European Management Review (2008) 5, 91–109

palgrave-journals.com/emr

University patenting and scientific productivity: a quantitative study of Italian academic inventors

Stefano Breschi[1], Francesco Lissoni[1,2], Fabio Montobbio[1,3]

[1]CESPRI, Department of Economics, Università L. Bocconi, Milan, Italy;
[2]Department of Industrial and Mechanical Engineering, Università degli studi di Brescia, Italy;
[3]Department of Economics, Università degli studi dell'Insubria, Varese, Italy

Correspondence: Stefano Breschi, CESPRI, Department of Economics, Università L. Bocconi, Via Sarfatti 25, Milan 20136, Italy.
Tel: +39 02 58363365;
E-mail: stefano.breschi@unibocconi.it

Abstract

Based on longitudinal data for a matched sample of 592 Italian academic inventors and controls, the paper explores the impact of patenting on university professors' scientific productivity, as measured by publication and citation counts. Academic inventors, that is, university professors who appear as designated inventors on at least one patent application, publish more and better quality papers than their non-patenting colleagues, and increase their productivity after patenting. Endogeneity problems are addressed by using instrumental variables and applying inverse probability of treatment weights. The beneficial effect of patenting on publication rates last longer for serial academic inventors. However, the positive effect of patenting on scientific productivity largely differs across scientific fields, being particularly strong only in pharmaceuticals and electronics.
European Management Review (2008) **5,** 91–109. doi:10.1057/emr.2008.9
Keywords: scientific productivity; university patents; technology transfer

Introduction

The increasing involvement of US universities in patenting and commercialization of research results is a well-documented phenomenon (Mowery *et al*., 2004), but recent studies have uncovered unexpected rates of university patenting also in Europe. In particular, new measurement efforts have highlighted that, although university-owned patents are still a relatively rare phenomenon in Europe, there exists a sizable and growing number of university-invented patents, that is inventions by one or more academic scientists whose intellectual property rights are assigned to business companies, governmental funding agencies or individual scientists. The most convincing evidence has been produced for Italy, France, Sweden, Finland and Germany (Meyer *et al*., 2003; Schmiemann and Durvy, 2003; Balconi *et al*., 2004; Lissoni *et al*., 2008), while case studies and survey research have suggested that a similar pattern could be found also in Spain, Belgium and the UK (Azagra Caro *et al*., 2003; Saragossi and van Pottelsberghe de la Potterie, 2003; Crespi *et al*., 2006). Taken together, university-owned and university-invented patents are now commonly referred to as 'academic patents'.

These findings have raised several concerns both in society and among scholars and practitioners about the potentially detrimental consequences of academic patenting on the pace of scientific and technological progress that might derive from (1) restrictions of access to the outcomes of (publicly funded) research and (2) changes in the academic scientists' incentives to carry out fundamental research. In this paper, we address in particular the second type of concerns by investigating the following questions. Does involvement in patenting affect academic scientists' research effort and productivity? Does it affect the direction of research, by diverting scientists' attention away from basic research and towards more applied fields of enquiry? Is there any detectable difference in the impact of patenting across technologies and disciplines? Does the nature of patent owners (academic *vs* business *vs* individual) bear an influence on how patenting affects research?

We attempt to provide an answer to such questions by building on our own previous research on Italian academic inventors (Balconi *et al*., 2004). Specifically, we examine the scientific productivity for a large sample of Italian academic inventors, that is, university professors who appear as designated inventors on patent applications at the European Patent Office (EPO), and compare it with a matched sample of non-patenting academic scientists. Our findings reveal that academic inventors tend to publish more and higher quality papers than their non-patenting colleagues, and increase further their productivity after patenting. The beneficial effect of patenting on publication rates last longer for serial inventors, that is, academic inventors with more than one patent. Moreover, results suggest no evidence of a shift away from basic towards applied research. However, the positive effect of patenting on the rate of scientific production seems to vary across disciplines, being particularly strong only in pharmaceuticals and electronics.

The paper is organized as follows. The next section provides a discussion of the controversial relationship between patenting and publishing, and puts forward some testable propositions. In the subsequent section, we describe the data set used and the procedures followed to construct our sample, while in the further section we introduce our basic econometric approach. The penultimate section reports our main results, discussing in detail the strategy used to cope with endogeneity issues and to identify causality in the relationship between patenting and publishing. The final section concludes.

Patenting–publishing: trade-off or complementarity?

The dramatic growth in academic patenting and licensing that has occurred over the last two decades or so has raised several concerns about the potentially negative effects that the commercialization of scientific discoveries could have on the conduct of academic research. In particular, it has been argued that financial incentives from patenting and licensing could shift the orientation of scientists away from basic and towards applied research and could also undermine their commitment to the norms of open science, thereby leading to undesirable behaviours, such as data withholding, secrecy and publication delays. At the opposite end of the spectrum, advocates of technology transfer via commercialization of research results have welcomed such trends, arguing that closer contacts between industrial and academic research will also bring benefits to the latter in the form of financial resources, access to costly instrumentation and sources of ideas. In this section, we discuss why and how patenting should be expected to affect an academic scientist's productivity. In particular, we try to summarize the main arguments that have been proposed in the literature on the detrimental or beneficial effects of patenting on the rate, quality and direction of scientific output, in order to derive a few empirically testable hypotheses.

Patenting and the rate of scientific production

Most of the expected negative effects associated with academic patenting and with a greater involvement of academic scientists in the commercialization of research findings are usually thought to derive from changes in the traditional incentive system of science. Patents are the results of an incentive system at odds with the one that has governed the scientific community over the past three centuries. Academic scientists build their careers chiefly upon reputation gained by claiming priority over scientific discoveries published on refereed journals (Merton, 1968; Cole, 1992). Such a reward system encourages scientists to disclose quickly and fully their research findings, via the publication of data, intense codification efforts of theories and methodologies, and repeated interaction and discussion with peers (Dasgupta and David, 1994). While also rewarding priority over technical inventions, the patent system differs from the norms of open science in that it promotes only incomplete and selective disclosure of knowledge. The amount of knowledge disclosed is limited to the minimum necessary for achieving patent protection over patentable results, while secrecy is heavily used to appropriate the returns from the non-patentable aspects of the invention (Cohen *et al*., 2000). This incentive mechanism entails the possibility of a trade-off between patenting and publishing. In its mildest form, the trade-off can appear as a publication delay dilemma: researchers concerned with patenting may be forced to wait before publishing any news on their discoveries, and keep them secret until the patent application has been filed, in order to avoid burning the novelty step of the application. Not only submitting a paper to a journal but also discussing it at conferences and workshops could invalidate the effort to obtain a patent.[1] Yet, to the extent that publication delay is the only potential adverse effect of academic patenting on the rate of publication, one should not be too worried. Any negative impact is likely to be only temporary, with no persistent change in the overall publication trend of a scientist. However, the trade-off between patenting and publishing can also show up in a stronger and potentially more negative way. As long as the patent holder imposes restrictions on the diffusion of related non-patentable knowledge assets, academic inventors may be forced to avoid publishing and keep secret those research results that contribute to such assets. Evidence on these restrictions is well established for sponsored research in the life sciences (Blumenthal *et al*., 1996; Bekelman *et al*., 2003). Moreover, Thursby and Thursby's (2002) survey of businesses who license from universities shows that 50% of the respondents reported that their contract contained clauses requiring delays of publication and rights to delete information. Walsh and Hong (2003) also show an increasing tendency for scientists to become more secretive and less willing to discuss and disseminate openly their research results in view of their perspective commercial exploitation. This line of arguments leads therefore to the following testable hypothesis:

> **H1A:** Publication restrictions will determine a decline in the academic inventor's rate of publication as a consequence of patenting, especially when the patent holder is a private for-profit organization concerned with secrecy.

Despite such concerns, however, there are several reasons to expect that, on the contrary, academic scientists who contribute to patenting will not sacrifice their publishing activity. In the first place, in many cases patenting is no more than an occasional event in an academic career, one from which no persistent impact is expected. Informal interviews with Italian academic inventors suggest that scientists are primarily devoted to the production of scientific papers, and treat patent matters as a secondary and rather infrequent issue. They consider patenting first and foremost as an incidental obligation arising from research or consultancy contracts or sponsorship agreements, especially with business companies. Moreover, scientists who do not entertain strong relationships with industry may also end up patenting, but only when they foresee potential applications of specific results arising from research, also when the latter is funded by public agencies or university block grants. Scientists of this kind will sometimes patent in their own name or in the name of either the funding agency or their university, depending on the provisions of the funding scheme. As long as patenting responds to this logic, we would expect a patenting–publishing trade-off to occur only for those scientists who are heavily and systematically committed to patenting with industry (i.e. serial inventors), whereas no significant effect should be found for scientists with one or very few patents and/or with patents held by public research organizations. In this respect, the previous hypothesis should be modified as follows:

> **H1B:** Patenting will determine a decline in the rate of publication only for serial academic inventors, especially those strongly involved in patenting with industry.

There are also several reasons to expect a positive impact of patenting on the scientific productivity of individual scientists. First, several studies have shown that interaction with industry may be a source of fertile research questions. Solutions to technical problems posed by industry may involve original scientific work, up to the point of opening up entirely new research avenues and provide the basis for the establishment of new scientific disciplines (Rosenberg, 1982; Mansfield, 1995, 1998; Siegel *et al*., 2003). Second, collaboration with industry allows scientists to gain access to data and instruments that may reduce the cost of conducting research or increase its productivity. More generally, Owen-Smith and Powell (2001) suggest that academic scientists may choose to engage in patenting for a variety of reasons, many of which positively feed back on their scientific production: increasing the odds of receiving financial funds from deans and firms, thanks to increased visibility and prestige; attracting Ph.D. and post-doctoral students by helping them to get jobs as developers of the patented inventions; or using patents as bargaining chips to enter research fields where other scientists or firms may have strong IPR positions. At the same time, one may argue that the positive effect arising from the access to such resources will be stronger when the relationship between scientists and industry develops over a long time, and it is not just the result of an occasional research contract. In other words, one may expect a positive association between patenting and publishing particularly for serial inventors, that is, inventors signing several related patents, rather than a single occasional patent.

> **H2:** Cognitive, technological and financial resources will determine an increase in the rate of publication of academic inventors, in particular of serial ones.

Patenting and the direction of scientific research

Even conceding that positive resource effects could counterweigh any negative impact deriving from an increased propensity to secrecy, some critics argue that a greater involvement of scientists in patenting could still bring detrimental consequences by diverting the allocation of time from basic towards applied research. This diversion would be especially likely to occur when the patented invention is no more than a prototype or proof of concept. In this case, after the patent filing, the academic inventor would be called or tempted to devote time and efforts to solve the many technical problems that stand in the way of a successful commercial application. While basic research can be portrayed as the unconstrained exploration of nature and theory, the focus of applied research is closer to industrial and potentially patentable applications. Lower levels of commitment towards basic research might therefore result either in a lower rate of publications in journals dedicated to fundamental research or in less ambitious publications, which will receive less citations from subsequent articles because of their narrower focus or lack of depth. Despite the intuitive appeal of this kind of arguments, there is still little empirical and theoretical research on the relationship between patenting and licensing, on the one hand, and the direction of research efforts, on the other. Thursby and Thursby (2007) examine the research profile of 3241 faculty members from six major US universities from 1983 to 1999, and find that the fraction of research that is published in 'basic' scientific journals has remained fairly constant, notwithstanding a tenfold increase in the probability of patenting. Moreover, they also show that both patenting and publishing rise and fall with age, following a life cycle such as the one envisaged by Stephan and Levin (1992). On the theoretical side, Thursby *et al.* (2007) propose a dynamic model of faculty behaviour consistent with these findings. In their model, scientists choose the amount of time to devote to applied and basic research and the amount of time to take as leisure, facing a fixed teaching load, in order to maximize (the net present value of) utility over their entire career. Utility is a positive function of scientific research output, market goods, leisure and the net present value of financial assets at retirement. In turn, the dynamics of financial assets is governed by current salary, which depends positively on the stock of knowledge, that is, the cumulated research output, and license income, which depends positively on the time spent on applied and basic research and on the stock of knowledge. Basic and applied research efforts may be either complementary or substitutes in the production of publications, while they are assumed to be complementary among each other and with the stock of knowledge in the production of licensable inventions. Time paths of research efforts and productivity are obtained through simulations of different scenarios. In particular, three basic scenarios are simulated. At one end of the spectrum, a scenario is simulated in which the only input in the production of publications is represented by basic research, while license output requires only applied research. Under these rather extreme assumptions, results show the presence of real effects of patenting and licensing: scientists divert time from basic to applied research and, as a consequence, the publication rate decreases compared to the case in which academic scientists do not derive any income from licensing. At the other end of the spectrum, a scenario is simulated in which applied research also produces publishable output, and basic and applied research are complements in both the research and license production functions. In this case, the ratio of applied to basic research effort increases in the presence of licensing compared to the case in which academic scientists do not derive any income from patenting; the level of basic and applied research, and therefore total research effort and output, are higher with licensing than without it. The intuition of this apparent paradox is that, since applied and basic research are complements in the production of both licensing and research, financial incentives associated with licensing induce researchers to shift the allocation of time from leisure to both types of research activities. Finally, a third scenario is also simulated in which basic and applied research are complements in the production of license output, but they are substitutes in the production function of publications. In this case, basic research efforts are always higher than applied ones regardless the licensing regime, and total research output does not decrease in a licensing regime compared to the case in which no income from licensing is allowed. In other words, the effect of patenting and licensing on the level of basic research effort and on the total research output seems not depend on the assumption that applied and basic research are either complements or substitutes in the production function of publications. The crucial point is whether or not applied research involved in patenting leads also to publishable output and thus adds to the stock of knowledge of a scientist. Overall, the model shortly outlined above suggests two alternative hypotheses with respect to the effect of patenting on the rate and direction of publication output.

> **H3A:** To the extent that applied research does not contribute to the stock of knowledge of a scientist (but is simply spent to develop licensable inventions), patenting will have a negative effect on the production of 'basic' scientific publications and on total research output.

> **H3B:** If both basic and applied research contribute to the stock of knowledge of a scientist, patenting will have a positive effect on the production of 'basic' scientific publications and on total research output.

Data and methodology

The data used in this paper come from the combination of two data sets: the EP-INV data set and the list of Italian professors provided by the Ministry of Education (MIUR). The EP-INV database contains all patent applications to the EPO signed by at least one inventor with an Italian address, from 1980 to 2000. Overall, the EP-INV database contains information on 30,243 inventors and 38,868 patent applications. The MIUR data set contains the names, age, affiliation, academic ranking (assistant, associate and full professor) and scientific discipline of all Italian professors who, in 2000, held a tenured position in the hard sciences (27,844 individuals).[2] By matching the two data sets at the level of individual researchers, we were able to identify what we call 'academic inventors', namely university professors whose name appears among the inventors on one or more patent applications. Overall, the matching procedure returned a total of 919 academic inventors responsible for 1475 patents and active in various scientific disciplines. For this paper we selected 296 scientists responsible for 729 patents and active in four scientific disciplines: chemical engineering, biology, pharmacology and electronic engineering. The reasons for focusing on this subset are twofold. First, collecting publication data at the individual level is a very time-consuming exercise, which requires frequent manual checks of the records; resource constraints suggested us to limit the sample at a manageable size, in order not to compromise data quality. Second, the four selected disciplines are those with the highest intensity of academic patenting, as measured by the percentage of academic inventors over all university professors in the discipline (Table 1).

The distribution of patents across the 296 selected academic inventors is highly skewed; most professors in our sample have signed only one patent, and very few more than five (Table 2). Moreover, most patents signed by academic inventors belong to business companies, as a result of contractual funding, with little meaningful differences across disciplines. The only exception is represented by biology, which records a higher number of both individual patents and patents owned by 'open science' institutions, namely universities and public funding agencies or labs (Table 3).

The next step of our methodology involved the construction of a control sample of non-patenting scientists. To this end, each academic inventor was matched to a control professor in the same scientific discipline, with the same academic ranking and of a similar age.[3] Controls were selected among professors who were never designated as inventors of patents applied for either at the EPO or at the United States Patent and Trademark Office (USPTO). Moreover, controls were chosen among the academic inventors' department colleagues or from universities of similar size and importance, or from the same region.[4]

Table 1 Italian academic inventors, selected scientific fields

	Professors active in 2000	*Of which academic inventors (percentage)*
Chemical engineering	355	63 (17.7)
Pharmacology	613	83 (13.5)
Biology	1359	78 (5,7)
Electronic engineering	630	72 (11.4)
Total	2957	296 (10.0)

Academic inventors are university professors who appear as designated inventors of patent applications by universities, business companies and individuals. The academic disciplines listed in the table are those with the highest number of academic inventors. The second column reports the total number of Italian professors active in 2000 in the selected scientific fields. The third column reports the number and fraction (in brackets) of professors who have signed patents over the period 1980–2000.

Table 2 Distribution of academic inventors by number of patents

	Number of patents		
	1	*2–5*	*>5*
Chemical engineering	60.9	32.8	6.3
Pharmacology	63.1	28.6	8.3
Biology	70.5	23.1	6.4
Electronic engineering	56.2	31.5	12.3
Total	62.9	28.8	8.3

The table reports the percentage distribution of academic inventors in the selected scientific fields by total number of patents signed over the period 1980–2000 and scientific field. The table shows that most academic inventors are listed on just one patent application, while few appear on more than five.

Table 3 Ownership of academic inventors' patents

	Business	*Open science*	*Individuals*
Chemical engineering	127 (76.0)	22 (13.2)	18 (10.8)
Pharmacology	200 (75.2)	32 (12.0)	34 (12.8)
Biology	88 (48.6)	57 (31.5)	36 (19.9)
Electronic engineering	200 (78.1)	40 (15.6)	16 (6.3)
Total	615 (70.7)	151 (17.4)	104 (11.9)

Patents by academic inventors can be owned either by their universities as well as public research organizations, business companies or one or more individuals (among whom the inventors themselves); co-ownership by two different organizations is also possible, although not frequent. The table reports the number and fraction (in brackets) of the academic patents considered in this paper across three different types of owners: *business* companies, which include both Italian and foreign companies, '*open science*' institutions, which comprise universities, public research organizations and government agencies, and *individuals.* Patents by more than one applicant were counted more than once. The table shows that: (i) the largest part of academic patents is in the hands of business companies; (ii) some variance exists across disciplines, with biology recording a larger-than-average share of academic patents owned by 'open science' institutions and individuals, and electronics recording the lowest share of individual ownership.

For academic inventors and their controls we then collected data on scientific publications from 1975 to 2003 using the ISI Science Citation Index (SCI). For each article, we also collected the number of citations received up to 2003, which we used for weighing the quality of each professor's publications. Finally, we classified articles according to their level of 'basicness' by using a reclassification of ISI-SCI journals proposed by CHI Research Inc. (Hamilton, 2003), which assigns them to four categories ranging from very basic to very applied research.[5] The publication data set we assembled following these criteria has two main characteristics, which we addressed with a few rules of thumb. First, it is left-censored: we do not have information on when our professors started their research careers, as we do not know when they completed their studies or got their first academic job.[6] Thus, a zero publication record in a given year can be explained either by a low productivity, or because the scientist had not started her research career yet. To solve this problem, we decided to include in each year's sample only those inventors who were already 25 or older, and the related controls (e.g. a professor born in 1965 is included, with its control, in the sample in 1990). Italian students leave high school at 19 and most degrees requires 5 years of courses, and a final dissertation; this makes 24 the earliest possible graduation age. We also set an upper age limit at 70: this is because a few professors active in 2000 were already near that age, and we presume them not to be any more active later on. It is important to point out that these rules bear no relationship with the inventor/control status, therefore they should not affect the conclusions we reach. A second issue is related to the use of ISI-SCI as a source of information, both at a general level and for the Italian case. In general, we may expect the number of publications per-capita to change over time, simply because the number and composition of journals included in the ISI-SCI database change. For Italy, until recently the academic career system used to put a low prize on high publication rates, and even less on publications in international journals; in addition, until the 1980s it was still a rare event for Italian scientists to complete their Ph.D. abroad, and publish their first papers there. These two facts conjure up to make the ISI-SCI data quite a reliable source of information on the scientific productivity of later cohorts of professors, but not so much for the early ones.

Descriptive statistics show that academic inventors in our sample are more productive than their controls (Table 4). The mean and median number of publications of academic inventors over the period 1975–2003 is significantly higher than the corresponding figures for control professors. Moreover, the same results (not reported here) hold when citations or publications in basic science journals are considered. In the next section, we provide a detailed description of our econometric approach.

Econometric approach

The main question addressed by this paper is whether patenting activity increases or decreases the yearly number of scientific publications of academic inventors. Empirically, we model the problem by including a variable that is equal to zero in the years when the individual has not yet patented and is equal to one for all the periods after the year of the first patent. For example, if a scientist is present in our database from 1980 to 1999 and his/her first patent is in 1995, this variable is 0 in the period 1980–1994 and 1 in the period 1995–1999. This type of variable is known in the literature as a *treatment* and has been widely used for the evaluation of policy programmes and, in medicine, for the evaluation of the effect of specific prescriptions or doctor visits (e.g. Wooldridge, 2002: Chapter 18). In our context, patents represent a *treatment* given only to some individuals, that is, academic inventors, while control professors are *non-treated* individuals. In order to evaluate the effect of such a treatment we use a difference-in-difference (DID) econometric estimation. The main intuition behind DID estimation is the following. First, we calculate yearly changes in the number of papers published by academic

Table 4 Scientific productivity of academic inventors and controls

	N	*Mean*	*Std. dev.*	*Median*
Academic inventors				
Chemical engineering**	63	2.0	1.75	1.5
Pharmacology *	83	2.2	1.21	2.0
Biology*	78	2.5	2.10	2.0
Electronic engineering	72	1.7	1.04	1.4
All fields	296	2.1	1.60	1.8
Control professors				
Chemical engineering	63	1.3	1.10	1.1
Pharmacology	83	1.7	1.11	1.6
Biology	78	1.8	1.27	1.5
Electronic engineering	72	1.3	1.18	1.0
All fields	296	1.6	1.28	1.3

The table compares the number of publications authored by the academic inventors in our sample with those authored by a matched control sample of professors in the same disciplines and institutions, and of similar age. It can be seen that: (i) in both samples the distribution of publications is asymmetrical, as suggested by the literature (the mean is always larger than the median, which suggest a long right tail); (ii) academic inventors are more productive than their colleagues (both the mean and the median of their distribution are higher); (iii) both (i) and (ii) hold across the four disciplines considered in the paper.
The * and ** symbols indicate that the distribution of publications for academic inventors is significantly different from controls, respectively, at the 90 and 95% levels using a Kolmogorov–Smirnov test.

inventors and their colleagues. Second, we compare productivity changes for academic inventors after the filing of their first patent with contemporary changes in the productivity of their non-inventing colleagues (control group). A significant difference between academic inventors and the control group suggests that patenting has an effect on the individual scientific productivity, which may be positive or negative depending on the sign of the coefficient. Since our data allow us to observe the same individual for many years, we can tell apart the impact of the treatment variable from whatever individual characteristics affecting scientific productivity, which are constant over time (such as the individual scientist's biographical details or unobservable skills). That is, under specific conditions to be discussed below, the longitudinal nature of our data helps us avoiding to confound the treatment effect with unobserved individual effects constant in time.

Regressions

We estimate the following regression equation for the whole sample of treated and control scientists ($i=1,\ldots,592$ refers to individuals in the panel,[7] $t=1980,\ldots..1999$ indexes time):

$$y_{it} = \mathbf{x}_{it}\beta + \mathbf{z}_i\theta + Postpat_{it}\gamma + \alpha_i + u_{it} \qquad (1)$$

where y_{it} is a measure of scientific productivity at time t, $\mathbf{x}_{it}$ is a vector of time-varying explanatory variables, including time dummies and other time-varying covariates that may account for heterogeneous dynamics, $\mathbf{z}_i$ is a vector of individual-specific time invariant variables, α_t is an individual fixed effect and u_{it} is an idiosyncratic (i.e. time and individual specific) error term. $Postpat_{it}$ is the *treatment* variable; it takes value 1 in the years after the first patent (including the year of patenting) and 0 elsewhere. In order to assess the robustness of our results, in some specifications we replace the $Postpat_{it}$ variable with two other different variables that may capture a different type of impact of patenting on the academic inventors' publication activity:

1. $Yearpat_{it}$, which takes value 0 in all periods except in the year of patenting when the value is set to 1;
2. $Stockpat_{i,(t-1)}$, which is the individual stock of patents at time $(t-1)$.

In the first case the impact of the patent is assumed to be temporary and contemporaneous to publications; by using it, we test whether, in the year of patenting, scientists experience a productivity slowdown or increase. In the second case, we try to capture the effect that the cumulative patenting effort may have on the scientist's productivity; in particular, we expect that the existence of a substantial complementarity (trade-off) between basic and applied research to show up in a positive (negative) coefficient.

As far as the dependent variable y_{it} is concerned, we used three alternative indicators that describe different aspects of a professor's scientific productivity in any given year: Pub_{it}, which measures the total number of articles published in year t; Cit_{it}, which measures the total number of citations up to 2003 received by articles published in year t; and $Basic_{it}$, which measures the total number of articles in 'basic' scientific journals published in year t.[8]

All the regressions include a full set of year and age dummies. Concerning in particular age, four dummy variables have been used: Age_{30} if age is in the interval $[30,39]$, Age_{40} if age is $[40,49]$, Age_{50} if age is $[50,59]$ and, finally, Age_{60} if age $[60,70]$ (the base age is $[25,29]$).[9] Moreover, since patenting may signal a pre-existing collaboration activity with industry, the variable $Coaut_{i,(t-1)}$, which measures the number of publications that each individual professor has co-authored with researchers affiliated to a private company at time $(t-1)$ has been included in all specifications to control for the individual propensity to collaborate with industry.

Table 5 lists all the variables used in the subsequent regression analysis along with descriptive statistics, while Table 6 provides the correlation matrix for the same variables. Overall, we have an unbalanced panel of 592 individuals (296 academic inventors and 296 controls) from 1980 to 1999. The panel is unbalanced since professors enter and exit from the sample at different times. The total number of individual-year observations is equal to 10,696.

Identifying assumptions and implementation

The key identifying assumption in the estimation of (1) is that the average change in the scientific output of professors is presumed to be the same for both the inventors, if they had not been treated, and their controls. This means that there are no other (time varying) factors that we cannot control for and that affect differently the dynamics of scientific publications of the patenting and

Table 5 Variables used in regression analysis

Variable	*Description*	*Obs*	*Mean*	*Std. dev.*	*Min*	*Max*
Dependent variables						
Pub_{it}	Number of scientific articles published in year *t*	10,696	1.83	2.24	0	25
Cit_{it}	Number of citations received by articles published in year *t*	10,696	26.41	63.39	0	1459
$Basic_{it}$	Number of 'basic' scientific articles published in year *t*	10,696	1.33	1.96	0	25
Treatment variables						
$Postpat_{it}$	Year of the first patent and following years = 1; all other years = 0	10,696	0.26	0.44	0	1
$Yearpat_{it}$	Year of the first patent = 1; all other years = 0	10,696	0.05	0.22	0	1
$Stockpat_{i,(t-1)}$	Stock of patents in year $(t-1)$	10,696	0.57	1.76	0	29
Control variables						
$Coaut_{i,(t-1)}$	Number of scientific articles co-authored with a private company in year $(t-1)$	10,696	0.19	0.63	0	12
Age_{30}	If age = [30, 39] = 1; else = 0	10,696	0.32	0.47	0	1
Age_{40}	If age = [40, 49] = 1; else = 0	10,696	0.32	0.47	0	1
Age_{50}	If age = [50, 59] = 1; else = 0	10,696	0.2	0.4	0	1
Age_{60}	If age = [60, 69] = 1; else = 0	10,696	0.05	0.21	0	1

The table reports the variables used in regression analysis. Dependent variables measure scientific productivity in terms of number of publications, citations received by publications and number of publications in 'basic' scientific journals. Independent variables comprise *treatment* and control variables. Three treatment variables have been considered to capture different impacts of patenting on the publication activity of academic inventors: two dummy variables that capture, respectively, persistent and temporary effects of patenting, and a continuous variable that capture the cumulative effect of patenting. Finally, control variables include year and age dummies, and the number of publications that each individual professor has co-authored with researchers affiliated to a private company. The latter is meant to control for the individual propensity to collaborate with industry.

Table 6 Correlation matrix

	Pub_{it}	Cit_{it}	$Basic_{it}$	$Postpat_{it}$	$Yearpat_{it}$	$Stockpat_{i,(t-1)}$	$Coaut_{i,(t-1)}$	Age_{30}	Age_{40}	Age_{50}
Pub_{it}	1.00									
Cit_{it}	0.54	1.00								
$Basic_{it}$	0.87	0.55	1.00							
$Postpat_{it}$	0.19	0.05	0.12	1.00						
$Yearpat_{it}$	−0.02	−0.01	0.01	0.4	1.00					
$Stockpat_{i,(t-1)}$	0.22	0.08	0.14	0.48	0.38	1.00				
$Coaut_{i,(t-1)}$	0.42	0.32	0.35	0.16	−0.03	0.23	1.00			
Age_{30}	−0.06	−0.03	−0.06	−0.12	−0.002	−0.12	−0.04	1.00		
Age_{40}	0.07	0.06	0.06	0.05	0.02	0.05	0.05	−0.48	1.00	
Age_{50}	0.13	0.04	0.12	0.21	0.02	0.16	0.06	−0.34	−0.33	1.00
Age_{60}	0.03	−0.02	0.03	0.18	−0.002	0.10	−0.01	−0.15	−0.15	−0.11

The table reports the correlation matrix for the dependent and independent variables used in regression analysis.

non-patenting groups of professors. For example if, for any reason different from patenting, an academic inventor activates more research projects or increases her access to research funds or to new equipment relatively to her matched colleague, this would increase her publication activity relative to her control. In this case we would observe a positive impact of the patenting activity even if the patent is not the factor determining the increase in the number of yearly publications. Therefore, we have to assume that time-varying unmeasured factors (e.g. new ideas, new contacts with industry, new research funds, new kids in the scientist family, etc.) should affect both treated and non-treated individuals in similar ways. A violation of this assumption is more likely if academic inventors and controls are very different from each other. By using a matched sample, we make sure that control scientists are not much different from treated professors. In addition, we include in specification (1) the interaction between year dummies and time invariant variables such as professors' date of birth, gender, disciplinary field and size of the department (relative both to other departments in the same discipline in Italy and to other departments in the same university). Conditioning on these variables makes scientists even more similar and reduces the probability of violating the identifying assumption.

A non-random treatment would also violate the identifying assumption. In particular, if the likelihood to have a patent depends upon the previous publication activity this, in turn, would bias the estimated impact of patenting activity on publication activity. We deal with the resulting endogeneity issue at length in the section Treatment of endogeneity.

Results

The effects of patenting on publication activity

Tables 7–9 show the main results of the effects of patenting on the rate and direction of scientific publications. Each table corresponds to one of the three different dependent variables described in the section Regressions. In addition, each table contains estimates for the three different specifications of the treatment variable. Since publications are non-negative integers and the distribution of individual publications is highly skewed, with significant overdispersion and a large number of zeros, a Fixed Effects Negative Binomial model has been used to estimate (1) (Hausman *et al*., 1984). As a robustness check, for each measure of

Table 7 Fixed effects negative binomial and linear regressions of the number of publications (Pub_{it})

	Dependent variable: Pub_{it}				*Dependent variable: log(Pub_{it}+1)*		
	Fixed effects negative binomial				*Within*		*Within+ interactions*
$Yearpat_{it}$	0.04			0.05**			
	(0.035)			(0.02)			
$Postpat_{it}$		0.14***			0.14***		0.14***
		(0.03)			(0.02)		(0.03)
$Stockpat_{i,(t-1)}$			0.06***			0.06***	
			(0.012)			(0.008)	
$Stockpat^2_{i,(t-1)}$			−0.001***			−0.001**	
			(0.0006)			(0.0004)	
$Coauti_{i,(t-1)}$	0.03***	0.03**	0.02*	0.06***	0.05***	0.05***	0.05***
	(0.01)	(0.01)	0.01)	(0.009)	(0.009)	(0.009)	(0.009)
Age_{30}	0.77***	0.76***	0.78***	0.32***	0.32***	0.33***	0.19***
	(0.05)	(0.05)	(0.05)	(0.02)	(0.02)	(0.02)	(0.04)
Age_{40}	0.77***	0.76***	0.78***	0.29***	0.29***	0.29***	0.16***
	(0.07)	(0.07)	(0.07)	(0.04)	(0.03)	(0.03)	(0.05)
Age_{50}	0.69***	0.69***	0.70***	0.21***	0.21***	0.21***	0.18***
	(0.09)	(0.09)	(0.09)	(0.05)	(0.05)	(0.05)	(0.06)
Age_{60}		0.6					
	0.68***	7***	0.14**	0.68***	0.14**	0.14**	0.14*
	(0.12)	(0.12)	(0.12)	(0.07)	(0.07)	(0.07)	(0.08)
# of observations	10,673	10,673	10,673	10,969	10,696	10,696	10,696
# of researchers	590	590	590	592	592	592	592
R^2-within				0.14	0.15	0.15	0.50
F-test				67.77**	69.87***	68.27***	
Degrees of freedom				25, 10,079	25, 10,079	26, 10,078	269, 591
Log-likelihood	−14470.2	−14460.5	−14451.3				
Wald χ^2	1547.4***	1571.6***	1571.6***				
Degrees of freedom	29	29	30				

The table reports results of regression estimates where the dependent variable is the number of publications. Symbols *, **, and *** indicate coefficients statistically significant, respectively at the 10, 5 and 1% levels. Standard errors are in brackets. The first four columns provide estimates of a fixed effects negative binomial model, whereas the last three columns provide results based on linear fixed effects within estimators. The specification in the last column also contains interactions between year dummies and dummies that group individuals according to their dates of birth, gender, disciplinary field, and size of the department. It also includes robust standard errors and the regression is clustered on individuals. Year dummies included in all specifications. The results show that the estimated treatment effect $Postpat_{it}$ is always significantly positive, indicating a positive impact of patenting activity on the number of publications. The size of the estimated coefficient suggests that an academic inventor may expect a 16% higher increase in her publication activity after the first patent relative to their colleagues who do not patent. This result implies rejection of hypothesis H1A and support to hypothesis H2. Different treatment variables lead to similar results. In particular, the impact of the individual stock of patents $Stockpat_{i,(t-1)}$ is positive and statistically significant, suggesting once again the existence of a substantial complementarity between publishing and patenting, also for serial academic inventors (rejection of hypothesis H1B). However, there is also a weak quadratic effect, which suggests that if the stock of patents grow substantially the marginal benefits in terms of publications could eventually decrease.

Table 8 Fixed effects negative binomial and linear regressions of the number of citations (Cit_{it})

	Dependent variable: Cit_{it}				*Dependent variable: $\log(Cit_{it}+1)$*		
	Fixed effects negative binomial				*Within*		*Within +interactions*
$Yearpat_{it}$	0.17***			0.19***			
	(0.047)			(0.06)			
$Postpat_{it}$		0.15***			0.30***		0.31***
		(0.03)			(0.05)		(0.08)
$Stockpat_{i,(t-1)}$			0.06***			0.10***	
			(0.012)			(0.02)	
$Stockpat^2_{i,(t-1)}$			−0.002**			−0.002**	
			(0.0006)			(0.001)	
$Coaut_{i,(t-1)}$	0.13***	0.12**	0.11***	0.10***	0.09***	0.09***	0.10***
	(0.15)	(0.15)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
Age_{30}	0.91***	0.90***	0.91***	0.76***	0.75***	0.77***	0.38***
	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)	(0.09)
Age_{40}	0.90***	0.89***	0.89***	0.68***	0.67***	0.68***	0.35***
	(0.06)	(0.06)	(0.06)	(0.09)	(0.09)	(0.09)	(0.13)
Age_{50}	0.80***	0.79***	0.80***	0.53***	0.52***	0.53***	0.39***
	(0.07)	(0.07)	(0.07)	(0.13)	(0.13)	(0.13)	(0.16)
Age_{60}	0.71***	0.80***	0.70***	0.44**	0.44**	0.44**	0.30**
	(0.09)	(0.09)	(0.09)	(0.18)	(0.18)	(0.18)	(0.2)
# of observations	10,613	10,613	10,613	10,696	10,696	10,696	10,696
# of researchers	587	587	587	592	592	592	592
R^2-within				0.09	0.09	0.09	0.50
F-test				40.22	41.43	39.53	
Degrees of freedom				25, 10,079	25, 10,079	26, 10,078	269, 591
Log-likelihood	−32133.7	−32129.5	−32125.7				
Wald χ^2	1209.7***	1218.29***	1236.39***				
Degrees of freedom	29	29	30				

The table reports results of regression estimates where the dependent variable is the number of citations received by publications. Symbols *, ** and *** indicate coefficients statistically significant, respectively at the 10, 5 and 1% levels. Standard errors are in brackets. The first four columns provide estimates of a fixed effects negative binomial model, whereas the last three columns provide results based on linear fixed effects within estimators. The specification in the last column also contains interactions between year dummies and dummies that group individuals according to their dates of birth, gender, disciplinary field and size of the department. It also includes robust standard errors and the regression is clustered on individuals. Year dummies included in all specifications. The results show that the estimated treatment effect $Postpat_{it}$ is always significantly positive, indicating a positive impact of patenting activity on the number of citations received. This suggests that academic scientists who turn into inventors do not see a decline in the quality of publications as measured by their citation impact (rejection of hypothesis H3A; support to H3B). Similar results are obtained using different treatment variables. In particular, the impacts of the individual stock of patents $Stockpat_{i,(t-1)}$ and of the 1-year dummy $Yearpat_{it}$ are both positive and statistically significant. Moreover, the size of coefficients is slightly larger than the one found in the regression for the number of publications.

scientific productivity, we have also reported linear fixed effects within estimators, using as dependent variable the log of publications, citations and publications in basic scientific journals.

Our results show that the estimated treatment effect $Postpat_{it}$ is always significantly positive. In fact, there is a positive impact of patenting activity on the number of publications (Table 7), on the number of citations received (Table 8) and on the number of publications in basic science (Table 9). This suggests that academic scientists who turn into inventors do not suffer a decline in their scientific output (rejection of hypothesis H1A) and do not see a decline in the citation impact of their publications, nor they divert their efforts towards more applied research (rejection of hypothesis H3A; support to H3B). The size of the estimated coefficient in Table 7 suggests that an academic inventor may expect a 16% higher increase in her publication activity after the first patent relative to their colleagues that do not patent.[10] Estimated coefficients do not change much when using citation or publication counts as dependent variables, although the size of the coefficient is slightly smaller when the dependent variable is the count of publications in basic science journals (Table 9) and slightly higher when we use citations (Table 8). For linear fixed effects within estimators, we have also included a specification which contains the interactions between time dummies and dummies that group individuals according to their dates of birth, gender, disciplinary fields and relative size of the department in the disciplinary field and in the university.[11] Allowing for different time effects across groups of individuals reduces the probability of violating the identifying assumption. This comprehen-

Table 9 Fixed effects negative binomial and linear regressions of the number of publications in 'basic' scientific journals ($Basic_{it}$)

	Dependent variable: $Basic_{it}$				*Dependent variable: $log(Basic_{it}+1)$*		
	Fixed effects negative binomial				*Within*		*Within +interactions*
$Yearpat_{it}$	0.04			0.05**			
	(0.03)			(0.02)			
$Postpat_{it}$		0.10***			0.09***		0.10***
		(0.04)			(0.01)		(0.03)
$Stockpat_{i,(t-1)}$			0.05***			0.04***	
			(0.014)			(0.014)	
$Stockpat^2_{i,(t-1)}$			−0.0002			−0.0003	
			(0.0007)			(0.0003)	
$Coaut_{i,(t-1)}$	0.03**	0.03**	0.02	0.05***	0.05***	0.04***	0.05***
	(0.01)	(0.01)	(0.1)	(0.008)	(0.008)	(0.008)	(0.009)
Age_{30}	0.72***	0.71***	0.73***	0.21***	0.21***	0.21***	0.12***
	(0.05)	(0.05)	(0.05)	(0.02)	(0.02)	(0.02)	(0.03)
Age_{40}	0.71***	0.71***	0.73***	0.19***	0.19***	0.19***	0.12***
	(0.08)	(0.08)	(0.08)	(0.03)	(0.03)	(0.03)	(0.05)
Age_{50}	0.67***	0.67***	0.68***	0.16***	0.16***	0.16***	0.15***
	(0.10)	(0.10)	(0.10)	(0.05)	(0.05)	(0.05)	(0.06)
Age_{60}	0.68***	0.67***	0.69***	0.14**	0.14**	0.14**	0.13**
	(0.13)	(0.13)	(0.13)	(0.06)	(0.06)	(0.06)	(0.07)
# of observations	9338	9338	9338	10696	10696	10696	10696
# of researchers	510	510	510	592	592	592	592
R^2-within				0.09	0.09	0.09	0.57
F-test				40.21	41.32	40.87	
Degrees of freedom				25, 10,079	25, 10,079	26, 10,078	269, 591
Log-likelihood	−11231.8	−11228.4	−11217.0				
Wald χ^2	1004.78***	1031.45***	1041.65***				
Degrees of freedom	29	29	30				

The table reports results of regression estimates where the dependent variable is the number of publications in 'basic' scientific journals. Symbols *, ** and *** indicate coefficients statistically significant, respectively at the 10, 5 and 1% levels. Standard errors are in brackets. The first four columns provide estimates of a fixed effects negative binomial model, whereas the last three columns provide results based on linear fixed effects within estimators. The specification in the last column also contains interactions between year dummies and dummies that group individuals according to their dates of birth, gender, disciplinary field and size of the department. It also includes robust standard errors and the regression is clustered on individuals. Year dummies included in all specifications. The results show that the estimated treatment effect $Postpat_{it}$ is always significantly positive, indicating a positive impact of patenting activity on the number of publications in basic science. This suggests that academic scientists who turn into inventors do not divert their efforts towards more applied research (rejection of hypothesis H3A; support to H3B). This result is broadly confirmed if we use different treatment variables.

sive linear specification includes also robust standard errors. We have also corrected the standard errors using clustered regressions in order to allow residuals to be correlated within each individual block. The estimated treatment effect remains significantly positive with similar size. The goodness of fit is considerably improved.

Different treatment variables lead to similar results. The impact of the individual stock of patents $Stockpat_{i,(t-1)}$ is always positive on all measures of scientific productivity. This result shows once more the existence of a substantial complementarity between publishing and patenting, also for serial academic inventors (rejection of hypothesis H1B). However, there is also a weak quadratic effect, which suggests that if the stock of patents grow substantially the marginal benefits in terms of publications could eventually decrease. The 1-year dummy $Yearpat_{it}$ bears always a positive effect, although this is significant only when citations are used. Therefore, our scientists do not experience a productivity slowdown due to legal or time constraints in the year of patenting.

Our results, however, vary across disciplinary fields. Table 10 reports the outcome of separate regressions for the four disciplines covered by our data and shows that the results we described above hold only for two disciplines, namely electronics and pharmacology. The estimated coefficient of $Postpat_{it}$ is particularly high for electronics, but does not differ significantly from zero for biology and chemical engineering. It is worth noting also that more than one-third of the patents by electronic engineers are owned by ST microelectronics. This is the largest semiconductor company in Italy with a long lasting cooperation record with the University of Pavia (Balconi *et al*., 2004; Breschi *et al*., 2007).

Table 10 Treatment effects in different scientific fields, Fixed effects negative binomial, dependent variable: number of publications (Pub_{it})

	Pharmacology	*Biology*	*Chemical engineering*	*Electronic engineering*
$Postpat_{it}$	0.18***	0.04	0.02	0.45***
	(0.0562)	(0,051)	(0.076)	(0.077)
$Coaut_{i,(t-1)}$	0.001	0.04***	0.00	0.05
	(0.0231)	(0,014)	(0.028)	(0,034)
Age_{30}	0.96***	0.36***	0.72***	0.91***
	(0.089)	(0.080)	(0.120)	(0.103)
Age_{40}	0.89***	0.35***	0.69***	1.01***
	(0.123)	(0.112)	(0.165)	(0.149)
Age_{50}	0.83***	0.25	0.55**	0.91***
	(0.163)	(0.154)	(0.221)	(0.208)
Age_{60}	0.72***	0.40**	0.41	0.88***
	(0.213)	(0.194)	(0.288)	(0.324)
# of observations	2992	2918	2296	2467
#. of researchers	166	156	125	143
Log-likelihood	−4261.0	−4294.0	−2913.5	−2902.9
Wald χ^2	576.9	345.9	246.2	493.3
Model degrees of freedom	26	26	26	26

The table reports the outcome of separate regressions for the four scientific disciplines considered in the paper. The dependent variable is the number of publications and a fixed effects negative binomial model has been estimated. The results show that the positive impact of patenting on publication activity found for the whole sample holds only for two scientific disciplines, namely electronic engineering and pharmacology. The estimated coefficient of $Postpat_{it}$ is particularly high for electronic engineering, but does not differ significantly from zero for biology and chemical engineering. Symbols *, ** and *** indicate coefficients statistically significant, respectively at the 10, 5 and 1% levels. Standard errors are in brackets. Year dummies included.

Treatment of endogeneity

As argued above, a major limitation of the approach followed so far is that it requires us to assume that the treatment is randomly assigned to individuals in our sample. However, this is hard to believe, because the selection into treatment (the decision to patent) is the outcome of a deliberate choice by the scientists, which can be explained by both time variant and invariant variables. In particular, highly productive individuals may sometimes find themselves with enough material at hand to pursue both one or more publications *and* a patent; on the contrary, less productive scientists may have often hardly enough material for a single publication. In this case, patents would occur along with an increase in publication activity, but would not explain the latter. Past works have tried to solve this problem in different ways. Stephan *et al.* (2003) use a cross-section of individual researchers with instrumental variables. Fabrizio and DiMinin (2005) use fixed effect estimations. Azoulay *et al.* (2004) borrow from epidemiology a new class of marginal structural models, which makes use of the inverse probability of treatment weights (IPTW). Goldfarb *et al.* (2006) use a wide array of instrumental variables, among which the amount of venture capital investment in a scientist's research area, to be interpreted as an exogenous shift in the demand for commercial research. Each approach has some drawbacks. Stephan *et al.* (2003) may incur in difficulties with accounting for individual heterogeneity in the cross-sectional setting. Fabrizio and DiMinin (2005), by using fixed effects, eliminate the impact of time invariant individual heterogeneity but do not consider time-varying confounders that could violate the crucial identifying assumption. Goldfarb *et al.* (2006) are forced to use a limited number of observations.

In order to address this issue, we have first of all tested whether and to what extent pre-patent scientific productivity affects differently the post-patent publication activity of academic inventors and controls. To this purpose, we have cross-tabulated the average number of publications of academic inventors and their controls in the 3 years after the inventors' first patent by the number of publications of the same individuals before the first patent (Table 11). The descriptive evidence shows that the post-patent performance of the academic inventors is always higher than their controls. However, the difference between the average post-patent performance of the inventors relatively to their controls does not increase with the pre-patent individual productivity. In other words, pre-patent individual productivity does not seem to be a strong predictor of the average publication gap between academic inventors and controls.

Second, we checked the consistency of our identifying assumption by plotting the yearly average estimated residuals from the regression in Table 7, for the control group and for the academic inventors before their first patent. This has been done for the most comprehensive specification which includes all the interaction terms (last column of Table 7). We expect these residuals to be similar across the two groups. Figure 1 shows that the residuals are indeed similar up to 1993. The difference between the two groups of residuals is particularly evident at the end of the time span covered by the sample; after 1994 few non-treated inventors are left (82 in 1995, 73 in 1996, 52 in 1997, 16 in 1998 and just 1 in 1999). Here, we find that the residuals of

Table 11 Publications of inventors and controls in the 3 years after the first patent by their pre-patent scientific productivity

	Number of publications of individual inventors and controls before the first patent of the inventors				
	0–9	*10–19*	*20–29*	*30–39*	*≥40*
Total number of publications of academic inventors in the 3 years after the first patent (a)	537	474	384	291	578
Number of academic inventors (b)	105	78	42	30	41
Average number of publications (a/b)	5.11	6.08	9.14	9.7	14.1
Total number of publications of controls in the 3 years after the first patent (a)	372	418	197	209	271
Number of controls (b)	123	85	42	24	22
Average number of publications (a/b)	3.02	4.92	4.69	8.71	12.32

The table splits the sample of academic inventors and controls into five categories according to the level of scientific productivity before the first patent of inventors. For example, the first category refers to individuals with a pre-patent number of publications between 0 and 19, while the fifth category refers to individuals with a pre-patent number of publications greater than or equal to 40. For each category of individuals, the average scientific productivity after the first patent is then calculated, separately for academic inventors and their controls. The table shows whether previous scientific productivity affects differently the post-patent publication activity of the academic inventors and their controls. The evidence reported shows that the post-patent performance of the academic inventors is always higher than their controls. However, the difference between the average post-patent performance of the inventors relatively to their controls does not increase with the pre-patent individual productivity. Thus, pre-patent individual productivity does not seem to be a strong predictor of the average publication gap between inventors and controls.

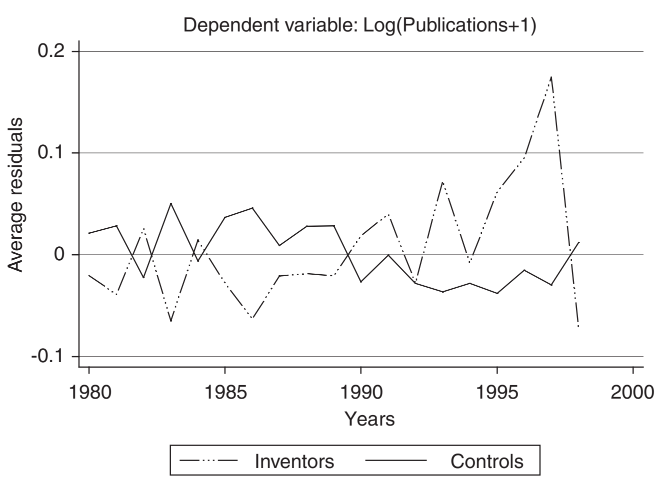

Figure 1 Estimated residuals for academic inventors and controls before the first patent. *Note*: The figure plots the yearly average estimated residuals from a regression where the dependent variable is the number of publications (see last column of Table 7), separately for the control professors and for the academic inventors before their first patent. To the extent that our identifying assumption, that is, that selection into treatment is randomly assigned to individuals in our sample, is correct, we would expect these residuals to be similar across the two groups. The plot shows that the residuals are indeed similar up to 1993. Yet, the plot also reveals that after 1994 there is an increase in the pre-patent residuals of the academic inventors relatively to their controls. Although few non-treated inventors are left after 1994 (82 in 1995, 73 in 1996, 52 in 1997, 16 in 1998 and just 1 in 1999), this result indicates a potential violation of our identifying assumption.

the inventors are regularly above the controls' ones. Table 11 suggests that pre-patent scientific productivity is not a strong predictor of the differences in the post-patent publication gap between the inventors and the control group. At the same time, Figure 1 shows that in recent years there is an increase in the pre-patent residuals of the inventors relatively to their controls. In a related publication, we show that publication counts may affect subsequent patenting and future publication activity (Breschi *et al*., 2005) and similar results are obtained by Azoulay *et al*. (2004). These results suggest that the lagged number of individual publications may act as a time-varying confounder. This problem seems to be more relevant than a possible sample selection bias due to individual productivity.

In order to solve this problem, we follow Azoulay *et al*. (2004) in using the IPTW, a technique first introduced in biostatistics (Hernan *et al*., 2000; Robins *et al*., 2000). A limitation of this methodology is the need to assume that there are no unmeasured confounders. Therefore, in the absence of a unique solution to the issue of endogeneity, we decided to run also a sets of regressions based on instrumental variables and to compare the ensuing results with those obtained by using IPTW.

Instrumental variables

The appreciative theory of individual professors' behaviour outlined in the second section can be used to come up with one or more instrumental variables that may affect the *treatment* choice, but not the *outcome*. The instrument should be a variable that does not belong to model (1) and therefore should not affect the scientific productivity of individual scientists, but should be correlated with the patenting activity (our suspected endogenous explanatory variable); it should also be uncorrelated with the error term (otherwise the instrument would have the same problem as the original predicting variable). In this paper, we used two different instruments: $Shareinv_{i,(t-1)}$, which measures the share of professors at time $(t-1)$ within the department of scientist i that are inventors, and $Stockcomp_{i,(t-1)}$, which captures the stock of patents at time $(t-1)$ of the organization that owns scientist i's first patent. Both variables may affect the individual scientist's propensity to patent and have no direct effect on her scientific publications and citations (other than through patents): colleagues who already have signed one or more patents may set an example or provide the would-be inventor with clues and help when it comes to drafting an application; likewise, applicants with a large stock of patents are more likely to file patents on inventions by the academic scientists they have sponsored, thus affecting the probability that those scientists will turn into academic inventors. These variables are correlated with $Postpat_{it}$ but they are not significantly different from zero when included in specification (1). This suggests that they are suitable instruments. Instrumental variable estimation is implemented through 2SLS and therefore the dependent variable is re-calculated as $k_{it} = \log(1 + y_{it})$, where y_{it} is either the number of scientific articles published at time t or the number of forward citations received by them.[12]

Inverse probability of treatment weights

The IPTW methodology is taken from the clinical and epidemiological studies where researchers are often interested in testing the impact of a treatment at improving survival. It happens often that the treatments are not randomly assigned. The treated group may be sicker, older or poorer than the non-treated group and may behave differently. Therefore differences in survival rates might depend upon different characteristics and behaviours of the groups that confound the effects of the treatment. The IPTW is a method to give a different importance to each observation in order to take into account that individuals have different probabilities of being treated. Intuitively, in our case, weights are built in such a way that scientists with a high predicted probability of patenting receive a lower weight, compared with other scientists with a low predicted probability of patenting. In doing so, the scientist with a low predicted probability of patenting, who actually patents, will have a more important role in the comparison with the group of professors that do not patent. In particular, we suspect that past publications affect the probability to patent and we observe that academic inventors publish more than their controls. Therefore, we want to give more weight to the academic inventors who have a lower probability to patent because they are more 'representative' of the control group.

IPTW methodology is implemented in two stages. First, estimates of the probability of patenting at time t are used to build weights. Second, a weighted fixed effect linear regression is used to estimate specification (1). In the Appendix, we explain in detail how the weights are calculated. Intuitively, IPTW corrects for endogeneity by making use of information on the probability that each observation has to be treated. If the value of $Postpat_{it}$ depends on the individual's past publications it follows that professors who have been more productive in the past have a higher probability to sign a patent at time t, that is to turn into academic inventors. Informally, therefore, the denominator of the weights is built as the probability to receive the observed treatment. Robins *et al.* (2000) show that under the assumption that all the confounding variables are correctly observed, the weights create a pseudo-population that is unconfounded by past publications. We estimate the probability of patenting by using a logit model (see the Appendix for more details):

$$logitprob[Postpat_{it} = 1] = \alpha + \mathbf{x}_{it}\beta + Pub_{i,(t-1)}\gamma + \delta_t + \nu_{it} \tag{2}$$

where the number of individual publications at time $(t-1)$ – that is, $Pub_{i,(t-1)}$ – is the confounding variable that creates the endogeneity problem. The set of other covariates $\mathbf{x}_{it}$ includes: $Stockpub_{i,(t-2)}$, which is the stock of individual publications at time $(t-2)$; $Coaut_{i,(t-1)}$, which is the number of publications with a co-author from industry at time $(t-1)$; $Shareinv_{i,(t-1)}$, which is the share of inventors in the same department at time $(t-1)$; and $Cumpatun_{i,(t-1)}$, which is the stock of patents held by scientist i's university at time $(t-1)$. Moreover, a full set of dummy variables are included: time dummies (δ_t), gender and disciplinary fields. Table 12 shows the results of this auxiliary regression and confirms that individual publications at time $(t-1)$, co-authorship

Table 12 Estimation of the probability of patenting

Dependent variable: $Postpat_{it}$, Logit estimation	
$Pub_{i,(t-1)}$	0.09***
	(0.005)
$Stockpub_{i,(t-2)}$	0.003
	(0.005)
$Coaut_{i,(t-1)}$	0.41***
	(0.09)
$Shareinv_{i,(t-1)}$	0.04***
	(0.008)
$Cumpatuni_{i,(t-1)}$	0.0007
	(0.001)
Intercept	−4.15***
	(0.3)
# of observations	8262
# of researchers	592
−2 Log-likelihood	2416
LR $\chi 2$	122.92***
Model degrees of freedom	27
Pseudo R2	0.05

The table reports the estimates of a logit model for the probability of patenting. The results of this auxiliary regression suggest that the number of publications at time $(t-1)$, that is, $Pub_{i,(t-1)}$, the number of papers co-authored with industry, that is, $Coaut_{i,(t-1)}$, and the share of other professors within the same department that have signed patents, that is, $Shareinv_{i,(t-1)}$, affect in a positive and statistically significant way the probability of patenting. In particular, the fact that the lagged value of the outcome variable, that is, $Pub_{i,(t-1)}$, significantly affects selection into *treatment* seems to confirm our suspects of potential endogeneity problems in regression estimates reported before. Symbols *, ** and *** indicate coefficients statistically significant, respectively at the 10, 5 and 1% levels. Standard errors are in brackets. Dummy variables included (not reported here): time, gender and disciplinary fields.

with industry and the presence of other inventors within the same department affect significantly the probability of becoming a patenter. In particular, it is important to stress that the lagged value of the outcome variable (publications) significantly affects selection into treatment. Accordingly, it is legitimate to suspect that regressions in Tables 7–9 fail to account for the dynamics of the selection process into patenting.

Instrumental variables and IPTW results

Tables 13 and 14 compare the results of the within estimations reported in Tables 7–9 with the results of IPTW and instrumental variables regressions. In both tables the dependent variable is defined as $k_{it} = \log(1 + y_{it})$, where y_{it} is, respectively, the number of scientific publications and the number of citations at time t. The estimated treatment effect ($Postpat_{it}$) is always significantly positive. Different methods do not produce different estimated coefficients for $Postpat_{it}$, both when the dependent variable consists of publication counts and when it consists of citations received. The only exception is regression IV(a), which produces a slightly higher coefficient than the others. Co-authorship with industry appears again to improve significantly the amount of publications and citations received by the academic inventors relative to controls.

Treatment effect varying over time and individual patterns

As a final step in our analysis, we try to open the black box of the estimated positive treatment effect on the basis of the following considerations. It is highly implausible that the effect of the treatment is instantaneous and constant as modelled by the variable $Postpat_{it}$. It is more likely that any positive or negative effect starts some time before the actual date of patenting, and fades away after some years. In fact, the publication rate may start increasing before the first patent is applied for, peak somewhere close to the patenting year, and then display different patterns depending on whether the scientist files more patents or not. Since a single-step variable is not sufficiently flexible to capture these effects, we add in the regression a set of dummies that are meant to capture the variation in the scientific publication in the three years before and after the first patent.[13]

It is also important to take into account that, as shown in Table 2, the distribution of patents across academic inventors is highly skewed. Therefore, we expect the treatment effect to vary with the number of patents signed by the inventor. For example, we expect that the impact on scientific publications could be negligible for inventors who patent once in their life and possibly more pronounced for serial inventors. To take this variation into account, we classified academic inventors in two distinct categories: *occasional* and *serial* inventors. The former are defined as those inventors who have signed just one patent, whereas

Table 13 Instrumental variable (IV) and inverse probability of treatment weights (IPTW) estimations of the number of publications (Dependent variable: log(Pub_{it} +1))

	Within	*IPTW with FE*	$IV^{(a)}$	$IV^{(b)}$
$Postpat_{it}$	0.14***	0.13***	0.17***	0.13***
	(0.02)	(0.02)	(0.06)	(0.05)
$Coautut_{i,(t-1)}$	0.05***	0.05**	0.05***	0.05***
	(0.01)	(0.01)	(0.009)	(0.009)
Age_{30}	0.32***	0.32***	0.32***	0.32***
	(0.02)	(0.02)	(0.02)	(0.02)
Age_{40}	0.28***	0.28***	0.28***	0.29***
	(0.04)	(0.04)	(0.04)	(0.04)
Age_{50}	0.21***	0.20***	0.20***	0.21***
	(0.05)	(0.05)	(0.05)	(0.05)
Age_{60}	0.14***	0.14***	0.14***	0.14***
	(0.07)	(0.07)	(0.07)	(0.07)
# of observations	10,696	10,696	10,696	10,696
# of researchers	592	592	592	592
R^2		0.51		
R^2-within	0.14		0.15	0.15
Wald χ^2			28,044***	28,051***
F-test	69.87***	27.9***		
F-test on excluded instruments			$F_{1,10,078} = 1050.18$***	$F_{2,10,078} = 711.68$***
Model degrees of freedom	(25, 10,079)	(616, 10,079)	25	25

The table compares the results of linear fixed effects within estimates with the results of IPTW and IV regressions where the dependent variable is the (log of) number of publications in year t. The first column replicates results reported in Table 7 (sixth column). The second column refers to IPTW estimates, while the third and fourth columns refer to IV estimates. Two IV estimates have been reported: [(a)] the instrument used is $Shareinv_{i,(t-1)}$, that is the share of professors at time (t−1) within the department of scientist i that are inventors; [(b)] in addition to $Shareinv_{i,(t-1)}$, $Stockcomp_{i,(t-1)}$, that is, the stock of patents at time (t−1) of the organization that owns scientist i's first patent, is also used as an instrument. The results show that the estimated treatment effect $Postpat_{it}$ is always significantly positive and that different estimation methods do not produce significantly different estimated coefficients. The only exception is regression $IV^{(a)}$, which produces a slightly higher coefficient than the others. Overall, these results confirm the beneficial effect of patenting on scientific productivity. Symbols *, ** and *** indicate coefficients statistically significant, respectively at the 10, 5 and 1% levels. Standard errors are in brackets. Year dummies included in all specifications.

Table 14 Instrumental variable (IV) and inverse probability of treatment weights (IPTW) estimations of the number of citations received (dependent variable: log(Cit_{it}+1))

	Within	*IPTW with FE*	*IV*$^{(a)}$	*IV*$^{(b)}$
$Postpat_{it}$	0.30***	0.30***	0.50***	0.30***
	(0.05)	(0.05)	(0.15)	(0.14)
$Coaut_{i,(t-1)}$	0.09***	0.09**	0.08***	0.09***
	(0.02)	(0.02)	(0.03)	(0.02)
Age_{30}	0.75***	0.75***	0.74***	0.75***
	(0.06)	(0.06)	(0.06)	(0.06)
Age_{40}	0.67***	0.67***	0.66***	0.67***
	(0.09)	(0.10)	(0.09)	(0.09)
Age_{50}	0.52***	0.52***	0.52***	0.52***
	(0.13)	(0.14)	(0.13)	(0.13)
Age_{60}	0.44***	0.44***	0.44***	0.44***
	(0.18)	(0.18)	(0.18)	(0.18)
# of observations	10,696	10,696	10,696	10,696
# of researchers	592	592	592	592
R^2		0.51		
R^2-within	0.09		0.09	0.09
Wald χ^2			23694***	23729***
F-test	41.43***	27.9***		
F-test on excluded instruments			$F_{1,10,078}=1050.18$***	$F_{2,10,078}=711.68$***
Model degrees of freedom	(25, 10,079)	(616, 10,079)	25	25

The table compares the results of linear fixed effects within estimates with the results of IPTW and IV regressions where the dependent variable is the (log of) number of citations received by articles published in year *t*. The first column replicates results reported in Table 8 (sixth column). The second column refers to IPTW estimates, while the third and fourth columns refer to IV estimates. Two IV estimates have been reported: [(a)] the instrument used is that is the share of professors at time (t–1) within the department of scientist *i* that are inventors; [(b)] in addition to $Stockcomp_{i,(t-1)}$, that is, the stock of patents at time (t–1) of the organization that owns scientist *i's* first patent, is also used as an instrument. The results show that the estimated treatment effect $Postpat_{it}$ is always significantly positive and that different estimation methods do not produce significantly different estimated coefficients. As found in the previous table, the only exception is represented by regression IV$^{(a)}$, which produces a slightly higher coefficient than the others. Symbols *, **, and *** indicate coefficients statistically significant, respectively at the 10, 5 and 1% levels. Standard errors are in brackets. Year dummies included in all specifications.

the latter have produced a stream of related patents after the first one. As a result we estimate the following specification by means of a fixed effects negative binomial model:

$$y_{it} = \mathbf{x}_{it}\beta + \mathbf{z}_i\theta + D_{i,-3}\gamma_{-3} + D_{i,-2}\gamma_{-2} + D_{i,-1}\gamma_{-1} + D_{i,-0}\gamma_{-0} + D_{i,+1}\gamma_{+1} + +D_{i,+2}\gamma_{+2} + D_{i,+3}\gamma_{+3} + \alpha_i + u_{it} \quad (3)$$

The variables included in the regression are the same as in specification (1). The treatment variable $Postpat_{it}$ is substituted with seven indicator variables where $D_{i,+j}$ ($D_{i,-j}$) is equal to 1 in the jth year after (before) the year of the first patent and 0 elsewhere. Moreover, the effects captured by $D_{i,+j}$ ($D_{i,-j}$) are estimated for the two groups of *serial* and *occasional* inventors. Table 15 illustrates the results from specification (3). In column (1) we notice that the dummy variables are negative until 2 years before the patenting event (even if they cannot be considered significantly different from zero), and positive afterward, with a peak on years, -1, $+1$ and $+3$. These results suggest that the regressions presented in sections The effects of patenting on publication activity and Treatment on endogeneity tend to overestimate academic inventors' publication activity until 2 years before the patenting event, and underestimate it 1 year before, 1 year after and 3 years after the patent. They also suggest the existence of either a publication delay effect, which causes the number of publications to decrease in the year before the patent, and to bounce back afterward, and/or a resource effect, which would contribute to explain the publication increase afterwards. In columns 3(a, b) and 4(a, b) estimations of specification (3), respectively for occasional and serial inventors, are presented. We examine more closely the dummies around the patenting year by interacting them with a dummy for 'occasional' inventors (for serial inventors, we consider only their first patent). The differences are quite relevant and suggest that the nature of the relationship between patenting and publishing is different in the two cases. Occasional inventors have a peak in their publications 1 year before the patent; most likely, their patents are a one-off by-product of a successful research project. On the contrary, serial academic inventors reach their publication peak at a later time, either 1, 2 or 3 years after the first patent. It looks as if the beneficial effect of patenting on publication rates lasted longer in this case, which is entirely consistent with the resource effect explanation (hypothesis H2) and probably associated with a continuous patenting activity over time. We also notice that serial inventors' estimated coefficients are significantly negative 3 years before the patent.

Table 15 Treatment effects varying over time

	Dependent variable: log (Pub_{it}+1) in columns 1, 3(a) and 3(b); Pub_{it} in columns 2, 4(a) and 4(b)					
	Within 1	*Fixed effects negative binomial* 2	*Within* 3(a)	*Fixed effects negative binomial* 4(a)	*Within* 3(b)	*Fixed effects negative binomial* 4(b)
D_{-3}	−0.03	−0.07	−0.004	−0.03	−0.09 +	−0.18+
	(0.03)	(0.06)	(0.04)	(0.07)	(0.06)	(0.10)
D_{-2}	0.06 +	0.06	0.11**	0.10 +	−0.04	−0.06
	(0.03)	(0.05)	(0.04)	(0.06)	(0.05)	(0.09)
D_{-1}	0.10***	0.11*	0.14**	0.16**	−0.009	−0.004
	(0.03)	(0.05)	(0.04)	(0.06)	(0.05)	(0.08)
D_0	0.06**	0.05†	0.08*	0.04	−0.02	−0.02
	(0.02)	(0.04)	(0.04)	(0.06)	(0.05)	(0.08)
D_{+1}	0.10**	0.09*	0.10**	0.09†	0.11*	0.08
	(0.03)	(0.05)	(0.04)	(0.06)	(0.05)	(0.08)
D_{+2}	0.08*	0.06	0.05	0.013	0.14**	0.13+
	(0.03)	(0.05)	(0.04)	(0.06)	(0.05)	(0.07)
D_{+3}	0.09**	0.09+	0.05	0.0003	0.16**	0.20**
	(0.03)	(0.05)	(0.04)	(0.07)	(0.05)	(0.07)
# of researchers	592	590	592	590	592	590
Years	1980–1999	1980–1999	1980–1999	1980–1999	1980–1999	1980–1999
# of observations	10,696	10,673	10,696	10,673	10,696	10,673

The table reports estimates of regressions where the dependent variable is the number of publications in year *t*. Compared to previous estimates, the treatment variable *$Postpat_{it}$* is substituted with seven *pulse* variables, which are meant to capture the variation in the scientific publication in the three years *before* and *after* the first patent. The seven variables are indicated with $D_{i,+j}$ ($D_{i,-j}$): each variable is equal to 1 in the *j*th year after (before) the year of the first patent and 0 elsewhere. Moreover, coefficients in 3(a), 3(b), 4(a) and 4(b) are obtained by interacting the pulse variables with a dummy variable for *occasional* and *serial* inventors. The former are defined as those academic inventors that have signed just one patent, whereas the latter are defined as those academic inventors that have signed a stream of related patents after the first one. Columns 3(a, b) refer to occasional inventors, while columns 4(a, b) refer to serial inventors. The results show that occasional inventors have a peak in their publications 1 year before the patent; most likely, their patents are a one-off by-product of a successful research project. On the contrary, serial academic inventors reach their publication peak at a later time, either 1, 2 or 3 years after the first patent. It looks as if the beneficial effect of patenting on publication rates lasted longer in this case, which is entirely consistent with the resource effect explanation (hypothesis H2) and probably associated with a continuous patenting activity over time. Symbols: **99% significance level; *95%; +90%; †85%. Standard errors are in brackets.

Conclusions

In this paper we have tested the impact of patenting activity on academic inventors' scientific productivity. We have shown that academic inventors become more productive after signing their first patent relatively to a matched sample of professors with no patents. The result holds also with citation-weighted publication data, and for publications on basic science journals. Yet, the positive effect of patenting on scientific productivity largely differs across disciplines, being particularly strong only in pharmacology and electronic engineering. Serial inventors exhibit a stronger and more persistent positive impact of patenting on publishing. These inventors are more likely to be found among those whose patents are held by business companies. Endogeneity problems arising from individual heterogeneity and the potential endogeneity of treatment have been dealt with instrumental variables and IPTW, but do not seem to affect significantly our estimates. We interpret our results as evidence that any possible trade-off between patenting and publishing, due to publication restrictions or an induced bias away from basic research, tends to be outweighed by complementarity effects, due to the increased availability of financial and cognitive resources accruing to scientists working on technologically relevant topics. These results are extremely close to those obtained very recently by Fabrizio and DiMinin (2005) and Azoulay *et al.* (2004) for the US case. Such a coincidence of results for the US and Italy suggests that scientists in the two countries benefit similarly from patenting their research results. Therefore, the well-known differences in university patenting intensity between the two countries may depend not so much on the scientists' characteristics, but on the institutional features of the two academic systems, and on the economic conditions under which academic research is undertaken, including the demand of science from the national industry. Our results also contribute more generally to the debate on the effects of patenting on the progress of science. According to Heller's and Eisenberg's (1998) anti-commons theory, extensive patenting of academic research results may interrupt the cumulative process of science advancement. Many scientists could face difficulties to conduct their research in areas where access to research tools becomes more costly because of extensive patenting and exclusive licensing, or where the danger of infringing patents is very high. If the anti-commons effect were proved to be quantitatively relevant (as suggested by Sampat, 2004; Murray and Stern, 2004), it may turn out to be amplified by the positive feedbacks from patenting to publishing at the individual level that we found in our paper. This is because it is precisely the output of a group of the most productive scientists, the academic inventors, that fellow scientists would find it hard to access. Finally, the evidence pointing at the superior productivity of academic inventors before their first patent, and at the further increase of their productivity afterward, suggests that academic patenting may end up strengthening the Matthew effect in science, as described by Merton (1968): according to it, more productive scientists increase their productivity over time thanks to increasing returns to reputation and visibility. The positive link between patenting and publishing could strengthen that effect for academic inventors. Our future research will address precisely the issue of whether the positive link between publishing and patenting at the individual level may combine with the anti-commons and the Matthew effects, and possibly result in a negative relationship at the systemic level.

Notes

1 In the US and Japan, the publication delay may be mitigated by the so-called 'grace period' rule. However, the European Patent Office does not allow for any grace period, so that any firm or inventor applying for a US or Japanese patent, but foreseeing to extend it to Europe, cannot exploit the rule (Kneller, 2001).

2 The MIUR list includes only those professors and researchers with tenured position. Thus, our data miss fixed-term appointees as well as all the Ph.D. students, post-doc fellows and technicians. In the current Italian system, assistant professor (called 'researcher') and associate professor positions, despite being only the first two steps of the academic career, are not offered as fixed-term appointments, but as tenured ones. The main differences with the position of full professor lie in wage and administrative power.

3 The choice of discipline, rank, and age as matching variables follow the best-established results of quantitative studies in the sociology of science (e.g. Long et al., 1993). In particular, for academic inventors born in between 1950 and 1970, we allowed for no more than 5 years of age difference with the controls. For professors born before 1950 the maximum age difference was 7 years. For academic inventors born after 1970 (just one) the maximum age difference reduced to 3 years. Exceptionally (no more than 10 cases) we matched a full professor (inventor) with an associate professor (control), or an associate professor with an assistant professor; in these cases the age criteria were stricter (maximum age difference: 3 and 5 years, respectively).

4 We decided not to adopt stricter matching rules at the level of university/department (such as choosing controls only from the same departments of the inventors), as they would have greatly reduced the sample.

5 The classification distinguishes between biomedical fields and all the other disciplines. In the first case, the scores correspond to the following definitions of the journals' contents: 1 = clinical observation/2 = clinical observation and investigation/ 3 = clinical investigation/4 = basic biomedical research. In the second case the correspondence is: 1 = applied technology/ 2 = engineering science -technological science/3 = applied research -targeted basic research/4 = basic scientific research.

6 Note also that most Italian professors born before 1960 had no Ph.D. at all: those pursuing an academic career simply served as teaching and research assistants to their B.A. thesis supervisors for a few years, while attending seminars and studying on their own. Thus, we may expect that older professors started publishing earlier than younger ones, who more frequently went on taking a Ph.D. after completing their undergraduate studies. In addition, the Italian system is such that young graduates may hold a research assistant and post-doc position (often as informal jobs) for a long while before securing an academic job such as those listed in our database.

7 Our sample comprises 296 *treated* academic inventors and 296 *non-treated* control professors, for a total of 592 individuals.

8 The year of publication as reported in the ISI-SCI data set has been used to date articles.

9 The use of dummy variables for age is necessary in order to avoid collinearity with the dummy variables representing calendar years. An alternative solution could be that of controlling by the year of birth of the scientists, but this is unfeasible in a DID estimation, the year of birth being a fixed effect that would be cancelled out.

10 Marginal effects are calculated using STATA following Long and Freese (2006).

11 Since we could not interact the time dummies with a dummy for each department and a dummy for each year of birth, we grouped departments in four categories according to the relative size of the department in the disciplinary field (at the national level) and according to its relative size in the university. We also built four equal cohorts using intervals in the years of birth.

12 A 1 is added since for a number of observations the number of publications (citations) is equal to 0 and the log is not defined.

13 These variables are often called *pulse* variables (Laporte and Windmeijer, 2005).

References

Azagra Caro, Joaquin M., Ignazio Fernández de Lucio and Gracia A. Gutiérrez, (2003), "University patents: Output and input indicators… of what?". *Research Evaluation*, **12**(2): 5–16.

Azoulay, Pierre, Waverley Ding and Toby Stuart, (2004), "The effect of academic patenting on (public) research output". Paper presented at the NBER Summer Institute: Academic Science and Entrepreneurship: Dual Engines of Growth?, July 23, Cambridge, MA.

Balconi, Margherita, Stefano Breschi and Francesco Lissoni, (2004), "Networks of inventors and the role of academia: an exploration of Italian patent data". *Research Policy*, **33**: 127–145.

Bekelman, Justine D., Yan Li and Cary P. Gross, (2003), "Scope and impact of financial conflicts of interest in biomedical research. A systematic review". *Journal of the American Medical Association*, **289**: 454–465.

Blumenthal, D., E. G. Campbell, N. Causino and K. S. Louis, (1996), "Participation of life-science faculty in research relationship with industry". *New England Journal of Medicine*, **335**: 1734–1739.

Breschi, Stefano, Francesco Lissoni and F. Montobbio, (2005), "From publishing to patenting: Do productive scientists turn into academic inventors?". *Revue d'Economie Industrielle*, **110**(2): 75–102.

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

Cohen, Wesley M., Richard R. Nelson and John P. Walsh, (2000), "Protecting their intellectual assets: Appropriability conditions and why firms patent or not?". NBER Working Paper 7552.

Cole, Stephen, (1992), *Making Science. Between Nature and Society*. Cambridge, MA: Harvard University Press.

Crespi, Gustavo, Aldo Geuna and Lionel Nesta, (2006), "Labour mobility of academic investors. Career decision and knowledge transfer". Paper presented at the workshop on Technology transfer from universities: A critical appraisal of patents, spin-offs and human mobility, Ecole Polytechnique Federale de Lausanne, September 29–30.

Dasgupta, Paul A. and Partha David, (1994), "Toward a new economics of science". *Research Policy*, **23**: 487–521.

Fabrizio, K. and A. Di Minin, (2005), "Commercializing the laboratory: The relationship between faculty patenting and publishing". Guizueta Paper Series GSM-OM-2005-04, Atlanta, GA.

Goldfarb, Brent, Gerald Marschke and Amy Smith, (2006), "Scholarship and inventive activity in the university: Complements or substitutes?". Paper presented at the workshop on Technology transfer from universities: A critical appraisal of patents, spin-offs and human mobility, Ecole Polytechnique Federale de Lausanne, September 29–30.

Hamilton, Kimberley S., (2003), "Subfield and level classification of journals". CHI No. 2012-R, CHI Research Inc.

Hausman, Jerry A., Bronwyn H. Hall and Zvi Griliches, (1984), "Econometric models for count data with an application to the patents-R&D relationship". *Econometrica*, **52**(4): 909–938.

Heller, Michael A. and Rebecca S. Eisenberg, (1998), "Can patents deter innovation? The anticommons in biomedical research". *Science*, **280**: 698–701.

Hernan, Miguel A., Babette Brumback and James M. Robins, (2000), "Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men". *Epidemiology*, **11**(5): 561–570.

Jensen, Richard and Marie Thursby, (2001), "Patent licensing and the research university". NBER working paper 10758.

Kneller, Robert, (2001), "Technology transfer: A review for biomedical researchers". *Clinical Cancer Research*, **7**: 761–774.

Laporte, Audery and Frank Windmeijer, (2005), "Estimation of panel data models with binary indicators when treatment effects are not constant over time". *Economics Letters*, **88**: 389–396.

Lissoni, Francesco, Patrick Llerena, Maureen McKelvey and Bulat Sanditov, (2008), "Academic patenting in Europe: New evidence from the KEINS database". *Research Evaluation* (forthcoming).

Long, J. Scott, Paul D. Allison and Robert McGinnis, (1993), "Rank advancement in academic careers: Sex differences and the effects of productivity". *American Sociological Review*, **58**(5): 703–722.

Long, J. Scott and Jeremy Freese, (2006), *Regression Models for Categorical Dependent Variables Using Stata*. Houston, TX: Stata Press.

Mansfield, Edwin, (1995), "Academic research underlying industrial innovations: Sources, characteristics, and financing". *Review of Economics and Statistics*, **77**: 55–65.

Mansfield, Edwin, (1998), "Academic research and industrial innovation: An update of empirical findings". *Research Policy*, **26**: 773–776.

Merton, Robert K., (1968), "The Matthew effect in science. The reward and communication systems of science reconsidered". *Science*, **159**(3810): 56–63.

Meyer, Martin, Tatiana Sinilainen and Jan T. Utecht, (2003), "Toward hybrid triple helix indicators: A study of university-related patents and a survey of academic inventors". *Scientometrics*, **58**: 321–350.

Mowery, David C., Richard R. Nelson, Bhaven Sampat and Arvidis Ziedonis, (2004), *Ivory Tower and Industrial Innovation: University-Industry Technology Transfer Before and After the Bayh-Dole Act in the United States*. Stanford: Stanford University Press.

Murray, Fiona and Scott Stern, (2004), "Do formal intellectual property rights hinder the free flow of scientific knowledge?: An empirical test of the anti-commons hypothesis". *Journal of Economic Behavior and Organization*, **63**(4): 648–687.

Owen-Smith, Jason and William W. Powell, (2001), "To patent or not: Faculty decisions and institutional success at technology transfer". *Journal of Technology Transfer*, **26**: 99–114.

Robins, James M., Miguel A. Hernan and Babette Brumback, (2000), "Marginal structural models and causal inference in epidemiology". *Epidemiology*, **11**(5): 550–560.

Rosenberg, Nathan, (1982), *Inside the Black Box*. Cambridge: Cambridge University Press.

Sampat, Bhaven N., (2004), "Genomic patenting by academic researchers: Bad for science?". Mimeo (http://mgt.gatech.edu/news_room/news/2004/reer/.).

Saragossi, Sarina and Bruno van Pottelsberghe de la Potterie, (2003), "What patent data reveal about universities: The case of Belgium". *Journal of Technology Transfer*, **28**: 47–51.

Schmiemann, Manfred and Jean-Noël Durvy, (2003), "New approaches to technology transfer from publicly funded research". *Journal of Technology Transfer*, **28**: 9–15.

Siegel, Donald S., David Waldman and Albert Link, (2003), "Assessing the impact of organizational practices on the relative productivity of university technology transfer offices: An exploratory study". *Research Policy*, **32**: 27–48.

Stephan, Paula E. and Sharon Levin, (1992), *Striking the Mother Lode in Science: The Importance of Age, Place, and Time*. New York: Oxford University Press.

Stephan, Paula E., Shif Gurmu, Albert J. Sumell and Grant Black, (2003), "Patenting and publishing: Substitutes or complements for university faculty?". In R. Ehrenberg, P.A. Stephan (eds.) *Science and the University*. Madison: University of Wisconsin Press.

Thursby, Jerry and Marie Thursby, (2002), "Who is selling the Ivory Tower? Sources of growth in university licensing". *Management Science*, **48**(1): 90–104.

Thursby, Jerry and Marie Thursby, (2007), "Patterns of research and licensing activity of science and engineering faculty". In R. Ehrenberg, P.A. Stephan (eds.) *Science and the University*. Madison: University of Wisconsin Press.

Thursby, Marie, Jerry Thursby and Svasti Gupta-Mukherjee, (2007), "Are there real effects of licensing on academic research? A life cycle view". *Journal of Economic Behavior and Organization*, **63**: 577–598.

John P. Walsh and Wei Hong, (2003), "Secrecy is Increasing in Step with competition". *Nature*, **422**: 801–802.

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

Appendix

Inverse probability of treatment weights

In order to calculate the weights the following procedure has been used (as in Azoulay *et al.*, 2004). A logit equation is estimated for the probability of patenting:

$$logitprob[Postpat_{it} = 1] = \mathbf{x}_{it}\beta + Pub_{i,(t-1)}\gamma + \alpha_i + \delta_t + \nu_{it} \qquad (A.1)$$

The above equation corresponds to the one reported in the text and estimates the probability of patenting as a function of a set of covariates and past publication activity. Let $\hat{p}_{it}$ denote the estimated probability. Under the strong assumption that no unobservable confounding variables exist, the weights can be simply calculated as

$$w_{it} = \frac{1}{\prod_{k=0}^{t}\left(1 - \hat{p}_{ik}\right)} \qquad (A.2)$$

if scientist i did not apply for any patent by year t, and

$$w_{it} = \frac{1}{\prod_{k=0}^{t-1}\left(1 - \hat{p}_{ik}\right)\hat{p}_{ik}} \qquad (A.3)$$

if scientist i applied for his first patent in year t. However, as suggested by Azoulay *et al.* (2004) and Robins *et al.* (2000) to reduce the variance of the IPTW estimator due to outliers, stabilized weights have to be computed. Such weights are defined as follows:

$$sw_{it} = \prod_{k=0}^{t} \frac{Prob(Postpat_{ik} = 1|\mathbf{x}_{ik})}{Prob(Postpat_{ik} = 1|\mathbf{x}_{ik}, Pub_{ik-1})} \qquad (A.4)$$

The denominator of the stabilized weights is based on the estimated probability of specification (4), while the numerator of the stabilized weights is based on the estimated probability of the following specification (i.e. the same as (4) without the time-varying confounders $Pub_{i,(t-1)}$):

$$logitprob[Postpat_{it} = 1] = \mathbf{x}_{it}\beta + \alpha_i + \delta_t + \nu_{it} \qquad (A.5)$$