THE JOURNAL OF INDUSTRIAL ECONOMICS 0022-1821
Volume LVII December 2009 No. 4

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

Pierre Azoulay†
Waverly Ding‡
Toby Stuart§

We examine the influence of faculty patenting on the rate, quality, and content of public research outputs in a panel dataset of 3,862 academic life scientists. Using inverse probability of treatment weights (IPTW) to account for self-selection into patenting, we find that patenting has a positive effect on the rate of publications and a weak positive effect on the quality of these publications. 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 misplaced.

I. INTRODUCTION

In the past few decades, universities and other public-sector research organizations have proactively patented scientific discoveries (see Henderson *et al*. [1998]; Jaffe and Lerner [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 considerable controversy, much of which has centered on the long-term effect of patenting on the development of future scientific knowledge.

*We thank Rajeev Dehejia, Kei Hirano, David Levine and Scott Stern for useful comments, and seminar audiences at the NBER, Columbia University, the Israel Institute of Technology, INSEAD, UC-Berkeley, the University of Arizona, MIT, Duke, NYU, 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.

†Authors' affiliations: MIT and NBER, Sloan School of Management, 50 Memorial Drive, E52-555 Cambridge, Massachusetts 02142, U.S.A.
e-mail: pazoulay@mit.edu

‡University of California, Haas School of Business, 545 Student Services #1900, Berkeley, California 94720, U.S.A.
e-mail: wding@haas.berkeley.edu

§Harvard Business School, 211 Rock Center, Soldiers Field, Boston, Massachusetts 02163, U.S.A.
e-mail: tstuart@hbs.edu

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 such issues. 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 suggested 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 [2008] 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.* [2007] exploited a survey of doctorate recipients 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; if anything, we find that the average quality of patenters' post-patent publications may be slightly higher than that of non-patenters. 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 more frequently appear 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 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. 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.

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.* [2007]). 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 are able to generate an extensive list of covariates to model the probability of selection into patenting.

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

II. 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 is now a locus of commercially-oriented innovation. However, this development is not without detractors; many observers have decried the emergence of academic patenting and other forms of commercial science for its potentially adverse effects on the 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 academic science is rooted in peers' acknowledgment of important research advances, the up-or-out promotion system, and the intrinsic satisfaction of solving challenging problems (Merton [1973]). How does patenting influence these traditional incentives to produce academic research? 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?

On one hand, time-related considerations may cause patenting to reduce publishing: critics suggest 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, at the expense of investigating questions of basic research, faculty members devote a substantial block of time to conduct the research that leads to patentable discoveries. If not the act of patenting *per se* or even of producing patentable research, a third possibility is that consulting and other remunerative opportunities that are born out of patenting will divert away from basic research a patenting faculty member's time.

On the other hand, there are a few facts that may mitigate the likelihood of crowding out. First, scientists are assisted in the patent application process by their university's technology transfer office (TTO), whose existence enables a division of labor between invention and commercialization activities (Hellman [2007]). If TTO's function well, the act of filing for a patent may require a negligible amount of faculty time. Second, qualitative evidence suggests that many patent applications are direct 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' (also see Thursby *et al.* [2007]). In other words, patents and publications may pertain to a nearly identical set of research findings.

Moreover, it is possible that patenting and publishing actually are complementary activities. First, an academic researcher's scientific reputation may be his/her most important currency in the effort to capitalize on intellectual property in the market for university-originated technology. The market for university inventions is rife with asymmetric information. Academic discoveries often require years of additional development to yield marketable products; there can be great 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, scientists' reputations are essential in the marketplace 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, technology transfer offices (TTO's) should be more likely to choose to file for patents on the discoveries of high-status scientists. Therefore, the *ex post* search, screening, and contracting problems in the market for ideas may 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.

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. A likely consequence of applying for a patent is that academic scientists become acquainted with researchers in companies. As these relationships develop, industry contacts might become sources of ideas for new research projects. Indeed, Agrawal and Henderson's [2002] interviews with MIT scientists suggest that ties with industry research in fact do play a role in idea generation. Thus, relationships that arise post-patenting may become pathways for spillovers between academic and industry researchers.[1]

Knowledge is not the only input to the research process that may transcend the university-industry divide; pecuniary spillovers between patenting and publishing may exist as well. 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

[1] A natural analogy to this argument is the complementarities frequently 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.

additional post-doctoral scientists or graduate students might result in higher output for a scientist's lab.[2]

Patenting and the direction of scientific advance. Independent of the effect of patenting on the *rate* of scientists' output, a second question concerns the effect of academic patenting on the *content* of faculty members' research.

Patenting may be incidental to the content of faculty members' research if in fact patents are merely byproducts of research programs that scientists would have pursued even without the option to file for intellectual property protection. Assuming that patenting scientists neither change their behavior coincident to self-selecting into patenting, nor do they alter their research programs in any way post-patenting, patenting may have no effect on the content of post-patent research.

It is possible, however, that patenting will cause a within-person shift in the content of scientific research. 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 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 (and, for that matter, between basic and commercial science), Owen-Smith and Powell [2001] 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.

If this characterization is accurate, scientists who choose to patent and thereby shift into the commercialist camp will begin to allocate their research time across a 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

[2] 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.

for which part of the pay-off for conducting the research will be a patent or some other form of commercial recognition. While the majority of a scientist's work may not shift, some share of it may be devoted to new-to-the-scientist research questions. In this way, patenting may be associated with a shift in scientists' focus toward exploring scientific questions with commercial application.

A second and possibly more meaningful mechanism through which 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 gain exposure to new (relative to their previous work) areas of commercially useful scientific inquiry. Exposure to new and diverse information from bridging the university-industry divide may, in addition to enhancing scientists' productivity, cause academic scientists to become intrigued by questions of interest to industry researchers.[3]

III. 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 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. This will be the case if the labor market matches scientists with similar commercial proclivities in the same institutions, and there are peer effects in patenting. These types of

[3] 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.

effects seem hard to rule out *a priori*. In what follows, we will simply assume that a good instrument is not available.

A second approach is to rely on within-scientist variation to identify the effect of patenting on publication output. Fabrizio and DiMinin [2008] 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, . . ., T. At each time t, for each individual i, an outcome of interest y_{it} is measured, and a discrete treatment $PATENT_{it} \in \{0, 1\}$ is chosen. Also measured at each point in time are a set of exogenous covariates X_{it} and time-varying confounders Z_{it}. (X, Z) are 'prognostic factors' for patenting – in our case, proxies for the costs and benefits of patenting from the individual faculty member's point of view. For any variable W, denote $\widetilde{W}_{it}$ its history up to time t.

Let y_{it}^{a} be the value of y that would have been observed at time t had i chosen treatment sequence $\underset{\sim}{\tilde{a}_{it}} = (a_{i0}, a_{i1}, \ldots, a_{it})$ rather than his observed treatment history $PA\widetilde{TENT}_{it}$. Note that, even if $a_{i\tau}$ is dichotomous in each year τ, there will be 2^{T+1} patenting histories and thus 2^{T+1} possible counterfactuals, only one of which is observed for each individual.

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

$$E[y_{it}^{\tilde{a}} \Big| PA\widetilde{TENT}_{it}, X_{it}] = \beta_0 + \beta_1' X_{it} + \beta_2 \sum_{\tau=0}^{t} \omega_\tau PATENT_{i,t-\tau} \tag{1}$$

where the vector $\tilde{\omega}$ defines a distributed lag of patenting choices. For example, if $\omega_\tau = 1$ in each time period τ, then it is the *stock* of patents that influences publishing rates. Conversely, if $\omega_t = 1$ and $\omega_\tau = 0, \tau = 0, \ldots, t-1$, then only the instantaneous *flow* of patents has a causal effect on outcomes. In the empirical work, we will experiment with various specifications for the distributed lag of treatment histories.

Following Robins [1999], 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:

$$y_{it}^{a} \coprod PATENT_{it} | \, PA\widetilde{TENT}_{i,t-1}, \widetilde{Z}_{i,t-1}, \widetilde{X}_{it}$$

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

$$y_{it} = \beta_0 + \beta_1' X_{it} + \beta_2 \sum_{\tau=0}^{t} \omega_\tau PATENT_{i,t-\tau} + \varepsilon_{it} \tag{2}$$

by weighted least squares, where the weights correspond to the inverse probability of following actual treatment history $PA\widetilde{TENT}_{it}$ 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 $y^{\tilde{a}}$.

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

$$w_{it} = \frac{1}{\prod_{\tau=0}^{t} Prob(PATENT_{i\tau} | \widetilde{PATENT}_{i,\tau-1}, \widetilde{Z}_{i,\tau-1}, \widetilde{X}_{i\tau})}$$

Each factor in the denominator is the probability that the researcher received her own observed treatment at time τ, conditional on past patenting history and her past history of 'prognosis factors' for patenting, whether time-varying or fixed overtime. Therefore, the denominator of w_{it} represents the conditional probability that an individual followed his or her own history of patenting 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 patenting and the causal association between patenting and outcome is the same as in the original population. We refer to $\hat{\beta}_2$ when equation (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. But 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 patenting effect.

The probabilities $Prob(PATENT_{ik} | \widetilde{PATENT}_{i,k-1}, \widetilde{Z}_{i,k-1}, \widetilde{X}_{ik})$ may vary greatly between subjects when time-varying confounders are strongly associated with patenting. 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}:

$$sw_{it} = \prod_{\tau=0}^{t} \frac{Prob(PATENT_{i\tau} | \widetilde{PATENT}_{i,\tau-1}, \tilde{X}_{i\tau})}{Prob(PATENT_{i\tau} | \widetilde{PATENT}_{i,\tau-1}, \tilde{Z}_{i,\tau-1}, \tilde{X}_{i\tau})}$$

Although this modification does not influence the consistency of IPTW estimators, it does increase their efficiency (Hernán *et al.* [2000]).[4]

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 they receive attractive offers to join commercial firms. Even if patenting 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 in this way is tantamount to estimating the causal effect of $PATENT$ 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:

$$sw^*_{it} = \prod_{\tau=0}^{t} \frac{Prob(EXIT_{i\tau}|\, PA\widetilde{TENT}_{i,\tau-1}, X_{i\tau})}{Prob(EXIT_{i\tau}|\, PA\widetilde{TENT}_{i,\tau-1}, \tilde{Z}_{i,\tau-1}, X_{i\tau})}$$

where 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 patenting 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 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 how the patenting treatment is defined. According to a first definition, treatment is a *flow*: $PATENT_{it} = 1$ whenever researcher i applies for at least one patent in year t, and 0 otherwise. This formulation implies that patenting does not necessarily have a lasting impact on the individual. In a second approach, the *regime* formulation defines $PATENT_{it} = 1$ for all years

[4] 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.

subsequent to the first patent application. Defining treatment this way implies a one-time shift in the outcome of interest, with subsequent patenting choices 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:

$$(3) \qquad Prob(PATENT_{it} = 1) = \alpha_0 + \alpha_1 PATENT_{i,t-1} + \alpha_2 Z_{i,t-1} + \alpha_3 \sum_{\tau=0}^{t-2} Z_{i\tau} + \alpha_4 X_{it} + \delta_t$$

where X_{it} includes exogenous 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. The effect of each time-varying confounder Z is modeled through the additive combination of a term for the one-year lagged value of the variable and a cumulative stock of the variables for the years $0,\ldots,t-2$. In practice, the vector Z includes publications, the number of past collaborations with industrial firms, a measure of the inherent patentability of the scientist's publications, along with various employer characteristics. Let T_1 denote the set of years in which scientist i gets at least one patent and T_2 the set of years during which i gets no patents. The estimate of the denominator of sw_{it} is $\prod_{t\in T_1} \hat{p}_{it} \prod_{t\in T_2} (1-\hat{p}_{it})$, where $\hat{p}_{it}$ refers to the predicted probability obtained after estimating equation (3). The numerator of sw_{it} stems from an almost identical specification, except that one omits the time-varying confounders Z from the list of covariates.

This approach must be modified when patenting is modeled as a regime shift rather than as a flow, because the probability of patenting 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, $PATENT_{i,t-1}$ is uniformly 0. To compute the denominator of sw_{it} we estimate a logit model for

$$(4) \qquad Prob(PATENT_{it} = 1) = \alpha_0 + \alpha_2 Z_{i,t-1} + \alpha_3 \sum_{\tau=0}^{t-2} Z_{i\tau} + \alpha_4 X_{it} + \delta_t$$

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

$$(5) \qquad Prob(PATENT_{it} = 1) = \alpha_0 + \alpha_4 X_{it} + \delta_t$$

Our estimate of the denominator of sw_{it} for scientist i in year t is $\prod_{\tau=0}^{t} (1-\hat{p}_{i\tau})$ if scientist i did not apply for at least one patent by year t, and $\prod_{\tau=0}^{t-1} (1-\hat{p}_{i\tau}) \times$

$\hat{p}_{it}$ if scientist i applied for his first patent in year t. Estimation of sw^{*}_{it} proceeds in the same fashion.[5]

Sequential Condional Independence: An Econometric Free Lunch? Like propensity score estimation, IPTCW relies on the assumption that selection into treatment solely occurs on the basis of factors observed by the econometrician. Moreover, IPTW estimates are just identified: the assumption of no unobserved determinants of selection cannot be tested. This will appear to many readers as a strong assumption – one that is unlikely to be literally true. Below, we provide a number of reasons why, despite the strength of the assumption, we consider the IPTCW estimates to be reliable in this case.

Past research in the program evaluation literature has shown that techniques that 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 Wahba [2002]; Smith and Todd [2005]). Conditions (2) and (3) are trivially satisfied here, but one might wonder about condition (1), as to the extent to which we account for the determinants of selection into patenting.

For an individual faculty member, the costs of patenting will depend on her employment and training context, including the level of support of her employer for academic entrepreneurship, and the inherent patentability of her field of research. The benefits might include one's taste for additional income, and other factors that are better thought of as outcomes, such as access to new research partners and additional resources, including industry funding and capital equipment. For most samples, determinants of this nature typically would be unobserved by the econometrician. However, we have chosen a study population with the specific goal of statisifying the selection on observables assumption required for IPTCW. In particular, we analyze a sample for which extensive archival information is available and in

[5] It merits note that Rosenbaum and Rubin [1983] refer to $Prob(PATENT_i = 1|X, Z)$ as the propensity score. Recently, Heckman *et al.* [1997] have combined the propensity score with difference-in-differences to estimate the causal effect of treatment and Abadie [2005] proposes a semiparametric difference-in-differences estimator that weights observations by the inverse probability of (own) treatment. 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 our setting and many others, however, selection into treatment can occur at different times and/or in several disjoint episodes. Matching on the propensity score is difficult in these cases because an untreated individual might be a good control for a treated subject in one period (i.e., the difference in their propensity scores is near 0) and a bad control for the same treated subject in another period. The advantage of IPTCW estimation is that it readily generalizes to the case of treatments that are staggered over time.

which we have invested a great deal of effort construct proxies for these determinants of selection into patenting. For example, we develop a keyword-based measure of the inherent patentability of very fine-grained fields of research, where previous analysts have relied on fixed effects for broad scientific fields (e.g., cell biology, pharmacology, etc.).

Despite this investment, the conclusion that our list of proxies for the determinants of selection is exhaustive is premature without further evidence. For instance, we do not observe levels of industry funding. While our qualitative evidence indicates that instances of funding from commercial sponsors seldom precede a patent application, it would obviously be better to control for such variation. Likewise, we can include in our specification of the propensity score the royalty rate on patent licensing income that each university pays to faculty members. This, along with other institutional characteristics, will capture variation in the *average* scientist's taste for income across universities if the scientific labor market sorts faculty members with similar entrepreneurial proclivities into the same institutions. But within an institution, there is likely to remain substantial variation among faculty members at a given point of time.

We are able to address concerns pertaining to residual selection on unobservables in three ways. First, in Appendix II, we provide detailed evidence that patenting is neither associated with the covariates that give the technique its empirical purchase in the weighted sample, nor is it associated with 'unused' covariates (variables that we did not include in the specification of the propensity score because they were made redundant by other variables we did include). In other words, we show graphically and in econometric tests that there are non-patenting faculty members who closely match the characteristics of patenting scientists, even when our model of selection predicts that they are either very likely or very unlikely to patent.

Second, we perform a sensitivity analysis. 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), residual heterogeneity likely leads IPTCW to overestimate the treatment effect. Our sensitivity analysis estimates the amount of (unobserved) heterogeneity that would be required for the effect to lose statistical significance, and shows that it is very high (these results are presented in the online appendix on the *Journal*'s editorial web site).

Third, we contrast the IPTCW estimates with fixed-effects estimates. Since patenting scientists have higher output (relative to their average in the pre-patenting regime) in the year immediately preceding their first patent application, the fixed effect estimator is likely to underestimate the effect of patenting on publishing rates. In combination, however, these two 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 most cases, these bounds are sufficiently tight to inform the policy debate surrounding academic patenting.

Yet, our conclusions must remain guarded because we cannot exploit true exogenous variation in the costs and benefits of patenting to identify the effect of interest. Whereas the precise magnitude of the effect of patenting remains an open question, at the very least, taken in their entirety, we believe that our results pin down its sign.

IV. DATA, SAMPLE CHARACTERISTICS, AND RESULTS

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 (see Table 1).

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,[7] 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

[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 Commission. We then stratified the random draw from UMI to correspond to the disciplines and Ph.D. years of firms' scientific advisors. For example, 22 per cent of biotechnology company scientific advisors hold biochemistry Ph.D.'s; we drew a corresponding proportion of biochemists into our sample. Table I lists the Top 15 disciplines from which scientists in our sample are selected.

[7] 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.

earned a Ph.D. until 1999, unless the individual exited academia.[8] The final panel contains 58,562 person-year observations between 1968 and 1999.

IV(i). *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).[9] 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 discovery to potential licensing partners in industry. According to Lach and Schankerman [2004], the average royalty rate among U.S. universities is about 40%, although many universities use non-linear schedules.

Research outputs. From the *Web of Science* we computed annual paper publication counts for each scientist. We count equally all (solo and coauthored) papers on which a scientist is listed as an author. 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 track, for each journal in which our scientists published, the relative prevalence of authors with corporate affiliations.[10] In particular, for each scientist-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.

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

[8] 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.

[9] Faculty members are contractually obligated to disclose potentially commercializable discoveries developed on university premises to the TTO; in theory, they do not have the option to patent university-originated discoveries without going through the official channels. The average TTO received 78 invention disclosures in 2003, but filed only 40 new patent applications (Stevens and Toneguzzo [2003]).

[10] 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*.

choice is motivated by a robust social norm in the life sciences that 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 is taken to be a measure of quality for the average article published by one of our scientists in a given year.

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 (details are presented in Appendix III). 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 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, so we include a 'scientist is female' indicator 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.

Institutional context has been shown to affect patenting both through the funding and experience of the technology licensing office, as well as through the presence of prominent peers who themselves are engaged in this activity (Di Gregorio and Shane [2003]; 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 is 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 to hold constant an individual scientist's research agendum. In contrast, we attempt to measure patentability directly by using the title words in scientists' publications to identify the areas in which they have conducted research, and then applying 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 (RP) score for scientist i in year t is defined as:

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

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,[11] n_{ijt} is the

[11] We used 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

number of times each of the keywords j has appeared in scientist i's articles published in year t, and ω^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 ω^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 ratio $\frac{n_{ijt}}{\sum_k n_{ikt}}$; the latter by the weights $\omega^i_{j,t-1}$. Because the benchmark in year $t-1$ is used to weight title words in year t, year-to-year changes in there search patentability score will only reflect actions of the scientist (through their choices of title keywords), rather than contemporaneous changes in the benchmark.

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:[12]

$$\begin{aligned} STOCK\_RP_{it} &= (1-\delta)STOCK\_RP_{i,t-1} + FLOW\_RP_{it} \\ &= \sum_{\tau=0}^{t} (1-\delta)^{t-\tau} \cdot FLOW\_RP_{i\tau} \end{aligned}$$

IV(ii). *Descriptive Statistics*

Of 3,862 scientists, we found 473 (12.2%) patenters who were listed on 1,372 patents. 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. An example is Richard J. Lagow, who has held professorships at MIT and the University of Texas Austin. Lagow

methodology used in the paper. We find high overlap between title words and keywords in the papers for which both are available.

[12] 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.

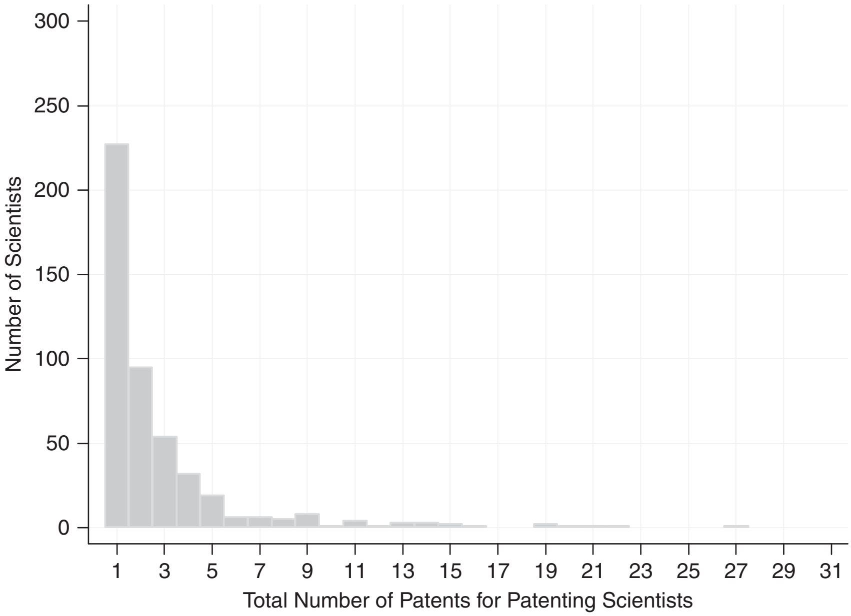

Figure 1
Distribution of Patent Count for Patenting Scientists

began patenting in 1973 and his patents have been assigned to MIT, University of Texas, and Exfluor Research Corporation. Among 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.[13]

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 down by their patenting status. Consistent with past findings that patenting is concentrated among the group of

[13] 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. This is consistent with results reported by Thursby *et al*. [2009].

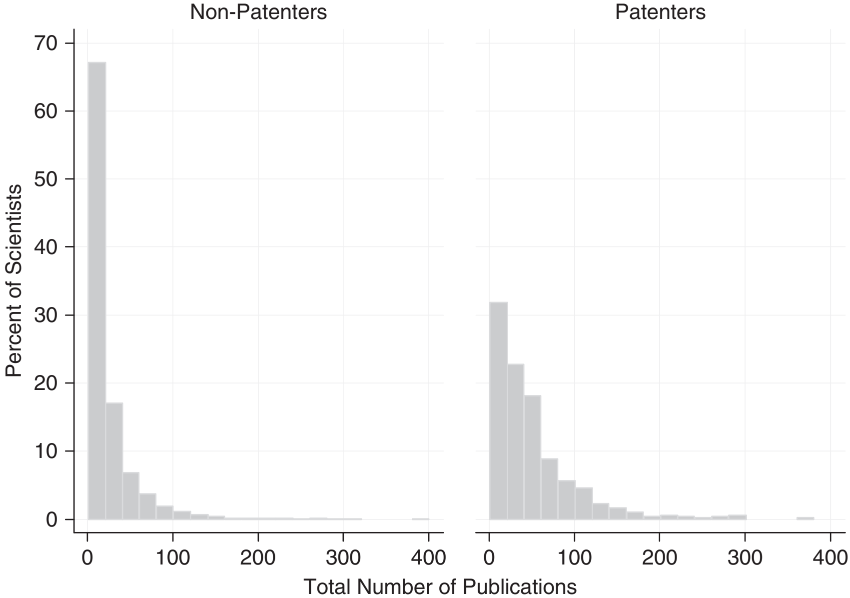

Figure 2
Distribution of Publication Count for Patenting and Non-Patenting Scientists

TABLE I
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%)

Table I 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.

academically productive scientists, the distribution for the patenter subsample is much less skewed than that for the non-patenter subsample.

Table II presents the summary descriptive statistics for variables used in our analysis. Table III reports, by scientists' patenting status, the mean

TABLE II
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
Number of Pubs. with Industry Coauthors (Flow)	0.141	0.552	0	13	58,562
Number of Pubs. with Industry Coauthors (Stock)	1.206	4.242	0	135	58,562
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
Employer royalty rate	0.452	0.056	0.150	0.969	9,455
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

research and employer characteristics measured at five career stages. Patenters are more productive at each career stage: they publish more research papers than those yet to have entered the patenting regime, and the papers they produce appear to be of marginally higher quality (as captured by average JIF). Scientists with patents are closer to commercial research than their non-patenting counterparts: they collaborate more with researchers in the private sector and the intrinsic patentability of their research is 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 III.

VI(iii). *Results*

We present four sets of results. Table IV 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 variables, 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 V), quality (Table VI), and direction (Table VII) of scientific output.

TABLE III

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

Table III 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.

**indicates that two-sample t test suggests the mean values of the patenters and non-patenters are different at 1% significant level;

*indicates difference significant at 5% level;

†indicates significant at 10% level.

TABLE IV
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]	0.128	0.195	0.154	0.239		
	(0.153)	(0.153)	(0.166)	(0.164)		
Experience = [9, 15]	0.203	0.347	0.291	0.432	0.215	− 0.006
	(0.155)	(0.151)*	(0.167)†	(0.162)**	(0.060)**	(0.057)
Experience = [16, 22]	0.003	0.218	0.230	0.401	0.127	− 0.264
	(0.174)	(0.162)	(0.196)	(0.180)*	(0.086)	(0.077)**
Experience = [23, 35]	− 0.393	− 0.097	− 0.386	− 0.232	0.386	− 0.122
	(0.215)†	(0.198)	(0.280)	(0.267)	(0.116)**	(0.101)
Female	− 0.641	− 0.675	− 0.663	− 0.700	0.146	0.243
	(0.132)**	(0.133)**	(0.153)**	(0.152)**	(0.054)**	(0.053)**
Patent Flow$_{t-1}$	1.946	2.048			0.288	
	(0.124)**	(0.128)**			(0.172)†	
Patent Stock$_{t-2}$	1.982	2.065			− 0.132	
	(0.092)**	(0.093)**			(0.102)	
Publications Flow$_{t-1}$	0.040		0.073		− 0.202	
	(0.017)*		(0.023)**		(0.031)**	
Publications Stock$_{t-2}$	0.004		0.002		− 0.013	
	(0.002)*		(0.003)		(0.003)**	
Patentability Flow$_{t-1}$	0.866		0.849		0.425	
	(0.289)**		(0.300)**		(0.290)	
Patentability Stock$_{t-2}$	0.216		0.344		0.005	
	(0.287)		(0.282)		(0.201)	
Pub Flow with Industry Coauth.$_{t-1}$	0.069		0.158		0.181	
	(0.056)		(0.070)*		(0.066)**	
Pub Stock with Industry Coauth.$_{t-2}$	− 0.018		− 0.026		− 0.002	
	(0.011)†		(0.019)		(0.010)	
Top 20 Employer	0.147		− 0.006		0.053	
	(0.113)		(0.119)		(0.059)	
TTO$_{t-1}$	0.133		0.018		− 0.049	
	(0.097)		(0.118)		(0.053)	
Employer Patent $_{t-1}$	− 0.007		0.088		0.031	
	(0.025)		(0.033)**		(0.016)†	
Top 20 Ph.D.	0.007	0.053	0.086	0.121	− 0.148	− 0.181
	(0.091)	(0.089)	(0.104)	(0.104)	(0.053)**	(0.053)**
Ph.D. Univ. Patent	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	− 6.114	− 5.968	− 6.144	− 6.039	− 4.346	− 4.533
	(0.291)**	(0.300)**	(0.302)**	(0.302)**	(0.139)**	(0.139)**
Observations	58,562		54,746		58,437	
Number of Researchers	3,862		3,862		3,862	
Log Pseudo-Likelihood	− 3,956.73	− 3,994.80	− 2,548.92	− 2,578.29	− 8,872.04	− 9,092.91
Wald χ^2	2,249.82	2,089.54	347.37	272.91	595.86	308.91
Number Of Variables	50	40	48	38	49	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; models 1a, 2a and 3a also control for zero output in lagged publication flow.
(3) Robust standard errors in parentheses, clustered around individual researchers.
(4)†significant at 10%; *significant at 5%; **significant at 1%.

TABLE V
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	0.208	0.161	0.200	0.205	0.162	0.206	0.208
	(0.018)**	(0.018)**	(0.018)**	(0.018)**	(0.018)**	(0.018)**	(0.018)**	(0.018)**	(0.018)**
Experience = [9, 15]	0.260	0.445	0.422	0.262	0.430	0.414	0.263	0.447	0.419
	(0.029)**	(0.030)**	(0.032)**	(0.029)**	(0.030)**	(0.032)**	(0.028)**	(0.030)**	(0.032)**
Experience = [16, 22]	0.229	0.554	0.438	0.228	0.521	0.425	0.229	0.548	0.425
	(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	0.340	0.085	0.487	0.331	0.082	0.494	0.325
	(0.050)†	(0.073)**	(0.070)**	(0.050)†	(0.073)**	(0.071)**	(0.050)†	(0.073)**	(0.070)**
Female		– 0.215	– 0.228		– 0.203	– 0.224		– 0.216	– 0.224
		(0.052)**	(0.048)**		(0.051)**	(0.049)**		(0.052)**	(0.048)**
Top 20 Ph.D. Univ.		0.067	0.058		0.063	0.056		0.070	0.056
		(0.042)	(0.040)		(0.042)	(0.041)		(0.042)†	(0.040)
Ph.D. Univ. Patent		0.046	0.051		0.043	0.051		0.047	0.050
		(0.048)	(0.047)		(0.047)	(0.047)		(0.048)	(0.047)
Patent Flow	0.165	0.539	0.294						
	(0.028)**	(0.057)**	(0.058)**						
Patent Regime				0.195	0.394	0.235			
				(0.031)**	(0.048)**	(0.047)**			
Patent Stock							0.016	0.045	0.054
							(0.010)	(0.012)**	(0.012)**
Constant		0.033	0.037		0.034	0.039		0.037	0.039
		(0.044)	(0.044)		(0.044)	(0.045)		(0.044)	(0.044)
Log pseudo-likelihood	– 78,109.5	– 120,275.9	– 116,680.6	– 78,070.0	– 119,953.1	– 116,764.9	– 78,126.2	– 120,266.5	– 116,594.7
Wald χ^2	2,897.67	1,234.53	968.43	2,966.37	1,301.65	949.21	2,858.59	1,228.92	972.35
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 VI
EFFECT OF PATENTING ON THE QUALITY OF PUBLICATIONS

	Model 1a	Model 1b	Model 2a	Model 2b
	Fractional Logit QML Proportion of First or Last-Authored Publications		Poisson Model QML Average JIF of Publications	
	Unweighted	IPTCW	Unweighted	IPTCW
Experience = [5, 8]	– 0.096	– 0.096	– 0.087	– 0.088
	(0.029)**	(0.029)**	(0.013)**	(0.013)**
Experience = [9, 15]	0.034	0.028	– 0.189	– 0.186
	(0.034)	(0.034)	(0.018)**	(0.018)**
Experience = [16, 22]	0.133	0.122	– 0.273	– 0.276
	(0.046)**	(0.046)**	(0.027)**	(0.027)**
Experience = [23, 32]	0.155	0.139	– 0.354	– 0.368
	(0.068)*	(0.070)*	(0.039)**	(0.040)**
Female	– 0.003	– 0.002	0.031	0.034
	(0.038)	(0.038)	(0.022)	(0.022)
Top 20 Ph.D. Univ.	0.050	0.048	0.135	0.131
	(0.033)	(0.033)	(0.021)**	(0.021)**
Ph.D. Univ.	0.049	0.044	0.086	0.092
Patent × 100	(0.042)	(0.043)	(0.030)**	(0.029)**
Patent Regime	0.026	0.003	0.077	0.053
	(0.048)	(0.051)	(0.029)**	(0.030)†
Constant	0.826	0.827	1.370	1.372
	(0.047)**	(0.047)**	(0.023)**	(0.023)**
Log pseudo-likelihood	– 22,238.9	– 21,803.9	– 91,867.7	– 90,034.1
Wald χ^2	272.6	270.7	642.1	678.1

Notes: (1) Number of observations = 38,007; number of researchers = 3,862; number of variables = 39.
(2) All models control for Ph.D. 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%.

Determinants of patenting activity. In Table IV, we begin by presenting results pertaining to the probability of applying for a patent in a given year (flow formulation, Models 1a and 1b) or for the first time (regime formulation, Models 2a and 2b). In each case, the first column includes time-varying confounders on the right hand side, whereas the second column excludes them. This split is important for the follow-on econometric exercise, since the denominator (respectively numerator) of the weights used in IPTCW estimation corresponds to a product of predicted probabilities based on estimates of the first (respectively second) column.[14]

[14] Recall that the list of independent variables and the risk set differ 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 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.

TABLE VII
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		Poisson Estimates Number of Pub. with Industry Coauthors		Fractional Logit QML Estimates Average Journal Commercial Score	
	Unweighted	IPTCW	Unweighted	IPTCW	Unweighted	IPTCW
Experience = [5, 8]	0.008	0.006	0.278	0.289	0.016	0.015
	(0.039)	(0.039)	(0.066)**	(0.071)**	(0.014)	(0.014)
Experience = [9, 15]	– 0.025	– 0.024	0.534	0.512	0.006	0.006
	(0.038)	(0.037)	(0.092**)	(0.095)**	(0.019)	(0.019)
Experience = [16, 22]	– 0.054	– 0.057	0.636	0.578	0.015	0.021
	(0.038)	(0.038)	(0.123)**	(0.123)**	(0.025)	(0.025)
Experience = [23, 32]	– 0.103	– 0.104	0.593	0.468	0.057	0.076
	(0.042)**	(0.043)*	(0.176)**	(0.186)*	(0.035)	(0.035)*
Female	– 0.023	– 0.023	– 0.263	– 0.323	– 0.007	– 0.005
	(0.022)	(0.023)	(0.125)*	(0.108)**	(0.017)	(0.017)
Top 20 Ph.D. Univ.	– 0.027	– 0.024	– 0.162	– 0.178	– 0.069	– 0.067
	(0.021)	(0.022)	(0.094)†	(0.094)†	(0.018)**	(0.018)**
Ph.D. Univ.	– 0.017	– 0.019	0.183	0.216	– 0.018	– 0.016
Patent × 100	(0.020)	(0.020)	(0.130)	(0.134)	(0.025)	(0.026)
Patent Regime	0.090	0.083	0.528	0.422	0.043	0.051
	(0.028)**	(0.029)**	(0.094)**	(0.096)**	(0.024)†	(0.025)*
Constant	– 5.700	– 5.693	– 3.845	– 3.855	– 2.491	– 2.494
	(0.353)**	(0.350)**	(0.160)**	(0.160)**	(0.024)**	(0.024)**
Log pseudo-likelihood	– 4,887.3	– 4,738.4	– 25,645.7	– 24,457.1	– 7,669.4	– 7,508.3
Wald χ^2	2,089.6	1,943.8	656.05	530.37	431.53	391.01

Notes: (1) Number of observations for models 1 and 3 = 38,007; number of observations for model 2 = 58,562; number of researchers = 3,862; number of variables = 39.
(2) All models control for Ph.D. 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%.

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 9.98% for the flow specification (column 1a) and by 19.97% for the regime specification (column 2a).[15]

This conditional correlation strikes us as being an important finding because it helps to distinguish competing interpretations of the association

[15] In a companion paper (Azoulay *et al.* [2007]), we confirm that this result is robust to much more flexible specifications of the lag structure.

between scientific productivity and involvement with the world of commerce. In one interpretation, commercialization activities correspond to attempts by academics to monetize established reputations and professional status. In the other 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.[16] The plausibility of this interpretation is reinforced by the fact that U.S. patent law 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.[17]

We also find that 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 standard deviation increase in research patentability score raises the probability of patenting by 4.3% (column 1a) and by 4.2% (column 2a).[18] 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 patents. But because it is the flow, and not the stock of this measure that matters, the evidence suggests that a patent application does not just constitute a response to changes in incentives faced by academic scientists over their careers, but also reflects the seizing of opportunities along a novel research trajectory. This interpretation is further supported by the effect of collaborations with industry: in the regime formulation (column 2a), the one-year lagged value shifts the probability of patenting upwards, whereas earlier collaborations do not seem to have much influence.

Our models for the probability of patenting are well-suited to techniques that rely on selection of observables, because the region of common support

[16] This interpretation is also consistent with Murray and Stern's [2007] analysis of paper-patent pairs and it suggests that this phenomenon is broader than the single journal 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.

[17] This result contradicts 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.

[18] 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.* [2007]).

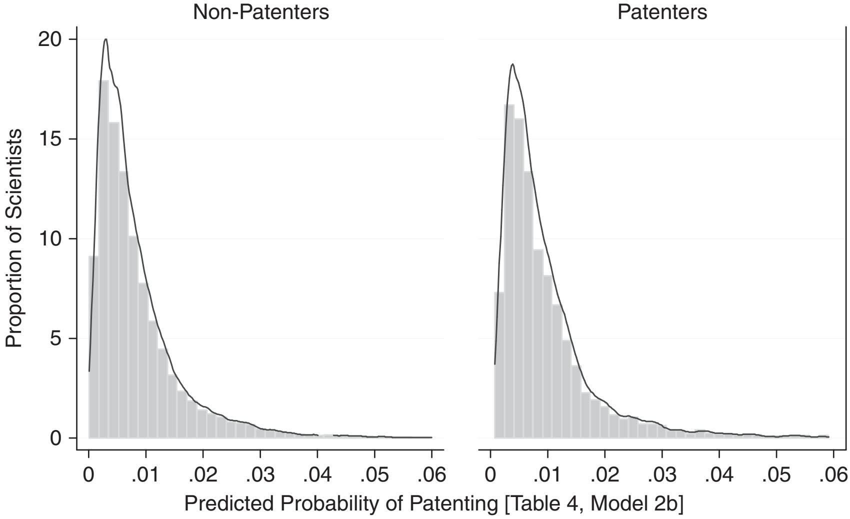

Figure 3
Common Support for the Propensity Score

includes almost all obervations. Figure 3 displays histograms of the predicted probabilities of patenting implied by the estimates in column 2b, for patenters and for non-patenters. The shapes of the two distributions are virtually indistinguishable, which means that one can find non-patenting counterparts to patenting scientists, even for those individuals whose predicted likelihood of patenting is very high or very low.

Determinants of exit from academia. Models 3a and 3b display the results corresponding to specifications modeling the probability of exiting from academia. *A priori*, one might imagine that academic scientists leave academia because they have poor publication records. 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 33.4% 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 V 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 of 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).[19] 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 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 V yields three robust results: (a) the estimated effect of patenting is positive and statistically significant in 8 out of 9 specifications; (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.

To provide direct evidence of the impact of pre-existing trends in the fixed effects model, we estimate a slight variant of Model 2a in which the treatment effect is interacted with a series of 10 indicator variables corresponding to 4 years before first patenting event; 3 years before first patenting event; . . .; 4 years after first patenting event; and 5 years or more after first patenting event. In Figure 4, we graph the incidence rate ratios corresponding to each interaction (solid line), along with the 95$^{\text{th}}$ confidence interval around the estimates (dotted lines). Even in years $t-4$ and $t-3$, we can detect an uptick in output, although the magnitudes are much smaller than in the post-patenting period. The effect becomes statistically and substantially significant in year $t-1$ and gradually increases in magnitude until year $t+3$, declining slightly thereafter. In light of these results, 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 these transitory dynamics.

[19] 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 the PQML estimator can be used for any non-negative dependent variables, whether integer or continuous (see Santos Silva and Tenreyro [2007]).

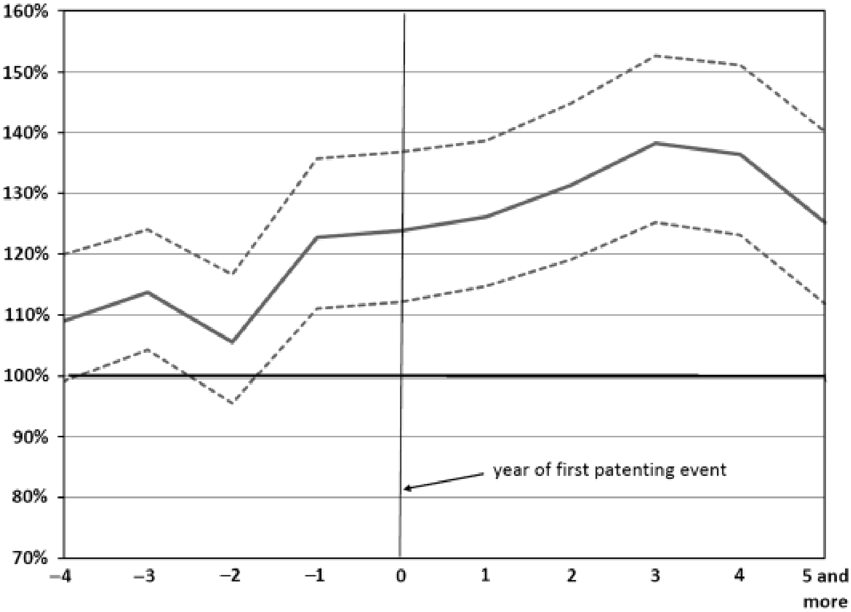

Figure 4
Preexisting Trends in Publication Output and Timing of First Patent Application

Effect of patenting on the quality of publication output. Table VI examines two measures of publication quality.[20] 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 versions of the model, and is statistically insignificant in both cases. This suggests that patenting has 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 V. 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 V does not come at the expense of the quality of these publications.

[20] 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.

Effect of patenting on the content of publication output. Measuring changes in the direction of scientific research is more challenging than merely assessing the quantity of output. In Table VII, we employ three distinct measures of 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 IV as the dependent variable, and we perform estimation using the Poisson QML estimator in columns 1a and 1b (since *RP* is a non-negative 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 1b imply that entering the patenting regime increases *RP* by a statistically significant 8.7%.

Models 2a and 2b provide a different angle on the same question by focusing on the institutional affiliations of scientists' coauthors. In the years of positive output, we compute the fraction 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 52.5%. 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, Lim [2004] computes the scientist-specific score by averaging these weights over all articles published in a given year.[21] 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. Patenting increases the rate of scientific output while (at worst) maintaining its quality, but may also modestly shift the content of these publications toward questions of commercial interest.

IV(iv). *Robustness Checks*

If the fixed effect specifications understate the causal effect of patenting, but the IPTCW specifications overstate it, Models 2a and 2c in Table V imply

[21] 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.

that the average treatment effect of patenting on publication rates, expressed as an elasticity, lies within the interval [0.215; 0.265]. For robustness purposes, we consider alternative possibilities that may draw into question the accuracy of both endpoints of this range.

Robustness of IPTCW estimates. In the absence of an instrument, it is important to gauge the sensitivity of the IPTCW results to a failure of our assumption that selection into patenting can be accurately captured by observable factors. We do so in two different ways. First, we report the results of formal covariate balance tests in Appendix II. The estimates show that IPTCW estimation balances the list of covariates that determine selection into patenting. Second, we ask how much selection on unobservables would there need to be for the confidence interval around our treatment effect to include 0? This approach requires us to parameterize the bias from unobserved confounding – a functional form choice that must be guided by 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 selection on unobservables. This is a multi-step analysis so estimation details and results are provided in the online appendix.

We reemphasize that none of these tests provides a formal test of sequential conditional independence. However, we can think of them contrapositively: had they been inconclusive, they would have cast doubt on the statistical approach we followed. Taken together, they are pieces of circumstantial evidence that help make the case for the usefulness of the IPTCW approach, especially in combination with fixed effects estimates.

Do the fixed-effects estimates bound the treatment effect from below? Suppose that 'big ideas' lead to both patents and publications, but patenting has no causal influence on public research outputs. The 'big ideas' hypothesis is consistent with the IPTCW results insofar as the observables (e.g., the pre-patent publication flurry, and the research patentability score) do not fully capture the importance of the idea. It is also consistent with the fixed effects results under a plausible alternative scenario, which is that most of the publications associated with the big idea post-date the filing of the patent application. If the patent application is coincident in timing to the occurrence of the big idea and most of the associated publishing takes place later, the fixed effects estimate could overstate, rather than understate, the treatment effect.

We therefore conduct a range of checks pertaining to the fixed effects specification of Tables V, Model 2a. Recall our argument: a 'big idea' generates a flurry of publications before a patent application. Patenting then has effects that go beyond the scientific payoffs associated with the initial idea, because the very act of patenting enables the scientist to attract/digest ideas and resources from industry. This timing is quite critical to our argument, and we believe it to be consistent with the data. First, each patent application is preceded by a flurry of publications in the year that precedes it; in contrast, no flurry can be observed in the year that precedes the *grant* of the first (or subsequent) patent.[22] This provides circumstantial evidence that the patent application in year t and the publication flurry in year $t-1$ correspond to the same set of scientific insights. Second, this result resonates with the experience and practice of the technology licensing officers with whom we have discussed it. Patent lawyers seem to be keenly aware of the one-year grace period in U.S. patent law. This also explains why very little evidence of publication delay associated with patenting has emerged to date. Nonetheless, we cannot be sure that the publications associated with the patented idea always precede the patent application itself.

To assess the robustness of the fixed effects estimates to departures from the timing between publications and patents, we conducted a simulation exercise (500 replications). With probability .5, we pushed forward by one year the date of first patent application; with probability .25, we pushed it forward by two years; with probability .25 we retained the original timing.

Our justification for this manipulation is as follows: if patenting is incidental to the birth of an important idea, but has no impact on the subsequent elaboration of this idea or the onset of subsequent scientific ideas, then its effect should disappear – or markedly attenuate – after the publications of the papers associated with the original idea. In the simulated data that randomly moves forward the timing of a patent application, the estimate of the treatment effect is .198 (*s.e.* = .040), versus .215 (*s.e.* = .039) in the original data. If the alternative sequence of events were correct, pushing the patent forward in time should locate it beyond the bump in scientific output associated with the idea. Our simulated results, however, contradict this explanation. Because the estimated treatment effect declines only slightly, the simulated results lead us to conclude that the effect of patenting outlasts the initial idea that led to the patent application.

V. DISCUSSION AND CONCLUSION

We find that academic scientists who patent produce more public scientific outputs than do otherwise equivalent non-patenters. Across three different measures of the content of scientific research, we also find that following patenting, scientists appear to modestly shift the focus of their research toward research topics with commercial application. These results depend on the maintained 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. It is obviously

[22] These results are available from the authors upon request.

better to include a large set of potential confounders to model the probability of selection, but we recognize that in practice, the Sequential Independence Assumption may still not be precisely or even approximately tenable. To buttress the credibility of these estimates, we perform a battery of robustness checks, including covariate balance tests and a formal sensitivity analysis. While not fully dispositive, these checks ensure that our estimates are not overly sensitive to specification choices. Whereas the precise magnitude of the effects of patenting remains an open question, at the very least, taken in their entirety, our results pin down the sign of these effects.

There are two other avenues 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 [2007], who provide evidence that scientific papers paired with patent applications are less likely to be cited after the patent is granted by the U.S. Patent Office (though the effect is modest in magnitude).

Academic patenting might also alter the career trajectories of the graduate students and post-doctoral fellows who 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.

APPENDIX I: KEYWORD WEIGHTS

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

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

where m_{sjt} denotes the number of times keyword *j* has appeared in articles published up to year *t* by scientist *s*, I_t^p is the subset of scientists in our sample that have already applied for one or more patents as of year *t*, and I_t^{np} 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 ω_{jt}^i, 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, $m_{sjt}/\Sigma_k m_{skt}$, 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 ω_{jt}^i 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, $\omega_{jt}^i = 0$ 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 ω_{jt}^i. 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 $\omega_{j,t-1}^i$, 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 AI 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: COVARIATE BALANCE TESTS

In the absence of genuine exogenous variation in the costs of patenting across the scientists in our sample, it is important to gauge the sensitivity of the IPTCW results to a failure of our assumption that selection into patenting can be accurately captured by observable factors. In regressions available upon request, we examine the robustness of the results to trimming the sample corresponding to observations with especially high and low weights ($5^{th}/95^{th}$ percentile and $1^{st}/99^{th}$ percentile). The magnitude of the treatment effects are very similar to those reported in Table V. Below, we report the results of formal covariate balance tests. Recall that there were three time-varying covariates that gave IPTCW estimation its bite: the flow of publications in the year preceding the first patenting event; the inherent patentability of these publications (proxied by our RP score); and coauthorships with industry partners. Table AII shows that in the weighted sample, the one year lagged values of these variables are NOT associated with patenting in year *t*. In the same spirit, we have at our disposal two 'unused observables,' the employer's royalty rate and the average journal commercial score of the publications associated with each scientist. These variables are 'unused' because they are redundant with other variables (such as RP, number of collaborations, and the employer's patent stock) that are included in our model of the propensity to patent (this evidence is available from the authors upon request). However, if there are no major left-out determinants of patenting left-out of the model, these covariates should also be balanced in the weighted sample. This is indeed the case: the association between each lagged covariate with patenting in year *t* in the raw data disappears once we weight the sample appropriately, as can be observed in Table AII.

APPENDIX III: NAME MATCHING

We first generated a stratified random sample of scientists from relevant life sciences disciplines in the Proquest Dissertation Database. Fortunately, this database records full names of authors of dissertations. With this list of names, we queried for publications by these authors in the Web of Science Database. Because Web of Science indexes each publication with only the combination of authors' last names and first and middle initials, we painstakingly screened all downloaded papers to make sure that (i) an author started publishing no earlier than the year he started his Ph.D. program and no later than 35 years after the grant of his Ph.D. degree, and (ii) that the journals in which an author publishes papers match the Ph.D. subject field in which he wrote his dissertation. After we cleaned the publications data, we used the affiliation information recorded with each published paper to determine scientists' career histories. To collect the patent application information, we queried the NBER patent database with the names in our dataset (in the format of last name plus first name and middle initials). We checked for mismatches with the affiliation information we have for each scientist from their published papers and the address field (specifically, the state of patent application) in the patent database. We also used the scientist's degree year to specify a range of time that he can patent (i.e., no more than five years before and no later than 35 years after his degree grant year). Of course, as all researchers who have worked on cleaning patent data know, it is hazardous to rely solely on automated routines for name matching. We thus engaged in another round of manual screening of all the patents in our data to make sure that the class of a patent application matches a scientist's field of expertise as revealed by his dissertation and research publications.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Table S1. Sensitivity Analysis: Table
Figure S1. Sensitivity Analysis: Figure
Table AI. Sample Title Keywords in 1999
Table AII. Covariate Balance Tests

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

REFERENCES

Abadie, A., 2005, 'Semiparametric Difference-in-Differences Estimators,' *Review of Economic Studies*, 72, pp. 1–19.

Agrawal, A. K. and Henderson, R. M., 2002, 'Putting Patents in Context: Exploring Knowledge Transfer from MIT,' *Management Science*, 48, pp. 44–60.

Azoulay, P.; Ding, W. W. and Stuart, T. E., 2007, 'The Determinants of Faculty Patenting Behavior: Demographics or Opportunities?,' *Journal of Economic Behavior and Organizations*, 63, pp. 599–623.

Blumenthal, D.; Causino, N.; Campbell, E. G. and Louis, K. S., 1996, 'Relationships between Academic Institutions and Industry in the Life Sciences – An Industry Survey,' *New England Journal of Medicine*, 334, pp. 368–73.

Blumenthal, D.; Gluck, M.; Louis, K. S.; Stoto, M. A. and Wise, D., 1986, 'University-Industry Research Relationships in Biotechnology: Implications for the University,' *Science*, 232, pp. 1361–66.

Campbell, E. G.; Clarridge, B. R.; Gokhale, N. N.; Birenbaum, L.; Hilgartner, S.; Holtzman, N. A. and Blumenthal, D., 2002, 'Data Withholding in Academic Genetics – Evidence from a National Survey,' *Journal of American Medical Association*, 287, pp. 473–80.

Dehejia, R. H. and Wahba, S., 2002, 'Propensity Score-Matching Methods for Nonexperimental Causal Studies,' *Review of Economics and Statistics*, 84, pp. 151–61.

Di Gregorio, D. and Shane, S., 2003, 'Why Do Some Universities Generate More Start-Ups Than Others?,' *Research Policy*, 32, pp. 209–27.

Fabrizio, K. and DiMinin, A., 2008. 'Commercializing the Laboratory: The Relationship between Faculty Patenting and Publishing,' *Research Policy*, 37 (5), pp. 914–931.

Gouriéroux, C.; Montfort, A. and Trognon, A., 1984, 'Pseudo Maximum Likelihood Methods: Applications to Poisson Models,' *Econometrica*, 52, pp. 701–720.

Hall, B. H.; Jaffe, A. B. and Trajtenberg, M., 2001. 'The NBER Patent Citations Data File: Lessons, Insights and Methodological Tools,' NBER Working Paper #8498.

Hausman, J.; Hall, B. H. and Griliches, Z., 1984, 'Econometric Models for Count Data with an Application to the Patents-R&D Relationship,' *Econometrica*, 52, pp. 909–38.
Heckman, J. J.; Ichimura, H. and Todd, P. E., 1997, 'Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme,' *Review of Economic Studies*, 64, pp. 605–654.
Hellman, T., 2007, 'The Role of Patents for Bridging the Science to Market Gap,' *Journal of Economic Behavior and Organizations*, 63, pp. 624–647.
Henderson, R. M.; Jaffe, A. B. and Trajtenberg, M., 1998, 'Universities as a Source of Commercial Technology: A Detailed Analysis of University Patenting, 1965–1988,' *The Review of Economics and Statistics*, 80, pp. 119–127.
Hernán, M. A.; Brumback, B. and Robins, J. M., 2000, 'Marginal Structural Models to Estimate the Causal Effect of Zidovudine on the Survival of HIV-Positive Men,' *Epidemiology*, 11, pp. 561–570.
Hernán, M. A.; Brumback, B. and Robins, J. M., 2001, 'Marginal Structural Models to Estimate the Joint Causal Effect of Nonrandomized Treatments,' *Journal of the American Statistical Association*, 96, pp. 440–48.
Jaffe, A. B. and Lerner, J., 2001, 'Reinventing Public R&D: Patent Policy and the Commercialization of National Laboratory Technologies,' *The RAND Journal of Economics*, 32, pp. 167–98.
Krimsky, S., 2003, *Science in the Private Interest: Has the Lure of Profits Corrupted Biomedical Research?* (Rowman & Littlefield, Lanham, Maryland, U.S.A).
Lach, S. and Schankerman, M., 2004, 'Royalty Sharing and Technology Licensing in Universities,' *Journal of The European Economic Association*, 22–3, pp. 252–264.
Levin, S. G. and Stephan, P. E., 1991, 'Research Productivity over the Life-Cycle – Evidence for Academic Scientists,' *American Economic Review*, 81, pp. 114–32.
Lim, K., 2004, 'The Relationship between Research and Innovation in the Semiconductor and Pharmaceutical Industries (1981–1997),' *Research Policy*, 33, pp. 287–321.
Merges, R., 1997, *Patent Law and Policy: Cases and Materials* (Michie Co, Charlottesville, Virginia, U.S.A).
Merton, R. K., 1973 [1942], 'The Normative Structure of Science,' in *The Sociology of Science: Theoretical and Empirical Investigations*, Merton, Robert K. (ed.) (The University of Chicago Press, Chicago, Illinois, U.S.A).
Mowery, D. C.; Nelson, R. R.; Sampat, B. N. and Ziedonis, A. A., 2001, 'The Growth of Patenting and Licensing by U.S. Universities: An Assessment of the Effects of the Bayh-Dole Act of 1980,' *Research Policy*, 301, pp. 99–119.
Murray, F., 2002, 'Innovation as Co-evolution of Scientific and Technological Networks: Exploring Tissue Engineering,' *Research Policy*, 31, pp. 1389–1403.
Murray, F. and Stern, S., 2007, 'Do Formal Intellectual Property Rights Hinder the Free Flow of Scientific Knowledge? An Empirical Test of the Anti-Commons Hypothesis,' *Journal of Economic Behavior and Organizations*, 63, pp. 648–687.
Owen-Smith, J. and Powell, W. W., 2001, 'To Patent or Not: Faculty Decisions and Institutional Success at Technology Transfer,' *Journal of Technology Transfer*, 26, pp. 99–114.
Papke, L. E. and Wooldridge, J. M., 1996, 'Econometric Methods for Fractional Responses with an Application to 401(k) Plan Participation Rates,' *Journal of Applied Econometrics*, 11, pp. 619–632.
Robins, J. M., 1999, '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*, Halloran, M. E. and Berry, D. (eds.) (Springer-Verlag, New York, U.S.A).
Robins, J. M.; Hernán, M. A. and Brumback, B., 2000, 'Marginal Structural Models and Causal Inference in Epidemiology,' *Epidemiology*, 115, pp. 550–60.

Rosenbaum, P. R. and Rubin, D. B., 1983, 'The Central Role of the Propensity Score in Observational Studies for Causal Effects,' *Biometrika*, 701, pp. 41–55.

Rosenberg, N., 1998, 'Chemical Engineering as a General Purpose Technology,' pp. 167–92 in *General Purpose Technologies and Economic Growth*, Helpman, E. (ed.) (MIT Press, Cambridge, Massachusetts, U.S.A).

Silva, J. M. C. Santos and Tenreyro, S., 2007, 'The Log of Gravity,' *The Review of Economics and Statistics*, 884, pp. 641–658.

Smith, J. A. and Todd, P. E., 2005, 'Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators?,' *Journal of Econometrics*, 125, pp. 305–353.

Stephan, P. E.; Gurmu, S.; Sumell, A. J. and Black, G., 2007, 'Who's Patenting in the University? Evidence from the Survey of Doctorate Recipients,' *Economics of Innovation and New Technology*, 16, pp. 71–99.

Stevens, A. J. and Toneguzzo, F. (eds.), 2003, *AUTM Licensing Survey, 2003* (The Association of University Technology Managers, Northbrook, Illinois, U.S.A.).

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

Stuart, T. E. and Ding, W. W., 2006, 'When Do Scientists Become Entrepreneurs? The Social Structural Antecedents of Commercial Activity in the Academic Life Sciences,' *American Journal of Sociology*, 92, pp. 92–144.

Thursby, J. G. and Thursby, M. C., 2002, 'Who Is Selling the Ivory Tower? Sources of Growth in University Licensing,' *Management Science*, 48, pp. 90–104.

Thursby, M. C.; Thursby, J. G. and Gupta-Mukherjee, S., 2007, 'Are There Real Effects of Licensing on Academic Research? A Life Cycle View,' *Journal of Economic Behavior and Organizations*, 63, pp. 577–598.

Thursby, J.; Fuller, A. and Thursby, M., 2009. 'U.S. Faculty Patenting: Inside and Outside the University,' *Research Policy*, 38, pp. 14–25.

Wooldridge, J. M., 2002, 'Inverse Probability Weighted M-Estimators for Sample Selection, Attrition and Stratification,' *Portuguese Economic Journal*, 12, pp. 117–139.