Inventor's Knowledge Set as the Antecedent of Patent Importance

ALI MOHAMMADI & CHIARA FRANZONI

Department of Management, Economics and Industrial Engineering, Politecnico di Milano, P. Leonardo da Vinci 32, Milan 20133, Italy

ABSTRACT   This paper investigates the antecedents of patent importance by looking at the prior knowledge set of academic inventors. Using independent methodologies, we distinguish between the scientific knowledge set and the technical knowledge set, and separate these from other kinds of prior expertise. We find that the patents of the inventors who have a prior scholarly knowledge of the topic are on average more important (more cited after 6 years). Conversely, we find an inverted U-shape relationship between prior technical relatedness and patent importance. These results are potentially useful to support the work of practitioners such as university technology managers, which often face difficulties in identifying the importance and perspective value of the disclosed inventions, amid high market and legal uncertainty and budget shortages.

1. Introduction

Scholarly works on Intellectual Property Rights (IPRs) have for long highlighted that market transactions concerning patents and other technological assets are constrained by problems of asymmetric information and moral hazard (Shane, 2002; Scotchmer, 2004). One critical problem is anticipating the prospective relevance of a new technology, which depends on a combination of technological opportunity, market demand and legal perils (Gittelman and Kogut, 2003). Concerning technological and market opportunity, the effective capacity of a technology to translate into functional products and generate

Correspondence Address: Ali Mohammadi, Department of Management, Economics and Industrial Engineering, Politecnico di Milano, P. Leonardo da Vinci 32, Milan 20133, Italy. Tel: +39 02 2399 4041. Email: ali.mohammadi@polimi.it

revenues depends on events that may or may not take place in the future, such as the possibility that a substitute or competing technology will prevail on the market. In addition to market uncertainty, transactions are complicated by problems of asymmetric information among the parties that are involved in the negotiation (Gallini and Wright, 1990). For example, the potential acquirer tends to have less information on the quality of a technology than the vendor, particularly if the technology embeds frontier knowledge for which it is even difficult to find an expert of the field to assist the transaction (Franzoni, 2007). Both issues—uncertain outcomes and asymmetric information—are even more severe for technologies that are developed and patented in universities (Shane, 2002), whose relevance is, therefore, even more difficult to anticipate.

In this paper, we focus on the antecedents of patent importance in academic inventions, and investigate whether and to what extent highly relevant patents are correlated to certain characteristics of the prior knowledge set of the inventor. In general, the literature on academic patents has documented that the most productive scientists are more likely to become inventors than their less-productive colleagues (Breschi *et al.*, 2007; Stephan *et al.*, 2007; Fabrizio and Di Minin, 2008). It has also documented that inventions are preceded by a flurry of scholarly publications (Azoulay *et al.*, 2007; Calderini *et al.*, 2007). This evidence seems to indicate that inventions are developed as a byproduct of a prolific research activity that the scientist has conducted for scientific purposes. In practice, to the best of the authors' knowledge, no prior study has established a clear link between the flurry of publications that precede an invention and the invention itself. In other words, we presume that there is an ideal continuity between *what* a scientist does in a certain year in his/her academic lab and *what* does he/she invent in the same year or shortly after. However, this presumption is based solely on time-contiguity and not on content relatedness. In order to observe content relatedness, standard techniques that rely exclusively on counts of occurrences (number of patents and number of publications in a given time span) are not sufficient. In this paper, we take a step further in this direction by using techniques of content analytics. We do so by starting from a sample of US patents that have been invented by faculty members in the field of physics. We measure patent importance by looking at the number of citations received 6 years after the year of patent priority. We characterize the inventors in terms of the related knowledge sets that they possess. We retrieve information on the inventor's scientific knowledge set by analyzing the content of the articles that he/she published in scientific outlets prior to the priority year of the focal patent. We retrieve information on the inventor's technical knowledge set by analyzing the content of the inventions that he/she filed prior to the priority year of the focal patent. This allows us to characterize the content of an invention in terms of its relatedness to two distinct knowledge sets of the academic inventor: (i) the scientific understanding of the subject, which originates from his/her scholarly research and
(ii) the technical understanding of technologies that originates from his/her prior inventions.

Our approach is novel in that it applies techniques of content analytics in order to investigate scientific and technical relatedness. We rely on a technique that allows computing the similarity of two or more documents (patents abstracts vs. prior publications abstracts, patents abstracts vs. prior patent abstracts), based on the co-occurrence of the same semantics (co-word) and on the same pairs (co word-pairs), if pairs exist. We use this technique to partition the inventor's publications and patents into items whose content is related to that of a focal patent and items whose content is unrelated. This methodological choice is novel and can potentially be replicated with ready-to-use and inexpensive commercial software.

Based on the current scholarly understanding, we formulate the hypotheses that highly cited patents would be those grounded on the scientific knowledge set and that further exploitation of an existing technical knowledge set would be associated with less-important patents.

We test our hypotheses on a sample of 295 academic patents[1] in all subfields of physics. Our results suggest that patents whose content is related to prior scholarly papers are disproportionately distributed among the more cited. On average, during the first 6 years of life, a patent invented in a prior area of scholarly investigation receives 50 per cent more citations than a patent invented in areas unrelated to the research scope of a scientist. Prior technical knowledge, as captured by prior patents whose content is related to that of a focal patent, has an inverted U-shape relationship with patent citations.

Although citations give but a blunt appreciation of patent quality, these results clearly point at a correlation between the inventor's knowledge set and the degree to which inventions are likely to have impact and be valuable. We discuss how these results may bring potentially relevant implications for the practice of technology management at universities, as well as for our understanding of the process that leads from scholarly investigation to the development of practical applications.

The paper is structured as follows. Section 2 reviews the findings of the literature on academic patenting with regard to patent importance and patent citations and defines two theoretical hypotheses for the empirical investigation. Section 3 describes the sample and the data collection procedure and defines the variables and the methodology used in the analysis. Section 4 presents the results of the econometric estimates and Section 5 discusses the implications.

2. Academic Patents

2.1 Assessing the Market Value of Disclosed Inventions at Universities

In a context of technology and market uncertainty, evaluating the prospective importance of a university patent is an especially difficult task. The scholars who have investigated technology transfer have extensively documented a number of common and widespread problems that make it difficult to assess the importance of academic inventions. First, a large share of the inventions disclosed by academic professors is still at the proof of concept stage, or anyway has a long way before they could become eventually marketable (Jensen and Thursby, 2001). Second, academic professors often lack business competences, have little understanding of the market conditions, few links to companies potentially interested in their inventions (Swamidass and Vulasa, 2008) and often overestimate the real market importance of their ideas (Owen-Smith and Powell, 2001).

Third, faculty members naturally are moved by sets of priorities and goals typical of academia, which are not always or not necessarily aligned to those of the potential investors (Jensen *et al.*, 2003). For example, some scientists are interested in the personal reputation that they think they can gain from patenting (Go¨ktepe-Hulten and Mahagaonkar, 2010). Other scientists may be more concerned that patenting may require delaying their scientific publications (Blumenthal *et al.*, 1997;Campbell*et al.*, 2002; Franzoni and Scellato, 2010) and

---

[1] Academic patents are defined here as patents invented by a university faculty member.

bring about other conflict of interests (Walsh *et al.*, 2005; Murray and Stern, 2007; Ambos *et al.*, 2008; Haeussler *et al.*, 2009).

These circumstances partly explain the difficulties encountered by universities in getting sizable revenues from the commercialization of technologies. The available figures point at a situation in which few of the many universities active in technology commercialization generate a substantial stream of income. Those that have income generate their revenues from a handful of highly successful patents (AUTM, 2011). For example, the University of California—known to be top for profit generation—makes about half of its revenue with just five patents, all in the medical or biotech sector (Farrell, 2008). In more common situations, technology managers strive to find interested potential buyers. A large survey conducted in eight European countries highlighted that the share of unlicensed patents at academic institutions is almost double than that of unlicensed patents at large and medium firms (EC, 2006). Amid budget shortages, many offices report having just enough resources to accomplish the legal and procedural aspects of patent filing, leaving too few for marketing and commercialization (Swamidass and Vulasa, 2008).

2.2 Importance of Academic Patents

Some early works on patent citations have been conducted between 1998 and 2003. The focus of these analyses was to determine if the institutionalization of the patenting activities from universities in the aftermath of the Bayh−Dole act led to a decline of patent quality (Sampat *et al.*, 2003). The purpose of this body of works is different from the scope of our analysis. Nonetheless, a summary of the basic findings provides a helpful background for our research. First, university patents (defined as those assigned to universities or having at least one scientist among the inventors) on average tend to be more highly cited and broader in scope than non-university patents (Henderson *et al.*, 1998; Mowery *et al.*, 2001; Czarnitzki *et al.*, 2011). Second, in terms of trends, there is a clear increase in the number of patents issued to universities, but it appears less clear if the above-average citation rate has remained unchanged. The most recent findings seem to suggest that there has been no sensible decline in citations received, although citations come at a slower pace, due to slowdowns in the patent application procedure (Sampat *et al.*, 2003).

Concerning non-US universities, the issue is more controversial, partly because many patents of university faculty members are assigned to companies and are more difficult to keep track of (Lissoni *et al.*, 2010). Recent works show similar citation patterns experienced by university and non-university patents in Europe (Crespi *et al.*, 2011) and Japan (Bacchiocchi and Montobbio, 2010). In terms of trends, Czarnitzki *et al.* (2011) find a decline in citations since the mid-1980 in a sample of German patents.

A parallel debate centered on whether or not pushing for commercial research in academe could backfire on the willingness or capability of scientists to continue to advance fundamental science, which traditionally cannot be commercialized in the short to medium term (Nelson, 2004). Several empirical tests have consistently shown that, in general, the association between inventive and scientific activity tends to be positive (Azoulay *et al.*, 2009), although with two caveats. First, a trade-off can exist at the tail of the distribution, i.e. for scientists who are highly productive in very fundamental areas of science (Calderini *et al.*, 2009). Second, scientists may be discouraged from pursuing those lines of research in

which the important research tools are agressively commercialized and protected by patents (Murray and Stern, 2007).

2.3 Scientific Knowledge Set as an Input to the Invention Process

A body of empirical investigations has convincingly established a link between scientific and inventive productivity, both at the level of firms and the level of single individuals. At the level of individuals, cross-sectional investigations have shown that the most productive scientists are more likely to become inventors than their less productive colleagues (Breschi *et al.*, 2007; Stephan *et al.*, 2007; Fabrizio and Di Minin, 2008). The correlation has also been documented in longitudinal studies, showing that inventions are preceded by a flurry of scholarly publications (Azoulay *et al.*, 2007; Calderini *et al.*, 2007). When looking at licenses of academic patents, Elfenbein (2007) showed the signalling nature of publications that correlates to increasing probability of licencing a new technology.

When looking at the level of firms, rather than at the level of the individuals, the positive relationship is confirmed. Gittelman and Kogut (2003) studied the effect of firms' scientific publishing on patent importance and provided evidence that patents based on scientific research are more likely to be cited, although highly cited publications are negatively correlated with good patents. Fleming and Sorensen (2004) find that patents citing publications are more likely to be broader in scope. Zucker *et al.* (2002) studied the role of collaboration with star scientists and its effect on patent importance. Their empirical results show that joint publications of firms' employees and star scientists are correlated with a high citation rate.

This evidence is consistent to a scenario in which scientific research is a valuable input for the production of innovation. Fleming and Sorenson (2004) interpret these findings in light of the theory of the invention as a process of knowledge recombination (Nelson and Winter, 1982; Kogut and Zander, 1992; Fleming, 2001), which argues that all inventions at a closer look are the result of a combination of pre-existing discrete elements and components. They maintain that the search for new components is conducted locally, in the sense that it usually occurs in areas that are proximal in cognitive and semantic terms to the prior experience of the inventor. Their view is that experience of scientific research helps inventors in orientating their local search toward more fruitful and less-exhausted combinations.

It is important to note that all these studies investigate numerical relationships between counts (or dummies) of publications, and counts of citations, but are incapable to account for the degree to which two documents (e.g. a patent and a paper) have similar or dissimilar content. As such, quantitative relationships have limitations. For example, if we observe that those inventors who publish more papers are more likely to patent and/or receive citations later on, we cannot tell if this happens because scientific research was an input to the inventive process, or simply because those inventors were going through an exceptionally fruitful period of their life, and/or had more resources, hence more output of any kind. In statistical terms, we cannot tell if scientific understanding is the antecedent of valuable patents or if both the scientific understanding and the valuable patents are jointly determined by some common external factor. If the scientific research conducted by an inventor and disclosed in articles is a valuable input, for example because it directs the search of new technical solutions toward unexplored pathways, then we should observe not only the co-occurrence of events

(publications and patent citations), but we should also see that the invention refers to the specific knowledge set that the inventor has developed in scientific articles prior to inventing. In statistical terms, our correlation would be driven by only those subset of co-occurrences that happen in the same semantic domain. This would not entirely rule-off the possibility that the variables are jointly determined by a separate set of unobservable variables. However, it clears off potential biases due to co-occurrences that are not happening in the same semantic domain. In order to do so, we code separately for documents of similar content. Our measure of relatedness is based on observing a high degree of semantic overlap between the patent and the published research by the same inventor/author. We distinguish this variable from the numerical count of prior publications that we use as a control, thus clearing off the impact of scientific research used as an input, from the effect of mere variance in the productivity of the inventor in a given period of time.

Our first research hypothesis is formulated as follows:

H1: Inventions that rely on a prior scientific knowledge set (are built on the results of the prior scientific research of the academic inventor) are more likely to be important than inventions that do not rely on a prior scientific knowledge set.

2.4 Technical Knowledge Set as an Input to the Invention Process

Many academic scholars produce inventions quite continuously during their career (Azoulay *et al.*, 2007). In line with the theory of invention as the product of knowledge recombination (Fleming, 2001), several scholars have investigated the impact that the background of experience accumulated by serial inventors has on the characteristics of the new patents invented. This literature has generally used patent classes to characterize the degree to which patents are similar/dissimilar, a technique that has encountered increasing criticism (Strumsky *et al.*, 2012). Two patents are considered similar if they are assigned to the same patent class, or to the same combination of patent classes. Fleming (2001) shows that patents that are based on non-familiar components or on a novel recombination of familiar components are more likely to be a breakthrough. Conversely, patents exploiting the same combination used in the past are less likely to be a breakthrough. Audia and Goncalo (2007) broaden these findings by showing that inventors who were more successful in the past are more likely to move along incremental lines, rather than experimenting on new trajectories. They speculate that people who experienced success are more productive of new ideas, but these ideas are less likely to be divergent, since people enjoy the payoffs from exploiting known areas instead of engaging in new exploration. Conti et al. (forthcoming) supported this argument by studying a rich database of European patents. They conclude that continuing to use known technological components increases the productivity in numerical terms, but reduces the probability of producing high quality inventions. In conclusion, there is growing evidence that serial inventors are likely to produce more patents, but the new patents would be of lower impact (Gambardella *et al.*, 2012).

In line with these findings, we formulate our second research hypothesis as follows:

H2: Inventions that rely on the prior technical knowledge set (are based on other inventions within the same technological domain) are less likely to be important than inventions that do not rely on a prior technical knowledge set.

3. Data and Methodology

3.1 Sample

We draw our analysis based on a data-set of 373 unique patent−academic inventor combinations. Academic inventors were identified as the scientists in all areas of physics who were appointed as "Fellows of the American Physical Society" (APS) prior to 2005, had a US academic affiliation at that date and had been inventors of at least one patent deposited at the US Patent and Trademark Office (USPTO) between 1992 and 2005.[2] Full patent records were retrieved from the archives of Thompson Innovation at the beginning of 2012, along with patent citation records. Information on the scholarly activity of the academic inventors was retrieved from the archives of ISI-Web of Knowledge and complemented with individual information on the inventor (gender, PhD year, affiliation, etc.). These were compiled from faculty web pages and CVs available on the web.

In order to avoid using patent documents that may constitute substantial duplicates, we worked carefully to clean away records from the same patent family as defined by those having at least one common priority code. In case of patent families, we choose to keep only the one record that had the largest number of citations and, with equal citations, the one issued at the earliest date. This resulted in 98 drops, leaving us with 295 unique patent− inventor combinations from 95 different inventors.

3.2 Dependent Variable

During spring 2012, we retrieved the citations received by the 295 patents and restrict the citation window to 6 years since the year of priority. Recall that our aim is to investigate correlations between the inventor knowledge set and the prospective importance of a patent. Prior contributions have indicated that citations are positively correlated to patent importance (Jaffe *et al.*, 1993) and patents that are cited more frequently are related to innovations with higher social (Trajtenberg, 1990) and private (Griliches, 1981; Hall *et al.*, 2005) value. It has also been shown that the citations received in the first years are a strong predictor of later citations (Fleming, 2001; Gittelman and Kogut, 2003). It seems therefore safe to adopt 6-years citation counts as a proxy for the importance of an invention.

Since our patents are deposited in different years, the window of observation of patent citations is dynamic. For example, for the patents deposited in 1992, we computed the citations received until the end of 1998. For those deposited in 1993, we computed the citations received until the end of 1999 and so on. For the most recent patents in our data-set, i.e. those issued in 2005, citations were computed until the end of 2011.

A description of the variables used in the analysis is provided in Table 1.

3.3 Independent Variables

We have two independent variables of prior scientific and technical knowledge set.
   (1) *Prior scientific knowledge set*: Because a standard methodology does not exist when it comes to accounting for the prior scientific knowledge set, one contribution of this

---

[2] Full information on the sampling procedure is given in Franzoni (2009).

Table 1. Variables construction

Variable	Description
Forward citation	Number of forward citation after 6 years
Prior scientific knowledge	Binary variable being 1 if the patent content is resonant with at least one prior publication of the scientists; 0 otherwise
Prior technical knowledge	Number of prior patent, which their content is in resonant with focal patent
Inventor technical productivity (log + 1)	The logarithm of prior patents plus one
Inventor scientific productivity (log + 1)	The logarithm of papers in the year of priority and two previous years plus one
Number of claims (log)	The logarithm of number of claims
Number of inventors (log)	The logarithm of number of inventors in the same patent
Count of cited patents (log + 1)	The logarithm of patents that have been cited by a patent (backward citations) plus one
Count of scientific articles cited (log + 1)	The logarithm of publications that have been cited in the patent document (backward citations) plus one
First patent	Binary variable being 1 if the inventor is at his/her first patent; 0 otherwise
Gender (male)	Binary variable being 1 if the inventor is a man; 0 if a woman
Career-age and career-age squared	Years since PhD was awarded at the patent priority year
Assignee type	Set of three binary variables based on different types of assignees: firm, university or others; 0 otherwise
Assignee's patenting experience	Categorical variable based on six intervals of prior patents count (0; 1−10; 11−100; 101−1,000; 1,001−10,000 and over 10,000).
Patent scope	Number of IPC classes assigned to the patent
Patenting trend	Number of patent granted by USPTO after 5 years
Spatial distribution	Number of patents granted by USPTO in the US state of assignee
IPC classes dummies	Set of 22 dummy variables based on two-digit main IPC class
Year dummies	Set of 14 dummy variables based on the priority year of the patent

paper is to fill this methodological gap. We explain the information collection and the methodology for treatment in this sub-section.

We begin by retrieving information on the scholarly activity of inventors. One advantage of using patents of faculty members is that the prior scientific knowledge set of the inventor is well documented. We searched the articles published on scientific outlets by the individuals between 1990 and 2005 in ISI Web of Knowledge. For each article, title and full abstract were coded in chronological order. Our aim is to study if the inventions that have emerged in domains of expertise that a scientist masters because of his/her scholarly knowledge are more cited on average than those that have emerged in other areas. Therefore, we want to know the degree to which the inventor is an expert (in scientific terms) of the domain employed in a certain invention. To gain this information, we compare the *content* of the patent to the *content* of the scholarly articles that the person has published until the time of patent priority. Comparison is based on semantic similarity. Rather than doing this screening manually, we use an algorithm of semantic analysis contained in a commercial software for scientific research, called Crawdad Text Analysis System 1.2. The software compares a single document against a group of documents (the set of all publications authored by the inventor) and outputs a score of resonance for each pair of documents based on the

co-occurrence of the same semantics and the co-occurrence of word pairs in the title and abstract of the patent and the title and abstract of each article. Based on Franzoni and Scellato (2010) who used the same algorithm for a similar purpose and validated the scores by means of a panel of experts, we choose a quite restrictive threshold of similarity (0.1). We code patents as related to the scholarly work of the inventor if the patent was similar ("resonant" in the wording of the software) to one or more scientific articles, or unrelated in case we find no scholarly publication with a similar content. Our threshold and consequent measure of relatedness are quite conservative. We find a total of 71 (24.07 per cent) patents that are related to at least one article and 224 (75.93 per cent) patents that resulted to be unrelated to prior scientific research.

(2) *Prior technical knowledge*: We use the same methodology in order to measure related prior technical knowledge set. The technical knowledge set has been compiled by retrieving information on the prior patenting activity of the inventors. For each inventor, we collected all patents issued prior to the years of observation. We characterize inventors in terms of the number of patents that are resonant with any prior patent, including patents issued before the window of observation 1990–2005. Like before, the abstracts are compared using Crawdad Text Analysis System 1.2. For robustness check, we also use the discounted stock of normalized number of prior related patents. Under the assumptions that more recent patents are more relevant than older ones, we use an annual discount rate of 10 per cent. This implies that patents filed 11 years ago have virtually no effect.

3.4 Control Variables

We controlled for a number of potentially confounding factors that may concur to determine the importance of a patent and/or to affect our measure of patent importance. Potentially confounding factors relate to characteristics of the inventors, characteristics of the assignee, patent characteristics and annual trends. In terms of individual characteristics, we control for gender and age, which were found to correlate—among other things—with patent performances (Azoulay *et al.*, 2007; Stephan *et al.*, 2007). We also control for the individual productivity in terms of patents (Gambardella *et al.*, 2012) and scholarly works (Fleming and Sorenson, 2004), whether or not the inventor is at his/her first patent, and how productive he/she has been in the last years in terms of both patents and scholarly articles. Concerning the characteristics of the assignees and in line with the findings of prior literature (Lissoni *et al.*, 2010), we control for the assignee type (whether it is a university; a firm; or a different residual type), and for the assignee's prior patenting experience. Concerning patent features, we control for a number of patent characteristics that may be correlated to unobserved heterogeneity. These are the number of claims, the number of backward citation (from patents and scientific articles separately),[3] the number of inventors and the broadness of scope of the patents [number of International Patent Classification (IPC) classes assigned to the patent] (Trajtenberg *et al.*, 1997). In order to account for trends in the number of patents issued after the current year, which affects the amount of potentially citing documents and hence of total citations, we control for the number of patents granted by the USPTO in the 5

---

[3] It has been shown patents which cite published material—peer review or commercial—receive more citations (Fleming and Sorenson, 2004).

years after the priority year. For robustness check, we account for heterogeneity caused by the geographic distribution of patents by controlling for the number of patents granted in the same year and in the same US state where the assignee is located. As a further check, we control for trends that may occur at a single patent class by considering the number of patents granted in the 5 years after a focal patent in the same one-digit patent class. We additionally control for the year in which the patent was filed (priority year dummies), and the technological domain to which the patent belongs (two-digit patent class dummies).

3.5 Descriptive Statistics

Summary statistics of all variables used in the analysis is reported in Table 2. Several variables are transformed in natural logarithm to correct for skewed distributions. A

Table 2. Descriptive statistics

Variables	Obs.	Mean	St. dev.	Min.	Max.
Forward citations	295	18.59	35.78	0	421
Prior scientific knowledge (= 1 if the patent content is resonant)	295	0.24	0.43	0	1
Prior technical knowledge	295	0.46	1.18	0	8
Inventor technical productivity	295	9.94	17.95	0	76
Inventor technical productivity (log)	295	1.52	1.23	0	4.34
Inventor scientific productivity	295	22.53	19.09	0	90
Inventor scientific productivity (log)	295	2.77	1.02	0	4.51
Number of claims	295	25.31	19.87	1	176
Number of claims (log)	295	2.97	0.76	0	5.12
Number of inventors	295	3.53	1.94	1	13
Number of inventors (log)	295	1.11	0.55	0	2.56
Count of cited patents	295	15.99	24.65	0	190
Count of cited patents (log)	295	2.27	1.01	0	5.25
Count of scientific articles cited	295	10.39	15.02	0	128
Count of scientific articles cited (log)	295	1.70	1.26	0	4.56
First Patent (= 1 if it is first patent of inventor)	295	0.30	0.46	0	1
Gender (= 1 if inventor is a man)	295	0.95	0.21	0	1
Assignee patenting experience	292	3.08	1.30	1	6
Patent scope	295	1.90	1.76	1	15
Patenting trend	295	760,008.7	101,443.8	523,070	875,789
Patenting trend (log)	295	13.53	0.15	13.17	13.68
Career age	295	21.98	11.35	4	55
Career age squared	295	611.60	642.57	16	3,025
University (= 1 if assignee is a university)	295	0.61	0.49	0	1
Firm (= 1 if assignee is a firm)	295	0.29	0.45	0	1
Others (= 1 if assignee is individual/government/both university and firm)	295	0.10	0.30	0	1
Spatial distribution	263	6,949.82	6,298.96	55	19,692
Spatial distribution (log)	263	8.34	1.15	4.01	9.89
Priority year	295	1,998.45	3.63	1992	2005

preliminary inspection of the data shows the expected high variance and skewness of forward citations. Thirty-two patents (10.58 per cent) have not received any citation in the 6 years after priority. Twenty-six had received one citation (8.8 per cent). Fifty-nine received between two and five citations (20.0 per cent). The median number of citation is 7 and the mean is 18.6. There are 9 patents that have more than 100 citations and some outliers with a maximum of 421 citations received by one patent.

The inventor experience is also highly variable, ranging from inventors with no prior patents and a maximum of 76 prior patents. Table 3 reports the distribution of patents among inventors. The scientists received their PhD between 1950 and 1991 and, accordingly, their career age at the priority year ranges from a minimum of 4 years to a maximum of 55. The scientists also vary considerably for their scientific production. On average, each scientist published 22.6 papers during the 3 years preceding the patent, with considerable variability. Regarding the assignee type, 61.1 per cent of patents are owned by universities (or intermediary institutions owned by universities). This result is aligned with prior studies of American universities (e.g. Thursby *et al.* (2009) showed that around 70 per cent of academic patents in the USA are owned by universities). About 29 per cent of patents in the sample are assigned to firms, and the rest is assigned to individuals (6.44 per cent), US governmental bodies (1.36 per cent) or both universities and a firm (2.03 per cent).

Concerning the number of patents that the assignees had prior to the current patent, we go from a low of zero, to a high of more than 10,000. Patents are coming from 21 different technological domains measured by the IPC class at the first two digits. Patents on average cited 15.99 prior patents and 10.39 scientific articles. Moreover, 95 per cent of inventors in the sample are men while only about 5 per cent are women.

Concerning the prior knowledge sets of the inventor, the majority of patents (67 per cent) are not directly related to technical and scientific knowledge of the inventor. About one in four

Table 3. Distribution of patents for each inventor ($n = 95$)

Prior patents	Freq.	Per cent	Cum.
0	27	28.42	28.42
1	21	22.11	50.53
2	15	15.79	66.32
3	6	6.32	72.63
4	5	5.26	77.89
5	2	2.11	80
6	2	2.11	82.11
7	2	2.11	84.21
8	1	1.05	85.26
9	4	4.21	89.47
10	3	3.16	92.63
12	2	2.11	94.74
13	1	1.05	95.79
21	1	1.05	96.84
22	2	2.11	98.95
76	1	1.05	100
Total	95	100	

Table 4. Patents by prior knowledge set of the inventor

Prior technical knowledge	Prior scientific knowledge		
	0	1	Total
0	200 (74.63 per cent)	68 (25.37 per cent)	231 (78.31 per cent)
1	53 (82.81 per cent)	11 (17.19 per cent)	64 (21.69 per cent)
Total	224 (75.93 per cent)	71 (24.07 per cent)	295

patents (24.07 per cent) is resonant with prior scientific knowledge, as visible in publications, and about one in five (22 per cent) is resonant to the prior technical knowledge, as visible in inventions.[4] The mean of technical relatedness is 0.46. Only 4 per cent of the inventions rely at the same time on the specific technical and scientific knowledge of the inventor. Table 4 shows the distribution of the technical and scientific knowledge sets of the inventors.

In the domains examined, we observe the expected general surge in the number of (utility) patents granted, with approximately three steps increases of 50,000 patents each: the first increase happened in 1993, the second in 1998 and the third in 2010 (Figure 1).

3.6 Methodology and Estimation

We model the importance of patents (proxied by forward citation) as a function of the scientific knowledge set (measured by the existence of publications whose content is related to the patent), and of the technical knowledge set (measured as number of prior patent whose content is related to the patent), plus a number of controls and an error component.

Since the independent variable is a positive integer, we model the data using count models. Given the high dispersion of our variable, the general assumptions underlying Poisson models that the conditional mean be equal to the variance is violated, suggesting to adopt either Negative Binomial or Quasi-Maximum-Likelihood (QML) Poisson. Both techniques lead to unbiased estimates in case of overdispersion and QML Poisson is deemed to imply fewer restrictions than Negative Binomial (Wooldridge, 2002). However, in our case, the Hausmann test indicates that QML Poisson is inconsistent ($p = 0.000$). Therefore, we chose to comment the result of Negative Binomial estimation and repeat the analysis with QML for comparison. The results are almost invariably consistent.[5] To cope with problems caused by having repetitive inventors, we use robust standard errors adjusted for clustering at the inventor level in order to allow for non-independence of the observations for the same inventor. This implies that the conditional mean is:

$$E(\text{FC}|X) = \exp\left[\beta_i \cdot X_i + \sum \alpha_i \cdot \text{Year\_Dummy}_i + \sum \gamma_i \cdot \text{IPC} - \text{Dummy}_i + \sum \delta_i \cdot Z_i\right],$$

where X_i is a vector of independent variables (*technical knowledge set* and *scientific knowledge set*) and Z_i is a vector of controls variables that is meant to clear off the effect of

---

[4] We find a maximum of eight prior patents related to current patent.

[5] For brevity, the QML estimates are not reported and are available upon request.
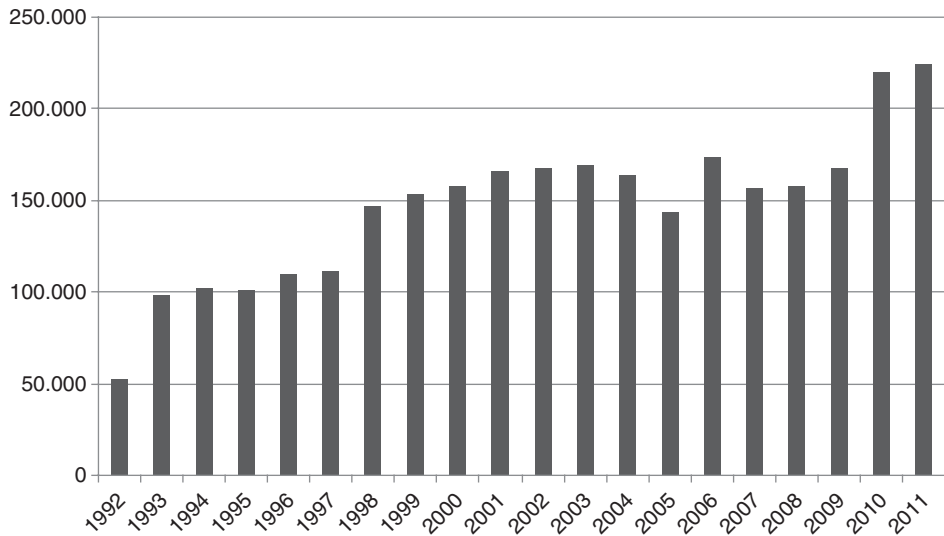
Figure 1. Yearly distribution of patents (utility) granted by USPTO (data extracted from USPTO: http://www.uspto. gov/web/offices/ac/ido/oeip/taf/reports_stco.htm)

potentially confounding factors. To control for the personal characteristics of the inventor, we include the career age (plain and squared, to account for known curvilinear effects), gender, prior experience of patents (log), count of scientific articles published during the last 3 years (log) and a dummy variable capturing inventors at their first patent. We control for patent characteristics that may capture unobserved heterogeneity such as the number of claims (log), the number of inventors (log), the number of backward citations (patents and scientific articles separately (log)) and the broadness of patent scope (number of IPC classes to which the patent has been assigned). We also control for the prior experience of the assignee with patents at the time of the deposit. The rationale for including these controls is that these characteristics may affect the degree to which patents get cited in ways that do not reflect the importance of the invention, but rather reflect the expertise of the assignee with patents, the accuracy with which the patent document has been prepared, etc.

Moreover, we control for changes in patenting trends, as captured by the number of patents granted by the USPTO (log) in the 5 years after the priority.[6] To account for potential effects of time and market trends, we include a set of single-year dummy variables, based on the year of priority (Year_Dummy). We finally include a set of 21 dummy variables capturing the propensity of patents in different technological classes to receive citations.

Table 5 reports the correlation matrix.

---

[6] For robustness check, alternatively we control for number of patents filed 5 years after priority year in the USPTO in the same patent class (one digit).

Table 5. Correlation matrix

	Forward citations	Prior scientific knowledge	Prior technical knowledge	Inventor technical productivity (log)	Inventor scientific productivity (log)	Number of claims (log)	Number of inventors (log)	Count of cited patents (log)	Count of scientific articles cited (log)	First patent	Gender	Assignee patenting experience	Patent scope	Patenting trend (log)	Career age	Career age squared	University	Firm
Forward citations	1.00																	
Prior scientific knowledge	0.18	1.00																
Prior technical knowledge	0.01	−0.12	1.00															
Inventor technical productivity (log)	0.09	−0.16	0.53	1.00														
Inventor scientific productivity (log)	0.03	0.15	0.02	−0.01	1.00													
Number of claims (log)	0.09	−0.12	0.02	0.02	0.00	1.00												
Number of inventors (log)	0.08	0.01	−0.14	−0.11	0.10	0.18	1.00											
Count of cited patents (log)	0.07	−0.08	0.06	0.10	−0.05	0.03	−0.04	1.00										
Count of scientific articles cited (log)	0.17	0.06	0.01	0.00	0.07	0.05	0.11	0.33	1.00									
First patent	−0.11	0.01	−0.19	−0.59	−0.09	0.00	0.06	−0.07	−0.03	1.00								
Gender	0.04	0.01	0.07	0.11	0.06	0.02	−0.04	−0.04	−0.08	−0.09	1.00							
Assignee patenting experience	0.05	−0.04	0.05	0.08	0.15	−0.05	0.05	−0.20	0.02	−0.02	−0.01	1.00						
Patent scope	−0.04	−0.10	−0.02	0.01	0.06	−0.03	−0.01	0.00	−0.02	−0.09	0.10	−0.14	1.00					
Patenting Trend (log)	0.02	0.01	0.04	0.08	0.18	0.15	0.17	0.00	0.03	−0.11	0.09	−0.10	0.06	1.00				
Career age	−0.04	−0.26	0.39	0.61	−0.18	0.01	−0.06	0.01	−0.10	−0.14	0.15	−0.04	0.04	0.15	1.00			
Career age squared	−0.02	−0.25	0.44	0.66	−0.16	−0.02	−0.09	0.00	−0.09	−0.14	0.14	0.01	0.01	0.07	0.97	1.00		
University	0.12	0.16	−0.03	−0.05	0.08	0.13	0.04	−0.11	0.20	−0.07	−0.08	0.19	0.05	0.23	0.01	0.00	1.00	
Firm	−0.08	−0.22	0.04	0.10	−0.05	−0.13	−0.06	0.18	−0.16	−0.01	0.07	−0.06	−0.09	−0.22	0.03	0.05	−0.80	1.00

4. Results

Table 6 shows the result of the estimates using Negative Binomial.

Model 1 reports the results for the full sample of 295 patents by including the prior scientific and technical knowledge set. In Model 2, we include the square of prior technical knowledge set in order to control for possible nonlinear relationship between technical knowledge set and patent importance. In Model 3, we use normalized discounted stock of patent as an alternative for prior technical knowledge set and considering the time difference of patents and Model 4 reports the results for a sub-sample of 237 patents in which we have omitted the patents of the inventors that had no prior patent experience. All econometric models strongly support our hypothesis 1 that patents which are related to the prior scientific knowledge set have on average more citations. By looking at the coefficient of the Negative Binomial model, we can see that patents which are related to prior publications receive on average 50 per cent more citations than the ones which are not related to prior publications.[7]

Concerning our hypothesis 2 that the prior technical experience of patents is associated to a reduced importance of the focal patent, the results of the unrestricted sample provide a more nuanced result than the one expected. We see in fact that there is evidence of an inverted U-shape relationship between prior technical knowledge and patent importance. The coefficient of the square of technical knowledge set is negative and significantly different from zero. The results are qualitatively similar in Model 3 when we use an alternative measure of technical knowledge set which take into account time difference between patenting activities. When (Model 4) we restrict the sample to include only the patents that had a prior track of technical knowledge (in other words we drop first patents), the results stay invariant.

Results of the control variables are aligned to those commonly established in the literature, although not always significant when different estimation methods are used. We can observe that (undifferentiated) experience with patents is associated to more citations. More claims are associated to higher importance. Men produce on average better cited patents than women. In addition, we can see that the number of patents and scientific articles cited in the focal patent are positively related to forward citations. Interestingly, patents assigned to universities on average receive more citations than the baseline. The baseline here is constituted of "other assignees types", and includes individual assignees, as well as governmental bodies. Similarly, patents assigned to firms exhibit a citation premium, compared to the baseline.

4.1 Robustness Check

In order to check the robustness of the estimates to the choice of the model, we have tried several alternative functional forms. Table A1 reports the results of these robustness checks.

First, we include an additional variable of geographic distribution in order to control for variation in patenting activities across US states. Since there are missing values in assignee address, the sample is reduced to 263 observations.

---

[7] $= \exp(0.408) - 1$.

Table 6. Negative Binomial models. Dependent variable: forward citation

	Model 1	Model 2	Model 3	Model 4
Prior scientific knowledge	0.408	0.400	0.399	0.437
	(0.197)**	(0.196)**	(0.198)**	(0.208)**
Prior technical knowledge	0.123	0.351		0.328
	(0.130)	(0.171)**		(0.175)**
Prior technical knowledge square		−0.046		−0.041
		(0.020)**		(0.020)**
Prior technical knowledge (normalized stock of patent)			4.263	
			(2.374)*	
Prior technical knowledge (normalized stock of patent) square			−9.018	
			(4.157)**	
Inventor's technical productivity (log)	0.180	0.141	0.158	0.224
	(0.124)	(0.116)	(0.113)	(0.133)*
Inventor's scientific productivity (log)	0.155	0.150	0.143	0.060
	(0.072)**	(0.073)**	(0.073)**	(0.099)
Number of claims (log)	0.328	0.325	0.315	0.351
	(0.087)***	(0.083)***	(0.087)***	(0.099)***
Inventor count (log)	0.258	0.274	0.258	0.335
	(0.157)	(0.157)*	(0.160)	(0.193)*
Count of cited patents (log)	0.115	0.120	0.124	0.078
	(0.077)	(0.072)*	(0.071)*	(0.083)
Count of scientific articles cited (log)	0.183	0.185	0.190	0.228
	(0.065)***	(0.066)***	(0.065)***	(0.071)***
First patent	0.039	0.012	0.018	NO
	(0.277)	(0.270)	(0.272)	NO
Gender (male)	0.585	0.578	0.604	1.118
	(0.485)	(0.472)	(0.469)	(0.422)***
Assignee experience	0.009	0.007	0.012	−0.037
	(0.061)	(0.060)	(0.060)	(0.069)
Patent scope	0.000	0.004	0.001	0.005
	(0.036)	(0.036)	(0.037)	(0.044)
Trend	−12.568	−11.047	−11.870	−1.581
	(5.380)**	(5.518)**	(4.913)**	(5.127)
Career_age	Yes	Yes	Yes	Yes
Career_sq	Yes	Yes	Yes	Yes
Assignee type	Yes	Yes	Yes	Yes
IPCdummy1	Yes	Yes	Yes	Yes
Years dummy 1−14	Yes	Yes	Yes	Yes
_cons	167.414	146.870	158.151	20.160
	(72.983)**	(74.795)**	(66.613)**	(69.290)
lnalpha	0.133	0.127	0.126	0.100
	(0.084)	(0.086)	(0.088)	(0.097)
N	295	295	295	237

Note: Clustered robust std. error in parentheses. $*p < 0.1$; $**p < 0.05$; $***p < 0.01$.

Second (Model 2), following the methodology of Fleming (2001), we add a control for relevant activities in technological domain (prior technological domain) by comparing the main IPC of the patent to those of prior patents by the same person. Third, we use the share of patent citations and scientific articles citations over the total, rather than the sheer counts (Model 3). Fourth, we exclude claims as a control variable since a recent article (Strumsky *et al.*, 2012) has criticized the use of the number of patent claims as antecedent of patent importance (Model 4). Fifth, we account for the potential effect of the changes in the disclosure policy of the USPTO occurred since November 2000, which mandates to publish the applications of pending patents after 18 months since the priority day. The policy change can impact citation trends in ways that are not captured by year dummies; therefore, in Model 5, we included a dummy which takes the value of one if the patent was filed after November 2000. Last, we control for variations in patenting activities in each patent class, which can affect forward citations. For this purpose in Model 6, we control for the number of patents filed at the USPTO in the 5-year period after the priority year of the focal patent. Results are robust to all these alternative specifications.

In Table A2, we test the robustness of our estimates to the choice of treatment for patent families. Recall that in the paper we choose to use a single focal patent in a family, as the one that had more citations or had an earlier priority date, in case the number of citations was the same. In order to check the robustness of the result to this strategy, we re-estimate the Models 1 and 2 of Table 6 including all patents in the family (373 in Model 1 and 348 in Model 2). Both models confirm the hypotheses 1, but in this case the result observed for the technical knowledge set is not robust.

5. Conclusion, Limitations and Discussion

The first aim of this paper was to understand the degree to which inventors rely on their prior scientific knowledge set, as opposed to relying on their prior technical knowledge set. Our results suggested that about one in four inventions issued to academics are in areas in which inventors have prior publications in scientific journals and about one in five are in areas of prior technical experience, as witnessed by prior patents. Around 17 per cent of patents are related with both prior scientific knowledge and prior technical knowledge. If the result were to be confirmed in further analysis, this would mean that a sizable, but largely minoritarian share of patents relates substantially to the scientific production of academics.

Our second aim was to study the correlation between prior knowledge sets and the relative importance of patents. We showed that patents based on prior scientific knowledge sets receive on average 50 per cent more citations than patents non-related to prior scientific knowledge. This is a robust and remarkable increase. Conversely, inventions based on areas in which the inventor had prior technical expertise have an inverted U-shape relationship with forward citations. This implies that technical expertise correlates to more important patents, but the relationship may become negative (patents receive fewer citations) as the inventor continues to exploiting the same area. This result is consistent to what emerged in prior studies (Gambardella *et al.*, 2012; Conti *et al.*, forthcoming), which highlighted the decreasing marginal returns of incremental inventive strategies. Overall, the results seem to suggest that there are high returns from re-using scientific knowledge, positive returns from re-using recent technical expertise, but decreasing returns from re-using many times the same technical expertise.

Our third aim was to propose and apply a novel methodology that looks at document relatedness by means of content analysis, rather than on temporal contiguity of events. We employed a content analysis software to identify semantic similarity and assess whether or not the content of a certain invention (as described in the patent abstract) is related to the prior scientific knowledge set of the academic inventor (as described in abstracts of scientific publications) and/or to the prior technical knowledge set (as described in the patent abstracts). This methodology can be replicated and applied to potentially broader areas of investigations. However, it is important to stress that our measure of relatedness is all but perfect, and the estimate may be affected by measurement biases in important ways. To mention just one potential source of biases, it is known that patent documents are edited by the attorneys. Therefore, the wording used in patents may not necessarily reflect the wording used in the scientific literature, and this would cause our software to underestimate the degree of relatedness in document content. Further analysis would be advisable in the area, possibly with the use of surveys. Furthermore, improvements in the methodology may certainly be advisable in terms of calibration (threshold of similarity) and in terms of the breadth of text analyzed. In this paper, we screened only the title and abstract of the documents, which simplifies the computational burden and offers the advantage of using documents of comparable length. A more refined text analysis could be obtained if length-normalized comparisons are drawn on the full texts.

The paper has several limitations that could be addressed in future research. We already mentioned two limitations that relate to the methodology of content analytics. In addition, our results are based on a limited sample of academic inventors working in the field of physics. It has been shown that the relationship between scientific productivity and patenting varies across fields (Stephan *et al.*, 2007) and even across sub-fields of the same area (Calderini *et al.*, 2009). Second, the inventors in our sample are all "Fellows" of the American Physical Society which implies that they are exceptionally brilliant scientist within their field. It is possible that further analyses conducted on different areas or more representative samples will lead to different results. Finally, in this paper we used a sample of academic inventions. It is important to acknowledge that the results should not be generalized outside of the realm of academe. On a more representative sample, the PatVal survey has shown that inventors consider prior "patent literature" and prior "scientific literature" among the top three sources of knowledge used in the invention process for relative importance, although no external source is rated as very relevant in absolute terms (Giuri *et al.*, 2007). Future analysis could be directed to investigate the extent to which scientific and technical knowledge sets are relevant in firms' innovation processes.

Our results have several potentially relevant implications for scholars and practitioners. First, over the last 20 years, there has been an increasing interest toward academic patenting among researchers in different fields of science. A large body of literature has investigated the impact of a steady growth in patenting by university researchers on their scientific output. In general, the literature on academic patents has documented that academic patents tend to be more valuable and general than non-academic ones. Furthermore, productive scientists are more likely to become inventors than their less-productive colleagues (Breschi *et al.*, 2007; Stephan *et al.*, 2007; Fabrizio and Di Minin, 2008,) and to produce a flurry of scholarly publications before patenting (Azoulay *et al.*, 2007; Calderini *et al.*, 2007). This evidence seems to indicate that inventions are developed as a byproduct of a prolific research activity that the scientist has conducted for scientific

purposes. In practice, however, the correlation observed in prior works was strictly numerical, meaning that no prior study has established a clear link between the content of the flurry of publications that precede an invention and the invention itself. Our findings advance this literature as they show that patents based on prior scientific research are on average more important (more cited) than those not based on prior scientific research, where the link is drawn not only on time-contiguity but also on *content*. The findings support the hypotheses that scientific knowledge is a valuable input to the patent process, at least for academic patents.

Second, our results are potentially of interest for agents involved in the technology transfer process at several levels. At the level of university administrators, our findings confirm that scientific research is a needed input to achieve success in technology transfer. Our results are also potentially valuable to technology managers called to assess and triage disclosures on new inventions, amid limited resources and strong market uncertainty. Our findings indicate that inventions backed by scientific research, all other things being equal, are among the most promising and should be marketed with special care. Conversely, patents that arise from incremental technical expertise are less likely to be valuable and should be the target of analysis for triage.

Acknowledgements

References

Ambos, T. C., Makela, K., Birkinshaw, J. and D'este, P. (2008) When does university research get commercialized? Creating ambidexterity in research institutions source, *Journal of Management Studies*, 45(8), pp. 1424–1447.

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

AUTM (2011) *U.S. Licensing Activity Survey: FY2010* (Deerfield, IL: s.l. Association of University Technology Managers).

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

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

Bacchiocchi, E. and Montobbio, F. (2010) International knowledge diffusion and home-bias effect. Do USPTO & EPO patent citations tell the same story? *Scandinavian Journal of Economics*, 112(3), pp. 441–470.

Blumenthal, D., Campbell, E. G., Anderson, M. S., Causino, N. and Louis, K. S. (1997) Withholding research results in academic life science – Evidence from a national survey of faculty, *Journal of the American Medical Association*, 277(15), pp. 1224–1228.

Breschi, S., Lissoni, F. and Montobbio, F. (2007) The scientific productivity of academic inventors: new evidence from Italian data, *Economics of Innovation and New Technology*, 16(2), pp. 101–118.

Calderini, M., Franzoni, C. and Vezzulli, A. (2007) If star scientists do not patent: the effect of productivity, basicness and impact on the decision to patent in the academic world, *Research Policy*, 36(3), pp. 303–319.

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

Campbell, E., Clarridge, B., Gokhale, N., Birenbaum, L., Hilgartner, S., Holtzman, N. and Blumenthal, D. (2002) Data withholding in academic genetics – Evidence from a national survey, *Journal of the American Medical Association*, 287(4), pp. 473–480.

Conti, R., Gambardella, A. and Mariani, M. (forthcoming) Learning to be Edison: inventors, organizations and breakthrough inventions, *Organization Science*.

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

Czarnitzki, D., Hussinger, K. and Schneider, C. (2011) Commercializing academic research: the quality of faculty patenting, *Industrial and Corporate Change*, 20(5), pp. 1403–1437.

EC (2006) *Study on Evaluating the Knowledge Economy What are Patents Actually Worth? The Value of Patents for Today's Economy and Society* (Brussels: European Commission).

Elfenbein, D. W. (2007) Publications, patents, and the market for university inventions, *Journal of Economic Behavior & Organization*, 63(4), pp. 688–715.

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

Farrell, M. (2008) *Universities that turn research into revenue*. Available online at http://www.forbes.com/2008/09/12/google-general-electric-ent techcx_mf_0912universitypatent.html (accessed 19 July 2012).

Fleming, L. (2001) Recombinant uncertainty in technological search, *Management Science*, 47(1), pp. 117–132.

Fleming, L. and Sorenson, O. (2004) Science as a map in technological search, *Strategic Management Journal*, 25(8–9), pp. 909–928.

Franzoni, C. (2007) Opportunity recognition in technology transfer organizations five case studies from UK and Italy, *International Entrepreneurship and Management Journal*, 3(1), pp. 51–67.

Franzoni, C. (2009) Do scientists get fundamental research ideas by solving practical problems? *Industrial and Corporate Change*, 18(4), pp. 671–699.

Franzoni, C. and Scellato, G. (2010) The grace period in international patent law and its effect on the timing of disclosure, *Research Policy*, 39(2), pp. 200–213.

Gallini, N. T. and Wright, B. D. (1990) Technology-transfer under asymmetric information, *Rand Journal of Economics*, 21(1), pp. 147–160.

Gambardella, A., Harhoff, D. and Verspagen, B. (2012) The economic value of patent portfolios. CEPR Discussion Paper 9264.

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(4), pp. 366–382.

Giuri, P., Mariani, M., Brusoni, S., Crespi, G., Francoz, D., Gambardella, A., Garcia-fontes, W., Geuna, A., Gonzales, R., Harhoff, D., Hoisl, K., Le Bas, C., Luzzi, A., Magazzini, L., Nesta, L., Nomaleri, O., Palomeras, N., Patel, P., Romanelli, M. and Verspagen, B. (2007) Inventors and invention processes in Europe: results from the PatVal-EU survey, *Research Policy*, 36(8), pp. 1107–1127.

Göktepe-hulten, D. and Mahagaonkar, P. (2010) Inventing and patenting activities of scientists: in the expectation of money or reputation? *Journal of Technology Transfer*, 35(4), pp. 401–423.

Griliches, Z. (1981) Market value, R&D, and patents, *Economics Letters*, 7(2), pp. 183–187.

Haeussler, C., Jiang, L., Thursby, J. and Thursby, M. (2009) Specific and general information sharing among academic scientists. NBER Working Paper, Issue 15315.

Hall, B. H., Jaffe, A. and Trajtenberg, M. (2005) Market value and patent citations, *Rand Journal of Economics*, 36(1), pp. 16–38.

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(1), pp. 119–127.

Jaffe, A. B., Trajtenberg, M. and Henderson, R. (1993) Geographic localization of knowledge spillovers as evidenced by patent citations, *Quarterly Journal of Economics*, 108(3), pp. 577–598.

Jensen, R. and Thursby, M. (2001) Proofs and prototypes for sale: the licensing of university inventions, *American Economic Review*, 91(1), pp. 240–259.

Jensen, R. A., Thursby, J. G. and Thursby, M. C. (2003) The disclosure and licensing of university inventions. National Bureau of Economic Research, Inc. NBER Working Papers 9734.

Kogut, B. and Zander, U. (1992) Knowledge of the firm, combinative capabilities, and the replication of technology, *Organization Science*, 3(3), pp. 383–397.

Lissoni, F., Montobbio, F. and Eri, R. (2010) Ownership and impact of European university patents. Working Paper.

Mowery, D. C., Nelson, R. R., Sampat, B. N. and Ziedonis, A. A. (2001) The growth of patenting and licensing by US universities: an assessment of the effects of the Bayh−Dole act of 1980, *Research Policy*, 30(1), pp. 99−119.

Murray, F. and Stern, S. (2007) Do formal intellectual property rights hinder the free flow of scientific knowledge? An empirical test of the anti-commons hypothesis, *Journal of Economic Behavior & Organization*, 63(4), pp. 648−687.

Nelson, R. R. (2004). The market economy, and the scientific commons. *Research Policy*, 33(3), pp. 455−471.

Nelson, R. and Winter, S. (1982) *An Evolutionary Theory of Economic Change* (Cambridge, MA: Belknap).

Owen-Smith, J. and Powell, W. W. (2001) To patent or not: faculty decisions and institutional success in academic patenting, *Journal of Technology Transfer*, 26(1), pp. 99−114.

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

Scotchmer, S. (2004) *Innovation and Incentives* (Cambridge, MA: MIT Press).

Shane, S. (2002) Selling university technology: patterns from MIT, *Management Science*, 48(1), pp. 122−137.

Stephan, P. E., Gurmu, S., Sumell, A. J. and Black, G. (2007) Who's patenting in the university? Evidence from the survey of doctorate recipients, *Economics of Innovation and New Technology*, 16(2), pp. 71−99.

Strumsky, D., Lobo, J. and Van Der Leeuw, S. (2012) Using patent technology codes to study technological change, *Economics of Innovation and New Technology*, 21(3), pp. 267−286.

Swamidass, P. M. and Vulasa, V. (2008) Why university inventions rarely produce income? Bottlenecks in university technology transfer, *Journal of Technology Transfer*, 34, pp. 343−363.

Thursby, J., Fuller, A. W. and Thursby, M. (2009) US faculty patenting: inside and outside the university, *Research Policy*, 38(1), pp. 172−187.

Trajtenberg, M. (1990) A penny for your quotes − Patent citations and the value of innovations, *Rand Journal of Economics*, 21(1), pp. 172−187.

Trajtenberg, M., Henderson, R. and Jaffe, A. (1997) University versus corporate patents: a window on the basicness of invention, *Economics of Innovation and New Technology*, 5(1), pp. 19−50.

Walsh, J. P., Cho, C. and Cohen, W. M. (2005) View from the bench: patents and material transfers, *Science*, 309(5743), pp. 2002−2003.

Wooldridge, J. M. (2002) *Econometric Analysis of Cross-Section and Panel Data* (Cambridge: MIT Press).

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

Appendix

Table A1. Robustness checks, Negative Binomial estimates

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Prior scientific knowledge	0.550	0.347	0.343	0.366	0.478	0.376
	(0.204)***	(0.192)*	(0.208)*	(0.205)*	(0.174)***	(0.198)*
Prior technical knowledge	0.328	0.389	0.368	0.372	0.398	0.397
	(0.192)*	(0.181)**	(0.162)**	(0.194)*	(0.148)***	(0.161)**
Prior technical knowledge square	−0.059	−0.053	−0.040	−0.048	−0.087	−0.055
Prior technological domain	(0.021)***	(0.021)***	(0.024)*	(0.021)**	(0.023)***	(0.020)***
		−0.811				
		(0.333)**				
Inventor's technical productivity (log)	0.109	0.114	0.118	0.206	0.126	0.146
	(0.118)	(0.118)	(0.117)	(0.121)*	(0.126)	(0.114)
Inventor's scientific productivity (log)	0.141	0.161	0.138	0.136	0.097	0.129
	(0.084)*	(0.072)**	(0.078)*	(0.080)*	(0.082)	(0.072)*
Number of claims (log)	0.277	0.308	0.294		0.332	0.357
	(0.089)***	(0.085)***	(0.083)***		(0.084)***	(0.080)***
Inventor count (log)	0.353	0.270	0.314	0.324	0.265	0.282
	(0.172)**	(0.153)*	(0.151)**	(0.157)**	(0.166)	(0.160)*
Count of cited patents (log)	0.185	0.129		0.111	0.165	0.094
	(0.083)**	(0.072)*		(0.075)	(0.082)**	(0.069)
Count of scientific articles cited (log)	0.225	0.192		0.189	0.218	0.171
	(0.065)***	(0.068)***		(0.067)***	(0.072)***	(0.066)***
First patent	−0.074	−0.053	−0.076	0.056	0.026	−0.060
	(0.296)	(0.268)	(0.269)	(0.267)	(0.284)	(0.277)
Gender (male)	0.553	0.594	0.517	0.673	0.480	0.276
	(0.514)	(0.462)	(0.486)	(0.415)	(0.507)	(0.557)
Assignee experience	−0.023	−0.002	−0.036	0.006	−0.039	0.025
	(0.074)	(0.059)	(0.062)	(0.064)	(0.066)	(0.061)
Patent scope	0.035	0.011	0.012	−0.024	−0.025	−0.001
	(0.036)	(0.037)	(0.037)	(0.034)	(0.031)	(0.031)
Trend	−0.903	−9.680	−10.313	−11.756	−0.087	
	(5.245)	(5.790)*	(4.485)**	(5.811)**	(0.688)	
Trend in patent class	0.017					
	(0.093)					
Share patent cited/total citations			−0.571			
			(0.264)**			
Policy change					−0.617	
					(0.216)***	
Spatial distribution (log)						−0.610
						(0.254)**
Career_age	Yes	Yes	Yes	Yes	Yes	Yes
Career_sq	Yes	Yes	Yes	Yes	Yes	Yes
Assignee type	Yes	Yes	Yes	Yes	Yes	Yes
IPCdummy1	Yes	Yes	Yes	Yes	Yes	Yes
Years dummy 1−14	Yes	Yes	Yes	Yes	No	Yes
	(71.189)	(78.493)	(60.582)**	(78.874)**	(9.299)	(3.529)
lnalpha	0.100	0.110	0.157	0.158	0.208	0.102
	(0.089)	(0.086)	(0.084)*	(0.087)*	(0.091)**	(0.089)
N	263	295	288	295	295	295

Table A2. Robustness check (patents belonging to patent family), Negative Binomial estimates

	Model 1	Model 2
Prior scientific knowledge	0.382 (0.193)**	0.392 (0.207)*
Prior technical knowledge	0.030 (0.158)	−0.003 (0.159)
Prior technical knowledge square	0.002 (0.031)	0.006 (0.033)
Inventor's technical productivity (log)	0.161 (0.113)	0.199 (0.135)
Inventor's scientific productivity (log)	0.113 (0.080)	0.019 (0.113)
Number of claims (log)	0.345 (0.092)***	0.369 (0.099)***
Inventor count (log)	0.280 (0.133)**	0.354 (0.156)**
Count of cited patents (log)	0.009 (0.077)	−0.048 (0.076)
Count of scientific articles cited (log)	0.145 (0.057)**	0.173 (0.063)***
First patent	0.009 (0.227)	
Gender (male)	0.561 (0.423)	0.990 (0.373)***
Assignee experience	−0.038 (0.077)	−0.066 (0.098)
Patent scope	−0.023 (0.034)	−0.020 (0.037)
Trend	−5.983 (4.675)	1.693 (5.239)
Career_age	Yes	Yes
Career_sq	Yes	Yes
Assignee type	Yes	Yes
IPCdummy1	Yes	Yes
Years dummy 1−14	Yes	Yes
_cons	0.903 (0.640)	1.636
lnalpha	0.217 (0.085)**	0.238 (0.095)**
N	373	310

Note: Clustered robust std. error in parentheses. *$p < 0.1$; **$p < 0.05$; ***$p < 0.01$.