Laboratory of Economics and Management
Sant'Anna School of Advanced Studies

Piazza Martiri della Libertà, 33 - 56127 PISA (Italy)
Tel. +39-050-883-343 Fax +39-050-883-344
Email: lem@sssup.it Web Page: http://www.lem.sssup.it/

LEM
Working Paper Series

Scientific and Technological Regimes in Nanotechnology: Combinatorial Inventors and Performance

Andrea BONACCORSI*[†]
Grid THOMA[†]

* University of Pisa, Italy
[†] Sant'Anna School of Advanced Studies, Pisa, Italy

2005/13 March 2006

ISSN (online) 2284-0400

Scientific and Technological Regimes in Nanotechnology: Combinatorial Inventors and Performance

Andrea Bonaccorsi‡ and Grid Thoma‡‡

‡School of Engineering, University of Pisa

‡‡Sant'Anna School of Advanced Studies[1]

December 2005

Abstract

Academics and policy makers are questioning about the relation between science and technology in the emerging field of nano science and technology (NST) and the effectiveness of different institutional regimes. We analyze the performance of inventors in the NST using multiple indicators. We clustered patents in three groups according to the scientific curricula of the inventors. The first two groups are composed by patents whose inventors all are authors of at least one scientific publication in the NST or none of them have obtained a scientific publication in that field respectively. Thirdly, we isolated those patents that have at least one inventor, who is also author of at least one scientific publication in the NST. The underlining presumption of this classification is that of a proxy of different institutional complementarities of the inventive collective action in NST.

Keywords: science-technology relation, emerging field, nanotechnology, patent quality, inventive productivity

[1] Corresponding author: Grid Thoma, Laboratory of Economics and Management, Sant'Anna School of Advanced Studies, Piazza Martiri della Libertà 33, 56127 Tel +39328 3837430 fax +39050 883344 Pisa, Italy. E-mail: thoma@sssup.it

Introduction

Nano Science and Technology (hereafter also, NST) regards the understanding and the control of the matter at the nano scale, which is a billionth of a meter. There is consensus in the scientific community that NST broadly involves; i) research and technology development at the atomic, molecular or macromolecular levels, in the scale of approximately 1 - 100 nanometer range; ii) creating and using structures, devices and systems that have novel properties and functions because of their small and/or intermediate size; iii) the ability to control or manipulate on the atomic scale.[2]

There is also consensus among scientists on a candidate birth date of NST, that is 1981 when the *Scanning Tunneling Microscope* (US Patent 4343993; hereafter also, STM) was invented by Gerd K. Binnig and Heinrich Rohrer at the IBM Research Laboratory in Zurich. In 1986 they received the Nobel Prize for that discovery. The STM enables atomic-scale images of metal and semiconductor surfaces, which could not be obtained by the so-called *Topografiner*, invented by Russell Young in the late 1960s. The range of materials that can be imaged with a scanning device increased with the invention of the *Atomic Force Microscope* (US Patents 4724318 and RE33387; hereafter also, AFM) by Gerd K. Binnig in 1986.[3]

These enabling instruments were invented at and with the support of the IBM Corporation, which was interested in the scientific advances within the semiconductor industry. Quickly they realized that the STM and the AFM could be used in a vast array of scientific and technological fields as chemistry, biology, biotechnology, telecommunications, and many others.

In this paper, we will try to make sense and measure the importance of scientific knowledge in fuelling technological inventions in the NST. We contribute to the literature on science-technology interactions and on the industrial dynamics of emerging fields, providing recent

---

[2] The definition is that of National Nanotechnology Initiative (www.nano.gov). We validated this definition in several interviews with nano scientists, that we had during 2004-05.

[3] The invention was filed in August 1986, while a related scientific article was published 6 months before (Cfr. Gerber, G. Binnig, H. Fuchs, O. Marti, and H. Rohrer *Scanning tunneling microscope combined with a scanning electron microscope* Rev. Sci. Instrum. 57, 221-224 (Feb 1986))

and large scale evidence on a young industry and developing new methods for measurement and quantitative analysis. In particular, we develop a methodology to match information on scientific publications and patents, leading to a characterization of several *communities of inventors*, that implement science-technology interactions in alternative ways. The paper is organized as follows. In the next Section we briefly review the literature, pointing to limitations in existing methodologies and substantive explanations. In Section 2 we introduce descriptive statistics on NST and develop the proposed methodology. In Sections 3 and 4 we investigate the performance implications of alternative ways of arranging science-technology interactions, as made visible by the matching analysis. We propose some measures of patent quality and test various hypotheses on differences between indicators across patents produced by different communities of inventors. Sections 5 and 6 add new dimensions of performance, analysing the productivity of individual inventors and the propensity to found a start-up company. Section 7 reviews the evidence and opens new research questions.

1. Literature background on science-technology interactions

In the '90s the notion that technological developments are increasingly dependent on advancements in science has been proposed repeatedly. On one hand, the scientometric literature has raised the attention to the sharp increase in the number and share of Non-Patent Literature citations (NPL) in patents (Narin and Olivastro 1992; Narin, Hamilton and Olivastro 1997), suggesting that inventors increasingly use directly inputs from published scientific research. Patents may be based not only on the prior art documented in other patents, but in part or fully on new scientific knowledge. Since published scientific research results can be used to illustrate the state of the art against which the application has to be evaluated, patent examiners will then search for relevant references in the scientific literature. The logic of these references is to document the material that is held against the application. Based on this metrics, a taxonomy of industries based on the dependence on science can be developed (Grupp 1992; Heinze and Schmoch 2004; Tijssen 2004). More recently, techniques for tracing back network of patent citations have been developed

(Popp 2005; Verspagen 2005). Tracing back the full ramification of citations from more recent patents to historical ones can give insights on the underlying dynamics of knowledge.

On the other hand, industry case studies on biotechnology (McKelvey 1996; Orsenigo 1990; Owen-Smith et al. 2002; Zucker and Darby 1996), chemical and electrical engineering (Kenney and Goe 2004; Mowery and Rosenberg 1998) semiconductor and laser (Klepper 2001), medical instruments (Trajtenberg 1989; 1990) illustrated important examples in which the very definition of industrial applications was made possible only due to the discovery of new physical properties of nature. In these fields the origin of entrepreneurship can often be traced back to scientists from the academic world or to scientists in large and technologically advanced companies.

The importance of this literature can be better understood in relation to the broader theoretical treatment of the relations between science and technology and, more generally, on the conditions for the productive use of knowledge. In fact, the critique of linear models carried out in the literature in the 1980s (Kline and Rosenberg 1986; Rosenberg 1982) had made clear that technological knowledge is subject to a specific internal dynamics, relatively independent of scientific advancements. Firms benefit from science only indirectly (Pavitt 1990) and the paths through which scientific research is used for industrial innovation are not mainly direct collaboration but rather the constitution of human capital (Cohen, Levin and Mowery 1987; Nelson 1986; Rosenberg 1990).

In parallel to this institutional treatment, a conceptualization of the nature of technological knowledge has taken place, beyond highly stylized representations. Design knowledge started to be characterized as a collection of highly specific rules for problem solving and selection of acceptable solutions (Klein 1985; Stankiewicz 2000; Vincenti 1990), which constitutes an autonomous body of knowledge. Design is not applied research and engineering is not applied physics.

Against this background of criticism to the linear model and of articulation of relative independence of different types of knowledge, the discovery of the sharp increase in the “scientific content” of patents, inventions, and companies still lacks a rigorous theoretical treatment. Is it an evidence that the linear model still holds? Or is it an evidence of a fundamental change in production technology so that the flows of knowledge between scientific research and technology are less mediated and more direct? Or is it just a transitory stage in the long term evolution of industries?

In other words, there is a strong need to go beyond the stylized evidence collected with patent data and industry case studies in the 1990s, and build up a more general framework for the analysis of the productive use of knowledge.

This task, however, is made difficult by a number of limitations in the existing literature. On a substantive basis, the critique of linear models has still to generate a stream of studies on the specific non-linear ways in which science and technology interact, providing evidence on the nature and intensity of feedback loops and iterations. As an example, Stankiewicz (1997) made the important distinction between discovery-driven and design-driven innovation. In the former case innovation critically depends on understanding of nature, while in the latter it is driven by internal technical issues, largely influenced by applications and demand considerations. The identification of these cases and of the possible transition between them in long term technological evolution is still an unexplored issue. The literature has still to explore to a great detail the micro-mechanisms of generation, validation and transmission of knowledge between science and technology. The epistemic foundations of science-technology interactions are still unclear, after the pioneering analyses offered by Callon, Courtial and Laville (1991). A preliminary investigation, which is part of a larger research agenda, has been proposed with the notion of search regime (Bonaccorsi, 2005b). According to this notion, in leading sciences such as biotech, materials and nanotech, scientific fields may grow exponentially and follow a pattern of increasing diversity, driven by the specific combination between deeper understanding of properties of matter at low levels or resolution and design objectives.

On the methodological side, several shortcomings of the existing measures should be recognized. First of all, Non Patent Literature (hereafter also, NPL) citations suffer from an important limitation: it is not clear to what extent they are assigned by inventors or by examiners. It is well known that inventors primarily introduce references in the USPTO, while in the European system they are introduced exclusively by the examiners. Breschi and Lissoni (2004) claimed that, at least in the US patent system – since references are assigned by different actors, who quote mainly US references for reasons of availability and for different purposes - there is a severe distortion in the interpretation of data. The full validity of information on cited patents has to be established, given that the motivations for a patent to cite another patents are rather intricate and call upon legal and strategic considerations. Thus we face both measurement and validity issues.

Second, NPL citations do not convey any information on the degree to which the scientific content was able to generate valuable innovation. Since we know that the distribution of patents by degree of usefulness is extremely skewed, it is possible that patents with a high number of non-patent references are among those that are never used, and so have limited economic value. One approach to mitigate this limitation is given by a careful analysis of patent quality, using the indicators proposed in the literature initiated by Trajtenberg (1989; 1990) and fully developed by Jaffe, Trajtenberg and Henderson (1993).[4] There is sufficient evidence in the literature that the economic value of patents is associated with the number and quality of citations received in other patents (Hall, Jaffe and Trajtenberg 2005; Harhoff et al. 1999; Jaffe and Trajtenberg 2002). Harhoff et al (2003) and Lanjouw and Schankerman (2001) have suggested a different metrics, i.e. the existence of litigation for patents, implying that patents for which assignees are willing to pay for defence against infringement, have larger economic value.

More fundamentally, existing methodologies identify science-technology interactions using documents, not individuals. A relation is said to be in place if and only if a paper trail can be identified. This largely ignores the variety of motivations that may lead to citations.

Therefore, a new approach is needed to capture the complexity of interactions between science and technology. In this paper we develop a new methodology for tracing and measuring

[4] For a survey of the literature see Jaffe and Trajtenberg (2000)

these interactions, based on matching algorithms of names of individuals between different datasets. We try to match the names of individuals between datasets , i.e. publications (individuals as authors), patents (individuals as inventors) and companies (individuals as entrepreneurs or partners). For all individuals in our final dataset we know whether and when he/she has published, invented, and created a company, and we know all the associated characteristics (and related indicators) for his/her papers, patents, and companies. This approach does not take the patent as unit of analysis but rather the individual, i.e. the author, the inventor, or the entrepreneur. The analysis focus on building up time series dataset and focus on the mobility and carriers of the individual across different institutional settings (Bozeman and Mangematin 2004; Gambardella, Giuri and Mariani 2005; Shane 2004). In combining different roles around the same individual we will be able to build up the full profile of inventors and to put forward several conjectures about the intrinsic dynamics of science and technology.

Assuming individuals, rather than patents or papers, as unit of observation has several advantages. First, it helps in identifying heterogeneities in the performance of inventors' action. Simple but robust classifications, such as the one we propose in this paper, are powerful enough to illuminate the existence of several paths to innovation followed by scientists and inventors during their professional life. In turn, this contributes to the identification of the dynamic nature of science-technology relations.

Second, the value of scientific production and of invention can be addressed directly. The literature on patent quality, emphasizing the relation between the value of the invention and successive citations in other patents, uses indicators that are internal to the patent system. Studies such as Shane (2004) try to trace forward the history of patents from academia down to the creation of startup firms, but do so with a limited focus (i.e. a large US university). Using our dataset we are able to disentangle the full path leading from scientific publications to patents (or viceversa) or from scientific publications and patents to the creation of a company, as well as any other relevant path. Different performance measures – scientific, technological and economical – are regressed over their initial characteristics and the carries paths, as in the studies of the industrial dynamics of firms. In spite of their great interest in understanding the very micro dynamics of the innovation process, these exercises face endogeneity problems of the explanatory variables to performance, which we have to discuss accurately.

Finally, focusing on inventors allows us to carry out longitudinal analysis at individual level, helping to identify and follow the steps of careers that lead to scientific activity, invention, and entrepreneurship. In this sense our analysis is a first preparatory work that might be followed by extensive examination of scientists' resumes, in order to trace the evolution of careers and the patterns of mobility.

2. Search regimes in Nanoscience and Technology

NST is an extremely interesting case in which the micro-mechanisms of science-technology interactions and the origins of entrepreneurship can be detected with high precision, due to the novelty of the field and the relatively richer documentation available. Following the notion of search regimes, scientific fields may be characterized with respect to a few stylized properties of their search process (Bonaccorsi 2005a; 2005b). First of all, the rate of growth in production of scientific results: regimes that exhibit exponential growth (or grow at significant larger rates than average) have completely different properties than regimes that grow linearly. Second, the degree of diversity of directions of research: in some areas all research programs converge along a few areas, usually associated with crucial experiments based on a commonly held body of theory, while in other areas the agreement on general theories generates a proliferation of (weakly or strongly) competing hypotheses and research programs, following a divergent dynamics. Third, the importance and nature of complementarities in the knowledge production process. Based on these dimensions, a number of disciplines, called new leading sciences, have been identified, including life sciences after the molecular biology revolution, computer science, materials science, and nanoscience. These broad disciplines share the following properties: they have been growing exponentially or much more than average for a long period, they follow a dynamic process of divergent search, and they are based on institutional and human capital complementarity. We will use these dimensions to characterize the emerging field of NST.

2.1. Data on Nano Patents and Publications

The search strategy for nanotechnology patents had to be mainly based on keywords, since the specific IPC-subclass B82B for this field was introduced in the year 2000 and does not cover former years. Therefore, it contributes only to a very small part of all documents identified (see Table 1).

Table 1 US Patents in IPC B8 Micro-Structural Technology and Nano-Technology

[Table 1 about here]

We used a keyword search strategy suggested by Fraunhofer ISI Institute in Karlsruhe, which we found to be the most complete and validated by experts among the static keywords methodologies (Fraunhofer-ISI 2002).[5]

In the following analysis we selected data from USPTO. Given the important role of the United States as a locus of technical change in the last decades, we think that this limitation to U.S. patenting activity does not constitute a serious drawback for a preliminary investigation of this kind. We executed the ISI strategy in the title and the abstract of a patent.[6]

Nanotechnology confirmed to be a significant phenomenon. We obtained a sample of 4828 patents granted before May 2004, classified by 1192 examiners in 331 three digit U.S. technological classes. The collective action in the NST involves more than 8000 inventors located in more 3000 cities in 52 countries. The patents are assigned to more 1900 companies, located in more than 800 cities in 37 countries.

[5] In order to circumvent the problem of an accidental selection of keywords given by experts, they listed all terms in the patent database beginning with "nano". An expert in the NST assessed for each term whether it is used in the context of nanotechnology and whether it indicates an unambiguous relation to this field. 40 keywords queries have been obtained, identifying singularly a field. See appendix 1 for more information.

[6] The source of data is constituted by the Delphion patent database (DPD), which is an on-line proprietary database, accessible from www.delphion.com. It includes data from different national Patents Offices. In particular, it offers a complete text and images of all patents issued by the US Patent and Trademark Office (USPTO) since 1971. It offers the possibility to ask in a very intuitive manner the remote database.

We performed the same search keywords strategy for publications, cleaning it from the references to technological classes. Follow up interviews with scientists in the field validated this keywords search strategy for publications (Beltram 2005). We searched both in the title and keywords of a publication. The data source is constituted by SCI and SSCI of ISI database for the years 1988-2001. We obtained a pool of 93.149 publications, authored by 119.640 individuals, affiliated to 13.752 institutes.[7]

More generally our dataset considers only a “seed”, from which a more complete dataset might be generated applying iteration techniques using for example citation links in order to include not only papers that show one of the keywords, but also papers cited in the seed that might be part of an emerging field.[8]

2.2. Rate of Growth of Nano Science and Nano Technology

In the case of Nanoscience, it is clear that not only individual fields (such as carbon nanotubes, or nanocoatings or nanobiotechnology) but the whole discipline had an impressive growth. In less than ten years an army of almost 120,000 scientists worldwide was mobilized around the new discipline. Several thousands new institutions worldwide entered the field. The scientific output of such collective action can be measured in about 100 000 publications.[9]

In the case of Nanotechnology, we can notice a less stable dynamic of inventive output in terms of growth rates. We have an impressive growth of production of patents, especially in the last years (1996-2002). The USPTO has patented several thousands of inventions in Nanotechnology, with around 6600 files at the end 2005. In Figure 2, data on USPTO patents (used throughout this paper) have been integrated with data on EPO patents and WIPO PCT patents, using the same keyword structure, in order to give a glance to other sources. While European patents are at the very

[7] The unit of analysis is constituted by the parent institute. Data on individuals have been cleaned correcting for classical distortions. Data on affiliations are raw data and should be treated carefully. The cleaning of affiliations is underway.

[8] In a latter of the PRIME project a new dataset, based on similar procedures, has been developed by Zitt (2006). In future works we plan to investigate such type of database.

[9] The database mentioned in footnote 8 is constituted by around 180 000 publications in the period 1991-2004.

low levels, patent cooperation treaty patents, that include Japan, stand very high. In particular, Japanese patents exceed 7000 in the period 1976 – 2004. (7469 up to 31/12/2004)

[Figure 1 about here]

Figure 1 Cumulate arrivals in Nano-science (Source: our elaboration from ISI)

[Figure 2 about here]

Figure 2 Cumulate arrivals in Nanotechnology, May 2004

Notes: The patents for the 2004 include only those that have been granted before May 2004.

What we observe is a dynamic process characterized by an average growth rate which is far larger than average in science and engineering for all the years in the period. To give an order of magnitude total publications in the SCI grew by 3% in the period 1990-93 and around 1% in 1998-2001. The peak rate was 14% in 2003, following a drop by 2% in 2002 (Our elaboration on WoS-ISI data).

However we do not observe exponential growth in publications, while the growth of applications in patents is exponential until 2002. This is in contrast with recent interesting results of Zucker and Darby (2003), who find exponential growth in publications in the period 1980-2000. We interpret this difference with respect to methodology and substance. Zucker and Darby's dataset was built using the generic word "nano*" while our dataset follows a static combination of keywords. Therefore the two datasets have different statistical properties. Our dataset may underestimate the production of papers that use only completely new keywords being based on a static keyword list. On the other hand, Zucker and Darby dataset may include false positives - for example nano-seconds, nanoplankton, nanoflagelate and others- and is subject to manipulations from authors that include the word "nano" only by fashion.[10] In addition, Zucker and Darby data

[10] The use of single keyword search strategy, in general, might have many drawbacks. For example, Zucker and Darby identify around 2000 USPTO granted patents in the period 1981-2000 using the word "nano" either in the title or in the description. Following the ISI Fraunhofer methodology, on the contrary, we identified no less than 2700 granted patents at the USPTO in the same period.
Trying to replicate their approach, we obtained different results. On one hand, we executed a search based only on the word "nano" in the USPTO archive (www.uspto.gov): (TTL/nano$ OR SPEC/nano$) AND ISD/1/1/1980->1/1/2001.

refer to the period 1980-2004, while our sample is limited to the period 1988-2001. As a matter of fact, the share of nano-articles per 1000 science articles started to grow significantly only after 1990 (Zucker and Darby 2003; p. 55 Figure 1).

From a substantive point of view, we may witness a catching-up effect in recent years, due to legitimating of keywords included in the sample (and hence a decrease in growth rates). Comparing publications with patents it is clear that the impressive growth in publications took place 5-7 years before the surge in patenting. This is a relatively short period for the real economic effects.

2.3. Degree of diversity

As we discussed in previously there are strong reasons to expect divergent dynamics in NST. Within the overall emerging field of NST there are several well-identified subfields. Fraunhofer ISI suggests a classication based on specific combinations of keywords. We used this classification for both publications and patents, and labelled the most important combinations.

Figure 3 shows the distribution of publications by subfield. The top 5 fields, which we labelled nano-biotechnology, nano-instrumentation, nano-electronics (fabrication and materials) and nano materials, account for around 60% of all publications. A somewhat lower degree of concentration is visible in patents (Figure 4), where a concentration ratio of 60% level is reached by the top seven fields.

This evidence is consistent with published reports on the main areas of nanoscience and technology. Basically, applications of nanostructures to life sciences, electronics, and new materials absorb the majority of published research, while instrumentation is a transversal area. The distribution of subfields in patents slightly differs, due to the relatively immaturity of patentable devices in nano-biotechnology field.

We found several tens of thousands patents (more than 50000) over the 1980 – 2000 period, clearly an unrealistic amount. More interestingly using the word “nano”, one cannot include the US04343993 patent granted in 1982 to Binning and Roehrer for the invention of scanning tunnelling microscope: we tested directly the suffix “nano” is missing in 11-page document of that patent.

[Figure 3 about here]

Figure 3 Nano-publications by field

[Figure 4 about here]

Figure 4 Nano-patents by field

An important dimension of diversity is dynamic: is diversity increasing or decreasing over time? Do research programmes move apart from others, or do they get closer as they proceed? Or, put in other terms, do we observe a pattern of divergence or a pattern of convergence among research programmes?

Following the literature on field delineation in bibliometrics and scientometrics, we would say that research programmes might be characterized by keywords and combination of keywords. A full appreciation of the dynamics of search would require the observation of the way in which keywords originate and cluster together over a long period. While this the object of future research, we offer here a preliminary exploration of this issue by looking at the industrial dynamics of keywords.

If a discipline is subject to a divergent search regime, there will be many new keywords appearing per unit of time. As we can notice in Figure 5, there is constant and linear growth in the number of entries of new keywords. We defined new keywords (new entrants) those that appear for the first time in the dataset at any point in time. If a keyword has been used at least once in any year before an observation, it is not labelled new. Consequently, each year the total number of keywords used includes new keywords and old ones, drawn from the set of keywords that appeared for he first time in any period before.

[Figure 5 about here]

Figure 5 Stocks, use, entry, and exit of keywords at the overall population by year

The ratio between number of new keywords and the total number of keywords used each year is an indicator of turnover. As it is shown in Figure 6, this ratio starts at very high levels at the beginning, given that, by construction all the keywords are new in the first year and the set of old keywords starts small. However, it is worth to notice that the ratio stands more than 40% even at the end of the period. Each year the scientific community is nanoscience is able to generate more than 10, 000 new words describing their, i.e. 40% of used keywords.

We find this ratio a remarkable indicator of turbulence, derived from a divergent dynamics of proliferation of new research programmes.

As we have shown in Bonaccorsi (2005b) similar levels of turnover can be found in computer science, where the ratio levels at 40% at the end of the observation period. However, in that case we examined the publications of the top 1000 scientists worldwide, not the whole community, and one might argue than top scientists have better that average ability generating new research topics on a continuous basis. Therefore we conclude that nanoscience fully meets the second requirement to qualify as leading science, i.e. divergent dynamics.

[Figure 6 about here]

Figure 6 Ratio between number of new keywords entered and total number of keywords used by year

2.4. Level of complementarities: interface between science and technology

In leading sciences institutional and human capital complementarities are crucial to the development of research. Institutional complementarities arise because researchers with different perspectives on the object of research are needed in order to generate discovery and invention. Because of different professional background, these researchers are usually affiliated to different institutional actors (e.g. public research, industry, hospital, public administration, regulatory bodies), bringing in the search process peculiar cognitive attitudes and operational practices. Human capital complementarities arise because the epistemic nature of discovery requires the deployment of several disciplinary competencies, even within the same team and/or the same institution.

Institutional and human capital complementarities are the fundamental mechanism for realizing effective science-technology interactions. In practice, however, there are several possible ways to implement these types of complementarities, for example, in terms of intensity of interaction between researchers, flows of communication, and pattern of mobility. Given the relation between discovery and design already discussed, in NST the fundamental complementarity is between industry and academia.

To investigate the relation between nano science and technology we clustered patents in three groups according to the scientific curricula of the inventors. The first two groups are composed by patents whose inventors are all authors of at least one scientific publication in the NST (only-authors) or on the contrary none of them have obtained a scientific publication in that field respectively (only-inventors). Thirdly, we isolated those patents that have at least one inventor, who is also author of at least one scientific publication in the NST (author-inventors).

This taxonomy is a first contribution to the development of new metrics of science and technology relations that are based on individuals and communities rather than on paper trails. Individual based indicators can be used only after appropriate matching procedures between different datasets. [11]

In this paper we present only a crude taxonomy based on the extreme discrete values and very sharp thresholds. Basically, we ask that all/none members of the group have a zero level of a selected variable. If patent inventors have zero publications they are labelled “only-authors”; if none have zero publications they are labelled “only-inventors”; if none of the two extreme cases is true, a residual group called “author-inventors” is created.[12]

[11] This types of matching cannot avoid the problem of synonymy or homonymy. In the past, the literature has faced the problem by collecting CV data and the matching is usually done by ad-hoc and tedious procedures. (For a set of contributions see Bozeman, B. and V. Mangematin (2004). "Editor's introduction: building and deploying scientific and technical human capital." Research Policy 33: 565-568.
Given the dimensions of our dataset this methodology is scarcely feasible. In future works we aim to develop automatic names matching algorithms, based on multiple indicators scores methods.

[12] We executed a simple matching procedure as follows:
- Inventor vs Inventor: Drexler K Eric = Drexler K$ Eric
- Author vs Author: Drexler K Eric = Drexler-KE
- Author vs Inventor: Drexler-KE = Drexler K$ Eric

Eric K. Drexler is considered one of the “fathers” of Nano Science and Technology. His main famous book was published in 1986, *Engines of Creation: The Coming Era of Nanotechnology*. New York: Anchor Press/Doubleday.

The outputs of the above matching procedure of the names of the individuals are collected in Figure 7.

[Figure 7 about here]

Figure 7 Composition of communities of inventors in nanotechnology

Summing up only-authors and author-inventors groups, it turns that the large majority of patents (above 66%) have at least one inventor that is also an active scientist. Therefore, we find evidence of a highly interconnected knowledge system, in which the transformation of scientific achievements into patentable results and of both into commercial ventures is very rapid, taking place through the multiple roles played by scientists themselves. Interestingly, the group of author-inventor grows more rapidly than the two other groups (see Figure 8)

[Figure 8 about here]

Figure 8 Entry of individuals by community

Surprisingly, this simple classification has strong validity and good predictive power. Although future research might develop more fine-grained taxonomies, this is a promising start. We interpret the three groups as approximations of different forms of institutional complementarities.[13]

Following our taxonomy, only-inventors patents originate from inventors that have never published in the field. Most likely, these inventors are industrial researchers for whom at least one the following propositions holds true: i) they do not come from the academic career; ii) they are not allowed to publish in the open literature; iii) they work on applications that could not be published in the scientific literature; or iv) they work in institutions for which publishing is encouraged but

[13] Remember that co-invention relations are stronger than, for example, co-authorships or in documents co-occurrence of keywords of co-citations. Patents are expensive, create exclusive rights, may originate streams of revenues. Being recognized as an inventor is a right that has full legal implications. Therefore, if more individuals are recorded as co-inventors of a patents, their mutual relation is likely to be institutionalized.

have no original results to submit, or publish in fields different than NST, or publish in non ISI journals, or are technical staff.

The literature on industrial research has noted that companies that allow their researchers to publish in the open literature do so because they want to gain access, visibility and reputation in the scientific community. Industrial researchers that do not have a track record of published articles have more difficulties in getting access to critical external knowledge and to built up absorptive capacity. Similarly, companies that want to access external knowledge may hire people with a track record in public research (Gittelman and Kogut 2003). Assuming that the case sub iv) is less important, we could conclude that this group identifies rather precisely industrial R&D that has no or very weak institutional complementarity with public research.

On the contrary, in the other two groups all or at least some of the inventors have published at least once in the literature. The only-authors group is interesting, because it is formed by individuals for whom at least one of the following propositions holds true: i) they all work in public research organisations and carry out both publishing and patenting; ii) they work in companies, but all of them are encouraged to publish in the open literature; iii) they work in companies, but all of them come from previous public research career; iv) any combination between proposition ii) and iii); v) part of them work in companies and are permitted to publish/have published in their early career, part of them work in public institutions, and they collaborate and patent together.

Basically, these three communities represent different ways of organizing the complementarity between industry and academia. only-inventors access scientific knowledge through codified channels, such as publications and conferences. They do not involve professional scientists in their inventor community, nor have inventors with a track record of scientific publications in their past. only-authors have full access to scientific knowledge, but have no structured and permanent access to knowledge of potential industrial applications, i.e. have difficulties in combining knowledge about physical structures (discovery) with knowledge about design. Since the knowledge about design is structurally more idiosyncratic and less codified than scientific knowledge, only-authors that come from academia implement the minimum level of complementarity. Finally, the community of author-inventors gets access to both discovery and design knowledge, either in the codified and in the embodied form, on a permanent and organized way. They realize the most of complementarity. We anticipate that the more intense the complementarity, the more effective the invention.

To explore the different institutional search regimes, we classified the assignees of our patent dataset in three groups, according if they are private companies, public research organizations (hereafter also, PROs) and individuals. We found that around 68% of the patents are assigned to private companies while PROs own 26% of them.

[Figure 9 about here]

Figure 9 Distribution of patents by community in relation to assignee type

Figure 9 allows us to conclude that there is a higher probability that the inventors-only are employed primarily in companies.

[Figure 10 about here]

Figure 10 Distribution of patents by assignee type in relation to community

At least one active scientist is involved in the invention for around 80% and 60 % of patents of PROs' patents and private companies, respectively.

3. The impact of complementarity on inventive performance

As we have seen from the matching exercise, there is considerable overlapping between different social roles of researchers in the production of public knowledge and of appropriable results. Two thirds of patents have been produced by a team of inventors in which at least one inventor has published in the field.

We propose that the specific pattern of interaction between different types of inventors has an impact on the quality of patents, as measured in the relevant literature (see below).

Patents in the only-author group originate from two different patterns: researchers in the public sector that publish and patent, or industrial researchers in companies that maintain close relations with academia. In future research we will disentangle the two communities, going down to individual data on patents and publications and tracing the career across various affiliations. For the time being, we predict that this community will produce patents with lower quality than the author-inventor community.

The latter will have the best performance, because here all groups of inventors benefit from at least one who has experience in publishing (currently or in the early career), while the others may well be pure industrial researchers. In this sense, the author-inventor community is characterized by the highest degree of institutional complementarity. In fact, in the only-author group it is required that all inventors are also active in publishing. Therefore, if the patent is done within a company, only industrial researchers that are all exposed to the scientific community will be involved, while if it is done in collaboration between industry and academia, public researchers collaborate only with those industrial researchers that also publish. In one word, in the only-author group there is low distance (cognitive, social, institutional distance) between co-inventors. In practice, these patents do not involve industrial researchers or technicians that, while not personally involved in publishing, may have deep knowledge of the technology. Therefore we expect that author-inventor community, based on interaction patterns that materialize high levels of institutional complementarity, will exhibit better quality in the inventive activity.

The relation between the two groups of only-author and only-inventor is a bit more complex. On one hand, patents originated by groups of inventors in which no one has published (only-inventors) are characterized by access to scientific knowledge through codified channels, leading to a prediction of poor quality. On the other hand, the community of only-authors may be heavily influenced by academic inventors, that do not have access to embodied design knowledge, also leading to a prediction of poor quality. Balancing these two effects is an empirical matter. We therefore do not put forward a specific proposition.

The proposition on superiority of author-inventor community extends to the overall quality of patents, and to the upper tail of the distribution of inventive productivity. We rank inventors by number of patents produced and investigate whether in the upper part of this distribution the community of author-inventor is more than proportionally present. This qualification is important, given the skewness of the distribution of inventive productivity.

Finally, we extend this proposition to one of the most important consequences of the quality of patents, namely the probability that an inventor becomes a founder of a new company.

We therefore put forward the following testable propositions:

Proposition 1

(a) The quality of patents of the inventor-only community will be lower than the quality of patents of the author-inventor community

(b) The quality of patents of the author-only community will be lower than the quality of patents of the author-inventor community

Proposition 2

The productivity of author-inventors’ inventive activity will be higher if counted by the number of the patents produced in the top percentiles of the distribution of patents/inventors.

Proposition 3

Given the higher technological permanence of author-inventors, we will observe more of them as founder of companies.

4. Patent Quality across Different Communities

In this section, we suggest some indicators that can be used in measuring the performance of inventive activity. Patents indicators have been widely used as proxy of intensity and quality of innovations in different fields of social sciences. In this paper we will not review that literature, but a broad survey can be found in Jaffe and Trajtenberg (2002). We will follow a multiple indicators approach as suggested by (Hall and Trajtenberg 2004; Henderson, Jaffe and Trajtenberg 1998; Lanjouw and Schankerman 2004).

One of the most used patent indicator are patent references or citations. In patents, citations serve an important legal function, since they delimit the scope of the property rights awarded by the patent. Thus, if patent B cites patent A, it implies that patent A represents a piece of previously existing knowledge upon which patent B builds, and over which B cannot have a claim. The applicant has a legal duty to disclose any knowledge of the prior art, but the decisions regarding which patents to cite ultimately rests with the patent examiner, who is supposed to be an expert in the area and hence to be able to identify relevant prior art that the applicant misses or conceals. The presumption is thus that citations are informative of links between patented innovations. Hence, citations made (or backward) may constitute a paper trail for spillovers, i.e., the fact that patent B cites patent A may indicate knowledge flowing from A to B.

Some scholars have suggested that large numbers of citations to others suggests that the particular innovation is likely to be more derivative in nature (Lanjouw and Schankerman 2004). This is more evident when citations are within the same technological field. Hence, dispersion indexes measures that take into account the distribution across technological classes have been elaborated like the originality index, suggested by (Henderson, Jaffe and Trajtenberg 1998)(Henderson, Jaffe and Trajtenberg 1998), which is 1 minus the Herfindal index of backward citation across technological classes.

As citations are a paper trail for spillovers, received (or forward) citations may be telling of the importance of the cited patent (Trajtenberg 1989; 1990). Received citations over the long term indicate an innovation has contributed to future research. Citations soon after patent application suggests rapid recognition of its importance as well as the presence of others working in a similar area, and thus the expectation of a valuable technological area. Due to the problem of citation lag we cannot trace forward citations over the long term in NST. We suggest a count indicator (FWcit5) of forward citations over the short term, in particular five years between the publication date of the cited and the application date of the citing. We expect FWcit5 to be positively correlated to the patent value.

The claims in the patent specification delineate the property rights protected by the patent. The principal claims define the essential novel features of the invention and subordinate claims describe detailed features of the innovation. The patentee has an incentive to claim as much as possible in the application but the patent examiner may require that the claims be narrowed before granting. The number of claims could be considered an indication that an innovation is broader and of greater potential profitability (Lanjouw and Schankerman 2004).

Patent family size, measured as the number of jurisdictions in which a patent grant has been sought, should be directly related to the expected (private) value of protecting an innovation and thus to the value of the innovation, since applying for protection in each country is costly (Lanjouw, Pakes and Putnam 1998). Hence, family size (*Family*) may be particularly well suited as an indicator of the economic value of patent rights.[14]

4.1. Empirical Evidence

In the following analysis we decided to consider only the utility granted patents with an application date comprised between 1988-1999. We decide to start from 1988 because the period of our dataset of scientific publications in the NST covers 1988-2001. Secondly, our patents' dataset is limited to 1999, since the FWD5 cannot be computed for the following years.

[14] We obtained the family size from Delphion-Derwent, a private vendor. Delphion extends the family data received from INPADOC to create the family unit and considers the set of patents filed with different patenting authorities that refer to the same invention.

Table 2 Descriptive Statistics of Patent Indicators over the period 1989-99

[Table 2 about here]

Table 2 reports the descriptive statistics of the suggested indicators. Some of them have a large spectrum of variation (max – min) and are strongly skewed, which fits with the previous findings regarding the patents value that most patents turn out to be of very little value, and only a handful have significant importance.

We compared the performance of the inventive activity across the three communities in three different ways.

Firstly, we compared the means of indicators for different groups. We can notice some statistically significant differences. The knowledge base is statistically different for only-authors patents with respect to the other two groups, that do not differ significantly with regards to Originality. In these terms, the only-authors rely less on interdisciplinary previous art.

The higher originality of authors-inventors patents is followed by a larger technological importance and patent scope, on which the other two groups do not differ significantly. only-inventors patents have received a larger protection, which could signal a higher economic impact of them. Authors-inventors produce patents that are significantly more original, and have wider scope and extected value than only-authors. With respect to only-inventors, patents produced by author-inventors are also more cited, in addition to larger scope. These results fully confirm Proposition 1(a) and 1(b).

On the other hand, patents produced by only-authors are significantly less original and have less expected value, than patents produced by only-inventors. We can interpret this finding in two ways. First, since patents produced by only-inventors are usually assigned to companies. (see Figure 10) it may be argued that they receive stronger protection, due to potential, strategic behaviour, and hence have smaller family size. On the other hand, it may be possible that only-authors are relatively more represented by academic inventors, that focus only on a limited technological areas. If this is true, it appears that accessing complementary knowledge is easier when it is formed by codified scientific knowledge than when it refers to idiosyncratic design knowledge.

Table 3 Comparison of the means: T-test statistics

[Table 3 about here]

Given the strong skewness of the quality of innovation the comparison between the means has well-known limits. Hence, it could be more interesting to look to the distribution of patents across different communities at the top percentiles of the distribution of the suggested patent indicators. Table 4 depicts forward citations and shows that author-inventors on average have more patents in those percentiles than the other two communities.

Table 4 Patents by different community at the top percentiles of the distribution of the patent indicator

[Table 4 about here]

Finally, we built a quality index with the mean of different patent indicators as suggested in Appendix 2. We estimated the specified factor model by maximum likelihood given the normality assumption over the unobserved factor. Table 5 presents the factors loadings for the specified model. Both the restriction of no common factor and on more factors are rejected.

It is worth noticing that the sign of the coefficients are positive as we expected, excluding Family size for only-inventors. We interpret this as over protection of their patents with respect to the value.

Table 5 Maximum likelihood factor loadings of one latent variable model

[Table 5 about here]

Table 6 reports the distribution of the quality index for different communities given the above factor loading. Again the authors-inventors exhibit the highest performance, while only-authors show a lower quality of the inventive activity with respect to the other two communities.

Table 6 Distribution of the quality index across communities

[Table 6 about here]

Similar results were obtained even conditioning the factor index for geographic origin of the inventors and application year of the patents. Table 7 We have reported the OLS regression of patent characteristic on the quality index: We found that patents produced by authors-inventors have a higher impact on the quality index than the other two groups. Moreover the interaction effects with assignee type are negative, in particular of both only-inventors and only – authors with PROs. This might be considered consistent with the above theoretical discussed: On one side the only-inventor group lacks adequate scientific knowledge, while the only-authors might not have structured and permanent access of potential industrial applications.

Table 7 OLS regression of the inventor type on the multidimensional quality index

[Table 7 about here]

5. Productivity of Publishing and Patenting for different communities

In the previous section, using multiple patent indicators we found that the inventive activity of only-authors has been of lower quality, while the authors-inventors the highest. In this section we explore the relation between patenting and publishing, in particular we want to test if Authors-Inventors have a higher participation rate in the top percentiles of the individuals distribution of patents and publications. The distribution is obtained ranking all authors or inventors by total number of articles or patents produced in the period. The findings show an interesting pattern. First, although authors-inventors do not out perform the only-authors in publishing distribution in top percentiles, the observed difference is very small. The top 1% of most productive scientists that have patented is formed (87 individuals, almost equally by inventors that cooperate only with other scientists (n=46) and by inventors that cooperate also with people without a publication record. This means that even highly productive scientist benefit from strong complementarities with inventors with different backgrounds. The higher quality of patenting is strongly reflected in the top list of inventors by count (see Table 8): here authors-inventors represent 87% of top 1% most productive inventors and 77% of top 5%. These results strongly confirm Proposition (2).

Table 8 Distribution of the inventors across different communities in top percentiles of the publishing distribution

[Table 8 about here]

Table 9 Distribution of the Inventors across different communities in top percentiles of the patenting distribution

[Table 9 about here]

6. Entrepreneurial productivity

The list of founders that we have used is provided by a questionnaire survey done by www.netinvestor.com. They have interviewed around 1000 companies that launched products based on Nano Science and Technology. We extracted a list of 425 founders of those companies. It turned that 67 of them holds at least one patent as defined in our database.

In Table 10 we classified those 67 by community membership according to the suggested taxanomy. As we can notice the 70 % of founders belong to the author-inventor community; the result holds even if we weight for the size of the community.

Table 10 Founders Distribution by Community

[Table 10 about here]

This confirms our Proposition (3). The past entry performance at these firms is an interesting research question, which will be addressed in future works.

7. Conclusions and Suggestions for Further Research

The importance of scientific discoveries in fuelling technological inventions has been widely documented in many fields. This is particularly evident in the NST where we found that the production of more than 2/3 of the nano patents involves an active scientist. To summarize, we have presented detailed and original evidence on an emergent field in which:

- the production of new knowledge grows much faster than the average for science and engineering;
- although the application areas seem quite well defined, within each area there is evidence of a divergent pattern of growth, following a process characterized by turbulent entry dynamics of new keywords;
- scientists have a tremendous impact on patenting activity in a variety of forms, and the whole field is characterized by high levels of institutional complementarity between industry and academia.

These features qualify NST as a new leading science, characterized by high growth, divergent dynamics, and new forms of complementarity. Some elements seem to suggest that in NST these elements are present with higher intensity and speed than in other leading sciences examined in their birth period (namely, biotech), although this proposition might be subject to rigorous testing in future research.

In order to give evidence on these aspects, we developed a battery of new indicators, namely indicators of entry and turnover of keywords, and individual-centered indicators based on the matching between publication and patent data.

Based on this descriptive and interpretative evidence, we developed some testable propositions that relate the institutional setting of NST research to the performance, as measured by a factor model of patent quality. In spite of the evidence of a highly interconnected knowledge production system, the transformation of scientific discoveries in economic welfare is not immediate and direct. A simple taxonomy of inventors revealed evident differences in the technological and economic performance of them, according to some standard indicators. In particular, communities characterised by the highest levels of institutional complementarity (author-inventors) perform better in both patenting and entrepreneurial activity, still maintaining a remarkable performance in publishing as well.

Further research is needed for a larger validation of the results and for combining the suggested framework with geographical and institutional contexts in which the inventors are embedded.

Acknowledgements: This paper has been originated in a PRIME Network of Excellence project on "Technological districts in nanotechnology", where we had intense discussions with Philippe Laredo, Eric Avenel, Vincent Mangematin, Rikard Stankiewicz, Arie Rip, Michel Zitt. There have been other interesting discussions with Mark Granovetter, Ian Cockburn, Francesco Lissoni, Stine Grodal and other colleagues from SiVNAP Project Group at the Department of Sociology of Stanford University, and from CESPRI at the Bocconi University. We gratefully recognize the contribution of Fabio Beltram, Director of the National Enterprise for nano Science and Technology at Scuola Normale Superiore, Pisa. The competent assistance of Donatella Caridi and Francesca Pierotti in data analysis is kindly acknowledged. All errors are ours, hopefully not too far from the nanoscale.

References

Beltram, Fabio. 2005. Interview to the authors on the discovery methods and instrumentation tools in Nano Science and Technology. Availaible under request to the authors.
Bonaccorsi, A. 2005a. Exploring the function space. in Paper presented to the CIPR Workshop: under review.
—. 2005b. Search regimes and the industrial dynamics of science. Paper presented to the PRIME General Assembly. Manchester: Under-review.
Bozeman, B. and Mangematin V. 2004. Editor's introduction: building and deploying scientific and technical human capital. Research Policy 33:565-568.
Breschi, S, and Lissoni. F 2004. Knowledge networks from patent data: Methodological issues and research targets. in Glänzel W, Moed H, and Schmoch U. (Eds), Handbook of Quantitative S&T Research Kluwer, Academic Publishers.
Callon, M., J. P. Courtial, and F. Laville. 1991. Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemsitry. Scientometrics 22:155.
Cohen, W. M., Levin R. C., and Mowery, D. C. 1987. Firm Size and R & D Intensity: A Re-Examination. The Journal of Industrial Economics 35:543-565.
Fraunhofer-ISI. 2002. Search methodology for mapping nano-technology patents. Karlsruhe, Germany: ISI.
Gambardella, A, Giuri, P and Mariani. M 2005. The Value of European Patents Evidence from a Survey of European Inventors. Final Report of teh PATVAL EU Project.
Gittelman, M, and Kogut. B. 2003. Does Good Science Lead to Valuable Knowledge? Biotechnology Firms and the Evolutionary Logic of Citation Patterns. Management Science 49:366-382.
Grupp, H. (Ed.). 1992. Dynamics of science based innovation: Springer.
Hall, B. H., Jaffe A., and Trajtenberg M. 2005. Market value and patent citations. RAND Journal of Economics 36:16-38.
Hall, B., and Trajtenberg M. 2004. Uncovering GPTs with Patent Data. in NBER Working Paper No. 10901. Cambridge, Mass.
Harhoff, D., Narin F., Scherer F. M., and Vopel K. 1999. Citation Frequency and the Value of Patented Inventions. Review of Economics and Statistics 81:511-515.
Harhoff, D., Scherer F. M., and Vopel K. 2003. Citations, family size, opposition and the value of patent rights. Research Policy 32:1343.
Heinze, S, and Schmoch. 2004 U. Opening the black box. in Glänzel W, Moed H, and Schmoch U. (Eds), Handbook of Quantitative S&T Research Kluwer, Academic Publishers.
Henderson, R., Jaffe A. B., and Trajtenberg M. 1998. Universities as a Source of Commercial Technology: A Detailed Analysis of University Patenting, 1965-1988. Review of Economics and Statistics 80:119-127.
Jaffe, A. B., and Manuel Trajtenberg. 2002. Patents, Citations, and Innovations: A window on the Knowledge Economy. Cambridge, Massachusetts, The MIT Press.
Jaffe, A. B., Trajtenberg M., and Henderson R. 1993. Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. The Quarterly Journal of Economics 108:577-598.
Jolliffe, I T. 2002. Principal Component Analysis, Second Edition. New York: Springer-Verlag.
Kenney, M, and Goe R. W.. 2004. The role of social embeddedness in professorial entrepreneurship: a comparison of electrical engineering and computer science at UC Berkeley and Stanford. Research Policy 33:691-707.
Klein, S. 1985. What is technology. Bulletin of Science, Technology and Society 1:215-218.
Klepper, S. 2001. Employee Startups in High-Tech Industries. Industrial and Corporate Change 10:639-674.

Kline, S. J., and Rosenberg N. 1986. An Overview of innovation. in Landau R. and Rosenberg N. (Eds) The positive sum strategy, Washington D.C: National Academy Press.
Lanjouw, J. O., Pakes A., and Putnam J. 1998. How to Count Patents and Value Intellectual Property:The Uses of Patent Renewal and Application Data. Journal of Industrial Economics 46:405-432.
Lanjouw, J. O., and Schankerman M. 2001. Characteristics of patent litigation: a window on competition. RAND Journal of Economics:Vol. 32 Issue 1.
—. 2004. Patent Quality and Research Productivity: Measuring Innovation with Multiple Indicators. Economic Journal 114:441-465.
McKelvey, M. 1996. Evolutionary Innovations. The business of Biotechnology: Oxford University Press.
Mowery, D., and Rosenberg N. 1998. Paths of Innovation: Technological Change in 20th-Century America. New York: Cambridge University Press.
Narin, F., and Olivastro D.. 1992. Status report: linkage between technology and science. Research Policy 21:237-249.
Narin, F., S. Hamilton K. S., and Olivastro D. 1997. The increasing linkage between U.S. technology and public science. Research Policy 26:317-330.
Nelson, R. R. 1986. Institutions Supporting Technical Advance in Industry. The American Economic Review 76:186-189.
Orsenigo, L. 1990. The emergence of biotechnology: Cambridge University Press.
Owen-Smith, J., Riccaboni M, Pammolli F., and Powell W. W. 2002. A Comparison of U.S. and European University-Industry Relations in the Life Sciences. Pp. 24 in Management Science: INFORMS: Institute for Operations Research.
Pavitt, K. 1990. What We Know about the Strategic Management of Technology. California Management Review 32:17.
Popp, D. 2005. They Don't Invent Them Like They Used To: An Examination of Energy Patent Citations Over Time. in NBER Working Paper No. 11415.
Rosenberg, N. 1982. Inside the Black Box: Cambridge University Press.
—. 1990. Why do firms do basic research (with their own money)? Research Policy 19:165-174.
Shane, S. 2004. Academic Entrepreneurship: University Spinoffs And Wealth Creation: Edward Elgar.
Stankiewicz, R. 2000. On the Concept of "Design Space. in Ziman J. (Eds) Technological Innovation as an Evolutionary Process, Cambridge University Press.
—. 1997. The Development of Beta-blockers at Astra-Hässle and the Technological System of the Swedish Pharmaceutical Industry. in Carlsson B (Eds) Technological Systems and Economic Preformance, Kluwer Publishers.
Tijssen, R. J. W. 2004. Measuring and evaluating Science and Technology Connections and Interactions. in Glänzel W, Moed H, and Schmoch U. (Eds), Handbook of Quantitative S&T Research Kluwer, Academic Publishers.
Trajtenberg, M. 1989. The Welfare Analysis of Product Innovations, with an Application to Computed Tomography Scanners. The Journal of Political Economy 97:444-479.
—. 1990. A Penny for Your Quotes: Patent Citations and the Value of Innovations. The Rand Journal of Economics 21:172-187.
Verspagen, B. 2005. Mapping Technological Trajectories as Patent Citation Networks. A Study on the History of Fuel Cell Research. in Merit Research Paper. Maastricht: Merit.
Vincenti, W. 1990. What Engineers Know and How They Know It. Baltimore: Johns Hopkins Press.
Zucker, L., and Darby. M 1996. Star Scientists and Institutional Transformation: Patterns of Invention and Innovation in the Formation of the Biotechnology Industry. in Proceedings of the National Academy of Sciences. Pp. 709-12, 716

Zucker, Lynne G., and Michael R. Darby. 2003. Grilichesian Breakthroughs: Inventions of Methods of Inventing and Firm Entry in Nanotechnology. in NBER working paper n. 9825.

Appendix: A multidimensional measure of patent quality

The construction of the multidimensional measure of patent quality relies on factor analysis. In factor models each series of data (quality indicator in our case) is decomposed into a common component and an idiosyncratic component. The common component is driven by only a few common shocks, denoted by $Q < N$, where N is the number of indicators. In a static factor model, the common shocks affect the indicators only contemporaneously. The basic model is given by $X = UB + E = K + E$, where X is the $(T \times N)$ matrix of observations on N series (indicators) of length T. The series are normalized to have mean 0 and variance 1. U is the $(T \times Q)$ matrix of Q common shocks and B is the $(Q \times N)$ matrix of factor loadings, which determines the impact of common shock q on series n. The common shocks and the factor loadings together make up the common component K. After the influence of common shocks has been removed, only the idiosyncratic component (E) remains. To estimate the common component we have to find a linear combination of the indicators in X that explains as much as possible the total variance of each indicator, minimizing the idiosyncratic component (for a technical discussion of factor models see (Jolliffe 2002)).

The parallel with least squares estimation is clear from this formulation, but the fact that the common shocks are unobserved complicates the problem. The standard way to extract the common component in the static case is to use principal component analysis. In principal component analysis the first Q eigenvalues and eigenvectors are calculated from the variance-covariance matrix of the dataset X. The common component is then defined as: $K = XVV'$, with $V = [p_1, \ldots, p_Q]$ and where p_i is the eigenvector corresponding to the ith largest ($i = 1 \ldots Q$) eigenvalue of the covariance matrix of X. This method does not ensure a unique solution. A further problem is that *ex ante* it is not known how many common shocks Q affect the series in X. Following the approach suggested by Lanjouw and Schankerman (2004), we use a multiple-indicator model with an unobserved common factor:

$$y_{k\,i} = \lambda_k q_i + \beta' X + e_{ki}$$

where y_{ki} indicates the value of the *k*th patent indicator for the *i*th patent; q is the common factor with factor loadings λ_k and normally distributed, and X a set of controls. The main underlining assumption is that the variability of each patent indicator in the sample may be generated by the variability of a common factor across all the indicators and an idiosyncratic part e_k not related to the other indicators and distributed $N(0,\sigma^2_k)$.

In our setting, the common factor is the unobserved characteristic of a patent that influences positively four quality indicators: backward citations, forward citations, number of claims, and family size. Estimation of common quality index q is based on information extrapolated from the covariance matrix of our four indicators. By assuming the normality of q_i and e_k we can estimate by maximum likelihood, which ensures an unique solution. Once the estimates of λ_k are obtained, the model is inverted to calculate q.

Table 1 US Patents in IPC B8 Micro-Structural Technology and Nano-Technology

INTERNATIONAL CLASS	USPTO	EPO
MICRO-STRUCTURAL DEVICES OR SYSTEMS, e.g. MICRO-MECHANICAL DEVICES (B81B)	37	20
PROCESSES OR APPARATUS SPECIALLY ADAPTED FOR THE MANUFACTURE OR TREATMENT OF MICRO-STRUCTURAL DEVICES OR SYSTEMS (B81C)	27	12
NANO-STRUCTURES; MANUFACTURE OR TREATMENT THEREOF (B82B)	14	1

Source: Our elaboration from USPTO and EPO, May 2005

Table 2 Descriptive Statistics of Patent Indicators over the period 1989-99

	sample	min	max	median	mean	std	skewness	kurtosis
Authors-Inventors								
Originality	1056	0.00	0.93	0.59	0.51	0.28	-0.77	2.39
FWD5	1056	0.00	73.00	3.00	5.34	7.43	3.27	20.10
Family	1056	1.00	422.00	7.00	12.76	24.68	8.71	108.59
Claims	1056	1.00	199.00	17.00	20.20	17.21	3.86	31.61
Only-Authors								
Originality	919	0.00	0.90	0.53	0.48	0.29	-0.61	2.08
FWD5	919	0.00	54.00	3.00	5.17	6.14	2.34	11.56
Family	919	1.00	468.00	4.00	7.64	18.02	19.01	469.15
Claims	919	1.00	155.00	14.00	17.97	14.70	2.76	16.90
Only-Inventors								
Originality	957	0.00	0.91	0.56	0.50	0.26	-0.77	2.55
FWD5	957	0.00	114.00	2.00	4.92	7.80	5.17	51.39
Family	957	1.00	884.00	8.00	14.01	33.82	18.41	460.48
Claims	957	1.00	236.00	14.00	18.06	16.06	4.39	43.30
Overall								
Originality	2932	0.00	0.93	0.56	0.50	0.28	-0.72	2.33
FWD5	2932	0.00	114.00	3.00	5.15	7.18	3.92	33.39
Family	2932	1.00	884.00	7.00	11.56	26.48	17.04	464.09
Claims	2932	1.00	236.00	15.00	18.80	16.11	3.79	32.64

Source: Our Elaboration

Table 3 Comparison of the means

	Authors-Inventors vs Only-Authors		Authors-Inventors vs Only-Inventors		Only-Authors vs Only-Inventors	
Originality	2.78	***	1.18		-1.69	***
FWD5	0.58		1.26	*	0.78	
Family	5.31	***	-0.94		-5.12	***
Claims	3.10	***	2.88	***	-0.13	

Source: Our Elaboration;

Notes: * statistical significant at 10% level; ** statistical significant at 5% level; *** statistical significant at 1% level;

Table 4 Patents distribution by different community at the top percentiles of the distribution of the patent indicator

Percentiles	75p	90p	95p	99p
Authors-Inventors	43.88 %	37.41 %	44.90 %	42.86 %
Only-Authors	30.61 %	31.97 %	12.59 %	28.57 %
Only-Inventors	25.51 %	30.61 %	42.52 %	28.57 %

Source: Our Elaboration

Table 5 Maximum Likelihood Factor Loadings of one Latent Variable Model

	Only-Authors	Only-Inventors	Authors-Inventors	Overall
Originality	0.39	0.21	0.33	0.29
FWD5	0.18	0.43	0.22	0.25
Family	0.20	-0.08	0.34	0.15
Claims	0.24	0.26	0.31	0.31

Source: Our Elaboration

Table 6 Distribution of the quality index across communities

percentiles	Only-Authors	Only-Inventors	Authors-Inventors	Overall
P25	-0.39	-0.32	-0.28	-0.32
P50	-0.02	-0.01	0.05	0.01
P75	0.30	0.31	0.40	0.34
P90	0.52	0.56	0.69	0.60
P99	0.97	1.03	1.16	1.06

Source: Our Elaboration

Table 7 OLS regression of the inventor type on the multidimensional quality index

	MODEL 1			MODEL 2			MODEL 3		
	Coef.	Std.Err.	Sign	Coef.	Std.Err.	Sign	Coef.	Std.Err.	Sign
Only Authors	-0.12	0.02	***	-0.06	0.02	***	-0.08	0.03	***
Only Inventors	-0.07	0.02	***	-0.07	0.02	***	-0.06	0.02	***
PROs * Only Authors							-0.04	0.03	
PROs * Only Inventors							-0.07	0.04	*
Control Variables									
Number of inventors				-0.01	0.05		-0.01	0.05	
Dummy 1989				0.05	0.07		0.05	0.07	
Dummy 1990				0.14	0.07	*	0.14	0.07	*
Dummy 1991				0.16	0.07	***	0.16	0.07	***
Dummy 1992				0.21	0.07	***	0.21	0.07	***
Dummy 1993				0.27	0.06	***	0.27	0.06	***
Dummy 1994				0.39	0.06	***	0.39	0.06	***
Dummy 1995				0.39	0.06	***	0.39	0.06	***
Dummy 1996				0.46	0.06	***	0.46	0.06	***
Dummy 1997				0.48	0.06	***	0.49	0.06	***
Dummy 1998				0.54	0.06	***	0.54	0.06	***
Dummy 1999				0.58	0.06	***	0.58	0.06	***
Austria				-0.44	0.5		-0.45	0.50	
Belgium				-0.41	0.46		-0.42	0.46	
Canada				-0.31	0.44		-0.3	0.44	
Checz Republic				-0.36	0.5		-0.36	0.44	
Croatia				-0.12	0.53		-0.4	0.43	
Denmark				-0.09	0.48		-0.1	0.48	
Finland				-0.34	0.5		-0.35	0.47	
France				-0.37	0.44		-0.34	0.50	
Germany				-0.39	0.43		-0.37	0.44	
Ireland				0.1	0.5		-0.29	0.44	
Italy				-0.35	0.45		-0.35	0.50	
Japan				-0.5	0.43		-0.13	0.53	
Netherlands				-0.31	0.45		0.091	0.50	
Norway				-0.48	0.53		-0.35	0.45	
Romenia				-0.39	0.61		-0.51	0.43	
Russia				-0.05	0.48		-0.31	0.45	
Spain				-0.35	0.47		-0.49	0.53	
Sweeden				-0.2	0.45		-0.39	0.61	
Switzerland				-0.36	0.44		-0.05	0.48	
Ukraine				-1.2	0.61	*	-0.21	0.45	*
UK				-0.29	0.44		-1.21	0.61	
US				-0.26	0.43		-0.25	0.43	
Costant				-0.04	0.44		-0.04	0.44	
Number of observations	2802			2802			2802		
Adjusted R-squared	2%			16%			17%		

Notes:
(1) *** 1% level significance; ** 5% level significance; * 10% level significance
(2) The likelihood –ratio test reject the null hypothesis across all models

Table 8 Distribution of the Inventors across different communities in top percentiles of the publishing distribution

Units				
top percentiles	Only Authors	Only Inventors	Author Inventors	Overall
1p	46	Nd	41	87
5p	209	Nd	226	435
10p	437	Nd	433	870
25p	1129	Nd	1146	2275
Shares				
top percentiles	Only Authors	Only Inventors	Author Inventors	Individuals
1p	53%	0,00	47%	100%
5p	48%	0,00	52%	100%
10p	50%	0,00	50%	100%
25p	50%	0,00	50%	100%

Sources: Our Elaboration

Table 9 Distribution of the Inventors across different communities in top percentiles of the patenting distribution

Units				
top percentiles	Only Authors	Only Inventors	Author Inventors	Overall
1p	5	6	76	87
5p	37	65	333	435
10p	102	151	617	870
25p	412	476	1387	2275
Shares				
top percentiles	Only Authors	Only Inventors	Author Inventors	Individuals
1p	6%	7%	87%	100%
5p	9%	15%	77%	100%
10p	12%	17%	71%	100%
25p	18%	21%	61%	100%

Sources: Our Elaboration

Table 10 Founders Distribution by Community

Founders	Only-Inventors	Only-Authors	Author-Inventors	Overall
Units	4	16	47	67
Share %	6%	24%	70%	100%
Normalization ratio	2,10	1,93	1,00	
Normalized Units	8	31	47	86
Normalized Share %	10%	36%	55%	100%

Notes: The normalization procedure takes into the account the fact that the Authors-Inventor community is larger than the other two communities. The normalization ratio adjusts for the size of the community.

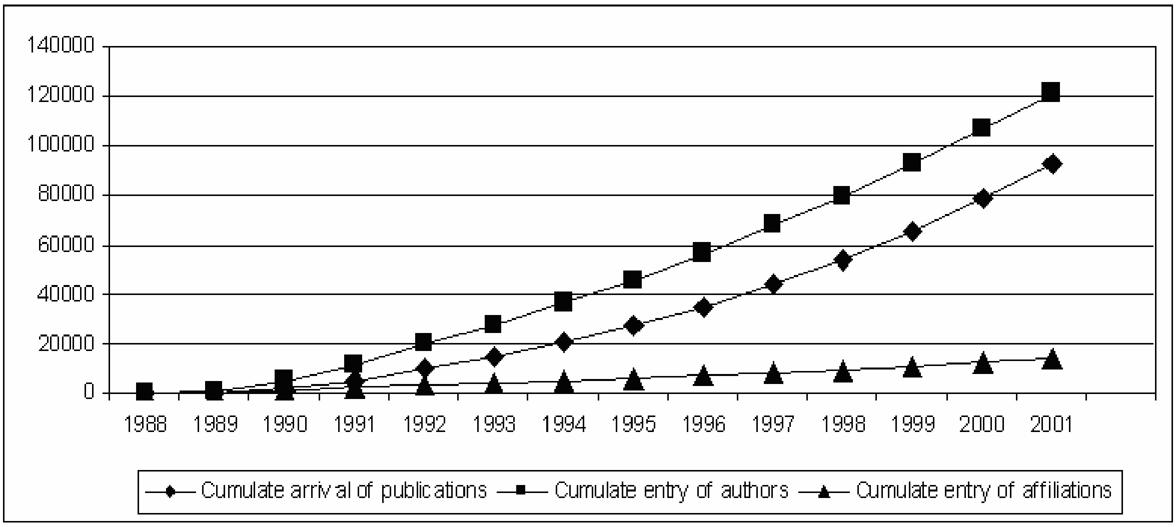


Figure 1 Cumulate arrivals in Nano-science (Source: our elaboration from ISI)

Source: Our elaboration

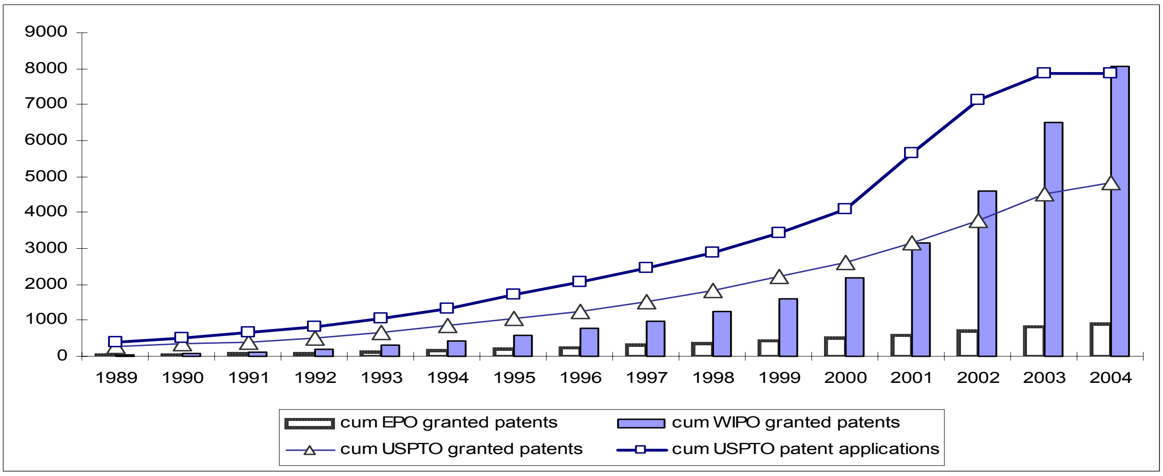


Figure 2 Cumulate arrivals in Nanotechnology, May 2004

Source: Our elaboration
Notes: The patents for the 2004 include only those that have been granted before May 2004.

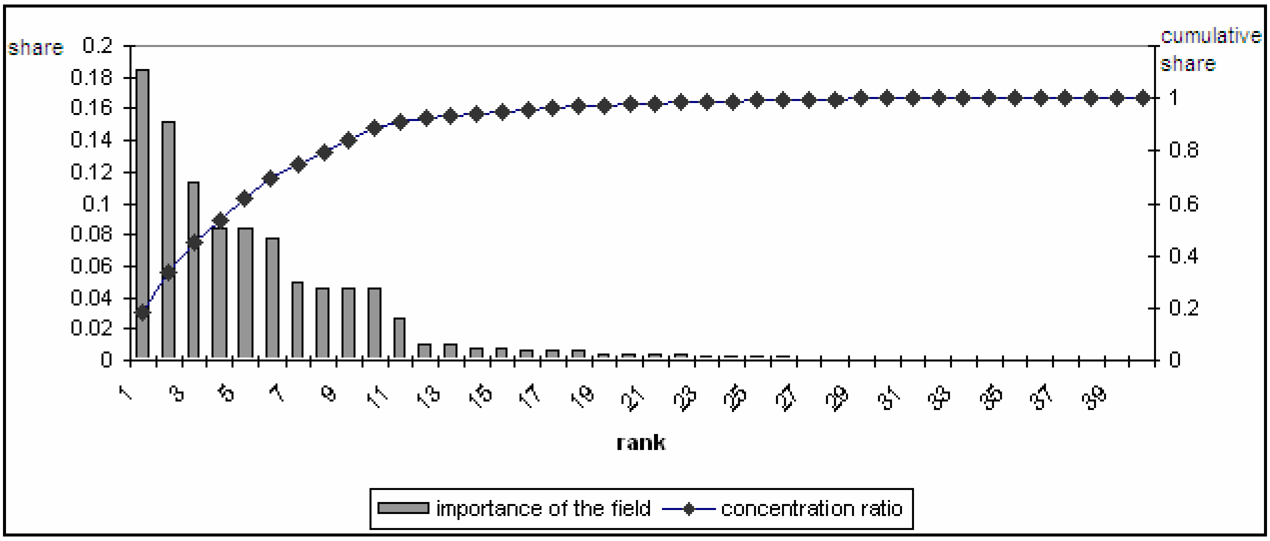


Figure 3 Nano-publications by field

Source: Our elaboration

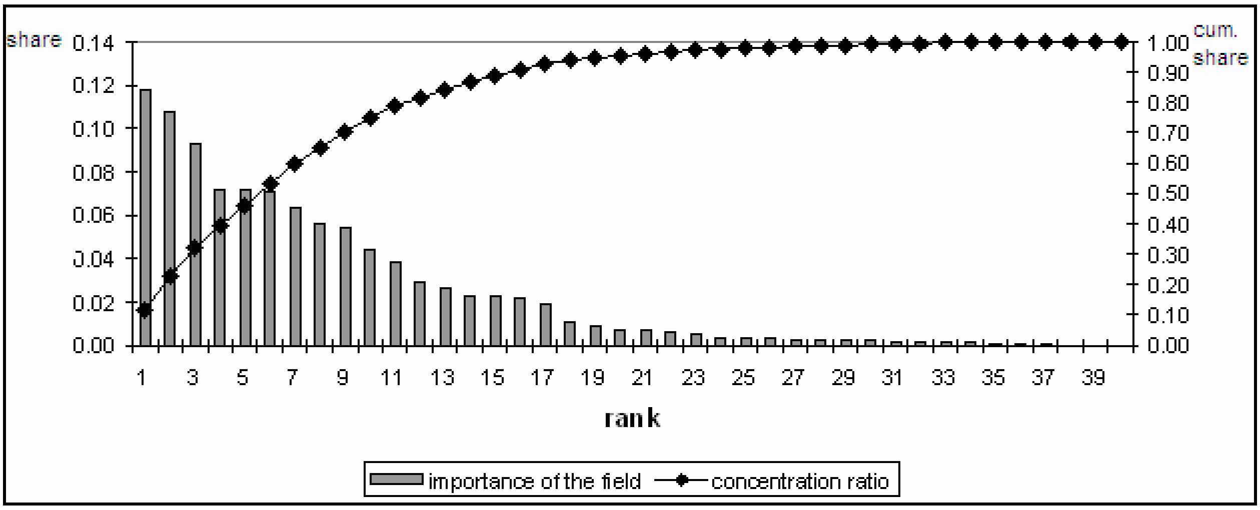


Figure 4 Nano-patents by field

Source: Our elaboration

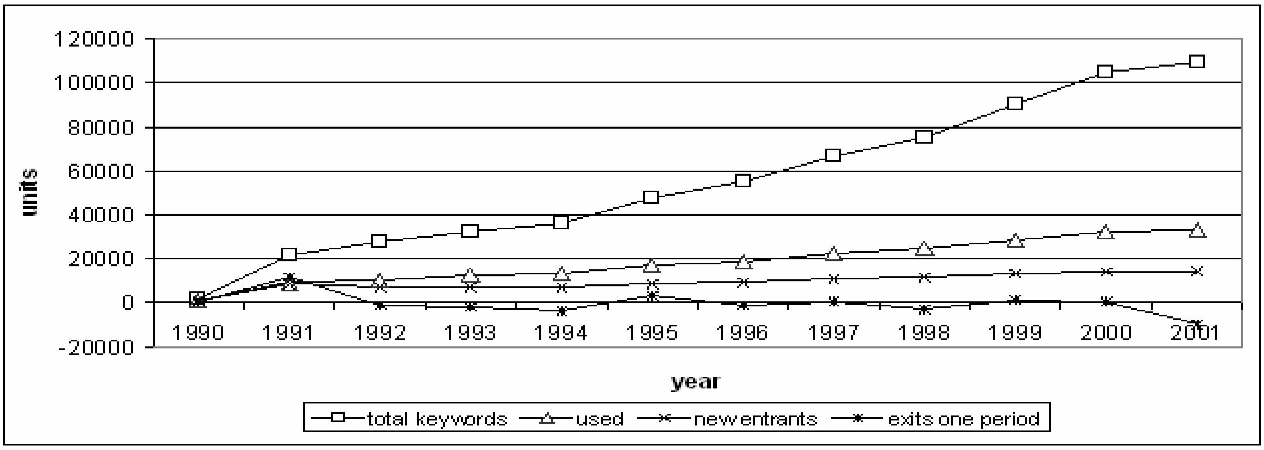


Figure 5 Stocks, use, entry, and exit of keywords at the overall population by year

Source: Our elaboration

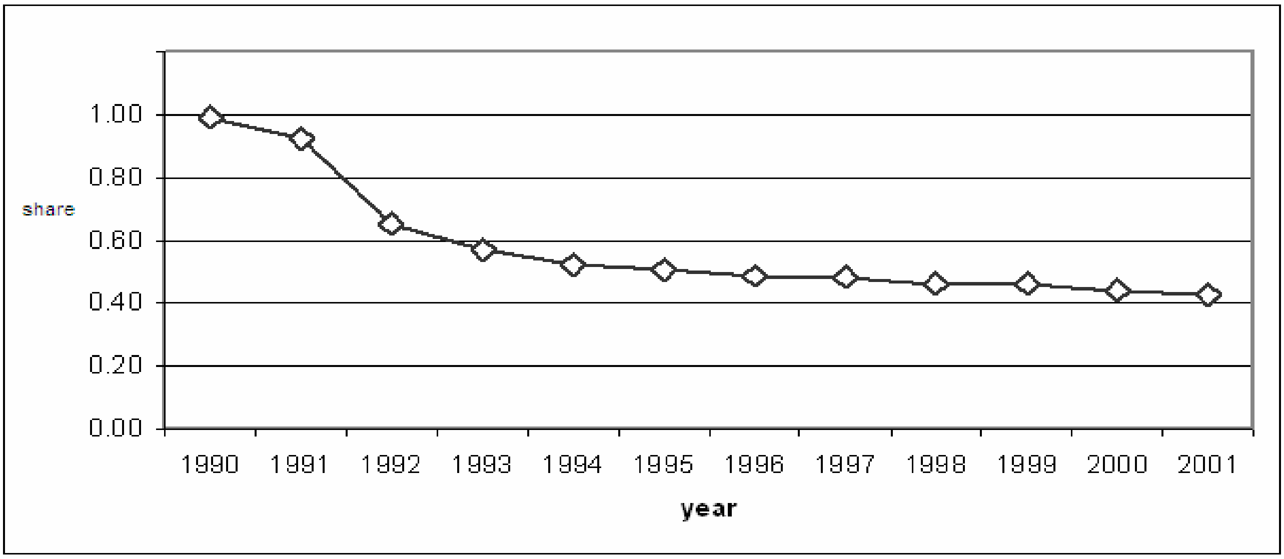


Figure 6 Ratio between number of new keywords entered and total number of keywords used by year

Source: Our elaboration

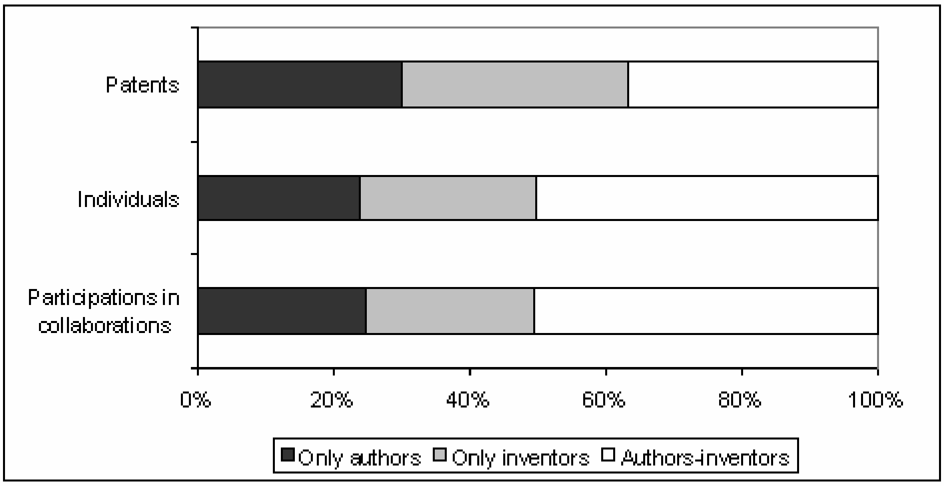


Figure 7 Composition of communities of inventors in nanotechnology

Source: Our elaboration
Notes: "Participations in collaborations" takes into account the number of times an individual has been a co-inventor of a patent; Source: our elaboration.

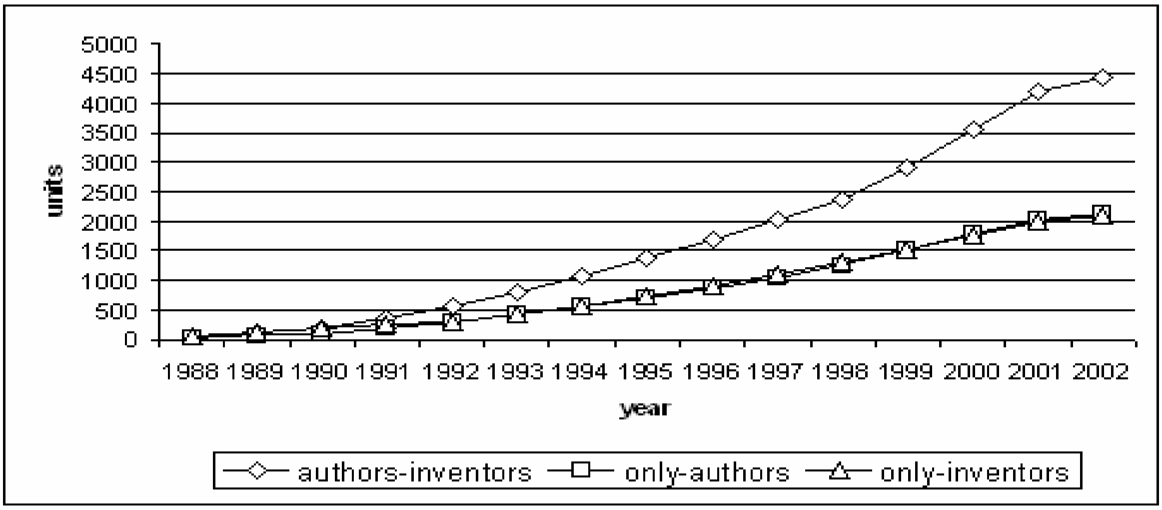
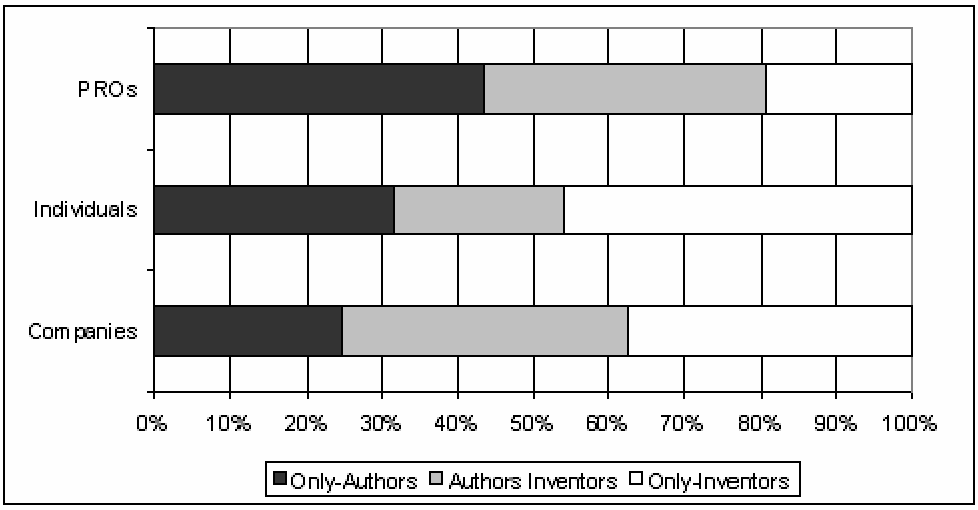


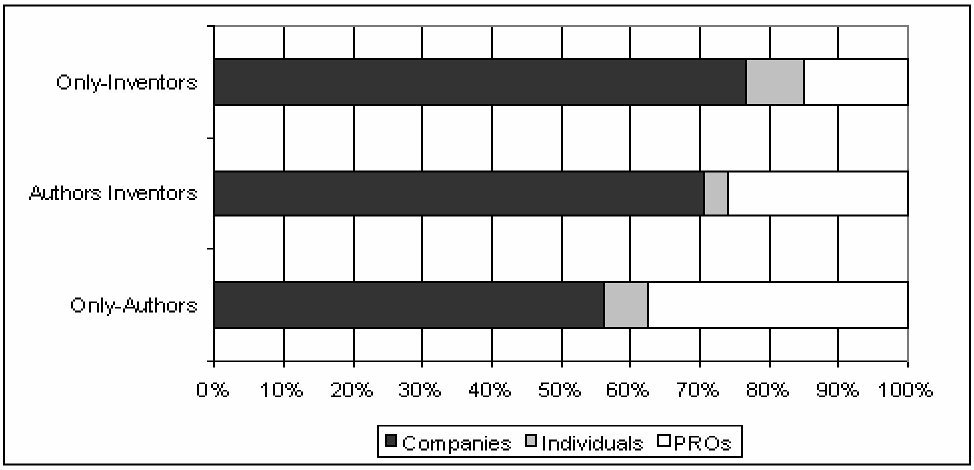
Figure 8 Entry of individuals by community

Source: Our elaboration



Source: Our elaboration

Figure 9 Distribution of patents by community in relation to assignee type



Source: Our elaboration

Figure 10 Distribution of patents by assignee type in relation to community