Research Policy 39 (2010) 14–26

Contents lists available at ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol

Tracing the links between science and technology: An exploratory analysis of scientists' and inventors' networks

Stefano Breschi[a,*], Christian Catalini[a,b]

[a] KITeS-Cespri, Università L. Bocconi, Milan, Italy
[b] Rotman School of Management, University of Toronto, Canada

ARTICLE INFO

Article history:
Received 25 September 2007
Received in revised form 9 November 2009
Accepted 22 November 2009
Available online 30 December 2009

JEL classification:
O31
D85

Keywords:
Science-technology
Research networks
Knowledge transfer

ABSTRACT

The paper provides an exploratory analysis of the research networks linking scientists working in an open science environment, and researchers involved in the private technology domain. The study combines data on scientific co-authorship with data on patent co-invention, at the level of individual researchers, for three science-intensive technology fields, i.e. lasers, semiconductors and biotechnology, in order to assess the extent of the overlap between the two communities and to identify the role of key individuals in the process of knowledge transfer. Our findings reveal that the extent of the connectedness among scientists and inventors is rather large, and that particular individuals, i.e. authors-inventors, who act as gatekeepers and bridge the boundaries between the two domains, are fundamental to ensuring this connectivity. These individuals tend to occupy prominent positions in the scientific and the technological networks. However, our results also show maintaining a very central position in the scientific network may come at the expense of being able to fill a similarly central position in a technological network (and vice versa). Finally, preliminary analysis of the institutional origins of authors-inventors shows that one characteristic, distinctive of Europe compared to the United States, is associated with the relatively lower involvement of corporate scientists at the intersection between the two worlds of science and technology.

1. Introduction

The relationships between science and technology have long been and continue to be the subject of intense debate both within academia and in the society. The idea that science and technology co-evolve and interact in rather complex ways has replaced the old linear model in which the progress of science was essentially exogenous and technological advances were merely the outcomes of applied research and development (R&D) efforts. Much of the empirical evidence collected since the mid 1980s supports this new thinking in documenting a tightening of the links between technological innovation and scientific research. Moreover, the perception that technological developments and scientific advancements are increasingly interdependent is affecting the design of public policies. Many governments around the world are looking for ways to encourage technology transfer from university to industry, through measures and instruments aimed at supporting academic scientists to assume more entrepreneurial attitudes, particularly through the enforcement of intellectual property rights over their discoveries.

Against this background, the development of new empirical approaches to capture the complex set of interactions between science and technology seems particularly important. Notwithstanding recent advances in the measurement of science-technology linkages, the extent of the connectivity between the communities of scientists and technologists has not been explored in detail. Although scientific and technological research networks represent distinctive social structures responding to different norms of behaviour and different reward systems, it has been noted that they can co-exist and interconnect in various ways (Dasgupta and David, 1994; David et al., 1999). In this respect, mapping the structure of these networks and assessing the degree of overlap between them may contribute greatly to our understanding of the processes underlying knowledge transfer and to the design of better policy instruments to support them. In spite of the recurrent use of network metaphors in much of the ongoing debate on university–industry relationships, there is a need for more systematic empirical analyses of the networks linking scientists working in the realm of open science and researchers involved in the private technology domain. The objective of this paper is to provide an exploratory analysis of the intersection and overlap between these networks and the simultaneous embeddedness of an individual researcher within them. We exploit a large scale data set containing full bibliographic information taken from patent applications and

---

* Corresponding author at: KITeS-CESPRI, Department of Economics, Università L. Bocconi, Via Sarfatti 25, 20136 Milan, Italy. Tel.: +39 02 58363365.
E-mail address: stefano.breschi@unibocconi.it (S. Breschi).

0048-7333/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.respol.2009.11.004

scientific publications cited in patents, for three science-intensive technology fields, i.e. lasers, semiconductors and biotechnology. We examine co-authorship and co-invention data, using the analytical tools of social network analysis, to assess the extent of connectedness among the two communities of researchers and to investigate the position of individual scientists in these research networks. The paper is organised as follows. Section 2 provides a short review of the empirical literature which attempts to trace the linkages between science and technology, with a focus on the role of patenting-publishing scientists. Section 3 describes the basic methodological framework and the strategy adopted to collect the empirical data. Section 4 illustrates and discusses the main findings, Section 5 concludes.

2. Background literature

2.1. Tracing the links between science and technology

The increasing interdependencies and interactions between science and technology have been measured and documented in several studies with the use of different empirical indicators. A line of research pioneered by Francis Narin and colleagues exploits information on references to scientific articles contained in patent documents, which shows an increasing reliance of private technology on public science (Narin and Noma, 1985; Narin et al., 1997; McMillan et al., 2000; Hicks et al., 2001; Tijssen, 2001; Branstetter and Ogura, 2005). A related line of enquiry investigates the patterns of scientific paper co-authorship among academic and corporate scientists, and provides evidence of increasing levels of collaboration across organisational boundaries (Hicks, 1995; Calvert and Patel, 2003; Tijssen, 2004).

A different stream of research attempts to assess the direct contribution of universities to the development of technology and industry competitiveness. In the wake of the Bayh-Dole Act, several US scholars have attempted to estimate the volume of academic patenting showing its impressive growth. The number of patent applications from universities yearly to the US Patent and Trademark Office (USPTO) has grown much faster than applications from business companies and individuals, from less than 100 in the 1960s to more than 3000 at the end of the 1990s (Henderson et al., 1998; Mowery et al., 2001; Sampat et al., 2003). From a different perspective and using a different methodology, some authors explore the contribution of university research to industrial innovation more directly, via interviews or surveys (Mansfield, 1995; Agrawal and Henderson, 2002). Alongside inter-industry differences in the relationship between university and industry, these studies generally show that industry respondents rate academic patents and licences as one of the least effective sources of knowledge, compared to scientific publications, conferences and informal interactions with academic researchers.

In recent times, the focus on university–industry relationships and knowledge transfer seems to have shifted somewhat from a macro (i.e. organisational and institutional) level research to a more micro (i.e. individual) level of analysis. In addition to the greater availability of data, there are two main reasons for this shift. First, there is increased concern over the adoption internationally of legislation emulating the Bayh-Dole Act and the accompanying need to test the effect of academic patenting on the scientific productivity of researchers. Second, scholars are becoming increasingly aware that an exclusive focus in technology transfer studies, on institutional characteristics, may preclude new insights into the channels through which knowledge flows from university to industry and the individual characteristics that may affect the choice to patent (Owen-Smith and Powell, 2001; D'Este and Patel, 2007).

On the first point, since the Bayh-Dole Act was passed in 1980, many observers have raised concerns about the potentially negative effects of the commercialisation of scientific discoveries for the conduct of academic research. It has been argued that the financial incentives from patenting and licensing could shift the orientation of scientists away from basic and towards applied research and could undermine their commitment to the norms of open science, thereby leading to secrecy and publication delays. In order to address these issues, many scholars have begun to compile large data sets matching inventor names with scientific author names, and to collect data on individual researchers' patenting and publication performance (Azagra-Caro et al., 2006; Azoulay et al., 2006; Van Looy et al., 2006; Breschi et al., 2007, 2008; Calderini et al., 2007; Fabrizio and Di Minin, 2008).[1] Although no consensus has been reached, there is evidence that there is no apparent trade-off between patenting and either quantity or quality of research output. Not only do scientists with better patenting performance tend to exhibit superior publication scores with no decrease in the quality of output, but also the most productive scientists are those most likely to become inventors.

A related stream of research focuses on the role of individual scientists in relation to knowledge spillovers from academic research and knowledge transfer from university to industry more broadly. The need for a better understanding of the mechanisms of knowledge flows has sparked several attempts to trace personal links between academic researchers and private firms. In this context, Cockburn and Henderson (1998) argue that the extent of connectedness to the community of open science is a key factor explaining the ability of firms to tap into scientific developments. Establishing linkages with the community of open science may crucially affect their capacity to recognise and effectively exploit upstream developments in basic research. Using qualitative information and data on scientific co-authorship, Cockburn and Henderson show that firms strive to develop this capacity by recruiting and rewarding researchers based on their ranking in the hierarchy of public-sector science and their ability to engage with the academic community. Along similar lines, Zucker et al. (1998, 2002) suggest that the most successful biotech companies are those that engage in co-authorship with university professors and show that their commercial success, in terms of numbers of products developed and commercialised, is positively related to the eminence of the researchers with shareholdings and scientific board membership. Gittelman and Kogut (2003) show the crucial importance to biotechnology firms of maintaining ties with the open science community via boundary-spanning 'gatekeepers' who facilitate access to socially embedded knowledge. Integrating scientific research at the level of individual scientists, i.e. including publishing scientists in teams of inventors, seems to have more of an impact on the quality of the innovative output than firm-level scientific capabilities, measured by the volume and quality of scientific publications.[2] The crucial role of corporate scientists in enabling flows of external knowledge to corporate researchers is emphasised by Furukawa and Goto

---

[1] A further motivation for European researchers to collect data on individual researchers is related to the institutional peculiarities of the academic system in Europe compared to the US. In particular, given the absence until recently of specific legislation promoting university patenting, and the general lack of administrative offices in European universities capable of handling patent applications, inventions by academic scientists are often produced in collaboration with and patented by private companies. Thus, the only way to assess the extent of academic patenting in Europe is to match the names of inventors listed in patent documents to the lists of academic researchers (Balconi et al., 2004; Lissoni et al., 2008).

[2] The results in Cassiman et al. (2008) are slightly different and show that invention-specific science linkages, as measured by citations in patents to scientific publications, are less important for the quality of innovation output than the firm's closeness to science, captured by the volume of scientific publications.

(2006). Their results show that core scientists, defined as corporate researchers whose scientific papers are frequently cited in other papers, while not responsible for a significantly higher number of patent applications than their fellow company researchers, do have a positive impact on the innovative productivity of their co-inventors.

Finally, another line of research has focused on the contractual mechanisms and individual motivations behind the involvement of academic scientists in collaborations with private companies. Stern (2004) suggests that there might be a labour cost advantage to firms hiring scientists as long as these individuals are willing to accept a lower wage in exchange for the opportunity to keep abreast of high quality basic research. Audretsch and Stephan (1996, 1999) and Jensen and Thursby (2001) argue that since scientists' knowledge is characterised by high degrees of natural excludability and, in many cases, academic inventions are disclosed at a proof-of-concept stage, firms need to recruit them as partners or stakeholders in order to gain access to the knowledge held by these individuals and to develop successful licensed inventions.

2.2. *Networks of scientists and inventors*

There is a long tradition in scientometrics of exploiting information on co-authorship of scientific papers to analyse knowledge exchange among researchers, both within and across individual companies and academic research groups, and to investigate social networks of academic scientists (e.g. Kretschmer, 1994; Melin and Persson, 1996; Persson and Beckmann, 1995). The most recent efforts in this tradition draw extensively on graph theory and social network analysis techniques, to show that the scientific co-authorship network is characterised by the structural properties of small world networks (Newman, 2000, 2001, 2004; Wagner and Leydesdorff, 2005). Broadly speaking, a small world network is represented by a graph where the nodes are grouped around tightly linked local cliques, but a relatively small number of steps will connect every node in the network to every other node. This type of structure is thought to be particularly important for both the generation and the diffusion of knowledge. The high degree of density and redundancy of the links within local cliques ensures the formation of a common language and communication codes that enhance reciprocal trust and support the sharing of complex and tacit knowledge among actors; the short cuts linking local cliques to different and weakly connected parts of the network ensure rapid diffusion and recombination of new ideas throughout the network and allow a degree of openness to new sources of knowledge, mitigating the risk of lock-in that could arise in the context of densely connected cliques (Cowan and Jonard, 2004).

Following the renewed interest in networks of collaboration, some studies examine the properties of networks of co-invention exploiting information contained in patent data. An important result from these studies is that social proximity among inventors in collaboration networks is a fundamental driver of knowledge flows, captured by patent citations (Breschi and Lissoni, 2004, 2005; Singh, 2005). At the same time, the co-invention network does not seem to exhibit the structural properties of a small world graph. The largest connected component in the network accounts for a very small fraction of the nodes and the network appears globally sparse, i.e. nodes are scattered across a relatively large number of disconnected components (Fleming et al., 2007).

Another approach taken by some scholars is to combine co-authorship and co-invention data in order investigate the extent to which the two communities – of academic scientists and industrial researchers – are linked. In an in-depth case study of tissue engineering, Murray (2002) proposes a novel methodology based on patent-publication pairs in order to analyse the co-evolution of the co-patenting and co-publishing network. Her results show that the scientific and technical networks remain distinctive and there is little overlap between them. Nevertheless, the contribution of key scientists to both domains is crucial through the creation of networks of co-inventors and co-authors and engagement in a wide range of activities, such as advising, consulting, licensing and establishing new firms. Bonaccorsi and Thoma (2007) investigate teams of inventors in the field of nano-technology, distinguishing among only-inventor patents, i.e. inventors with no scientific publication record, only-author patents, i.e. patents where all inventors have at least one scientific publication, and author-inventor patents, i.e. those named on the patent include inventors with no publications and publishing scientists. Bonaccorsi and Thoma show that author-inventor patents tend to outperform those in the other two categories in terms of patent quality. They interpret this finding as evidence that patenting teams with higher institutional and human capital complementarities are also the most effective at realising and exploiting interactions between science and technology.

Co-authorship and co-invention data should be interpreted with some caution since the rules determining authorship and inventorship can differ (Ducor, 2000). While the status of author of a scientific article is the result of a negotiation process perhaps involving numerous members of a research team and may vary according to the rules prevailing in the specific disciplinary field, the notion of inventorship, at least in principle, has a more precise legal meaning. At the same time, the number of author names on a scientific article is frequently higher than the number of inventors listed in a corresponding patent: Lissoni and Montobbio (2008) analyse this phenomenon empirically, examining 681 patent-publication pairs across different disciplinary fields. Their results show that the first and last named authors have a relatively lower probability of being excluded from the list of inventors on a patent and that senior researchers are also more likely to be retained in the team of inventors.

2.3. *Scope of this study*

Building on the literature reviewed above, we provide an exploratory analysis of the simultaneous embeddedness of researchers in scientific and technological networks. First, we identify the set of scientific papers relevant to a given technological field by exploiting information on citations to the scientific literature contained in patent documents; second, we use co-authorship and co-invention data to investigate the connectivity between the scientific and technological research networks. While most studies have focused on biotechnology and pharmaceuticals, or are limited to a specific national or institutional environment (i.e. academic inventors), the methodology proposed here allows us to generalise some of the previous findings to other, less well explored science-based industries, such as lasers and semiconductors, and to the global set of organisations. We also investigate the structural positions of patenting-publishing scientists in the scientific and technological networks. Given the importance of the role played by these scientists in connecting the two realms, it is crucial to understand their location in the overall web of relations. Are these individuals central in both types of networks, or are they prominent in only one (or none) of the two? How do patenting-publishing scientists compare with their non-patenting non-publishing peers in terms of their network locations? To what extent do patenting scientists play the role of knowledge brokers by spanning across structural holes between disconnected teams of corporate researchers? What are the institutional origins of patenting-publishing scientists? These questions are addressed in the succeeding sections.

3. Data and methodology

3.1. Data sources and matching procedures

The empirical analysis in this paper relies on a large, complex relational data set that combines information on the lists of inventors on patent documents and the lists of authors of the scientific publications cited in those patents. We use social network analysis of co-authorship and co-invention ties to test the degree of connectedness between the two communities. The construction of our data set started with the selection of technological fields for analysis of network links between scientists and inventors. We chose three technological fields, i.e. lasers, semiconductors and biotechnology, which are characterised by a strong reliance on scientific developments and, therefore (at least potentially) involve high levels of interaction among the individuals involved in science and those involved in industrial research. For each of these technology fields, we extracted all patent applications to the European Patent Office (EPO) registered in the period 1990–2003, on the basis of the primary International Patent Classification (IPC) code reported in the patent documents.[3]

For every patent application in these three technology fields, we identified and extracted citations to the so-called non-patent literature (NPL citations). These references derive from the search reports produced by patent examiners to assess the novelty of inventions and delimit the scope of their claims. NPL citations can include scientific articles/journals, books, technical bulletins and manuals or indeed any dated, written disclosure or publication which was made publicly available prior to a patent application (Michel and Bettels, 2001). Although the number of NPL citations in patent documents varies greatly across technology fields, the presence of such references may be taken as an indication of the knowledge indebtedness of the invention to the cited research.[4]

Given that NPL references can include items that are not considered scientific output, we implemented a procedure to identify and select the subset of NPL citations that refer to scientific articles. We used a matching algorithm to pair each NPL citation to the corresponding (if any) scientific article recorded in the Science Citation Index (SCI) data set.[5] Table 1 presents some summary statistics from the resulting data set. It should be noted that the percentage of patents citing scientific literature, and science intensity – defined as the ratio between total number of citations and total number of citing patents – differ across the three technology fields. Whereas approximately four in ten patents in lasers and biotechnology cite SCI papers, in the case of semiconductors the percentage is only 13.5%. This difference is probably due to the higher propensity for technical inventions in semiconductors to rely on types of codified knowledge (i.e. technical bulletins, standards, etc.), which are not typically considered to be ‘science’. While citations to the scientific literature represent a relatively large fraction of all NPL citations in the case of lasers and biotechnology (respectively, 53% and 45%), only some 20% of NPL citations in semiconductors refer to scientific articles.

For the set of patent applications and scientific articles described above, we compared the list of inventor names in the patent document with the list of authors on the cited scientific publications and matched them to identify individuals responsible for both a patented invention and a cited scientific publication.[6] In this respect, we had to deal with a major problem related to the fact that patent data report inventors' first names and last names, whereas the ISI-SCI data set records last name and initials of authors. There is a risk, therefore, in performing a simple matching by surname and first initial, of identifying different individuals as the same person, which could lead to an overestimation of the number of publishing-patenting scientists. To resolve this, we first standardised inventor and author names and surnames and conducted desktop research involving a lot of manual checking and the use of several sources of information. The primary source was author affiliation recorded on the publication and inventor affiliation recorded in the patent document. We also used sources such as SCOPUS,[7] the Internet, and university and company websites. We adopted a conservative approach, only matching two individuals (i.e. author and inventor) if we were reasonably confident that they were the same person. Table 2 reports the total number of inventors, scientific authors and author-inventors (i.e. publishing and patenting scientists) in our data set. Patenting-publishing scientists represent around 21% and 18% of all inventors in lasers and biotechnology respectively, but just 4.5% of all inventors in semiconductors. And the fraction of all scientific authors that also patent is 24% for lasers and 13% for biotechnology.[8]

---

[3] The selected IPC codes are: H01S for lasers, H01L for semiconductors and C12Q, G01N33 (/53,54,55,57,68,74,76,78,88,92) for biotechnology. For this last field, we have followed the definition in OECD (2005) which includes biotechnology applications for measuring, testing and diagnostics.

[4] Also, the average number of NPL references is sometimes seen as a fairly reliable indicator of the extent to which a technology field depends upon or is related to the scientific developments. Schmoch (1993) provides a detailed discussion of the reasons for citing scientific literature in patents. Meyer (2000) discusses the differences between scientific and patent citations.

[5] The SCI data set is a multidisciplinary database, produced by the Institute for Scientific Information (ISI-Thomson), that covers the most important journals in the natural and life sciences, providing information based on more than 5700 peer-reviewed international journals across 178 subject fields. For the present analysis, we used the online version of the database (www.isiknowledge.com/). Notwithstanding some limitations (for a detailed analysis see Callaert et al., 2006) the scientific publications recorded in this database are a satisfactory representation of internationally accepted high-quality “mainstream” basic and applied research, and have become a standard reference for most studies of science-technology linkages. The matching algorithm is rather complex and is decomposed in various stages. In the first stage, we implemented a parsing algorithm in order to separate the string of text containing the NPL reference into appropriate fields (i.e. article title, journal title, author names, etc.). In the second stage, we took the journal titles thus identified and matched them to the list of journal titles covered by SCI. In the third stage, all NPL references in the patents were matched with the source article (if any) contained in the SCI. For each technology field we considered only journals that had received at least 5 patent citations over the period 1990–2003.

[6] It should be noted that we used citations in patents to scientific articles to delimit the boundaries of the scientific community relevant to a certain technology field. Previous studies adopt different approaches to delimiting the set of scientific papers and authors to be associated to a given technology field. These include keyword search strategies (e.g. Meyer, 2006a,b; Bonaccorsi and Thoma, 2007), lexicographic approaches to pairing patents and publications (e.g. Bassecoulard and Zitt, 2004), and hybrid methods to match the titles of patents and scientific publications (e.g. Leydesdorff, 2004). The approach used in this paper is rather conservative. On the one hand, it could exclude a few authors and papers, which, although not cited in patents, have contributed to the generation of the patented technical inventions. On the other hand, by including only authors and papers cited in patents, we avoid the risk of considering authors and papers not related to the technology fields in question. We also conducted a benchmarking analysis in order to determine whether publications cited in patents are cited more widely, i.e. are also (highly) cited by other scientific articles. We compared the average number of citations to publications cited in patents from other publications, with the corresponding average for publications not cited in patents, controlling for publication year, journal title and subject field. Our results show that scientific publications cited in patents belonging to the three technology fields analysed, on average, receive a far larger number of citations from other scientific publications than articles that are not cited in patents. Overall, we believe this result validates our methodology by ensuring that we do not analyse a random and unchecked sample of publications, and are focusing upon probably the most important publications (and authors) in each field. Results of the benchmarking analysis are available at http://ec.europa.eu/invest-in-research/pdf/download_en/final_report_hcp.pdf.

[7] The SCOPUS database produced by Elsevier provides detailed information on institutional affiliations and full first names and last names of authors of scientific publications. It covers around 15,000 peer-reviewed journals, including most of those included in the SCI.

[8] It is worth remarking once again that these fractions should be interpreted in the light of the methodology described above in the text. In particular, it is

Table 1
Patent-publication data set: summary statistics.

	Lasers	Semiconductors	Biotechnology
Number of patent applications	4,057	26,778	13,192
Number of NPL citations	6,994	32,699	30,582
of which citations to scientific articles	3,756	6,700	13,940
Number of cited scientific articles	2,698	5,059	10,448
of which published before 1990	644	1,694	1,394
of which published after 1990	2,054	3,365	9,054
% of patents citing scientific articles	40.3	13.5	39.6
Science intensity·100 (all patents)	92.6	25.0	105.1
Science intensity·100 (only patents citing articles)	229.4	185.2	270.5

The table reports the overall number of patent applications registered in the period 1990–2003 for each of the three technology fields, as well as the total number of NPL citations, the total number of citations to scientific articles and the number of scientific articles cited by these patents. It also reports the fraction of patents citing science, and science intensity defined as the ratio between the total number of citations and the total number of citing patents. Patent applications are dated based on the priority year; cited scientific articles are dated based on publication year.

Table 2
Number of inventors, authors and authors-inventors.

	Lasers	Semiconductors	Biotechnology
Number of inventors	5,962	37,790	26,013
Number of scientific authors	5,115	10,200	36,600
of which authors-inventors	1,231	1,689	4,687
Authors-inventors as % of inventors	20.6	4.5	18.3
Authors-inventors as % of authors	24.1	16.6	12.8

3.2. *Measuring network linkages among authors and inventors*

In order to analyse the network linkages among scientists and inventors, we exploit information on co-authorship and co-invention from our data set. We assume that two inventors (authors) who have collaborated in the production of a patented invention (scientific publication) are connected by a tie, which means that they are linked by some kind of knowledge exchange and have a common knowledge base. Fig. 1 depicts an example of this main idea, which is of a hypothetical network of 17 scientific authors, 13 inventors and 3 author-inventors, identified respectively by the suffixes A, I and AI. Each individual is represented by a node, while the edges between two nodes indicate that these nodes (individuals) have collaborated over a scientific publication or a patented invention.[9] The network consists of two layers. The top layer consists of the authors of scientific publications and their links, i.e. refers to the co-authorship network. The bottom layer consists of the inventors of patents and their links, i.e. refers to the co-invention network. Connectivity between the two layers is realised by individuals (i.e. authors-inventors) connected to both networks through co-authorship and co-invention relationships. Three types of nodes can be identified: only-authors (i.e. individuals that participate in the co-authorship network only), only-inventors (i.e. individuals that participate in the co-invention network only) and authors-inventors (i.e. individuals active in both networks).

Author-inventors play a fundamental role in two respects. On the one hand, they connect the communities of scientists and inventors, and act as gatekeepers or knowledge brokers thereby ensuring more rapid diffusion of knowledge and ideas between domains. On the other hand, they bridge different communities of otherwise disconnected inventors. This idea is illustrated in Fig. 1: if we focus on the co-invention network (i.e. the bottom layer) and disregard the relations arising from co-authorship, we can see that the network of inventors is characterised by three disconnected components.[10] The smallest one is comprised of three nodes [I1,I2,AI3]; the next one has four nodes [I3,I4,I5,AI2], and the largest component has nine nodes [from I6 to I13 plus AI1]. However, if we consider both types of relations simultaneously, we can see that the three components in the co-invention network are connected, albeit indirectly, through the participation of author-inventors in the co-authorship network. Fig. 2 provides a graphical illustration of this, with the solid lines representing direct relations among inventors through co-invention, and the dashed lines representing indirect relations among authors-inventors via the co-authorship network. For example, the largest and the second-largest component identified above are indirectly connected through individuals AI1 and AI2 who have both co-authored with scientist A8.[11] Likewise, the smallest and the second-largest component in the inventors network are also indirectly connected since author-inventors AI2 and AI3 have co-authored, respectively, with authors A8 and A4. These latter have co-authored one or more publications, however, thus there is an indirect bridge between the two separate teams of inventors.

---

important to point out that authors-inventors are defined as researchers that have produced patents in the examined technological fields and that have been authors of scientific articles cited by patents in these fields. Some researchers identified as only-inventors might actually be authors of scientific articles, which are either not cited at all in patent documents or are cited by patents in other technological fields. Similarly, some researchers identified as only-authors might actually be inventors of patents, which however are classified in other technological classes than the ones considered here.

[9] Formally, Fig. 1 is a one-mode projection of a two-mode (affiliation) network (Wasserman and Faust, 1994). Note that the position of nodes, and the length of edges in the network are not significant.

[10] A component of a graph is a subset of nodes (i.e. a subgraph) such that a path (i.e. a sequence of distinct lines and nodes) exists between all pairs of nodes in the subset, but no path between nodes in the subset and other nodes not in the subset.

[11] Note that a similar representation, comprising indirect relations via co-invention, could be applied to the network of co-authorship. It is also important to note that in the example reported in the text, relations among disconnected components are ‘indirect’ in the strict sense of the term. E.g., the largest and the second largest components in the example are indirectly connected to each other because AI1 and AI2 have a common co-author (i.e. A8). In technical terms, the geodesic distance between the two individuals, and therefore between the components they help to connect, is equal to 2 (i.e. the number of edges separating them). Yet, in theory and in the real world, the connection between two inventor components via co-authorship may well be ‘direct’. This would be the case if the two individuals in our example, e.g. AI1 and AI2, had co-authored one or more publications.

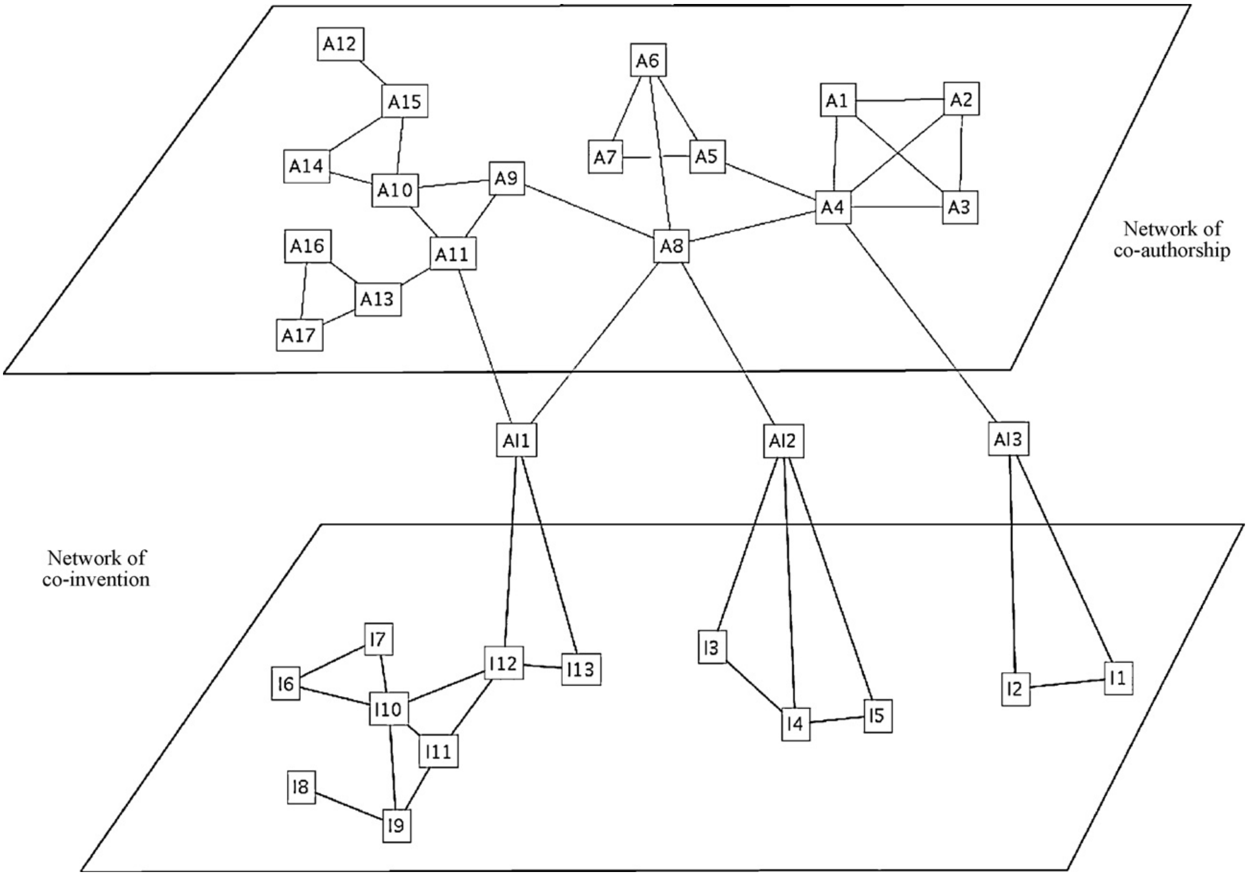

Fig. 1. A hypothetical network of inventors and scientific authors.

4. Structural properties of network

We constructed co-authorship and co-invention networks for each of the three technological fields by taking all patents (and their inventors) registered in the period 1990–2003, and all scientific articles (and related authors) cited by those patents and published in the same period of time. Our analysis focuses on the structural properties of the network which has evolved over the 14 year period covered by the data set.[12] In order to evaluate the extent of connectivity between the networks of co-authorship and co-invention, we constructed three networks: i) one that includes only co-invention ties (i.e. co-invention network); ii) one network that includes only co-authorship ties (i.e. co-authorship network); iii) one that includes simultaneous co-authorship and co-invention ties. For each of the three networks, we calculate three indicators of connectivity among nodes:

1. the largest connected component of the network—in absolute terms and as a percentage of the nodes in the network;
2. the percentage ratio between the number of reachable pairs of nodes and the total number of possible pairs of nodes in the network.[13] This index is defined formally as:

$$R = \frac{\sum_{i \geq j} r_{ij}}{(1/2)n(n-1)} \tag{1}$$

   where $r_{ij} = 1$ if node i can reach node j by a path of any length, and 0 otherwise. In a network with n nodes, the denominator is the total number of possible pairs of nodes;
3. the average geodesic distance among the nodes in the largest component. Average distance measures the number of steps required to connect two randomly selected nodes and is often used as a measure to quantify the efficiency of a network in terms of connecting nodes and facilitating flows of information (Cowan and Jonard, 2004).

Results are reported in Table 3. The first column reports the results of the connectivity analysis by considering only co-invention ties. Note first that the extent of connectivity among inventors is quite low, particularly in lasers and biotechnology where the largest connected components account for a small fraction of all inventors, respectively 13.9% and 9.3%. In other words, the co-invention network appears highly fragmented with a very large of number of small, disconnected components. This interpretation is corroborated by the very low fraction of reachable pairs of inventors, respectively, 1.0% and 2.6%. This is not so surprising given the institutional norms regulating the world of 'private technology'

---

[12] Given the objectives and the exploratory nature of the analysis, here we do not deal explicitly with network dynamics. Network relations formed through co-authorship or co-invention before 1990 are not considered. We also rule out the possibility that relations decay over time. Of course, to the extent that the effectiveness of collaborative ties as conduits of knowledge decays over time, we could have removed older patent applications (and publications) in order to construct a network of social linkages among inventors and scientists. However, in the absence of clear rules to establish the decay of social links, we assume here that all links formed during the period examined are equally effective as means of knowledge transfer. Given the relatively shortness of the time period examined and given the robustness of our empirical findings with respect to the definition of shorter time windows, we think that this assumption is reasonable.

[13] A pair of nodes is said to be reachable if the two nodes are connected by a path of finite length.

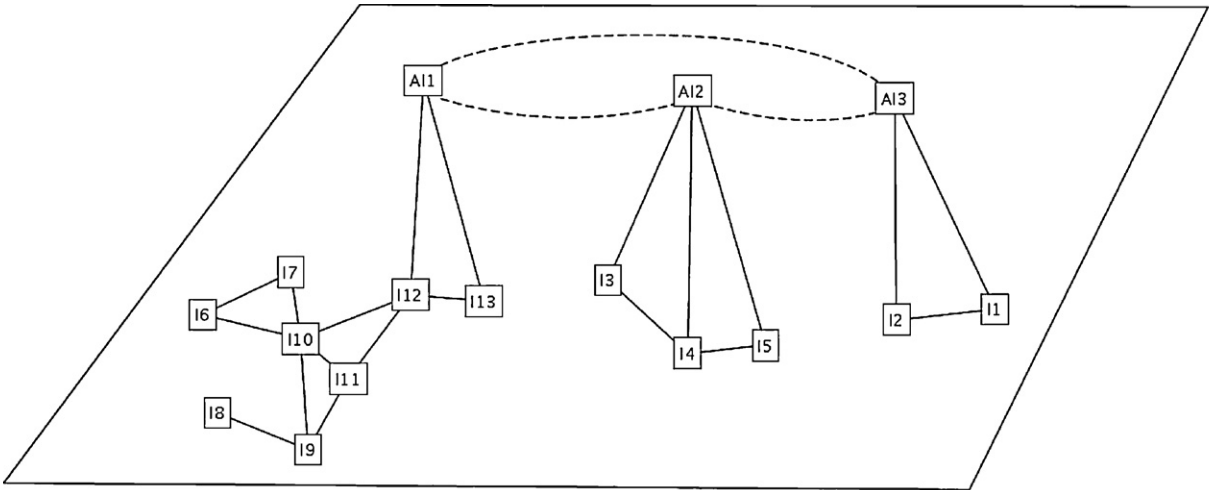

Fig. 2. Network of co-invention including indirect relations via co-authorship.

Table 3
Network connectivity.

	Co-invention	Co-authorship	Co-invention and co-authorship
Lasers			
Largest component (# of nodes)	830	3,185	5,103
as a % of all inventors	13.9		43.6
as a % of all authors		62.3	68.0
as a % of authors-inventors	24.8	70.1	79.4
Reachable pairs of inventors	2.6	–	22.4
Reachable pairs of authors	–	39.7	47.2
Average distance	9.9	7.9	8.2
Semiconductors			
Largest component (# of nodes)	12,652	5,803	20,478
as a % of all inventors	33.8		39.6
as a % of all authors		56.9	68.4
as a % of authors-inventors	49.1	54.3	79.5
Reachable pairs of inventors	13.5	–	18.5
Reachable pairs of authors	–	33.1	47.7
Average distance	13.7	8.9	11.4
Biotechnology			
Largest component (# of nodes)	2,370	22,900	31,789
as a % of all inventors	9.3		41.0
as a % of all authors		62.6	68.0
as a % of authors-inventors	11.1	63.9	75.9
Reachable pairs of inventors	1.0	–	18.4
Reachable pairs of authors	–	39.5	46.6
Average distance	11.4	7.3	7.9

(Dasgupta and David, 1994). The strategic importance to private companies of keeping knowledge mainly proprietary is likely to discourage both cross-organisational collaboration among inventors and mobility across teams of researchers. Given the weakness of such 'connection' mechanisms, the fragmentation characterising the co-invention network appears a rather natural consequence. A partial exception to this pattern is the field of semiconductors where the largest connected component accounts for around 34% of all inventors, although the average distance among them is quite large. There are two possible factors explaining the relatively greater connectivity in the co-invention network of semiconductors. First, in this field, overall patenting activity and, thus, teams of inventors are more concentrated in a smaller number of companies than in lasers and biotechnology. Second, the rate of mobility of engineers across companies and teams of inventors is quite high in semiconductors (Almeida and Kogut, 1999; Palomeras, 2005).[14] Taken together, these two factors ensure that disconnected teams of inventors are more likely to become connected by the mobility of inventors both within and across organisations.

For the co-authorship network, our results show that this is far better connected than the co-invention network (column 2 in Table 3). The largest component comprises a large fraction of all scientific authors—57% in semiconductors, around 62% in lasers and 64% in biotechnology. Likewise, the fraction of reachable pairs of authors is quite high in all three fields. Again, this broad picture is consistent with other studies of scientific networks (Newman, 2001) and arguments in the new economics of science (Dasgupta and David, 1994). The existence of one giant component connecting most of the nodes seems to be the expected outcome of the incentive structure and the norms of openness and collaboration prevailing in the realm of open science. At the same time, it is interesting to note that the vast majority of author-inventors in the

---

[14] In a sample of IBM semiconductor engineers, Palomeras (2005) shows that movers represent 15% of all inventors in the sample, but account for 33% of all patents. This implies that the most productive engineers are also those most likely to move within and across firms. As highly productive engineers tend to have a larger number of connections with other inventors, this likely explains the relatively high connectivity in the co-invention network.

fields examined are included in the largest connected component of the co-authorship network, thereby indicating that patenting-publishing scientists are well integrated within the open science community.

The last column in Table 3 reports the measures of network connectivity for the graph including simultaneous co-invention and co-authorship ties. The most important point here is that the extent of connectedness among inventors increases dramatically, particularly in lasers and biotechnology. The share of all inventors in the largest connected components goes from 13.9% to 43.6% in lasers, from 9.3% to 41.0% in biotechnology, and from 33.8% to 39.6% in semiconductors. Similarly, the fractions of reachable pairs of inventors increase from 2.6% to 22.4% in lasers, from 1.0% to 18.4% in biotechnology, and from 13.5% to 18.5% in semiconductors. Overall, these results seem to indicate the existence of a much higher degree of connectivity and a much larger set of communication channels among teams of industrial researchers than might be assumed. The key point is that this connectivity takes place through the indirect linkages created via the co-authorship network, and the existence of this knowledge transfer mechanism has been hidden thus far by the fact that the co-invention and co-authorship networks have been examined separately. Given the recent interest in the small world properties of the co-invention network (e.g. Fleming et al., 2007; Bettencourt et al., 2007), the message that emerges from this result is that the extent of the social links among inventors measured by co-invention ties is likely to grossly underestimate the extent of the social relations in which industrial researchers are embedded. Thus, accounting explicitly for the indirect linkages created via the co-authorship network may greatly improve our ability to capture the channels through which knowledge flows occur.

There are two other points to note. First, the fractions of both scientific authors and authors-inventors in the largest connected component increase when co-invention ties are added to co-authorship links. This implies that these types of nodes are also indirectly linked via co-invention (see fn. 11). However, the additional connectivity created through these indirect links is relatively lower than that observed for inventors when we account for co-authorship ties, particularly in lasers and biotechnology. For example, in the case of lasers, we observe that 70.1% of all author-inventors are included in the largest component of the co-authorship network. If we also take into account co-invention ties, this fraction increases only to 79.4%. Second, the average distance among nodes in the largest component of the network comprising both co-authorship and co-invention ties, presents the generally low values typical of small world networks (Newman, 2001), especially in lasers and biotechnology.[15] Overall, these results suggest the existence of a relatively high degree of connectedness between the two communities of researchers: scientific authors and industrial inventors are not only connected to each other, the distance between them is relatively short, thereby ensuring (at least potentially) a rapid diffusion of knowledge from one realm to the other.

4.1. Network position of authors-inventors

In this section, we examine the structural positions in the network of specific types of nodes. In particular, we want to test to what extent author-inventors occupy a more prominent position than the other two groups (i.e. only-authors and only-inventors) in the two networks in which they are simultaneously embedded. The results reported above suggest that author-inventors play a crucial role in ensuring a high degree of connectivity between the communities of scientific and technological researchers. Individuals that both publish scientific articles and patent new inventions bridge the academic and industrial worlds enabling access by their industry partners, to new scientific knowledge and methods. By connecting with authors-inventors, industrial researchers (i.e. inventors) can keep track of the scientific advances relevant to their activities. Thus, we would expect that author-inventors' embodiment of stocks of valuable, tacit knowledge will make them more attractive partners for other inventors. In social network analysis terms, we would expect that author-inventors will be more central and more 'in between' than only-inventors. A similar mechanism may be at work on the side of science. By co-inventing with industrial researchers, author-inventors may gain access to a larger pool of resources, both financial and material (i.e. costly equipment and instrumentation), and ideas, which may positively affect their ability to build larger teams of scientific researchers and attract other scientists for collaboration. To test these hypotheses, we computed two measures of network centrality (for details, see Wasserman and Faust, 1994):

1. betweenness centrality: this is an index that is widely used to assess the extent to which a node occupies a central position in the information flows within a network. Formally, the betweenness centrality of a node i is defined as the share of the shortest paths connecting each pair of nodes j and k that pass through i:

$$C_B(n_i) = \sum_{j<k} \frac{g_{jk}(n_i)}{g_{jk}} \tag{2}$$

   where g_{jk} is the number of shortest paths linking nodes j and k, and $g_{jk}(n_i)$ is the number of such paths between j and k that contain node i. Betweenness centrality measures how many times a node lies "between" two others, such that it must be activated to enable knowledge exchanges among them. It is a measure of 'gatekeeping' as it captures the importance of a node to all other nodes, as a channel of information. As the value of betweenness centrality depends on the number of nodes in each connected component, we standardised this index to values between 0 and 100.[16]

2. closeness centrality: this is defined as the inverse of the average distance between a node and all other nodes reachable from it. Formally:

$$C_C(n_i) = \frac{(n-1)}{\sum_{j=1}^{n} d(n_i, n_j)} \times 100 \tag{3}$$

where $d(n_i,n_j)$ is the geodesic distance (i.e. shortest path) linking nodes i and j and $(n-1)$ is the number of vertices that are reachable from node i. The index captures the proximity of an actor to all other actors in the network and can be interpreted as a measure of how long it takes for information to spread from a given node to other reachable nodes in the network (and vice versa). The value of the

---

[15] In the 3 fields examined here, the average distance is slightly larger than might be expected in a random graph with the same number of vertices and average degree of nodes. However, it should be noted that the shortest possible distance among only-inventors and only-authors, by construction, is always equal to 2, as both types of actors are necessarily connected through the intermediation of an author-inventor.

[16] If s_k indicates the number of nodes in component k, the maximum theoretical value of betweenness centrality is given by $(s_k-1)(s_k-2)/2$, which arises when an actor falls on all geodesics linking all other actors in that component. Standardised betweenness is computed therefore by dividing the value of (2) by $(s_k-1)(s_k-2)/2$ and multiplying it by 100. The range of the index is from 0 to 100. It is useful here to refer to the example reported in Fig. 1 which shows that the overall network, considering both co-invention and co-authorship ties, contains one largest connected component. However, if we focus only on the co-invention network there are three distinct components. The size of these components is $s_1 = 9$, $s_2 = 4$ and $s_3 = 3$. Taking node I9, we can see that its betweenness centrality is equal to 7, as it lies on all the shortest paths connecting node I8 to all other nodes. The standardised betweenness centrality of node I9 is therefore equal to $(7/28)\cdot 100 = 25$.

Table 4
Betweenness and closeness centrality in co-invention and co-authorship networks.

		Authors-inventors	Only inventors	Only authors	Wilcoxon–Mann–Whitney *z*-score
Lasers					
Co-invention	Betweenness	7.95 (0.00)	2.53 (0.00)	–	10.387
	Closeness	45.31 (39.61)	38.48 (36.44)	–	4.592
Co-authorship	Betweenness	1.02 (0.002)	–	0.33 (0.00)	15.669
	Closeness	20.58 (13.88)	–	17.22 (13.18)	6.334
Semiconductors					
Co-invention	Betweenness	4.10 (0.006)	0.58 (0.00)	–	17.627
	Closeness	27.86 (9.18)	12.55 (7.91)	–	13.044
Co-authorship	Betweenness	1.84 (0.00)	–	0.47 (0.00)	9.432
	Closeness	31.12 (12.94)	–	18.75 (12.20)	7.826
Biotechnology					
Co-invention	Betweenness	7.54 (0.00)	1.94 (0.00)	–	18.892
	Closeness	58.18 (57.14)	42.48 (36.79)	–	21.405
Co-authorship	Betweenness	1.15 (0.00)	–	0.16 (0.00)	33.005
	Closeness	23.53 (14.47)	–	17.86 (14.40)	4.415

The table reports average values for betweenness and closeness centrality for each of the co-invention and the co-authorship networks. Median values of the two indicators are reported in parentheses.

index lies between 0 and 100; the value is highest when an actor is adjacent to all other actors and the denominator is thus equal to $(n-1)$. When interpreting the results, it is important to note that higher values of the index mean that an actor is closer to, i.e. less distant from, all other nodes in the network.[17]

We calculated separate centrality measures for the co-invention and co-authorship networks. From Fig. 1 it is clear that authors-inventors – almost by definition – are more central than other types of nodes when we consider the overall network formed by co-authorship and co-invention ties. For this reason, we calculate the centrality of each node with reference to the specific component in which it is located in either the co-invention or the co-authorship network. In making this calculation, we consider only the subset of nodes that are included in the largest connected component in the overall network of co-invention and co-authorship ties.

Table 4 reports the average values for betweenness and closeness centrality respectively, for the authors-inventors and the only-inventors and only-authors groups. Since the variables in both cases are not normally distributed and tend to be highly skewed, we ran the non-parametric Mann–Whitney–Wilcoxon test to assess whether the sample of author-inventors comes from the same distribution as the only-inventors or only-authors. The results in the last column of Table 4 reject this hypothesis in all cases, at the conventional levels of statistical significance, thereby suggesting that author-inventors, on average, are more ‘in between’ and more centrally positioned than their peers in either the co-invention or the co-authorship networks.

To provide further support for this interpretation, we examined the shape of the distribution. For each network, we ranked the nodes according to their betweenness and closeness centrality values and grouped them into percentiles, focusing on nodes in the top 75%, 90% and 95% of the distribution. For each percentile class, we evaluated to what extent authors-inventors are disproportionately represented among the most central nodes by computing the ratio of share of authors-inventors on all nodes included in that percentile class, and share of authors-inventors on all nodes. An index greater than 1 indicates that author-inventors are over-represented in the nodes in a given percentile class of the centrality distribution compared to their overall share of nodes in the network. The results are reported in Table 5 and broadly confirm that author-inventors are far more represented in the more central nodes in both the co-invention and co-authorship networks than might be expected based on their overall weight. Moreover, the extent of over-representation increases with the upper percentiles of the distribution, suggesting that author-inventors are more likely to be found in the right-tail of the distribution (i.e. among the most prominent nodes) than only-inventors or only-authors. Overall, this suggests that author-inventors not only play a crucial role in bridging the gap between the two communities of scientists and inventors, they also occupy strategically important positions within each community. Far from being peripheral and marginalised actors, author-inventors are highly central players in both domains, contributing to rapid diffusion of knowledge and ideas between domains.

However, this conclusion requires some qualification since the previous analysis does not tell us whether it is the same individuals that are central in both networks. It is possible that, although author-inventors are disproportionately more represented in the most central nodes in each network, the author-inventors occupying these positions in the two networks are different. To test this idea, we computed a simple rank correlation coefficient of authors-inventors according to the value of betweenness and closeness

Table 5
Authors-inventors in the top percentiles of the centrality distribution.

		p75	p90	p95
Lasers				
Co-invention	Betweenness	1.47	1.70	1.94
	Closeness	1.41	1.44	1.42
Co-authorship	Betweenness	1.00	1.92	2.03
	Closeness	1.31	1.48	1.49
Semiconductors				
Co-invention	Betweenness	1.82	2.58	3.43
	Closeness	1.86	2.96	3.96
Co-authorship	Betweenness	1.00	1.54	1.95
	Closeness	1.49	2.22	2.36
Biotechnology				
Co-invention	Betweenness	1.43	1.88	2.13
	Closeness	1.52	1.63	1.63
Co-authorship	Betweenness	1.00	2.16	2.46
	Closeness	1.16	1.76	2.24

For each percentile class the table reports the ratio between share of authors-inventors in all nodes in that class and share of authors-inventors in all nodes. An index greater than 1 indicates that authors-inventors are over-represented among the nodes in a given percentile class.

---

[17] Referring again to Fig. 1, the closeness centrality of node I9 in the co-invention network is equal to $(8/15)\cdot 100 = 53.3$.

Table 6
Rank correlation coefficient of betweenness and closeness centrality in the co-invention and in the co-authorship networks.

	Betweenness (a)	Betweenness (b)	Closeness
Lasers	0.181	−0.186	−0.145
Semiconductors	0.082*	−0.299	−0.087
Biotechnology	0.182	−0.200	−0.034*

The table reports the rank correlation coefficient of betweenness and closeness centrality in the co-invention and co-authorship networks of authors-inventors. Column (a) includes all authors-inventors, column (b) excludes authors-inventors with zero betweenness centrality in both networks. All correlations are statistically significant at the 1% level, except * significant at the 5% level.

centrality in the co-invention and in the co-authorship networks. The results are reported in Table 6. The evidence is mixed and does not lend itself to easy interpretation. Column 1 reports the value of the correlation coefficient for betweenness centrality considering all authors-inventors; the positive and statistically significant value of the coefficient seems to indicate the existence of a broadly positive, although weak, association between the rankings of authors-inventors in the two networks, at least for lasers and biotechnology. However, if we exclude from the sample those author-inventors with zero betweenness centrality in both networks (column 2), the rank correlation coefficient of the two variables becomes negative and statistically significant, indicating that author-inventors that are highly ranked in one of the networks tend to be relatively lower ranked in the other. In terms of closeness centrality, the evidence indicates the existence of a weakly negative rank correlation between the positions of authors-inventors in the co-invention and co-authorship networks. Overall, these results would suggest that maintaining a highly central position in the scientific network may come at the expense of being able to occupy a similarly very central position in the technological network (and vice versa). It is plausible that this may be related to the institutional affiliations of different author-inventors. More specifically, we could test the hypothesis that corporate author-inventors tend to be more centrally located in the co-invention network than author-inventors from public-sector science. However, for the reasons explained below, this type of analysis is quite complex and is beyond the scope of this paper. Thus, in what follows, our analysis is limited to an exploratory investigation of the institutional affiliations reported in the articles of the most prominent authors-inventors.

4.2. *Institutional affiliation of authors-inventors*

Given the importance of authors-inventors in bridging between science and technology, we are interested in their institutional affiliations. Assigning institutional affiliation to the authors-inventors sample is not a trivial task. First, the affiliations reported in patent documents may reflect the patent right assignees, rather than the actual affiliation of the inventor, especially in the case of Europe (Lissoni et al., 2008). Second, the SCI database reports the affiliations for all authors of an article, but it does not link individual authors to institutions. Although there are partial solutions to this problem, such as assuming that an affiliation applies to an author if she is listed as the corresponding author, the sole author, or only one affiliation is indicated, this still entails a degree of arbitrariness.[18] Thus, we avoided making any assumptions about author-inventors' institutional affiliations. Instead, we looked at the list of all the institutions reported in their scientific articles categorising them into: companies, universities and public research organisations (PROs).[19] This allowed us to classify each scientific article by an author-inventor according to the mix of institutional categories reported in the list of affiliations. Also, we restricted attention to scientific articles produced by European and US authors-inventors and, among these, to the subset of articles by the top 25% authors-inventors based on betweenness centrality in the overall network formed by co-invention and co-authorship ties.[20] This is because these are the scientists who are most likely to act as gatekeepers between the realms of science and technology.

The results in Table 7 show that the share of articles by European authors-inventors where the only reported affiliation is a company, is much lower than the corresponding share for US authors-inventors. The picture is much the same if we include articles produced jointly by companies and other institutions: the share of articles by European authors-inventors with at least one company affiliation is 48.6%, 31.2% and 18.5%, respectively for lasers, semiconductors and biotechnology. The corresponding figures for the US are 69.4%, 65.2%, and 41.3%. Conversely, the share of the public scientific research system, i.e. universities and PROs, is significantly higher for Europe than the US: the share of articles by European authors-inventors affiliated to either a university or a PRO is 50.8%, 60.4%, 57.6%, respectively, for lasers, semiconductors and biotechnology, while the corresponding figures for the US are 30.1%, 31.0%, and 40.7%. In particular, large PROs in Europe – such as Commissariat à l'Énergie Atomique (CEA), Centre National de la Recherche Scientifique (CNRS), Max-Planck Gesellschaft, Fraunhofer Gesellschaft, Interuniversity Microelectronics Center (IMEC), Institut national de la santé et de la recherche médicale (INSERM), Institut Pasteur, and European Molecular Biology Laboratory (EMBL) – seem to account for a remarkably larger share of authors-inventors than their counterparts in the US, particularly in semiconductors and biotechnology. To the extent that the ability of private companies to profit from scientific output generated in the scientific community depends on the level of absorptive capabilities and, especially, on the existence of boundary-spanning individuals, the fact that a remarkably large fraction of the scientific activity of European authors-inventors takes place within the boundaries of universities and PROs might be a major factor explaining the apparent weakness of Europe in achieving effective knowledge transfer from science to industrial applications. However, we are reluctant to draw this conclusion for several reasons.[21] Our findings may reflect the relatively smaller size of

---

[18] The SCOPUS database (http://www.scopus.com/) includes the institutional affiliations of authors of scientific articles. However, this affiliation is their current affiliation and since our analysis refers to 1990–2003, we cannot exclude the possibility that the affiliations of the individuals in our sample may have changed.

[19] This last group includes government laboratories, public and private non-profit research organisations, non-teaching hospitals, etc. Although this classification method encompasses some arbitrariness, we believe that it does not have a major impact on the results. In only very few cases we were unable to classify the type of institution; for the vast majority of papers the type was unambiguous.

[20] To locate authors-inventors, we used the inventor addresses reported in patent documents, as these are most likely to reflect the place of work of the authors-inventors. Europe includes the EU-15 member states; US includes the United States and Canada. We also include in the analysis only articles by authors-inventors with reported affiliations in either Europe or the US. In other words, articles with authors with affiliations including Europe *and* the US were excluded. These latter types of co-authored articles represent less than 5% of all articles by European and US authors-inventors in lasers and semiconductors. Biotechnology is somewhat of an exception with around 20% of the articles by European and US authors-inventors in our sample being internationally co-authored. This is another reason for some caution in interpreting these results.

[21] We want to thank two anonymous referees for drawing attention to alternative interpretations of our results: first, we should point out that our analysis is based on European patent data. Thus, the results might simply be a reflection of the fact that US-based universities and PROs have fewer incentives than US-based companies to extend their patent applications to Europe (although recent evidence shows that US universities tend to dominate their European counterparts also in terms of owned EPO patents, see Van Looy, 2009) Second, we should take account of the different

Table 7
Distribution of scientific articles of the most central authors-inventors by institutional affiliation.

	Lasers	Semiconductors	Biotechnology
Europe			
Company	45.7	21.0	5.9
Company-University	2.3	6.6	6.1
Company-PRO	0.6	3.6	2.6
PRO	9.8	24.5	29.5
PRO-University	0.6	8.4	23.8
University	41.0	35.9	28.1
Company-PRO-University	–	–	3.9
Total	100.0 (77/173)	100.0 (138/167)	100.0 (303/538)
United States			
Company	63.8	55.3	23.1
Company-University	4.7	8.7	10.5
Company-PRO	0.9	0.7	2.8
PRO	4.4	1.6	11.4
PRO-University	0.5	3.9	18.0
University	25.7	29.4	29.3
Company-PRO-University	–	0.5	4.9
Total	100 (107/428)	100.0 (190/439)	100.0 (560/1140)

The table reports the percentage distribution of scientific articles produced by the top 25% of authors-inventors based on betweenness centrality, in the largest component of the overall network formed by co-authorship and co-invention ties, according to the mix of institutional affiliations reported in their articles. Numbers of authors-inventors/numbers of articles analysed are reported in parentheses.

major European corporate players compared to US ones, the lack of integration between the different national innovation systems and differences in labour mobility. Thus, although the preliminary findings reported here suggest a relationship with the controversial issue of the alleged inability of European companies to translate high-quality scientific output into profitable innovations (Dosi et al., 2006), we would recommend further research to provide more robust conclusions.

5. Discussion and conclusions

The preliminary findings from this exploratory analysis of the research networks linking the communities of scientists and inventors potentially help to improve our understanding of the processes driving the transfer of knowledge from science to industry, and open up several avenues for further research. First, our results show that, in spite of the different objectives and incentive structures, the two communities of researchers are connected to a relatively large extent. In particular, the structural properties of the network formed by co-authorship and co-invention ties may facilitate the spread of knowledge from one realm to the other. However, it should be pointed out that a pure structural analysis, such as described in this paper, is not informative about the extent to which knowledge actually travels via the social ties examined here. In this regard, a fruitful avenue for future research might be to investigate whether and to what extent the ability of firms to exploit scientific knowledge depends on the position of their researchers within the scientific and technological research network. At the same time, our results provide a caveat to the use of data on co-invention ties. Given the recent interest in the small world properties of the co-invention network, an important message from our results is that the extent of the social linkages among inventors measured by co-invention ties, is likely to grossly underestimate the extent of the social relations in which industrial researchers are embedded. By neglecting the indirect links among inventors created via the co-authorship network, structural analyses of co-invention networks are likely to provide a severely distorted picture of the extent of connectedness among industrial researchers.

Second, certain individuals, i.e. authors-inventors, play a key role in connecting the scientific and technological research communities, by acting as gatekeepers that bridge the boundaries between the two domains. The analysis in this paper shows that such individuals occupy prominent positions in both the scientific and the technological networks, fulfilling the crucial function of knowledge brokers between the two domains. However, our findings also show that maintaining a highly central position in the scientific network comes at the expense of being able to locate in a similarly central position in the technological network (and vice versa). In this respect, our analysis also suggests that a major weakness in Europe in achieving more effective knowledge transfer might be related to the relatively weak involvement of European corporate inventors in the world of open science and to a lack of researchers able to span the boundaries between science and technology. In order to test this conjecture, however, a further data collection and analysis effort would be needed to check the institutional affiliations of authors-inventors. A promising avenue for further research might be to examine when and under what circumstances publishing scientists are likely to be involved in the development of patented inventions.

Finally, we would stress the limitations of our study. The first, perhaps not so obvious limitation refers to the cost of conducting a large scale analysis of co-authorship and co-invention networks. Collecting and cleaning data on a large number of individual researchers involves huge amounts of time and resources. Examining a small sample of individuals may not be feasible if the purpose is to analyse the structure of the entire network. A very important concern in social network studies is how to define the network boundaries, i.e. which actors to include. The key point here is that if the boundaries are defined in a too restrictive way, in order, e.g., to reduce the size of the sample, this will risk excluding some important actors and ties and could produce a biased picture. The present study is not completely free of these problems. Although justifiable on methodological and theoretical grounds, we cannot rule out the possibility that the strategy we adopted to delimit the boundaries of the relevant networks, i.e. by analysing only scientific papers cited by patents in a specific technological field, excludes some relevant articles and scientific authors. Finally, it is important to point out that social network analysis based on bibliometric indicators

---

practices between Europe and the US, concerning the way that author-inventors assign institutional addresses in their publications. For instance, if US academic and PRO scientists who conduct research leading to a patent, in a firm's facilities, use the company address in a resulting publication (which also avoids problems related to the ownership of intellectual property rights), this might explain the larger fraction of corporate articles in the US than in Europe (Zucker and Darby, 2001).

cannot substitute for more in-depth analysis of the different, often informal mechanisms through which science and technology interact. At the same time, we believe that careful use of the tools offered by social network analysis and equally careful interpretation of results could make an important contribution to the design of more qualitative studies aimed at capturing the subtleties involved in the interactions between the two worlds.

Acknowledgements

We gratefully acknowledge financial support from Bocconi University, and from the DG Research, European Commission, service contract PP-CT-M2-2004-0005. This paper is based on the results of our research and analysis and does not necessarily reflect the views of the European Commission. Thanks are due to Gianluca Tarasconi, Paolo Guatta, Hrannar Johnson and Lorenzo Novella who provided invaluable research assistance. Helpful comments were received from three anonymous referees. Any remaining errors are our own.

References

Agrawal, A., Henderson, R., 2002. Putting patents in context: exploring knowledge transfer from MIT. Management Science 48 (1), 44–60.
Almeida, P., Kogut, B., 1999. Localization of knowledge and the mobility of engineers in regional networks. Management Science 45 (7), 905–917.
Audretsch, D.B., Stephan, P.E., 1996. Company-Scientist locational links: the case of biotechnology. American Economic Review 86 (3), 641–652.
Audretsch, D.B., Stephan, P.E., 1999. Knowledge spillovers in biotechnology: sources and incentives. Journal of Evolutionary Economics 9 (1), 97–107.
Azagra-Caro, J., Carayol, N., Llerena, 2006. Patent production at a european research university: exploratory evidence at the laboratory level. The Journal of Technology Transfer 31 (2), 257–268.
Azoulay, P., Ding, W., Stuart, T., 2006. The impact of academic patenting on the rate, quality, and direction of (Public) research output. National Bureau of Economic Research Working Paper Series, No. 11917.
Balconi, M., Breschi, S., Lissoni, F., 2004. Networks of inventors and the role of academia: an exploration of Italian patent data. Research Policy 33 (1), 127–145.
Bassecoulard, E., Zitt, M., 2004. Patents and publications: the lexical connection. In: Glanzel, W., Moed, H., Schmoch, U. (Eds.), Handbook of Quantitative S&T Research: The Use of Publication and Patent Statistics in Studies of S&T Systems, Springer Verlag, Berlin, pp. 665–694.
Bettencourt, L.M., Lobo, J., Strumsky, D., 2007. Invention in the city: increasing returns to patenting as a scaling function of metropolitan size. Research Policy 36 (1), 107–120.
Bonaccorsi, A., Thoma, G., 2007. Institutional complementarity and inventive performance in nano science and technology. Research Policy 36 (6), 813–831.
Branstetter, L., Ogura, Y., 2005. Is academic science driving a surge in industrial innovation? Evidence from patent citations. National Bureau of Economic Research Working Paper Series, No. 11561.
Breschi, S., Lissoni, F., 2004. Knowledge networks from patent data: methodological issues and research targets. In: Glanzel, W., Moed, H., Schmoch, U. (Eds.), Handbook of Quantitative S&T Research: The Use of Publication and Patent Statistics in Studies of S&T Systems, Springer Verlag, Berlin, pp. 613–643.
Breschi, S., Lissoni, F., 2005. Cross-firm inventors and social networks: localised knowledge spillovers revisited. Annales d'Economie et de Statistique 79–80, 1–29.
Breschi, S., Lissoni, F., Montobbio, F., 2007. The scientific productivity of academic inventors: new evidence from Italian data. Economics of Innovation and New Technology 16 (2), 101–118.
Breschi, S., Lissoni, F., Montobbio, F., 2008. University patenting and scientific productivity: a quantitative study of Italian academic inventors. European Management Review 5 (2), 91–109.
Calderini, M., Franzoni, C., Vezzulli, A., 2007. If star scientists do not patent: the effect of productivity, basicness and impact on the decision to patent in the academic world. Research Policy 36 (3), 303–319.
Callaert, J., Van Looy, B., Verbeek, A., Debackere, K., Thijs, B., 2006. Traces of Prior Art: an analysis of non-patent references found in patent documents. Scientometrics 69 (1), 3–20.
Calvert, J., Patel, P., 2003. University-industry research collaborations in the UK: bibliometric trends. Science and Public Policy 30 (2), 85–96.
Cassiman, B., Veugelers, R., Zuniga, P., 2008. In search of performance effects of (in)direct industry science links. Industrial and Corporate Change 17 (4), 611–646.
Cockburn, I.M., Henderson, R.M., 1998. Absorptive capacity, coauthoring behavior, and the organization of research in drug discovery. Journal of Industrial Economics 46 (2), 157–182.
Cowan, R., Jonard, N., 2004. Network structure and the diffusion of knowledge. Journal of Economic Dynamics and Control 28 (8), 1557–1575.
Dasgupta, P., David, P.A., 1994. Toward a new economics of science. Research Policy 23 (5), 487–521.
David, P., Foray, D., Steinmueller, E.W., 1999. The research network and the new economics of science: from metaphors to organizational behaviors. In: Gambardella, A., Malerba, F. (Eds.), The Organization of Economic Innovation in Europe. Cambridge University Press, Cambridge, pp. 303–342.
D'Este, P., Patel, P., 2007. University-industry linkages in the UK: what are the factors underlying the variety of interactions with industry? Research Policy 36 (9), 1295–1313.
Dosi, G., Llerena, P., Labini, M.S., 2006. The relationships between science, technologies and their industrial exploitation: an illustration through the myths and realities of the so-called [']European paradox'. Research Policy 35 (10), 1450–1464.
Ducor, P., 2000. Coauthorship and coinventorship. Science 289 (5481), 873–875.
Fabrizio, K.R., Di Minin, A.D., 2008. Commercializing the laboratory: faculty patenting and the open science environment. Research Policy 37 (5), 914–931.
Fleming, L., King, C., Juda, A.I., 2007. Small worlds and regional innovation. Organization Science 18 (6), 938–954.
Furukawa, R., Goto, A., 2006. The role of corporate scientists in innovation. Research Policy 35 (1), 24–36.
Gittelman, M., Kogut, B., 2003. Does good science lead to valuable knowledge? Biotechnology firms and the evolutionary logic of citation patterns. Management Science 49 (4), 366–382.
Henderson, R., Jaffe, A.B., Trajtenberg, M., 1998. Universities as a source of commercial technology: a detailed analysis of university patenting, 1965–1988. Review of Economics and Statistics 80 (1), 119–127.
Hicks, D., 1995. Published papers, tacit competencies and corporate management of the Public/Private character of knowledge. Industrial and Corporate Change 4 (2), 401–424.
Hicks, D., Breitzman, T., Olivastro, D., Hamilton, K., 2001. The changing composition of innovative activity in the US—a portrait based on patent analysis. Research Policy 30 (4), 681–703.
Jensen, R., Thursby, M., 2001. Proofs and prototypes for sale: the licensing of university inventions. American Economic Review 91 (1), 240–259.
Kretschmer, H., 1994. Coauthorship networks of invisible colleges and institutionalized communities. Scientometrics 30 (1), 363–369.
Leydesdorff, L., 2004. The university-industry knowledge relationship: analyzing patents and the science base of technologies. Journal of the American Society for Information Science and Technology 55 (11), 991–1001.
Lissoni, F., Llerena, P., McKelvey, M., Sanditov, B., 2008. Academic patenting in Europe: new evidence from the KEINS database. Research Evaluation 17, 87–102.
Lissoni, F., Montobbio, F., 2008. Inventorship and authorship in Patent-Publication pairs: an enquiry into the economics of scientific credit. Cespri Working Papers, 224.
Mansfield, E., 1995. Academic research underlying industrial innovations: sources, characteristics, and financing. The Review of Economics and Statistics 77 (1), 55–65.
McMillan, G.S., Narin, F., Deeds, D.L., 2000. An analysis of the critical role of public science in innovation: the case of biotechnology. Research Policy 29 (1), 1–8.
Melin, G., Persson, O., 1996. Studying research collaboration using co-authorships. Scientometrics 36 (3), 363–377.
Meyer, M., 2000. Does science push technology? Patents citing scientific literature. Research Policy 29 (3), 409–434.
Meyer, M., 2006a. Are patenting scientists the better scholars?: An exploratory comparison of inventor-authors with their non-inventing peers in nano-science and technology. Research Policy 35 (10), 1646–1662.
Meyer, M., 2006b. Knowledge integrators or weak links? An exploratory comparison of patenting researchers with their non-inventing peers in nano-science and technology. Scientometrics 68 (3), 545–560.
Michel, J., Bettels, B., 2001. Patent citation analysis. A closer look at the basic input data from patent search reports. Scientometrics 51 (1), 185–201.
Mowery, D.C., Nelson, R.R., Sampat, B.N., Ziedonis, A.A., 2001. The growth of patenting and licensing by U.S. universities: an assessment of the effects of the Bayh-Dole act of 1980. Research Policy 30 (1), 99–119.
Murray, F., 2002. Innovation as co-evolution of scientific and technological networks: exploring tissue engineering. Research Policy 31 (8–9), 1389–1403.
Narin, F., Hamilton, K.S., Olivastro, D., 1997. The increasing linkage between U.S. technology and public science. Research Policy 26 (3), 317–330.
Narin, F., Noma, E., 1985. Is technology becoming science? Scientometrics 7 (3), 369–381.
Newman, M., 2004. Who is the best connected scientist? A study of scientific coauthorship networks. Lecture Notes in Physics, 337–370.
Newman, M.E., 2000. Models of the small world. Journal of Statistical Physics 101 (34), 819–841.
Newman, M.E.J., 2001. The structure of scientific collaboration networks. Proceedings of the National Academy of Sciences of the United States of America 98 (2), 404–409.
OECD, 2005. A Framework for Biotechnology Statistics. Organisation for Economic Co-operation and Development, Paris.
Owen-Smith, J., Powell, W.W., 2001. To patent or not: faculty decisions and institutional success at technology transfer. The Journal of Technology Transfer 26 (1–2), 99–114.
Palomeras, N., 2005. Markets for inventors: examining mobility patterns of engineers in the semiconductor industry. DTEW Research Report 0513, K.U. Leuven.
Persson, O., Beckmann, M., 1995. Locating the network of interacting authors in scientific specialties. Scientometrics 33 (3), 351–366.

Sampat, B.N., Mowery, D.C., Ziedonis, A.A., 2003. Changes in university patent quality after the Bayh-Dole act: a re-examination. International Journal of Industrial Organization 21 (9), 1371–1390.

Schmoch, U., 1993. Tracing the knowledge transfer from science to technology as reflected in patent indicators. Scientometrics 26 (1), 193–211.

Singh, J., 2005. Collaborative networks as determinants of knowledge diffusion patterns. Management Science 51 (5), 756–770.

Stern, S., 2004. Do scientists pay to be scientists? Management Science 50 (6), 835–853.

Tijssen, R.J.W., 2001. Global and domestic utilization of industrial relevant science: patent citation analysis of science-technology interactions and knowledge flows. Research Policy 30 (1), 35–54.

Tijssen, R.J.W., 2004. Is the commercialisation of scientific research affecting the production of public knowledge? Global trends in the output of corporate research articles. Research Policy 33 (5), 709–733.

Van Looy, B., Callaert, J., Debackere, K., 2006. Publication and patent behavior of academic researchers: conflicting, reinforcing or merely co-existing? Research Policy 35 (4), 596–608.

Van Looy, B., 2009. The role of entrepreneurial universities within innovation systems: an overwiev and assessment. Review of Business and Economics 54 (1), 62–81.

Wagner, C.S., Leydesdorff, L., 2005. Network structure, self-organization, and the growth of international collaboration in science. Research Policy 34 (10), 1608–1618.

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

Zucker, L.G., Darby, M.R., 2001. Capturing technological opportunity via Japan's star scientists: evidence from Japanese firms' biotech patents and products. The Journal of Technology Transfer 26 (1–2), 37–58.

Zucker, L.G., Darby, M.R., Armstrong, J., 1998. Geographically localized knowledge: spillovers or markets? Economic Inquiry 36 (1), 65–86.

Zucker, L.G., Darby, M.R., Armstrong, J.S., Armstrong, J.S., 2002. Commercializing knowledge: university science, knowledge capture, and firm performance in biotechnology. Management Science 48 (1), 138–153.