International Journal of Industrial Organization 36 (2014) 83–94

Contents lists available at ScienceDirect

International Journal of Industrial Organization

journal homepage: www.elsevier.com/locate/ijio

Intellectual property rights and the evolution of scientific journals as knowledge platforms

Daniel C. Fehder [a], Fiona Murray [a], Scott Stern [a,b,*]

[a] MIT Sloan School of Management, Massachusetts Institute of Technology, 100 Main Street e62-370, Cambridge, MA 02142, United States
[b] National Bureau of Economic Research, Massachusetts Institute of Technology, 100 Main Street e62-470, Cambridge, MA 02142, United States

ARTICLE INFO

Available online 21 August 2014

JEL classification:
D02
I23
O30
O34
L33
D83

Keywords:
Scientific journals
Patents
Intellectual property
Communities
Platforms
Innovation

ABSTRACT

The objective of this paper is to evaluate the role that formal intellectual property rights (IPR) play in shaping the downstream demand for knowledge that is initially disclosed through scientific publication in fields where research is generated and utilized across different institutional settings (i.e., academia versus industry). For scientific discoveries with potential commercial applicability, researchers (or their funders) may also seek to establish formal intellectual property protection (e.g., patents); choosing to establish a "patent–paper pair" allows researchers to influence follow-on access to knowledge disclosed in a given scientific journal. This paper evaluates the relationship between scientific journal publication and patenting in research communities with significant public and private authorship by examining the incidence and impact of patent–paper pairs in two journals founded in the late 1990s/early 2000s, Nature Biotechnology and Nature Materials. Using a differences-in-differences framework that exploits the delay between publication and patent grant, we document a range of findings about the impact of patent grant across time and across research populations. First, we find that the negative impact of patent grant is concentrated in the first few years after a journal's founding and eventually becomes positive. Second, patent grant positively impacts follow-on citations from private authors more than from public authors. Finally, we observe an assortative matching pattern where intellectual property grant increases forward citations from authors sharing the same institutional affiliation (e.g., public authors citing public papers) more than research across institutional lines (e.g., public authors citing private papers).

1. Introduction

In 2004, researchers from the University of Manchester and the Institute for Microelectronics in Chernogolovka (Russia) developed the first feasible approach for fabricating and characterizing graphene—a complex single-layered carbon crystal structure. This fundamental scientific discovery, confirming a central theoretical prediction at the intersection of physics and materials science, was published promptly in *Science* (Novoselov et al., 2004) and almost immediately spawned extensive follow-on research. These follow-on studies are wide-ranging: from explorations of the role of graphene in addressing fundamental questions of physics, to considerations of how to exploit these structures across a range of commercial applications such as electronics, chemical engineering and even aircraft design. The simultaneous impact of graphene research on both basic scientific research and commercial applications typifies the dual purpose of research which Stokes and subsequent scholars have classified as residing in "Pasteur's Quadrant"[1] (Murray, 2002; Stokes, 1997). Indeed, the fact that the fabrication and characterization of graphene resolves a basic research question while immediately being applicable to important practical problems was at the heart of the decision to award these researchers the Physics Nobel Prize a mere six years after publication (Swedish Royal Academy, 2010). Importantly, while a small number of the most important contributions to graphene research are published in the two leading general-interest journals—*Science* and *Nature*—the single most important outlet for graphene research publication is the relatively young *Nature* satellite journal, *Nature Materials* (see for example Geim and Novoselov, 2007).

Given the "dual" nature of this research, it is perhaps not surprising that many graphene researchers chose not only to disclose their

* Corresponding author at: National Bureau of Economic Research, Massachusetts Institute of Technology, 100 Main Street e62-470, Cambridge, MA 02142, United States.
E-mail addresses: dfehder@mit.edu (D.C. Fehder), fmurray@mit.edu (F. Murray), sstern@mit.edu (S. Stern).

[1] Pasteur's Quadrant refers to the quadrant in Stoke's taxonomy of science where a piece of research has both basic theoretical impact and direct and immediate practical application. It takes Pasteur's name because the germ theory of disease is a seminal example of this type of science.

http://dx.doi.org/10.1016/j.ijindorg.2014.08.002
0167-7187/

findings through publication in elite scientific journals but also to seek formal intellectual property rights (IPR) (see Murray, 2002; Stokes, 1997).[2] Indeed, IPR has become so prevalent a form of disclosure in graphene research that recent editorials in *Nature Materials* considered whether the complicated web of patents from a broad range of both public and private research labs might eventually become a "patent thicket" (Shapiro, 2001; Tannock, 2011). This discussion is reminiscent of policy discussions in the 1990s about the proliferation of IP in biotechnology and its potential anti-commons effect (Cohen et al., 2000; Heller and Eisenberg, 1998). Like contributors to any institution designed for knowledge sharing and exchange, scientists and the editorial staff of the newly formed journal *Nature Materials* were concerned that IP might restrict access and undermine the nature and functioning of their scientific journal. This worry was particularly salient for *Nature Materials* since its mission was to create "an invaluable resource for scientists, in both academia and industry, who are active in the process of discovering and developing materials."[3]

The objective of this paper is to evaluate the role that formal intellectual property rights (IPR) play in shaping the downstream demand for dual-use knowledge that is initially disclosed through scientific publication in fields where research is generated in different institutional settings (i.e., academia versus industry). Research on standard setting organizations (SSOs) has shown that the interactions of actors with differing incentives and preferences shape an institution's success in fostering knowledge disclosure and cumulative innovation (Simcoe et al., 2007). Analogously, we examine how the scientific journal acts as a key point of interaction binding potentially diverse constituencies into a research field. Specifically, we examine how a researcher's decision to pursue intellectual property rights (IPR) impacts the demand for the article from follow-on researchers and how this impact varies across time and between public and private researchers.

Our empirical analysis uses the publications (and associated patents) linked to two new journals, *Nature Biotechnology* and *Nature Materials*, which share common "platform rules" including editorial policies, access policies, etc. For *Nature Biotechnology*, we observe publications from its founding in 1997 through 2005. For *Nature Materials*, we observe publications from its founding in 2002 through 2006. For each of these publications, we establish whether that publication is associated with a US patent (i.e., form a patent–paper pair). Finally, we observe detailed bibliometric data about each publication (e.g., author affiliations, citations, etc.), as well as detailed bibliometric data about each follow-on paper that cites one of our focal articles (through the end of 2010).

Building on the empirical approach of past research, we first document a range of findings that highlight the interaction between a journal article's follow-on citations and formal IP rights. In the cross section, patented articles are associated with higher forward citations. Building on the differences-in-differences framework in Murray and Stern (2007), we find that the average impact of IPR grant on follow-on citations was positive. Using non-parametric patent grant year cohorts to explore the time dynamics, we find that the impact of patent grant is negative in a journal's early years then subsequently becomes positive. Thus, formal IPR seems to facilitate a "market for ideas" once a journal's reputation is established.

We then turn to the heart of our analysis and focus on whether the public versus private institutional affiliations of both the cited and the citing researcher drive the magnitude of IPR grant's impact on forward citations. First, we find that IPR grant increased forward citations from both populations but has a larger positive impact on forward citations from private authors than from public authors. Next, when we account for the institutional affiliation of both the cited and citing author, we observe an assortative matching pattern where intellectual property grant increases forward citations from authors sharing the same institutional affiliation as the cited author (e.g., public authors citing public papers) more than research across institutional lines (e.g., public authors citing private papers). For example, IPR grant increases annual private citations to private articles by 50% while it decreases annual public citations to private articles by nearly 10%. In addition, when we account for citations coming from *Nature* publications versus other journals, we find suggestive evidence that the identity of the journal publishing the forward citations mediated this institution-based assortative matching process. Thus, IPR grant seems to increase the potential for incremental research projects where the cited and citing authors have matching institutional affiliations.

If IPR not only changes demand for a paper but also induces an assortative matching process shifting the institutional locus of subsequent research, IPR grant impacts not only the amount of downstream cumulative research but also the organization of downstream projects through their institutional location (public versus private). As a bundle of control rights and disclosed knowledge, a patent grant can change both a follow-on researcher's access to the prior work and their beliefs about the value of building upon it, but these two factors can have very different implications for welfare. While disclosure might change the ex ante tradeoffs between different potential projects, contracting frictions between researchers across an institutional divide might inhibit projects with high potential value from going forward. If, for instance, public researchers face differentially easier access to patented research from other public researchers (perhaps because of institutional norms promoting open science) than from private researchers (because the technology licensing offices of universities license IP only in one direction), then patented scientific discoveries made in private research labs may not receive the optimal mix of "exploratory" versus "exploitative" follow-on projects (Aghion et al., 2008; Bikard, 2012).

The paper is organized as follows: Section 2 motivates the analysis by a review of the relevant literature. Section 3 describes the empirical context and approach we use to explore both the supply and demand side characteristics of journal platforms. Section 4 outlines our identification strategy, the data and summary statistics. Section 5 presents the empirical results, and Section 6 concludes.

2. The role of IP on the diffusion of knowledge in scientific communities with heterogeneous actors: theory and evidence

Central to knowledge exchange and accumulation is the ability of individuals to learn from others who disclose ideas, share knowledge and provide sufficient access to enable replication, validation and follow-on innovation (Collins, 1974; Mokyr, 2004). Scientific journals play a foundational role in such cumulative exchange and learning. Indeed, scholarly journals are arguably the single most important mechanism facilitating the process of knowledge transfer across researchers and different research domains, especially in fields where researchers are drawn from significantly different institutional settings.

When scientific knowledge is useful, it is possible for the authors to disclose their findings in the form of a patent as well as via publication through a journal. With the grant of private property rights, knowledge that was otherwise exchanged only on the basis of the norms of the publication platform is now subject to the control rights associated with the IPR. The impact of institutions that offer authors an additional set of control rights beyond those specified by the infrastructure of the journal and its editorial board is a contentious area of debate among legal scholars, economists, policymakers and the scientific community.

One set of scholars considers how the increased use of intellectual property rights might undermine the established institutions of academic science. In particular, scholars are concerned with how the expansion of patents on scientific knowledge and the growing

[2] Interestingly, Novoselov and Geim (and the University of Manchester) chose to simply publish their 2004 *Science* findings and did not additionally seek formal IP over their breakthrough. And in an interview, Geim noted that he had refrained from filing patents in the graphene area because of concerns over potential lawsuits from "a major electronics company" and over a lack of specific industrial applications and industrial partners for his developments (Brunfield, 2010).

[3] http://www.nature.com/nmat/authors/index.html.

enforcement of these patents influence markets for knowledge traditionally shaped solely by the scientific community (Heller and Eisenberg, 1998). More specifically, it is argued that formal IP rights and the concomitant threats of legal enforcement throughout the community challenge established norms (David, 2004; Lessig, 2004). Some strong qualitative evidence suggests that patenting does rapidly shape demand-side practices as IP owners send “cease and desist” letters to potentially infringing individuals and their organizations; unwilling or unable to respond to these requirements, knowledge workers rapidly reduce their participation in the community (Murray, 2010).

Other work has grounded the impact of intellectual property in its specific functioning in a particular institutional context. Overall, it is important to consider that IP rights are facilitative, providing their owners with a variety of legal rights that they may or may not use (Edelman and Suchman, 1997). Thus, the effects of IP is often linked to specific legal events which clarify the particular rights and agreements for a specific community: examples include incorporating patents into technical standards (Rysman and Simcoe, 2008), the signature of an access agreement over research mice as occurred between DuPont and the National Institutes of Health (Murray et al., 2009), or the formulation of national and institution-specific standards for material transfer agreements (Mowery and Ziedonis, 2007; Walsh et al., 2007). Taken together, these studies highlight that the impact of IPR on knowledge exchange can vary significantly across different institutional locations.

Another set of research more formally conceptualizes the impact of patenting on the potential negotiation between upstream and downstream researchers in specific settings. Williams (2010) interprets the lower downstream use of Celera genes as a result of increased transaction costs in a setting with private information (Williams, 2010). On the other hand, researchers have stressed that it is also possible that formal IP actually facilitates technology transfer across research generations by enabling the market for ideas (Arora et al., 2004; Gans and Stern, 2000; Hellmann, 2007; Kitch, 1977). Theoretical work has explored the complementarities between publication and patents in the context of contracting between scientists and commercial firms (Gans et al., 2008). Taken together, these empirical and theoretical elements suggest that IP may have a limited influence on knowledge shared through journals or even allow for increased participation in knowledge exchange (Kieff, 2005).

Recent work has linked these debates to the design of institutions promoting cumulative innovation from a two-sided platform perspective. In the context of SSOs, researchers have found that the use of formal intellectual property varies significantly across different types of contributors, and across different platforms (Rysman and Simcoe, 2008; Simcoe et al., 2007). Analogously, a number of papers have suggested that scientific journals can usefully be placed into the theoretical framework of two-sided platforms and have explored how editorial policies and selection criteria influence the characteristics of researchers both supplying and utilizing research on the platform (Lerner and Tirole, 2006; McCabe and Snyder, 2005, 2007). Unlike the platform-informed approach to SSOs, empirical researchers have yet to examine journals as a critical site of exchange between heterogeneous types of researchers (e.g., public versus private) and the consequences of these exchanges for the functioning of academic journals. By focusing on the differential impact of IPR on the downstream use of scientific knowledge, our analysis follows an emerging body of research that clarifies the causal channels through which intellectual property impacts the disclosure and cumulative use of technical information through specific institutions (Furman and Stern, 2011; Murray and Stern, 2007; Williams, 2010).

3. Empirical framework

We now focus on the role of scientific journals in the market for knowledge and their influence over the policies and practices of scientists (arising both as a result of the professional editorial staff and the editorial practices of prominent scientists). This section overviews our empirical context and analytical strategy which allow us to explore how IPR impacts the cumulative use of scientific knowledge in scientific journals and how this varies by the institutional affiliation of the cited and citing researchers.

3.1. Empirical approach

Our research design focuses on the demand-side publishing activities for two high quality research journals which share significant scientific and institutional similarities, *Nature Biotechnology* and *Nature Materials*. Both of these journals were established by the same high-quality publishing house—Nature Publishing Group (NPG). While the NPG also established a variety of other journals in this period, our specific choice of *Nature Biotechnology* and *Nature Materials* is grounded in their similar intellectual origins and similar mission: research exchanged in both journals is grounded in chemistry—a discipline that has served as an important foundation for spawning powerful new research communities dedicated to particular application arenas. In the life sciences, especially those aspects covered by *Nature Biotechnology*, much of the work on new tools to interrogate (and reshape) the chemical machinery of the human body is grounded in chemistry and bio-chemistry in particular (see McCormick, 1990). Likewise, in *Nature Materials*, research into new tools and methods to transform the chemical machinery of the material world are grounded in chemistry and in particular physical chemistry.

In both instances, the journals were initiated in response to a growing awareness of the need for high-quality research platforms for the exchange of knowledge in Pasteur's Quadrant in both of these arenas. The editorial mission of Nature Biotechnology emphasizes this need: “*to publish high-quality original research that describes the development and application of new technologies in the biological, pharmaceutical, biomedical, agricultural and environmental sciences, and which promise to find real-world applications in academia or industry*”. Likewise, the editorial policy of Nature Materials focuses (in part) on issues in Pasteur's Quadrant, stating that “*Nature Materials covers all applied and fundamental aspects of the synthesis/processing, structure/composition, properties and performance of materials, where “materials” are identified as substances in the condensed states (liquid, solid, colloidal) designed or manipulated for technological ends*”.

What is particularly striking about both of these journals is the rapidity with which both developed a highly-active market for knowledge with high average annual citation counts, as captured in the Journal Impact Factor (JIF). In 1998, *Nature Biotechnology* had an impact factor of 8.09 which rose to 31.04 in 2010. Similarly, *Nature Materials* reached 29.92 in 2010 (up from only 10.78 in 2003) rapidly placing it in first place among materials science journals and across all primary research journals in physics and chemistry (see Editorial 2003).[4] We use publication citations in our study to trace the flow of ideas over time and across institutional settings (de Solla Price, 1965; Garfield and Merton, 1979; Hall et al., 2001; Jaffe et al., 1993).[5]

In addition to their success as impactful journals serving research communities at the intersection of academic and commercial research, both *Nature Biotechnology* and *Nature Materials*, as NPG journals, also use professional editorial boards rather than academic editors to administer the journal and select articles for publication. Professional editorial boards remove the risk that “political gatekeepers” govern the selection of published articles and ensure consistent application of editorial policies. They also allow NPG to promulgate uniform policies

[4] The impact factor measures the average number of citations per paper in a year based on articles published in the previous two years.

[5] More broadly, the sociological literature has articulated the importance of publication citations as a form of recognition for knowledge exchange in the scientific community (Hagstrom, 1975; Schubert and Braun, 1993).

across their holdings with regards to editorial policies, screening, and review processes. As described above, these editorial design choices influence how a journal functions in its research community.[6]

Our empirical approach exploits the existence of patent–paper pairs to explore the impact of patenting as an alternative exchange mechanism that can take place alongside the publication-mediated exchange (Ducor, 2000; Gans et al., 2008; Murray, 2002).[7] A patent–paper pair embodies a decision to additionally embed a piece of knowledge within the patent system which makes it available for exchange through an entirely different set of property-based rules and norms. In addition, the substantial gap between the date of scientific publication and the date of patent grant provides a natural experiment to measure the impact of IPR grant (see Murray and Stern, 2007). While papers in the Nature Biotechnology and Nature Materials are typically published rapidly from the time of submission to the journal (within 3–6 months), grant of the paired patent takes approximately four years. It is important to emphasize that patent grant delay is more than simply a pro forma administrative glitch. The formal process of patent application produces successive reductions in uncertainty about the extent of an application's formal IP rights. Each of these reductions shifts the uses of a potential patent in knowledge exchange, culminating with the threat of infringement lawsuits after a patent grant.

3.2. Analytical approach

By following the supply and demand-side activities of the population of *Nature Biotech* and *Nature Materials* research articles—a fraction of which have paired patents (covering the same knowledge)—we are able to evaluate two main questions regarding how IPR impacts the downstream use of scientific knowledge disclosed in academic journals. First, we can build upon prior research by assessing how patent grants change the average demand for papers (as captured in annual citations) in our sample by determining how the rate of forward citations changes after patent grant. Second, we can see how the impact of IPR grant on forward citations varies not only by the institutional affiliation of the citing author but also the match between citing and cited author (e.g. public authors citing public papers). Additionally, we can examine whether this institution-based matching process varies for downstream citations within the *Nature* family versus all other publications.

Our empirical specification examines the annual count of forward citations in publications. This dependent variable took the form of count data skewed to the right. We used a negative binomial model of the annual citations for each of the publications in our dataset. For our first set of regressions, we use random effects models to establish the cross-sectional behavior of both article suppliers and articles users for each journal and how these behaviors relate to patenting. In our second set of models, we estimate the causal impact of a patent grant on annual forward citations using a difference-in-difference framework where we control for the variation in article quality and impact of individual papers using article fixed effects. In all models, we use additional controls. To account for potential correlations between annual forward citations and the effects of particular calendar years, we include journal-specific calendar year fixed effects in all models in this paper (called Journal-Citation Year fixed effects in our empirical tables). To account for the age-related patterns in the citation year data, we include journal-specific article age fixed effects in all models discussed in this paper (called Journal-Age Fixed Effects in our empirical tables). When we observe the number of citations a paper receives before and after the grant of a patent, we are able to identify how the average pattern of citations to a paper changes after the introduction of a patent.

By identifying whether each follow-on citation has either entirely public authors or at least one private author, we can examine how the impact of IPR varies by the institutional affiliation of the follow-on researchers. Our key regression results used a difference-in-difference framework to estimate the causal impact of patent grant on annual forward citations. The specific form of each regression presented in our results is discussed in detail in Appendix B, but our key regression results build upon the results from Murray and Stern (2007) by exploring the different marginal responses to patent grant in two critical populations in Pasteur's Quadrant: private and public authors. We simultaneously estimate the impact of a patent grant on follow-on citations by public and private authors. Specifically, we jointly estimated:

$$CITES_{i,pubyear(j),t,public} = f\left(\varepsilon_{i,j,t,p};\ \gamma_i + \beta_{tp} + \delta_{t-pubyear,p} + \psi_{public}\ \ POST{-}GRANT_{i,t}\right)$$

$$CITES_{i,pubyear(j),t,private} = f\left(\varepsilon_{i,j,t,p};\ \gamma_i + \beta_{tp} + \delta_{t-pubyear,p} + \psi_{private}\ \ POST{-}GRANT_{i,t}\right)$$

β_{tp} is a separate year effect for public and private citations, $\delta_{t-pubyear,p}$ captures the separate effect for the age of the article by citation type, and POST-GRANT is a dummy variable equal to one only for articles which have received a patent.[8] Conditional on the vector of fixed effects and the treatment variable, the number of forward citations for each article's citation year is modeled as being drawn independently from a common negative binomial distribution (i.e. $\varepsilon_{i,j,t,p}$ is assumed to be i.i.d.). In this specification, ψ_{public} measures the impact of patent grant on the downstream demand behavior of researchers with public affiliations, while $\psi_{private}$ measures the impact of a patent grant on the downstream behavior of researchers with private affiliations. Thus, ψ measures the impact of patent grant accounting for fixed differences in the citation rate across articles and relative to the non-parametric trend in citation rates for articles with similar characteristics. The difference-in-difference framework implicitly compares treated and untreated articles from the same journal publication year across the same citation years ensuring that our causal estimates are not affected by potential truncation problems. Overall, we are interested in assessing whether $\psi_{public} = \psi_{private}$ or whether there are heterogeneous treatment effects across the two subpopulations. Also, we want to examine whether ψ_{public} and $\psi_{private}$ each varied for citations to public versus private articles.

In our results, we extend our analysis of the two institutional populations in a number of additional ways. We provide a brief summary here, but provide a detailed account of our estimation techniques in Appendix B. First, we examine whether the marginal impact of patenting for each population varies by journal. Next, we introduce a set of non-parametric cohort effects into the above regression to identify the time varying marginal impact of patenting by journal across the different citing populations. We used patent grant year cohorts because these allowed us to identify how the causal impact of patent grants varied over time. Lastly, we examined how the impact of IPR grant varied for authors in Nature publications across the two different institutional populations.

4. The data

4.1. Sample definition

Our sample was composed of 1450 published scientific research papers disclosing potentially patentable knowledge drawn from two

[6] Explicit common policies shared between all NPG journals (including *Nature* itself) that are relevant to a journal's ability to shape knowledge exchange include statements on author responsibilities, copyright, embargo policies, availability of materials and data, refutations, complaints and corrects, plagiarism and fabrication. See http://www.nature.com/nmat/authors/index.html for a detailed and also for links to the policy specifics in each of these areas.

[7] Scientists produce "pairs" when they choose to disclose the same novel research results in both scientific publications and patent applications.

[8] Similar to Murray and Stern (2007), this baseline analysis assumes that age fixed effects associated with citations do not depend on whether a paper receives a patent. In particular, a key assumption of our base model is that patented articles are not simply "shooting stars"—articles that, for exogenous reasons, experience a high rate of early citation followed by a rapid decline.

related top-tier journals, *Nature Biotechnology* and *Nature Materials*. Our data set begins with the founding of each journal (1997 for *Nature Biotechnology* and 2002 for *Nature Materials*). Our choice of time period and journals allows us to examine the impact of IPR in two different research communities. Prior research has explored the differing motivations for intellectual credit and financial gain and how they influence a scientist's decision making (Dasgupta and David, 1994). By choosing two journals which quickly became focal institutions in research communities squarely in Pasteur's Quadrant, we expect to see a rich set of interactions between researchers with different institutional affiliations.

For each of the 1450 articles in our sample, we established whether or not an associated patent had been granted by the USPTO (thus generating a pair). We conducted searches on the USPTO database using a series of decreasingly restrictive combinations of author names and geographical location. We then hand-coded all patents returned from these searches to establish whether or not they represented a part of a pair by comparing abstracts and other patent content. Using this method, we identified 525 patents (36%) that were associated with a paper to form a patent–paper pair.

For each of these 1450 articles and 525 patents, we then collected a range of variables describing the observable characteristics of the papers and patents (see Table 1). For each publication, we then created a variable PATENTED equal to one if the article was part of a pair and set to 0 otherwise. Additionally, we coded the date in which the patent application was filed (APPLICATION YEAR) and the year in which the patent was granted (GRANT YEAR). We then generated a variable from the difference of GRANT YEAR and APPLICATION YEAR (PATENT LAG) which represented the random latency between filing and USPTO action.

We then collected data on the forward citations to the 1450 articles from Thomson ISI Web of Science for the years 1997 to 2010 (188,126 articles in total). Using the ISI data, we coded a range of variables characterizing the citing team's institutional affiliations and geographical location. In particular, we coded a set of dummy variables to describe whether all of the citing institutions were public entities (CITE PUBLIC) or if at least one of the citing institutions was a for-profit company (CITE PRIVATE). These dummy variables were then used to generate marginal citation year counts for each publication that we will use to explore the dynamic interactions effects of intellectual property grants on different sub-populations of our sample.

From the citation-year level data, we constructed one final set of variables. Our main dependent variable was the total number of citations an article received in a calendar year (YEAR) which we defined as ANNUAL FORWARD CITATIONS. For each citation-year observation, we defined PUBLICATION AGE as YEAR–PUBLICATION YEAR. We then defined an indicator variable, PATENT POST GRANT, set to one for each paper's citation-years in which the patent has already been granted (i.e., when YEAR–GRANT YEAR > 0). Using the indicator variables described in the paragraph above (CITE PUBLIC and CITE PRIVATE), we were also able to construct a set of counts of annual citations by different subgroups. ANNUAL FORWARD CITATIONS PUBLIC captures all those forward citations in a given year with all sector authors. Similarly, ANNUAL FORWARD CITATIONS PRIVATE measures the number of citations in a year that had at least one private sector author (i.e. working for a for-profit firm). Taken together, our variables allow us to characterize both the supply and demand side behavior on similar two-sided platforms. In addition, the structure of our data allows us to make comparisons across subpopulations, across journals and across time.

Table 1
Variable definitions.

Variable	Definition	Source
Publication characteristics		
PUBLICATION YEARj	Year in which article is published	NB, NM
# AUTHORSj	Count of the number of authors of Article *j*	NB, NM
PATENTEDj	Dummy variable = 1 if Article is associated with a patent issued by the USPTO prior to October, 2003	USPTO
TOTAL CITATIONSj	# of FORWARD CITATIONS from publication date to 12—2005	SCI
Patent characteristics		
APPLICATION YEARj	YEAR in which PATENT was applied for	USPTO
GRANT YEARj	YEAR in which PATENT has been granted	USPTO
# INVENTORSj	Count of the number of inventors listed in the granted patent associated with Article *j*	USPTO
Citation year characteristics		
ANNUAL FORWARD CITATIONSjt	# of Forward Citations to Article *j* in Year *t*	SCI
ANNUAL FORWARD CITATIONS PUBLICjt	# of Public Forward Citations to Article *j* in Year *t*	SCI
ANNUAL FORWARD CITATIONS PRIVATEjt	# of Private Forward Citations to Article *j* in Year *t*	SCI
PATENT POST-GRANTjt	Dummy variable = 1 if PATENTED = 1 & CITATION YEAR > GRANT YEAR	USPTO
YEARjt	Year in which FORWARD CITATIONS are received	SCI

USPTO—United States Patent Office; NB—Nature Biotechnology; NM—Nature Materials; SCI—Science Citation Index.

4.2. Summary statistics

Both *Nature Biotechnology* and *Nature Materials* are highly regarded in their fields with higher than average journal impact factors (31.085 and 29.897, respectively in 2010) making them both the highest-ranked journals in their category by this metric. Given their prestige and broad audience, it is unsurprising that they each show high ANNUAL FORWARD CITATIONS: 14.11 (std 20.97) and 21.25 (std 27.33) for *Nature Biotechnology* and *Nature Materials* respectively (see Tables 1 and 2).

Overall, these articles also demonstrated a high degree of collaboration. In *Nature Biotechnology*, we observed an average of 6.86 (std 4.45) authors per paper and 5.77 (std 2.82) authors per paper for *Nature Materials*. While the rates of authorship were high for published papers in both journals, we observed similarly reduced numbers of inventors for patents in our pairs. The average #INVENTORS was 3.3 (std 1.76) for *Nature Biotechnology* and 3.9 (std 1.9) for *Nature Materials*.

We also observed some important differences in article characteristics between the two journals, most notably the rate of patenting between articles in these two journals. Overall, the rate of patenting in *Nature Biotechnology* was 47% in our sample while the rate was 18% in *Nature Materials*. The rates of patented articles by publication year for

Table 2
Summary statistics.

	Mean	Std. Dev	Min	Max
Panel A: publication characteristics				
PUBLICATION YEAR	2002.06	2.79	1997	2006
# AUTHORS	6.46	4.18	1	43
PRIVATE AUTHOR	0.24	0.42	0	1
PATENTED	0.36	0.48	0	1
Observations	1,450			
Panel B: patent characteristics				
GRANT YEAR	2004.91	3.96	1996	2011
PATENT LAG	1468.80	642.97	238	3714
# INVENTORS	3.42	1.86	1	15
Observations	525			
Panel C: citation-year characteristics				
ANNUAL FORWARD CITATIONS	15.96	23.95	0	453
CITATION YEAR	2006.04	3.02	1998	2010
ANNUAL FORWARD CITATIONS PUBLIC	14.16	21.60	0	425
ANNUAL FORWARD CITATIONS PRIVATE	1.79	3.57	0	63
ANNUAL FORWARD CITATIONS US	6.55	10.70	0	166
PATENT POST GRANT	0.23	0.42	0	1
Observations	11,507			

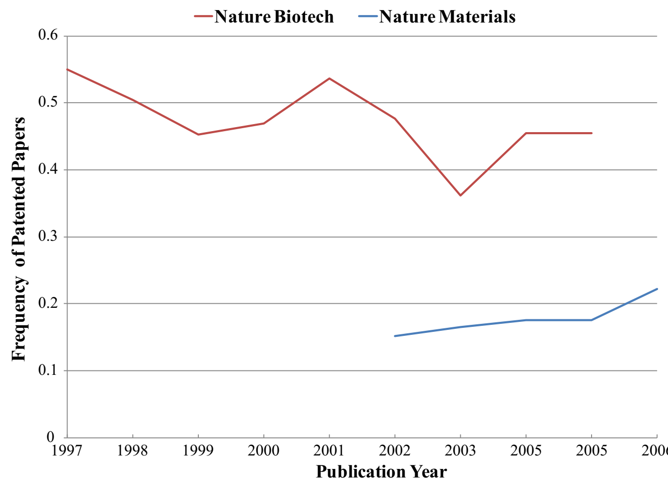

Fig. 1. Frequency of patented papers by publication year and journal.

each journal are displayed in Fig. 1. While patent–paper pairing fluctuates across years in *Nature Biotechnology*, it remains consistently higher than the rate in *Nature Materials*. There are also significant differences between the characteristics of patented and unpatented articles and these differences are heterogeneous across journal (see Table 3).

The compositions of authorship teams varied for patented and unpatented articles (in addition to varying across journal). For *Nature Biotechnology*, 35% of patented articles had at least one private author compared to 26% for unpatented articles. Fig. 2 shows how the rate of private authorship has varied across publication years for both *Nature Biotechnology* and *Nature Materials*. While there is variation across time for both journals, neither shows a clear trend. Overall, there are noticeable differences across significant margins for articles supplied to the two journals.

Above, we have focused on the characteristics of the articles published in the two journals (the supply side of the platform), but there are also some important differences in the characteristics of the citing articles across journals (the demand side). A similar trend to the differences in patenting explored above is observed when we aggregated yearly citation counts for each journal by the different institutional affiliations of the citing articles (Public versus Private). The mean ANNUAL FORWARD CITATION PUBLIC was 12.21 (std 19.18) for *Nature Biotechnology* and the mean ANNUAL FORWARD CITATION PRIVATE was 1.9 (std 3.67). For *Nature Materials*, the mean ANNUAL FORWARD CITATION PUBLIC was 19.76 (std 26.59) while ANNUAL FORWARD CITATION PRIVATE was 1.48 (std 3.24). When we analyzed the frequency of private forward citations by citation year, we also observed notable time dynamics (see Fig. 3).

The frequency of private forward citations has been decreasing sharply for *Nature Biotechnology* throughout our period of observation, moving from a high of 0.195 in 1998 to a low of 0.0975 in 2010.

Table 3 explores how annual forward citations vary by patent status for both journals. For both journals, patented articles receive higher mean annual forward citations as well as higher mean public and private annual forward citations than non-patented articles. Table 4 shows the impact of patenting on annual forward citations broken down by the institutional affiliation of the cited article's authors. Interestingly, ANNUAL FORWARD CITATION PRIVATE is highest when citing patented private authored papers and higher for articles citing unpatented private papers than patented public papers for both *Nature Biotech* and *Nature Materials*. Similarly, the highest level of ANNUAL FORWARD CITATION PUBLIC is observed when citing public patented articles. Thus, we observe an assortative pattern between the institutional affiliation of cited and citing author that we will return to below.

Table 3
Summary statistics by journal and patent status.

	Nature biotech			Nature material science		
	Overall	Patented	Not Patented	Overall	Patented	Not Patented
Panel A: publication characteristics (N = 1450 original publications)						
N	916	430	486	534	95	439
PUBLICATION YEAR	2000.69	2000.49	2000.86	2004.42	2004.58	2004.39
# AUTHORS	6.86	6.76	6.95	5.77	6.23	5.67
PRIVATE AUTHOR	0.30	0.35	0.26	0.12	0.20	0.10
PATENTED	0.47	–	–	0.18	–	–
Panel B: patent characteristics						
N	430	–	–	95	–	–
GRANT YEAR	2004.1	–	–	2008.5	–	–
PATENT LAG	1474.9	–	–	1441.3	–	–
# INVENTORS	3.3	–	–	3.9	–	–
Panel C: citation year characteristics (N = 11,507 original citation-year observations)						
N	9445	4519	4926	3512	610	2902
CITATIONS	14.11	15.50	12.83	21.25	27.30	19.98
CITATION YEAR	2005.51	2005.42	2005.59	2007.58	2007.65	2007.56
ANNUAL FORWARD CITATIONS PUBLIC	12.21	13.27	11.22	19.76	25.26	18.62
ANNUAL FORWARD CITATIONS PRIVATE	1.90	2.23	1.61	1.48	2.04	1.36
PATENT POST GRANT	0.27	0.564	–	0.044	0.251	–

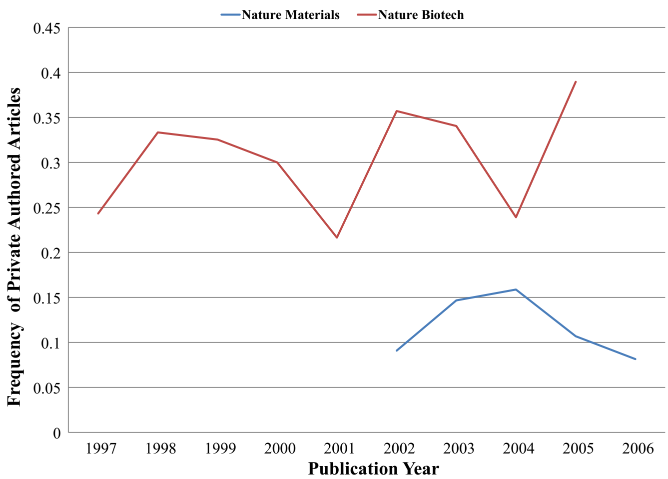

Fig. 2. Frequency of private authored articles by publication year and journal.

5. Results

The empirical core of this paper explores two issues. First, we build upon the prior literature by establishing the overall relationship in our sample between IPR grant and forward citations. We use random effects negative binomial models to evaluate the cross-sectional relationship and then follow with two difference-in-difference regressions to establish baseline results similar to prior work such as Murray and Stern (2007). Second, we move on to the heart of our analysis and examine how the institutional affiliation of both citing and cited authors drives the impact of IPR grant on follow-on citations. Throughout our exposition of the results of our analysis, we report exponentiated coefficients, commonly referred to as incidence-rate ratio (IRR) coefficients. We will focus on the IRR as the coefficient of interest because of its intuitive interpretation as the multiplicative effect on the expected yearly citations to an article.

5.1. Overall relationship between IPR and follow-on citations

Our regression results begin in Table 5 with a series of random effects negative binomial models where ANNUAL FORWARD CITATIONS is the dependent variable. For all the models in Table 5, we use a full set of journal-specific, article-age fixed effects as well as full set of journal specific citation year fixed effects so that we fully control for idiosyncratic calendar year effects for each journal and thus address potential truncation bias. For the first two models in the table, the random effects

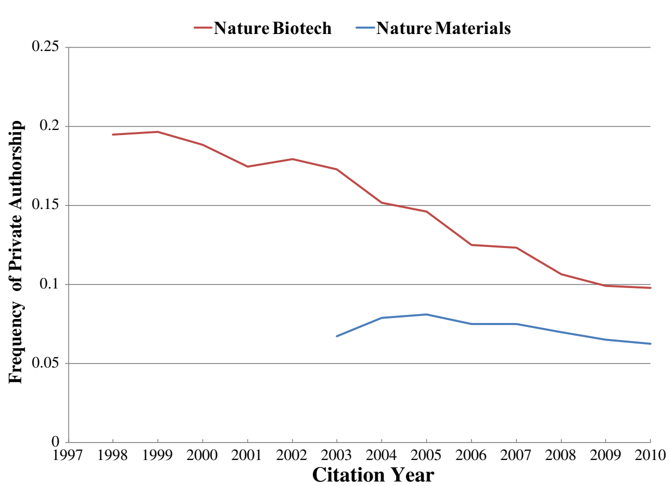

Fig. 3. Frequency of private authorship in forward citations by citation year.

Table 4
Citation-year demand by author demographics and patent status.

	Private, patented	Private, no patent	Public, patented	Public, no patent
Panel A: nature biotech				
CITATIONS	14.65	14.16	15.95	12.37
ANNUAL FORWARD CITATIONS PUBLIC	11.71	11.66	14.10	11.08
ANNUAL FORWARD CITATIONS PRIVATE	2.94	2.51	1.85	1.30
Panel B: nature materials				
CITATIONS	29.93	26.32	26.61	19.22
ANNUAL FORWARD CITATIONS PUBLIC	25.79	23.05	25.12	18.09
ANNUAL FORWARD CITATIONS PRIVATE	4.14	3.27	1.49	1.14

specifications depart from the conditional independence of the fixed effects models described in Section 3.2 and developed throughout the rest of the paper. By introducing the correlated error term of the random effects models, these models allow us to characterize the full relationship between patenting and forward citations (both sorting and treatment effect) before we measure the causal impact of patenting. For more details, see Appendix B. Model (5-1) provides a baseline measure of the association between patenting and ANNUAL FORWARD CITATIONS similar to baseline regressions in prior work on the impact of IPR (Williams, 2010). We find a positive and statistically significant coefficient suggesting that a patent is associated with a 9% increase in yearly citations across all years. The last random effects model (5-2) decomposes the overall patenting association into journal-specific effects of patenting using journal-specific patent grant dummies. We find statistically insignificant but positive associations between patenting and annual forward citations for both journals. Our random effects models capture the cross-sectional demand behavior across journals, but they do not provide a clear causal account of the impact of patent grants on demand–side behavior.

The rest of the models in Tables 5 show the results of our conditional fixed effects models. Each of these models provides a set of

Table 5
Baseline random effects and fixed effects models.[a]

Dep Var = FORWARD CITATIONS	(1) Patenting effect (RE)	(2) Journal specific patenting association (RE)	(3) Journal specific patenting effect (FE)
PATENTED	[1.090*] 0.086* (0.045)		
NB PATENT		[1.083] 0.08 (0.049)	
NM PATENT		[1.115] 0.109 (0.097)	
NB PATENT POST GRANT			[1.068***] 0.065*** (0.021)
NM PATENT POST GRANT			[1.058] 0.056 (0.042)
Observations	11,507	11,507	11,507
Individual article fixed effects	N	N	Y
Journal-age fixed effects	Y	Y	Y
Journal-citation year fixed effects	Y	Y	Y
Log likelihood	−35,057.05	−35,057.01	−26,676.15

* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.
[a] Incident rate ratios in square brackets. (Robust coefficient standard errors in parentheses).

difference-in-difference estimators that makes use of the randomly varying length of patent review allows us to estimate the effects of patent grants on the level of forward citations for the two journals. Model (5-3) provides a baseline estimate of the journal-specific effect of patent grant on ANNUAL FORWARD CITATIONS. The coefficient estimate for the effect of a patent grant on *Nature Biotechnology* articles (NB PATENT POST GRANT) shows a positive, statistically significant effect of patenting across time periods (1.068) corresponding to a 6.8% increase in forward citations resulting from a patent grant. For *Nature Materials*, the coefficient estimate (NM PATENT POST GRANT) is also positive, similar in magnitude, but not statistically significant. While this baseline regression accords substantially to the structure of key regressions in prior research on the impact of IPR (Murray and Stern, 2007; Williams, 2010), our results show evidence for an overall positive impact of patent grant compared to prior results which established a negative effect.

The source of these differences can be seen in Fig. 4 where we explore the time varying-effects of patent grants for articles in each journal. We observe statistically significant negative effects of patenting (less than one in the graph) for articles whose patents were granted in the earliest cohorts of our sample for each journal. Given the delay between patent application and approval, these coefficients show that patented articles published in the earliest issues of both journals were subject to a significant decline in ANNUAL FORWARD CITATIONS resulting from the patent grant. This negative effect of patenting early in the sample trends to a positive (greater than one) but statistically insignificant effect for both journals. Thus, the negative impact of IPR grant on forward citations is largest for an article published in the early years of both journals, before they established a reputation for selecting high-quality science (and, in the case of *Nature Biotech*, corresponding to the Murray and Stern (2007) sample).

5.2. Impact of institutional affiliation and IPR on follow-on citations

We now examine the differing impact of patent grants on researchers with private and public institutional affiliations for both the life science and materials science communities. While the results in Table 5 provide useful evidence showing differences in the impact of patents over time on the aggregate life sciences and materials science communities, our detailed micro-data allowed us to evaluate these issues more precisely by comparing the impact of patents for the demand behavior of follow-on researchers disaggregated across these key subsets of research communities in Pasteur's Quadrant. Specifically, we were interested in whether PATENT POST GRANT has a different impact on the different subpopulations of potential citers in our sample (i.e. whether $\psi_{public} = \psi_{private}$).

The results of this analysis are presented in Table 6. All the regressions have journal-specific article-age and citation-year fixed effects for each subpopulation which addresses potential forward citation truncation bias. Model 6-1 examines how the impact of patenting varies across our two populations. While both public and private downstream research demand show a statistically significant positive increase after the grant of a patent, the magnitude of the increase is larger for private forward citations than public (a 21% increase versus a 5.8% increase). The difference between these coefficients is statistically significant ($p < 0.001$). This result suggests that IPR grant facilitates greater follow-on research from private researchers than public.

Model 6-2 explores whether these effects vary by journal. Similar to the previous regression, we find that patented articles published in *Nature Biotechnology* are associated with statistically significantly higher levels of annual forward citations, but the effect of patenting on private author citations is larger in magnitude (1.22 for articles from private authors compared to 1.066 for articles from public authors). The difference between these coefficients is statistically significant ($p < 0.001$). The coefficients estimating the impact of a patent grant on annual forward citations in *Nature Materials* were also positive, but not statistically significant. The positive coefficients for *Nature Biotechnology* across both subpopulations shows that use of IP in the life-sciences community seems to have a facilitating effect for both sub-populations of the community but provides greater facilitation for researchers from private institutions. Fig. 5 provides time varying estimates of the impact of patent grant across both journals and across the public/private margin. It shows a negative and statistically significant impact of patenting on public citations in the early years of both journals, but a smaller and statistically insignificant effect on private citations.

Table 7 builds upon model 6-1 to ask whether the impact of patent grant on public and private forward citations depends upon the institutional affiliation of the cited author (i.e. whether ψ_{public} and $\psi_{private}$. each vary by type of cited author). Model 7-1 estimates the impact of patent grant for each possible combination of institutional affiliation of cited and citing authors (e.g. private authors citing private papers, private

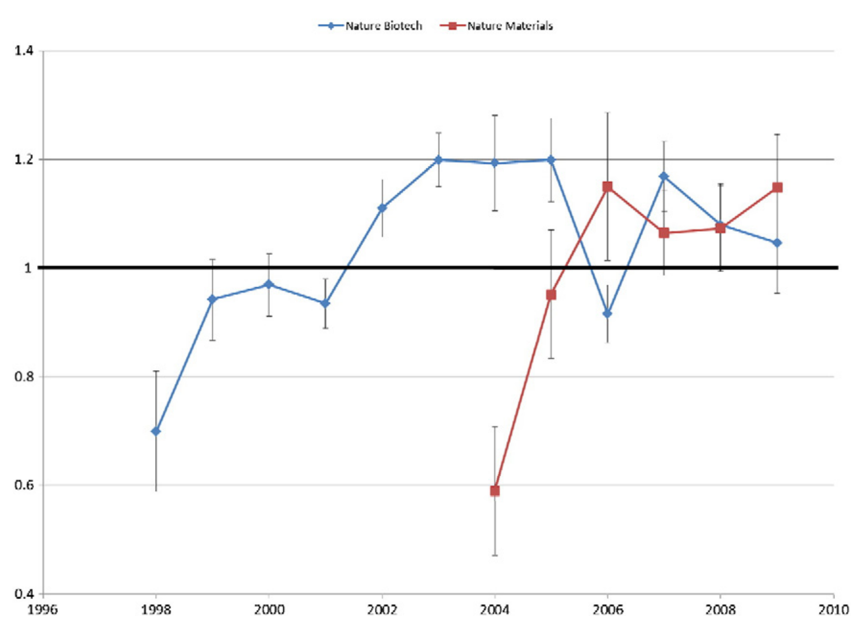

Fig. 4. Coefficient estimates for journal specific patent grant cohort patenting effects.

Table 6
Public/private author margin fixed effects models.[a]

Dep Var = FORWARD CITATIONS	(1) Margin specific patenting effect		(2) Margin specific patenting effect by journal	
	Public citations	Private citations	Public citations	Private citations
PATENT POST GRANT	[1.058**] 0.056** (0.019)	[1.210***] 0.190*** (0.036)		
NB PATENT POST GRANT			[1.061**] 0.059** (0.021)	[1.221***] 0.199*** (0.038)
NM PATENT POST GRANT			[1.044] 0.0438 (0.043)	[1.116] 0.11 (0.107)
Observations	23,014		23,014	
Journal-age fixed effects	Y		Y	
Journal-citation year fixed effects	Y		Y	
Log likelihood	−49,253.06		−34,581.50	

* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.
[a] Incident rate ratios in square brackets (Robust coefficient standard errors in parentheses).

Table 7
Public/private margin fixed effects model controlling for cited paper's institutional affiliation.[a]

Dep Var = FORWARD CITATIONS	(1) Public and private citations by cited author type	
	Public citations	Private citations
PATENT POST GRANT FOR PUBLIC AUTHORS	[1.149***] 0.139*** (0.023)	[1.047] 0.046 (0.037)
PATENT POST GRANT FOR PRIVATE AUTHORS	[0.908***] −0.097*** (0.026)	[1.503***] 0.408*** (0.062)
Observations	23,014	
Journal-age fixed effects	Y	
Journal-citation year fixed effects	Y	
Log likelihood	−43,251.85	

* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.
[a] Incident rate ratios in square brackets (Robust coefficient standard errors in parentheses).

authors citing public papers, etc.). Overall, we observe a strongly assortative pattern where intellectual property grant increases forward citations from authors sharing the same institutional affiliation as the cited author (e.g., public authors citing public papers) than follow-on research across institutional lines (e.g., public authors citing private papers). IPR grant increases the count of private authors citing private papers by 50% while decreasing citations from public researchers by nearly 10% (both coefficients and their difference were significant at the 0.001 level). Similarly, IPR grant increases the annual count of public authors citing public papers by a statistically significant 15% compared to a statistically insignificant 4.7%. The difference between these coefficients is statistically significant ($p < 0.001$). These results show that the positive coefficient estimates in 6-1 mask a great deal of heterogeneity. Patent grant seems to induce a matching process where follow-on research from a patented article seems to draw more heavily from

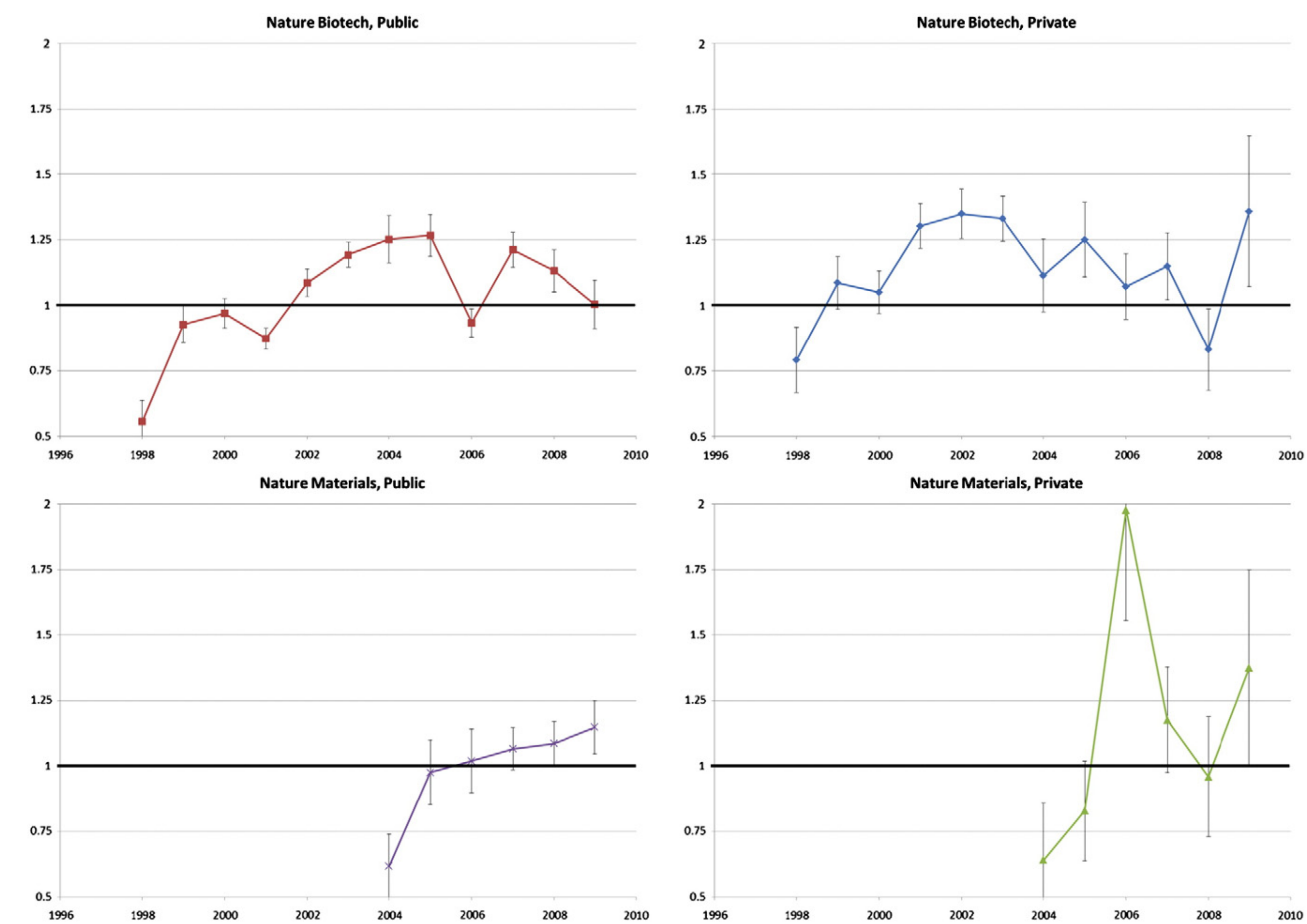

Fig. 5. Coefficient estimates for journal-cohort effect of patenting by public/private margin.

Table 8
Nature platform margin models.[a]

Dep Var = FORWARD CITATIONS	(1) Public and private authored citations broken out by *Nature*			
	Public forward citations		Private forward citations	
	Nature	Not nature	Nature	Not nature
PATENT POST GRANT FOR PUBLIC AUTHORS	[1.218*]	[1.120***]	[1.124]	[1.039]
	0.197*	0.114***	0.117	0.0386
	(0.104)	(0.0238)	(0.187)	(0.0386)
PATENT POST GRANT FOR PRIVATE AUTHORS	[0.844]	[0.910**]	[1.703**]	[1.517***]
	−0.169	−0.0946**	0.532**	0.417***
	(0.106)	(0.0271)	(0.299)	(0.0648)
Observations	25,914			
Journal-age fixed effects	Y			
Journal-citation year fixed effects	Y			
Log likelihood	−57,943.124			

* $p < 0.05$.
** $p < 0.01$.
*** $p < 0.001$.
[a] Incident rate ratios in square brackets (Robust coefficient standard errors in parentheses).

researchers with the same institutional affiliation as the authors of the cited paper.

Next, we asked whether the matching process observed in Table 7 is different for cumulative knowledge exchange within *Nature* publications versus other publications. To do so, we classify each of the forward citations measured in the baseline regressions described in Table 5 as either published in a Nature journal or published in any other journal. We then measure the differences in forward citation rates from articles published in Nature journals versus forward citations from all other journals. Our results are shown in Table 8. Overall, we found the same assortative matching pattern as observed in Table 7: IPR increased forward citations from citing authors who shared their institutional affiliation with the cited paper. Interestingly, the magnitude of this effect seems larger for *Nature* citations. For example, IPR grant increases the count of private authors citing private papers by 70% within *Nature* publications compared to a 50% increase for all other journals. Due to the small size of the sample of *Nature* articles, the Wald Test comparing the size of these coefficients is statistically insignificant. Similarly, IPR grant increases the count of public authors citing public papers by 22% for *Nature* articles compared to 12% for all other journals.

Our results suggest that IPR grant not only changes the demand for a piece of knowledge (as measured by follow-on citations) but creates greater incentives for research that is an "institutional match" (i.e. both citing and cited research have the same institutional affiliation). Furthermore, the resulting matching process seems particularly acute for follow-on contributions in *Nature* publications. By shaping a community of researchers across institutional settings through their editorial policies and reputation for quality, journals mediate the impact of IPR on downstream use across the public/private institutional divide.

6. Conclusions & discussion

This paper provides an empirical examination of the relationship between IPR and the exchange of scientific knowledge across a critical institutional divide (public versus private researchers) for fields in Pasteur's Quadrant. In particular, we examine the differing impact of IPR grant on the follow-on citations from public and private researchers. We document how the differing impact of IPR grant on the rate of follow-on citations from these different subgroups is driven by the institutional affiliation of the cited authors of the original paper: IPR grant leads to an assortative matching process whereby the patent grant increases forward citations from authors sharing the same institutional affiliation more than research across institutional lines.

These results suggest IPR grant impacts not only the amount of downstream cumulative research but also the organization of downstream projects through their institutional location (public versus private). The welfare implications of this matching process hinge largely on the specific channels through which it shifts project choices of researchers in private and public settings. A patent grant provides newly disclosed information to follow-on researchers and new control rights to the original author. New information changes the ex ante tradeoffs between potential new projects whereas new control rights could introduce new contracting frictions in a marketplace for ideas which might inhibit a high value project from being undertaken. If, for example, public scientists more easily access patented research from a university (because of the norms of open science) than from a private firm (because university legal resources are specialized in licensing IP out of rather than into the University), then a patented discovery made in a private lab might experience a very different mix of follow-on research than it would have if it were discovered in a public lab (Aghion et al., 2008; Bikard, 2012). We remain cautious in the interpretation of our results since we are not able to identify which of these mechanisms are at play in our setting.

Our results are relevant for scholars in a growing literature on the interaction of intellectual property and institutions with platform characteristics in contexts ranging from SSOs, to the growth of exchange traded financial products, to the impact of biological resource centers (Furman and Stern, 2011; Lerner and Tufano, 2011; Rysman and Simcoe, 2008). While these studies draw from wide-ranging phenomena, they attempt to measure the impact of innovation institutions on the formation of cumulative knowledge while addressing the importance of complementary institutions (like the patent system) for shaping the long-run success of these institutions. Thus, our findings are informative for the broader debate about the role that intellectual property plays in shaping the structure and impact of knowledge platforms.

We also view this article as the first thrust in a larger effort to characterize the time-varying dynamics of journals in important scientific fields located squarely in Pasteur's Quadrant. While our choice of two related NPG journals allowed us to control for a good deal of unobserved heterogeneity in order to observe some of the basic dynamics of journals as two-sided platforms for knowledge diffusion, our study design limited us to a focus on the demand side of the platform for our key regressions. We intend follow-on studies to remedy this gap by broadening our perspective away from just NPG journals to the key journals serving as conveyors of knowledge, simultaneously basic and applied, in the fields of materials science and biotechnology. In so doing, we hope to make more precise statements about the ways in which the mechanisms outlined in the two-sided platform literature shape the growth of cumulative knowledge in Pasteur's Quadrant.

Appendix A. Example of a patent–paper pair

It is useful to consider a particular patent–paper pair to provide a sense of the relationship between publications and patents in practice. Consider the following patent–paper pair from biotechnology drawn directly from the description in Murray and Stern (2007):

> "A method has been developed for control of molecular weight … during production of polyhydroxyalkanoates in genetically engineered organisms by control of the level and time of expression of one or more PHA synthases… The method was demonstrated by constructing a synthetic operon for PHA production in *E. coli* … Modulation of the total level of PHA synthase activity in the host cell by varying the concentration of the inducer …was found to affect the molecular weight of the polymer produced in the cell.".
> [(Snell; K. D.; Hogan; S. A.; Sim; SJ; Sinskey; A. J.; Rha; C.. 1998, Patent 5,811,272)]

"A synthetic operon for polyhydroxyalkanoate (PHA) biosynthesis designed to yield high levels of PHA synthase activity in vivo was constructed .... Plasmids containing the synthetic operon ... were transformed into *E. coli* DH5 alpha and analyzed for polyhydroxybutyrate production... Comparison of the enzyme activity levels of PHA biosynthetic enzymes in a strain encoding the native operon with a strain possessing the synthetic operon indicates that the amount of polyhydroxyalkanoate synthase in a host organism plays a key role in controlling the molecular weight and the polydispersity of polymer."

[(Sim SJ, Snell KD, Hogan SA, Stubbe J, Rha CK, Sinskey AJ, Nature Biotechnology 1997)]

There are clear similarities between the technical language in these two passages. Their similarity was the final characteristic in a carefully constructed algorithm to determine whether a paper had a paired patent. For each paper in our sample, we searched to find all patents within a time window which were filed with an inventor with the same name and geographic location as one of the authors (state for US authors and Country otherwise). Next, we ranked each of these patents according to a similarity metric using a bag-of-words approach weighted by the inverse frequency of the words in the abstracts of the articles in our sample. Next, we carefully compared the wording of the patent and the article moving from most similar to least, stopping when a pair was found.

Appendix B. Detailed description of regression models

Our empirical specification predicted the annual count of forward citations in publications in our two journal sample over the period 1997–2010. This dependent variable took the form of count data skewed to the right. Observations in our data were defined at the citation-year level for each publication. We coded a set of variables to test our hypotheses. We defined an indicator variable, PATENT POST-GRANT, equal to one in years after the patent grant year for each paper that received a patent. When we observe the number of citations a paper receives before and after the grant of a patent, we are able to identify how the average pattern of citations to a paper changes after the introduction of a patent.

In the random-effects models, we estimated (RE 1):

$$CITES_{i,j,pubyear(j),t} = \mathrm{f}\left(\eta_i, \varepsilon_{i,j,t};\ \gamma + \beta_t + \delta_{t-pubyear} + \psi\ PATENTED_{i,t}\right)$$

β_t is a year effect, $\delta_{t-pubyear}$ captures the age of the article, and PATENTED is a dummy variable equal to one only for articles which have received a patent.[9] In our random effects models, we include two error terms (η_i and $\varepsilon_{i,j,t}$) to represent the two sources of error in the maximum likelihood estimator. In these models, η_i represents the error associated with the random effect, modeled by the beta distribution, which enters the marginal portion of the likelihood function (loosely, the between article variation). In contrast, $\varepsilon_{i,j,t}$ captures the conditional portion of the likelihood function (roughly, the within article variation) (Cameron and Trivedi, 2013; Hausman et al., 1984). In the fixed effects models that follow, η_i is conditioned out of the model through the use of the article-level fixed effects. For a more detailed account of the derivation of the model, see especially the treatment in Cameron and Trivedi (2013). The coefficient on PATENTED (ψ) in this model provides a baseline measure of the association between patenting and annual forward citations.

[9] This baseline analysis assumes that age fixed effects associated with citations do not depend on whether a paper receives a patent. In particular, a key assumption of our base model is that patented articles are not simply "shooting stars"—articles that, for exogenous reasons, experience a high rate of early citation followed by a rapid decline.

To examine the potentially different effects of patenting across the two journals, we modified specification (RE 1) to account for journal specific differences. Specifically, we estimated (RE 2):

$$CITES_{i,j,pubyear(j),t} = \mathrm{f}\left(\eta_i, \varepsilon_{i,j,t};\ \gamma + \beta_t + \delta_{t-pubyear} + \psi_j PATENTED_{i,t,j}\right)$$

In this equation, ψ_j measures the specific association between patenting and annual forward citations for each journal. ψ_j represents an aggregate measure of the relationship between patenting and forward citations across years for each journal.

In the next section of regressions, we moved from random effects models to fixed effects models as we attempted to make more precise statements about the varying impact of patent grants over time. We began the analysis by adapting (RE 2) into a fixed effects specification. Specifically, we looked at (FE 1):

$$CITES_{i,j,pubyear(j),t} = \mathrm{f}\left(\varepsilon_{i,j,t};\ \gamma_i + \beta_t + \delta_{t-pubyear} + \psi_j POST-GRANT_{i,t,j}\right)$$

Note that here that γ_i is a fixed effect for each article. As discussed above, this equation measures the impact of a patent grant on annual forward citations. With the inclusion of fixed effects, we have created a difference-in-difference estimator. Here, the coefficient on POST-GRANT (ψ) estimated the marginal impact of the intervention on the set of treated articles. Thus, ψ measured the impact of patent grant accounting for fixed differences in the citation rate across articles and relative to the non-parametric trend in citation rates for articles with similar characteristics. We introduced a set of non-parametric cohort effects into (FE 1) in order to identify the time-varying marginal impact of patenting. We used patent-year cohorts to estimate (FE 2):

$$\begin{aligned} CITES_{i,j,pubyear(j),t} = \mathrm{f}(&\varepsilon_{i,j,t};\ \gamma_i + \beta_t + \delta_{t-pubyear} \\ &+ \sum_{patyear=1998}^{2009} \psi_{j,patyear}\ NB\ POST-GRANT_{i,t,patyear} \\ &+ \sum_{patyear=2004}^{2009} \psi_{j,patyear}\ NM\ POST-GRANT_{i,t,patyear}) \end{aligned}$$

In this regression, the set of $\psi_{j,patyear}$ coefficients identity the separate marginal effects of patenting or each of the journal-specific patent-year cohorts in our sample. We used patent grant year cohorts because these allowed us to identify how the causal impact of patent grants varied over time.

Our next set of regressions explores the different marginal responses to patent grant in two critical populations in Pasteur's Quadrant: private and public authors. We simultaneously estimate the impact of a patent grant on follow-on citations by public and private authors. Specifically, we estimated (PUBLIC/PRIVATE AUTHOR MARGIN 1):

$$CITES_{i,j,pubyear(j),t,margin} = \mathrm{f}\left(\varepsilon_{i,j,t,margin};\ \gamma_i + \beta_t + \delta_{t-pubyear} + \psi_{margin}\ POST-GRANT_{i,t}\right)$$

For this and each of the subsequent regressions with a subscript "margin", we add this subscript to denote the simultaneous estimation of the marginal impact of patenting on two different populations. In this specification, ψ_{margin} measures the impact of patent grant on the downstream demand behavior of researchers with public and private affiliations.

Building upon this regression, we next explore whether the downstream demand behavior of these two groups varies across the research communities utilizing the two journals. To do so, we separately estimate the marginal effect of patenting on the two groups for both journals. Specifically, we simultaneously estimate (PUBLIC/PRIVATE AUTHOR MARGIN 2):

$$CITES_{i,j,pubyear(j),t,margin} = \mathrm{f}\left(\varepsilon_{i,j,t,margin};\ \gamma_i + \beta_t + \delta_{t-pubyear} + \psi_{margin,\ j}\ POST-GRANT_{i,t}\right)$$

In this regression, $\psi_{margin,\ j}$ captures the potentially heterogenous impact of patent grants for each of the four downstream research populations defined by the two journals and the two types of authors.

Next, we explored the potential time variation in the public/private margin across research communities by creating a set of journal specific patent year cohorts for both the public and private margins. Specifically, we simultaneously estimate (PUBLIC/PRIVATE AUTHOR MARGIN 3):

$$CITES_{i,j,pubyear(j),t,margin} = f(\varepsilon_{i,j,t};\ \gamma_i + \beta_t + \delta_{t-pubyear} + \sum_{patyear=1998}^{2009} \psi_{j,patyear,margin}\ NB\ POST-GRANT_{i,t,patyear} + \sum_{patyear=2004}^{2009} \psi_{j,patyear,margin}\ NM\ POST-GRANT_{i,t,patyear})$$

In this regression, the point estimate of $\psi_{j,patyear,margin}$ gives the potentially heterogenous time varying impact of patenting on the downstream demand by journal by researchers from different backgrounds.

Next, we explored the potential differences in downstream demand to patented research originating from different institutional locations. We did this by creating a dummy variables, PUBLIC PATENT POST GRANT, equal to one for articles with only public authors in years after the patent grant year for each paper that received a patent and a dummy variable, PRIVATE PATENT POST GRANT, equal to one for articles with at least one private author in years after the patent grant year for each paper that received a patent. To understand if there is downstream demand for the different types of articles supplied varies by researcher type, we next jointly estimate the impact of patent grant for downstream citations from private and public authors for articles with different supply characteristics. Specifically, we estimate the regression equation:

$$CITES_{i,j,pubyear(j),t,margin} = f(\varepsilon_{i,j,t,margin};\ \gamma_i + \beta_t + \delta_{t-pubyear} + \psi_{margin,\ public}\ PUBLIC\ POST-GRANT_{i,t} + \psi_{margin,\ private}\ PRIVATE\ POST-GRANT_{i,t})$$

This regression equation identifies different marginal responses for each of the four supply and demand cases. This regression allows us to understand if downstream researchers respond differently to different types of research depending on their institutional location.

Lastly, we explore whether there are differences in the marginal response to patenting for researchers attempting to publish in Nature journals versus other journals. Specifically, we code each downstream citation for whether or not it appeared in a Nature Journal. We then estimated the regression equation (NATURE PLATFORM MARGIN 1):

$$CITES_{i,j,pubyear(j),t,margin} = f\left(\varepsilon_{i,j,t,margin};\ \gamma_i + \beta_t + \delta_{t-pubyear} + \psi_{margin,\ j}\ POST-GRANT_{i,t}\right)$$

Here, the different estimates of $\psi_{margin,\ j}$ identify the impact of patenting for downstream citations on and off the Nature platform.

References

Aghion, P., Dewatripont, M., Stein, J.C., 2008. Academic freedom, private-sector focus, and the process of innovation. RAND J. Econ. 39 (3), 617–635.

Arora, A., Fosfuri, A., Gambardella, A., 2004. Markets for Technology: The Economics of Innovation and Corporate Strategy. The MIT Press.

Bikard, Michaël, 2012. Is knowledge trapped inside the Ivory Tower? Technology spawning and the genesis of new science-based inventions. Sloan Working Paper.

Cameron, A. Colin, Trivedi, Pravin K., 2013. Regression Analysis of Count Data. Cambridge university press, p. 53.

Cohen, W.M., Nelson, R.R., Walsh, J.P., 2000. Protecting Their Intellectual Assets: Appropriability Conditions and Why US Manufacturing Firms Patent (or Not). National Bureau of Economic Research.

Collins, H.M., 1974. The TEA set: tacit knowledge and scientific networks. Sci. Stud. 4 (2), 165–185.

Dasgupta, Partha, David, Paul, 1994. Toward a new economics of science. Res. Policy 23, 487–521.

David, Paul, 2004. Can 'open science' be protected from the evolving regime of IPR protections? J. Inst. Theor. Econ. Z. Die Gesamte Staatswiss. 160 (1), 9–34. http://dx.doi.org/10.1628/093245604773861069.

De Solla Price, D.J., 1965. Networks of scientific papers. Science 149 (3683), 510–515.

Ducor, P., 2000. Intellectual property—coauthorship and coinventorship. Science 289 (5481), 873–875. http://dx.doi.org/10.1126/science.289.5481.873.

Edelman, L.B., Suchman, M.C., 1997. The legal environments of organizations. Annu. Rev. Sociol. 23, 479–515. http://dx.doi.org/10.1146/annurev.soc.23.1.479.

Furman, J.L., Stern, Scott, 2011. Climbing atop the shoulders of giants: the impact of institutions on cumulative knowledge production. Am. Econ. Rev. 101 (5), 1933–1963.

Gans, Joshua, Stern, Scott, 2000. Incumbency and R&D incentives: licensing the gale of creative destruction. J. Econ. Manag. Strateg. 9 (4), 485–511. http://dx.doi.org/10.1162/105864000567945.

Gans, Joshua, Fiona Murray, and Scott Stern. 2008. "Patents, Papers, Pairs & Secrets: Contracting over the Disclosure of Scientific Knowledge". Unpublished.

Garfield, E., Merton, R.K., 1979. Citation Indexing: Its Theory and Application in Science, Technology, and Humanities. vol. 8. Wiley, New York.

Geim, A.K., Novoselov, K.S., 2007. The rise of graphene. Nat. Mater. 6 (3), 183–191. http://dx.doi.org/10.1038/nmat1849.

Hagstrom, W.O., 1975. The Scientific Community. Southern Illinois University Press, Carbondale.

Hall, B.H., Jaffe, A.B., Trajtenberg, M., 2001. The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools. National Bureau of Economic Research.

Hausman, Jerry A., Hall, Bronwyn H., Griliches, Zvi, 1984. Econometric Models for Count Data with an Application to the Patents–R&D Relationship. National Bureau of Economic Research, Cambridge, Mass., USA, (http://www.nber.org/papers/t0017).

Heller, M., Eisenberg, R.S., 1998. Can patents deter innovation? The anticommons in biomedical research. Science 280 (5364), 698.

Hellmann, T., 2007. The role of patents for bridging the science to market gap. J. Econ. Behav. Organ. 63 (4), 624–647.

Jaffe, A.B., Trajtenberg, M., Henderson, R., 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. Q. J. Econ. 108 (3), 577–598.

Kieff, F., 2005. IP Transactions: On the Theory & Practice of Commercializing Innovation. Washington U. School of Law Working Paper, Stanford Law and Economics Olin Working Paper. No. 311.

Kitch, E.W., 1977. The nature and function of the patent system. J. Law Econ. 20, 265.

Lerner, J., Tirole, J., 2006. A model of forum shopping. Am. Econ. Rev. 96 (4), 1091–1113.

Lerner, J., Tufano, P., 2011. The Consequences of Financial Innovation: A Counterfactual Research Agenda. National Bureau of Economic Research.

Lessig, L., 2004. The creative commons. Mont. L. Rev. 65, 1.

McCabe, Mark, Snyder, Christopher, 2005. Open access and academic journal quality. Am. Econ. Rev. 95 (2), 453–458. http://dx.doi.org/10.1257/000282805774670112.

McCabe, Mark, Snyder, Christopher, 2007. Academic journal prices in a digital age: a two-sided market model. B.E. J. Econ. Anal. Policy 7 (1).

Mokyr, J., 2004. The Gifts of Athena: Historical Origins of The Knowledge Economy. Princeton Univ Press.

Mowery, David, Ziedonis, Arvids, 2007. Academic patents and materials transfer agreements: substitutes or complements? J. Technol. Transf. 32 (3), 157–172. http://dx.doi.org/10.1007/s10961-006-9011-1.

Murray, Fiona, 2002. Innovation as co-evolution of scientific and technological networks: exploring tissue engineering. Res. Policy 31 (8–9), 1389–1403. http://dx.doi.org/10.1016/S0048-7333(02)00070-7.

Murray, Fiona, 2010. The oncomouse that roared: hybrid exchange strategies as a source of distinction at the boundary of overlapping institutions. Am. J. Sociol. 116 (2), 341–388. http://dx.doi.org/10.1086/653599.

Murray, Fiona, Stern, Scott, 2007. Do formal intellectual property rights hinder the free flow of scientific knowledge?: an empirical test of the anti-commons hypothesis. J. Econ. Behav. Organ. 63 (4), 648–687.

Murray, Fiona, Aghion, P., Dewatripont, M., Kolev, J., Stern, Scott, 2009. Of Mice and Academics: Examining the Effect of Openness on Innovation. NBER Working Paper.

Novoselov, K.S., Geim, A.K., Morozov, S.V., Jiang, D., Zhang, Y., Dubonos, S.V., Grigorieva, I.V., Firsov, A.A., 2004. Electric field effect in atomically thin carbon films. Science 306 (5696), 666–669. http://dx.doi.org/10.1126/science.1102896.

Rysman, Marc, Simcoe, Timothy, 2008. Patents and the performance of voluntary standard-setting organizations. Manag. Sci. 54 (11), 1920–1934. http://dx.doi.org/10.1287/mnsc.1080.0919.

Schubert, A., Braun, T., 1993. Reference standards for citation based assessments. Scientometrics 26 (1), 21–35.

Shapiro, Carl, 2001. Navigating the Patent Thicket: Cross Licenses, Patent Pools, and Standard Setting. Innovation Policy and the Economy. vol. 1. MIT press, pp. 119–150 (http://www.nber.org/chapters/c10778.pdf).

Simcoe, Timothy, Graham, S.J.H., Feldman, M., 2007. Competing on Standards? Entrepreneurship, Intellectual Property and the Platform Paradox. National Bureau of Economic Research.

Stokes, D.E., 1997. Pasteur's Quadrant: Basic Science and Technological Innovation. Brookings Inst Pr.

Swedish Royal Academy, 2010. The Nobel Prize in Physics 2010. Press Release, (http://www.nobelprize.org/nobel_prizes/physics/laureates/2010/press.html).

Tannock, Q., 2011. Exploiting carbon flatland. Nat. Mater. 11 (1), 2–5.

Walsh, John P., Cohen, Wesley M., Cho, Charlene, 2007. Where excludability matters: material versus intellectual property in academic biomedical research. Res. Policy 36 (8), 1184–1203. http://dx.doi.org/10.1016/j.respol.2007.04.006.

Williams, H., 2010. Intellectual property rights and innovation: evidence from the human genome. NBER Working Paper.