Do Formal Intellectual Property Rights Hinder
the Free Flow of Scientific Knowledge?
An Empirical Test of the Anti-Commons Hypothesis[*]

Fiona Murray
MIT Sloan School

Scott Stern
Northwestern University and NBER

June 2005

Prepared for the
NBER Academic Science and Entrepreneurship Conference

Do Formal Intellectual Property Rights

Hinder the Free Flow of Scientific Knowledge?

A Test of the Anti-Commons Hypothesis

Fiona Murray
Scott Stern

June 2005

ABSTRACT: While the potential for intellectual property rights to inhibit the diffusion of scientific knowledge is at the heart of several contemporary policy debates, evidence for the "anti-commons effect" has been anecdotal. A central issue in this debate is how intellectual property rights over a given piece of knowledge affects the propensity of future researchers to build upon that knowledge in their own scientific research activities. This article frames this debate around the concept of dual knowledge, in which a single discovery may contribute to both scientific research and useful commercial applications. A key implication of dual knowledge is that it may be simultaneously instantiated as a scientific research article and as a patent. Such patent-paper pairs are at the heart of our empirical strategy. We exploit the fact that patents are granted with a substantial lag, often many years after the knowledge is initially disclosed through paper publication. The knowledge associated with a patent-paper pair therefore diffuses within two distinct intellectual property environments – one associated with the pre-grant period and another after formal IP rights are granted. Relative to the expected citation pattern for publications with a given quality level, anti-commons theory predicts that the citation rate to a scientific publication should fall after formal IP rights associated with that publication are granted. Employing a differences-in-differences estimator for 169 patent-paper pairs (and including a control group of publications from the same journal for which no patent is granted), we find evidence for a modest anti-commons effect (the citation rate after the patent grant declines by between 9 and 17%). This decline becomes more pronounced with the number of years elapsed since the date of the patent grant, and is particularly salient for articles authored by researchers with public sector affiliations.

I.   Introduction

In the early 1980s, Professor Phil Leder, recently recruited to head the new Genetics department at the Harvard Medical School, developed the first genetically engineered mouse, dubbed the Oncomouse. Leder and his post-doc Tim Stewart had used novel transgenic techniques to insert an oncogene into a mouse embryo; the result was a mouse that was highly susceptible to cancer. Using the mice to examine the importance of genes in the onset of cancer, Leder came to recognize that "it could serve a variety of different purposes, some purely scientific others highly practical" (Kevles, 2002, p. 83). This research was published in *Cell* in 1984, and, in 1988 a broad patent for the Oncomouse was granted by the US Patent Office (USPTO). The Oncomouse patent was more controversial than most; not only was the Oncomouse the first living mammal to be patented, but Harvard's licensee Dupont aggressively enforced the property rights, including demands for "reach-through" rights and review of publications that used the Oncomouse in further scientific research. The generation of scientific ideas such as the Oncomouse - ideas that are simultaneously of value as a scientific discovery and as a useful, inventive construct - is increasingly widespread, particularly in the disciplines that underpin modern biotechnology (Stokes, 1997)

. Dual-purpose ideas provide their originators with multiple disclosure choices, including an option to both publish research in the scientific literature *and* obtain intellectual property rights (IPR) over that knowledge. In academia, an increasing number of scientists haven chosen this path of dual knowledge disclosure – which we describe as patent-paper pairs (Murray, 2002; Ducor 2000). Patent-paper pairs are linked scientific articles and individual patents which disclose (and serve as a property right over) the same underlying "piece" of knowledge.   Patent-paper pairs are thus more than simply a reflection of the rise in patenting in academia over knowledge unrelated (or only tangentially related) to scientific research. Rather, by embedding the same piece of knowledge in two distinct institutional regimes, they instantiate the expansion of formal intellectual property rights over knowledge that was traditionally disclosed solely through scientific publication.

The increased use of intellectual property rights (IPR) in scientific research has sparked a vigorous academic and policy debate over what is known as the "anti-commons effect." At its core, the anti-commons debate is an argument over whether such expansion of IPR (in the form

of patents and/or copyrights) is "privatizing" the scientific commons, and limiting the benefits from scientific progress (Heller and Eisenberg, 1998; Argyres & Liebskind, 1998; David, 2001b). Specifically, the anti-commons hypothesis states that IPR may inhibit the free flow and diffusion of scientific knowledge and the ability of researchers to cumulatively build on each other's discoveries (Heller & Eisenberg, 1998; David, 2003, 2000; Lessig, 2002; Etzkowitz, 1998; Krimsky, 2003). Conversely, IPR may facilitate the creation of a market for ideas and mitigate disincentives to disclose and exchange knowledge which might otherwise remain secret (Merges & Nelson, 1990, 1994; Arora, Fosfuri, and Gambardella, 2001; Gans and Stern, 2003).

Building on several key case examples in the literature, this article frames the anti-commons debate around the concept of dual knowledge as instantiated in patent-paper pairs. Building on this framework we then report a novel empirical strategy to evaluate the salience of the anti-commons effect on the cumulative impact of scientific knowledge. Our strategy exploits the existence of patent-paper pairs as not only the instantiation of the expansion of intellectual property rights over dual knowledge but also a concrete and specific starting point from which to identify the impact of such rights on the rate of scientific knowledge diffusion.

Our approach is to compare patterns of scientific citations to scientific articles that are part of patent-paper pairs, relative to citation patterns for articles that are *not* part of patent-paper pairs (but are similar along other dimensions). This allows us to evaluate several key hypotheses at the center of the anti-commons debate. First, we evaluate whether citation patterns are different for scientific research which is ultimately also patented. In other words, to what extent does published scientific knowledge disclosed as a patent-paper pair differ in its future cumulative impact on public domain research (as measured by forward citations to the publication) to papers that are similar in topic, published in the same journal in the same time period, but never receive IPR? Second, we take advantage of *patent grant delay*. While publication lags are usually modest (on the order of a few months), patent grant delays are substantial (in most cases IPR is usually granted two to four years after initial application). Consequently, scientific knowledge associated with a patent-paper pair diffuses under two distinctive institutional environments – a pre-grant period where no IP rights are present and a post-grant period in which specific property rights have been granted. This difference allows us to ask: how does the *grant* of formal patent rights over such knowledge influence the trajectory

of forward citations and therefore the impact of the scientific research findings in the public domain?

The "experiment" afforded by the combination of patent-paper pairs and patent grant delay allows for a set of precise tests motivated by the anti-commons hypothesis: if the grant of intellectual property hinders the ability of researchers to build (in the public domain) on a given piece of knowledge, then the citation rate to the scientific publication disclosing that knowledge should be lower than for scientific publications with no IP and should fall after formal property rights are granted. Of course, such an analysis must control for the fact that citation patterns vary with the underlying quality of the article and with the time elapsed since publication. Our use of patent grant delay allows us to do so. Specifically, by observing a *given* piece of knowledge in two different institutional environments, we are able to evaluate how differences in the institutional environment affect the diffusion of a given piece of knowledge. In our empirical work, we therefore implement a differences-in-differences estimator, with fixed effects for each article in our sample. To evaluate the anti-commons hypothesis, we examine how the grant of IPR *changes* the citation rate to scientific articles, accounting for fixed differences in citation rates across articles and relative to the trend in citation rates for articles with similar characteristics.

Our sample is composed of 340 peer-reviewed scientific articles appearing between 1997 and 1999 in *Nature Biotechnology*, perhaps the leading publication for research exhibiting knowledge duality in the life sciences. The incidence of patent-paper pairs is quite high within this sample: For just under 50% of the scientific articles in our sample, a US patent has been granted over the knowledge covered in that publication. As well, for those articles which ultimately receive a patent, there is a significant lag between scientific publication and patent grant (on average, 37.5 months). We exploit these data to establish three core findings. First, published articles also associated with formal IP are more highly cited than those whose authors choose not to file for patents; however this can largely be accounted for by observed characteristics such as author location and number of authors on the article. Second, there is robust evidence for a quantitatively modest but statistically significant anti-commons effect; across different specifications, the article citation rate declines by 9 to 17% after a patent grant. Thirdly, the anti-commons effect is particularly salient for articles with US and public sector co-authors. While we are cautious in our interpretation of the precise mechanisms that give rise to

this dampening effect on citations, our evidence suggests that, at a broad level, the anti-commons effect seems to have a sound empirical basis, although the size of the effect may be modest.

The remainder of the paper is as follows. The next section reviews the economic foundations of the anti-commons hypothesis, framed around the nature and institutional foundations of dual knowledge. Section III develops the empirical test and predictions of the anti-commons hypothesis. After a review of the data in Section IV, Section V presents our empirical findings. We conclude with implications for future empirical work and policy development.

II.   The Anti-commons Effect

The anti-commons debate focuses on the impact of intellectual property rights over knowledge that was traditionally placed in the public domain through scientific disclosure.[1] According to the traditional literature a clear distinction should be drawn between basic and applied research, with basic research being focused on questions of fundamental scientific interest, and applied research focused on questions of usefulness and applications (Stokes, 1997). Scholarship in the "new" economics of Science compares alternative institutional arrangements associated with these distinctive types of knowledge, and assumes strongly exogenous institutional regimes and disclosure choices (Dasgupta and David 1994). Basic research is undertaken in the institutions of Open Science and disclosed through publication; applied research takes place in a Private Property regime and is disclosed through patents (to the extent that disclosure takes place).

Knowledge production in the Open Rights regime ("Science") is characterized (in the stylized case) by a distinctive set of economic incentives for cumulative knowledge production, including the adoption of norms that facilitate full disclosure and diffusion of knowledge. This system includes the recognition of scientific priority and a system of public (or coordinated) expenditures to reward those who contribute to cumulative knowledge production over the long term (Merton, 1973; Dasgupta and David, 1994). By premising career rewards (such as tenure) on disclosure through publication, Open Science leverages the public goods nature of basic research and therefore promotes cumulative innovation and "standing on the shoulders of

---

[1] In contrast, much knowledge produced in industry is often maintained as a trade secret, or only disclosed through the patent system. However, an increasing amount of industry-produced knowledge is also disclosed through scientific publication, often to serve specific strategic purposes.

giants."[2] In contrast, the incentives that govern the private property rights regime pay little attention to the basis of a researcher's ultimate impact on follow-on research, but rather depend on the degree to which a researcher can *exclude* others and so appropriate some of the value created by their knowledge through the commercialization of new technology (Nelson 1959; Arrow 1962; Levin et al. 1987; Kremer 1997; Scotchmer 1987).

While concise in its formulation of the relationship between the nature of knowledge and incentives for its production and distribution, this analytical framework fails when knowledge has *both* basic and applied value. By highlighting its *duality,* Stokes (1997) reformulated the traditional distinction between basic and applied research; a single discovery could simultaneously possess both applied and basic characteristics. Figure A illustrates this essential insight. Stokes' formulation allows for "use-inspired basic research," as exemplified by the research activities of scientists such as Pasteur (hence the term Pasteur's Quadrant). Like Leder's Oncomouse research in the 1980s, Pasteur's fundamental insights into microbiology served both as a foundation for the germ theory of disease and had practical application for cholera and rabies (Geison 1995; Stokes 1997).

For knowledge generated in Pasteur's Quadrant disclosure becomes endogenous: scientists have a well-defined set of disclosure choices, including publication in scientific journals and application for protection through formal intellectual property rights. Moreover, these choices are *not* mutually exclusive – a given piece of knowledge can both be disclosed through scientific publication and be protected by intellectual property rights. The phenomena of "dual knowledge – dual disclosure" is perhaps most clearly instantiated in the context of "patent-paper pairs" – linked scientific articles and individual patents which disclose (and serve as a property right over) the same underlying "piece" of knowledge.[3] Patent-paper pairs are thus more than simply a reflection of the rise in patenting in academia over knowledge unrelated (or only tangentially related) to scientific research. Rather, they embed the same piece of knowledge

---

[2] While closely associated with university research, Open Science is also feasible (and profitably adopted) by private firms, including many within industries dependent on the life sciences (Cockburn and Henderson, 1998; Zucker et al., 1998; Stern, 2004; Murray, 2002).

[3] There is also evidence for increased secrecy and delay in publication as a result of increased economic incentives in academia, the disclosure requirements of patent applications in the US and Europe, and the strictures placed on academics when their research is funded by private industry (Blumenthal et al. 1996, Campbell et al. 2000). The growing privatization of basic research may also be associated with refocusing of research agendas (Thursby and Thursby, 2003; David 2003) and an increased potential for bias in research results (Nelkin 1984; Krimsky 2003).

in two distinct institutional regimes, each of which is governed by specific (and potentially opposing) rules for access and attribution.

The Anti-Commons Debate in Molecular Biology

While not limited to life sciences, many of the key issues that animate the anti-commons discussion surround the interaction between public and private knowledge exploitation in areas related to biotechnology (Kenney, 1986; Orsenigo, 1989; Powell et al., 1996; Gambardella, 1995). The shift towards dual knowledge as embodied in dual institutions reflects at least three related forces: The expanding promise of biotechnology, reductions in the costs of academic patenting, and increases in the scope of IPR over knowledge produced in the life sciences. The rise in dual knowledge in this field can be traced back to the early 1970s; most key scientific milestones in the field, but for noted exceptions, have been disclosed as patent-paper pairs[4]:

- The techniques of recombinant DNA which provide insights into the cellular machinery of the cell but also laid the foundation for the production of recombinant therapeutic proteins (Cohen et al., 1973),
- The Oncomouse simultaneously provided great insight into the sources of cancer while becoming a model for investigating cancer therapies (Stewart et al. 1984),
- The discovery of RNA interference represented a further step towards explaining DNA replication but also the foundation of a potentially new therapeutic category (Zamore et al. 2000),
- The development of embryonic stem cells teaching us how cells develop but also having the potential to serve as novel therapeutics or the foundations of organ replacement (Thomson et al. 1998).

At the same time, policy shifts encouraged academics to claim IPR over their dual knowledge. Prior to this time, patent applications filed by universities on behalf of investigators required case-by-case negotiation of the assignment of patent rights and their subsequent licensing. The

---

[4] Perhaps the most interesting exception to this pattern concerns the Nobel Prize-winning work on the development of hybridomas that allowed understanding of the immune systems and also allowed the creation of monoclonal antibodies (Kohler and Milstein, 1975). As Kohler and Milstein submitted their ground breaking findings to *Nature*, they also submitted the manuscript to their funding agency (the Medical Research Council) with a proposal to file for a patent. However, the request was refused, on the basis that "It is certainly difficult for us to identify any immediate practical applications which could be pursued as a commercial venture, even assuming that publication had not already occurred." (http://www.path.cam.ac.uk/~mrc7/mab25yrs/index.html last accessed March 14, 2005)

1980 Bayh-Dole Act assigned IPR to Universities along with a duty to license the patents and facilitate their translation and commercialization (Mowery and Sampat 2001). This clarification in institutional practice likely increased the propensity for dual knowledge to be disclosed through both scientific publications and patents. Nevertheless, individual researchers and institutions varied widely in their response to this new environment[5].

Third, there was a significant expansion in the *scope* of IPR across the life sciences. After the 1980 *Diamond vs. Chakrabarty* decision and the granting of the Oncomouse patent in 1988, IPR comprehensively covered the domain of genetically modified living organisms – from bacteria to mammals (Kevles 2002). In combination with the developments in the biotech industry, "universities were literally propelled into an awareness of the potential economic value of the technology that was being generated in their research programs" (Bremer 2001).

Thus the ground work for the anti-commons debate was laid: increasingly dual knowledge and the widening scope of patents meant that many scientific discoveries now contained patentable inventions. This potential for dual disclosure together with more streamlined institutional rules for patenting and changes in university culture shifted the likely set of disclosure decisions to more frequently include patenting. In the period between 1989 and 1999 US Research One universities received over 6,000 life science patents (Owen-Smith and Powell 2002) and patent-paper pairs became an important disclosure phenomenon (Murray 2002). In short, the life science commons were increasingly littered with intellectual property rights.

Increasingly, economists, law and technology scholars, and policymakers have focused on key cases to highlight concerns over the impact of this expansion in IPR on the growth and diffusion of scientific knowledge (Heller and Eisenberg, 1998; David, 2001a; Campbell et al. 2002; Strauss et al. 2002; Walsh et al. 2002). Consider the Oncomouse (Murray 2005). The knowledge described by Leder's *Cell* publication was also covered by a "paired" Oncomouse patent (4,736,866). This patent was licensed exclusively to DuPont Corporation who, as the donor of a broad research gift to support Leder's lab, actually participated in the decision to

---

[5] In this context, the traditional justification of IP rights over Open Science relates not to scientific knowledge accumulation but rather to cumulative commercial innovation. IP provides incentives for further commercial investment. IPR may also facilitate a "market for ideas" by encouraging the exchange and trade of knowledge particularly with private sector researchers (Merges and Nelson, 1994; Arora et al, 2001; Gans and Stern, 2000; Gambardella, 1995). For knowledge generated in academia the third leg of the traditional argument for IP - as an inducement for disclosure over secrecy - does not apply since there is a presumption of and evidence for disclosure in publications.

patent. Soon after the patents were filed, DuPont imposed their property rights through a series of stringent licensing clauses: "reach-through" clauses that specified royalties for DuPont for any discoveries or products that might be developed using the Oncomouse, oversight and control of scientific publications, and limits on informal sharing of these mice among scientific colleagues. Each of these practices represented a dramatic shift in scientific norms.

These limits were seen as a severe impediment to scientific research and a contravention of the norms of the scientific community. Scientists were vocal in their opposition to DuPont; led by Harold Varmus, senior scientists in mouse genetics met at Cold Spring Harbor in 1992 and began talking of revolution. In response, the NIH stepped in and sanctioned a non-profit central facility (The Jackson Laboratory) to serve as a repository for genetically altered mouse strains. However, while the Jackson Labs could distribute the mice, the Oncomouse was still constrained by restrictive licensing terms. For example, DuPont continued to contact scientists and demand review of their articles before publication and "nobody was able to exchange materials" (Marshall, 2000). When Varmus became the Director of the NIH, he furthered his campaign to limit DuPont's actions on the basis that they "seriously impede further basic research" and in 1999 an agreement was reached between NIH and DuPont which outlined the terms under which the technology could be used in research supported by the NIH. DuPont commented that they "deeply appreciate the importance of wide dissemination of tools for basic research and are committed to making this tool available to the academic community…while retaining our commercial rights" (NIH News Release January 19, 2000). While lifting key restrictions in the license, all non-profit recipients were still required to have an overarching agreement in place with DuPont, while commercial entities are required to purchase a commercial license.

The anti-commons debate was further sharpened by controversies over patenting of gene fragments known as expressed sequence tags (ESTs) - small fragments of DNA generated to identify regions of a gene using well known research methods. Like the Oncomouse, researchers wanted to file patent applications on gene sequences that would also be placed into the public domain. This raised considerable debate among the scientific and legal communities; some argued that EST patents would distort research priorities, encourage researchers to search for and patent ESTs rather than focus on characterizing full-length genes, and preclude further research on entire genes (Kimball, 2001). Furthermore, it was anticipated that these patents would lead to

complex and inefficient litigation and ownership disputes over full genes (Eisenberg, 1996). Others suggested that these patents would, like many other patents on novel ideas, spur innovation. Although no longer considered patentable by the USPTO on the basis of a lack of utility, the EST episode highlights the potential for complex transactions in the use of genes for biomedical research and development (Merges and Nelson 1990; Lawson, 2002) and the complexity of the debate over IPR over research conducted in "Open Science" that goes beyond the anti-commons issues.

Overall the anti-commons hypothesis suggests that intellectual property rights over scientific knowledge traditionally disclosed only through publication may have a negative impact on the capability for cumulative progress within the *scientific* community. Varmus and others in the transgenic mouse community felt that the means through which DuPont imposed their property rights impeded scientific enquiry and impinged upon their ability to "stand on the shoulders of giants" such as Leder and others. By its very nature, scientific knowledge is non-rivalrous, so that the diffusion of that knowledge can serve repeatedly (and with little additional cost) as an input into future knowledge production and hence cumulative innovation. Because IP can serve to exclude follow-on researchers from exploiting scientific discoveries, the anti-commons hypothesis posits that the use of IP over areas traditionally maintained in the public commons can undermine the process of cumulative scientific discovery (Heller, 1998; Heller and Eisenberg, 1998; David 2003). In other words, the rise of patent-paper pairs with their effective privatization of the scientific commons imposes a "tax" on the use of the knowledge and may restrict the diffusion of that knowledge and so lower the equilibrium level of research productivity.[6] Finally it is useful to note that the policy concerns surrounding the anti-commons effect is most salient for knowledge produced in public settings such as universities and/or funded from public expenditures.

III. **The Empirical Framework**

Anti-commons theory suggests that while the scientific community as a whole may benefit from the free dissemination and diffusion of knowledge, individual researchers have strong incentives to take advantage of the protections afforded by formal IP rights. If protected

---

[6] Furthermore, to the extent that IPR is narrow in scope and highly dispersed across individuals and institutions, fragmentation can impose a further tax in the form of significant transaction costs (Eisenberg, 1996; Shapiro, 2001; Hall and Ziedonis, 2001; Ziedonis, 2004).

by IPR, the impact of an individual piece of knowledge on follow-on research by others is diminished, resulting in a lower equilibrium level of on-going cumulative research productivity. These theoretical ideas have observable empirical implications: Published ideas that are also covered by IPR (such as the Oncomouse) will be expected to have lower levels of cumulative scientific *impact*, relative to the impact that would be realized in the absence of IPR. Moreover, we expect the salience of this effect to be higher for certain types of researchers (e.g., for patent-pairs associated with public sector research) and certain types of discoveries (e.g., for patent-paper pairs over research tools). This section first develops an empirical approach for evaluating the anti-commons effect, and then links this framework to a number of specific theoretical predictions of the anti-commons theory.

The empirical challenge in documenting the anti-commons effect is straightforward: How can we compare the impact a given piece of knowledge would have *with* IPR versus the impact of that knowledge without IPR? For a given piece of knowledge *with* IPR, one cannot observe the counterfactual impact of that knowledge in the event that the IPR had been waived. For example, knowledge protected by IPR may tend to have a higher (or lower) intrinsic scientific value than knowledge that is not protected by IPR. A simple comparison between patented and non-patented knowledge may therefore be biased by unobserved heterogeneity. The key empirical challenge in adjudicating the impact of disclosure regimes is thus an identification problem – from an experimental perspective, the econometrician would ideally like to observe a given piece of knowledge in two distinct institutional environments (e.g., a non-patent versus patented environment), and compare the impact on future research for each case.

Our approach exploits the existence and nature of patent-paper pairs. A patent-paper pair is the dual instantiation of a given piece of knowledge as both a scientific research article and a patent. Consider the following example of research undertaken in the biology department at the Massachusetts Institute of Technology in the field of bacterial genetics and metabolism:

> "A method has been developed for control of molecular weight and molecular weight dispersity during production of polyhydroxyalkanoates in genetically engineered organisms by control of the level and time of expression of one or more PHA synthases in the organisms. 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 effect the molecular weight of the polymer produced in the cell." (Snell; Kristi D. (Belmont, MA); Hogan;

Scott A. (Troy, MI); Sim; Sang Jun (Seoul, KR); Sinskey; Anthony J. (Boston, MA); Rha; Chokyun (Boston, MA) 1998, Patent No. 5,811,272)

"A synthetic operon for polyhydroxyalkanoate (PHA) biosynthesis designed to yield high levels of PHA synthase activity in vivo was constructed …by positioning a genetic fragment … behind a modified synthase gene containing an Escherichia coli promoter and ribosome binding site. 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)

As outlined in these brief excerpts, the research described in both documents is based on a specific genetic modification of a bacterium (E. Coli) designed to control the type and amount of particular chemicals (PHA) the bacteria would ordinarily produce. From the scientific perspective, the publication emphasizes that these experiments deepen our understanding of the genes that regulate particular chemical pathways in bacteria. However, as highlighted in the patent, they also provide practical techniques for the manipulation of bacteria and the optimization of their use as a source of useful biomaterials. In other words, this single discovery has been instantiated as both a publication emphasizing its scientific contribution and as a patent disclosure emphasizing its utility.

Our empirical approach exploits three key aspects of the phenomenon associated with the production of patent-paper pairs:

(a) Within the scope of published scientific research which is potentially patentable, a significant fraction of researchers choose to forego formal IPR, and

(b) Among those who pursue formal IPR, there is usually a significant delay between the first publication of that knowledge within the scientific literature and the granting of formal IPR over that knowledge and

(c) Future scientific citations are a noisy but useful measure of the impact of a scientific article on future scientific research. Combining these factors, we exploit the existence of patent-paper pairs to evaluate the impact of IPR on the diffusion of scientific knowledge.

The first step in our approach is to collect a sample of published scientific research articles which are of roughly similar "quality" (though we will account for quality variation among articles in our empirical framework) and which disclose knowledge that is potentially patentable (whether or not the researchers choose to apply for IPR). We draw a set of research articles, within a specific time frame, from a top-tier research journal that specializes in publishing dual knowledge discoveries, *Nature Biotechnology* (we discuss this choice further in Section IV). By limiting our sample to articles published within a narrow time window within a specialized journal, we ensure that all articles share some affinity in topic, and are subjected to the same peer-review process. By choosing a journal whose editorial policy is to focus on research that is simultaneously of scientific and commercial interest, we consider the publications in our dataset to be "at risk" of being associated with a USPTO patent and thus forming a patent-paper pair. As we discuss further in the next section, we are able provide qualitative support for our assumption of comparability between the articles within our sample with analysis of the article's scientific content and review of the potential patentability of the knowledge disclosed in the non-patented articles by an experienced intellectual property lawyer. Of course, the decision to patent a piece of dual knowledge is endogenous to the specific circumstances of individual researchers, including factors such as their institutional affiliation and their gender (Azoulay, Ding and Stuart, 2005; Ding, Murray, and Stuart, 2005; Markiewicz and Diminin 2004; Agrawal and Henderson 2002). Our procedure simply assumes that the impact of intellectual property on the use of knowledge by *follow-on* researchers is (conditionally) independent of the patent filing decision.

Within our sample of articles (only a subset of which are associated with a patent-paper pair), we then exploit the fact that for any patent-paper pair, there exists a patent grant delay – a substantial gap between the date of scientific publication and the date at which the associated patent is granted.[7] And, in many cases, follow-on researchers are unaware of whether or not a particular discovery will ultimately be associated with a patent-paper pair. This empirical technique exploits the insight that while publication in the scientific literature often occurs within

---

[7] The specifics of patent law regarding the precise timing of disclosure is complex, varies across countries and has been subject to change. However, broadly speaking, under US patent law, inventors have a grace period of twelve months between public disclosure (for example in an academic publication or presentation) and filing for patenting covering that knowledge. Thus, the timing of the publication submission and patent application can vary among patent applications with some filed before publication and some after. In the Oncomouse case, a patent was filed on June 22, 1984, the article was submitted on August 21, 1984, the article was published in October, 1984 and patent grant and publication occurred on April 12, 1988.

six months (or less) after initial submission to a journal, the delay between the initial application and receipt of a patent is often many years (in most cases a 2-4 year time window). It is important to emphasize that patent grant delay is more than simply a matter of the timing of a *pro forma* administrative decision. During the time between application and grant, applicants and examiners undertake detailed negotiations about the scope and extent of the patent grant, and so there is significant uncertainty about the extent of IPR prior to grant (Cockburn, Kortum, and Stern, 2003; Jaffe and Lerner, 2004). Perhaps more saliently, prior to the patent grant date, the patent applicant holds no formal IPR, and, in nearly all cases, cannot sue for infringement for activities undertaken during the pre-patent grant period. Finally, until 2001 (and thus for nearly all of the cases within our empirical work), USPTO patent applications remained *secret* until granted. In other words, for any given patent-paper pair, we observe the same "piece" of knowledge in two distinct institutional regimes: one associated with the pre-patent grant period and then a regime shift into the post-patent grant period.

Of course, the behavioral response to the patent grant date will depend importantly on the degree to which the patent grant serves as "news": To what extent were researchers aware of the impending grant, and how does patent grant (and associated post-grant enforcement activities) change behavior by raising the perceived "price" of building on a prior discovery? If follow-on researchers believe that a patent-paper pair is likely (as might occur, for example, for articles published by researchers from a for-profit company), the impact on patent grant on behavior is likely to be modest (since researchers will anticipate the potential for IPR in advance and incorporate these potential costs into their research decisions). On the other hand, if only a minority of university researchers engage in patenting behavior, the potential for a post-patent grant "surprise" is quite high. For the purpose of the current analysis, we assume that at least some potential follow-on researchers may experience the patent grant as "news," though of course the level of news may depend on factors such as the institutional affiliations or locations of the originating authors, and the amount of time elapsed since the patent grant date.[8]

We combine these elements into an empirical model by taking the annual rate of citation to an article in follow-on scientific research articles as a measure of the cumulative impact of that

---

[8] It is also theoretically possible that follow-on researchers will exploit the "window" between publication date and the patent grant date to take advantage of the absence of IPR over knowledge which will ultimately be protected. While we acknowledge this theoretical possibility, our fieldwork strongly mitigates against this strategy. Few laboratories are able to predict the precise timing of their research results, and so are unlikely to strategically enter and then exit a research area that will come under patent protection at an uncertain date in the future.

article through time. Though scientific citations are by no means a perfect measure of the impact of a specific article, they provide a useful (if noisy) index of the relative salience of research in follow-on research which is also disclosed in scientific publications (Merton 2000; Cole 2000). More precisely, if the granting of IPR raises the cost of building on a specific piece of knowledge, then the citation rate to IPR-linked scientific publication should decline post patent grant. Of course, measuring the impact of scientific research using citations implies that we must account for its form as count data skewed to the right (and likely over-dispersed relative to Poisson). Therefore, except where noted, we employ a negative binomial model of the citations produced per year for each scientific article in our dataset. However, the impact of a given piece of research, as measured by citations, will vary considerably with the underlying importance of the research discovery, with the time elapsed since initial publication, and with the year for which the citations are being considered. To control for these factors, our empirical specifications will account for individual publication quality (through article fixed effects) and for the effects of publication age and the overall rate of citation in a given year (through age and citation year fixed effects).[9] Taken together, our baseline empirical test for the anti-commons hypothesis is therefore:

$$CITES_{i,t} = f(\varepsilon_{i,t}; \gamma_i + \delta_{t-pubyear} + \beta_t + \psi POST-GRANT_{i,t}) \qquad (1)$$

where γ_i is a fixed effect for each article, $\delta_{t-pubyear}$ captures the age of the article, β_t is a fixed effect for each citation year, and POST-GRANT is a dummy variable equal to one only for years in which a patent is in force for an individual article.[10] Because we observe citations to a scientific publication both before and after the patent is received (and because we observe a control group of similar publications which never receive a patent) we are able to identify how the temporal pattern of citations to a scientific publication changes as the result of patent grant.[11]

---

[9] Several subtle issues, including an incidental parameters problem, arise in incorporating multiple fixed effect vectors into a negative binomial specification. We experimented with a range of alternative approaches, including the conditional negative binomial estimator (Hausman, Griliches, and Hall, 1984) and the fixed effects estimator (Allison and Waterman, 2002). All of our qualitative findings are unchanged across these different procedures; building on the simulation results reported by Allison and Waterman (2002), we report the results associated with the fixed effects procedure, which also facilitates straightforward computation of robust standard errors.

[10] Of course, it is possible to constrain this specification in order to, for example, obtain a separate estimate of the impact of unobserved heterogeneity between patented versus unpatented articles versus the impact of patent grant itself. In this case, one could simply replace the article fixed effect with a dummy variable for those articles which ultimately receive a patent: $CITES_{i,t} = f(\varepsilon_{i,t}; \delta_{t-pubyear} + \beta_t + \lambda PATENTED_i + \psi POST \quad GRANT_{i,t})$

[11] This baseline analysis does assume that the age fixed effects associated with citation depend on whether a paper receives a patent. In particular, a key assumption of our base model (which we later relax) is that patented articles

This test goes beyond the potentially biased test of whether patented publications are more or less highly cited that those that are not associated with patents. In other words, we test for the anti-commons effect by calculating how the citation rate for a scientific publication *changes* after patent rights are granted, 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.

At its core, anti-commons theory predicts that if the grant of intellectual property hinders the ability of researchers to cumulatively build on a given piece of knowledge, then the citation rate to the scientific publication disclosing that knowledge should fall after formal IP rights are granted over that knowledge. In addition to this baseline prediction, theory and the details of our empirical setting suggest a number of more nuanced hypotheses. First, as we discussed earlier, the ability to infer the impact of IPR from the date of patent grant (and not before) depends importantly on the fact that the patent grant represents "news" to at least some follow-on researchers. As such, we expect our proposed test to be more salient among researchers with public sector author affiliations (e.g., government or university researchers). When a purely private research team publishes a scientific result, potential follow-on researchers may (rationally) anticipate that the knowledge and/or tools described in the article will eventually be covered by IPR. Accordingly, in contrast to the "news" associated with public sector patent grant, researchers may be deterred at the outset from research areas in which private sector researchers are publishing, whether or not the patent grant has yet been issued. Furthermore, we would expect that the "news" from a particular patent grant likely diffuses among researchers over time, particularly in the few years after the grant. Taken together, we expect that the impact of patent grant changes with the time elapsed since patent grant:

$$CITES_{i,t} = f(\varepsilon_{i,t}; \gamma_i + \delta_{t-pubyear} + \beta_t$$
$$+ \sum_{k=1....r} \psi_{PRE\_k} PRE - GRANT(k)_{i,t} + \sum_{l=1....s} \psi_{POST\_l} POST - GRANT(l)_{i,t})$$

This specification allows us to evaluate both whether the impact of patent grant was anticipated (in terms of changes of citation behavior that preceded the actual patent grant date) as well as how the impact of patent grant changes with the years elapsed since the patent grant.

are not simply "shooting stars" – articles that, for exogenous reasons, experience a high rate of early citation followed by a rapid decline. In part, this characterization is counterfactual to the most well-documented pattern of scientific citation, the so-called Matthew Effect, in which articles with a high rate of early citation tend to continue to receive an ever-higher rate of citation after a favorable early record (Merton, 1973). Also, in our robustness analysis, we actually rely exclusively on a sample of *patented* articles (with varying patent grant lag times), with no change in our core results.

Anti-commons theory also offers several additional predictions that provide precision in terms of the types of knowledge most likely to be associated with the anti-commons effect. First, the effect should be greatest for patent-paper pairs with particular patent characteristics – namely patents associated with significant prior art, broad scope, and specifically focused on research tools. The anti-commons hypothesis is premised on the difficulty of using contracting methods to access to knowledge covered by patent protection. While contractual difficulties may be relatively unimportant when there is only a single patent or area of prior art, transaction costs (and bargaining break-down) are likely to be more important when there are multiple competing claims limiting access to knowledge and materials. Therefore, patents with significant patent prior art and non-patent prior art might be expected to be associated with denser patent thickets and hence stronger anti-commons effects. In addition, broad patents represent the broadest definition of the patent-paper pair and the strongest rights over the published knowledge. As such they are likely to have the greatest anti-commons impact. Likewise, patents on research tools rather than compositions of matter, pathways etc. – tools like the Oncomouse, cell lines, gene probes – would also be predicted to have a larger anti commons effect. They represent ideas that are of broad relevance to many researchers and are also subject to greater transparency in their use; cell deposits, material transfer agreements, and other institutional arrangements facilitate their dissemination but potentially also alert patent holders to possible infringement.

Finally, the impact of patent grant may be strongest for publications by US-based authors. For articles written exclusively by non-US authors, the process of applying for and having a US patent granted is likely a very heterogeneous process, and enforcement activities may be uneven. As a result, we expect to observe a more salient effect associated with grant behavior for articles with a US domestic author. Overall, our empirical framework provides a rich setting for evaluating a number of detailed hypotheses implied by the anti-commons argument. The remainder of this paper focuses on evaluating these empirical questions.

IV. The Data

IV.A. Sample Definition

The data for this study is based on the entire population of peer-reviewed research articles published in the journal *Nature Biotechnology* over the period 1997 to 1999. While the journal publishes scholarly material in a variety of formats, we confine our data to research articles

which are defined by the editorial policies of the journal as "a substantial novel research study" (see *Nature Biotechnology*, A Guide to Authors). Under these criteria, the dataset consists of 340 unique research articles.

Our sample population was chosen to focus on research exhibiting the duality emphasized by Stokes (1997). As noted above, biotechnology is a particularly salient arena for dual knowledge. In *Nature Biotechnology's* first issue in 1996, the editorial mission of the journal was described as: *"*cover[ing] business, financial, and regulatory matter: not to do so would be perverse and self-defeating. But its emphasis will be unashamedly on research and development, the fuel for biotechnology's fire." (Nature Biotechnology, 1996).[12] The publishing policy adopted by *Nature Biotechnology* explicitly aimed at research with potential applications to biotechnology: "[the journal] aims 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. We also have a strong interest in research that describes the application of existing technologies to new problems or challenges, and basic research that reports novel findings that are directly relevant and/or of interest to those who develop biology into technology." Since its inception, *Nature Biotechnology* has established itself as the leading outlet (in terms of measured scientific impact) for refereed scientific research relating to biotechnology and continues to play this role. In other words, research published in *Nature Biotechnology* is both "at risk" of serving as a foundation for future scientific studies and for commercial exploitation.

For each of the 340 articles article, we determined whether a patent associated (or paired) with the article had been granted by the USPTO. Using the USPTO search engine, we defined a series of searches for each article. The basic search included i) the first, last and corresponding authors for the article and ii) the list of institutions found in the article "address field" in the Web of Science database. For some institutions, specific name variations were used to account for the fact that some institutions patent under distinctive institutional names: For example, patents assigned to the University of Oxford are listed under the name of its separately incorporated technology licensing office, ISIS Innovation. Different combinations of authors and/or

---

[12] The new journal was not entirely new. It "picked up the torch from Bio/Technology", a journal founded in 1983 to explore and publish leading edge science in biotechnology.

institutions were used from the most to the least inclusive in order to identify all issued patents associated with the authors and institutional affiliations whose research appeared in *Nature Biotechnology*. For example, since some patents were assigned to individuals (rather than an institution), the search procedure examined whether each author for each article received a patent within the time frame in question. After establishing the set of patent grants received by individuals and institutions represented in the articles, patent abstracts and claims were read to establish the presence of a patent-paper "pair." To do so, we verified whether the material described in the abstract of the article was incorporated into the description, claims and/or examples of the granted patent.[13] By checking the precise content of patents granted to those whose research is published in *Nature Biotechnology*, our procedure provides a consistent way to identify the subset of articles within our overall sample which are also patent-paper pairs. Using this procedure, 169 of the 340 articles were found to be associated with a patent as of October, 2003. In other words, approximately half of all publications in Nature Biotechnology are associated with a patent-paper pair within five years of publication.

Since, as noted above, our empirical work relies on the fact that a) the entire sample of articles is initially "at risk" of being patented, and b) but for being patented the articles do not vary systematically along an unobserved dimension impacting the citation time trend we checked the comparability assumption in several ways. First, the sample design ensures that the articles are comparable insofar as all are drawn from the same (reasonably specialized) high-quality journal. All articles have undergone a similar refereeing process, and editorial decisions are presumably made with the journal's editorial mission in mind. Second, for a subset of 34 of the non-patented articles, we undertook a detailed evaluation of their innate patentability. The standard for patentability (defined in 35 U.S.C. Section 101) is defined as "new and useful processes, machines, manufactures, compositions of matter; and any new or useful improvements thereof….subject to the conditions of patentability" where conditions of patentability include novelty, non-obviousness and utility (see, for example, Merges, 2003, for an overview of the law and economics of the patent system). It is important to note that this standard excludes important categories of knowledge which might be reflected in scientific

---

[13] One of the authors (Murray) holds a PhD in Applied Sciences and has conducted detailed qualitative research on the scientific content of contemporary biotechnology and applied microbiological research (Murray, 2002). The criteria used to assign a patent-paper pair was conservative insofar as there had to be a direct connection between the disclosures in the article abstract and patent record. In the vast majority of cases, the presence (or not) of a patent-paper pair was unambiguous.

research articles, such as the discovery of new scientific principles, abstract ideas, and the identification of naturally occurring materials. However, novel "research tools" and "compositions of matter" are patentable, and constitute the bulk of patented technologies in the sample. An experienced patent attorney (graciously) undertook an examination of the publication abstracts and was asked to make a "conservative" determination of whether the research findings included a potentially patentable discovery. Of the articles submitted for review, more than 75% (27 out of 34) were considered to be obviously patentable; of the remaining, most contained at least some potential for patentability (i.e., while they failed the conservative test we requested, they likely would have passed a more lenient (but still plausible) standard for patentability). In particular, most of the articles not considered patentable under our test reported research results using standard techniques on pre-existing materials, and so the abstract did not include a description of a novel research tool or composition of matter. While these evaluations do not constitute a formal legal opinion, this check does provide support for the assumption that most (if not entirely all) articles within the sample are at risk of being patented.

Third, we directly compared the similarity of articles within the sample. The MedLine database includes a feature allowing the identification of "similar" articles, based on keyword matching. For each patented article, we identified the "most similar" non-patented article within the sample and qualitatively evaluated a subset of these for comparability. By and large, matched articles were found to be qualitatively similar, both in terms of their underlying scientific content and their potential for patentability.[14]

Finally we gathered a series of variables associated with the number of authors, the number and type of institutional affiliations, and the author age, rank and gender. On the one hand, there are some (marginally) significant differences between the unpatented and patented articles. For example, articles with private sector affiliations tend to have a higher patenting rate

---

[14] Consider the following matched example. One article, published in June, 1997, describes a research study (by researchers at John Hopkins University) of a novel method of using bacteriophage which express ligands on their surface to detect the interaction between key proteins, thus allowing "a powerful approach to the molecular studies of protein-protein interactions" (Li, 1997). The second article, published in December, 1999, by researchers at Sugen (a biotechnology firm) describes a novel display technique to examine the interaction among specific proteins using a library of DNA fragments that contain specific mutations used to reduce non-specific binding that will render results imprecise and difficult to analyze. The method was applied to key signal transduction pathways and the authors suggest could be "a rapid and efficient tool for elucidating protein networks" (Zozulya et al. 1999). Though both articles are concerned with extremely similar scientific issues (methods for identifying protein interactions) and both are clearly describing (patentable) research tools, only the first is associated with a USPTO patent grant.

than those with exclusively public sector affiliations; articles with US-based authors tend to have a higher rate of USPTO patenting than those articles with exclusively foreign authors. Indeed, at one level, this institutional variation is precisely what we are exploiting in our core empirical specification. More importantly, there were no significant correlations between patenting and any aspect of the *nature* of the scientific research contribution (e.g., article length, the number of authors, or the number of backward references). Overall, given the nature of the sampling process, and the qualitative comparisons, we assume that articles within our sample are each "at risk" of being patented (in a technical sense), and both observable and unobservable institutional factors are important in determining this outcome. In most of the empirical work, we therefore assume that, after controlling for the overall quality an article, the sample population is composed of articles at risk of being patented and, but for a patent grant, the time trend for all articles follows the same stochastic process.

The dataset is drawn from three distinct sources, each noted in Table 1. Article-specific characteristics are gathered from *Nature Biotechnology*; the date of publication, the number of authors, and the location and institutional affiliation of each author (available from the address list provided for each article). For each article we identified the lead author (the so-called Principal Investigator) for whom we gathered individual and institutional information from institutional websites and author resumes. The citation counts for each article (through October, 2003) are calculated using the Science Citation Index Expanded (SCI).[15] For each article associated with a USPTO patent, we collected a number of patent characteristics from the USPTO public database.

IV.B.   Summary Statistics

For the variables used in our analysis, Table 1 provides variable names and definitions and Tables 2 and 3 reports summary statistics. For each article in the dataset, we track citations beginning in the year in which the article was published and continuing until the end of 2002. The total number of article-year observations is 1688.

The key dependent variable in our analysis is FORWARD CITATIONS, the number of articles that reference the focal article in a given year. Not surprisingly given the prestige and

---

[15] Maintained by the Institute for Scientific Information (ISI), SCI records reference and citation information for nearly six thousand scientific and technical journals in approximately 150 disciplines.

quality of *Nature Biotechnology*, the average level of annual citations received by articles in this dataset is quite high, relative to randomly selected academic articles (mean = 9.35), and, by the end of 2002, the average article had received more than 54 total cumulative citations. Consistent with prior citation analysis studies, the distribution of citation counts is quite skewed, with nearly 20% of the citation-years receiving either 0 or 1 citation, but also including one publication with an annual citation count equal to 181 (Figure B). Because we observe article-years from 1997 through 2002 (but only observe articles published in 1998 or 1999 for a shorter set of years), the average CITATION YEAR is at the margin of the 2000 calendar year, and the average AGE observed within the sample is just a little over 2.0 years.

While the heart of the analysis incorporates article fixed effects to account for differences between articles in terms of their impact and overall quality, we examined a number of article-specific characteristics. These characteristics include the number of authors (# AUTHORS, mean = 5.89) as well as their institutional affiliations. For example, US AUTHOR is a dummy variable measuring whether *at least* one of the authors lists a US address (mean = 0.59). We assign university and government researcher affiliations as "public sector" institutions and pharmaceutical and biotechnology affiliations as "private sector" organizations. We then define two dummy variables, PUBLIC SECTOR AUTHOR and PRIVATE SECTOR AUTHOR, which are equal to one if *at least* one author is associated with a public sector or private sector organization, respectively. Interestingly, nearly 90% of the articles in the sample have at least one public sector author, and more than 30% have at least one private sector author.

Finally, our data includes information about the 169 patents associated with each patent-paper pair. The average date of patent grant (weighted by article) is mid-2000 with the average lag between the patent application and patent grant date is just over 3 years).[16] Interestingly, while the patent grant lag is correlated with some traditional factors (such as the number of backward references in the patent), the overall patent grant delay is not robustly correlated with institutional affiliations (PUBLIC SECTOR AUTHOR, PRIVATE SECTOR AUTHOR, or US AUTHOR). PATENTED, POST-GRANT is the dummy variable indicating whether the citation-year is linked to a patent-paper pair, and whether the patent has been issued and is in

---

[16] Some patents have been issued as late as 2003 (mostly associated with 1999 publications). Inclusion or exclusion of these 23 articles from the analysis does not change any of the qualitative findings (primarily because our strategy relies on articles where we observe a change in the IP regime during those years where we observe a citation count).

force.[17] We also measure several patent-specific characteristics which along with article-specific characteristics can be interacted with PATENTED, POST-GRANT to assess whether the impact of patent grant on the citation rate depends on observable characteristics of the article or associated patent.

First, similar to PUBLIC SECTOR AUTHOR, PUBLIC SECTOR ASSIGNEE (mean = 0.65) is a dummy variable equal to one if there is *at least* one assignee from the public sector. In addition, when the research upon which a patent is based is funded (even in part) by the Federal government, the applicant must disclose a Federal "interest" (indeed, the ability to retain patent rights in most cases despite this Federal interest is at the heart of the Bayh-Dole Act). Of the 169 articles associated with patents in our dataset, 49 report a Federal interest. To the extent that contemporary policy proposals related to the anti-commons effect would most likely impact Federally funded research, the impact of patent grant on citation is particularly interesting for those patents for which GOVT FUNDED = 1. We then measure the number of listed inventors (# INVENTORS, mean = 3.04); interestingly, the average for # INVENTORS is just over half that associated with # AUTHORS. We also calculate two measures of the level of prior art cited by the patent, including the number of citations to prior patents (PATENT BACK CITATIONS) and the number of citations to prior non-patent references (PATENT BACK REFERENCES). As the number of prior art references increases, the potential for a "patent thicket" increases (Shapiro, 2001); in the spirit of Heller and Eisenberg (1998), the presence of a patent thicket may exacerbate the anti-commons effect and result in a greater decline in the post-grant citation rate of patented articles. Interestingly, relative to the overall means for citations made to patented prior art by "biotechnology" patents reported by Allison and Lemley (2002), the averages for both PATENT BACK CITATIONS (mean = 7.26) and PATENT BACK REFERENCES (mean = 28.19) are somewhat high. We also include two measures of the type and scope of the patent. # CLAIMS is simply the number of allowed claims which is an extremely imperfect (and noisy) measure of the strength of patent rights, but does provide at least a weak measure of patent scope thus an increase in # CLAIMS should intensify the impact of patent grant on article citation rates. TOOLS PATENT is a dummy variable equal to one if the primary class for the patent is within the 435 and 800 patent classes (mean = 0.58). Out of the 11 3-digit patent classes

---

[17] PATENT, POST-GRANT is equal to one from the year of the patent grant date. All of the qualitative results are robust either to including the years associated with the year of the patent grant in the "pre-grant" category or dropping these article-year observations from the analysis.

represented across the patents within the sample, these two 3-digit classes are most closely associated with processes and tools. Since research tools have been of particular concern within the anti-commons debate (relative to composition of matter patents), the TOOLS PATENT dummy will allow us to assess whether the impact of patent grant on citation is greater for patents covering research tools and methods.

IV.C.   *Patented Versus Non-Patented Articles*

Table 3 compares the means of patented and unpatented articles within the sample. A few notable differences stand out. First, the average rate of citation is relatively similar across the two groups, with the patented articles receiving, on average, an additional citation per article-year over the sample. However, this 10% average difference masks more substantial differences that manifest themselves over time. In Figure C, the average FORWARD CITATIONS are plotted by AGE (years since publication). During the year of publication and in the subsequent three years, PATENTED articles have a significant citation advantage, equivalent to an approximately 15-18% "boost" over the citations rates for non-patented articles. However, in the fourth and fifth years after disclosure in the literature, patented articles converge to the citation rate associated with non-patented articles. As we explore further in the next section, it is during these later years in which patented articles are in the post-grant phase, perhaps helping to explain the decline relative to trend experienced by patent-paper pair publications. In addition to these overall differences in citation rates, it is important to recognize that there are also differences in article characteristics. Relative to non-patented articles, patented articles have a significantly higher chance of having at least one US author, or at least one author from a private sector organization (the differences in means in # AUTHORS and PUBLIC SECTOR AUTHOR are not significant).[18]

These data suggest that the number of citations and article characteristics vary across the margin of whether or not an article is part of a patent-paper pair. While the drop-off in the citation advantage associated with patented articles is consistent with the presence of an anti-commons effect, such an effect could result from differences in the characteristics of articles

---

[18] Appendix A explores these conditional means in more detail by breaking them out according to whether a public sector author is associated with article. Notably, among patented articles, there is a distinct citations advantage associated with those articles with PUBLIC SECTOR AUTHOR equal to 0.

represented in the different age-cohort categories. In order to precisely identify the source of these differences, the next section undertakes a systematic econometric analysis of these patterns.

V.   Empirical Analysis

The empirical analysis proceeds in several stages. We first compare the cross-sectional differences in citation rates between patented and unpatented articles allowing for controls for other article characteristics. We then turn to the principal empirical exercise, examining how the citation rate changes with the grant of formal intellectual property under a variety of control structures and examining how the effect of IP manifests itself over time. Finally, we examine how our diffs-in-diffs estimates vary with article and patent characteristics. As noted above, except where noted, all specifications employ a negative binominal regression. Coefficients are reported as *incident rate ratios*, and so should be interpreted as a multiplicative effect on the expected number of citations received in a given year resulting from a one unit change in a regressor (i.e., the null hypothesis of no effect yields a coefficient of 1.0).

We begin in Table 4 with two negative binomial specifications, each of which includes YEAR and AGE fixed effects. While PATENTED articles are associated with a higher rate of citation without any additional controls, this effect is reduced in magnitude and statistically insignificant when controls for # AUTHORS, US AUTHOR, and PUBLIC SECTOR AUTHOR are included. In other words, the overall citation advantage observed in the conditional means can be (primarily) explained by differences in observable article characteristics.[19] This brief cross-sectional analysis motivates our main empirical analysis of the impact of *patent grant conditional on article quality*. We begin in (5-1) with a regression, in which the dependent variable is equal to the natural log of FORWARD CITATIONS + 1; this OLS regression includes fixed effects for each calendar year, article age, and article fixed effects. Though this specification does not account for the nature of citations as skewed count data, the results provide support for nearly a 10% decline in the rate of citation after patent grant (significant at the 10% level). The remainder of this table turns back to negative binomial regression. In (5-2) and (5-3), we estimate both the overall difference between patented and non-patented articles and the marginal impact of being in the post-grant phase. For each of these specifications, we

---

[19] Though it is not the principal focus of the analysis, several unreported specifications confirm this qualitative finding. Though PATENTED is significant when controls are not included, the estimated effect becomes both smaller and imprecise when even a single observable article characteristic is included.

include the same article controls included in (4-2); as well, whereas (5-2) employs a parametric treatment for the impact of AGE (including a linear and quadratic term), (5-3) includes a complete set of age and year fixed effects. In both cases, patented articles enjoy approximately an 18% overall citation boost; however, the receipt of a patent – establishment of a granted patent-paper pair - effectively erases this advantage, resulting in an estimated 15-17% decline in the expected citation rate. Finally, in our differences-in-differences estimate in (5-4), we include a separate fixed effect for every article (and a complete set of year and age fixed effects); as such, these estimates are identified exclusively off the within-article contrasts between pre-grant and post-grant citation levels (and after accounting for the impact of age and year). According to this specification, the estimated post-grant decline is estimated at just over 9% with a high level of statistical significance. Moreover, these results are robust to alternative specifications and sample definitions.[20] At face value, these estimates provide concrete evidence for the existence of an anti-commons effect. Simply put, the impact of a given piece of scientific research on scientific research declines after IP rights are granted. At an intuitive level, the results suggests that between 1 in 11 and 1 in 6 researchers building on a given paper may forego a specific research project requiring citation to the patent-paper pair after IPR is granted.

The evidence for an anti-commons effect is even stronger when we examine the relative citation rate for an article in the years preceding and following the patent grant. To do so, we estimate a fixed effects negative binomial regression including specific dummy variables for each year preceding and following the patent grant date. In Figure D, we display the coefficients from that specification. Overall, patented articles tend to have an overall higher citation rate (relative to the baseline of 0 which is the value of the "average" fixed effect in the sample). Earlier, we had attributed this overall difference to differences in observable characteristics such as author affiliations. More importantly, Figure D provides evidence consistent with the idea that the anti-commons effect grows in the time elapsed since patent grant (though individual coefficients are estimated imprecisely). Though articles receive an up-tick in citations in the year prior to patent grant, there is an intermediate decline in the year of patent grant which continues steadily through four years after the date of patent grant. The difference between the pre-grant average and the average four years after patent grant is more than 25%, suggesting that

---

[20] For example, the results are similar (though a bit more noisy) if broken out by individual years of publication, or if (the incomplete record for) 2003 citations are included as an additional year of data.

the impact of IPR accumulates over time; taking these pattern at face value, the "size" of the anticommons effect becomes a sizable deterrent after IPR are in force for several years.

The Role of Institutional Affiliations, Location and Funding

As noted above, the focus of the anti-commons debate is on scientific research conducted with US government funds that involves university or other public sector researchers. In Table 6, we move beyond the baseline analysis to examine how the effect of patent grant on citations varies with the affiliations of article authors. Each specification consists of a negative binomial with fixed effects for articles, article age, and calendar year (as in (5-4)). In (6-1) and (6-2), we examine how the impact of patent grant varies with whether an article has any public authors and whether it has any private authors. The results are striking. While there is an insignificant effect when there is at least one private author (or all private authors), the decline in citations is most closely associated with the case where there are *no* private authors. It is possible that articles associated with private sector researchers are *assumed* to have intellectual property associated with them, and so the actual receipt of a patent has little impact on behavior by those who might build on that research. In contrast, until the patent grant date, researchers put a lower probability on purely public sector authored papers having a patent associated with them; the effect of IP in this case is to chill follow-on research endeavors. Interestingly, there is much less difference in the citation decline depending on whether the article is associated with at least one US author (6-3); moreover, (6-4) demonstrates that the anticommons effect is identified even if one *excludes* all articles without at least one US author. In other words, though the impact of patent grant is muted for those articles without at least some public sector involvement (perhaps due to expectations about IPR by follow-on researchers), these results do not provide evidence for the hypothesis that the sensitivity of citation to patent grant depends on author location.

In our final analysis of the impact of institutional affiliation on the anticommons effect, we turn to the affiliations as reported on the patent rather than the publication and assess whether the source of funding for research mediates the sensitivity of the citation rate to patent grant. In (7-1), we compare the coefficients associated with patents for which at least one of the assignees is a public sector organization versus those that are exclusively assigned to private firms. In (7-2), we divide out the impact of patent grant according to whether the patent applicant acknowledges Federal funding for the research upon which the patent is based. In both cases,

one cannot distinguish a separate impact according to the source of reported research funding (in the patent). However, since policy interventions aimed at the anticommons effect would be primarily focused on research funded by public sector organizations or the Federal government itself, it is important to emphasize the anti-commons effect may exist in organizations beyond the direct province of the most likely policy interventions.

Does the Impact of Patent Grant Depend on Patent Type?

At its heart, the anti-commons hypothesis is premised on the difficulty of using contracting methods to overcome the limitations on access to knowledge covered under patent protection. As noted above we would expect contractual difficulties or the imposition of patent rights to be strongest for patents with multiple competing claims (extensive prior art), broad claims, in areas of broad applicability such as tools. Table 8 explores this hypothesis by examining the interaction between the POST-GRANT dummy and our two measures of the degree of prior art cited in the patent (PATENT BACKWARD CITATIONS and PATENT BACKWARD REFERENCES). In both cases, there is a quantitatively large and statistically significant interaction. A one standard deviation increase in the level of prior patent art citation made by a patent-paper pair is associated with an additional 6% decline in the expected citation rate after the patent is issued. Table 9 shifts the focus from the extent of the patent thicket to the scope and nature of the patent grant. Though # CLAIMS is an extremely imperfect (and noisy) measure of the strength of patent rights evidence for the anticommons effect is strengthened by the fact that the predicted decline of citations as the result of patent grant is larger as the number of patent claims increases. In contrast, though a key tenet of the anticommons theory is that the effects are particularly salient for research tool patents, there is no evidence that the impact of patent grant depends on the nature of the invention.

Exploiting Variation in Patent Grant Delay

Though encouraging, it is important to emphasize that the results so far have relied on the presence of the control group of non-patented articles. While we have emphasized the comparable nature of our control group, and included variables to account for observable differences, the potential for unobserved differences in the two groups which correlate to the pattern of forward citations remains a possibility. We examine this by exploring an empirical

approach relying exclusively on a sample composed of articles which are ultimately associated with a patent-paper pair. In the absence of a control group, it is difficult to disentangle the impact of patent grant from the impact of age on the citation rate, particularly if one simultaneously controls for fixed quality differences across articles. Indeed, if the amount of time elapsed between publication and patent grant were constant across articles, the impact of post-grant would not be separately identified from a set of age fixed effects.

Nonetheless, in Table 10 we the estimate of the impact of patent grant relying solely on differences across articles in the amount of time between publication and patent grant. The first column of Table 10 suggests that the anticommons effect is indeed identified by this variation. This specification controls for article quality by including #AUTHORS, US AUTHOR, and PUBLIC SECTOR AUTHOR, as well as a full set of age and year fixed effects. The coefficient on POST-GRANT is quantitatively and statistically significant, with a magnitude similar to that found in the second and third columns of Table 5 (patent grant is associated with a 14-17% decline in citation rate). Following the analysis in Table 6, the last two columns of Table 10 address include article fixed effects with the difference between the two specifications being that, whereas (10-2) includes the entire sample, (10-3) only focuses on articles which have at least one US author and at least one author with a public sector affiliation.

VI.   Discussion

This paper provides the first systematic empirical test of the anti-commons hypothesis. Our empirical approach exploits the fact that the duality of knowledge is captured in the phenomenon of patent-paper pairs. Specifically, we exploit variation in both the occurrence of patent-paper pairs among a group of similar scientific publications and the patent grant delay thus allowing us to examine pieces of knowledge in distinctive institutional settings. Our evidence suggests that knowledge duality in general and patent-paper pairs in particular are an important phenomenon (nearly 50% of articles published in *Nature Biotechnology* are associated with a pair). And, that for scientific knowledge subject to both Open and Private Property institutional regimes, the granting of IPR is associated with a significant but modest decline in knowledge accumulation as measured by forward citations. Moreover, the decline in forward citations becomes more pronounced with the number of years elapsed since patent grant, and is greatest for knowledge produced with university or public institutions for which patent grant is

likely to be "news". We also find that the decline in forward citations is greatest for scientific articles paired with patents that cite a high levels of both patent and non-patent (publication-based) prior art. Overall, we are able to reject the null hypothesis that IP rights have no impact on the diffusion of scientific knowledge.

These patterns provide a novel perspective on the economic consequences of the privatization of the scientific commons. Rather than simply serving to facilitate a "market for ideas," IP may indeed restrict the diffusion of scientific research and the ability of future researchers to "stand on the shoulders of giants". However, erecting a (property rights) barrier to the accumulation of knowledge does not eliminate all Open Science use of that knowledge. The demand for specific published findings by follow-on researchers is relatively inelastic, limiting the empirical impact of the anti-commons effect in our estimates to between one in six and one in eleven research projects. Furthermore, disentangling the differential anti-commons effect among different types of future researchers – in the public versus the private sector, in the same or different geographic area as the initial publishing authors (be it local, regional or national), or among those who themselves generate patent-paper pairs or not – remains an open empirical question and the subject of future research. The magnitude and distribution of the anti-commons effect among distinctive groups of researchers has important implications for policy and for the degree to which anti-commons issues should be used to animate and justify changes in patent policy over dual knowledge generated in academia. For example, the policy implications of our findings would depend considerably on whether the "missing" citations are associated with independent public sector researchers versus being centered exclusively on follow-on research by private sector teams.

Moreover, it should be emphasized that our evidence for the anti-commons effect captures only one aspect of the impact of IP on dual knowledge. IPR may enhance incentives for (unobserved) research (particularly by private sector organizations) or lead to more effective commercialization (which is far more costly than the basic research component itself). Without a detailed accounting of the size of these effects versus the effect identified here, it is impossible to infer the elements of an optimal policy.

Despite these caveats, the importance of adjudicating this debate should not be understated. The production of dual-use knowledge is increasingly central in scientific research and of great interest to those who fund research. Biotechnology and the life sciences have grown

enormously as a share of overall research activities, and similar questions arise in the context of areas such as nanotechnology and open source software. Because biotechnology and related disciplines simultaneously offer the potential for fundamental scientific discoveries *and* commercial breakthroughs, our traditional understanding of the role of Private Property institutions and the norms of Open Science is open to question.   By providing a window into the impact of IP rights on the diffusion and accumulation of scientific knowledge, this paper offers some insights into the policy tradeoffs associated with embedding scientific knowledge in two distinctive and potentially opposing institutional regimes.

References

Adams, J. (1990) "Fundamental Stocks of Knowledge and Productivity Growth," *Journal of Political Economy*, 98, 673-702.

Allison, Paul D. and Richard Waterman. 2002. "Fixed effects negative binomial regression models." Ross Stolzenberg (ed.), *Sociological Methodology* 2002. Boston: Basil Blackwell.

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

Argyres, Nicholas. and Julia. Liebskind (1998). "Privatizing the Intellectual Commons: Universities and the Commercialization of Biotechnology Research," *Journal of Economic Behavior and Organization..*

Arora, Ashish. 1995. Licensing Tacit Knowledge: Intellectual Property Rights And The Market For Know-How", Economics of Innovation and New Technology, 1995, Vol. 4, 41-49.

Arora, Ashish, Andrea Fosfuri, Alfonso Gambardella (2001) "Markets for Technology: Economics of Innovation and Corporate Strategy", Cambridge, MA: MIT Press, December 2001

Arrow, Kenneth. "Economic Welfare and the Allocation of Resources for Invention." In *The Rate and Direction of Inventive Activity*, edited by Richard R. Nelson, 609-25. Princeton, NJ: Princeton University Press, 1962.

Azoulay, Pierre, Waverly Ding, and Toby Stuart. (2005). "The Impact of Academic Patenting on (Public) Research Output. Working Paper 2005.

Blumenthal, David, Causino, Nancy, Eric Campbell, and Karen Seashore Louis (1996). "Relationships between Academic Institutions and Industry in the Life Sciences — An Industry Survey" *New England Journal of Medicine*, Volume 334:368-374.

Bremmer, H. The First Two Decades of the Bayh-Dole Act as Public Policy. Address to the National Association of State Univrsities and Land Grant Colleges, November 11, 2001.

Bush, V. (1945). Science: The Endless Frontier, A Report to the President. Washington, D.C., U.S. Government Printing Office.

Campbell, Eric G., Brian R. Clarridge, Manjusha Gokhale, Lauren Birenbaum, Stephen Hilgartner, Neil A. Holtzman, and David Blumenthal. 2002. Data Withholding in Academic Genetics: Evidence from a National Survey. *Journal of the American Medical Association* 287 (4):473 - 479.

Campbell EG, Weissman JS, Causino N, Blumenthal D. Data-withholding in academic medicine: Characteristics of faculty denied access to research results and biomaterials. <u>Research Policy</u> 2000;29:303-12.

Cockburn, Iain, and Rebecca Henderson "Absorptive Capacity, Coauthoring Behavior, and the Organization of Research in Drug Discovery." *Journal of Industrial Economics*, 1998, 46(2), pp. 157-182.

Cockburn, Iain, Samual Kortum and Scott Stern. 2002. "Are All Patent Examiners Created Equal? Examiners, Patent Characteristics and Litigation Outcomes." Intellectual Property in the Knowledge-Based Economy, National Academy of Sceinces, 2003.

Cohen, S., Chang, A., Boyer, H., & Helling, R. 1973. Construction of biologically functional bacterial plasmids *in vitro*. <u>Proc. Natl. Acad. Sci. USA</u>, Vol. 70: 3240-3244.

Crespi, R. S. 'Patents on Genes: Can the Issues be clarified?' (1999/2000) 5 BioScience Law Review, 199 – 204.

Dasgupta, Partha. and David, Paul. 1994. "Towards a New Economics of Science" Research Policy (23), 487-521.

David, Paul. 2003. "Can 'Open Science' be Protected from the Evolving Regime of IPR Protections?" Stanford Working Paper #03-011.

David, Paul. 2001a. "From Keeping 'Nature's Secrets' to the Institutionalization of 'Open Science" Stanford Working Paper #01-006.

David, Paul. 2001b. "Will Building Good Fences really make Good Neighbors in Science" Stanford Working Paper #01-005.

David, Paul. 2000. "The Digital Technology Boomerang: New Intellectual Property Rights Threaten Global "Open Science" Stanford Working Paper #00-016.

Ding, Waverly, Fiona Murray and Toby Stuart. 2005. "An Empirical Study of Gender Differences in Patenting in the Academic Life Sciences", Working Paper.

Ducor P., 2000. Intellectual property. Coauthorship and coinventorship. Science 289(5481), 873-875.

Eisenberg, R. 1996. "Public Research and Private Development: Patents and Technology Transfer in Government-Sponsored Research." *Virginia Law Review* :1663-1727.

Etzkowitz, H., 1998. The norms of entrepreneurial science: cognitive effects of the new university-industry linkages. Research Policy 27, 823-833.

Galambos L. and J. Sewell. 1995. Networks of Innovation: Vaccine Development at Merck, Sharp & Dohme, and Mulford, 1895-1995. (New York: Cambridge University Press.

Gambardella, A. 1995. Science and Innovation. Cambridge University Press, Cambridge UK.

Gans, J. and S. Stern 2000. Incumbency and R&D Incentives: Licensing the Gale of Creative Destruction," (with. J. Gans), *Journal of Economics and Management Strategy*, 2000

Geison, G. L. 1995. The Private Science of Louis Pasteur. Princeton, NJ: Princeton University Press.

Hall, B.H. and R.H. Ziedonis (2001). "The Patent Paradox Revisited: An Empirical Study of Patenting in the US Semiconductor Industry, 1979-95." *RAND Journal of Economics,* 32(1): 101-128.

Hardin, G. 1968. "The Tragedy of the Commons," *Science*, 162(1968):1243-1248.

Hausman, Jerry A., Zvi Griliches, and Broonwyn Hall. 1984. "Econometric Models for Count Data with an Application to the Patents-R&D Relationship. *Econometrica* 52 (July), pp. 909-937.

Heller, M. 1998. "The Tragedy of the Anticommons: Property in the Transition from Marx to Markets," 111 *Harvard Law Review* 621 (1998).

Heller, M. and R. Eisenberg. 1998. "Can patents deter innovation? The anticommons in biomedical research", *Science*, vol. 280.

Henderson, R., A. Jaffe, M. Trajtenberg. 1994. Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *Quarterly Journal of Economics*, August 1993, Vol. 434, pp. 578-598

Jaffe, Adam B. and Josh Lerner. 2004. Innovation and Its Discontents: How our broken patent system is endangering innovation and progress, and what to do about it. Princeton: Princeton University Press.

Jensen, Richard and Marie C. Thursby. 2001. "Proofs and prototypes for sale: The licensing of university inventions." *American Economic Review*, 91:1, pp. 240-59.

Kenney M., 1986. Biotechnology: The University-Industrial Complex. New Haven: Yale University Press.

Kimball, D. 2001. "Patenting Genes: Risks and Rewards versus Politics and Policies", Christie Parker & Hale LLP Publications.

Kline S, and Nathan Rosenberg, 1986, "An Overview of Innovation," in *The Positive Sum Strategy: Harnessing Technology for Economic Growthy,* eds. R. Landau and N. Rosenberg, National Academy Press, Washington (DC).

Köhler, G. & Milstein, C. 1975. 'Continuous cultures of fused cells secreting antibody of predefined specificity.' Nature 256, 495-497.

Kremer, M. 1997. "Patent Buy-Outs: A Mechanism for Encouraging Innovation," NBER, Working Paper no. 6304.

Krimsky, Sheldon. 2003. Science and the Private Interest. Rowman-Littlefield Publishing Co., August 2003.

Lawson, Charles. 2002. Patenting genes and gene sequences and competition: patenting at the expense of competition. *Federal Law Review* 30:97-133.

Lemley, M.A. and Allison, J.R. (2002). "The Growing Complexity of the United States Patent System," *Boston University Law Review,* 82, p. 77-144.

Lessig, L. 2002. The Future of Ideas: The Fate of the Commons in a Connected World. Random House.

Levin, R., A. Klevorick, R. Nelson, S. Winter. 1987. Appropriating the returns from industrial R&D. Brookings Papers Econom. Activity 3 783-831.

Li, M. 1997. Use of a modified bacteriophage to probe the interactions between peptides and ion channel receptors in mammalian cells. Nature Biotechnology 15:6 (559-563)

Markiewicz, Kira and Alberto DiMinin. 2004. "Commercializing the Laboratory: The Relationship between Faculty Patenting and Publishing". Working Paper.

Marshall, Elliott. 2000. "Property Claims: A Deluge of Patents Creates Legal Hassles for Research". Vol. 288, p. 255.

Merges, R., Mennell, P.S. and M. Lemley (2003), *Intellectual Proprety in the New Technological Age,* Aspen Publishers.

Merges, R. and Nelson, R. R. 1990. "On the Complex Economics of Patent Scope," *Columbia Law Review*, Vol. 90.

Merges, R.P. and Nelson, R.R. (1994). On Limiting or Encouraging Rivalry in Technical Progress: The Effect of Patent Scope Decisions. Journal of Economic Behavior and Organization, 25 (1), 1-24.

Merton, R. The Sociology of Science: Theoretical and Empirical Investigations. Edited by Norman Storer. Chicago: University of Chicago Press, 1973.

Mowery, David, Richard Nelson, Bhaven Sampat, and Arvids Ziedonis. 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 30:99-119.

Murray, F. 2002. "Innovation as overlapping Scientific and Technological trajectories: exploring tissue engineering", Research Policy, Vol. 31.

Narin, F., Olivastro, D., 1992. Status Report: Linkage Between Technology and Science. Research Policy Vol 21(3).

Nature Biotechnology 1996. Editorial: A Broad and Inclusive Enterprise. March 1996 Volume 14 Number 3 p235.

Nelkin, D. 1984. Science as Intellectual Property: Who Controls Scientific Research? New York: MacMillan Publishing Co.

Nelson, R.R. (1959) "The Simple Economics of Basic Scientific Research," *Journal of Political Economy* 67(3), 297-306.

Orsenigo, L. 1989. The Emergence of Biotechnology: Institutions and Markets in Industrial Innovation.", Pinter Publishers, London.

Powell. W., Koput, K., Smith-Doerr, L. 1996. Inter-organizational Collaboration and the locus of Innovation: Networks of learning in Biotechnology. Administrative Science Quarterly 41, 116-145.

Romer, Paul M, 1990. "Endogenous Technological Change," *Journal of Political Economy*, Vol. 98 (5) pp. S71-102.

Rosenberg, N. (1974). "Science, Invention, and Economic Growth." The Economic Journal 84(333): 90-108

Shapiro, C. (2001), "Navigating the Patent Thicket: Cross Licenses, Patent Pools, and Standard Setting," in A.B. Jaffe, J. Lerner and S. Stern (eds.), *Innovation Policy and the Economy*, Vol. 1.

Stern, Scott. 1999. "Do Scientists Pay to be Scientists". NBER Working Paper w7410.

Stokes, Donald. 1997. Pasteur's Quadrant: basic science and technological innovation. Washington D.C.: The Brookings Institution.

Straus, Joseph, Henrik Holzapfel, and Matthias Lindenmeir. 2002. Empirical Survey on "Genetic Inventions and Patent Law". Munich.

Thomson, J.A. et al. (1998). *Science* 282, 1145-1147.

Thursby, G., Thursby M. 2003. University Licensing Under Bayh-Dole: What are the Issues and Evidence? Working Paper accessed from opensource.mit.edu/papers/Thursby.pdf.

Walsh, John P., Ashish Arora, and Wesley M. Cohen. 2002. The Patenting and Licensing of Research Tools and Biomedical Innovation: US National Academies' Science, Technology and Economic Policy Board.

Weitzman, M. "Free Access vs. Private Ownership as Alternative Systems for Managing Common Property." *Journal of Economic Theory* 8, no. 2 (June 1974): 225-234.

Williams-Jones, B. (2002) "History of a Gene Patent: Tracing the Development, Marketing, and Application of Commercial BRCA Testing in Canada" *Health Law Journal*, Vol. 10, p. 121-144.

Zamore, P.D., Tuschl, T, Sharp, P.A., and Bartel, D.P. (2000). "RNAi: Double-stranded RNA directs the ATP-dependent cleavage of mRNA. *Cell*, 101(1), 25-33.

Ziedonis, R. (2004). "Don't Fence Me In: Fragmented Markets for Technology and the Patent Acquisition Strategies of Firms," *Management Science*, June.

Zucker, L., Darby, M., Brewer, M., 1998. Intellectual Human Capital and the Birth of U.S. Biotechnology Enterprises American Economic Review 88, 290-306.

Zozulya, S., Lioubin, M., Hill, R.J., Abram, C., Gishizky, M. 1999. Mapping Signal transduction pathways by phage display. Nature Biotechnology, 17:2 (1193-1198).

TABLE 1
VARIABLES & DEFINITIONS

VARIABLE	DEFINITION	SOURCE
CITATION-YEAR CHARACTERISTICS		
FORWARD CITATIONS$_{jt}$	# of Forwad Citations to Article j in Year t	Science Citation Index (SCI)
YEAR$_t$	Year in which FORWARD CITATIONS are received	SCI
AGE$_{jt}$	YEAR – PUBLICATION YEAR	Nature Biotechnology (NB)
PUBLICATION CHARACTERISTICS		
PUBLICATION YEAR$_j$	Year in which article is published	NB
# AUTHORS$_j$	Count of the number of authors of Article j	NB
US AUTHOR$_j$	Dummy variable equal to 1 if *at least* one of the institutional affiliations associated with Article j is in the US; 0 otherwise	NB
PUBLIC AUTHOR$_j$	Dummy variable equal to 1 if *at least* one of the institutional affiliations associated with Article j is a university or government organization; 0 otherwise	NB
PRIVATE AUTHOR$_j$	Dummy variable equal to 1 if *at least* one of the institutional affiliations associated with Article j is a biotechnology or pharmaceutical firm; 0 otherwise	NB
TOTAL CITATIONS$_j$	# of FORWARD CITATIONS from publication date to 2003	SCI
PATENT CHARACTERISTICS		
PATENTED$_j$	Dummy variable equal to 1 if Article is associated with a patent issued by the USPTO prior to October, 2003	USPTO
GRANT YEAR$_j$	YEAR in which PATENT has been granted	USPTO
PATENT, POST-GRANT$_{jy}$	Dummy variable equal to 1 if PATENTED = 1 and YEAR > GRANT YEAR	USPTO
# INVENTORS	Count of the number of inventors listed in the granted patent associated with Article j; 0 if PATENTED = 0.	USPTO
PATENT LAG$_j$	Days elapsed between patent priority application and grant dates, 0 if PATENTED = 0	USPTO
PATENT BACK CITATIONS$_j$	Count of the number of citations to *patented prior art* included in the granted patent associated with Article j; 0 if PATENTED = 0	USPTO
PATENT BACK REFERENCES$_j$	Count of the number of citations to *non-patent prior art* included in the granted patent associated with Article j; 0 if PATENTED = 0	USPTO
# CLAIMS	Count of the number of allowed claims; 0 if PATENED = 0	USPTO
TOOLS PATENT$_j$	Dummy variable equal to 1 if primary patent classes are associated with *research tools* (as opposed to composition of matter patents); 0 if PATENTED = 0	USPTO; author verification
PUBLIC SECTOR ASSIGNEE$_j$	Dummy variable equal to 1 if *at least* one of the assignees on the patent associated with Article j is a university or government organization; 0 otherwise	USPTO; author verification
GOVT. FUNDED$_j$	Dummy variable equal to 1 if patent disclosure reports a Federal interest, indicating Federal funding of research upon which patent is based; 0 otherwise	USPTO

TABLE 2

MEANS & STANDARD DEVIATIONS

VARIABLE	N	MEAN	STANDARD DEVIATION	MIN	MAX
CITATION-YEAR CHARACTERISTICS					
FORWARD CITATIONS	1688	9.35	12.29	0	184
TOTAL CITATIONS	1688	54.74	58.84	2	523
CITATION YEAR	1688	1999.95	1.52	1997	2002
AGE	1688	2.05	1.52	0	5
PUBLICATION CHARACTERISTICS *(N=340 total articles)*					
PUBLICATION YEAR	340	1998.03	0.83	1997	1999
# AUTHORS	340	5.89	3.20	1	20
US AUTHOR	340	0.59	0.49	0	1
PUBLIC SECTOR AUTHOR	340	0.90	0.30	0	1
PRIVATE SECTOR AUTHOR	340	0.32	0.47	0	1
PATENT CHARACTERISTICS (N = 340 total articles, 169 articles associated with USPTO patents)					
PATENTED	340	0.50	0.50	0	1
GRANT YEAR[*]	169	2000.54	1.71	1996	2002
PATENT, POST-GRANT[^]	1688	0.24	0.43	0	1
# INVENTORS[*]	169	3.04	1.59	1	8
PATENT LAG[*]	169	1126.07	480.10	238	3714
PATENT BACK CITATIONS[*]	169	7.26	13.10	0	79
PATENT BACK REFERENCES[*]	169	28.19	37.25	0	226
# CLAIMS[*]	169	21.12	15.00	2	94
TOOLS PATENT[*]	169	0.58	0.49	0	1
PUBLIC SECTOR ASSIGNEE[*]	169	0.65	0.48	0	1
GOVT FUNDED[*]	169	0.29	0.45	0	1

* Summary statistics for these measures is calculated only for those article for which PATENTED = 1 and is weighted by Article (i.e., N = 169).

^ Summary statistics for PATENT, POST-GRANT is calculated over all articles, weighted by citation year

TABLE 3

MEANS CONDITIONAL ON PATENT STATUS

	NO PATENT	PATENTED
# Publications	171	169
FORWARD CITATIONS	8.96	10.04
# AUTHORS	5.76	6.03
US AUTHOR	0.53	0.65
PUBLIC SECTOR AUTHOR	0.93	0.86
PRIVATE SECTOR AUTHOR	0.25	0.38

TABLE 4

CROSS-SECTIONAL RESULTS

	NEGATIVE BINOMIAL Dep Var = FORWARD CITATIONS (Coefs reported as incident rate ratios)	
	(4-1) Baseline Model[*]	**(4-2)** With Publication Controls[*]
PATENTED	**1.156** **(0.063)**	1.085 (0.056)
# AUTHORS		**1.034** **(0.008)**
US AUTHOR		**1.266** **(0.064)**
PUBLIC SECTOR AUTHOR		0.930 (0.108)
Parametric Restrictions		
Age FEs = 0	**# Restrict 5** χ^2 **219.83** **p-value 0.000**	**# Restrict 5** χ^2 **236.83** **p-value 0.000**
Year FEs = 0~	**# Restrict 5** χ^2 **20.80** **p-value 0.001**	**# Restrict 5** χ^2 **16.76** **p-value 0.005**
R-squared		
Log-likelihood	-5257.86	-5230.73
P-value of Chi	0.00	0.00
# of Observations	1688	1688

 * Robust standard errors are in parentheses.
 ~ Year FEs included for 1998-2002 (1997 is excluded).

TABLE 5
THE IMPACT OF PATENT GRANT:
DIFFERENCES-IN-DIFFERENCES ESTIMATES

	OLS: DepVar = ln(FORWARD CITATIONS)	NEGATIVE BINOMIAL Dep Var = FORWARD CITATIONS (Coefs reported as incident rate ratios)		
	(5-1) Marginal Impact, with Article FEs	**(5-2)** Selection and Marginal Effects w/ Controls	**(5-3)** Selection and Marginal Effects w/. AGE & YEAR FE	**(5-4)** Marginal Effects, with Article FEs
PATENT CHARACTERISTICS				
PATENTED		**1.177** **(0.078)**	**1.181** **(0.081)**	
PATENTED, POST-GRANT	**-0.097** **(0.056)**	**0.834** **(0.069)**	**0.841** **(0.070)**	**0.911** **(0.039)**
CONTROL VARIABLES				
# AUTHORS		**1.035** **(0.008)**	**1.035** **(0.008)**	
US AUTHOR		**1.282** **(0.066)**	**1.272** **(0.065)**	
PUBLIC SECTOR AUTHOR		0.906 (0.108)	0.922 (0.065)	
AGE	**0.959** **(0.033)**	**3.151** **(0.199)**		
AGE*AGE	**-0.160** **(0.006)**	**0.834** **(0.011)**		
PUBLICATION YEAR = 1998		**1.251** **(0.073)**		
PUBLICATION YEAR = 1998		**1.263** **(0.089)**		
Parametric Restrictions				
Article FEs = 0	**# Restrict 340** **χ^2 10882.97** **p-value 0.00**			**# Restrict 338** **χ^2 11038.17** **p-value 0.00**
Age FEs = 0			**# Restrict 5** **χ^2 237.92** **p-value 0.00**	**# Restrict 5** **χ^2 480.30** **p-value 0.00**
Year FEs = 0˜			**# Restrict 5** **χ^2 17.40** **p-value 0.004**	# Restrict 5 χ^2 8.97 p-value 0.11
Regression Statistics				
Log-likelihood		-5266.35	-5227.61	-4028.87
P-value of Chi		0.00	0.00	0.00
R-Squared	0.75			
# of Observations	1688	1688	1688	1688

* Robust standard errors are in parentheses.

˜ Year FEs included for 1998-2002 (1997 is excluded).

TABLE 6
DIFFS-IN-DIFFS RESULTS
BY INSTITUTIONAL OR NATIONAL AFFILIATION*

	NEGATIVE BINOMIAL REGRESSIONS Dep Var = FORWARD CITATIONS			
	(6-1) No Public Author v. Public Author	**(6-2)** No Private Author v. Private Author	**(6-3)** No US Author v. US Author	**(6-4)** US Author Articles Only
ARTICLE CHARACTERISTICS				
PATENTED, POST-GRANT				**0.900** **(0.049)**
PATENTED, POST-GRANT * NO PUBLIC AUTHOR	0.952 (0.106)			
PATENTED, POST-GRANT * PUBLIC AUTHOR	**0.903** **(0.041)**			
PATENTED, POST-GRANT * NO PRIVATE AUTHOR		**0.882** **(0.045)**		
PATENTED, POST-GRANT * PRIVATE AUTHOR		0.955 (0.062)		
PATENTED, POST-GRANT * NO US AUTHOR			0.913 (0.058)	
PATENTED, POST-GRANT * US AUTHOR			*0.910* *(0.047)*	
CONTROL VARIABLES				
Article Pair FEs = 0	**# Restrict 338** **χ^2 10963.05** **p-value 0.00**	**# Restrict 338** **χ^2 11205.55** **p-value 0.00**	**# Restrict 338** **χ^2 11022.10** **p-value 0.00**	**# Restrict 198** **χ^2 9398.12** **p-value 0.00**
Age FEs = 0	**# Restrict 5** **χ^2 475.44** **p-value 0.00**	**# Restrict 5** **χ^2 469.29** **p-value 0.00**	**# Restrict 5** **χ^2 469.89** **p-value 0.00**	**# Restrict 5** **χ^2 328.69** **p-value 0.00**
Year FEs = 0˜	# Restrict 5 χ^2 2.51 p-value 0.77	# Restrict 5 χ^2 1.00 p-value 0.96	# Restrict 5 χ^2 1.34 p-value 0.93	# Restrict 5 χ^2 3.75 p-value 0.59
Regression Statistics				
Log-likelihood	-4028.70	-4028.32	-4028.87	-2429.43
P-value of Chi	0.00	0.00	0.00	0.00
# of Observations	1688	1688	1688	994

* Robust standard errors are in parentheses.

˜ Year FEs included for 1998-2002 (1997 is excluded).

TABLE 7
DIFFS-IN-DIFFS RESULTS
BY NATURE OF RESEARCH FUNDING *

	NEGATIVE BINOMIAL Dep Var = FORWARD CITATIONS (Coefs reported as incident rate ratios)	
	(7-1) No Public Sector Assignee Versus Public Sector Assignee	**(7-2)** Non Govt. Funded Versus Govt. Funded
PATENTED, POST-GRANT * NO PUBLIC SECTOR ASSIGNEE	*0.899* *(0.056)*	
PATENTED, POST-GRANT * PUBLIC SECTOR ASSIGNEE	0.921 (0.049)	
PATENTED, POST-GRANT * NON-GOVT. FUNDED		*0.913* *(0.046)*
PATENTED, POST-GRANT * GOVT. FUNDED		0.907 (0.061)
Parametric Restrictions		
Article Pair FEs = 0	**# Restrict 338** χ^2 **11053.87** **p-value** **0.00**	**# Restrict 338** χ^2 **10919.50** **p-value** **0.00**
Age FEs = 0	**# Restrict 5** χ^2 **517.05** **p-value** **0.00**	**# Restrict 5** χ^2 **471.38** **p-value** **0.000**
Year FEs = 0˜	**# Restrict 5** χ^2 **48.68** **p-value** **0.07**	**# Restrict 5** χ^2 **11.37** **p-value** **0.05**
Log-likelihood	-4028.82	-4028.86
P-value of Chi	0.00	0.00
# of Observations	1688	1688

* Robust standard errors are in parentheses.

˜ Year FEs included for 1998-2002 (1997 is excluded).

TABLE 8
DIFFS-IN-DIFFS RESULTS
WITH PATENT CHARACTERISTIC INTERACTIONS *

	NEGATIVE BINOMIAL REGRESSIONS Dep Var = FORWARD CITATIONS		
	(8-1) PATENT LAG Interaction	**(8-2)** PATENT BACKWARD CITATIONS Interaction	**(8-3)** PATENT BACKWARD REFERENCE Interaction
"DIRECT" EFFECT OF PATENT GRANT			
PATENTED, POST-GRANT	1.043 (0.097)	0.952 (0.045)	0.987 (0.049)
INTERACTION EFFECTS			
PATENTED, POST-GRANT * PATENT LAG	**0.999 (0.0000)**		
PATENTED, POST-GRANT * PATENT BACKWARD CITATIONS		**0.995 (0.001)**	
PATENTED, POST-GRANT * PATENT BACKWARD REFERENCES			**0.997 (0.0000)**
CONTROL VARIABLES			
Article Pair FEs = 0	**# Restrict 340** χ^2 **10965.59** **p-value 0.00**	**# Restrict 340** χ^2 **10838.08** **p-value 0.00**	**# Restrict 340** χ^2 **11070.68** **p-value 0.00**
Age FEs = 0	**# Restrict 5** χ^2 **535.25** **p-value 0.00**	**# Restrict 5** χ^2 **480.83** **p-value 0.00**	**# Restrict 5** χ^2 **471.55** **p-value 0.00**
Year FEs = 0~	**# Restrict 5** χ^2 **76.93** **p-value 0.000**	**# Restrict 5** χ^2 **18.24** **p-value 0.00**	# Restrict 5 χ^2 0.89 p-value 0.97
Regression Statistics			
Log-likelihood	-4027.97	-4026.88	-4024.95
P-value of Chi	0.00	0.00	0.00
# of Observations	1688	1688	1688

  * Robust standard errors are in parentheses.
  ~ Year FEs included for 1998-2002 (1997 is excluded).

TABLE 9
DIFFS-IN-DIFFS RESULTS
BY TYPE OF IPR*

	NEGATIVE BINOMIAL Dep Var = FORWARD CITATIONS (Coefs reported as incident rate ratios)	
	(9-1) NUMBER OF CLAIMS Interaction*	(9-2) No Tools Versus Research Tools Patents*
PATENTED, POST-GRANT	1.017 (0.063)	
PATENTED, POST-GRANT * NUMBER OF CLAIMS	**0.995** **(0.002)**	
PATENTED, POST-GRANT * NON-RESEARCH TOOL		*0.899* *(0.055)*
PATENTED, POST-GRANT * RESEARCH TOOL		0.918 (0.048)
Parametric Restrictions		
Article Pair FEs = 0	# Restrict 338 χ^2 11139.06 p-value 0.00	# Restrict 340 χ^2 11037.62 p-value 0.00
Age FEs = 0	# Restrict 5 χ^2 484.17 p-value 0.00	# Restrict 5 χ^2 481.60 p-value 0.000
Year FEs = 0~	# Restrict 5 χ^2 10.15 p-value 0.07	# Restrict 5 χ^2 11.00 p-value 0.05
Log-likelihood	-4026.73	-4028.83
P-value of Chi	0.00	0.00
# of Observations	1688	1688

* Robust standard errors are in parentheses.

~ Year FEs included for 1998-2002 (1997 is excluded).

TABLE 10
PATENTED ARTICLES ONLY

	NEGATIVE BINOMIAL Dep Var = FORWARD CITATIONS (Coefs reported as incident rate ratios)		
	(10-1) Marginal Effects w/. AGE & YEAR FE	**(10-2)** Marginal Effects, with Article FEs (Full Sample)	**(10-3)** Marginal Effects, with Article FEs (US Public Sector Authors Only)
PATENT CHARACTERISTICS			
PATENTED, POST-GRANT	**0.857** **(0.079)**	0.978 (0.048)	**0.877** **(0.053)**
CONTROL VARIABLES			
# AUTHORS	**1.043** **(0.010)**		
US AUTHOR	**1.430** **(0.103)**		
PUBLIC SECTOR AUTHOR	0.893 (0.131)		
Parametric Restrictions			
Article FEs = 0		**# Restrict 167** χ^2 5740.68 **p-value 0.00**	**# Restrict 167** χ^2 4784..79 **p-value 0.00**
Age FEs = 0	**# Restrict 5** χ^2 124.19 **p-value 0.00**	**# Restrict 5** χ^2 321.26 **p-value 0.00**	**# Restrict 5** χ^2 235.05 **p-value 0.00**
Year FEs = 0~	# Restrict 5 χ^2 4.28 p-value 0.510	**# Restrict 5** χ^2 51.01 **p-value 0.00**	**# Restrict 5** χ^2 93.56 **p-value 0.00**
Regression Statistics			
Log-likelihood	-2667.07	-2039.791	-1099.151
P-value of Chi	0.00	0.00	0.00
# of Observations	849	849	462

* Robust standard errors are in parentheses.

~ Year FEs included for 1998-2002 (1997 is excluded).

APPENDIX A
CONDITIONAL MEANS
BY PATENTED & AUTHOR AFFILIATION *

	No Patent		Patented	
	No Public Author	Public Author	No Public Author	Public Author
# PUBS	12	159	23	146
FORWARD CITATIONS	9.38	8.92	12.02	9.73
# AUTHORS	7.75	5.61	7.61	5.78
US AUTHOR	0.67	0.52	0.83	0.62

FIGURE A
THE STOKES MODEL

		Consideration of Use?	
		NO	YES
Quest for Fundamental Understanding?	NO		Pure Applied Research (Edison)
	YES	Pure Basic Research (Bohr)	Use-inspired / translational Basic research (Pasteur)

FIGURE B
CITATION DISTRIBUTION

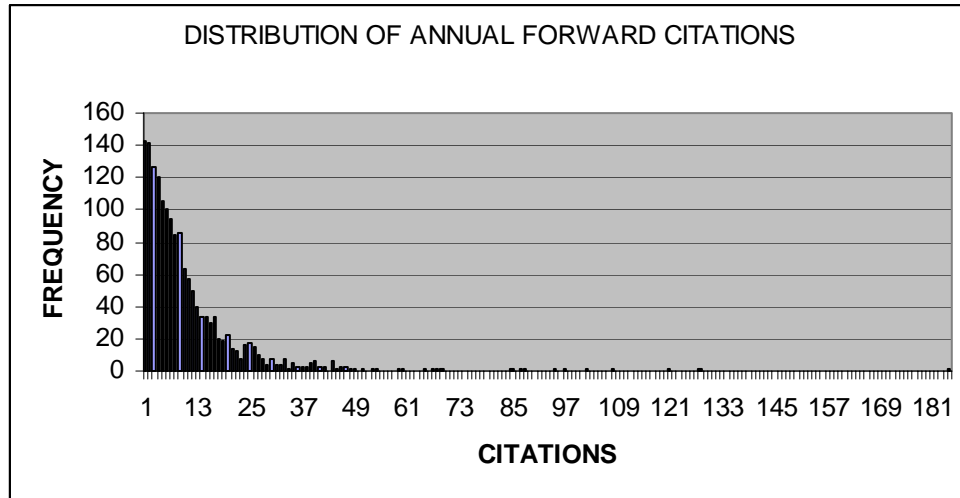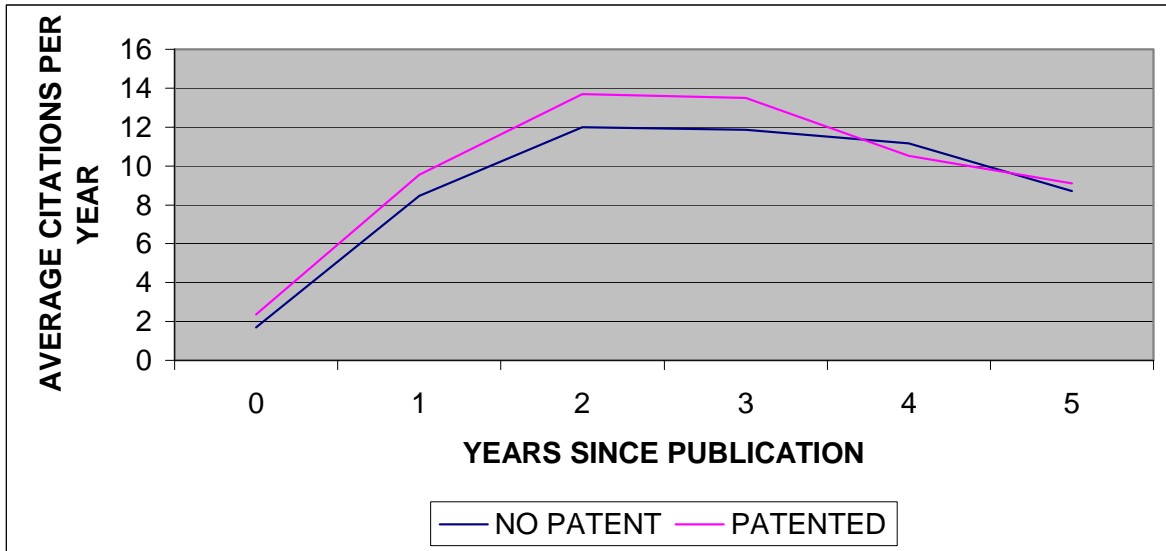DISTRIBUTION OF ANNUAL FORWARD CITATIONS

FIGURE C
CITATIONS BY TYPE BY AGE

FIGURE D
IMPACT OF PATENT GRANT ON FORWARD CITATIONS,
BY YEAR BEFORE AND AFTER PATENT GRANT
(NEGATIVE BINOMIAL WITH ARTICLE FEs)