Research Policy 42 (2013) 1378–1388

Contents lists available at SciVerse ScienceDirect

Research Policy

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

Bridges or isolates? Investigating the social networks of academic inventors

Enrico Forti[a,b,*], Chiara Franzoni[c,1], Maurizio Sobrero[b,2]

[a] University College London, Department of Management Science & Innovation, Gower Street, WC1E 6BT London, United Kingdom
[b] University of Bologna, Department of Management, Via Capo di Lucca 34, 40126 Bologna, Italy
[c] Politecnico di Milano, DIG – Dipartimento di Ingegneria Gestionale, Piazza Leonardo da Vinci 32, 20133 Milano, Italy

ARTICLE INFO

Article history:
Received 27 April 2012
Received in revised form 29 March 2013
Accepted 3 May 2013
Available online 14 June 2013

Keywords:
Academic patenting
Social networks
Knowledge generation
Brokerage
Cohesive networks
Technology transfer

ABSTRACT

We analyze the acquaintances of a sample of academic inventors and their paired controls to investigate the contribution of social networks to the generation of inventive ideas in academe. Prior to patenting, inventors work in networks of similar dimension and structure as those of their colleagues who do not invent. The ego-networks of the inventors are however more cohesive (denser), a circumstance that is often seen as associated to the exchange of more fine-grained information and to a greater climate of trust which facilitates long-term relationships and learning. Over time, both inventors and non-inventors extend their networks and become more central. In general, we found no evidence that after patenting inventors isolate or close their networks.

1. Introduction

In recent years academic patenting has become the subject of extensive academic investigation (Baldini et al., 2006, 2007; Stephan et al., 2007; Azoulay et al., 2007, 2009; Calderini et al., 2007, 2009; Fabrizio and Di Minin, 2008; Breschi et al., 2008; Crespi et al., 2011). In this paper we contribute to this debate by investigating the social network dimension behind academic inventorship (Allen, 1977; Etzkowitz, 1983; Balconi et al., 2004; Murray, 2004; Lissoni, 2010). The hypotheses that underlie this paper are grounded on the theories of socially constructed knowledge and on the power of weak ties (Granovetter, 1983; Lee, 2009). Networks channel the knowledge and information that each scientist receives and recombines into their research. The accomplishments of a scientist are therefore affected by the power of their network to convey rich information. Networks that are larger in size, keeping all other things constant, convey more ideas to exploit, more complementary knowledge to make research successful and a larger group of supporters of one's own ideas (Sobrero, 2000; Lissoni, 2010). Networks with denser or sparser nodes convey information of different quality and individuals may be more or less capable to benefit from this information, depending on the position they occupy within the network (Phelps et al., 2012).

The knowledge network of academic scientists can be investigated by means of co-authorship in articles. The large incidence of homonyms in publication databases, however, creates data reliability issues, since the number of records in need of name-matching scales up at the power law for social network analysis purposes. To the best of the authors' knowledge our work is unique in its kind because it takes advantage of recent developments in name disambiguation techniques to ensure reliability of the publication data. Thanks to cooperation with Elsevier–Scopus, we could perform a network analysis based on 9997 authors of 283,280 scientific articles with reasonable certainty of very limited homonyms bias.

The analysis offered herein investigates the potential impact of patented inventions on the network's structure and the ego's position within the network by comparing pre-event measures across inventors and controls. This comparison allows us to speculate on the characteristics of networks that are associated to the inventive activity. The latter part of the analysis is also relevant to uncovering potentially changing patterns of collaboration behaviour in the aftermath of academic patenting. In principle, closer proximity to the exploitation realm may alter the role of academic inventors within their scientific community, making them more secluded and distant from their non-patenting peers (Toole and Czarnitzki, 2010). For example, they may become more prone to relying on closer and more independent relational sets, ultimately

* Corresponding author at: Department of Management Science & Innovation, University College London, 2 Malet Place, London WC1E 7JG, United Kingdom. Tel.: +44 020 7388 3666.
E-mail addresses: e.forti@ucl.ac.uk, enrico.forti@unibo.it (E. Forti), chiara.franzoni@polimi.it (C. Franzoni), maurizio.sobrero@unibo.it (M. Sobrero).
[1] Tel.: +39 02 2399 4823; fax: +39 02 2399 2730.
[2] Tel.: +39 051 2098076; fax: +39 051 2098074.

0048-7333/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.respol.2013.05.003

diminishing the overall social returns of their scientific discovery. Conversely, the inventive activity may stimulate new interests and/or expand and diversify their networks of co-authors in directions that comprise a more variegated setting.

The analysis is useful also to help disentangle team size effects from network effects. Prior works have taken team size (given by the average number of co-authors) into account, and found that inventors generally work in larger groups than non-inventors (Czarnitzki et al., 2009). Here we distinguish between individuals who repeatedly work with a large team and individuals who cooperate with many diverse co-authors in different studies.

The paper is organized as follows. In Section 2, we develop the hypotheses that will drive the empirical investigation. In Section 3, we describe the research design, the dataset, the matching procedure used to create the paired samples and the measures of social networks used in the analysis. Section 4 presents the empirical evidence and discusses the results. We conclude by highlighting the contributions of our paper and some open questions for future research in Section 5.

2. Academic patenting and social network effects

2.1. Explaining the research accomplishments of academic inventors

We have learned from recent works on academic patenting that inventors represent a small share of the population of academics. Even in the subfields in which patenting is relatively common, like biotech and chemistry academic inventors never seem to exceed 10–15% of the scholars (Agrawal and Henderson, 2002; Breschi et al., 2008). Several works have also consistently shown that the most productive and accomplished individuals in science are overrepresented within the sample of academic inventors (Fabrizio and Di Minin, 2008; Stephan et al., 2007). Furthermore, when data is analyzed on a longitudinal timespan, patents seem to be preceded by a burst of publications (Azoulay et al., 2007; Calderini et al., 2007) and tend to boost productivity in the years immediately after patenting (Azoulay et al., 2009; Breschi et al., 2008; Calderini et al., 2009).

The fact that a few productive authors in science are disproportionally responsible for a large share of the publications has been well documented since the early '60s (de Solla Price, 1963; Allison and Stewart, 1974). Still, it is at first counterintuitive that an even smaller proportion of scholars seems to be capable of simultaneously producing advances in the scientific understanding of principles, phenomena and new technologies suitable for industrial application.

This circumstance has raised a question about what capabilities form the basis of academic patenting and if there are common drivers that explain the success of a scientist in the academic and industrial worlds. Although in the traditional view of science as a speculative activity, scientific inquiry and practical application are seen as antonyms, at a closer look several considerations suggest that this vision is oversimplified and obscures the true nature of research. Scholars have suggested multiple potential explanations for the positive correlation between publications and patents.

First, there are areas of investigation (the so-called "Pasteur's Quadrant") in which fundamental understanding and practical applications can be pursued at the same time and other areas of investigation in which this is not the case (Stokes, 1997). In the first case, the pursuit of scientific and technological goals can be combined, and the two activities can generate positive feedback for one another. This happened for instance in the early years of biotechnology, when many eminent scientists became famous for their technological advances while maintaining a leading position in science (see, for instance, Zucker et al., 1998; Davies, 2001; Feldman et al., 2005). A first possible explanation for the correlation between scientific and technical achievements is that we are observing areas in which the trade-off is less severe.

Second, success in research often requires the solution of technical problems that constrain scientific investigation. Scholars who study the creativity of scientists maintain that the rate-limiting factor for progress in science is not the pace at which new ideas come to researchers but the pace at which those ideas can be transformed into feasible operations on the bench (Holmes, 2004). Since a large proportion of the inventions that academic scientists produce relate to improved research technologies, the event of producing a patent precedes success in research (Franzoni, 2009).

Third, successful scientists are often described as individuals who are entrepreneurial by nature (Allen, 1977; Etzkowitz, 1983). Success in science requires extensive organizational skills as well as the capacity to raise funds to support a line of research. This is especially true in recent years, as proven by the steadily increasing sizes of research teams (Adams et al., 2005; Wuchty et al., 2007) and the enlarged budgets that need to equip fully functional research labs (Stephan, 2012). A successful scientist needs to be skilled at envisioning funding opportunities, establishing collaborations, brokering research scope and uncovering market needs. These abilities are also likely to underlie success in developing technologies for the market (Murray, 2004; Baldini et al., 2007; Franzoni and Lissoni, 2009).

2.2. The effect of social networks on inventive activities

In this paper we investigate the social network of inventors – prior and after patenting – in search of explanations for why scientific and market achievements are correlated. This explanation is grounded on the theory of socially constructed knowledge and hypothesizes that a larger and richer social network is the basis for superior performance by scientists in both research and inventive accomplishments.

Extensive studies on social network theory have emphasized the relevance of the social network dimension in the creation and diffusion of knowledge (Coleman, 1988; Freeman, 1991; Ahuja, 2000). The importance of relational capital depends on the circumstance that knowledge is only partially codifiable and remains largely tacit and bound to individuals (Nelson and Winter, 1982). This highlights the importance of face-to-face (or somehow socially-channelled) collaborations to enable the circulation and exchange of novel ideas in research. The characteristics and structure of the network of collaborations in which a person works and the position of a specific node within the network should therefore concur to explain the extent to which a single node would be productive of new ideas, such as those leading to innovation (Sobrero, 2000; Nerkar and Paruchuri, 2005).

In this section, we build on the contributions of the literature to formulate hypotheses regarding the correlation between a number of characteristics of one person's network and her propensity to produce inventions. We focus on three elements: (i) network dimension; (ii) network position; and (iii) ego-network structure building hypotheses on how these features of relational capital may be associated to a greater propensity to become inventors.

Network dimension (*size*). Based on the theory of knowledge recombination, each individual at a given moment owns a certain endowment of knowledge accumulated during prior experience. When individuals interact with other individuals (for example they co-author a work), they exchange and recombine their respective knowledge sets, producing new combinations. Knowledge recombination may not always be easy or successful but when a successful recombination occurs, this generates a novel idea, solution or insight (Hargadon and Sutton, 1997; Fleming, 2001; Murray and O'Mahony, 2007). Collaboration enables a faster pace of progress because it is not constrained by the speed of individual learning. Furthermore, the recombination of a network's knowledge generates more chances to produce creative results than the recombination of a single individual's knowledge, because the contribution of each person is different. Scholars who work on creativity in science have proposed a theory of chance and creativity, called Chance Combination Theory, which helps us to explicate this issue (Simonton, 2004). According to this theory, creativity results from the ability of the scientist to associate and combine pieces of knowledge and information in ways that are both original (never tried before) and useful. The probability of a successful combination is tiny and cannot be foreseen in advance, but it increases with the number of times that novel combinations are tried. If we consider people to be repositories of idiosyncratic and tacit knowledge, and co-authorship as the link through which knowledge travels from one individual to another, the number of unique combinations that each individual can make increases with the pool of knowledge that she can access and hence with the size of her network (Perry-Smith and Shalley, 2003). Several studies have confirmed this idea, for example, by finding positive correlations between the size of a research team and various indicators of the quantity and quality of its publications (De Beaver and Rosen, 1979; Kretschmer, 2004; Defazio et al., 2009).

A second argument suggesting a positive correlation between the extent to which a scientist relies on a broad relational capital and her productivity relates to the internal approach to the world of science, which was extensively explored by Robert K. Merton and colleagues (see for instance Merton, 1957; Hagstrom, 1965). Scientific theories and findings, especially new and disruptive ones, spread out and become affirmed when they are known, discussed and agreed upon by the scholarly community. Therefore, numerous and frequent relationships with a large community of colleagues favour the acceptance of a scholar's work (Allen, 1977), while isolation reduces the probability of success.

Both of these arguments – that relational capital supports the generation of new ideas, as well as their diffusion and acceptance – lead us to expect a positive correlation between the size of the social network to which the scientist has access and success in inventive activities. Individuals who are working in larger networks should therefore be more likely to generate new ideas, simply because larger networks have greater capacity to convey more numerous information, all other things being equal. Our Hypothesis 1 will therefore be as follows:

H1 *(:).* Inventors should be found in larger proportions among scientists with larger networks.

Network position (*centrality and brokerage*). The nodes that compose a network can be characterized in terms of their position. One prominent feature of a node position relates to how essential it is for connecting the other nodes of a network. This property is captured by network centrality. Nodes connecting otherwise disconnected individuals are more critical to convey information exchange. Scholars of social network theory have maintained that nodes in more central positions have not only more direct ties (hence more occasion to get first-hand information), but also enjoy controlling privileges over valuable information exchanged by their acquaintances (Nerkar and Paruchuri, 2005). Bridges are crucial to enable knowledge transmission throughout the entire network and consequently enjoy a higher status by virtue of their position (Langlois, 1977; Burt, 1992). In social network analysis, individuals who are central in their network, not only provide means to connect individuals, but also make the connection that they bring less redundant with respect to all other connections. As a consequence, they should exhibit a comparative advantage over less central individuals. In this paper, consistently with prior studies (Reagans and McEvily, 2003; Nerkar and Paruchuri, 2005; Lee, 2009) we are interested in investigating whether the brokering position of a scholar associates with inventive behaviour.

In general, a brokerage position is assumed to enhance the capability to generate fresh knowledge (Audia and Goncalo, 2007; McFadyen and Cannella, 2004). The relationship may not necessarily be linear, as diseconomies may emerge when networks grow very large (McFadyen and Cannella, 2004). However, for the relatively limited size that characterizes scientific co-authorship networks, we can formulate the following:

H2 *(:).* Inventors should be found in larger proportions among scientists that lay in central and brokering positions.

Ego-network structure (*density and constraint*). The predictions of the literature concerning the link between knowledge generation and dimension and network position are quite straightforward. In contrast, social network scholars have for long been divided concerning whether ego-networks in which individuals interact more intensely and repeatedly are more or less useful for knowledge generation than ego-networks with less redundant links and structural holes (Phelps et al., 2012). More cohesive networks, where individuals are more interdependent and share the same acquaintances, increase trust, permit to build a reputation and enable a more fine-grained diffusion of knowledge (Reagans and McEvily, 2003). However, when individuals continue to interact with the same group of people, they are prone to receiving the same information several times (Burt, 1992; Lee, 2009). Intuitively, redundant information is less relevant than fresh information and exhausts creative power. Conversely, weak ties and sparser connections might bring novel, non-redundant information (Granovetter, 1983; Nahapiet and Ghoshal, 1998).

Some scholars proposed a contingency view of ego-networks, suggesting that different types are more useful for certain tasks than for others. Because high cohesion supports productivity and trust – and is therefore important to exchange detailed or highly sensitive knowledge – dense networks function better in sustaining learning processes (Morgan and Sørensen, 1999). Cohesion may also promote an easier spread and adoption of new ideas (Fleming et al., 2007). Conversely, weak ties bring in more diversity and tend to function better in sustaining creative processes (Uzzi et al., 2007), because individuals with very diverse endowments have more chances of producing novel, non-duplicative combinations (Perry-Smith and Shalley, 2003; Uzzi and Spiro, 2005).

Some scholars have suggested that the degree to which the external environment is stable serves as a moderating factor. For example, Rowley et al. (2000) study alliance networks in the steel and semiconductor industries and find that high density is more important in relatively stable competitive markets, while weak ties would be more advantageous in uncertain environments. Given that the predictions of the literature are inconclusive, we formulate two mutually exclusive hypotheses.

H3a *(:).* Denser and less constrained networks promote the exchange of finer-grained information and more effective learning and would be associated to greater propensity to invent.

H3b *(:).* Sparser (less dense) and more constrained networks foster creative processes and would be associated to greater propensity to invent.

2.3. The impact of inventive activities on scientists' networks

A final issue we want to explore in our analysis relates to the participation of academic inventors in networks and communities in the post-invention period. A very rich debate has flourished in the last decade about the positive and negative implications of favouring the patenting activities of academic scholars. A summary of this extensive debate would exceed the scope of this paper.[3] Here, we limit our discussion to recalling that isolation in secluded scholarly communities has been reported among the potential risks associated with excess attention for research commercialization of academic scientists (Toole and Czarnitzki, 2010).

In this paper, we contribute to this debate by investigating the structure of the social network and the respective position of academic inventors five years after patenting. To highlight variations, we compare network indicators before and after the invention. In order to account for potentially confounding factors posed by life cycle effects (e.g., networks tend to grow as the scientist ages, then level-off towards the end of the career), we further look at comparable indicators computed for a matched colleague, paired as a control.

If academic inventors become more prone to isolate themselves from the academic community in order to pursue commercial research, then we should observe deterioration in their networks along the different dimensions highlighted above.

H4a. Smaller networks or less central or brokering positions after patenting may indicate that academic inventors become more secluded from the scientific community after patenting.

H4b. Larger networks and more central or brokering positions after patenting may support the view that inventions do not undermine the capacity/willingness of the scientist to interact with the academic world.

In addition, or in combination to the previous, we would be interested to observe whether the ego-network structure (density and constraint) of an academic inventor changes after patenting. As mentioned earlier we do not have a unique prediction as of which structure would be more desirable. Consequently, it would be difficult to tell ex-ante which specific change should be regarded as an improvement or rather a deterioration. Accordingly, we propose two alternative hypotheses, each subordinated to whether in the pre-invention period we had found confirmation to H3a or H3b. If H3a is confirmed, then we will investigate the following hypothesis:

H5a. Less dense and more constrained networks after patenting may indicate that academic inventors lose part of the knowledge generation capacity of their network in the post-invention phase.

Conversely, if 3b is confirmed, then we will investigate the following hypothesis:

H5b *(:)*. More dense and less constrained networks after patenting may indicate that academic inventors lose part of the knowledge generation capacity of their network in the post-invention phase.

3. Research design

We investigate whether differences in the network structure and position of scientists are individual-specific and fixed, or if they change over time. By including the time dimension, we also ask whether the circumstance of being inventive contributes to alterations in the network and the inventor's position within it. Empirically, answering this question requires addressing the issue of time-varying confounding factors. Typically, as for all scientific indicators, the most relevant of these confounding factors is the life-cycle (or career age) effect.

Productivity is in fact known to grow (more than linearly) throughout most of a scientist's career but generally declines in the last years (Levin and Stephan, 1991). Therefore, when comparing scientific achievements before and after a specific year (e.g., the year of the invention), there is a general tendency to register growth. Problems arise in determining to what extent this tendency is to be attributed to merely the effects of time rather than to the event of interest. Difference-in-differences methods have been shown as an effective strategy to address issues of spurious correlation (Bertrand et al., 2004). Accordingly, we control for the life-cycle dependence and other confounding factors by assessing differences in network characteristics and positions across matched pairs of scientists with similar confounding factors.

3.1. Data

Studies on the communities of scientists often highlight profound international differences in the scientific labour markets as well as differences among academic institutions operating in the same country (Bonaccorsi and Dario, 2007; Czarnitzki et al., 2009; Lee, 2009). To avoid potential interference of institutional factors, we focus on one single country, Italy, with a homogeneous university system and homogeneous training, hiring and promotion practices during the time of observation. We begin by building a sample of academic inventors and sample of academic scholars who did not invent.

Academic inventors were identified using the Patiris database (Forti and Sobrero, 2013), which includes all patent applications by Italian universities, filed nationally and abroad, both directly or as extensions of patents filed elsewhere. This database includes all of the inventors appearing in patents filed by a university or national research agency based in Italy. However, it does not include all Italian academic inventors, as many academic inventors have filed patents with companies or under their own name, given the academic privilege that was in force during most of the timespan considered (Balconi et al., 2004). We will keep this in mind when interpreting the results.

To avoid disciplinary differences, we focused exclusively on the field of Chemistry.[4] This restricts the sample to 59 inventors between 1982 and 2006. Patent incidence is consistent to those found in other independent studies (Breschi et al., 2008).

Controls were taken from the population of Italian chemists who did not appear among the inventors of academic patents. To ensure that these scholars had never been inventors of patents assigned to other institutions or individuals, we further restricted the sample to those who explicitly denied patenting experience in a questionnaire distributed to Italian academic scholars (Baldini, 2004). This gave us a sample of 85 non-inventors in the field of Chemistry for drawing matched pairs.[5]

[3] See for example Siegel and Wright (2007) for a review.

[4] All Italian professors are affiliated with a unique subfield within the classification of the Italian Ministry of University and Research. The list of classes (http://www.miur.it/UserFiles/116.htm) and the professors affiliated with them (http://cercauniversita.cineca.it/php5/docenti/cerca.php) is publicly available. This data are updated on a yearly basis and our study is based on professors listed at the time of data collection.

[5] We have no reason to believe that our controls could have had any motivation to avoid reporting patenting activity. Therefore, we are confident that they didn't become inventors within or before the observation window. It is important however to note that the control and the treated may have partially overlapping networks, thus potentially causing indirect treatment on the controls via their network. In terms of magnitude the incidence of network overlap at the second level of acquaintances is however very limited. Considering the whole set of papers co-authored by inventors and non-inventors and their co-authors, we found that 0.02% of all articles published during their entire career overlapped to some extent. We appreciate the comments of an anonymous reviewer in suggesting that we clarify this point in explaining our use of the data.

3.2. Matching procedure

We run a pair matching procedure starting with 144 scientists: 59 academic inventors and 85 academic scholars who did not hold inventions. We regarded the event of becoming an inventor as a "treatment" and held the non-inventors as "untreated" individuals to serve as controls. The observable pre-treatment differences were accounted for by looking at a set of exogenous demographics (age, gender, university location) and other exogenous personal variables (has a PhD, obtained PhD in Italy vs. abroad, which subfield of Chemistry does the person belong). We calculated an individual propensity score of patenting as the predicted coefficient of a probit estimate in which the probability of experiencing the event depended only on the two sets of covariates.[6]

We then paired each inventor to the scholar whose propensity score was more proximal by means of a one-to-one nearest-neighbour matching procedure without replacement. Under a reasonably strict Calliper (Cochran and Rubin, 1973) of 0.70, we dropped four academic inventors because there was no good match available and another four were dropped for name disambiguation concerns. The final matched pairs had a mean difference probability of 19.52% (standard deviation 0.24), indicating a satisfactory matching Calliper of 0.67. The final group of scholars for which we perform the network analysis is composed of 106 scientists, evenly split between inventors and controls.

It is important to clarify that we intentionally avoided using variables based on publication when predicting propensity scores. Since our links are based on co-authorship in articles, forcing matching on article-related measures would cause selection on those variables that we wish to observe later in the network analysis. In other words, it would cause endogeneity of our observables due to selection. Rather, we chose to include predictors that are exogenous to our network measures, but are good predictors of productivity of publications. As we will see in the next sub-sections, this strategy works fairly well. In fact the paired individuals result to have very similar productivity and co-authorship records.

3.3. Construction of the social networks

We begin with two sets of egos, each composed of 53 scientists. Ego-network data require three elements to be determined: the egos, in our case the 106 academic inventors and controls; the alters, in our case all the individuals with which one of the egos (either an inventor or a control) co-authored their papers within a certain time frame; and finally, a measure of ego-alter relationships, in our case the co-publication frequency – i.e., the number of articles co-authored by egos and alters (Wasserman and Faust, 1994, p. 42).

Alters were identified by retrieving the scientific publications of all egos listed in the Scopus database within a moving window of 11 years. This time period varies for each individual and is centred on the year of patent priority for each inventor and the same year for the inventor's matched control. Inventors start to be observed five years before the invention[7] and are observed for five years after the invention, which gives an 11-year window of observation. Controls are observed over the same time window of their paired inventor.

Name matching is a challenging endeavour in most publication-based research and presents peculiar issues for collecting network data. To make sure that listed authors were consistently the same person, we controlled for homonyms and common names. Additionally, for each individual in our sample we checked affiliation data, publication subject and consistency between the time of publication and the age of the scientist. However, while these checks could be performed on the focal sample of inventors and non-inventors (egos), we could not run the same checks on all of their co-authors (alters) due to a lack of knowledge of the attributes of those individuals. To avoid generating unreliable network data for alters, we base our data collection on the latest release of Elsevier Scopus, which tracks individuals by their author identifiers, rather than by first and last name (or last name and initial).[8] These are unique numerical author-IDs assigned by Scopus after a process of name disambiguation. Ultimately, we use this information to identify all authors in our sample and match them to publications' IDs.

We began our data collection with manual searches using the names and surnames of our 106 focal authors. A database was compiled to match authors' names with the corresponding Scopus author-ID. Using the author-IDs of co-authors we then retrieved all the publications of the alters and the author-IDs of the co-authors of the alters. This methodological choice raises the caveat that our analysis will generate unbiased results only to the extent that biases and mistakes in the assignation of author-IDs by Elsevier Scopus should apply randomly to both the alters of the inventors and to those of their controls. Fortunately, this circumstance is likely to be satisfied.

The publications of all egos in the time windows of observation resulted in a total of 4971 articles written by the egos along with 6871 co-authors.[9] Building on this first set of authors, we retrieve via author-IDs the publications of the 6871 alters. This step is required to take into account possible direct connections between the alters that do not involve the egos. Failing to do so would produce seriously biased network indicators, based only on a small fraction of the underlying data.

At this point, our dataset contains 283,280 scientific articles written by a total of 9997 authors. Since we limit our study to a moving window of 11 years (centred on the year of patent priority for an inventor and the same year for its matched control), after slicing and dicing data outside of the relevant windows, our final dataset contains a total of 59,457 articles and 6157 authors. This breaks down into 1927 publications authored by the 53 inventors along with their 3057 co-authors, 1824 by the 53 controls along with 3303 co-authors and 55,659 articles published by 4486 co-authors of either inventors and controls that do not include the 53 inventors nor the 53 controls. Inventors and controls shared 297 co-authors and were co-authors in 54 papers (co-authored with 106 common co-authors).

We divided this matrix into 4 sub-matrixes mapping the connections of inventors and controls, before and after their first patent application. These matrixes include direct and indirect co-authorship ties for all the egos and all direct ties of the alters. We used Pajek to convert each 2-mode matrix (author-by-publication) into a 1-mode one (author-by-author) where each intersection contains the number of publications in common between any dyad of

[6] The probit estimate was omitted for brevity, but is available upon request.

[7] This is the patent priority year. Please note that in Italy, as well as within the European Patent Office countries, the inventors are required to file a patent application before disclosing the invention in any form. This results in quick patent filing. Patent filing and priority hence give a close and fairly good approximation of the timing of an invention.

[8] We are grateful to the Elsevier–Scopus team for custom data-retrieval and assistance with the use of author's identifiers.

[9] Between 1982 and 2006, 54 publications were jointly co-authored by inventors and controls, 2581 articles were written by the 53 inventors along with 2947 co-authors and 2444 were published by the controls along with 4333 co-authors. Inventors and controls shared 410 co-authors.

Table 1
Descriptive statistics of the final sample (106 Obs.).

Variable	Mean: inventors	Mean: non-inventors
Year of birth	1952	1950
	(9.48)	(9.26)
Dummy age: 35–45	0.24	0.2
	(0.43)	(0.4)
Dummy age: 46–55	0.29	0.29
	(0.46)	(0.46)
Dummy age: 56–65	0.33	0.36
	(0.47)	(0.49)
Dummy age: over 65	0.15	0.15
	(0.36)	(0.36)
Gender	0.78	0.89
	(0.42)	(0.32)
Geographic location: north	0.51	0.58
	(0.51)	(0.5)
Geographic location: centre	0.33	0.2
	(0.47)	(0.4)
Geographic location: south	0.16	0.22
	(0.37)	(0.42)
Phd	0.35	0.24
	(0.48)	(0.43)
Phd in Italy	0.33	0.22
	(0.47)	(0.42)
Phd Abroad	0.02	0.02
	(0.14)	(0.14)

Standard error in parentheses.

Table 2
Pre- and post-patent between- and within-group differences.

	Time	Obs.	Mean	Standard deviation	Min	Max
Inventors						
Publications	Before	53	15.22	14.25	0	71
Publications	After	53	30.19	33.23	0	161
Authors per paper	Before	616	5.43	2.49	1	20
Authors per paper	After	1421	6.01	4.50	1	71
Size	Before	53	22.49	23.13	0	126
Size	After	53	53.77	51.37	6.00	307
Norm. betweenness	Before	53	33.99	25.38	0	84.21
Norm. betweenness	After	53	47.92	23.73	0	88.69
Norm. brokerage	Before	53	0.52	0.33	0	0.92
Norm. brokerage	After	53	0.73	0.23	0	0.96
Density	Before	53	32.98	28.25	0	100
Density	After	53	27.25	22.52	3.65	100
Constraint	Before	53	0.27	0.22	0	0.97
Constraint	After	53	0.23	0.15	0.04	0.64
Controls						
Publications	Before	53	15.48	11.01	0	51
Publications	After	53	24.35	18.13	0	81
Authors per paper	Before	650	5.43	2.36	1	23
Authors per paper	After	1264	6.00	2.65	1	17
Size	Before	53	24.58	24.12	0	114
Size	After	53	50.66	40.40	0	213
Norm. betweenness	Before	53	39.91	27.81	0	88.29
Norm. betweenness	After	53	51.50	20.97	0	85.28
Norm. brokerage	Before	53	0.56	0.33	0	0.94
Norm. brokerage	After	53	0.76	0.19	0	0.95
Density	Before	53	23.24	20.06	0	100
Density	After	53	21.92	16.35	0	100
Constraint	Before	53	0.22	0.20	0	1.05
Constraint	After	53	0.20	0.15	0.04	0.83

authors. Network indicators for inventors and controls were finally computed using UCINET 6 (Borgatti et al., 2002).

3.4. Descriptive statistics

Table 1 gives an account of the characteristics of the egos split between the two samples of inventors and non-inventors. Around 22% of the scientists sampled are between 35 and 45 years old. Consistently with studies on age and productivity, this age class exhibits the peak of productivity. In agreement with the country-level age distribution of university professors, 34% are rather senior (56 years old or older). The majority of the scientists sampled work in the north of the country, less than 20% work in the south, and the remaining are based in central Italy. This distribution is again consistent with the geographical distribution of universities in the country. About one-third of the scientists have a PhD, while the majority have a second-level university degree. Again, this is consistent with the age profile when considering that PhD programmes were first introduced in Italy in 1984 and full training abroad was quite unusual in the seventies.

3.5. Network indicators

General statistics of the indicators computed for the two samples – inventors and controls – separately before and after the year of their first patent application are reported in Table 2. Before the focal year, inventors and controls have approximately the same productivity (15.2 vs. 15.5 publications). The number of co-authors per paper is also very much the same (5.43). This additionally reassures us on the effectiveness of our matching procedure.[10]

We characterize individual networks by five properties: *size*, *density*, *betweenness centrality*, *constraint* and *brokerage*. First, a network's size grows with the total number of individuals directly related to each other even without the mediating role of the relevant ego. Accordingly, we measure *size* to evaluate the dimension of the community to which a scientist belongs. The size of a network is computed by counting the number of nodes (Wasserman and Faust, 1994). Since each scientist can deploy limited resources and capabilities for building and maintaining ties, network size becomes a useful indicator to assess the structure of scientists' relations. Second, networks where all logically possible ties are actually present are empirically rare. Therefore, it is useful to look at how close a network is to realizing this potential – that is, to examine its density (Wasserman and Faust, 1994). A network's *density* relates to the intensity of the relationship linking the authors based on the frequency of their interactions. Since overall density is affected by network size, we compute a standardized measure of density – the relative network density – instead of an absolute density (Friedkin, 1981).

Third, investigating the division of scientists into groups and sub-structures as well as redundancies in scientists' connections may have important implications for the social dynamics of scientific communities. For instance, when colleagues connected to a focal scientist are not themselves connected to one another, a brokerage opportunity exists for the focal scientist. Burt defines "structural holes" as non-redundant relationships among actors and describes the use of a measure of *constraint* to formally evaluate whether an actor's network is rich or poor in structural holes by capturing the extent to which an actor's "time and energy are concentrated in a single group of interconnected colleagues – which means no access to structural holes" Burt (2010, p. 294).

The constraint indicator has been validated by Burt (e.g., 1992, 2004) and other scholars in a variety of empirical contexts (e.g., Lee, 2009). Constraint captures the degree to which an ego enhances the cohesion among sub-groups of a network and is expressed as a negative value. The greater is the opportunity for a scientist to act as a bridge among subgroups (that may themselves be more or less densely knit), the lower will be the constraint indicator. We

[10] Note that in our construction of the data the sample of inventors and controls are drawn from populations of different sizes. The fact that our chosen controls are only very productive individuals mitigates the potential problems caused by excluding the null relationships in a true global matrix (virtually composed of all Italian scientists in chemistry). We appreciate the comments of an anonymous reviewer that we provide more information about this approach.

Table 3
Summary of variables and data sources.

Variable	Description	Formula
Size	Size reflects the number of ties ego has and is generally computed by counting the number of individual nodes. In our case is computed as the number of individual co-authors x of the focal scientist i plus the scientist i itself (Wasserman and Faust, 1994)	$S_i = x_i + 1$
Density	Networks where all logically possible ties are actually present are empirically rare. It is useful to look at how close a network is to realizing this potential. That is, to examine its density. The network density of a scientist i is measured as the ratio between the lines that are actually present in scientist i's ego network (L_i) and the maximum number of lines that can be present $g(g-1)/2$ (Wasserman and Faust, 1994, p. 101)	$\Delta_i = \frac{2L_i}{g(g-1)}$
Normalized betweenness	Nodes that occur on many shortest paths between other vertices have higher betweenness than those that do not. Betweenness centrality is defined for each node k as the share of times that a node i needs a node k in order to reach a node j via the shortest path. Specifically, g_{ij} is the number of geodesic paths from i to j, and g_{ikj} is the number of these geodesics that pass through node k. We normalize the betweenness indicator dividing simple betweenness by its maximum value (Wasserman and Faust, 1994; Borgatti, 2005)	$B_k = \sum_i \sum_j \frac{g_{ikj}}{g_{ij}}, \quad i \neq j \neq k$
Constraint	The network constraint of i captures redundancies in a scientist's connections where p_{ij} is the proportion of time and energy that ego i invested in ego j (p_{iq} and p_{qj} are defined analogously). Hence, the network constraint of i is the proportion of ego i's relationship that are invested in connection with ego j. Lower values on this indicator imply that scientist i occupies a less constrained position, thereby brokering more extensively in the network (Burt, 1992; Lee, 2009)	$C_i = \sum_j \left(p_{ij} + \sum_{q \neq i \neq j} p_{iq} p_{qi} \right)^2$
Normalized brokerage	When the actors participating in a focal scientist's ego network are not connected directly to one another, the ego may be a broker if she/he is the only intermediary between a pair of actors. Since we are interested in measuring how many times pairs of actors in a scientist's network are not directly connected (i.e., the number of paths on which the focal scientist lies) we compute a scientist j's absolute brokerage following the method developed by Gould and Fernandez (1989, p. 101, Eq. (2)). We normalize this indicator by dividing the total brokerage score t_j by the number of pairs that can be present in scientist j's network (Borgatti et al., 2002). Higher values on this indicator imply that scientist j is frequently the only intermediary between two other actors	$t_j = b_{0j} + b_{i0j} + b_{0Ij} + w_{0j} + w_{Ij}$

compute *constraint* scores for each inventor and non-inventor in our sample before and after patenting to assess changes in the level of cohesion that a scientist brings in her network. The measures are computed for all nodes in the network, treating each one as ego and considering the ego network as if the rest of the network did not exist, so that ties beyond alters have no effect. Hence we only consider alter–alter ties as originally suggested by Burt (2004).

Fourth, *betweenness centrality* measures are used to characterize the position of an ego within its own network and to assess the strength with which the individual acts as an intermediary in the circulation of information (Freeman, 1991; Borgatti, 2005). Centrality increases when the individual is a necessary link connecting two pairs of nodes from a larger network (Borgatti, 2005). Technically, information centrality indexes are based on the geodesic distance between actors and measure the shortest path connecting a pair of actors (Wasserman and Faust, 1994). The more an individual lays between others on their geodesics, the greater her importance in keeping the nodes in contact. We normalize the betweenness indicator by dividing simple betweenness by its maximum value.

Finally, in line with the work of other innovation scholars (Reagans and McEvily, 2003; Nerkar and Paruchuri, 2005; Lissoni, 2010), we compute *brokerage* to further characterize a scientist's position in its knowledge network. Complementing the information offered by betweenness centrality, the value of a brokering position stems from the relative redundancy of network ties. The intuition is that scientists have different gains from connecting to colleagues who are already in connection through others as opposed to connecting with colleagues who would otherwise be disconnected (Gould and Fernandez, 1989). Scientists acting as brokers can bridge gaps in the social fabric of science by helping and/or exploiting the flow of information and opportunities (Burt, 2004). As little or no new information becomes available through redundant ties, scientists who act as the only intermediary between pairs of colleagues can benefit from a superior network position for inventive activities (Lee, 2009; Fleming et al., 2007; Stovel and Shaw, 2012). Since we are interested in measuring the number of paths on which the focal scientist lies we characterize a scientist's absolute brokerage following the method developed by Gould and Fernandez (1989).[11] Table 3 reports all these measures and how they are computed

4. Empirical analysis

Our analysis is based on paired-tests. We want to test the null hypothesis that the mean difference of the observed variable between each paired-sample equals zero. In order to do so, in this and in all subsequent comparisons, we take the following approach: for normally distributed paired differences, a simple *T*-test was performed under the null hypothesis of zero mean difference of the variable. Because the distribution of paired differences sometimes is skewed in important ways, we performed a Kolmogorov–Smirnov test, where the null hypothesis is that the sum of paired differences obeyed a normal distribution with the mean equal to the variance. This is equivalent to testing whether or not the values of the variables in the two samples have a similar distribution. Normality is evaluated with the Shapiro–Wilk test.

We look at between and within group comparisons. The between-group comparisons highlight differences between inventors and controls. Recall that paired samples are meant to remove the effects of confounding variables (age, gender, geographic location, PhD) that would otherwise concur to cause differences in the indicators of inventors and controls. In fact, as we desired, the inventors and controls do not statistically differ before the treatment in terms of number of publications and number of co-authors per paper (Table 4, left). Recall that our propensity score matching was not based on publication measures, to avoid issues of

[11] Gould and Fernandez's (1989) approach to brokerage allows to measure the absolute brokerage capacity of an actor and to further differentiate between five types of brokers (liaisons, itinerants, coordinators, gatekeepers, and representatives) corresponding to different triadic configurations of actors and intermediaries. In our paper we are simply interested in measuring the extent to which a scientist's role is that of a broker, therefore, we characterize brokerage for each focal scientist using Gould and Fernandez's absolute brokerage score.

Table 4
Differences and significance.

Variable	Between groups (Inventors vs. controls)		Within groups (After vs. before)	
	5 years before patent	5 years after patent	Inventors	Controls
Publications	−0.47 (1.50)	6.21 (3.67)*	17.26 (4.96)*	11.53 (3.00)*
Authors per paper	0.37 (0.31)	−4.69 (5.42)	19.42 (0.33)*	24.49 (5.43)*

Standard error in parentheses.
* $p<0.05$.

endogeneity. The fact that our paired individuals exhibit similar productivity is a further confirmation that our predictors worked well at identifying fairly similar controls. Suppose for example that inventors are exceptionally good scientists; the matching procedure would work to pair these individuals with another set of exceptionally good scientists, although with no patent, rather than on just an average group of scientists. In line with prior evidence (Azoulay et al., 2009; Calderini et al., 2009), our results also confirm that the inventors published more than the controls after patenting.

The within-group comparisons highlight differences between the pre and the post patent period within the subgroup of inventors and controls separately. As stated before, we should be aware that most of the indicators of scientific performance tend to grow over time, consistent with the life cycle effect (the scientists in our sample are on average 53 and therefore experience monotonic growth on average). Here we use the same time-window for inventors and controls to account for the potentially confounding effect of time and life-cycle in inflating the values of the indicators as the scientist ages. In fact, as we expect, both groups of scientists experience increase in the number of articles published over time (Table 4, right). Note also that both groups experience an increase of co-authors. This evidence can be explained in part by life-cycle effects, and in part by the general trend towards expanding research teams observed in recent years (Adams et al., 2005; Wuchty et al., 2007). Whatever the reason, we want to test our hypotheses net of this impact. Therefore we wish to explore whether inventors experience a change in the network structure and position that could not just be explained by the effect of time.

4.1. *Relational capital and ego position before patenting*

In Table 5 we report the results of the tests performed between groups, i.e., inventors versus controls, relative to the network dimension, position and ego-network structure of our scientists. The positive sign indicates that inventors exhibit a higher value of the indicator and the negative sign indicate that the controls exhibit a higher value of the indicator. In the first raw of the table we see that inventors have slightly smaller networks than controls prior to inventing, although the difference is not statistically significant. In general, contrary to the prediction of H1, inventors and non-inventors were immersed in networks of similar *size* prior than the focal year. Network dimension is therefore not likely to be associated to a greater propensity to produce inventions, at least in our sample.

Table 5
Differences and significance – between groups.

Variable	(Inventors minus controls)	
	5 years before patent	5 years after patent
Size	−2.09 (3.10)	3.11 (6.06) ψ
Betweenness (normalized)	−5.92 (4.03)* ψ	−3.58 (4.13) ψ
Brokerage (normalized)	−0.04 (0.04) ψ	−0.03 (0.03)* ψ
Density	9.74 (4.62)*	5.34 (3.81)
Constraint	0.02 (0.04) ψ	−0.11 (0.03)* ψ

Standard error in parentheses. ψ: non-normally distributed; Kolgomorov–Smirnov test.
* $p<0.05$.

In the second and third row of Table 5 we report the average difference in the indicators of network structure in terms of centrality and brokerage of inventors and controls. Here, we observe that the inventors were less central in their networks than non-inventors prior to patenting (they exhibit a lower *betweeness centrality* than controls and a comparable level of *brokerage*). We therefore find no support to our H2 of a positive association of the inventive activity and a more central or more *brokering* position towards otherwise disconnected nodes.

In the third and forth row of Table 5 we report the mean difference values for the ego-network indicators of density and constraints. Here we see that inventors' ego-networks were more cohesive than the ego-networks of non-inventors, a circumstance that is highlighted by a superior *density*, with virtually the same level of *constraint*. This evidence supports our H3a that denser networks might be associated to greater inventive capacity, possibly because frequent and redundant connections enhance the exchange of fine-grained information, and/or promote trust, cooperation and learning among the network participants.

In conclusion, when we compare inventors and controls prior to the event of patenting, we find no clear differences concerning network dimension, and network position. We do find evidence that, prior to patenting, inventors were immersed in denser, more cohesive networks than non-inventors. In other words the inventors and their co-authors were more likely to be interconnected among each other with multiple ties than the controls.

4.2. *Relational capital and ego position after patenting*

Scientific performance is strongly dependent on life and career effects (Levin and Stephan, 1991). Furthermore, a general trend of increasing cooperation in science has been documented in recent years (Adams et al., 2005; Wuchty et al., 2007). Hence, comparing measures derived from scientific outcomes (articles published, citations, co-authorship) in a certain period against similar measures derived from scientific outcomes at a later period can be misleading and result in spurious correlations. In this paper we mitigate this risk by comparing the indicators calculated in different time periods in light of the within groups differences (Table 6). Within group differences highlight the growth of the indicators over time within the group of the inventors and within the group of the paired controls separately.

In Table 6, a positive sign of the inventor's value means that after patenting the indicator is higher than prior to patenting. A negative sign means that after patenting the value is lower. Similarly, for the controls, the positive sign means that over time the indicator has increased and a negative sign means that over time the indicator has decreased. The top row of the table shows that both the inventors and the controls increase their network dimension (*size*) over

Table 6
Difference and significance – within groups.

Variable	(After minus before)	
	Inventors	Controls
Size	31.28 (5.16)* ψ	25.22 (3.98) ψ
Betweenness (normalized)	13.93 (3.63)* ψ	11.44 (3.57)* ψ
Brokerage (normalized)	0.21 (0.04)* ψ	0.20 (0.04)* ψ
Density	−5.73 (4.79)	−1.36 (2.93)
Constraint	−0.05 (0.21)* ψ	−0.02 (0.23) ψ

Standard error in parentheses. ψ: non-normally distributed; Kolgomorov–Smirnov test.
* $p < 0.05$.

time. We saw in the prior section that inventors had slightly smaller networks than controls prior to patent. When we look at the same difference 5 years after patent (Table 5, right panel), we appreciate that inventors network size over time have levelled-off and in fact it is now slightly larger than the network size of the controls, although the difference is still not statistically significant.

The second and third row of Table 5 (right panel) show that inventors, who were in general less central prior to inventing, after patenting lay in positions more similar than those of the controls. When we look at the differences within (Table 6), again we see a general tendency of both types of individuals to become more central in their networks and act more as brokers as the time passes. This is consistent to common wisdom. Although our data cannot observe directly so, we presume that individuals progress in their career, create or expand their lab and group of PhD students and post-docs and these changes result in greater network centrality and more brokerage in networks. However the interesting thing is that over time the inventors, who were more peripheral prior to patenting, partially catch-up to non-inventors, so that in the aftermath of patents there's no substantial difference among inventors and controls concerning their network *centrality* (Table 5, right panel). Conversely, as the time passes, both inventors and controls become less necessary in their networks to connect nodes that would otherwise be disconnected, as witnessed by evidence of decreasing *brokerage* (Table 6). There is evidence that inventors act less as a broker of relationships than non-inventors in the post-patenting period, but this difference is very small in magnitude and only weakly significant (Table 5, right panel). In general, we find no support for either H4a or Hypotheses 4a or 4b that patenting is associated to important alterations in the network structure of the academics.

We now move on to considering ego-network position. Recall that, consistently with Hypothesis 3a, inventors tended to have more cohesive networks before patenting. We therefore investigate Hypothesis 5a, which would consider less dense and more constrained networks as evidence of a deterioration in ego-network structure in the post-invention period. We see that over time inventors expand their connections by establishing new relationships more markedly than non-inventors, so that the *size* of the network increases, with a comparable intensity of relations (*density*). They however become more likely to act as cohesive forces in their network, by connecting subgroups of otherwise disconnected alters. This tendency is visible from the lower network *constraint* (Table 5, right), which captures the degree to which ties are redundant within a network. We see from Table 6 that network constraint tends to decrease with time in both groups, but the decrease registered in the group of the inventors is sharper so that, after patenting, inventors are less constrained than controls. Therefore we find no support for the deterioration Hypothesis formulated in 5a.

In general, the broader evidence points towards denying that inventorship deteriorates the knowledge generation capacity of the networks of academic inventors. None of the evidence we have provided shows that after patenting inventors become more peripheral or secluded form the scientific community.

5. Conclusions

We investigated the co-authorship networks of academic inventors prior and after patenting. Because there are clear effects of life-cycle and trend that make inventors more productive over time and more central in their communities, comparing measures based on the publications of a group of scientists over time may cause spurious associations. To overcome this problem, our research strategy was to assemble a paired sample of scientists who have never obtained a patent and use this sample as a benchmark group. The assumption is that the control group would be affected by life-cycle and trend effects in ways similar to those affecting the inventors. We would therefore interpret as significant not the levels or the variations (higher or lower) of a certain indicator over time, but the differences between the levels of the inventor group against its target. Comparability across the samples was achieved by means of a one-to-one pair-matching strategy based on the estimation of a propensity score. The latter was made to depend only on predetermined and exogenous variables. The sample we used offers multiple advantages, including being geographically confined and disciplinarily homogeneous.

Overall, we did not find support for our hypotheses that inventive activities are associated with a broader network in the pre-invention phase or with a more central and brokering position of the inventors. The networks of the inventors were of comparable size to those of the controls prior and after patenting and inventors were in fact even slightly more peripheral in their networks than controls. Inventors were conversely found to be in denser networks with more redundant links than those of non-inventors. Our results suggest a positive correlation of denser networks and inventorship, but do not imply causality. Based on the contributions of the knowledge generation capacity of networks, we can state that a positive association might be consistent with the view that denser networks are more useful in conveying fine-grained information and in promoting trust and learning among nodes.

We also studied the network dimension, network position and ego networks of the inventors in the period that follows the invention of a patent, to inquire whether inventions are associated with alterations in the knowledge generation capacity of networks, with a specific focus on looking for any evidence of increased separation of the inventors from the academic community. The evidence we show points at denying any hypothesis of separation or isolation of academic inventors. We saw that in the years after patenting, inventors keep extending their networks particularly towards otherwise disconnected subgroups. We see this evidence as complementary to prior studies of social networks of academic inventors (Lissoni, 2010; Toole and Czarnitzki, 2010).

6. Limitations and avenues for future research

Our work presents some limitations. First, we observe only inventors of patents assigned to their respective universities of affiliation. We are aware from previous studies (Balconi et al., 2004, Breschi et al., 2008) that this is not the prevailing invention strategy in Italy. While most studies focusing on the possible rivalry between academic patenting and publishing patterns do not find support for such trade-off, Czarnitzki et al. (2009) using a German sample show that patents assigned to non-profit organizations (incl. individual ownership of the professors themselves) complement publication quantity and quality, whereas patents assigned to corporations are negatively related to quantity and quality of publication output.

Second, we broadly considered patent applications filed in different countries. We are aware that this approach might not control for the heterogeneity regarding the number of countries covered by a patent, as well as the different filing strategies. Future studies should investigate further in this direction and explore the differences, if any, in the corresponding social network structures.

More generally, as is often the case in social network analysis, accurate sampling at the individual level might not necessarily lead to stronger external validity with respect to the network results. This is particularly true when, as in our case, one looks at longitudinal data. While we followed all of the standard procedures normally used in these cases and normalized all indexes and defined comparable network structures over the different time intervals, some problems with the computation of the different indexes may still remain. This reasoning also led us to explicitly rely on a limited number of indicators that have been proven to be less sensitive to computational problems.

An additional problem relates to the possible presence of direct and indirect links of inventors with the treatment group. While we cannot rule out completely this potential source of bias, we can reassure the reader that, if a bias exists, it is indeed extremely limited. Inventors and non-inventors jointly co-authored the 1.1% (54) of their papers. Common co-authors account for 1.6% (106) of total coauthors, while common acquaintances – collaborations among co-authors – account for 3.1% (191).

Name matching is source of potential bias in most publication- and patent-based research, and it is particularly challenging in social network analysis, where the number of acquaintances scale-up at the power law. To minimize the incidence of mismatching we relied on Scopus' numerical identifiers – assigned directly by Scopus, after running disambiguation algorithms – to retrieve publications for our focal scientists and all of their co-authors. As these techniques become more precise and widely available, more extensive analysis could be performed to verify the robustness of our findings.

Finally, to distinguish between inventor and non-inventors we relied on a self reported indication of the latters. While there were no reason to deny inventorship when asked and indeed several others responding to the questionnaire used did so although they were never listed as inventors in Academic patents, we could formally rule out only that our controls were not listed as inventors in Italian academic patents.

Despite these shortcomings, we believe that our paper offers a unique and original contribution to the debate about the characteristics and behaviour of academic inventors. Future work should extend our comparative analyses to include multivariate modelling of social network structure in the immediacy of the inventing event. It should also extend the investigation to more scientific disciplines and to different institutional environments.

Acknowledgements

The authors acknowledge support from the Italian Ministry of Education, University and Research, under the programme FIRB RBNE037AWA_003 Italian Research in Innovation Systems (IRIS) and the Italian Ministry of Innovation and Economic Development (Direzione Generale Lotta alla Contraffazione, Ufficio Italiano Brevetti e Marchi) under the project "Patiris" (http://2.228.75.154/eprpatentsopendata). The sample of academic inventors was drawn from a prior work of Nicola Baldini and Maurizio Sobrero. The publication records used in the final version of the study were obtained by courtesy of Scopus–Elsevier. The authors wish to thank M'hamed el Aisati and Massimiliano Bearzot for assisting in the preparation of the data. The authors wish to thank Daniela Defazio, Carlo Vercellis, Soon Moon Kang, Rajesh K. Chandy, Simone Santoni and three anonymous referees for insightful comments.

References

Adams, J., Black, G.C., Clemmons, J.R., Stephan, P.E., 2005. Scientific teams and institutional collaborations: evidence from U.S. universities, 1981–1999. Research Policy 34, 259–285.

Agrawal, A., Henderson, R., 2002. Putting patents in context: exploring knowledge transfer from MIT. Management Science 48 (1), 44–60.

Ahuja, G., 2000. Collaboration networks, structural holes, and innovation: a longitudinal study. Administrative Science Quarterly 45 (3), 425–457.

Allen, T.J., 1977. Managing the Flow of Technology: Technology Transfer and the Dissemination of Technological Information within the R&D Organization. MIT Press, Cambridge, MA.

Allison, P.D., Stewart, J.A., 1974. Productivity differences among scientists: evidence of cumulative advantages. American Sociological Review 39, 596L 606.

Audia, P.G., Goncalo, J.A., 2007. Past success and creativity over time: a study of inventors in the hard disk drive industry. Management Science 53 (1), 1–15.

Azoulay, P., Ding, W., Stuart, T., 2007. The determinants of faculty patenting behavior: demographics or opportunities? Journal of Economic Behavior & Organization 63 (4), 599–623.

Azoulay, P., Ding, W., Stuart, T., 2009. The impact of academic patenting on the rate, quality, and direction of (public) research output. The Journal of Industrial Economics 57 (4), 637–676.

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.

Baldini, N., Grimaldi, R., Sobrero, M., 2006. Institutional changes and the commercialization of academic knowledge: a study of Italian universities' patenting activities between 1965 and 2002. Research Policy 35, 518–532.

Baldini, N., 2004. Cambiamenti istituzionali e processi innovativi: valorizzazione della ricerca universitaria italiana attraverso i brevetti. University of Bologna, Bologna, Italy, Unpublished doctoral dissertation.

Baldini, N., Grimaldi, R., Sobrero, M., 2007. To patent or not to patent? A survey of Italian inventors on motivations, incentives and obstacles to university patenting. Scientometrics 70 (2), 333–354.

Bertrand, M., Duflo, E., Mullainathan, S., 2004. How much should we trust differences-in-differences estimates? The Quarterly Journal of Economics 119 (1), 249–275.

Bonaccorsi, A., Dario, C., 2007. Universities and Strategic Knowledge Creation. Edward Elgar, Cheltenham, UK.

Borgatti, S., 2005. Centrality and network flow. Social Networks 27 (1), 55–71.

Borgatti, S., Everett, M., Freeman, L., 2002. Ucinet for Windows: Software for Social Network Analysis. Analytic Technologies, Harvard, MA.

Breschi, S., Lissoni, F., Montobbio, F., 2008. University patenting and scientific productivity: a quantitative study of Italian academic inventors. European Management Review 5, 91–109.

Burt, R.S., 1992. Structural Holes: The Structure of Competition. Harvard University Press, Cambridge, MA.

Burt, R.S., 2004. Structural holes and good ideas. American Journal of Sociology 110 (2), 349–399.

Burt, R.S., 2010. Neighbor Networks. Oxford University Press, Oxford.

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, 303L 319.

Calderini, M., Franzoni, C., Vezzulli, A., 2009. The unequal benefits of academic patenting for science and engineering research. IEEE Transactions on Engineering Management 56 (1), 16–30.

Cochran, W., Rubin, D.B., 1973. Controlling bias in observational studies: a review. Sankyha 35, 417L 446.

Coleman, J.S., 1988. Social capital in the creation of human capital. American Journal of Sociology 94, 95L 120.

Crespi, G., D'Este, P., Fontana, R., Geuna, A., 2011. The impact of academic patenting on university research and its transfer? Research Policy 40 (1), 55–68.

Czarnitzki, D., Glänzel, W., Hussingere, K., 2009. Heterogeneity of patenting activity and its implications for scientific research? Research Policy 38 (1), 26–34.

Davies, K., 2001. Cracking the Genome. Inside the Race to Unlock Human DNA. The Free Press, New York, NY.

De Beaver, D.B., Rosen, R., 1979. Studies in scientific collaboration Part II. Scientific co-authorship, research productivity and visibility in the French scientific elite. 1799–1830. Scientometrics 1 (2), 133–149.

Defazio, D., Lockett, A., Wright, M., 2009. The impact of collaboration and funding on productivity in research networks. Research Policy 38, 293–305.

de Solla Price, D.J., 1963. Little Science, Big Science. Columbia University Press, New York.

Etzkowitz, H., 1983. Entrepreneurial scientists and entrepreneurial Universities in American Academic Science. Minerva 21, 198–233.

Fabrizio, K.R., Di Minin, A., 2008. Commercializing the laboratory: faculty patenting and the open science environment. Research Policy 37 (5), 914–931.

Feldman, M., Colaianni, A., Liu, K., 2005. Commercializing Cohen'Boyer 1980-1997. Druid Working Paper no. 05-21, Copenhagen, DK.

Forti, E., Sobrero, M., 2013. Patiris: Permanent Observatory on Italian Academic Patenting, http://2.228.75.154/eprpatentsopendata

Fleming, L., 2001. Recombinant uncertainty in technological search. Management Science 47, 117–132.

Fleming, L., Mingo, S., Chen, D., 2007. Collaborative brokerage, generative creativity, and creative success. Administrative Science Quarterly 52 (3), 443–475.
Franzoni, C., 2009. Do scientists get fundamental research ideas by solving practical problems? Industrial and Corporate Change 18 (4), 671–699.
Franzoni, C., Lissoni, F., 2009. Academic entrepreneurship: definitional issues, policy implications and a research agenda. In: Varga, A. (Ed.), Academic Entrepreneurship and Regional Development. Edward Elgar, London.
Freeman, C., 1991. Networks of innovators: a synthesis of research issues. Research Policy 20 (5), 499–514.
Friedkin, N.E., 1981. The development of structure in random networks: an analysis of the effects of increasing network density on five measures of structure. Social Networks 3 (1), 41–52.
Gould, R.V., Fernandez, R.M., 1989. Structures of mediation: a formal approach to brokerage in transaction networks. Sociological Methodology 19, 89–126.
Granovetter, M., 1983. The strength of weak ties: a network theory revised. Sociological Theory 1, 201L 233.
Hagstrom, W.O., 1965. The Scientific Community. Basic Books Inc., New York, London.
Hargadon, A., Sutton, R., 1997. Technology brokering and innovation in a product development firm. Administrative Science Quarterly 42 (4), 716–749.
Holmes, F.L., 2004. Investigative Pathways. Patterns and Stages in the Careers of Experimental Scientists. Yale University Press, New Haven & London.
Kretschmer, H., 2004. Author productivity and geodesic distance in bibliographic coauthorship networks, and visibility on the Web. Scientometrics 60 (3), 409–420.
Langlois, S., 1977. Les Reseaux Personnels et la Diffusion des Informations sur les Emplois. Recherches Sociographiques 2, 213–245.
Lee, J.J., 2009. Heterogeneity, brokerage, and innovative performance: endogenous formation of collaborative inventor networks. Organization Science 21 (4), 804–822.
Levin, S.G., Stephan, P.E., 1991. Research productivity over the life cycle: evidence for academic scientists. The American Economic Review 81 (1), 114–132.
Lissoni, F., 2010. Academic inventors as brokers. Research Policy 39 (7), 843–857.
McFadyen, M.A., Cannella, A.A., 2004. Social capital and knowledge creation: diminishing returns of the number and strength of exchange relationships. Academy of Management Journal 47 (5), 735–746.
Merton, R.K., 1957. Priorities in scientific discovery: a chapter in the sociology of science? American Sociological Review 22 (6), 635–659.
Morgan, S.L., Sørensen, A.B., 1999. Parental networks, social closure, and mathematics learning: a test of coleman's social capital explanation of school effects. American Sociological Review 64 (5), 661–681.
Murray, F., 2004. The role of academic inventors in entrepreneurial firms: sharing the laboratory life. Research Policy 33, 643L 659.
Murray, F., O'Mahony, S., 2007. Exploring the foundations of cumulative innovation: implications for organization science. Organization Science 18 (6), 1006–1021.
Nahapiet, J., Ghoshal, S., 1998. Social capital, intellectual capital and the organizational advantage. Academy of Management Review 23 (2), 242–266.
Nelson, R.R., Winter, S.G., 1982. An Evolutionary Theory of Economic Change. The Belknap Press of Harvard University Press, Cambridge, Massachusetts and London, England.
Nerkar, A., Paruchuri, S., 2005. Evolution of R&D capabilities: the role of knowledge networks within a firm. Management Science 51 (5), 771–785.
Perry-Smith, J.E., Shalley, C.E., 2003. The social side of creativity: a static and synamic social network persoective. The Academy of Management Review 28 (1), 89–106.
Phelps, C., Heidl, R., Wadhwa, A., 2012. Knowledge, networks, and knowledge networks: a review and research agenda. Journal of Management 38 (4), 1115–1166.
Reagans, R., McEvily, B., 2003. Network structure and knowledge transfer: the effects of cohesion and range. Administrative Science Quarterly 48 (2), 240–267.
Rowley, T., Behrens, D., Krackhardt, D., 2000. Redundant governance structures: an analysis of structural and relational embeddedness in the steel and semiconductor industries. Strategic Management Journal 21, 369L 386.
Siegel, D.S., Wright, M., 2007. Intellectual property: the assessment. Oxford Review of Economic Policy 23 (4), 529–540.
Simonton, D.K., 2004. Creativity in Science. Chance, Logic, Genius, and Zeitgeist, Cambridge.
Sobrero, M., 2000. Structural constraints, strategic interactions and innovative processes: measuring network effects in new product development projects. Journal of Management & Governance 4 (3), 1–25.
Stephan, P., Gurmu, S., Sumell, A.J., Black, G., 2007. Who's patenting in the university? Evidence from a survey of doctorate recipients. Economics of Innovation and New Technology 16, 71–99.
Stephan, P., 2012. How Economics Shapes Science. Harvard University Press, Cambridge, MA, London, England.
Stokes, D.E., 1997. Pasteur's Quadrant. Basic Science and Technological Innovation. Brookings Institution Press, Washington, DC.
Stovel, K., Shaw, L., 2012. Brokerage. Annual Review of Sociology 38 (1), 139–158.
Toole, A., Czarnitzki, D., 2010. Commercializing science: is there a university brain drain from academic entrepreneurship? Management Science 56 (9), 1599–1614.
Uzzi, B., Spiro, J., 2005. Collaboration and creativity: the small world problem. American Journal of Sociology 111 (2), 447–504.
Uzzi, B., Amaral, L.A.N., Reed-Tsochas, F., 2007. Small-world networks and management science research: a review. European Management Review 4, 77L 91.
Wasserman, S., Faust, K., 1994. Social Network Analysis. Methods and Applications. Cambridge University Press, Cambridge; New York.
Wuchty, S., Jones, B.F., Uzzi, B., 2007. The increasing dominance of teams in production of knowledge. Science 316, 1036–1039.
Zucker, L.G., Darby, M.R., Brewer, M.B., 1998. Intellectual human capital and the birth of US biotechnology enterprises. American Economic Review 88, 290L 306.