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

March 2005

Prepared for the
NBER Academic Science and Entrepreneurship Conference
Santa Fe, NM

* We thank Jeff Furman, both for extremely insightful comments and suggestions as well as for generous assistance in obtaining the citation data. We also thank Iain Cockburn, Wes Cohen, Shane Greenstein, Rebecca Henderson, Roberto Fernandez, Roberto Rigobon and seminar participants at the SPRU Conference in Honor of Keith Pavitt, the 2004 International Industrial Organization Conference, the NBER Academic Science and Entrepreneurship Conference, the Georgia Tech Roundtable on Engineering Entrepreneurship Research, the MIT Economic Sociology Seminar, the MIT Industrial Organization Seminar, Duke, and the University of Michigan for extremely useful comments. We also thank Kranthi Vistakula and Mitun Ranka for their research assistance. All errors are our own.

Do Formal Intellectual Property Rights

Hinder the Free Flow of Scientific Knowledge?

A Test of the Anti-Commons Hypothesis

Fiona Murray
Scott Stern

March 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 "anticommons effects" has been anecdotal. A central issue in this debate is how intellectual property rights over a given piece of knowledge affect the propensity of future researchers to draw upon that knowledge in their own scientific research activities. In this paper, we develop and implement an empirical approach to uncover this effect, exploiting two key aspects of the process of disclosing and protecting scientific knowledge: (a) scientific knowledge receiving formal IP often appears also in the form of scientific research articles (a phenomena we refer to as a "patent-paper pair") and (b) 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, anticommons 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 other publications from the same journal for which no patent is granted), we find evidence for a modest anticommons effect. While publications linked to patent grants are associated with a higher overall citation rate, 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.

Journal of Economic Literature Classification Numbers: O300, O330, O340, L330
Keywords: anti-commons, intellectual property, academic science, Pasteur's Quadrant.

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 newly emerging 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 oncogenes 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 1987, a broad patent for the Oncomouse was granted by the USPTO. The Oncomouse patent was more controversial than most; in 1987, not only was the Oncomouse the first living animal to be patented, but Harvard's licensee Dupont aggressively enforced those 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 disciplines such molecular biology that underpin modern biotechnology. 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 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) -- leading to an expansion of formal intellectual property rights over knowledge that is first disclosed within the scientific literature.

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 "anticommons effect." At its core, the anticommons debate is an argument over whether the expansion of IPR (in the form of patents and/or copyrights) over scientific knowledge that would traditionally have been maintained in the public domain is "privatizing" the intellectual commons, and limiting the benefits from scientific progress (Heller and Eisenberg, 1998; Argyres & Liebskind, 1998; David, 2001b). Specifically, the anticommons 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).

This paper reports findings from a novel empirical approach to evaluate the salience of the anti-commons effect on the diffusion of scientific knowledge. Our strategy exploits the existence of patent-paper pairs, a phenomenon that is central to evaluating the anti-commons effect. The dual disclosure of a given "piece" of knowledge in scientific research articles and also in the patent record provides a concrete and specific starting point from which to examine the impact of intellectual property rights on the diffusion of scientific knowledge.

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 IP environments – a pre-grant period where no rights are present and a post-grant period in which specific 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 Anticommons 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 IPR 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 anticommons 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 the location and number of authors on the article. Second, there is robust evidence for a quantitatively modest but statistically significant anticommons effect; across different specifications, the article citation rate declines by 9 to 17% after a patent grant. Thirdly, the anticommons effect is particularly strong for articles with US and public sector co-authors. While we are cautious in our interpretation, our evidence suggests that the anticommons 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, including the role of knowledge duality. Section III develops the empirical test to assess the anticommons 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 Anticommons Effect

Intellectual Property Rights and the Anticommons Effect

At the heart of the anticommons debate lies the rapid expansion of intellectual property rights over knowledge that was traditionally placed in the public domain without restrictions or formal property rights. The debate is therefore largely focused on scientific knowledge generated in academia and traditionally disclosed through scientific publication.[1] The scientific knowledge under discussion is therefore focused on "basic" knowledge that addresses questions of fundamental scientific interest. The traditional literature on innovation draws a clear distinction 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, 1986). Building on the classical distinction between "free access" and "private property rights" emphasized by Weitzman (1974) and research in the sociology of science (Merton, 1973), scholarship in the "new" economics of Science focuses on comparisons of alternative institutional arrangements and incentive regimes associated with these distinctive types of knowledge and knowledge production and largely assumes strongly exogenous institutional regimes and disclosure choices (Dasgupta and David 1994). The new economics of science links basic research with an Open Rights regime and applied research with a Private Property Rights regime.

Scientific 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 by future scientific generations, the importance of demonstrating experimental replicability, 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

---

[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.

the shoulders of giants."[2] In contrast, the incentives that govern applied knowledge production typically map to a quite distinctive private property rights regime: rather than rewarding researchers on the basis of their ultimate impact on follow-on research, private research incentives 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). The classical rationale for the patent system is to provide incentives for innovation while simultaneously achieving the disclosure of technological knowledge for long-term knowledge accumulation (Scotchmer 1987).

Overall, these two institutional regimes with their associated modes of disclosure have traditionally mapped to distinctive forms of knowledge and knowledge production. Basic research that explores topics of scientific interest is undertaken in the institutions of Open Science and is disclosed through publication; applied research takes place a Private Property institutional regime and is disclosed through patents (to the extent that disclosure takes place). While this allows for a concise formulation of the relationship between the nature of knowledge and the incentives provided for its production and distribution, this analytical framework fails when knowledge has *both* basic and applied value. Stokes (1997) reformulated the traditional distinction between basic and applied research by highlighting the *duality* of research; rather than being empirically distinct, a single discovery could simultaneously possess both applied and basic characteristics. Figure A illustrates this essential insight. Stokes' formulation also 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).

When knowledge has this dual character, the institutional regimes and mode of its disclosure becomes endogenous: scientists have a well-defined set of disclosure choices, ranging from 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. For much knowledge generated within academia, the dilemma is not so much a

[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, 1999; Murray, 2002).

shift away from publication and towards patenting, but rather a shift towards simultaneous publication and patenting. 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 more than simply reflecting a rise in research and patenting in academia over knowledge unrelated (or only tangentially related) to scientific research. Rather, patent-paper pairs disclose the same knowledge in two different "documents," reflecting the interests of researchers in having an impact on both future scientific research as well as commercial exploitation. However, by both publishing and patenting the same piece of knowledge, that knowledge becomes embedded in two distinct institutional regimes, each of which is governed by specific (and somewhat opposing) rules for access and attribution.

Origins of patent-paper pairs in Molecular Biology

On the one hand, patent-paper pairs as a phenomenon are not restricted to molecular biology and modern biotechnology; moreover, the anticommons debate, as a key implication of dual knowledge and patent-paper pairs, is not conceptually restricted to these fields. However, the debate over the anti-commons effect has largely focused on this arena. Since its inception, biotechnology was seen as making a contribution both to scientific progress and technological progress, leading to a complex interaction between the institutions of open Science and those of Private Property Rights (Kenney, 1986; Orsenigo, 1989; Powell et al., 1996; Gambardella, 1995). Patenting of knowledge that was traditionally disclosed solely through publication and that remains disclosed through publication has become particularly salient. This shift in the rise of dual knowledge in life sciences research 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 molecular biology and genetics.

The rise in dual knowledge in molecular biology can be traced back to the early 1970s. Most key scientific milestones in the field, but for a few noted exceptions, have been disclosed as patent-paper pairs: the techniques of recombinant DNA which provide insights into the cellular

[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).

machinery of the cell but also laid the foundation for the production of recombinant therapeutic proteins, the Oncomouse (which simultaneously provided great insight into the sources of cancer while becoming a model for investigating cancer therapies), the discovery of RNA interference (a further step towards explaining DNA replication but also the foundation of a potentially new therapeutic category), tissue engineered constructs (teaching us how cells develop in three dimensions in response to growth signals but also having the potential to serve as organ replacements), and others (Murray 2002). Interestingly, for each of these leading examples of knowledge that has both basic and applied interest, the research was both published in a leading scientific journal and ultimately covered by intellectual property rights.[4]

At the same time as this revolution in biotechnology occurred, policy shifts also encouraged academic IPR by clarifying ownership of IP generated in academia using Federal funding. Prior to this time, if patent applications were filed by universities on behalf of investigators the assignment of patent rights and their subsequent licensing required case by case negotiation with the funding agency. The 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). While not fully responsible for the rise in IP particularly in the life sciences, this clarification in institutional practice made the disclosure decision over dual knowledge more likely to be skewed towards patent filing and the generation of patent paper pairs rather than simply publication alone. Nevertheless, individual researchers and institutions varied widely in how they behaved in this new environment.

Third, there was a significant expansion in the scope of IPR across the life sciences. After the 1980 *Diamond vs. Chakrabarty* decision, followed by the granting of the Oncomouse patent in 1987, IPR had expanded comprehensively into 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

---

[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* for publication, 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)

the potential economic value of the technology that was being generated in their research programs" (Bremer 2001). Thus the ground work for the anticommons debate was laid: increasingly dual knowledge and the widening scope of patents meant that many scientific discoveries now contained patentable inventions; the 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 include patenting. In the period between 1989 and 1999 US Research One universities received over 6,000 life science utility patents (Owen-Smith and Powell 2002) and patent-paper pairs became an important disclosure phenomenon (Murray 2002). In short, the molecular biology intellectual commons became increasingly littered with intellectual property rights.

The anticommons debate in Molecular Biology

While not conceptually limited to molecular biology, it is hardly surprising that this arena has become the fulcrum of the anticommons debate. Largely through the National Institutes of Health (NIH) the United States federal government spends over US$20 billion per year on public research in the health sciences. As is clear from its mission to "foster fundamental creative discoveries… as a basis to advance significantly the Nation's capacity to protect and improve health [and]…expand the knowledge base in medical and associated sciences in order to enhance the Nation's economic well-being…" (NIH Almanac, 2005), the agency is particularly concerned that the decisions regarding disclosure, dissemination and use of the knowledge generated from its funding serve to promote cumulative innovation in both fundamental understanding and useful application. Throughout the 1980s, following the passage of the Bayh-Dole Act, the dominant perspective on how to accomplish these broad objectives was that patents could serve an important purpose in ensuring cumulative innovation towards commercialization, with little or no impact on the other side of the cumulative equation – scientific knowledge accumulation. However with the dramatic rise in life science patenting, and the sharply increased prevalence of patent-paper pairs over key discoveries, several highly publicized cases came to light regarding patent rights that had served to impinge on the work of academic scientists building on knowledge which had been publicly disclosed through scientific publication. Increasingly, economists, law and technology scholars, and policymakers have focused their concerns over the impact of intellectual property rights on the growth and diffusion of new scientific knowledge on this arena (Heller and Eisenberg, 1998; David, 2001a; Campbell et al. 2002; Strauss et al. 2002; Walsh et al. 2002).

The knowledge described by Leder's *Cell* publication was also covered by a "paired" Oncomouse patent (4,736,866), licensed exclusively to DuPont Corporation who, as the donor of a broad research gift to support Leder's lab, had actually participated in the decision to patent. Soon after the patents were filed, DuPont imposed their property rights through a stringent series of 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 that used the Oncomouse in the research, and limits on informal material exchange and sharing of these mice between scientific colleagues. Each of these practices represented a dramatic shift in the publication and exchange norms typically practiced by individuals in this field. The patent-paper pair that allowed the imposition of property rights over ideas that Leder and colleagues had described in their *Cell* publication was 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, a Nobel Prize winner in oncogene research, senior scientists in mouse genetics met at Cold Spring Harbor in 1992 where they "began talking of revolution" (Science 2 April 1993, p. 23). As a reflection of the concern for anti-commons effects on public science, the NIH stepped in and sanctioned a non-profit central facility (The Jackson Laboratory) to serve as a repository for genetically altered mouse strains in an attempt to give researchers equal access to genetic research tools (Science V. 288, p.255, April 2000). However, while the Jackson Labs could distribute mice, the Oncomouse and another DuPont licensed mouse Cre-loxP were still covered by the unpopular licensing terms. DuPont continued to contact scientists and demand review of their articles before publication and "nobody was able to exchange materials" (David Einhorn quoted in Science, 2000). In 1997, Varmus, now in his new role as Director of NIH, wrote to DuPont that they could "seriously impede further basic research" and started to negotiate terms to set limits on the reach-through royalty clauses and to set standards. The Working Group for Mouse Genomics and Genetics made a series of joint recommendations and, in 2000, an agreement was reached between NIH and DuPont which outlined the terms under which the technology can 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). With this agreement, the limitations on publication and other reach-through rights were lifted. Dissemination was also permitted under a Materials Transfer Agreement structured to meet the needs of different types of uses. All non-profit recipients are required to have an overarching agreement in place with DuPont, while commercial entities are required to purchase a commercial license.

While the Oncomouse has become the canonical example of the potential for the growing coverage of intellectual property rights over scientific knowledge that was traditionally disclosed through publication to impinge upon cumulative scientific innovation, the debate was raised more sharply in the prospective debate 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 intellectual property over research conducted in "Open Science" that goes beyond the anticommons issues.

At its most expansive the debate over patenting is an attempt to examine the balance between the role of patents in cumulative commercial innovation on the one hand and cumulative scientific innovation on the other. 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 by providing incentives through monopoly rights. Thus patents might encourage DuPont to further develop the Oncomouse as a model for cancer therapeutics, or encourage Celera or others to develop new therapies or diagnostics. 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). In the context of knowledge generated in academia and its instantiation in patent-paper pairs, the third leg of the traditional argument for IP that support the role of IPR as an inducement for disclosure over secrecy does not apply since there is a presumption of and evidence of disclosure in publications.

In contrast, the anti-commons hypothesis suggests that there are unforeseen negative consequences having intellectual property rights over scientific knowledge that was traditionally disclosed through publication, and that those consequences impinge specifically upon cumulative scientific progress. 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 anticommons hypothesis posits that the use of IP over areas traditionally maintained in the public commons can undermine the process of cumulative scientific discovery. As illustrated by experience with the Oncomouse, Heller and Eisenberg (1998) suggest that the assignment of IPR to scientific research provides researchers with a control right to *exclude* others using that knowledge for the traditional purpose of cumulative knowledge production. 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 disclosed by the publication and patent and may restrict the diffusion of that knowledge. Instead of raising incentives for discovery, the use of IPR over knowledge which has been traditionally associated with Open Science can lower the equilibrium level of research productivity and lead to the under use of the "knowledge commons" – hence the anti-commons (Heller, 1998; Heller and Eisenberg, 1998; David 2003). Furthermore, to the extent that IPR is narrow in scope and highly dispersed across individuals and institutions, fragmentation can impose significant transaction costs and limit research progress (Eisenberg, 1996; Shapiro, 2001; Hall and Ziedonis, 2001; Ziedonis, 2004).

Thus the anti-commons debate focuses on knowledge generated in the public domain, disclosed through publication and now subject to IPR over that knowledge (patent-paper pairs). It hinges on the tradeoff the positive impact on commercial accumulation (translation) on the one hand and the negative anticommons effect on further scientific accumulation on the other. It is useful to note that the related debate over IP for whole gene patents (such as those for the breast cancer genes BRCA-1 and BRCA-2) is rendered more complex knowledge. In particular, though similar issues arise, this example does not fall within the precise boundaries of the anticommons debate (William-Jones, 2002, Cho et al. 2004). In the breast cancer case, a portfolio of over a dozen US and international patents gave biotechnology firm Myriad exclusive rights to commercialize laboratory testing services, diagnostic test kits and therapeutic products that use the BRCA1/2 DNA sequences. This is a case in which knowledge generated in *industry* is patented and involve the imposition of patent rights by private companies rather than by academic institutions (Merges and Nelson, 1990; Crespi 2000). Here we might argue that IP at least encourages disclosure and we cannot assume that this knowledge would, in the absence of IP, have been disclosed (or even generated in the first instance). Similarly, the imposition of newly granted IP rights in the Axel case by Columbia University over its industry licensees while of concern for cumulative commercialization does not impinge upon academic researchers and therefore does not directly impact scientific knowledge accumulation.

III. Empirical Implications & an Empirical Test of the Anticommons Hypothesis

Empirical Implications

A detailed understanding of the canonical examples of patent-paper pairs leading to anticommons effects together with the theory as described above allows us to develop key empirical predictions regarding the impact of formal IPR on the free flow and diffusion of scientific knowledge and the ability of researchers to cumulatively build on each other's discoveries through the mechanisms of Open Science[5].

Specifically, the theory predicts that while the scientific community as a whole may benefit from the free dissemination and diffusion of knowledge, individual researchers have

[5] In an notable exception, Blumenthal et al (1996) provide a useful survey suggesting that scientists perceive that the increased use of intellectual property rights are limiting their access to research resources. However, extant studies have yet to disentangle whether such perceptions are associated with the precise type of knowledge under consideration (basic or applied) or the property rights regime associated with that knowledge (non-patented versus patented).

strong incentives to take advantage of the protections afforded by formal IP rights. If protected by IPR in the form of a patent-paper pair, 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. Therefore we would predict that for published ideas that are also covered by patents as was the case with the Oncomouse, future cumulative scientific innovation – as measured by published literature citing the original publication – would be stifled and would be lower relative to the case where the Oncomouse or another piece of scientific knowledge were not patented.

Second, and more precisely, we would expect that this stifling effect would have a clear temporal structure. We would predict that lower equilibrium cumulative research productivity would be precisely associated with the grant of the patent – in other words, DuPont or another patent assignee (or licensee) cannot impose IP rights until after the patent is granted. It is only after grant that researchers engaging in scientific knowledge accumulation might be impacted by stringent restrictions. Thus we would predict that the anticommons effect as measured by forward citations in the published literature would drop, relative to non-patent pairs *post patent grant.*

Third, we would only expect to see this temporal structure for published ideas published by public sector researchers funded from public sources. As a concrete example, in cases such as Myriad's, researchers would *anticipate* that knowledge disclosed in Myriad's scientific publications would also be the subject of a patent application and likely be associated with the imposition of patent rights. As such researchers might be less likely to work in these areas even prior to patent grant. Thus while citations might be lower overall, the patent grant would not represent "news" and we would not see the post-grant decline in cumulative scientific knowledge.

Fourth, we would expect the impact of patent grant to be strongest for publications with US-based authors. For an article written exclusively by non-US authors, the process of applying for and having a US patent granted is likely a very heterogeneous process; it is also likely that the scope of rights granted and enforcement activity is less likely to occur in a timely manner after rights have been granted. As a result, we expect to observe a more salient effect from grant behavior for articles with a domestic author.

Finally, we would anticipate that the anticommons effect would be the strongest for patent-paper pairs with particular patent characteristics – namely patents of significant prior art, broad scope, and specifically on research tools. At its heart, the anticommons 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 at issue, 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 anticommons effects. In addition, patents that are broad in scope 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 anticommons 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.

Patent-Paper Pairs and Patent Grant Delay

While the theoretical predictions of the anti-commons effect are relatively straightforward, systematic empirical evidence for the anticommons effect is quite limited. Empirical research has been hampered by a fundamental inference problem. Specifically, for a given piece of knowledge *with* IPR, one cannot observe the counterfactual impact that knowledge would have had if 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 on cumulative innovation is thus one of identification – 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 observe the cumulative innovation that follows.

We are able to use patent-paper pairs to provide the critical empirical leverage we need to evaluate key predictions of the anti-commons hypothesis. In particular, they allow us to develop a procedure that takes direct advantage of the dual institutional environments implied by the "patent-paper pairs" phenemonon. The starting point for our analysis is a sample of "pieces" of scientific knowledge that have been disclosed through publication in a top-tier peer-reviewed scientific journal. While any reader of the academic literature is aware that articles vary in quality, we limit the underlying heterogeneity of ideas by limiting our sample to articles published in a narrow time window in one journal; as a result, 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. We confirm the inherent patentability of all these research articles by working with a leading intellectual property lawyer. As expected we find that some articles are in fact instantiated as both a scientific publication and a patent – they constitute "patent-paper pairs." While articles are not randomly assigned to being pairs or not (as we might ideally like in a truly controlled experiment), qualitative interviews and other research in this area has suggested that conditional on patentability (as confirmed), the dual disclosure decision arises from a complex interplay between the characteristics of individual scientists and their institutions rather than systematically with underlying article quality or article characteristics. Thus we build a sample of similar scientific publications some of which are pairs and some of which are not. Indeed, as described in the next section, approximately 50% of the publications in our dataset in fact result in a patent-paper pair.

With the sample of publications we then exploit the fact that for a given patent-paper pair, there can exist a substantial gap between the date of scientific publication and the date at which the associated patent is granted. While publication in the scientific literature often occurs within 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, resulting in a 2-4 year time window). For any patent-paper pair, we can observe therefore observe a pre-patent grant period and a post-patent grant period. It is important to emphasize that our exploitation of patent grant delay allows us to capture more than simply examining articles at different points in time (indeed, simple "timing" effects are captured by the use of fixed age effects). Instead, patent grant shifts a given piece of publicly disclosed scientific knowledge into an alternative institutional regime and allows us to examine a piece of knowledge in two distinctive institutional regimes. In particular, until 2001, patent applications were kept secret until granted. As well, when patents are pending, patent applicants face substantial uncertainty over the scope of the rights they will receive. Finally, patent rights only commence with patent grant. In other words, during the pre-grant period, researchers building on findings that are disclosed in the scientific literature are unrestricted by (and often unaware of) pending patent applications. In the post-grant period, the knowledge in the publication is covered by IP rights.

Obviously this assumption depends upon the degree to which patent-paper pairs truly represent distinctive instantiations of the same knowledge. Careful examination of each patent-paper pairs confirms this. Consider the following example as an illustration:

> "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)

With this empirical structure we then estimate the impact of knowledge in different institutional regimes as a precise test of the anticommons hypothesis. As our measure of impact we use the annual rate of citation to the initial article in follow-on scientific research articles (publications). Such an analysis must account for the fact that we are measuring the impact of scientific research using citations, which come in the form of count data skewed heavily to the right. Therefore, except where noted, we employ a negative binomial model of the citations produced per year for each scientific article in our dataset. 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 pattern of citations over time to a scientific publication changes as the result of the receipt of a patent. 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. More precisely, if the grant of intellectual property hinders the ability of researchers to build on a given piece of knowledge, then the citation rate to the scientific publication disclosing that knowledge should fall after formal IP rights over that knowledge is granted. However, the impact of a given piece of research will vary considerably with its underlying importance and with the time elapsed since initial publication. 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 scientific maturity (through age and vintage effects). Taken together, our baseline empirical test for the anticommons hypothesis is therefore:

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

where γ_i is a fixed effect for each article, $\beta_{pubyear}$ is a cohort effect, $\delta_{t - pubyear}$ captures the age of the article, and POST-GRANT is a dummy variable equal to one only for years in which a patent is in force for an individual article.[6] In other words, we test for the anticommons 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.

In addition to estimating the overall average impact of patent grant, we can gain additional insight into the anti-commons hypothesis by exploring several variations of (1). First, it is likely that the impact of a patent grant increases over time, particularly in the first few years

[6] As well, it is possible to separately identify the degree of unobserved heterogeneity versus the impact of patent grant itself. In particular, one can replace the full set of fixed effects in the specification with an overall effect capturing the difference between patented and non patented articles, (i.e., $CITES_{i,pubyear,t} = f(\varepsilon_{i,pubyear,t}; \beta_{pubyear} + \delta_{t-pubyear} + \lambda PATENTED_i + \psi POST-GRANT_{i,t})$).

after the grant. As such, we can estimate whether the impact of patent grant changes with the time elapsed since patent grant. In addition, by interacting POST-GRANT with a range of characteristics, we can examine the various implications of the anticommons hypothesis described above specifically, how the effect of patent grant varies with technological, researcher and institutional factors associated with the scientific publication and the patent itself.

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).[7] 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

[7] 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 largely from academics but also from the newly founded biotechnology firms.

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 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.[8] 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

[8] 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.

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 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.[9] 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 and again found no systematic differences between the patented and unpatented articles. Overall then given the nature of the sampling process and the qualitative comparisons the assumption of comparability seems plausible. As such, the bulk of the empirical work assumes that, after controlling for overall quality, 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).[10] 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.

---

[9] 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.

[10] 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.

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 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 over 1100 days (i.e., just over 3 years).[11] PATENTED, POST-GRANT is the dummy variable indicating whether the citation-

[11] Some patents have been issued to articles in 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 the differences-in-differences strategy relies on those articles where we observe a change in the IP regime during those years where we observe a citation count).

year is linked to a patent-paper pair, and whether the patent has been issued and is in force.[12] 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 anticommons 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 anticommons 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

[12] 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.

within the 435 and 800 patent classes (mean = 0.58). Out of the 11 3-digit patent classes 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 anticommons 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).[13]

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

---

[13] 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.

citation advantage associated with patented articles is consistent with the presence of an anticommons effect, such an effect could result from differences in the characteristics of articles 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.[14] 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)

[14] 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.

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 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.[15] At face value, these estimates provide concrete evidence for the existence of an anticommons 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 that specific research direction after IPR are granted over that research.

The evidence for an anticommons 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. The pattern is encouraging (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 the impact of IPR accumulates over time; taking the pattern in Figure D at face value, the "size" of the

[15] 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.

anticommons effect becomes a sizable deterrent to future research efforts once patent protection is 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 anticommons 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. One way to examine this is the explores the impact of excluding the control group and 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. Conclusions

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 phenomena of patent-paper pairs (and that the granting of patents occurs only after a substantial lag). Our evidence suggests that knowledge duality is a quite important phenomenon (nearly 50% of articles published in *Nature Biotechnology* are associated with a pair) and that for scientific knowledge exhibiting this duality and being subject to both Open and Private Property regimes, the granting of IPR is associated with a significant but modest decline in knowledge accumulation as measured by forward citations. Moreover, the decline associated with patent grant becomes more pronounced with the number of years elapsed since the grant, specifically for papers more closely associated with university or public research, and for patents that cite a high level of prior art. Overall, we are able to reject the null hypothesis that IPR have no impact on the diffusion and use of associated scientific knowledge. However, we also find evidence that articles which are associated with patenting do continue a high level of citation, similar to the level associated with articles that never receive patent protection.

These patterns provide a novel perspective on the economic consequences of increasing the scope of IPR to cover knowledge which had traditionally been produced under the norms of Open Science. Rather than simply serving to facilitate a "market for ideas," IP may indeed restrict the diffusion of scientific research. However, erecting a barrier to the costless use of knowledge does not eliminate all use of that knowledge. The demand for specific research findings by follow-on researchers is relatively inelastic, limiting the empirical impact of the anti-commons effect.

Moreover, it should be emphasized that our evidence for the anti-commons effect captures only one aspect of the impact of IP on dual use 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.

The importance of adjudicating this debate should no be understated. Rather than a fringe activity, dual-use knowledge is increasingly the norm in scientific 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, traditional justifications for IPR and for the norms of Open Science are open to question. By providing a window into the impact of IP rights on the diffusion of scientific knowledge, this paper offers some insight into the policy tradeoffs associated with the interaction between public science and private commercialization.

References

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

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

Arora, A. 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, A., 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.

David Blumenthal, Nancyanne Causino, 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. Research Policy 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.

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.

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.

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

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)

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.

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.

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

FIGURE B

CITATION DISTRIBUTION

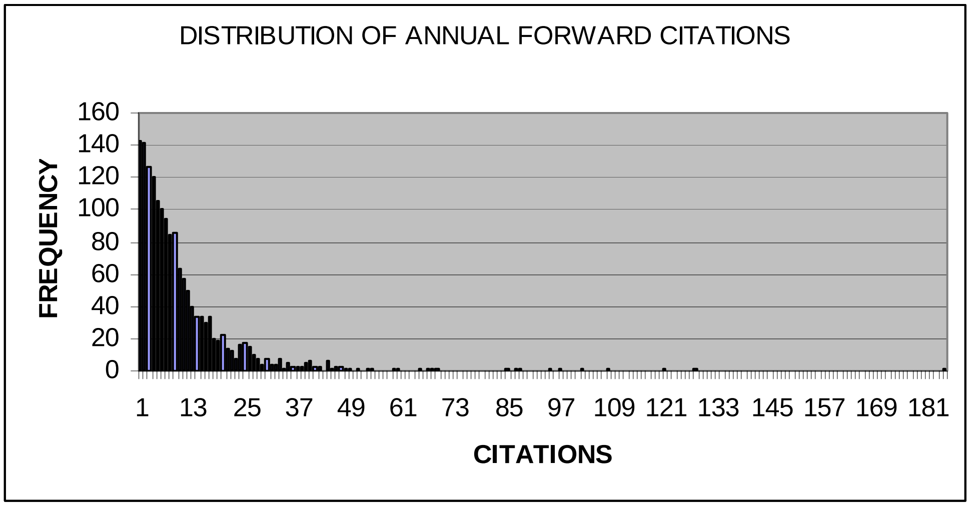

FIGURE C
CITATIONS BY TYPE BY AGE

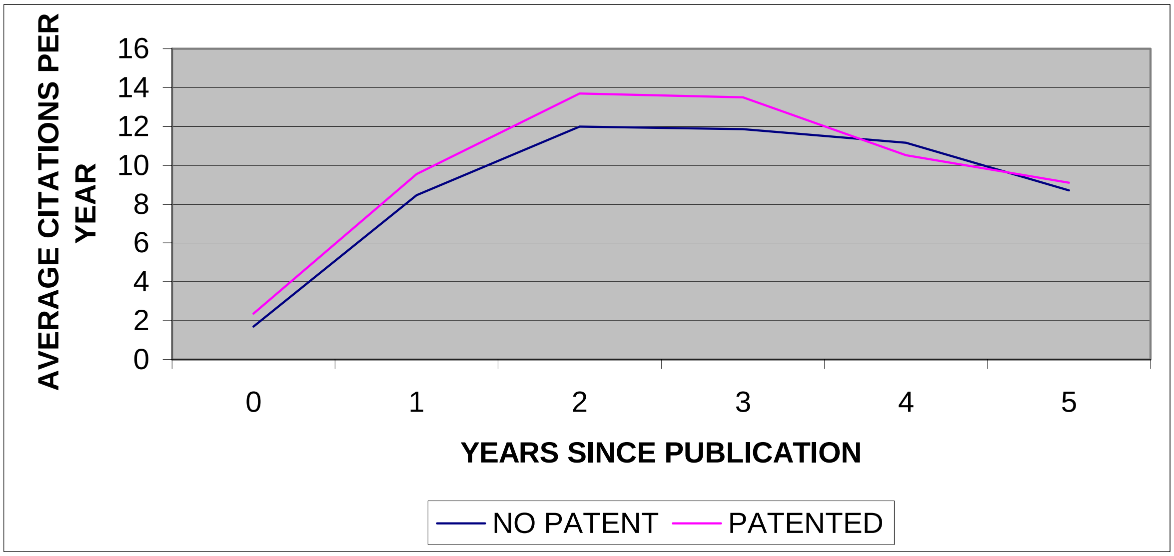

FIGURE D
IMPACT OF PATENT GRANT ON FORWARD CITATIONS,
BY YEAR BEFORE AND AFTER PATENT GRANT
(NEGATIVE BINOMIAL WITH ARTICLE FEs)