NBER WORKING PAPER SERIES

THE IMPACT OF ACADEMIC PATENTING ON THE RATE, QUALITY, AND DIRECTION OF (PUBLIC) RESEARCH

Pierre Azoulay
Waverly Ding
Toby Stuart

Working Paper 11917
http://www.nber.org/papers/w11917

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2006

Send all correspondence to pa2009@columbia.edu. We thank Thomas Hellman, Kei Hirano, Scott Stern, Alberto Abadie, and Rajeev Dehejia for useful discussions, as well as seminar audiences at the NBER, Columbia University, the Israel Institute of Technology, INSEAD, UC-Berkeley, the University of Arizona, Stanford GSB, and Harvard Business School. We also thank the Eugene Lang Center for Entrepreneurship at Columbia Business School for financial support. The usual disclaimer applies. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

©2006 by Pierre Azoulay, Waverly Ding, and Toby Stuart. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

The Impact of Academic Patenting on the Rate, Quality, and Direction of (Public) Research Output

Pierre Azoulay, Waverly Ding, and Toby Stuart

NBER Working Paper No. 11917
January 2006, Revised June 2006
JEL No. O31, O32, O34

ABSTRACT

We examine the influence of faculty patenting activity on the rate, quality, and content of public research outputs in a panel dataset spanning the careers of 3,862 academic life scientists. Using inverse probability of treatment weights (IPTW) to account for the dynamics of self-selection into patenting, we find that patenting has a positive effect on the rate of publication of journal articles, but no effect on the quality of these publications. Using several measures of the "patentability" of the content of research papers, we also find that patenters may be shifting their research focus to questions of commercial interest. We conclude that the often-voiced concern that patenting in academe has a nefarious effect on public research output is, at least in its simplest form, misplaced.

Pierre Azoulay
Columbia University
Graduate School of Business
3022 Broadway, Uris Hall 704
New York, NY 10027
and NBER
pa2009@columbia.edu

Waverly Ding
University of California
Haas School of Business
545 Student Services, #1900
Berkeley, CA 94720
wding@haas.berkeley.edu

Toby Stuart
Columbia. University
Graduate School of Business
3022 Broadway, 710 Uris Hall
New York, NY 10027
toby.stuart@columbia.edu

1 Introduction

In the past few decades, universities and other public-sector research organizations have become much more proactive in their efforts to commercialize scientific discoveries (see Henderson et al., 1998; Jaffe and Lerner, 2001; Jensen and Thursby, 2001; Mowery et al., 2001; Thursby and Thursby, 2002). Underlying this well documented upswing in university patenting has been a sharp increase in the number of individual academic scientists who are listed as inventors on patents. As its incidence has increased, however, academic patenting has generated intense controversy, much of which has centered on the long-term effect of patenting on the development of future scientific knowledge.

At this juncture, every available indicator suggests that a growing number of university faculty will become involved in the commercialization of scientific research. As the literature shifts to evaluating the consequences of faculty patenting for the traditional research process, a number of questions will require investigation. In this paper, we focus on two issues that have come to the fore in debates about academic patenting. First, in what direction and to what degree does faculty patenting affect the rate of production of public scientific outputs? Second, does patenting directly influence either the quality or the content of the subsequent-to-the-patent research performed by the scientist?

These questions are important and, we believe, largely unresolved. On one hand, surveys of academic scientists have found that patenting skews scientists' research agendas toward commercial priorities, causes delay in the public dissemination of research findings, and crowds out effort devoted to producing public research (Blumenthal et al., 1996; Campbell et al. 2002; Krimsky, 2003). In stark terms, this work has portrayed a tradeoff between patenting and the progress of academic science. On the other hand, the few studies that have econometrically assessed the scientist-level relationship between patenting and publishing have come to a very different conclusion. Agrawal and Henderson (2002) estimated fixed-effect regressions of the effect of patenting in a 15-year panel of 236 scientists in two MIT departments. They found that patenting did not affect publishing rates. Fabrizio and DiMinin (2005) constructed a sample of 166 academic patenters that were matched to an

equivalent number of non-patenting scientists. In a fixed effects specification, they found a statistically positive effect of researchers' patent stocks on their publication counts. In a third study, Stephan et al. (forthcoming) exploited a survey of doctorate recipients and used instrumental variables to estimate the cross-sectional relationship between patenting and publishing; they found that patenting and publishing relate positively.

Our findings concur with — and significantly extend — this latter set of results. With careful adjustment for selection into patenting, we find that both the flow and the stock of scientists' patents are positively related to subsequent publication rates. Moreover, this increase in output does not come at the expense of the quality of the published research; we find some evidence that the average quality of patenters' publications is higher than that of non-patenters, but this effect is neither especially large in magnitude nor always statistically significant. However, we present three distinct pieces of evidence which indicate that patenting induces a moderate shift in the content of scientists' research. First, faculty holding patents are more likely to coauthor papers with researchers in firms. Second, patenters' publications appear more frequently in journals that have a higher proportion of company-affiliated authors. Finally, we develop a measure of the latent "patentability" of research based on the title keywords of articles and find it to be significantly higher in the subsequent-to-the-patent papers of patenting scientists.

At a minimum, we interpret our results as refuting the simple form of the claim that academic patenting has a deleterious effect on the production of public science. However, our findings do confirm that patenting has had real effects on the direction of scientific progress. Although it is legitimate to ask whether the continued migration of commercial interests into universities will further induce scientists to select research projects on the basis of their perceived value in the private sector, assessing the welfare implications of this change will require a more refined understanding of the relationship between research outputs that are "applied" (*i.e.,* less likely to become an important foundation for subsequent scientific research) versus those that are "patentable" (*i.e.,* focused on questions of industrial usefulness). In the context of the life sciences, for example, it is not a priori clear that there

is a trade-off between the academic influence and the patentability of a research project (see Stokes, 1997).

In addition to presenting findings pertinent to an ongoing policy debate, our study makes two other contributions. First, we have assembled a comprehensive, longitudinal dataset: it is a prospective, 3,862-person random sample drawn from the population of life scientists in academia between 1968 and 1999. For the individuals in the sample, we have reconstituted entire career histories, including patent and publication information, as well as many employer-level variables. We believe that this is the most inclusive dataset available for assessing the relationship between patenting and public research productivity among academic scientists.

Second, we attempt to disentangle correlation from causality in the assessment of the effect of patenting. As we will show, patent holders differ from other researchers on many observable characteristics (see also Stephan et al., forthcoming). More accomplished researchers are much more likely to patent, and controlling for the stock of past publications, scientists with a recent good run are also more likely to patent. This evidence calls into question the ability of traditional fixed-effect specifications to consistently estimate causal effects, since patenters and non-patenters do not appear to follow similar trends in publication rates *before* the initiation of patenting. We use Inverse Probability of Treatment Weighted (IPTW) estimation (Robins et al., 2000; Hernán et al., 2001) to account for the dynamics of self-selection of researchers into patenting. This methodology, which generalizes the propensity score to settings in which treatment is staggered over time, accounts for selection into patenting on the basis of observable characteristics, including (in our case) lagged productivity and the latent patentability of a scientist's research trajectory. While this approach naturally cannot rule out selection based on unobservable factors, we were able to generate an extensive list of covariates to model the probability of selection into patenting.

In addition to these two primary contributions, the paper indirectly relates to the literature on the tension between applied and basic research (Cohen and Levinthal, 1989;

Rosenberg, 1990; Henderson and Cockburn, 1994; Ding, 2005). This group of studies has sought to understand why for-profit firms fund basic research. It has generally concluded that basic and applied research are complements, although several distinct mechanisms are at play, and not all of them correspond to true spillovers of knowledge across activities (Stern, 2004). This work bears an obvious similarity to our effort to assess the nature of the relationship between basic and commercial scientific projects conducted by individual scientists.

The rest of the paper proceeds as follows. In the next section, we provide an overview of the controversies surrounding academic patenting. Section 3 presents our econometric methodology. Section 4 describes the construction of the sample and data sources, presents descriptive statistics, and reports our econometric results. Section 5 concludes.

2 Basic, Applied, and Commercializable Research: Where Do We Stand?

Both the current level and the trend line for academic patenting leave little doubt that the contemporary research university has become a locus of commercially-oriented innovation. However, this development is not without controversy; many observers have decried the emergence of academic patenting and other forms of commercial science for its potentially adverse effects on the further advancement of science (Krimsky, 2003). Among critics' concerns, the most fundamental revolves around the potential effect of academic patenting on the traditional incentives in science. It is commonly acknowledged that the reward system in science is rooted in peers' acknowledgment of important research advances (Merton, 1973; Dasgupta and David, 1994). Scientists' incentives to create and quickly publish research findings are clear when promotions, salary increases, and professional accolades are awarded on the basis of contributions to the corpus of public scientific findings. Seen in this light, the relevant question about university patenting becomes, to what degree does the availability of the option to patent alter the incentive or ability of scientists to contribute public (i.e., non-excludable) advances to the scientific literature?

We are not persuaded that the presence of the option to patent meaningfully decreases scientists' incentives to invest in the production of public science. In fact, for two reasons we foresee the possibility that patenting and publishing are complementary activities. First, an academic researcher's scientific reputation is his/her most important currency in the effort to capitalize on intellectual property in the market for university-originated technology. Second, with respect to the production of new scientific knowledge, there are likely to be (intra-person) scope economies that emerge when a scientist is involved in the development of both academic and commercial science. As we describe below, there are, though, more compelling reasons to believe that the emergence of academic patenting has and will continue to cause a shift in the *content* of scientists' research.

Scientist reputation, patents, and the market for university inventions. Scientists are thought to be strongly motivated by the intrinsic satisfaction of solving vexing problems (Stern, 2004), the search for fame and status (Merton, 1973 [1942]), and the up-or-out promotion rule inherent to the tenure system (Carmichael, 1988). How does patenting influence these traditional incentives to produce academic research? While the direct effect of intellectual property rights is probably small, academic patenting could still influence incentives in subtle ways. In particular, the market for university inventions is rife with asymmetric information. As the literature frequently notes, academic discoveries often require years of additional development to yield marketable products; there is likely to be a great deal of uncertainty surrounding the commercial and scientific merit of discoveries at this primitive stage; and exhaustive due diligence regarding the value of a discovery is costly.

Because of these information problems, we argue that scientists' reputations are essential in the market for university technology. By acting as a signal of invention quality, the prominence of a patenting faculty in the community of science diminishes the search and screening costs that potential licensees must incur in the process of identifying promising university technology. Furthermore, university technology transfer officers are aware of the certification role of scientific eminence. Other things equal, because the discoveries of prominent scientists are more marketable in industry, TTOs are more likely to choose to file for

patents on the discoveries of high-status scientists.[1] Therefore, the ex post search, screening, and contracting problems in the market for ideas increase faculty's ex ante incentives to maintain their reputation on the scientific labor market, as doing so enhances both the odds of finding an industrial match for their inventions, and the value of their patents conditional on a match.[2]

Non-pecuniary and pecuniary spillovers. A likely consequence of applying for a patent is that academic scientists become acquainted with researchers in companies. As these acquaintances develop into relationships, we expect that industry contacts might become sources of ideas for new research projects. The notion that connections with researchers in industry serve as fruitful sources for unearthing interesting research questions emerges in Agrawal and Henderson's (2002) interviews with MIT scientists. In addition, there is a natural analogy to the complementarities observed between applied and basic research in industrial firms. Rosenberg (1998), for example, documented that innovations born out of contact with commercial enterprises in the applied field of chemical engineering ushered a new era of basic discoveries in chemistry. The possibility of within-scientist economies of scope is also consistent with evolutionary theories of technological and scientific progress in which major advances are understood to represent insightful combinations of disparate pieces of knowledge (e.g., Hull, 1988; Weitzman, 1998). Insofar as access to diverse information facilitates the development of new and fruitful lines of scientific inquiry, patenting

[1] Along these lines, Shane and Khurana (2003) showed that startup firms are more likely to be founded to capitalize on university technology if the intellectual property was created by full professors. Elfenbein (2004) found that discoveries made by scientists with extensive publication records were more likely to find a licensing partner.

[2] Some critics would counter that there is an automatic tradeoff between patenting and publishing because it is time consuming to disclose inventions and flesh out patent applications. In addition, crowding out would occur if faculty members devote a substantial block of time to conduct the research that leads to patentable discoveries. However, two facts mitigate these concerns. First, scientists are assisted in the patent application process by their university's technology transfer office, whose existence enables a division of labor between invention and commercialization activities (Hellman, 2005). Second, qualitative evidence suggests that patent applications are often byproducts of traditional scientific efforts, and that patents and scientific articles routinely encode related pieces of knowledge. For example, in her study of tissue engineering, Murray (2002) shows that many scientists choose the path of dual-knowledge disclosure, a practice whose output she labels "paper-patent pairs." We therefore doubt the claim that patenting necessarily requires substantial amounts of time and always crowds out traditional scientific research (also see Thursby et al., 2005).

may facilitate the creation of the professional ties that productively broaden researchers' information networks.

Knowledge is not the only input to the research process that may transcend the university-industry divide; it is also possible to envisage pecuniary spillovers between patenting and publishing. Useful commercial discoveries often lead to industrial sources of funding for the laboratory of the patenting scientist. Even without access to new pools of knowledge, the ability to hire additional post-doctoral scientists or graduate students might result in higher scientific output for a scientist's lab. A related point is that many seminal scientific achievements have been made possible only by technological advances in instrumentation. In the biomedical fields and other areas of science, technological and scientific advances are therefore interdependent: new understandings are often beholden to progress in instrumentation. If patenting scientists are more likely to be in a position to negotiate access to state-of-the-art equipment in corporate laboratories (Owen-Smith and Powell, 2001a), or if they are more likely to have developed the technical expertise to understand and modify research equipment, complementarities between the capital stock of their laboratory and that of their industrial partners might also increase publication output.[3]

Patenting and the *direction* of scientific advance. While we expect to find that patenting has at worst a neutral effect on the rate and quality of scientists' publication output, there is also a case to be made that it will influence the content of the output.

In formulating this argument, it is useful to begin with an over-simplified description of the controversy surrounding the commercialization of university science. Suppose that there are two types of academic scientists: purists, who disapprove of commercial encroachments into the university and select research topics solely on the basis of scientific merit, and commercialists, who participate in university patenting and frequently associate with firms

[3]Note that whether the relevant spillovers are technological or pecuniary, it is not the act of seeking intellectual property rights that, in itself, changes the nature and quantity of output produced by a scientist. Rather, patenting, by making the scientist's research visible to new constituencies, will lead to collaborations (intellectual or financial) that would not have occurred in the absence of the patent application, and between individuals with potentially complementary scientific backgrounds or access to non-overlapping social networks. It should be clear that any spillovers of this type will arise over time, not contemporaneously.

in industry. Scientists in this latter camp investigate two kinds of research questions: like purists, they explore issues of basic scientific relevance. In addition, they allocate some fraction of their time to investigating discoveries with patentable, commercial application. Although this characterization may exaggerate the actual level of difference between purists and commercialists in some institutions, Owen-Smith and Powell (2001b) present qualitative evidence that there is in fact a division along these lines in many academic departments: traditional scientists who, like Nobel Prize winner John Sulston, oppose the convergence of academe and commerce represent the purist pole, and serial patenters and entrepreneurs constitute the other (Sulston, 2003).

If this characterization is approximately accurate, scientists that choose to patent and thereby shift into the commercialist camp will begin to allocate their research time across a different, wider set of research questions than they had done when they were purists. Once a scientist accepts the label of commercialist, we can expect a within-person change such that a scientist will be more likely to pursue projects for which part of the pay-off for conducting the research will be a patent or some other form of commercial recognition. We do not anticipate that all or even a majority of a scientist's work will shift, but rather that some share of it will be focused on new (to the scientist) research questions. Thus, we expect to discover that patenting is associated with a shift in scientists' focus toward exploring scientific questions with commercial application.

A second and possibly more meaningful mechanism for why patenting may result in a shift in scientists' research foci relates to our previous assertion that patents are a form of translational publication that facilitates the formation of relationships between academic scientists and members of the industrial research community. Through the university's efforts to commercialize their technologies, patenting scientists gain visibility in industry circles. As this visibility leads to associations with researchers in corporate laboratories, academic scientists become interested in and familiar with scientific questions of central importance to industry. These contacts presumably expose university scientists to new (relative to their previous work) areas of commercially-useful scientific inquiry. As we have argued above, ceteris paribus, exposure to new and diverse information may enhance scientists'

productivity. In addition, one should expect the nature of the work that arises after a scientist bridges the university-industry divide to be different from what preceded it, if only because academic scientists are likely to become intrigued by questions of interest to industry researchers.[4]

We will remain agnostic regarding the welfare implications of this potential change in research agenda. Among critics of the increasing dependence of universities on private-sector funding, distortion in the choice of topics is a frequently assumed and vigorously lamented consequence. An implicit assumption of these policy discussions is that engagement with the world of commerce will necessarily produce applied research, in the sense that future generations of researchers will be less likely to build on it because of its narrow focus. Increasingly, however, scholars of technological change recognize that ideas might simultaneously have high scientific value and important commercial potential (Stokes, 1997). In the conclusion, we briefly discuss how the present study could be expanded to adjudicate between the optimistic and pessimistic interpretations of the results we present below.

3 Econometric Considerations

Estimating the causal effect of academic patenting on research output must confront a basic selectivity problem: researchers choose whether, when, and how much to patent. As a result, traditional econometric techniques, which assume that exposure to "treatment" occurs randomly, cannot recover causal effects. The standard econometric approach for this type of problem is instrumental variable estimation. Yet, the credibility of IV estimates hinges on the validity of the associated exclusion restriction(s). Unfortunately, academic science is not a setting that provides many (or in fact *any*) sources of exogenous variation in the costs of patenting across researchers and/or universities. For instance, characteristics of the

[4]Reliable evidence of a shift in research priorities is still scant. The most systematic data come from Blumenthal et al. (1986). They surveyed academic life scientists, asking whether respondents had considered commercial potential when choosing research projects. 30% of life science faculty with industry funding replied affirmatively, compared to just 7% of faculty without private sector funding. This correlation suggests that industry funding (often associated with patenting) skews scientists' research agenda, but the causality could just as easily flow in reverse, from researchers' interests to funding sources.

scientist's university (such as the presence of a TTO, or the propensity of scientists to patent in other departments) are certainly correlated with individual scientists' decision to patent, but might also affect their productivity directly. In what follows, we will simply assume that a good instrument is not available.

An second approach is to rely on within-scientist variation to identify the effect of patenting on publication output. Fabrizio and DiMinin (2005) use a fixed effects specification in a panel dataset of matched patenting and non-patenting researchers. In so doing, they purge their estimates from any influence of unobserved heterogeneity that is constant over time. However, it is well-known that for difference-in-differences estimation to be valid, it must be the case that the average outcome for the treated and control groups would have followed parallel paths over time in the absence of treatment. This assumption is implausible if pretreatment characteristics that are thought to be associated with the dynamics of the outcome variable are unbalanced between treatment and control units. Below, we provide strong evidence that selection into patenting is influenced by transitory shocks to scientific opportunities. In this respect, estimating the causal effect of academic patenting on research output presents similar challenges to that of estimating the effect of a job training program on wages. In the job training example, treated individuals have lower earnings on average (relative to their pre-treatment average) in the year immediately preceding enrollment into the program; therefore, the fixed effects estimator is likely to overestimate the treatment effect. Conversely, we will show that patenting scientists have higher output (relative to their average in the pre-patenting regime) in the year immediately preceding their first patent application; as a result, the fixed effect estimator is likely to *underestimate* the effect of patenting on publishing rates.

To overcome these challenges, we make use of a novel approach that has recently gained acceptance in biostatistics: Inverse Probability of Treatment Weighted (IPTW) estimation (Robins et al., 2000; Hernán et al., 2001). These estimators are akin to propensity-score matching techniques (Rosenbaum and Rubin, 1983; Dehejia and Wahba, 2002) in that they make the (untestable) assumption that selection into treatment is based on variables that are observable to the econometrician, but extend it to the case of time-varying treatments.

In particular, IPTW estimation allows one to recover average treatment effects even in the presence of time-varying confounders, i.e., time-varying variables that (1) are correlated with future values of the dependent variable; (2) predict selection into treatment; and (3) are themselves predicted by past treatment history. As we will show below, this applies to the case of academic patenting, since publication rates are strongly auto-correlated, the probability of patenting increases after a recent flurry of publications, and past patenting history influences future publication rates.

Consider a study in which treatment decisions are made in $T+1$ distinct periods $0, 1, \ldots, T$. At each time t, for each individual i, "prognostic factors" Z_{it} and an outcome of interest y_{it} are measured, and a discrete treatment $TREAT_{it} \in \{0, 1\}$ is chosen. For any variable W, denote $\widetilde{W}_{it}$ its history up to time t.

Let $y_{it}^{\widetilde{a}}$ be the value of y that would have been observed at time t had i chosen treatment sequence $\widetilde{a}_{it} = (a_{i0}, a_{i1}, \ldots, a_{it})$ rather than his observed treatment history $\widetilde{TREAT}_{it}$. Note that, even if a_{ik} is dichotomous in each year k, there will be 2^k treatment histories and thus 2^k possible counterfactuals, only one of which is observed for each individual.

By definition, the average treatment effect of treatment history $\widetilde{a}$ on the outcome y is the difference $E[y^{\widetilde{a}}] - E[y^{\widetilde{0}}]$, the average difference between outcomes when following $\widetilde{a}$ and outcomes when never treated. We model the mean of $y^{\widetilde{a}}$ conditional on treatment and exogenous covariates X as:

$$E[y_{it}^{\widetilde{a}} | \widetilde{TREAT}_{it}, X_{it}] \quad = \quad \beta_0 + \beta_1' X_{it} + \beta_2 \Psi(\widetilde{TREAT}_{it}) \tag{1}$$

where $\Psi(.)$ is a "dose-response function." For example, if $\Psi(.)$ puts a weight of 1 on $TREAT_{it}$ in each time-period, then it is the *stock* of patents that influences publishing rates. Conversely, if $\Psi(.)$ puts a weight of 1 on $TREAT_{it}$ and a weight of 0 on $TREAT_{ik}, k = 0, \ldots, t-1$, then only the instantaneous *flow* of patents has a causal effect on outcomes. In the empirical work, we will experiment with various specifications for $\Psi(.)$.

Following Robins (1999b), we introduce the Sequential Conditional Independence Assumption (SCIA), which provides a formal way to extend the assumption of selection on

observables to the case of dynamic treatments:

$$y_{it}^{\widetilde{a}} \amalg TREAT_{it} | TREAT_{i,t-1}, Z_{i,t-1}, X_{it}$$

for all i and t, where the $\amalg$ sign denotes statistical independence. Robins (1999b) shows that under SCIA, the average treatment effect β_2 is identified and can be recovered by estimating

$$y_{it} \quad = \quad \beta_0 + \beta_1^{'} X_{it} + \beta_2 \Psi(\widetilde{TREAT}_{it}) + \varepsilon_{it} \tag{2}$$

by weighted least squares, where the weights correspond to the inverse probability of following actual treatment history $\widetilde{TREAT}_{it}$ up to time t for individual i. Note that (2) differs from (1) in that the observed outcomes y have been substituted for the counterfactual outcomes $y^{\widetilde{a}}$.

Implementing IPTW estimation is relatively straightforward. Under SCIA, the selection bias can be removed by weighting the regression by:

$$w_{it} = \frac{1}{\prod_{k=0}^{t} Prob(TREAT_{ik} | \widetilde{TREAT}_{i,k-1}, \widetilde{Z}_{i,k-1}, \widetilde{X}_{ik})}$$

Each factor in the denominator is the probability that the researcher received her own observed treatment at time k, conditional on past treatment history and her past history of "prognosis factors" for treatment, whether time-varying or fixed over time. Therefore, the denominator of w_{it} represents the conditional probability that an individual followed his or her own history of treatment up to time t. Suppose that all relevant time-varying confounders are observed and included in Z_{it}. Then, weighting by w_{it} effectively creates a pseudo-population in which Z_{it} no longer predicts selection into treatment and the causal association between treatment and outcome is the same as in the original population. We refer to $\widehat{\beta}_2$ when eqn. (1) is weighted by w_{it} as the Inverse Probability of Treatment Weighted (IPTW) estimator of β_2.

At this juncture, it is useful to pause and ask, why, if selection is assumed to depend only on observables, would it be invalid to just include all determinants of selection on the right-hand side of the outcome equation and to proceed with estimation by ordinary least squares? The answer is twofold. First, weighting the outcome equation by the inverse

probability of treatment controls for these factors without making strong functional form assumptions; it can be thought of as regressing outcomes on treatment and a very flexible function of the variables in the selection equation. In the presence of staggered treatments and time-varying confounders, there is another important consideration. Under the usual assumption regarding orthogonality of the regressors to the error term, β_2 can be estimated consistently. However, such an estimate will not correspond to any causal parameter of interest, because the time-varying confounders are themselves affected by past treatment history. In this situation, controlling directly for intermediate outcomes (for instance by including a lagged dependent variable as a regressor) would lead to an underestimate of the magnitude of the treatment effect.

The probabilities $Prob(TREAT_{ik}|\widetilde{TREAT}_{i,k-1}, \widetilde{Z}_{i,k-1}, \widetilde{X}_{ik})$ may vary greatly between subjects when time-varying confounders are strongly associated with treatment. This variability can result in extremely large outlying values for w_{it}. These outliers will contribute heavily to the pseudo-population, and the resulting IPTW estimator will have a very large variance. This problem can be alleviated by replacing w_{it} by a "stabilized" weight sw_{it}:

$$sw_{it} = \prod_{k=0}^{t} \frac{Prob(TREAT_{ik}|\widetilde{TREAT}_{i,k-1}, \widetilde{X}_{ik})}{Prob(TREAT_{ik}|\widetilde{TREAT}_{i,k-1}, \widetilde{Z}_{i,k-1}, \widetilde{X}_{ik})}$$

Although this modification does not influence the consistency of IPTW estimators, it does increase their efficiency (Hernán et al., 2000). Despite its simplicity and intuitiveness, IPTW estimation also presents some significant drawbacks. First and foremost, the assumption of no unobserved confounding is a strong one. Past research in the program evaluation literature has shown that techniques assume selection on observables perform well (in the sense of replicating an experimental benchmark) when (1) researchers use a rich list of covariates to model the probability of treatment; (2) units are drawn from similar labor markets, and (3) outcomes are measured in the same way for both treatment and control groups (Dehejia and Waba, 2002; Smith and Todd, 2005). All of these conditions would appear to be met in our setting and data, but this should not lead researchers to believe that IPTW estimation represents a universal solution for endogeneity problems. A second limitation is that IPTW estimates are just identified: the assumption of no unobserved determinants

of selection into treatment cannot be tested; neither can misspecification of the selection equation used to estimate the weights. Third, the causal effect estimated by IPTW models is the population average treatment effect (ATE). In social science applications, however, the average treatment effect on the treated (ATET) might also be of policy interest. However, Lechner and Miquel (2005) show that ATET is not identified without further assumptions on the joint distribution of the counterfactual outcomes.[5]

Informative censoring. Although we focused the first part of the discussion on the problem of non-random selection into patenting, a second problem arises because some subjects might exit the sample for endogenous reasons. For instance, scientists might leave academia because of low scientific productivity, or because they receive attractive offers to join commercial firms. Even if treatment was randomly allocated across units, this type of informative censoring could jeopardize the validity of the statistical estimates. We deal with this problem by treating censoring as just another time-varying treatment. As Robins et al. (2000) note, from this point of view, adjusting for censoring is only to say that our interest lies in estimating the causal effect of $TREAT$ on y if, contrary to the fact, all subjects had remained in the sample rather than having followed their censoring history. We model the exit decision as a function of constant and time-varying observable factors, and compute weights corresponding to the probability of exit given these observables:

$$sw^*_{it} = \prod_{k=0}^{t} \frac{Prob(EXIT_{ik}|\widetilde{TREAT}_{i,k-1}, X_{ik})}{Prob(EXIT_{ik}|\widetilde{TREAT}_{i,k-1}, \widetilde{Z}_{i,k-1}, X_{ik})}$$

sw^*_{it} is the inverse of the ratio of a scientist's probability of exiting academia up to year t divided by that probability calculated as if there had been no time-dependent determinants of censoring except past treatment history and X. Hernán et al. (2001) shows that consistent estimates for β_2 can be obtained by combining the weight corresponding to the inverse probability of treatment sw_{it} and the weight corresponding to the inverse probability of censoring sw^*_{it}. The denominator of the final weight, $sw^*_{it} \times sw_{it}$, is the probability that a

[5] One might worry about performing statistical inference using "second stage" IPTW estimates, since the weights that are used as input in the outcome equation are themselves estimated. In contrast to two-step selection correction methods, Wooldridge (2002) has shown that the standard errors obtained in this case are conservative.

subject would have followed his own treatment and censoring history up to year t, conditional on observables. As a result, we label this methodology *Inverse Probability of Treatment and Censoring Weighted* (IPTCW) estimation in the rest of the paper.

Estimation of the weights. The procedure followed to compute the weights depends on the way in which treatment is defined. According to a first definition, treatment is a *flow*: $TREAT_{it} = 1$ whenever researcher i applies for at least one patent in year t, and 0 otherwise. This formulation implies that treatment does not necessarily have a lasting impact on the individual. In contrast, the *regime* formulation defines $TREAT_{it} = 1$ for all years subsequent to the first patent application. Defining treatment this way implies a one-time shift on the outcome of interest, with subsequent treatment episodes having no effect on the dependent variable.

In the flow formulation, the weights are computed by estimating pooled cross-sectional logit specifications on the whole dataset. To compute the denominator of sw_{it}, one estimates a logit model for:

$$Prob(TREAT_{it} = 1) = \alpha_0 + \alpha_1 TREAT_{i,t-1} + \Phi(\widetilde{Z}_{i,t-1}, \alpha_2) + \alpha_3 X_{it} + \delta_t \tag{3}$$

where $\Phi(\widetilde{Z}_{i,t-1}, \alpha_2)$ corresponds to a parametric function of past values for time-varying confounders, X_{it} includes both time-varying and fixed-over-time characteristics of individuals in the sample (such as years of experience, gender, characteristics of the Ph.D-granting institution, etc.), and δ_t represents calendar year effects. In practice, we specify Φ as a linear function of publication flow in year $t-1$ and stock of publications up to year $t-2$, the number of past collaborations with industrial firms, patentability of the scientist's flow of publication in year $t-1$ and its stock up to year $t-2$, and employer characteristics. Let T_1 denote the set of years in which scientist i gets at least a patent and T_2 the set of years during which i gets no patents. The estimate of the denominator of sw_{it} is $\prod_{t \in T_1} \widehat{p}_{it} \prod_{t \in T_2} (1 - \widehat{p}_{it})$, where $\widehat{p}_{it}$ refers to the predicted probability obtained after estimating eqn. (3). The numerator of sw_{it} stems from an almost identical specification, except that one omits the term $\Phi(\widetilde{Z}_{i,t-1}, \alpha_2)$ from the list of covariates.

This approach needs to be slightly modified when treatment is modeled as a regime shift rather than as a flow, because the probability of getting treatment remains constant and equal to one once a scientist enters the patenting regime. As a result, it is only necessary to fit the model on a subset of the data, that of scientist-year observations up to the year when the scientist applies for his/her first patent. In this risk set, $TREAT_{i,t-1}$ is uniformly 0. To compute the denominator of sw_{it} we estimate a logit model for

$$Prob(TREAT_{it} = 1) = \alpha_0 + \Phi(\widetilde{Z}_{i,t-1}, \alpha_2) + \alpha_3 X_{it} + \delta_t \tag{4}$$

and to compute the numerator of sw_{it} we estimate a logit model for

$$Prob(TREAT_{it} = 1) = \alpha_0 + \alpha_3 X_{it} + \delta_t \tag{5}$$

Our estimate of the denominator of sw_{it} for scientist i in year t is $\prod_{k=0}^{t}(1 - \widehat{p}_{ik})$ if scientist i did not apply for at least one patent by year t, and $\prod_{k=0}^{t-1}(1 - \widehat{p}_{ik}) \times \widehat{p}_{it}$ if scientist i applied for his first patent in year t. Estimation of sw^*_{it} proceeds in the same fashion.

Relationship of IPTCW estimation with propensity-score matching methods. Rosenbaum and Rubin (1983) refer to $Prob(TREAT_i = 1|X, Z)$ as the propensity score, and show how to use this probability to estimate treatment effects when selection into treatment depends only on observables. Recently, Heckman et al. (1997) have combined the propensity score with difference-in-differences to estimate the causal effect of undergoing a job training program. Abadie (2005) proposes a semiparametric difference-in-differences estimate that weights observations by the inverse probability of (own) treatment. Although the goals of these earlier papers resemble ours, we follow a different approach because the structure of our data differs significantly from the typical program evaluation setup. Labor econometricians generally study programs for which a "before" and "after" period can be unambiguously delineated for both treated and untreated units. In contrast, in our setting and many others, selection into treatment can occur at different times and/or in several disjoint episodes. Matching on the propensity score is difficult in these cases. Intuitively, an untreated individual might be a good control for a treated subject in one period (in the sense that the difference in their propensity scores is close to 0) and a bad control for the

same treated subject in another period. In contrast, IPTCW estimation readily generalizes to the case of treatments that are staggered over time.

Robustness to unobserved heterogeneity. If our selection model does not capture some relevant determinants of the patenting decision, and these omitted factors influence research output directly, IPTCW estimates will be biased. Since there is overwhelming evidence of positive selection in the cross-sectional dimension of the data ("better" scientists are both more likely to patent and publish heavily), such unobserved heterogeneity likely leads us to *overestimate* the treatment effect. Therefore, neither fixed effects nor the IPTCW approach provides a fully satisfactory solution to the problem of estimating the causal effect of patenting on publishing rates. In combination, however, these estimators implicitly define a confidence interval, with the fixed effects estimate providing a lower bound, and the IPTCW estimate providing an upper bound. The evidence presented below will show that, in practice, these bounds are sufficiently tight to inform the policy debate surrounding academic patenting.

4 Data and Sample Characteristics

We examine the association between patenting and publishing in a panel dataset of academic life scientists employed at universities and non-profit research institutes. This area was chosen because the biomedical fields have accounted for the preponderance of university patenting and licensing activity (Mowery et al., 2001). While we have not selected scientists because they have patented, we have sampled from scientific disciplines that we know to have significantly contributed to a vibrant area of technological development. We began by drawing 12,000 doctoral degree recipients from UMI Proquest Dissertations, which lists Ph.D. recipients from more than one thousand universities. In forming the sample, we randomly selected individuals, but only those with Ph.D.s in scientific disciplines that have

informed commercial biotechnology.[6] This assures a random sample of Ph.D.s in areas in which academic research may have significant, short-term commercial value.

Given our focus on the life sciences, one might question whether our results generalize to other academic fields, such as mechanical or electrical engineering. One should note, however, that our definition of life sciences is expansive. For example, our data include scientists holding Ph.D's in chemistry, chemical engineering, materials engineering, plant biology, veterinary sciences, and food science. The life sciences, broadly construed, represent such a large slice of the academic patenting phenomenon that the issue of generalizability does not loom particularly large.[7]

Next, we obtained scientists' publication records from the ISI's *Web of Science* database. Because the *Web of Science* includes authors' affiliations, we were able to identify Ph.D. graduates who pursued careers outside of academe. After removing individuals that (i) had no publications in any post-graduate year, (ii) published exclusively under corporate affiliations, or (iii) exited academe early in their careers,[8] we were left with 3,862 scientists, all of whom we know to have been employed at U.S. universities or public research institutions. Each scientist is observed from the year after he or she earned a Ph.D. until 1999, unless the individual exited academia.[9] The final panel contains 58,562 person-year observations between 1968 and 1999.

[6]To identify the scientific disciplines that have been most important to biotechnology, we coded the educational backgrounds of the Ph.D.-holding, university-employed scientific advisory board members of all publicly traded biotechnology firms. The source of information on scientific advisors' degrees was the IPO prospectuses of the 533 U.S.-based biotechnology firms that were filed with the U.S. Securities and Exchange Committee. We then stratified the random draw from UMI to correspond to the disciplines and Ph.D. years of firms' scientific advisors. For example, 22 percent of biotechnology company scientific advisors hold biochemistry Ph.D.s; we drew a corresponding proportion of biochemists into our sample. Table 1 lists the Top 15 disciplines from which scientists in our sample are selected.

[7]In a related paper, one of the authors assembled a dataset of "superstar" academic patenters, which were defined to be US-based academics with more than 17 patents between 1976 and 2004 (this corresponds to scientists above the 99th percentile of the patent count distribution). Among the 544 such scientists, he found only 138 (25.37%) that did not fit our definition of "life scientists."

[8]Ph.D.s with academic affiliations lasting less than five years were dropped from the dataset to exclude post-doctoral fellows that later moved to jobs in industry.

[9]We assume a researcher has exited academia when he or she fails to publish for five consecutive years, or in fewer instances, when the scientist begins to publish almost exclusively under a corporate affiliation. In either case, we censor observation in the year in which a scientist last publishes under a university affiliation.

4.1 Variables

A brief description of the patenting process in academia is useful to interpret the results we will present. The process begins when a faculty member discloses an invention to the university's Technology Transfer Office (TTO).[10] The commercial potential of this invention is then evaluated by the TTO, which may decide to seek patent rights on the invention. Concurrently, the TTO will market the innovation to potential licensing partners in industry. A typical licensing agreement specifies a 40% royalty rate for the individual faculty inventor, to be assessed on the gross licensing revenues the invention accrues.

Research outputs. From the *Web of Science* we computed annual paper publication counts for each scientist. We count all papers on which a scientist is listed as an author (in other words, we treat sole authored and coauthored papers as equivalents). Second, we used the affiliation data available in the *Web of Science* to identify all instances in which a scientist wrote a paper that was coauthored with one or more individuals in a corporate research and development lab. We consider the rate of publication of papers with coauthors in industry as an indicator of the degree to which scientists are engaging in commercially-oriented research. We also keep track, for each journal in which our scientists published, of the relative prevalence of authors with corporate affiliations.[11] In particular, for each scientist and in each year, we compute, following Lim (2004), an average Journal Commercial Score (JCS) by weighting each publication by the proportion of corporate authors who publish in the corresponding journal, summing the weights corresponding to all the articles published by the scientist in a given year, and dividing this sum by the (unweighted) number of articles he/she published during the year.

[10] Faculty members are contractually obligated to disclose potentially commercializable discoveries developed on university premises to the TTO; they do not have the option to patent university-originated discoveries without going through the official channels. On average, TTO received 78 invention disclosures in 2003, but filed only 40 new patent applications (AUTM, 2003). Of course, these numbers vary widely across institutions depending on whether involvement with the world of commerce corresponds to a well-established culture within the institution.

[11] For example, 35.7% of the affiliations for the authors publishing articles in the *Journal of Medicinal Chemistry* correspond to corporations. In contrast, the number is only 1.60% for the *Journal of General Physiology*.

We use a two-pronged approach to measure the quality of the articles published. First, we make use of the order of authorship, computing the proportion of articles in which the scientist appears in first or last position. This choice is motivated by a robust social norm in the life sciences which systematically assigns last authorship to the principal investigator (generally the head of the laboratory), first authorship to the junior author who was responsible for the actual conduct of the investigation, and apportions the remaining credit to authors in the middle of the authorship list, generally as a decreasing function of the distance from the extremities of the list. In the second approach, we make use of the Journal Citation Reports, published yearly by the Institute for Scientific Information. ISI ranks journals by impact factor (JIF) in different scientific fields. The impact factor is a measure of the frequency with which the "average article" in a journal has been cited in a particular year. We weight each article published by the scientists in our sample by the corresponding journal's JIF, sum these weights for all the published output in a given year, and divide by the yearly publication count. The resulting variable can be thought of as a measure of quality for the average article published by one of our scientists in a given year.[12]

Patents. The patents of the academic scientists in our data were assembled from the NBER patent database (Hall et al., 2001). To identify academic patenters, we matched the scientists in our dataset to the list of inventors in the NBER patent database. Matches were done on the basis of first and last names, and we used information on assignee (university) and geographic region to eliminate false matches. For each scientist in our data, we generated flow and stock measures of patent applications, as well as corresponding dummy variables.

Control variables. Following a number of studies of the determinants of scientists' productivity, we were also able to construct a rich set of control variables to account for individual

[12]Basically a ratio between citations and recent citable items published, JIFs suffer from built-in biases: they tend to discount the advantage of large journals over small ones, of frequently-issued journals over less frequently-issued ones, and of older journals over newer ones. Nonetheless, they convey quite effectively the idea that the *New England Journal of Medicine* (Impact Factor = 23.223 in 1991) is a much more influential publication than the *Journal of General Internal Medicine* (Impact Factor = 1.056 in 1991). In an ideal world, rather than assigning an identical weight to all publications appearing in a given journal, we would instead weight each publication by the number of citations it garnered from other scientists. At the present time, querying the *Web of Science* to collect this information is prohibitively time-consuming since this database does not provide time-varying citation data.

and institutional attributes that may influence rates of publication and patenting. To account for life-cycle effects (Levin and Stephan, 1991), we include the number of years since a scientist earned his or her Ph.D. An extensive literature in the sociology of science has documented gender differences in productivity (e.g., Long and Fox, 1995), so we include a "scientist is female" dummy variable. Because the time involved in publishing scientific research varies across fields, the regressions include a set of dummies for researchers' dissertation subject areas. Some of the regressions control for quality differences among researchers through the inclusion of scientist fixed effects. In specifications without fixed effects, we enter a dichotomous measure of the quality of a scientists' Ph.D.-degree granting institution — a dummy variable indicating whether or not a scientists' doctoral program was ranked in the Top 20. Specifically, we collected Gourman Report rankings for all institutions in our dataset. Gourman rankings for graduate schools were issued for the first time in 1980. We assigned universities their original rating for all years prior to 1980 and updated them every other year for the subsequent period. We also included in the models the stock of patents issued to the Ph.D-granting institution in the five years preceding the doctorate, to further control for the "imprinting" of norms regarding commercial activities during graduate training.

From previous research, we know that institutional context has an effect on the propensity to commercialize research, either in the form of a well-funded technology licensing office, or through the presence of prominent peers who themselves are engaged in this activity (Di Gregorio and Shane 2003; Lach and Schankerman 2004; Stuart and Ding, 2006). As a result, we also include in our models a number of employer-level variables. These covariates are updated each year and when scientists change employers. First, we include a quality rank dummy variable analogous to the one constructed for Ph.D.-granting institutions. There are a variety of reasons why scientists at prominent universities are likely to be more productive, including the availability of more resources and easy access to high quality colleagues. Second, we used the AUTM surveys to create a technology transfer office (TTO) dummy variable, which is set to one in all years in which a scientist's employing university has an

active TTO. Third, a university's stock of patents is entered in the model, among other things to further control for institutional differences in support for patenting.

Patentability. In the regressions for selection into patenting used to construct the inverse probability of treatment weights, it would obviously be desirable to account for differences among scientists in the inherent "patentability" of their research. In past studies, latent patentability was thought to be unobservable, and researchers used field fixed effects as controls in order to hold constant individual scientists' research agendum. In contrast, we attempt to measure patentability directly. To construct such a measure, we use the title words in scientists' publications to identify the areas in which they have conducted research, and then apply weights to theses areas based on an (endogenous-to-the-sample) measure of the extent to which other scientists working in these areas have patented their discoveries. Intuitively, we use the publications of scientists that have already applied for patent rights as the benchmark for patentable research, and then compare the research of each scientist in our dataset to this benchmark to generate a research patentability score for each scientist-year. Specifically, the research patentability score for scientist i in year t is defined as:

$$PATENTABILITY_{it} = \sum_{j=1}^{J} w^{i}_{j,t-1} \frac{n_{ijt}}{\sum_k n_{ikt}}$$

where $j = 1, \ldots, J$ indexes each of the scientific keywords appearing in the titles of the journal articles published by scientist i in year t,[13] n_{ijt} is the number of times each of the keywords j has appeared in scientist i's articles published in year t, and w^{i}_{jt} is a weight for each keyword that measures the frequency with which word j is used in the titles of articles published by scientists who have entered the patenting regime in year t or earlier, relative to those who have not entered the patenting regime as of year t (the calculation of w^{i}_{jt} is detailed in Appendix I). Intuitively, the patentability of a scientist's research can change because of a change in the direction of the research of that scientist, or because other patenters' research increasingly comes to resemble that of the scientist. The former effect is captured by the

[13] We relied on title words in journal articles instead of journal- or author-assigned keywords because the *Web of Science* database did not begin to include keyword descriptors until 1992. However, the titles of biomedical research papers typically indicate the research area and the methodology used in the paper. We find high overlap between title words and keywords in the papers for which both are available.

ratio $\frac{n_{ijt}}{\sum_k n_{ikt}}$, the latter by the weights $w^i_{j,t-1}$. Because the benchmark in year $t-1$ is used to weight title words in year t, year-to-year changes in the research patentability score will only reflect actions of the scientist (through their choices of title keywords), rather than contemporaneous changes in the benchmark.[14]

Finally, to capture the idea that the inherent patentability of past research might still influence the current propensity to patent, we compute a depreciated stock of the research patentability score using a perpetual inventory model. Through the impact of the depreciation rate δ, this formulation captures the fact that the recent substantive research orientation of a scientist's research should influence current behavior more strongly than scientific trajectories that unfolded in the more distant past:[15]

$$STOCK_RP_{it} = (1-\delta)STOCK_RP_{i,t-1} + FLOW_RP_{it} = \sum_{\tau=0}^{t}(1-\delta)^{t-\tau} \cdot FLOW_RP_{i\tau}$$

4.2 Descriptive Statistics

Out of a population of 3,862 scientists, we found 473 (12.2%) patenters who were listed on 1,372 patents. Out of these patents, 342 were assigned to corporate entities (of which 31 were co-assigned to a university and a corporation), even though the inventors of interest were academically affiliated based on information revealed in other patent applications filed by the inventor or in publication records. Most of these corporate patents have multiple inventors and a university scientist could be listed as one of the inventors for his advice during the process of invention. A typical example is Richard J. Lagow, who obtained a Ph.D. in inorganic chemistry from Rice University in 1970 and subsequently held professorships at MIT and the University of Texas Austin. Lagow began patenting in 1973 and his patents have been assigned to MIT, University of Texas, and Exfluor Research Corporation. Among

[14]Previous researchers have developed other measures of proximity in technological space. For instance, Jaffe (1986) used a cosine-based measure to assess the proximity between the R&D portfolio of any given pair of firms. While this approach works well for measuring technological distance between dyads, it is not well suited to our setting, since we need to measure the distance between the scientific trajectory of any given scientist relative to that of a benchmark *group* of (patenting) scientists.

[15]We set δ equal to 0.15 — the Griliches constant — which has been used by many innovation researchers on whose work this paper builds. We verified that our core results are not sensitive to this arbitrary choice.

the 31 patents for which Exfluor is the assignee and Lagow is an inventor, 28 involved multiple inventors and 3 listed Lagow as the sole inventor. Based on the data sources available to us, it is not possible to determine the exact role of Lagow in developing these inventions and what type of arrangement Lagow has with University of Texas, but from the titles and abstracts of the Exfluor patents it is clear that the patented inventions are based on knowledge closely related to Lagow's research. Therefore, our data suggests that a non-trivial portion of faculty patenting activity may occur without the official involvement of their employing university's technology transfer office.

In Figure 1, we plot the distribution of patents for the patenting researchers in our sample. The histogram illustrates a rapid drop off after one — most patenters are listed on 1 or 2 patents throughout their career, and very few scientists in our data receive more than 10 patents. Figure 2 displays the distribution of scientists' total publication counts by the end of our observation period, broken out by their patenting status. Consistent with the conventional wisdom that patenting is concentrated among the group of academically productive scientists, the distribution for the patenter subsample is much less skewed than that for the non-patenter subsample.

Table 2 presents the summary descriptive statistics for variables used in our analysis. Table 3 reports, by scientists' patenting status, the mean research and employer characteristics measured at five career stages. Researchers who have sought and received patent rights for their discoveries are more productive at each career stage: they publish more research papers as those who have not yet entered the patenting regime, and those papers appear to be of marginally higher quality (as captured by average JIF). Scientists who have applied for patent rights are closer to commercial research than their non-patenting counterparts, especially at the beginning of their career; they collaborate more often with researchers in the private sector and the intrinsic patentability of their research appears higher. However, these differences vanish at later career stages. Finally, patenters are more likely to work in settings where a technology transfer office exists and patenting activity is intensive. Of course, these univariate comparisons are subject to "static" omitted variable bias in addition to the dynamic selection bias mentioned in section 3.

4.3 Results

We present four sets of results. Table 4 focuses on the antecedents of selection into patenting, and on the determinants of exit from academia. It provides evidence on the importance of time-varying confounding, and displays the specifications from which our probability of treatment and censoring weights are derived. Using these weights as inputs, the following tables present results pertaining to the effect of patenting on the rate (Table 5), quality (Table 6), and direction (Tables 7) of scientific output.

Determinants of patenting activity. We begin by presenting results pertaining to the probability of applying for a patent in a given year (flow formulation) or for the first time (regime formulation). The results are displayed in Table 4. It is important to note that the list of independent variables and the risk set differ significantly across the flow and regime models. In the former, all scientist-year observations are included, and the list of independent variables includes a lag structure for patenting in order to address the possibility of structural state dependence. In the latter, the observations corresponding to years subsequent to the year of the first patent application are not part of the risk set; consequently, no lag structure for the dependent variable can be part of the set of right-hand side variables.

The econometric analysis confirms that time-varying confounders are important determinants of patenting activity for these scientists. First, controlling for the stock of publications up to year $t-2$, the probability of patenting in year t is significantly increasing in the flow of publications in year $t-1$: at the mean of the data, a standard deviation increase in the flow of lagged publications increases the probability of patenting by 10.40% for the flow specification (column 1a) and by 20.3% for the regime specification (column 2a).[16]

This conditional correlation strikes us as being an important finding, for it can help distinguish between competing interpretations of the association between scientific productivity and involvement with the world of commerce. In the first interpretation, commercialization activities correspond to attempts by academics to monetize established reputations and

[16] In a companion paper (Azoulay et al., 2005), we confirm that this result is robust to much more flexible specifications of the lag structure.

professional status. In the second interpretation, publications and patents are co-occuring outputs that encode the same set of scientific insights; patents, just like publications, reflect genuine shocks to scientific opportunities. We see the correlation between the onset of patenting and the lagged flow, but not the stock, of publications as much more consistent with the latter interpretation.[17] The plausibility of this interpretation is reinforced by a peculiar aspect of US patent law, which grants inventors a one-year grace period from the date of publication for the filing of a patent application (Merges, 1997, p. 226). In other words, an academic inventor wishing to maximize the effective life of a patent would apply to the USPTO exactly 364 days after the date of publication, provided that he/she is willing to forego patent protection in foreign jurisdictions.[18]

We also find that previous ties to industry in the form of coauthorships, and the stock of patents for the university where the scientist obtained his/her doctorate increases the likelihood of patenting activities. Similarly, scientists working in areas of science that are inherently more amenable to patenting are, unsurprisingly, more likely to patent. At the mean of the data, a high (in the top quartile) research patentability score increases the probability of patenting by 36.20% (column 1a) and by 39.90% (column 2a).[19] Just as in the case of publications, the onset of patenting appears simultaneous with a change in the content of a scientist's research in a direction that makes it more similar to that of scientists who have already applied for patent rights. But because it is the flow, and not the stock of this measure that seems to matter, the evidence is consistent with the idea that a patent application does not constitute merely a response to changes in the formal and informal

[17]This interpretation is also consistent with Murray and Stern's (2005) analysis of paper-patent pairs, but it suggests that this phenomenon is not confined to the single journal whose articles they analyze. Of course, since we do not examine the actual content of patents and papers, we can only provide circumstantial evidence in favor of a substantive linkage between these two forms of output. In practice, it seems likely that patentable claims will be spread over a number of papers revolving around a common theme, some published before, some after the filing of the patent application.

[18]This result also provides strong evidence against the crowding-out hypothesis, at least in its simplest form. If the patent application process carried a high opportunity cost of time, one would expect this to be reflected in the output of patenting scientists before their first patent application. The opposite is true.

[19]This conclusion is not altered when using a more flexible functional form to model the distributed lag of the latent patentability score (Azoulay et al., 2005).

incentives faced by academic scientists over their careers, but also reflects the seizing of opportunities along a novel research trajectory.

In light of the results above, the shortcomings of fixed-effects estimation strategies become clearer. Selection into patenting is influenced by transitory shocks to outcome variables of interest, such as publications and their commercial content. While scientist fixed effects purge econometric estimates from selection bias stemming from immutable characteristics, they will fail to account for the transitory dynamics highlighted above.

Determinants of exit from academia. Models 3a and 3b display the results corresponding to specifications modeling the probability of exit from academia. A priori, one might imagine that academic scientists leave academia because they do not achieve success in the publication game. One might also conjecture that very productive academics are more likely to be poached by the private sector, leading to a premature exit from the academic ranks. We find support for both stories. Even controlling for the stock of past publications, a dry spell in academic productivity significantly increases the likelihood of exit. The stock of patents up to year $t-2$ and research patentability are found to have no meaningful effect, but a patent application in year $t-1$ is associated with a 34.9% increase in the probability of exit — although the effect is only marginally significant (column 3a).

Effect of patenting on the rate of publication output. Table 5 is divided into three sets of results, corresponding to three definitions of the patenting effect: flow (Models 1a, 1b, and 1c), regime (Models 2a, 2b, and 2c), and stock (Models 3a, 3b, and 3c). Within each set, the first column reports on the determinants the rate of publication using the conditional fixed effect poisson model of Hausman et al. (1984). As noted earlier, these estimates are likely to understate the causal effect of patenting. The second column is a "naïve" specification for the count of research publications, using Poisson Quasi-Maximum Likelihood Estimation (PQMLE).[20] The corresponding estimates are likely to be biased upwards by unobserved heterogeneity. The third column is identical to the second except that it also incorporates

[20]Because the Poisson model is in the linear exponential family, the coefficient estimates remain consistent as long as the mean of the dependent variable is correctly specified (Gouriéroux et al., 1984). Further, "robust" standard errors are consistent even if the underlying data generating process is not Poisson. In fact

our inverse probability of treatment and censoring weights. Under the sequential conditional independence assumption, these estimates correspond to the average treatment effect of patenting. Table 5 yields three robust results: (a) regardless of the method employed, the estimated effect of patenting is positive and statistically significant; (b) the IPTCW estimates are always higher than the conditional fixed effect estimates; and (c) in the cross-section, the magnitude of the effect is much lower once we account for self-selection into patenting. The formula $(e^{\beta} - 1) \times 100\%$ (where β denotes an estimated coefficient) provides a number directly interpretable in terms of elasticity. For example, the estimates in columns 2a, 2b, and 2c imply elasticities of publishing with respect to patenting equal to .215, .483 and .265, respectively.

Effect of patenting on the quality of publication output. Table 6 uses two distinct measures of publication quality.[21] The first is the proportion of publications in which the researcher appears in first or last position in the authorship list (Models 1a and 1b). We estimate the model using the quasi-maximum likelihood fractional logit estimator of Papke and Wooldridge (1996). The estimated effect is small in magnitude, flips sign between the unweighted and weighted version of the model, and is statistically insignificant in both cases. This suggests that patenting has very little impact on authorship position.

Our second measure is the average journal impact factor for the articles published in a given year (Models 2a and 2b). Estimation is performed using the Poisson QML approach as in Table 5. Here, we do find a positive and statistically significant effect, although it is quite small in magnitude (with an elasticity of about .05). From this mixed set of results, we conclude that the publication boost estimated in Table 5 does not come at the expense of the quality of these publications.

Effect of patenting on the content of publication output. Measuring the direction of scientific progress is always more challenging than merely measuring scientific output. In

the PQML estimator can be used for any non-negative dependent variables, whether integer or continuous (see Santos Silva and Tenreyro, Forthcoming).

[21] These two measures are not defined whenever a scientist has no output in a given year. As a result, the estimation sample shrinks by about a third.

Table 7, we propose three distinct ways of measuring the commercial content of scientists' publications, and we show that our conclusions are not sensitive to the choice of measure. We begin by using the research patentability score described in section 4 as the dependent variable, and we perform estimation using the PQML estimator in columns 1a and 1b (since RP is a non-negative, albeit continuous, dependent variable). Patenting increases modestly the latent patentability of the research published in the post-patenting regime, even when we adjust for confounding (our weights take into account the fact that a shock to patentability in period $t-1$ is associated with an increased likelihood of patenting at time t). For example, the estimates in Model 2b imply that entering the patenting regime increases RP by a statistically significant 8.8%.

Models 2a and 2b provide a different angle on the same question by focusing on the institutional affiliations of our scientists' coauthors. In the years of positive output, we compute the fraction of of total publications accounted for by articles in which at least one coauthor has an industry affiliation. At the mean of the data, the IPTCW estimates imply that entering the patenting regime increases this proportion by a statistically significant 29.4%. The naïve cross-sectional estimate is of a similar magnitude.

Finally, Models 3a and 3b use the average Journal Commercial Score (JCS) as the dependent variable. Starting from a journal-specific index that measures the proportion of authors publishing in the journal that have an industry affiliation, we compute the scientist-specific score by averaging these weights over all articles published in a given year.[22] Patenting appears to increase the average JCS in a statistically significant fashion, but the magnitude of the effect is modest: at the mean of the data, the IPTCW estimates correspond to 4.2% increase in average JCS for patenting scientists.

Taken together, however, these results paint a consistent picture whereby patenting increases the rate of scientific output while maintaining its quality, but also changes the content of these publications by connecting them more tightly to the world of commerce.

[22]Note that this measure has the advantage of not conflating the effect of patenting on the content of publications with its effect on the quantity of publication. As in the case of the average JIF, however, it suffers from the shortcoming that it is not defined whenever a scientist does not publish in a given year.

Sensitivity analysis. If the fixed effect specifications understate the causal effect of patenting, but the IPTCW specifications overstate it, Models 2a and 2c in Table 5 imply that the average treatment effect of patenting on publication rates, expressed as an elasticity, lies within the interval $[0.215; 0.265]$. In order to gauge the robustness of inverse probability of treatment-weighted estimation to confounding by unobserved variables, we conduct a sensitivity analysis. Following Robins (1999a) and Brumback et al. (2004), we ask how much unmeasured confounding would there need to be for the confidence interval around our treatment effect to include 0? This approach compels us to parameterize the bias from unobserved confounding — a functional form choice that is guided by little else than intuition regarding the cause and direction of bias. As such, the exercise does not provide a formal specification test for our results. Yet, its results are reassuring in the sense that our estimates appear robust to substantial amounts of unmeasured confounding. Estimation details and results are provided in Appendix II.

5 Discussion and Conclusion

While past research had established that commercialists are disproportionately recruited from the ranks of elite scientists and institutions (Zucker et al., 1998), our results build on this prior literature by showing that patenting is often accompanied by a flurry of publication activity in the year preceding the patent application, even after accounting for the lagged stock of publications. This result highlights the fact that academic patenting, rather than merely reflecting the influence of time-invariant demographic factors, also responds to variation in scientific opportunities (Azoulay et al., 2005). We also find that academic scientists who patent are more productive than otherwise equivalent scientists that are not listed as inventors on patents, but that publication quality appears relatively similar in the two groups. Thus, the evidence appears to reject the assertion that the increase in patenting in academe has come at the cost of diverting researchers' time, interest, and attention from their traditional focus on standard scientific research. However, we also find that scientists

alter the content of their research after they patent in ways that make their output more relevant to questions of commercial interest.

These results depend on the strong assumption that the outcomes we examine be independent of patenting conditional on the history of observables. As in all observational studies, this assumption cannot be tested from the data. It is obviously better to include a large set of potential confounders to model the probability of selection, but we recognize that in practice, the assumption of selection on observables may still not be precisely or even approximately tenable. We take solace in the results of a sensitivity analysis showing that our core result is in fact robust.

There are two other avenues, all outside the scope of this analysis, through which patenting in academic science could yet have a significant — and possibly deleterious — effect on the advancement of scientific knowledge. As a result, beyond the first-order effect of a scientist's decision to patent on his or her individual productivity, our conclusions must remain tempered.

First, as patenting within a department or research area continues to grow, is there a point at which a negative effect on the collective output sets in, either because researchers are deterred or blocked by intellectual property rights held by others, or because concerns about intellectual property rights diminish open communications among scientists? This "tragedy of the anti-commons" has recently been investigated by Murray and Stern (2005), who provide evidence that scientific papers paired with patent applications are less likely to be cited after the patent is granted by the USPTO (though the effect they uncover is modest in magnitude). In the context of this paper, we present evidence that patenting changes the content of individual scientists' research trajectory, but there is a strong leap from our results to normative statements regarding the welfare implications of this change in research agenda.

Academic patenting might also alter the career trajectories of the graduate students and post-doctoral fellows that work in patenters' laboratories. For instance, patenters may have much thicker and more diverse relationships with researchers in firms than non-patenting

scientists, which may in turn facilitate apprentice scientists' job searches in the private sector. Therefore, patenters may (perhaps unintentionally) encourage their students to select private-sector careers above academic posts. Conversely, if patenters enlist the help of scientists-in-training in the research streams that lead to patents, and if these projects are different from the research topics that intrigue non-patenters, apprentices training under patenters may be less appealing to academic departments searching for new faculty. In short, the most significant impact of patenting on public research output may well lie in the consequence of the behavior for non-patenting and soon-to-be scientists. We plan to investigate this topic in future research.

References

Abadie, Alberto. 2005. "Semiparametric Difference-in-Differences Estimators." *Review of Economic Studies*, 72:1, pp. 1-19.

Agrawal, Ajay K. and Rebecca M. Henderson. 2002. "Putting Patents in Context: Exploring Knowledge Transfer from MIT." *Management Science,* 48:1, pp. 44-60.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy*, 113:1, pp. 151-184.

AUTM Licensing Survey. 2003. *A Survey Summary of Technology Licensing (and Related) Performance for US and Canadian Academic and Nonprofit Institutions, and Technology Investment Firms*. Ashley J. Stevens and Frances Toneguzzo, eds.

Azoulay, Pierre, Waverly W. Ding, and Toby E. Stuart. 2005. "The Determinants of Faculty Patenting Behavior: Demographics or Opportunities?" NBER Working Paper #11348.

Blumenthal, David, Nancyanne Causino, Eric G. Campbell, and Karen Seashore Louis. 1996. "Relationships between Academic Institutions and Industry in the Life Sciences — An Industry Survey." *New England Journal of Medicine*, 334:6, pp. 368-73.

Blumenthal, David, Michael Gluck, Karen Seashore Louis, Michael A. Stoto, and David Wise. 1986. "University-Industry Research Relationships in Biotechnology: Implications for the University." *Science*, 232:4756, pp. 1361-66.

Brumback, Babette A., Hernán, Miguel A., Haneuse, Sébastien J.P.A., and James M. Robins. 2004. "Sensitivity Analyses for Unmeasured Confounding Assuming a Marginal Structural Model for Repeated Measures." *Statistics in Medicine*, 23:5, pp. 749-767.

Campbell, Eric G. , B. R. Clarridge, N. N. Gokhale, L. Birenbaum, S. Hilgartner, N.A. Holtzman, and D. Blumenthal. 2002. "Data Withholding in Academic Genetics — Evidence from a National Survey." *JAMA*, 287:4, pp. 473-80.

Carmichael, H, Lorne. 1988. "Incentives in Academics: Why Is There Tenure?" *Journal of Political Economy*, 96:3, pp. 453-472.

Cohen, Wesley M. and Daniel A. Levinthal. 1989. "Innovation and Learning — The Two Faces of R&D." *Economic Journal,* 99:397, pp. 569-96.

Dasgupta, Partha and Paul A. David. 1994. "Toward a New Economics of Science." *Research Policy*, 23:5, pp. 487-521.

Dehejia, Rajeev H. and Sadek Wahba. 2002. "Propensity Score-Matching Methods for Nonexperimental Causal Studies." *Review of Economics and Statistics*, 84:1, pp. 151-61.

Di Gregorio, Dante and Scott Shane. 2003. "Why Do Some Universities Generate More Start-ups Than Others?" *Research Policy*, 32:2, pp. 209-27.

Ding, Waverly W. 2005. "Why Do For-Profit Firms Adopt Open Science? Assessing the Impact of Founder Imprinting, Niche Crowding and Competitive Influence." Working Paper, Haas School of Business, University of California.

Elfenbein, Daniel W. 2005. "The Market for Embryonic Technologies: Lessons from University Licensing." Working Paper, University of California.

Fabrizio, Kira and Alberto Diminin. 2005. "Commercializing the Laboratory: The Relationship between Faculty Patenting and Publishing." Working Paper, Emory University.

Gouriéroux, Christian, Montfort, Alain, and Alain Trognon. 1984. "Pseudo Maximum Likelihood Methods: Applications to Poisson Models." *Econometrica,* 52:3, pp. 701-720.

Hall, Bronwyn H., Adam B. Jaffe, and Manuel Trajtenberg. 2001. "The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools." NBER Working Paper #8498.

Hausman, Jerry, Bronwyn H. Hall, and Zvi Griliches. 1984. "Econometric Models for Count Data with an Application to the Patents-R&D Relationship." *Econometrica,* 52:4, pp. 909-38.

Heckman, James J., Hidehiko Ichimura, and Petra E. Todd. 1997. "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme." *Review of Economic Studies,* 64:4, pp. 605–654.

Hellman, Thomas. 2005. "The Role of Patents for Bridging the Science to Market Gap." NBER Working Paper #11460.

Henderson, Rebecca M. and Iain Cockburn. 1994. "Measuring Competence? Exploring Firm Effects in Pharmaceutical Research." *Strategic Management Journal,* 15, pp. 63-84.

Henderson, Rebecca M., Adam B. Jaffe, and Manuel Trajtenberg. 1998. "Universities as a Source of Commercial Technology: A Detailed Analysis of University Patenting, 1965-1988." *Review of Economics and Statistics*, 80:1, pp. 119-127.

Hernán, 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, pp. 561-570.

Hernán, Miguel A., Babette Brumback, and James M. Robins. 2001. "Marginal Structural Models to Estimate the Joint Causal Effect of Nonrandomized Treatments." *Journal of the American Statistical Association*, 96:454, pp. 440-48.

Hull, David I. 1988. *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science*. Chicago: University of Chicago Press.

Jaffe, Adam B. 1986. "Technological Opportunity and Spillovers of R&D: Evidences from Firms' Patents, Profits and Market Value." *American Economic Review*, 76:5, pp. 984-1001.

Jaffe, Adam B. 1989. "Real Effects of Academic Research." *American Economic Review*, 79:5, pp. 957-970.

Jaffe, Adam B. and Josh Lerner. 2001. "Reinventing Public R&D: Patent Policy and the Commercialization of National Laboratory Technologies." *The RAND Journal of Economics,* 32:1, pp. 167-98.

Jensen, Richard and Marie C. Thursby. 2001. "Proofs and Prototypes for Sale: The Licensing of University Inventions." *American Economic Review*, 91:1, pp. 240-59.

Krimsky, Sheldon. 2003. *Science in the Private Interest: Has the Lure of Profits Corrupted Biomedical Research?* Lanham, MD: Rowman & Littlefield.

Lach, Saul and Mark Schankerman. 2004. "Royalty Sharing and Technology Licensing in Universities." *Journal of The European Economic Association,* 2:2-3, pp. 252-264.

Levin, Sharon G. and Paula E. Stephan. 1991. "Research Productivity over the Life-Cycle — Evidence for Academic Scientists." *American Economic Review*, 81:1, pp. 114-32.

Lim, Kwanghui. 2004. "The Relationship between Research and Innovation in the Semiconductor and Pharmaceutical Industries (1981–1997)." *Research Policy,* 33:2, pp. 287-321.

Long, J. Scott and Mary Frank Fox. 1995. "Scientific Careers: Universalism and Particularism." *Annual Review of Sociology,* 21, pp. 45-71.

Merges, Robert. 1997. *Patent Law and Policy: Cases and Materials.* Charlottesville, VA: Michie Co.

Merton, Robert K. 1973 [1942]. "The Normative Structure of Science." In *The Sociology of Science: Theoretical and Empirical Investigations.* Robert K. Merton, ed. Chicago, IL: The University of Chicago Press.

Mowery, David C., Richard R. Nelson, Bhaven N. Sampat, and Arvids A. Ziedonis. 2001. "The Growth of Patenting and Licensing by US Universities: an Assessment of the Effects of the Bayh-Dole Act of 1980." *Research Policy*, 30:1, pp. 99-119.

Murray, Fiona. 2002. "Innovation as Co-evolution of Scientific and Technological Networks: Exploring Tissue Engineering." *Research Policy,* 31:8-9, pp. 1389-1403.

Murray, Fiona and Scott Stern. 2005. "Do Formal Intellectual Property Rights Hinder the Free Flow of Scientific Knowledge? An Empirical Test of the Anti-Commons Hypothesis." NBER Working Paper #11465.

Papke, Leslie E. and Jeffrey M. Wooldridge. 1996. "Econometric Methods for Fractional Responses with an Application to 401(k) Plan participation Rates." *Journal of Applied Econometrics,* 11:6, pp. 619–632.

Newey, Whitney K. and Daniel McFadden. 1994. "Large Sample Estimation and Hypothesis Testing." Chapter 36 in *Handbook of Econometrics, vol. 4.* Robert F. Engle and Daniel L. McFadden, eds. Amsterdam: North-Holland.

Owen-Smith, Jason and Walter. W. Powell. 2001a. "Careers and Contradictions: Faculty Responses to the Transformation of Knowledge and its Uses in the Life Sciences." *Research in the Sociology of Work,* 10, pp. 109-40.

Owen-Smith, Jason and Walter W. Powell. 2001b. "To Patent or Not: Faculty Decisions and Institutional Success at Technology Transfer." *Journal of Technology Transfer,* 26:1-2, pp. 99-114.

Robins, James M. 1999a. "Sensitivity Analysis for Selection Bias and Unmeasured Confounding in Missing Data and Causal Inference Models." pp. 1-92 in *Statistical Models in Epidemiology: The Environment and Clinical Trials.* M.E. Halloran and D. Berry eds. Springer-Verlag: New York.

Robins, James M. 1999b. "Marginal Strucural Models versus Structural Nested Models as Tools for Causal Inference." pp. 95-134 in *Statistical Models in Epidemiology: The Environment and Clinical Trials.* M.E. Halloran and D. Berry, eds. Springer-Verlag: New York.

Robins, James M., Miguel A. Hernán, and Babette Brumback. 2000. "Marginal Structural Models and Causal Inference in Epidemiology." *Epidemiology*, 11:5, pp. 550-60.

Rosenbaum, Paul R. 2002. *Observational Studies,* 2nd ed. New York: Springer.

Rosenbaum, Paul R. and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika*, 70:1, pp. 41-55.

Rosenberg, Nathan. 1990. "Why Do Firms Do Basic Research (with Their Own Money)?" *Research Policy*, 19, pp. 165-74.

Rosenberg, Nathan. 1998. "Chemical Engineering as a General Purpose Technology," in *General Purpose Technologies and Economic Growth*. E. Helpman, ed. Cambridge: MIT Press, pp. 167-92.

Santos Silva, J.M.C. and Silvana Tenreyro. 2005. "The Log of Gravity." Forthcoming, *Review of Economics and Statistics.*

Shane, Scott and Rakesh Khurana. 2003. "Bringing Individuals Back In: The Effects of Career Experience on New Firm Founding." *Industrial and Corporate Change,* 12:3, pp. 519-543.

Smith, Jeffrey A. and Petra E. Todd. 2005. "Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators?" *Journal of Econometrics,* 125, pp. 305-353.

Stephan, Paula E., Shiferaw Gurmu, A.J. Sumell, and Grant Black. 2004. "Who's Patenting in the University? Evidence from the Survey of Doctorate Recipients." Forthcoming, *Economics of Innovation and New Technology.*

Stern, Scott. 2004. "Do Scientists Pay to Be Scientists?" *Management Science*, 50:6, pp. 835-53.

Stokes, Donald. 1997. *Pasteur's Quadrant: Basic Science and Technological Innovation.* Washington, D.C.: The Brookings Institution.

Stuart, Toby E. and Waverly W. Ding. 2006. "The Social Structural Determinants of Academic Entrepreneurship: An Analysis of University Scientists' Participation in Commercial Ventures." Forthcoming, *American Journal of Sociology.*

Sulston, John. 2003. "Beyond Release: The Equitable Use of Genomic Information." *The Lancet*. 362, pp. 400-402.

Thursby, Jerry G. and Marie C. Thursby. 2002. "Who Is Selling the Ivory Tower? Sources of Growth in University Licensing." *Management Science*, 48:1, pp. 90-104.

Thursby, Marie C., Jerry G. Thursby, and Swastika Mukherjee. 2005. "Are There Real Effects of Licensing on Academic Research? A Life Cycle View." NBER Working Paper #11497.

Weitzman, Martin L. 1998. "Recombinant Growth." *Quarterly Journal of Economics*, 113:2, pp. 331-360.

Wooldridge, Jeffrey M. 2002. "Inverse Probability Weighted M-Estimators for Sample Selection, Attrition, and Stratification." *Portuguese Economic Journal,* 1:2, pp. 117–139.

Zucker, Lynne G., Michael R. Darby, and Marilynn B. Brewer. 1998. "Intellectual Human Capital and the Birth of U.S. Biotechnology Enterprises." *American Economic Review*, 88:1, pp. 290-306.

Table 1
Top 15 Scientific Disciplines in the Sample

UMI Subject Code	UMI Subject Description	Frequency	
487	Biochemistry	855	(22.1%)
306	Biology, General	563	(14.6%)
410	Biology, Microbiology	466	(12.1%)
419	Health Sciences, Pharmacology	239	(6.2%)
490	Chemistry, Organic	212	(5.5%)
786	Biophysics, General	210	(5.4%)
369	Biology, Genetics	191	(4.9%)
433	Biology, Animal Physiology	170	(4.4%)
982	Health Sciences, Immunology	167	(4.3%)
307	Biology, Molecular	102	(2.6%)
301	Bacteriology	61	(1.6%)
287	Biology, Anatomy	54	(1.4%)
571	Health Sciences, Pathology	52	(1.3%)
349	Psychology, Psychobiology	37	(1.0%)
572	Health Sciences, Pharmacy	33	(0.9%)

Legend: Table 1 reports the Top 15 disciplines from which the sample was drawn and the number and proportion of scientists in each of the 15 scientific disciplines. The table also reports the frequency and the proportion of scientists in our sample for each of these 15 scientific disciplines.

Table 2
Descriptive Statistics

	Mean	Std. Dev.	Min.	Max.	*N*
Patent Flow (=1 if one or more patent app. in year)	0.017	0.131	0	1	58,562
Patent Regime (=1 after first patent app.)	0.073	0.261	0	1	58,562
Patent Stock	0.184	1.175	0	57	58,562
Research Publication Flow	1.729	2.379	0	35	58,562
Research Publication Stock	17.563	26.759	0	386	58,562
Fraction of First or Last Authored Publications (Flow)	0.619	0.397	0	1	38,007
Average JIF of Publications (Flow)	3.956	3.101	0.005	30.334	38,007
Average Journal Commercial Score of Pubs. (Flow)	0.076	0.055	0.001	1	38,007
Fraction of Pubs. with Industry Coauthors (Flow)	0.075	0.223	0	1	38,007
Research Patentability Score (Flow)	0.022	0.049	0	4.173	58,562
Research Patentability Stock	0.111	0.142	0	4.201	58,562
Employer Graduate School in Top 20	0.231	0.422	0	1	58,562
Employer Has TTO	0.488	0.500	0	1	58,562
Employer Patent Stock	71.80	145.18	0	2,189	58,562
Experience (Career Age)	10.201	7.122	1	32	58,562
Calendar year	1986	7.741	1968	1999	58,562
Female	0.183	0.387	0	1	3,862
Ph.D. Univ. Grad. School in Top 20	0.308	0.462	0	1	3,862
Ph.D. Univ. 5-year Patent Stock	18.983	40.906	0	566	3,862
Scientist Has One or More Patents	0.122	0.328	0	1	3,862

Table 3
Mean Scientist and Employer Characteristics at Five Career Stages, by Patent Application Status

	Experience = 5		Experience = 10		Experience = 15		Experience = 20		Experience = 25	
Scientist ever applied for a patent right	*Yes*	*No*	*Yes*	*No*	*Yes*	*No*	*Yes*	*No*	*Yes*	*No*
(1) Count of Research Publications (Flow)	1.563	1.290	2.524	1.821	3.208	2.036	3.513	2.215	3.395	2.179
	(1.648)	(1.801)	(2.945)	(2.228)	(3.276)	(2.629)	(4.029)	(2.888)	(4.023)	(2.955)
(2) Count of Research Publications (Stock)	6.760	5.832	19.066	14.996	35.389	24.429	50.974	37.227	74.386	48.098
	(5.971)	(6.668)	(16.753)	(14.819)	(28.251)	(23.490)	(40.143)	(34.069)	(60.078)	(45.535)
(3) Fraction of First or Last Authored Pubs.	0.625	0.605	0.604	0.628	0.568	0.623	0.617	0.577	0.654	0.566
	(0.404)	(0.416)	(0.390)	(0.394)	(0.366)	(0.383)	(0.362)	(0.392)	(0.345)	(0.389)
(4) Average JIF of Research Publications	5.257	4.107	4.441	3.901	4.161	3.800	4.021	3.586	4.244	3.417
	(4.133)	(3.368)	(3.586)	(2.982)	(2.785)	(2.765)	(3.004)	(2.616)	(2.751)	(2.660)
(5) Average JCS of Publications	0.070	0.077	0.074	0.077	0.084	0.075	0.068	0.073	0.062	0.075
	(0.039)	(0.056)	(0.053)	(0.060)	(0.075)	(0.051)	(0.037)	(0.050)	(0.031)	(0.057)
(6) Fraction of Pubs. with Industry Coauthors	0.145	0.052	0.102	0.077	0.089	0.085	0.105	0.114	0.108	0.099
	(0.306)	(0.196)	(0.250)	(0.230)	(0.225)	(0.233)	(0.251)	(0.260)	(0.225)	(0.239)
(7) Research Patentability Score (Flow)	0.024	0.016	0.043	0.023	0.037	0.027	0.047	0.032	0.037	0.036
	(0.028)	(0.050)	(0.133)	(0.047)	(0.032)	(0.032)	(0.048)	(0.041)	(0.029)	(0.038)
(8) Research Patentability Score (Stock)	0.078	0.052	0.178	0.113	0.230	0.157	0.289	0.209	0.293	0.245
	(0.074)	(0.091)	(0.183)	(0.142)	(0.181)	(0.130)	(0.166)	(0.144)	(0.122)	(0.175)
(9) Employer Grad. School in Top 20	0.323	0.264	0.313	0.219	0.250	0.200	0.197	0.181	0.175	0.170
	(0.470)	(0.441)	(0.465)	(0.413)	(0.434)	(0.400)	(0.399)	(0.385)	(0.382)	(0.376)
(10) Employer has TTO	0.531	0.384	0.620	0.486	0.694	0.595	0.719	0.688	0.825	0.738
	(0.502)	(0.486)	(0.487)	(0.500)	(0.462)	(0.491)	(0.450)	(0.463)	(0.382)	(0.440)
(11) Employer Patent Stock	107.4	53.6	159.4	64.6	143.0	75.9	134.4	110.2	172.3	120.8
	(206.8)	(136.7)	(307.3)	(133.7)	(224.1)	(116.4)	(185.1)	(155.1)	(238.6)	(163.7)
Number of scientists (rows 1, 2 and 7-11)	96	3,610	166	2,429	216	1,621	228	1072	114	519
Number of scientists (rows 3-6)	69	2,278	128	1,646	176	1,108	198	738	87	355

Legend: Table 3 reports the mean and standard deviation (in parentheses) of scientist research and employer characteristics measured at five career ages: the 5th, 10th, 15th, 20th and 25th year after a scientist was granted a Ph.D. At each professional age, the table is further broken out by whether a scientist has applied for at least one patent right throughout his career. For example, if a scientist applied for a patent right during the 20th year after he was granted a Ph.D., he contributed to the mean values of the "no" category of experience = 5, 10 and 15, and to the mean values of the "yes" category of experience = 20 and 25.

Table 4
Probability of Patenting and Exiting Academia

	Model 1a	Model 1b	Model 2a	Model 2b	Model 3a	Model 3b
Dependent Variable	*Patent Flow*		*Patent Regime*		*Exit Academia*	
	Denominator	*Numerator*	*Denominator*	*Numerator*	*Denominator*	*Numerator*
Experience = [5, 8]	0.141	0.195	0.166	0.239		
	(0.153)	(0.153)	(0.166)	(0.164)		
Experience = [9, 15]	0.219	0.347	0.305	0.432	0.206	-0.006
	(0.155)	(0.151)*	(0.168)†	(0.162)**	(0.060)**	(0.057)
Experience = [16, 22]	0.022	0.218	0.252	0.401	0.116	-0.264
	(0.174)	(0.162)	(0.196)	(0.180)*	(0.087)	(0.077)**
Experience = [23, 35]	-0.357	-0.097	-0.343	-0.232	0.371	-0.122
	(0.213)†	(0.198)	(0.278)	(0.267)	(0.116)**	(0.101)
Female	-0.649	-0.675	-0.663	-0.700	0.147	0.243
	(0.130)**	(0.133)**	(0.153)**	(0.152)**	(0.054)**	(0.053)**
Patent Flow $_{t-1}$	1.971	2.048			0.299	
	(0.093)**	(0.128)**			(0.174)†	
Patent Stock $_{t-2}$	1.945	2.065			-0.128	
	(0.124)**	(0.093)**			(0.103)	
Publications Flow $_{t-1}$	0.042		0.083		-0.215	
	(0.016)**		(0.022)**		(0.024)**	
Publications Stock $_{t-2}$	0.003		-0.001		-0.013	
	(0.002)		(0.002)		(0.003)**	
High Research Patentability $_{t-1}$	0.309		0.336		-0.097	
	(0.093)**		(0.112)**		(0.068)	
Research Patentability Stock $_{t-2}$	0.129		0.247		0.017	
	(0.309)		(0.300)		(0.203)	
Has Industry Coauthors $_{t-1}$	0.076		0.061		0.055	
	(0.093)		(0.113)		(0.061)	
Employer Grad. School in Top 20	0.143		-0.014		0.054	
	(0.113)		(0.119)		(0.059)	
Employer has TTO $_{t-1}$	0.137		0.012		-0.050	
	(0.096)		(0.118)		(0.053)	
Employer Patent Stock $_{t-1}$ × 100	-0.007		0.090		0.031	
	(0.026)		(0.033)**		(0.016)†	
Ph.D. Univ Grad. School in Top 20	0.011	0.053	0.089	0.121	-0.151	-0.181
	(0.092)	(0.089)	(0.104)	(0.104)	(0.053)**	(0.053)**
Ph.D. Univ. 5-Year Patent Stock	0.001	0.001	0.001	0.002	-0.001	-0.001
	(0.001)	(0.001)†	(0.001)	(0.001)*	(0.001)	(0.001)
Constant	-6.043	-5.968	-6.098	-6.039	-4.383	-4.533
	(0.295)**	(0.300)**	(0.304)**	(0.302)**	(0.139)**	(0.139)**
Observations	58,562	58,562	54,746	54,746	58,437	58,437
Number Of Researchers	3,862	3,862	3,862	3,862	3,862	3,862
Log Pseudo-Likelihood	-3,956.36	-3,994.80	-2,549.11	-2,578.29	-8,878.77	-9,092.91
Wald X^2	2,263.35	2,089.54	348.72	272.91	564.09	308.91
Number Of Variables	48	40	46	38	47	37

Notes:

(1) Models 2a-2b exclude observations after a researcher has filed for his or her first patent application. Models 3a-3b exclude observations after a researcher has accumulated 30 years' professional experience (at which point he or she is no longer considered at risk of exiting academia).

(2) All models control for Ph.D. subject and calendar year dummies.

(3) Robust standard errors in parentheses, clustered around individual researchers.

(4) † significant at 10%; * significant at 5%; ** significant at 1%.

Table 5
Effect of Patenting on the Rate of Publications: Poisson Models

	Model 1a	Model 1b	Model 1c	Model 2a	Model 2b	Model 2c	Model 3a	Model 3b	Model 3c
Scientist Fixed Effects	*Yes*	*No*	*No*	*Yes*	*No*	*No*	*Yes*	*No*	*No*
IPTC Weights	*No*	*No*	*Yes*	*No*	*No*	*Yes*	*No*	*No*	*Yes*
Experience = [5, 8]	0.160	0.205	0.209	0.161	0.200	0.206	0.162	0.206	0.209
	(0.018)**	(0.018)**	(0.018)**	(0.018)**	(0.018)**	(0.019)**	(0.018)**	(0.018)**	(0.018)**
Experience = [9, 15]	0.260	0.445	0.428	0.262	0.430	0.420	0.263	0.447	0.425
	(0.029)**	(0.030)**	(0.033)**	(0.029)**	(0.030)**	(0.033)**	(0.028)**	(0.030)**	(0.033)**
Experience = [16, 22]	0.229	0.554	0.443	0.228	0.521	0.427	0.229	0.548	0.430
	(0.041)**	(0.049)**	(0.047)**	(0.041)**	(0.049)**	(0.047)**	(0.041)**	(0.049)**	(0.047)**
Experience = [23, 32]	0.085	0.521	0.348	0.085	0.487	0.335	0.082	0.494	0.332
	(0.050)†	(0.073)**	(0.069)**	(0.050)†	(0.073)**	(0.070)**	(0.050)†	(0.073)**	(0.069)**
Female		-0.215	-0.230		-0.203	-0.224		-0.216	-0.225
		(0.052)**	(0.049)**		(0.051)**	(0.049)**		(0.052)**	(0.049)**
PhD Univ. Grad School in Top 20		0.067	0.056		0.063	0.052		0.070	0.054
		(0.042)	(0.041)		(0.042)	(0.041)		(0.042)†	(0.041)
PhD Univ. 5-yr Patent Stock × 100		0.046	0.046		0.043	0.048		0.047	0.046
		(0.048)	(0.047)		(0.047)	(0.047)		(0.048)	(0.047)
Patent Flow	0.165	0.539	0.300						
	(0.028)**	(0.057)**	(0.057)**						
Patent Regime				0.195	0.394	0.235			
				(0.031)**	(0.048)**	(0.047)**			
Patent Stock							0.016	0.045	0.055
							(0.010)	(0.012)**	(0.011)**
Constant		0.033	0.038		0.034	0.041		0.037	0.040
		(0.044)	(0.044)		(0.044)	(0.045)		(0.044)	(0.044)
Log pseudo-likelihood	-78,109.5	-120,275.9	-117,023.9	-78,070.0	-119,953.1	-117,057.9	-78,126.2	-120,266.5	-116,923.5
Wald χ^2	2,897.67	1,234.53	975.28	2,966.37	1,301.65	948.59	2,858.59	1,228.92	983.47
Number of covariates	29	39	39	29	39	39	29	39	39

Notes:
(1) Number of observations = 58,562; number of researchers = 3,862.
(2) All models control for calendar year dummies; all cross-sectional models also control for Ph.D. subject dummies.
(3) All cross-sectional models report robust standard errors in parentheses, clustered around researchers.
(4) † significant at 10%; * significant at 5%; ** significant at 1%.

Table 6
Effect of Patenting on the Quality of Publications

	Model 1a	Model 1b	Model 2a	Model 2b
	Fractional Logit QML Estimates Proportion of First or Last-Authored Publications		Poisson Model QML Estimates Average JIF of Publications	
	Unweighted	*IPTCW*	*Unweighted*	*IPTCW*
Experience = [5, 8]	-0.096 (0.029)**	-0.096 (0.029)**	-0.087 (0.013)**	-0.088 (0.013)**
Experience = [9, 15]	0.034 (0.034)	0.029 (0.034)	-0.189 (0.018)**	-0.186 (0.018)**
Experience = [16, 22]	0.133 (0.046)**	0.122 (0.046)**	-0.273 (0.027)**	-0.275 (0.027)**
Experience = [23, 32]	0.155 (0.068)*	0.137 (0.070)†	-0.354 (0.039)**	-0.366 (0.040)**
Female	-0.003 (0.038)	0.0003 (0.038)	0.031 (0.022)	0.033 (0.022)
PhD Univ. Grad School in Top 20	0.050 (0.033)	0.047 (0.033)	0.135 (0.021)**	0.131 (0.021)**
PhD Univ. 5-yr Patent Stock × 100	0.049 (0.042)	0.041 (0.043)	0.086 (0.030)**	0.094 (0.029)**
Patent Regime	0.026 (0.048)	-0.004 (0.051)	0.077 (0.029)**	0.052 (0.030)†
Constant	0.826 (0.047)**	0.827 (0.047)**	1.370 (0.023)**	1.371 (0.023)**
Log pseudo-likelihood	-22,238.9	-21,846.2	-91,867.7	-90,193.4
Wald χ^2	272.6	268.9	642.1	680.8

Notes:
(1) Number of observations = 38,007; number of researchers = 3,862; number of variables = 39.
(2) All models control for PhD subject and calendar year dummies.
(3) Robust standard errors are reported in parenthesis, clustered around researchers.
(4) † significant at 10%; * significant at 5%; ** significant at 1%.

Table 7
Effect of Patenting on the Commercial Content of Publications

	Model 1a	Model 1b	Model 2a	Model 2b	Model 3a	Model 3b
	Poisson Models QML Estimates Research Patentability		Fractional Logit QML Estimates Proportion of Pub. with Industry Coauthors		Fractional Logit QML Estimates Average Journal Commercial Score	
	Unweighted	*IPTCW*	*Unweighted*	*IPTCW*	*Unweighted*	*IPTCW*
Experience = [5, 8]	0.008	0.005	0.102	0.099	0.016	0.016
	(0.039)	(0.039)	(0.069)	(0.070)	(0.014)	(0.014)
Experience = [9, 15]	-0.025	-0.024	0.130	0.124	0.006	0.006
	(0.038)	(0.037)	(0.086)	(0.086)	(0.019)	(0.019)
Experience = [16, 22]	-0.054	-0.054	0.122	0.128	0.015	0.019
	(0.038)	(0.038)	(0.111)	(0.111)	(0.025)	(0.025)
Experience = [23, 32]	-0.103	-0.104	0.087	0.083	0.057	0.076
	(0.042)**	(0.043)*	(0.154)	(0.155)	(0.035)	(0.035)*
Female	-0.023	-0.023	-0.070	-0.066	-0.007	-0.005
	(0.022)	(0.023)	(0.091)	(0.092)	(0.017)	(0.017)
PhD Univ. Grad School in Top 20	-0.027	-0.025	-0.313	-0.329	-0.069	-0.067
	(0.021)	(0.022)	(0.084)**	(0.086)**	(0.018)**	(0.018)**
PhD Univ. 5-yr Patent Stock × 100	-0.017	-0.018	0.133	0.113	-0.018	-0.018
	(0.020)	(0.020)	(0.098)	(0.091)	(0.025)	(0.026)
Patent Regime	0.090	0.085	0.222	0.278	0.043	0.052
	(0.028)**	(0.029)**	(0.088)*	(0.097)**	(0.024)†	(0.026)*
Constant	-5.700	-5.700	-3.831	-3.827	-2.491	-2.494
	(0.353)**	(0.352)**	(0.153)**	(0.153)**	(0.024)**	(0.024)**
Log pseudo-likelihood	-4,887.3	-4,750.6	-9,099.0	-8,901.8	-7,669.4	-7,524.1
Wald χ^2	2,089.6	1,939.8	305.47	295.21	431.53	394.01

Notes:
(1) Number of observations = 38,007; number of researchers = 3,862; number of variables = 39.
(2) All models control for PhD subject and calendar year dummies.
(3) Robust standard errors are reported in parenthesis, clustered around researchers.
(4) † significant at 10%; * significant at 5%; ** significant at 1%.

Figure 1:
Distribution of Patent Count for Patenting Scientists

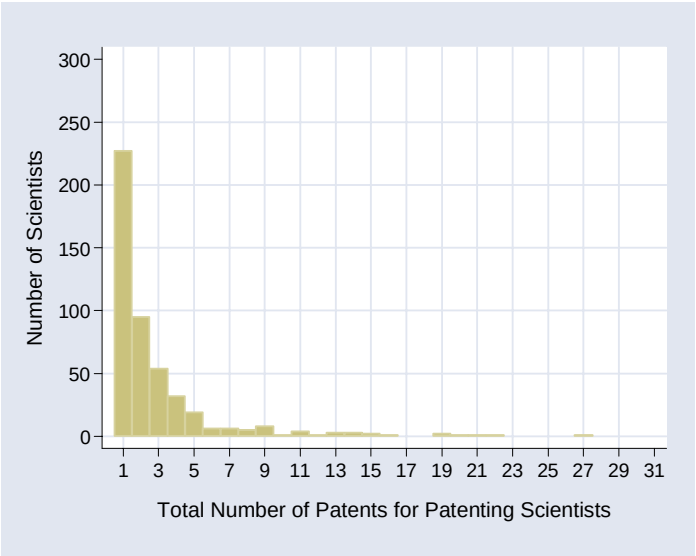
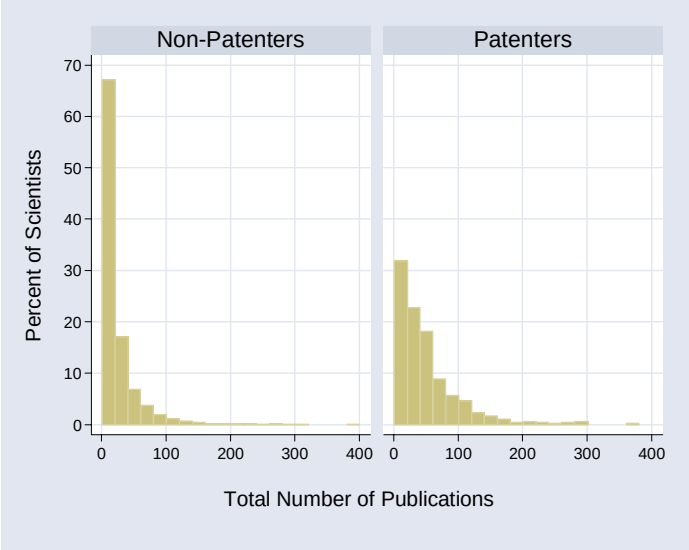


Figure 2:
Distribution of Publication Count for Patenting and Non-Patenting Scientists



Appendix I: Keyword Weights

w^i_{jt}, the patentability weight for each keyword j in year t is defined as:

$$w^i_{jt} = \frac{\sum_{s \in I^p_t - \{i\}} \frac{m_{sjt}}{\sum_k m_{skt}}}{\sum_{s \in I^{np}_t - \{i\}} m_{sjt}}$$

where m_{sjt} denotes the number of times keyword j has appeared in articles published up to year t by scientist s, I^p_t is the subset of scientists in our sample that have already applied for one or more patents as of year t, and I^{np}_t is the subset of scientists in our sample that have not yet applied for any patent as of year t. The weight is also indexed by scientist i, because i's publications are taken out of the set of articles used to compute the formula above.

To create the numerator of w^i_{jt}, we first create a row-normalized matrix with each scientist in the patenting regime listed in a row and each of the keywords used to describe their papers up to year t listed in a column. The sj^{th} cell in the matrix, $[m_{sjt}/\sum_k m_{skt}]$, corresponds to the proportion of title keywords for scientist s that corresponds to keyword j. We then take the column sums from this matrix, i.e., we sum the contributions of individual patenting scientists for keyword j. Turning next to the denominator, we proceed in a similar manner, except that the articles considered only belong to the set of scientists who have not applied for patents as of year t. The numerator is then deflated by the frequency of use for j by non-patenters (in the rare case of keywords exclusively used by patenters, we substitute the number 1 for the frequency).

The weights w^i_{jt} are large for keywords that have appeared with disproportionate frequency as descriptors of papers written by scientists already in the patenting regime, relative to scientists not yet in the patenting regime.

Two things should be noted about the construction of these weights. First, $w^i_{jt} = 0$ for all keywords that have never appeared in the titles of papers written by scientists that have patented before t. Second, the articles written by scientist i him/herself do not contribute at all to the weights w^i_{jt}. Therefore, no scientist can directly influence year-to-year changes in these weights.

The final step for each scientist i in the dataset is to produce a list of the keywords in the individual's papers published in year t, calculate the proportion of the total represented by each keyword j, apply the appropriate keyword weight $w^i_{j,t-1}$, and sum over keywords to produce a composite score. The resulting variable increases in the degree to which keywords in the titles of a focal scientist's papers have appeared relatively more frequently in the titles of other academics who have applied for patents. This score is entered in the regressions to control for the research patentability of scientists' areas of specialization.

To illustrate the construction of the research patentability measure, Table A1 lists some representative keywords, along with their patentability weights in the year 2000. Consider the keyword "ubiquitin" (italicized in the table) in group 1. In 1999, it had previously appeared 55 times as a keyword in one or more articles of scientists who had patented prior to 1999. Among them is Keith D. Wilkinson, professor of biochemistry at Emory University School of Medicine, who is listed as an inventor on a patent filed in 1992. To compute the numerator of the patentability weight for this keyword, we begin with the fraction of Wilkinson's research using "ubiquitin" in the title. In his 43 ISI-listed research papers published between 1977 (when he was granted a Ph.D.) and 1999, 133 unique keywords have been used a total of 330 times. The word "ubiquitin" was used 24 times, hence the fraction of Wilkinson's research stock devoted to "ubiqutin" is 0.073. This procedure is repeated for the other eight patenting scientists who have used the word. The sum of these fractions taken over all patenting scientists is reported in column (2) of the table. Next, to compute the denominator in the above equation, we examine the keywords of all scientists who had not yet received a patent by 1999 for the appearance of the word ubiquitin. In the research publications of 3,854 such scientists, this keyword has appeared on 30 occasions. The patentability weight for each keyword is obtained by dividing the sum of proportions of keyword use by patenting scientists (column 2) by the count of the use of the keyword by non-patenting scientists (column 3).

Appendix II: Sensitivity Analysis

We present a sensitivity analysis for unmeasured confounding, following Robins (1999a) and Brumback et al. (2004). The basic approach is similar to that adopted by other researchers who have evaluated the robustness of non-experimental estimators in the presence of selectivity (see for example Altonji et al. (2005) for the case of Heckman two-step estimators and Rosenbaum (2002, chap. 4) for the propensity score). Namely, we attempt to quantify the amount of unmeasured confounding that would be necessary for our results to lose statistical significance. The exercise assumes a particular functional form for the bias from unmeasured confounding, generates new outcomes adjusted for the bias, and then performs IPTCW estimation by substituting these adjusted outcomes for the actual outcomes. For clarity, we begin by presenting the method in the case of a binary treatment with only two periods: baseline and end-of-study. The results are then generalized to the more relevant case of repeated treatments in a panel context.

Cross-sectional case. Let y be the outcome of interest (in our case, publications) at the end of the study, A be a binary treatment administered at baseline (*e.g.*, patenting), and V be a set of observables measured at baseline. Let y_i^a denote the outcome that person i would have experienced if he/she had, possibly contrary to the fact, received treatment a ($a = 0$ or $a = 1$). The method proposed by Robins (1999a) quantifies unmeasured confounding through the function Δ defined by:

$$\Delta(a, v) = E[y_i^a | A = a, V = v] - E[y_i^a | A = 1 - a, V = v]$$

Δ represents, among the set of individuals with baseline characteristics $V = v$, the average difference in potential outcomes between those treated with $A = a$ and those treated with $A = 1 - a$. For example, $\Delta(1, v)$ is the difference between the average outcome experienced by the treated when they are treated and the average outcome experienced by the untreated had they been treated instead. Note that the assumption of no unmeasured confounding is synonymous with $\Delta(a, v) = 0$ for all a and v. The effect of interest can be written $\theta = E[y_i^1 - y_i^0 | V = v]$. A naïve estimator for the treatment effect is:

$$\theta_{naive} = E[y | A = 1, V = v] - E[y | A = 0, V = v] \tag{AII$_1$}$$

The function Δ links θ and θ_{naive} through the identity

$$\theta = \theta_{naive} + \Delta(1, v) \cdot Prob[A = 0 | V = v] - \Delta(0, v) \cdot Prob[A = 1 | V = v] \tag{AII$_2$}$$

If $\Delta(1, v) > 0$ but $\Delta(0, v) < 0$, then on average, treated individuals will have higher potential outcomes to both treatment and no treatment than untreated individuals (*i.e.*, more talented scientists are more likely to patent). This seems to correspond to the type of bias that threatens the validity of the estimates presented in Section 4. Therefore we choose the following functional form for Δ:

$$\Delta(a, y_{i0}) = \alpha \cdot (2a - 1) \cdot y_{i0}$$

where α is a sensitivity parameter and y_{i0} denotes baseline outcome. With $\alpha > 0$, θ_{naive} is biased upward by $\alpha \cdot y_{i0}$. Robins (1999a) shows that for any user-specified α, one can obtain unbiased estimates of the treatment effect by replacing each observed outcome y_i with the adjusted outcome $y_i^\alpha = y_i - \Delta(A, V) \cdot Prob[1 - A | V]$ and recalculating θ_{naive} according to:

$$\bar{\theta}_{naive}^{\alpha} = E[y_i^\alpha | A = 1, V] - E[y_i^\alpha | A = 0, V] \tag{AII$_3$}$$

Intuitively, $\bar{\theta}_{naive}^{\alpha}$ subtracts the bias term in eqn. [AII$_2$] from θ_{naive}. To counteract bias from unmeasured confounding, we adjust the outcome of a treated individual downward by subtracting $\alpha \cdot y_{i0} \cdot Prob[A = 0 | V]$, and that of an untreated individual by adding $\alpha \cdot y_{i0} \cdot Prob[A = 1 | V]$.

Generalization to the panel case. The method of adjustment described above generalizes as follows for time-varying confounding. First, we compute the adjusted outcome

$$y_{it}^{\alpha} = y_{it} - \sum_{k=0}^{t-1} \Delta_{ik} \cdot Prob[1 - TREAT_{ik} | TREAT_{i,k-1}, \widetilde{Z}_{i,k-1}, \widetilde{X}_{i,k-1}] \tag{AII$_4$}$$

where $\Delta_{ik} = \alpha y_{ik}(2 \times TREAT_{ik} - 1)$. Second, we re-estimate the causal model with y_{it}^{α} in place of y_{it}. Subject to weak regularity conditions, the resulting augmented IPTCW estimator is consistent, assuming of course that the confounding function Δ is correctly specified. In what follows, we will focus on the case where treatment is specified as a "regime," *i.e.*, once the scientist has patented for the first time, the treatment variable never again reverts to 0. For such monotonic treatments, the sum above simplifies to

$$\alpha y_{im} Prob[TREAT_{im} = 0 | \widetilde{Z}_{i,m-1}, \widetilde{X}_{i,m-1}] - \alpha \sum_{k=0}^{m-1} y_{ik} Prob[TREAT_{ik} = 1 | \widetilde{Z}_{i,k-1}, \widetilde{X}_{i,k-1}]$$

for scientists who initiate patenting in period $m \leq t-1$, whereas for those initiating patenting at $m \geq t$, it remains

$$-\alpha \sum_{k=0}^{t-1} y_{ik} Prob[TREAT_{ik} = 1 | \widetilde{Z}_{i,k-1}, \widetilde{X}_{i,k-1}]$$

Choice of α and estimation details. From eqn. [AII_4], α is measured on the same scale as the dependent variable. We vary delta from 0.00 to 1.00 publications, in increments of 0.10. A priori, we regard negative values for α as implausible, since $\alpha < 0$ implies that non-patenters have higher potential outcomes regardless of whether they patent. One can think of increasing α as progressively stacking the deck against the finding of a positive effect of patenting on the rate of publication. $\alpha = 0$ corresponds to the estimate found in Table 5, Model 2c. The corresponding 95% confidence interval is $[0.137; 0.323]$.

IPTCW estimation proceeds in two steps, where estimates of the first step (from which the weights derive) are used as inputs in the second step. As Newey and McFadden (1994) have shown, the standard errors that result from the second step when the first-step estimates are assumed known instead of estimated will be biased. In the specific case of IPTCW estimation, Robins (1999b) has shown that the naïve asymptotic standard errors are upward-biased. However, this does not necessarily hold true for the augmented model. Since computing the standard errors analytically can be very cumbersome in these circumstances, we have used the block bootstrap (based on 500 replications) where we treat each scientist as a sampling unit, thus allowing for arbitrary serial correlation and heteroskedasticity. We account for the two-stage nature of IPTCW estimators by bootstraping the entire estimation procedure.

Results. As can be seen on Table A2, the threshold value of α such that the 95% confidence interval around the treatment effect falls just short of containing 0 is about 0.56. If one is willing to use a one-tailed test, the threshold value of α is larger, about 0.65. This might be warranted if one holds a strong prior regarding the direction of bias from unmeasured confounding.

Table A1: Sample Title Keywords in 1999

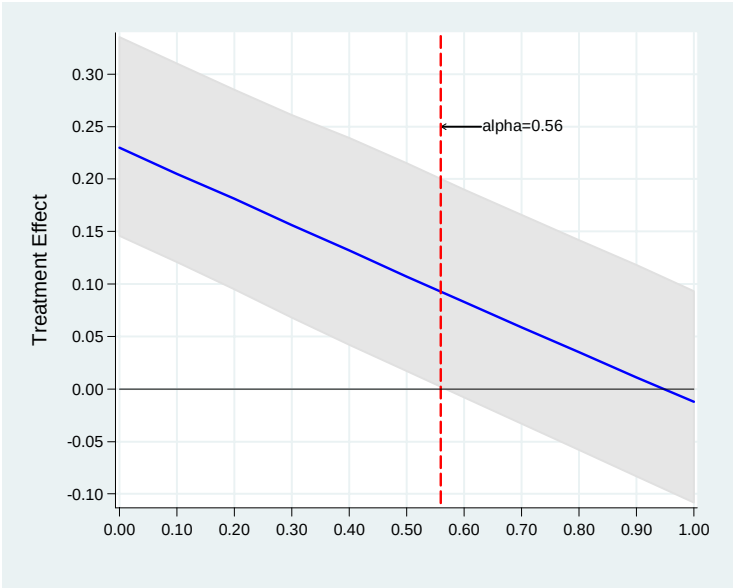
	(1)	(2)	(3)	(4)
	Number of times keyword used by patenting scientists	**Sum over all patenting scientists of keyword's proportion of total keywords used**	**Number of times keyword used by non-patenting scientists**	**Keyword weight: Column (2) / Column (3)**
	$\sum_{s \in I_t^p - \{i\}} m_{sjt}$	$\sum_{s \in I_t^p - \{i\}} \frac{m_{sjt}}{\sum_k m_{skt}}$	$\sum_{s \in I_t^{np} - \{i\}} m_{sjt}$	$\frac{\sum_{s \in I_t^p - \{i\}} \frac{m_{sjt}}{\sum_k m_{skt}}}{\sum_{s \in I_t^{np} - \{i\}} m_{sjt}}$
Group 1				
HIV-inhibitory	24	0.011	1	1.100
Ribozyme	32	0.074	15	0.493
Ubiquitin	55	0.145	30	0.483
Glycosylase	22	0.037	10	0.370
Aldose	36	0.059	16	0.369
Vitronectin	23	0.076	23	0.330
Glaucoma	30	0.069	25	0.276
Telomere	37	0.094	35	0.269
Melatonin	40	0.114	44	0.259
Lymphokine-activated	83	0.084	33	0.255
Spirochete	24	0.039	16	0.244
Coronavirus	28	0.066	28	0.236
Dendritic	43	0.178	83	0.214
E1A	37	0.066	32	0.206
Pheromone	66	0.119	58	0.205
Group 2				
Receptor	1161	2.270	4134	0.055
Antigen	494	1.094	1789	0.061
Antibody	425	1.043	1587	0.066
T-Cell	424	0.900	1242	0.072
Peptide	403	1.098	1511	0.073
Group 3				
Carnitine	1	0.0004	60	0.0007
Aromatase	1	0.0006	70	0.0009
Adenovirus-mediated	1	0.0004	37	0.001
Bismuth	1	0.0003	33	0.001
Endothelium-dependent	1	0.0007	51	0.001

Legend: To illustrate the construction of keyword weights, we have chosen representative words in three categories. Group 1 keywords are typical of those that appear frequently in the work of patenting scientists, and infrequently in the work of non-patenting scientists. These words receive high patentability weights. Group 2 comprises keywords that occur frequently in the journal articles of both patenting and non-patenting scientists. Words in this group garner intermediate weights. Group 3 contains keywords that are very common in the research of non-patenting scientists but uncommon in the work of patenters. In consequence, these keywords receive low weight.

Table A2
Sensitivity Analysis

α	*Estimate of the Treatment Effect*	*naïve SE*	*bootstrap SE*	*95% Confidence Interval (based on bootstrap SE)*	
0.00	0.230	0.047	0.046	0.146	0.335
0.10	0.205	0.048	0.047	0.121	0.310
0.20	0.181	0.048	0.047	0.095	0.285
0.30	0.156	0.048	0.047	0.068	0.261
0.40	0.132	0.049	0.048	0.042	0.239
0.50	0.107	0.049	0.048	0.017	0.215
0.60	0.083	0.049	0.048	-0.008	0.190
0.70	0.059	0.050	0.049	-0.033	0.166
0.80	0.035	0.050	0.049	-0.058	0.142
0.90	0.011	0.050	0.049	-0.083	0.118
1.00	-0.012	0.051	0.050	-0.108	0.093

Figure A2
Sensitivity Analysis



Effect estimates and 95% pointwise confidence bands are graphed versus the scale parameter $\boldsymbol{\alpha}$.