency of geographic matching between cited and citing (control) patents relates to the way patents are assigned to locations. Patent documents report the town/city and postal address of each inventor[14]. However, patents can have multiple inventors, each one with a different address. Therefore, the location of patents in geographic space cannot be resolved in an unequivocal way. In case of multiple inventors, JTH assigned each patent to the country/state in which pluralities of inventors resided, with ties assigned arbitrarily. Here, we take a slightly different approach and argue that two patents match geographically to the extent that they share *at least* one inventor's location.

3.4 Data sampling

Following as closely as possible the methodology developed by JTH, we have selected for this study three cohorts of *originating* patents, consisting respectively

12 Companies in the EP-CESPRI database are identified by name and address. Groups and companies with multiple names and addresses are identified, for 40% of the applications, by matching names to Dun&Bradstreet codes. As for EP-INV, inventors are identified by checking raw names for misspellings, use of initials and second names. Moreover, a round of e-mailing and phone calls helped identifying homonyms. *Ad hoc* checks have been made on multiple names and the group structure of applicant of patents by cross-firm inventors. For a fuller description of EP-INV and for some more descriptive statistics on the resulting network of inventors, see BALCONI, BRESCHI and LISSONI [2004]. For more information on an early version of EP-CESPRI see BRESCHI *et al.* [2003].

13 Two points have to be remarked. First, the network of inventors has been constructed in a cumulative way by adding each year new nodes (i.e. inventors) and new linkages (i.e. patents). To the extent that the importance of linkages among inventors decays over time, one could alternatively think to remove old patent applications in order to construct the network of social linkages among inventors. Second, when considering a citing patent at year T the existence of a social linkage among the inventors behind that patent and the inventors behind the cited patent has been evaluated on the basis of the network of inventors as at year T-1, i.e. without including patent applications at year T. The reason for this choice is that if social linkages matter as vehicles for transmitting knowledge then the existence of a linkage among inventors should be evaluated at some time *before* the date of application of the citing patent. We think that one-year lag is a reasonable window of time for that purpose.

14 The Nomenclature of Statistical Territorial Units (NUTS) has been used here to define the spatial units of analysis. The city level corresponds to the so-called "comuni" (NUTS4), of which there are 8,100. Moreover, there are 95 provinces (NUTS3) and 20 regions (NUTS2).

of the 1987, 1988 and 1989 patent applications[15]. For each cohort, we considered all patent applications in the EP-INV database that received at least one subsequent citation by the end of 1996. The 1987 originating cohort contains 699 patents that had received a total of 1,631 citations by the end of 1996. The 1988 originating cohort contains 843 patents that had received a total of 1,784 citations by the end of 1996. The 1989 originating cohort contains 779 patents that had received a total of 1,615 citations by the end of 1996.

For each cohort of originating patents, we eliminated all applications that either received citations only from foreign organizations, or whose applicant was an Italian organization, *but did not* report any Italian inventor[16]. It must be pointed out that the choice of excluding citations from foreign companies implies that our study investigates the extent of intra-national localization of patent citations, and it is unable to say anything about the extent of international localization. This choice has been mainly dictated by data constraints, as the inclusion of citations coming from foreign organizations would have implied the construction of the worldwide network of inventors. At the same time, given that the basic intuition behind the notion of localized knowledge spillovers is that the strength of spillovers should fade with distance, our choice should not have major effects upon the results.

For each originating patent, we then took all patents that subsequently cited them as prior art. For the construction of the citing sample, we considered only patent applications made before 1996 included. Moreover, since we are interested in knowledge spillovers, we removed all observations in which citing and originating patents have been assigned to the same organization (i.e. company self-citations).

Finally, for each citing patent we took the primary classification code at the 4-digit level of the International Patent Classification (IPC) and used this to construct a sample of control patents. Specifically, for each citing patent we identified all patents in the same patent class with the same application year. We then chose from that set a control patent whose application date was as close as possible to that of the citing patent, and that did not cite the same originating patent. The resulting data set therefore consists of all originating patents, for which there is a matching of citing and control patents. In turn, each citing patent is paired with a specific control patent within the same technological class and with (approximately) the same application date. The final sample consists of 366 originating patents, which have received 483 citations from other Italian organizations.

4 Descriptive statistics

This section reports a few descriptive statistics concerning our sample of citing, cited and control patents as well as the overall network of inventors from the EP-INV database, which has been used to assess the degree of social connectedness between pairs of patents.

15 The priority year has been used to date patent applications.

16 The nationality of inventors has been derived by the address reported in patent documents. It is worth pointing out that the share of patent applications by Italian organizations made exclusively by non-Italian inventors is negligible (around 2% for each cohort of originating patents). On the other hand, the share of originating patents receiving citations *only* from foreign organizations is high (around 60% for each cohort of originating patents).

4.1 Overall network

Table 1 reports some descriptive statistics for the one-mode network of inventors. By construction, the size of the network grows as new inventors start patenting (in the absence of more information, we are forced to assume that social ties arising from co-invention are never severed). At the same time, the average number of inventors in each component also grows, as previously disconnected inventors from different firms get in touch with cross-firm inventors, and each firm's pools of inventors gets wider.

TABLE 1
Evolution of the One-Mode Network of Italian Inventors (1978-95)

Year	Number of inventors	Number of edges[(a)]	Number of components of size ≥ 2[(b)]	Average size of components[(c)]	Size of largest component[(d)]
1978-1986	6,670	5,203	1,084	3.7 (8.0)	164
1978-1987	8,058	6,534	1,287	3.8 (20.3)	723
1978-1988	9,554	7,912	1,487	3.9 (26.9)	1,032
1978-1989	11,117	9,557	1,661	4.1 (33.4)	1,359
1978-1990	12,951	11,474	1,878	4.2 (43.7)	1,885
1978-1991	14,613	13,294	2,100	4.3 (48.1)	2,194
1978-1992	16,412	15,421	2,329	4.4 (52.6)	2504
1978-1993	18,048	17,514	2,508	4.5 (58.0)	2,858
1978-1994	19,725	19,437	2,731	4.6 (61.7)	3,166
1978-1995	21,526	21,593	2,969	4.6 (64.8)	3,449

(a) Total number of edges in the one-mode network of inventors.
(b) Number of socially connected components including at least two inventors.
(c) Average number of inventors in components with at least two inventors (standard deviation).
(d) Number of inventors in the most numerous component.

In Table 2 we report a simple calculation on the extent of inventors' cross-firm patenting activity over the period 1978-95. Specifically, the table reports the distribution of all inventors included in the EP-INV database according to the number of different organizations for which they have signed at least one patent. As one can see, the distribution is highly skewed. The vast majority of inventors (85%) registered patents for only one organization. At the same time, we also note that the frequency of cross-firm inventors is not negligible: 2,350 inventors (around 11% of all) have signed patents for two different organizations, while more than 800 (4%) have signed patents for three or more organizations.

A key feature of our population of inventors is the extent of geographic mobility. In table 3, we report the percentage of cross-firm inventors that never changed location, at the levels of city, province and region. The striking result emerging from the table is the high degree of geographic immobility that characterizes the

TABLE 2
Inventors' Activity across Assignees (1978-1995)*

Number of different assignees	Number of inventors	Percentage
1	18,353	85.3
2	2,350	10.9
3	549	2.6
4	161	0.7
5	65	0.3
6	25	0.1
7	11	0.1
>7	12	0.1
Total	21,526	100.0

* *The table reports the distribution of all Italian inventors according to the number of different assignees for which they have recorded patents over the period 1978-1995. The calculation includes patents registered by individual inventors, i.e. not assigned to organizations. The calculation includes also patents that have been co-assigned to different organizations.*

TABLE 3
Geographic Location of Cross-Firm Inventors* (1978-1995)

Number of different assignees	Percentage of cross-firm inventors that never changed geographic location		
	City	Province	Region
2	90.6	95.8	97.2
3	84.6	92.7	94.5
4	76.7	87.7	92.3
5	70.3	87.5	95.3
6	76.0	88.0	92.0
7	63.6	72.7	81.8
>7	75.0	75.0	83.3

population of Italian inventors. About 90% of all inventors that worked for two different employers never changed the city of residence. That percentage increases to 97% at the level of region. Even though the extent of geographic immobility is slightly lower for inventors that patented across more than two companies, the propensity to change location remains remarkably low.

An important consequence one may expect from the data in table 3 is the concentration in space of social network of inventors. Inventor teams from different companies are linked one another by cross-firm inventors. To the extent that the latter are found to be scarcely mobile in space, we expect a high percentage of the teams linked through them to be located in the same geographical areas. As calculations of the spatial dispersion of social network of inventors prove to be quite complicated, we will to check this intuition only on our sample data.

4.2 Sample data

First of all, we note that our sample of citing, control and cited patents includes for the most part patents assigned to private companies. Of the 483 citing and control patents, 472 and 473, respectively, come from business companies[17]. Similarly, of the 366 cited patents, 359 have been assigned to business companies. Secondly, we also note that the number of patents co-assigned to different companies is very low. Of the 483 citing and control patents, only 13 and 18, respectively, have more than one assignee. Similarly, of the 366 cited patents, only 14 have more than one assignee[18].

Table 4 reports the composition of our final sample, as well as some summary statistics concerning the number of inventors included in it, and the distribution of citing-cited and control-cited pairs of patents according to their social connectedness. In particular, the table shows that the absolute number of personally connected patent pairs is much higher for the citing-cited sample than for the control-cited sample, while the opposite holds for socially connected pairs.

TABLE 4
Sample of Cited, Citing and Control Patents: Number of Inventors and Linkages

	N. of patents	N. of inventors	N. of inv. *per patent*	N. of inventor pairs*(a)*	N. of connected patents		
					Total*(b)*	*Personal connection(c)*	*Social connection(d)*
Cited	366	572	1.8 (1.2)	-			
Citing	483	721	2.0 (1.4)	1,789	132	*76*	*56*
Control	483	726	1.9 (1.2)	1,927	100	*17*	*83*

(a) The column reports the total number of pairs of cited inventors – citing (control) inventors summed up over all patents in the sample [Note: a patent signed by n inventors citing a patent signed by m inventors generate n × m pairs of inventors].

(b) The column reports the absolute number of citing (control)-cited patent pairs in which inventors from the two teams belong to the same component, i.e. are directly or indirectly linked in the social network. [minimum geodesic distance finite].

(c) The column reports the absolute number of citing (control)-cited patent pairs which are personally connected (the same inventor appears in both patents) [minimum geodesic distance = 0].

(d) The column reports the absolute number of citing (control)-cited patent pairs which are socially connected according to the social network of inventors [minimum geodesic distance finite and > 0].

---

17 This is not surprising given the weakness of the Italian system of public research and the low propensity of Italian universities and public research organizations to file for patents. For the few patents in our sample that have not been assigned to private organizations, the largest applicants are the National Research Council and the Ministry for Research and Education.

18 In case of co-assigned patents, the identification of company self-citations has been made by comparing all the assignees behind the citing and the cited patents.

Socially connected patents within the control sample outnumber those in the citing sample. However, the two samples differ with respect to their social distance from the cited patents. As shown in table 5, citing patents host a much higher percentage of patents whose distance from cited patents is just one. As a consequence, the mean distance of citing patents from cited ones is much shorter than the distance of control patents[19].

TABLE 5
Socially Connected Citing (Control) Patents: Geodesic Distances from the Cited Patent

Geodesic dist.	N. of citing patents (%)	N. of control patents(%)
1	23 (41.1)	17 (20.5)
2	8 (14.3)	9 (10.8)
3	4 (7.1)	9 (10.8)
4	5 (8.9)	10 (12.0)
> 4	16 (28.6)	38 (45.8)
Total	56 (100)	83 (100)
Mean (sd)§	*3.30 (2.960)*	*4.59 (3.044)*

§ t-test of mean difference = -2.4881.

This sampling outcome is consistent with the results obtained in a companion paper, which shows that the probability to observe a citation tie between two patents declines with the social distance between them (BRESCHI and LISSONI, [2004]). The citation probability is highest for personally connected patent pairs, and it is much higher for socially connected pairs than for unconnected ones. However, as the social distance between patents increases, the citation probability declines sharply: for paths longer than four geodesic steps, no meaningful difference is any more observed between socially connected and un-connected patents.

Finally, we have calculated the geographical distribution of both the citing and the control patent samples, with respect to the originating patents. As expected, personally connected patent pairs, both in the citing and the control sample, are highly co-localized with the cited ones, up to 100% at the regional level (table 6, box 6.1).

Socially connected patent pairs are also highly localized, but co-location percentages are higher for citing patents, especially at the regional level (table 6, box 6.2). These differences are explained by reading jointly table 7, which reports the geographic matching between cited patents and their socially connected patents (cited and controls), and table 5. Table 7 suggests that connected patent pairs separated by just one geodesic step are much more likely to be co-located than other connected pairs; table 5, as we said, indicate that patent pairs at such short geodesic distance are more frequent in the citing sample.

19 We also notice that 33 out of 83 socially connected control patents match as many personally linked citing patents, that is are more weakly linked to the same originating (cited) patent. On the contrary, only 3 out of 56 socially connected citing patents are matched to personally linked controls. If we drop from our sample all patent pairs where either the control or the cited patent are personally linked, the citing and control sample turn out to host pretty much the same number of socially connected patents (54 and 50 respectively), and the differences in the mean distance to cited patents increase further.

TABLE 6
Geographic Matching: % Frequency for Personally/Socially Connected Patents

Matching by	Citing *(n. of patents)*	Control *(n. of patents)*
	6.1 Personal connections	
City	92.1	88.2
	(70)	*(15)*
Province	98.9	100.0
	(75)	*(17)*
Region	98.9	100.0
	(75)	*(17)*
	6.2 Social connections	
City	35.7	34.9
	(20)	*(29)*
Province	60.7	50.6
	(34)	*(42)*
Region	73.2	62.6
	(41)	*(52)*

Size of samples for personally connected patents: 76 citing/17 controls.
Size of samples for socially connected patents: 56 citing/83 controls.

TABLE 7
Geographic Matching for Socially Connected Patents, by Geodesic Distance

Geodesic distance	Total patents (citing + controls)	Geographic match % (regional level)
1	40	77.5
2	17	64.7
3	13	61.5
4	15	60.0
>4	54	63.0

Total connected inventors: 139 obs (56 citing + 83 controls).

5 Results

As a first step in our analysis, we simply replicate the JTH's exercise on our data. Box 8.1 in table 8 reports the percentage of citing patents that are co-located with the cited ones, at the city, province and regional level. The second column reports the same percentage for the control sample. The third column reports the value of a z test for the difference between the two proportions, which is distributed according to a standard normal; a 1-tail test is performed, on the hypothesis that $p_c > p_{nc}$

(where p_c is the co-location probability of citing patents, and p_{nc} the co-location probability of controls). Finally, the fourth column reports the 95% confidence interval for the estimated odds ratio[20].

Our results successfully replicate those found by JTH with reference to the United States. The proportion of citing patents co-located with cited ones is significantly greater than the proportion of control patents, at all geographical levels. As suggested by the odds ratio confidence interval, citing patents are at least 15% more likely then the control ones to be co-localized with the cited patents; the same value for the regional level is 31%. The upper bound values suggest that citing patents may even be as twice more likely than controls to come from the same area of the originating (cited) patent.

The descriptive statistics we presented in section 4, however, suggest that the different composition of the citing and control sample may be responsible for this result. We know that the citing patents are much more likely than the control ones to be personally connected to the cited ones, and that the cross-firm inventors responsible for the connection are most often immobile in space. Hence we expect that controlling for personal connections may reduce the observed differences in co-location rates between citing and control patents.

As the citing patent sample contain a higher number of socially connected patents at geodesic distance 1 (which we know to be highly co-located with cited ones), we expect to observe a similar result when controlling for social connections.

Box 8.2 of table 8 shows that by excluding all the personally connected patents from both the citing and the control samples, the observed differences in the co-location percentages of the citing and control samples disappear at the city and

---

20 We define:

$$z = \frac{\hat{p}_c - \hat{p}_{nc}}{\sqrt{\hat{p}(1-\hat{p})(1/n_c + 1/n_{nc})}},$$

where $\hat{p}_c$ and $\hat{p}_{nc}$ are the sample proportion estimates for the citing and the control patents, respectively; n_c, n_{nc} are the size of the citing and the control samples (in our case: $n_c = n_{nc}$); and

$$\hat{p} = \frac{colocated_c + colocated_{nc}}{n_c + n_{nc}},$$

where $colocated_c$ and $colocated_{nc}$ are the number of co-located citing patents and controls, respectively.

The 1-tail test of our interest calculates the probability attached to values higher than z from a standard normal distribution (in table 8: $P > z$). As z^2 is distributed as a Chi-square with one degree of freedom, a similar test can be performed using that distribution.

The odds of co-location for a citing patent are defined as $O_c = p_c/(1 - p_c)$. Let O_{nc} to be the corresponding odds of co-location for a non-connected citing patent. The odds ratio is defined as $OR = O_c/O_{nc}$, and it is also distributed according to a Chi-square with one degree of freedom, from which confidence intervals can be derived (Fleiss, [1981]). *OR* values higher than 1 suggest a higher colocation probability of citing patents, as opposed to non-citing ones.

JTH's test of proportion is slightly different, as it is based on a *t*-distributed statistic (JTH, p. 589):

$$t = \frac{\hat{p}_c - \hat{p}_{nc}}{\sqrt{\left[\hat{p}_c\left(1-\hat{p}_c\right) + \hat{p}_{nc}\left(1-\hat{p}_{nc}\right)\right]/n}}$$

where $\hat{p}_c$ and $\hat{p}_{nc}$ have the same meaning as above, and n is the size of both the citing and the cited samples ($n = n_c = n_{nc}$).

TABLE 8

Geographic Matching % Frequency and Test of Proportions (citing vs control)

Matching by	Citing *(n. of patents)*	Control *(n. of patents)*	*z*-statistic $(P > z)$	Odds ratio 95% conf. interval
	8.1 All observations (JTH experiment)[(a)]			
City	25.1	17.4	2.91	1.15-2.20
	(121)	*(84)*	*(0.00)*	
Province	38.7	29.8	2.92	1.13-1.96
	(187)	*(144)*	*(0.00)*	
Region	53.8	40.6	4.13	1.31-2.22
	(260)	*(196)*	*(0.00)*	
	8.2 Excluding personal connections[(b)]			
City	12.5	12.5	0.00	0.65-1.56
	(50)	*(50)*	*(0.50)*	
Province	27.8	25.8	0.64	0.80-1.53
	(111)	*(103)*	*(0.26)*	
Region	46.1	37.8	2.367	1.05-1.88
	(184)	*(151)*	*(0.01)*	
	8.3 Excl. both personal & social connections[(c)]			
City	8.5	10.9	-1.052	0.43-1.31
	(28)	*(36)*	*(0.85)*	
Province	21.8	22.7	-0.281	0.65-1.39
	(72)	*(75)*	*(0.61)*	
Region	40.2	33.8	1.690	0.95-1.83
	(133)	*(112)*	*(0.05)*	
	8.4 Excl. personal conn. & social connections with geodesic ≤ 2[(d)]			
City	8.8	12.1	-1.455	0.42-1.16
	(32)	*(44)*	*(0.92)*	
Province	23.1	24.8	-0.522	0.64-1.30
	(84)	*(90)*	*(0.70)*	
Region	42.2	36.9	1.517	0.92-1.71
	(154)	*(134)*	*(0.06)*	

(a) 483 obs. in both the citing and the control samples.

(b) 399 obs. in both the citing and the control samples. The new samples result from those in box 8.1, after dropping all personally connected citing-cited pairs (and the related controls) and all personally connected control-cited pairs (and the related citing patents).

(c) 331 obs. in both the citing and the control samples. The new samples result from those in box 8.2, after dropping all socially connected citing-cited pairs (and the related controls) and all socially connected control-cited pairs (and the related citing patents).

(d) 363 obs. both the citing and the control samples. The new samples result from those in box 8.2, after dropping socially connected citing-cited pairs (and the related controls) and socially connected control-cited pairs (and the related citing patents) with geodesic distances ≤ 2.

province level, and diminish sharply at the regional level[21]. Co-location at the city level is now exactly the same for the citing and control samples, while differences at the province level are no more significant. The lower bound of the odds ratio confidence interval at the regional level is still higher than one, but barely so (as one can expect from the sharp decrease of the *z*-statistic).

These figures suggest that the results obtained by applying the original JTH methodology can be largely explained by the importance of cross-firm inventors. Cross-firm inventors are responsible for a large number of citations, but are scarcely mobile in space: they move or diffuse their knowledge across different firms, but not across different localities. As we remove the patent pairs linked personally by those inventors, the JTH experiment does not confirm anymore that citations are localized at the city and province level; and it suggests much weaker localization effects at the regional level.

We then proceed in a similar fashion to check the effect of socially connected patents. When we exclude from our exercise also the socially connected patents (box 8.3, table 8), we observe that the co-location frequencies for both the citing and the control patents drop further at all geographical levels, but once again the effect is more visible for citing patents. Differences at the regional level diminish: significance of the observed difference in proportions drops from 99% to 95%, and the lower bound for the estimated odds ratio slips under one.

If we restrict our attention only to social connected patents with geodesic distance up to 2, we notice that co-location differences at the city level turn out to be significantly negative at 90% ($P < z$ equals 0.08), and the confidence interval for the estimated odds ratio at the regional level slips further down (box 8.4, table 8)[22].

These results show how also social links between inventors affect the original results obtained by applying the JTH methodology, through a combination of high citation probability and low geographical dispersion. Social networks of inventors spread knowledge across firms, but keep it quite bounded in space.

Once we drop personal self-citations, short social distances between inventors become the most likely origin of a citation link. At the same time, closely connected inventors are also more likely than others to be co-located at any geographical level. As we drop socially connected inventors from our samples, what remained of the original JTH results after excluding for personal self-citations receive a further blow. This effect is more visible if we drop only the most closely connected inventors, such as those separated by no more than two geodesic steps.

In order to further test the effects associated to different types of social connectedness, we have estimated a logit model in which the dependent variable is the geographic match/no match between the citing and the cited patents[23]. As independent variables, we included three sets of dummy variables. First, we introduced two dummy variables to capture the type of connectedness existing between the citing and the cited patent: the variable "social" takes value one if there is a social connection between the inventors in the two patent teams, and zero else; the variable "personal" takes value one if the same cross-firm inventor appears in both patent teams, and zero otherwise. The reference modality is therefore represented by the

21 More precisely, we exclude all the citing patents that turn out to be personally connected, and the related controls; we also exclude all the control patents that turn out to be personally connected, and the citing patents they were meant to match.

22 When removing all personally connected patent pairs from our samples, the citing patents turn out to host not only more patents at geodesic distance 1, but also more patents at distance 2.

23 This exercise is once again similar to that performed by JTH (1993, p. 593).

case of non-connected patents. We expect social and personal connectedness to predict accurately the probability of co-location of citing and cited patents. Second, in order to control for industry location effect, we included a dummy variable that takes value one if the control patent that corresponds to a given citing patent matches geographically the cited one. Third, we introduced dummies for the social connectedness between the control and the cited patents, to control for the possibility that similar technologies exhibit similar social network patterns.

One important reason for conducting this final exercise is that the geographic distribution of patents in Italy is rather asymmetric. In particular, of the 366 originating patents considered in our analysis, 165 (45%) are located in the Lombardy region, and 113 (31%) of them are located in the Milan province. Although the inclusion of control patents aims precisely to capture this pre-existing geographic concentration, in the regressions reported here we have further controlled for this peculiarity of the Italian innovation system. Specifically, for the analysis at the city and province levels, we also included a dummy variable that takes value one if the cited patent is located in the Milan province, while for the analysis at the region level we included a dummy variable that takes value one if the cited patent is located in the Lombardy region.

Results of logit estimations are reported in Table 9. The importance of controlling for the localization of technological activities is confirmed: the co-location of control patents bears always a positive and statistically significant influence at all geographic levels. However, the personal and social connections between cited and citing patent are clearly the most important factor affecting the geographic matching between the citing and the cited patents. Parameters for these variables have a positive and statistically significant value at all geographic levels. Moreover, the magnitude of the coefficients is also remarkable.

Computing the odds-ratios from the estimated logit coefficients indicates that the odds of co-location for citing patents linked to cited ones through the same inventor are more than one hundred times larger than the same odds for non-connected patents, other things being equal[24]. Although social connections bear less influence than cross-firm activity, they also matter a lot: citing patents connected to cited ones in this way exhibit co-location odds four to five times larger than non-connected ones[25]. Controlling for the social connection of control patents gives instead mixed results. On the one hand, the social connection between the control and cited patents is never statistically significant. On the other hand, the parameter for personal connections of controls is significant only in regressions at the region and province levels and it presents a negative sign. This result suggests that when the control and the cited patents are connected because the same inventor appears in both patents it is quite unlikely that the citing and the cited patents match geographically.

Finally, the dummy variable for the Lombardy region is positive and statistically significant, suggesting that citations directed to patents from this region are twice more likely than others to be located in the same region. On the contrary, the dummy variable for the Milan province is only marginally significant in the regression at the province level, and it is not significant in the regression at the city level.

24 The odds ratios can be computed from estimated coefficients produced by logit as follows: $OR = \exp(b_{sc})$, where b_{sc} is the estimated coefficient of the variable for social connection. Odds-ratios are not reported here, but can be easily computed from the data reported.

25 One can also calculate the odds-ratio for personal *vs.* social connection, and notice that the former kind of connection makes the co-location probability at least 20 times more likely than the other.

TABLE 9

Probability of Geographic Matching between Citing and Cited Patents (logit estimates)

	City	Province	Region
Intercept	-2.637 *a)*	-1.716 *a)*	-0.931 *a)*
	(0.225)	(0.170)	(0.163)
Geographic matching control-cited	1.291 *a)*	1.234 *a)*	0.656 *a)*
	(0.362)	(0.265)	(0.224)
Connectedness citing-cited			
Social	1.728 *a)*	1.725 *a)*	1.173 *a)*
	(0.453)	(0.413)	(0.410)
Personal	4.840 *a)*	5.803 *a)*	4.997 *a)*
	(0.525)	(1.063)	(1.089)
Connectedness control-cited			
Social	-0.303	-0.294	0.210
	(0.481)	(0.447)	(0.430)
Personal	-1.424	-2.114 *b)*	-2.165 *b)*
	(0.891)	(1.005)	(0.920)
Dummy for Milan (City and	0.377	0.482 *c)*	0.800 *a)*
Province) / Lombardy (Region)	(0.326)	(0.270)	(0.215)
Number of observations	483	483	483
Number of geographic matches	121	187	260
Log-likelihood	-154.8	-213.1	-259.1
Likelihood ratio statistic χ^2	233.9	218.4	148.4
% of correctly classified cases*	87.1	76.6	61.5

The dependent variable takes value one when citing and cited patents match geographically, zero else. Standard errors among brackets.
a) statistically significant at the 99% level;
b) statistically significant at the 95% level;
c) statistically significant at the 90% level.
* *Cut-off probability value for determining whether an observation has a predicted positive outcome is 0.9.*

6 Conclusions

This paper has brought JTH's recommendation to use patent citations as a paper trail for tracking knowledge flows to some extreme consequences. Patent data have been used to control for the role of "cross-firm inventors", who can both generate personal self-citations (by moving across companies, or by doing research for different companies) and help creating social links across companies, by entering the various teams of inventors.

When controlling for personal self-citations, no localization of knowledge flows remains to be seen at the city or province level. This is because a disproportionate amount of patent citations (even after controlling for self-citations at the company level) come from inventors who may move across firms, but remain immobile in space.

At the regional level, what remains of the difference of co-location frequencies between citing and control patents diminishes sensibly after controlling also for the social ties between inventors from different companies. This happens because social networks of inventors are both localized in space and responsible for many patent citations.

Our results also raise a few substantive issues that deserve to be further discussed and investigated.

In the first place, our results qualify the original intuition of those economists and sociologists that first stressed the tacit content of technological knowledge: knowledge always travel along with people who master it. If those people move away from where they originally learnt, researched, and delivered their inventions, knowledge will diffuse in space. Otherwise, access to it will remain constrained in bounded locations. That is, knowledge flows are localized to the extent that cross-firm activity and the resulting social networks also are localized. Why Italian cross-firm inventors do not relocate in different regions is of course an important question, but one which goes beyond the scope of the present study. We will investigate it by applying our methodology to other countries.

In addition, our results suggest that, as far a large percentage of localized citations are personal self-citations, interpretations of localized knowledge flows as spillovers, that is externalities, ought to be regarded with more caution, and carefully checked. On the one hand, one may suspect that cross-firm inventors may not diffuse the knowledge they master, but move it around the companies they happen to work for. On the other hand, none of our cross-firm inventors is entirely responsible for both the cited and the citing patent in any given pair, as the teams joined by our cross-firm inventors when moving across firms are never the same. This suggests that cross-firm inventors may indeed give up some of their knowledge assets to the companies they work for.

Finally, we deal with the natural objection that our definition of social network includes no more than a tiny subset of all the relevant contacts enabling inventors to achieve their results. We reply to this by suggesting that the population of inventors is more than a tiny and unchecked sample of all the individuals who can influence inventors themselves. Rather, it is the most immediate and influential social environment from which inventors draw ideas and information, at least for the technical contents of their patents.

A limitation of our results comes from the tiny dimension of the Italian innovation system: the patent samples and social networks we may derive from it are quite small, and may be disproportionately affected by a few highly influential individuals, such as our cross-firm inventors. Targeting larger countries, and in particular the US, would allow for a more direct evaluation of our methodological refinement against the original JTH test. ■

References

ALMEIDA P., KOGUT B. (1999) – « Localization of Knowledge and the Mobility of Engineers in Regional Networks », *Management Science*, 45(7), pp. 905-917.

AUDRETSCH D.B. (1998). – « Agglomeration and the Location of Innovative Activity », *Oxford Review of Economic Policy*, 14(2): pp. 18-29.

AUDRETSCH D.B., FELDMAN M.P. (1996). – « R&D Spillovers and the Geography of Innovation and Production », *American Economic Review*, 86(3), pp. 630-640.

BALCONI M., BRESCHI S., LISSONI F. (2004). – « Networks of Inventors and the Location of Academic Research: An Exploration of Italian Data », *Research Policy*, 33(1): pp. 127-145.

BRESCHI S., LISSONI F. (2001). – « Knowledge Spillovers and Local Innovation Systems: A Critical Survey », *Industrial and Corporate Change*, vol. 10, n. 4, pp. 975-1005.

BRESCHI S., LISSONI F. (2004). – « Knowledge Networks from Patent Data: Methodological Issues and Research Targets », in W. Glänzel, H. Moed, U. Schmoch (eds.), *Handbook of Quantitative S&T Research*, Kluwer Academic Publishers.

BRESCHI S., LISSONI F., MALERBA F. (2003). – « Knowledge Relatedness in Firm Technological Diversification », *Research Policy*, 32/1, pp. 69-87.

FELDMAN M.P. (1999). – « The New Economics of Innovation, Spillovers and Agglomeration: A Review of Empirical Studies », *Economics of Innovation and New Technology*, 8, pp. 5-25.

FLEISS J.L. (1981). – « Statistical Methods for Rates and Proportions », *Wiley&Sons*, New York.

GRILICHES Z. (1992). – « The Search for R&D Spillovers », *Scandinavian Journal of Economics*, 94, Supplement, pp. 29-47.

JAFFE A.B. (1989). – « Real Effects of Academic Research », *American Economic Review*, 79, pp. 957-970.

JAFFE A.B., TRAJTENBERG M. (2002). – « Patents, Citations, and Innovations: A Window on the Knowledge Economy », *MIT Press*.

JAFFE A.B., TRAJTENBERG M., FOGARTY M.S. (2000). – « Knowledge Spillovers and Patent Citations: Evidence from a Survey of Inventors », *American Economic Review*, 90(2), pp. 215-218.

JAFFE A.B., TRAJTENBERG M., HENDERSON R. (1993). – « Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations », *Quarterly Journal of Economics*, 10, pp. 577-598.

KRUGMAN P. (1991). – « Geography and Trade », *MIT Press*, Cambridge MA.

MOEN J. (2000). – « Is Mobility of Technical Personnel a Source of R&D Spillovers », *NBER Working Paper*, n. 7834.

MOWERY D.C., ZIEDONIS A.A. (2001). – « The Geographic Reach of Market and Non-Market Channels of Technology Transfer: Comparing Citations and Licenses of University Patents », *NBER Working Papers*, n. 8568.

NEWMAN M.E.J. (2000). – « Who Is the Best Connected Scientist? A Study of Scientific Coauthorship Networks », *SFI Working Paper*, 00-12-64, Santa Fe.

NEWMAN M.E.J. (2001). – « The Structure of Scientific Collaboration Networks », *Proceedings of the National Academy of Science USA*, 98, pp. 404-409.

THOMPSON P., FOX-KEAN M. (2005). – « Patent Citations and the Geography of Knowledge Spillovers: A Reassessment? », *American Economic Review* 95/1, pp. 450-460.

WASSERMAN S., FAUST C. (1994). – « Social Network Analysis: Methods and Applications », *Cambridge University Press*.

ZUCKER L.G., DARBY M.R., ARMSTRONG J. (1998). – « Geographically Localized Knowledge: Spillovers or Markets? », *Economic Inquiry*, 36 (January), pp. 65-86.