Patents, Papers, Pairs & Secrets: Contracting over the disclosure of scientific knowledge[*]

JOSHUA S. GANS
Melbourne Business School
University of Melbourne
200 Leicester Street, Carlton VIC 3053
j.gans@unimelb.edu.au

FIONA E. MURRAY
MIT Sloan School of Management
Cambridge, MA 02142
fmurray@mit.edu

SCOTT STERN
NBER & Kellogg School of Management
Northwestern University
Evanston, IL 60208
s-stern2@northwestern.edu

PRELIMINARY & INCOMPLETE
(9th September, 2008)

[*] Thanks to Chris Snyder and participants at the 2008 Organisational Economics Workshop (Sydney), Workshop on Scholarly Communication (Michigan) and seminars at Colorado, Stanford, Hong Kong, Melbourne, Queensland, and the Australian National University. JG acknowledges an ARC Discovery Grant for financial assistance. FM and SS acknowledge a National Science Foundation Science of Science Policy Grant. Responsibility for all errors lies with us. All correspondence to Joshua Gans, Melbourne Business School, 200 Leicester Street, Carlton VIC 3053 Australia, E-mail: J.Gans@unimelb.edu.au.

Patents, Papers, Pairs & Secrets: Contracting over the disclosure of scientific knowledge

Abstract

Motivated by the central role of voluntary knowledge disclosure in ensuring cumulative progress, this paper evaluates the conditions supporting the disclosure of privately funded new knowledge through scientific publication, patenting, or both. We ground our analysis in the incentives facing researchers and their funders: scientists have incentives to disclose discoveries through scientific publication, while firms have incentives to protect their ideas through patenting or secrecy. When negotiating their compensation, researchers and firms bargain over whether (and how) knowledge will be disclosed. We evaluate four different disclosure regimes: secrecy, commercial science (where the only disclosures result from patenting), open science (where the only disclosures occur through scientific publication) and patent-paper pairs (where the firm discloses along both dimensions). Our model derives conditions under which each of these outcomes emerges as the result of the strategic interaction between researchers and research funders, and offers a number of novel insights into the determinants of the *disclosure strategy* of a firm. Among other findings, our analysis highlights the subtle interdependency between patenting and publication: for example, as the knowledge required for patenting and publication converge, we will observe either secrecy or patent-paper pairs, to the exclusion of uni-dimensional disclosures (i.e., Commercial Science or Open Science). *Journal of Economic Literature* Classification Number: O34.

1. Introduction

Since the seminal work of Nelson (1959) and Arrow (1962), economists have sought to understand the economics of knowledge production. To the extent that scientific knowledge serves as a non-rivalrous input into technological innovation, its production plays a critical (and much-studied) role in economic growth (Romer, 1990). In recent years, however, the tradeoff between the production of non-rivalrous scientific knowledge widely disclosed through publications and the generation of commercially relevant knowledge excludable through intellectual property rights has received widespread attention (Dasgupta & David, 1994; Lessig, 2002; Heller 2008). In his seminal formulation of endogenous growth, Romer (1990) provides a resolution of this tradeoff by distinguishing the non-rival component of (scientific) knowledge produced by a research project from the rival commercial (technological) component. While traditionally thought of as arising from distinctive *types* of research investments, Romer assumes that a single investment by a private firm can produce a non-rival good that is also partially excludable. This formulation assumes that knowledge has a dual character: knowledge outputs can simultaneously serve as the basis for currently useful products but also provide a non-rival input for future knowledge production.

A concrete manifestation of the Romer formulation would be new knowledge that is simultaneously an important scientific discovery (e.g., one that might yield a Nobel Prize for the lead scientist) and consequential technology (e.g., one that might be associated with significant profits for the research funder). However, for dual knowledge to contribute to growth it must also be disclosed. Consider, for example, the US$30 billion perfume industry (devoted to fine fragrances that you might find in a department store). It may be surprising to learn that a company like US-based International Fragrance and Flavors (IFF) whose expertise ranges from fine and functional fragrances and snack food flavorings, invests over US$200M each year in R&D (about 9% revenues). However, much of the knowledge generated in their R&D projects remains the subject of trade secrets and undertaken by highly paid but little known “noses” whose expertise in the production of complex chemical mixtures is the stuff of legend. Thus, while these firms generate knowledge, the powerful role of secrecy limits the contribution of that knowledge to economic growth. This choice is not simply a matter of the type of knowledge and its patentability (Levin et al., 1987). Other aspects of perfume R&D including the chemist of many of perfume’s key high-value ingredients are the subject of extensive patenting. Beyond the traditional production of corporate patents, however, lies a wealth of knowledge disclosed in peer-reviewed publications. Far from a scientific backwater, perfumes, their complex chemistry and their interaction with the olfactory system have been the subject of three Nobel Prizes in the past century. In each case, the scientists worked within (or collaborated closely with) industry and having their knowledge published and serve as the basis for numerous patents. Patterns such as these suggest that at the heart of our understanding of endogenous growth lies not only the recognition that growth is driven by dual knowledge, but also that its power to drive the engine of growth lies in disclosure.

Our focus on disclosure is framed by recent literature highlighting the empirical significance of an alternative to traditional disclosure via papers, patents or secrets, known as *patent-paper pairs* (Ducor, 2000; Murray, 2002; Murray and Stern, 2007). For disclosure in a patent-paper pair, knowledge generated in a single research project must contribute to both scientific research and useful commercial (technical) applications. If it meets this test, then researchers can disclose their knowledge simultaneously in both documents; garnering property rights on the knowledge through patents but also making it available as an input for future generations of researchers. While they may, initially, appear to be a peculiar anomaly, as we will show, patent-paper pairs constitute the disclosure choices of both private firms and in academia, and across a wide range of disciplines and are an important instantiation of dual knowledge. This paper considers the implications of such dual knowledge, its implications for knowledge disclosure including patent-paper pairs, and the economic and strategic interactions mediating alternative disclosure choices.

In particular, we argue that for many projects, dual knowledge production supports a choice among four alternative disclosure regimes; secrecy, patenting (Commercial Science), publication (Academic Science), and patent-paper pairs. Moreover, these disclosure choices are grounded in the negotiation between researchers and the firms who fund their projects that mediates disclosure choices. By developing a simple economic framework to describe these negotiations, we can map more precisely the conditions under which each disclosure choice is likely to dominate. In doing so, we provide the basic microeconomic foundations for the disclosure choices that are central to our conceptualization of endogenous growth, and a framework within which to explore a range of empirical findings over disclosure.

At the core of our model lies the insight that there exists significant potential for conflict over disclosure between researchers and their funders. While this is a complex problem, at its most simple level, researchers want to publish the knowledge they produce in order to gain scientific recognition (often referred to as kudos) and are willing to take reduced wages in order to garner these reputational benefits (Stern, 2004; Murray & Hsi, 2007). On the other hand, funders want to retain the knowledge as a secret – a strategy at odds with disclosure through publication, or to use patent protection to allow for disclosure but to serve as a barrier to entry for potential rivals. With disclosure through patenting comes the potential to resolve the conflict between researchers and funders and disclose knowledge as patent-paper pairs. The resolution of this conflict arises through a simple negotiation between a single researcher and a private funder.

Our approach lays out a simple negotiation between a single researcher and a private funder. We assume, initially, a single period model with researchers only deriving benefits from disclosure through publications and negotiating over the level of that disclosure. Firms make a choice between patenting or secrecy, also negotiating over the disclosure level in publications.[1] The benefits to firms derives from the barriers to entry provided by their ultimately disclosure choices. We find that disclosure strategies map to the balance between marginal benefit to marginal cost ratio for disclosure through patenting and the marginal benefit to marginal cost ratio for disclosure through publication. We demonstrate that there are three drivers of a complementarity between patenting and publishing. First, to the extent that disclosures in each path overlap, once you have engaged in disclosures through one route, the marginal cost of disclosing through the other falls. Second, a patent (partially at least) protects a firm against the consequences of disclosures through the patent or publication. Thus, through this protection, taking out a patent reduces the marginal cost of publications. Finally, when scientist wages can fall to zero (as they might if scientific kudos is relatively important), disclosures through publication are the route by which the scientist is rewarded. Thus, to the extent that a patent raises commercial returns, this may increase publication disclosures as a means of compensating the scientist for their participation.

We then extend our model to a dynamic setting with overlapping generations of scientist

---

[1] Mukherjee & Stern (2008) consider a model whereby the only choices are between open science and secrecy and do not model the patenting decision and hence, does not address the main question of this paper that involves the interaction between the two disclosure paths.

and firm pairs. In this model, we consider the role of citation (in providing scientific kudos) and licensing (if providing a return to patenting). In our specification, publication can diminish the returns to licensing as it allows future pairs to work-around, more easily, the patent should license negotiations breakdown. This introduces substitutability between patenting and publishing in that taking out the former increases the costs associated with more of the latter. While this captures some of the concerns of scientists regarding the negative impact of commercial orientation on scientific practices, we demonstrate that the zone whereby we observe open science (publication without patenting) could increase.

The simple economic foundations of knowledge disclosure elaborated in our paper highlight the central role of knowledge production and disclosure in the economy. Our approach provides insights into a range of disclosure choices and allows us to make key predictions about the empirical variations in disclosure that we would anticipate. As such, our model is closely linked to an observable and critical phenomenon, one that, up to now, has been treated as a series of descriptions or puzzles rather than as a more systematic and strategic set of economic choices. The model we present also has a number of implications for the economics of knowledge, the institutional foundations of knowledge production, and for our understanding of the role of knowledge in economic growth. As we consider and justify expanding investments in scientific and technical knowledge production based on economic growth, we should pay close attention to their associated disclosure requirements. While researchers have highlighted the importance of freedom in providing for the effective production of knowledge (Aghion, Dewatripont, & Stein, 2005), disclosure is also critical. Even in cases where knowledge might be produced in the private sector, it may be that public funding, if contingent upon more unconstrained disclosure, is more effective. Likewise, while recent studies emphasize the potentially damaging nature of strong IP rights (Heller & Eisenberg, 1998), our results suggest that this is not necessarily the case. Indeed, IP rights might be associated with increased levels of patent-paper pairs and therefore more disclosure than commonly associated with scholarly journals.

The paper proceeds as follows. In Section 2, we elaborate the intellectual foundations upon which this research is built, current empirical studies describing the disclosure patterns across different types of knowledge and funding sources and the way in which our approach is at odds with much of the literature in this area. We then turn to the model and in Section 3 outline the foundations of our modeling approach, our assumptions and the basic framework, deriving predictions for the different disclosure equilibria: secrecy, patenting, publishing and patent-paper pairs. Section 4 then introduces our dynamic model extension while a final section concludes and suggests directions for future research.

2. The Disclosure of Scientific Knowledge

The importance of knowledge disclosure for economic growth and the challenges that disclosure poses to innovators and the tradeoff between disclosure and secrecy are, of course, well-known (Romer 1990, Arrow 1962). Beyond well-known motivating examples such as the secret formula for Coca-Cola, however, very little is known theoretically or empirically about the economic and strategic drivers shaping the margin between publically disclosed knowledge on the one hand and knowledge that is maintained a secret on the other (Lemley 2007).

At the broadest level, an implicit assumption links knowledge disclosure to both the type of knowledge and the organizational environment of knowledge production. According to Dasgupta and David (1994), basic and applied knowledge map to two distinctive institutional environments for disclosure (Dasgupta & David, 1994). Basic knowledge, produced in academia with public funding, is disclosed in the “Open Science” regime in peer-reviewed academic publications. Not only does publication allow for the adjudication of knowledge quality, the incorporation of knowledge into the research of others, and its later application to useful problems, it also rewards to scientists via kudos instead of intellectual property rights. The establishment of norms that facilitate full disclosure is grounded in institutional regime that recognizes scientific priority and a system of public (or coordinated) expenditures rewarding those who contribute to public knowledge production and disclosure over the long term (Merton, 1973). By premising career rewards (such as tenure) on disclosure through publication, the “Open Science” regime of disclosure through publication leverages the public goods nature of basic research and promotes cumulative innovation.

In contrast, the incentives that govern the “Commercial Science” disclosure regime – most notably patents – are typically mapped to applied science performed in the for-profit sector applied science. These 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; Scotchmer 1996). While the precise details of the two systems, and the ways in which secrecy may substitute for either patenting or publication, are complex, these broad distinctions cannot be overlooked: it is impossible to win a Nobel Prize without disclosing the discovery!

Elegant in its conceptualization, however, the simple mapping from the organizational setting of production, to the type of knowledge, and finally to disclosure is too simplistic. An emerging body of evidence points to a much richer, more nuanced set of disclosure strategies across organizational settings and within for given types of knowledge.

First, while some organizational arrangements are more likely to be associated with particular disclosure strategies, universities' patent portfolios and the growing contribution to the scientific literature made by industrial scientists suggests that no simple relationship exists. Patenting has been on the rise on university campuses around the world; today U.S. universities own over 50,000 patents. While these numbers continue to be dwarfed by the number of publications produced by university faculty at a ratio of at least 25 to 1, university faculty file more than 3,000 patents each year on the basis of at least 30,000 invention disclosures (USPTO, 2005; Ding, Murray & Stuart 2007; Azoulay, Ding & Stuart 2007). While considering changes on the university campus, a growing literature has also pointed to the rise in secrecy among academics suggesting that even in this organizational setting, full disclosure is unlikely to be the only disclosure strategy that prevails (Blumenthal, 1997). In contrast, while we, traditionally, consider industry to maintain its scientific knowledge as trade secrets or to serve as the engine of patenting, corporate publishing is also now regarded as a significant activity for corporate scientists. For example, in the period from 1985 to 1997, scientists at Intel produced 665 publications and 1700 patents (a ratio of 0.4:1). More surprising is the 4:1 ratio of publications to patents for scientists at Merck in the same period (Lim 1999). Moreover, some firms are now eschewing patenting altogether and using publications for some of their scientific knowledge (Bar-Gil & Pamovsky 2003) and increasingly seeking to disclose their ideas in pursuit of so-called open innovation strategies (Chesbrough 2003).

While organizational distinctions do not predict disclosure strategies, simple distinctions between types of knowledge also fail to define disclosure patterns. Nonetheless, following Vannevar Bush's prescription that research could be categorized into pure (basic) and applied, with a "perverse law governing research" under which "applied research invariably drives out the pure" (Bush, 1945), this separation has been the dominant paradigm guiding science policy. However, in his analysis of the motivation of scientists, Stokes (1997) argued that while "a great deal of research is wholly guided by one or the other goals of understand and use, some studies … show that the successive choices of research are influenced by both goals" (p. 12).

Taking Louis Pasteur as his eponymous example, Stokes suggests that for some scientists, and some research projects, basic and applied goals are blended (Geison 1995). Likewise, the joint goals of science and "a deeply industrial view" inspired Lord Kelvin's basic insights into physics (Smith & Wise 1989). Not only are research projects subject to the mixed motivations attributed to Pasteur, recent studies argue that knowledge from a single research project can be both basic and applied in its nature. In other words, the outcome of such research jointly contributes to fundamental knowledge and to commercial application, a situation much more akin to the role of knowledge in Romer's conceptualization of technology-driven economic growth (Murray & Stern 2007). To the extent that a given research investment (made in academia or industry) can yield knowledge of various types – both basic and applied, then it is amenable to disclosure in multiple channels including publications and patents– a disclosure strategy we refer to patent-paper pairs (Murray 2002). Pairs recognized the multifaceted nature of most research projects undertaken by researchers in academia and industry.

Knowledge produced in the course of these projects is often amenable for submission to peer-reviewed publications bringing academic kudos to its authors. At the same time, it may also meet the criteria for an invention making it possible for scientists to disclose the outputs of their knowledge production as both patents and papers. Fundamental scientific breakthroughs such as the development of the first genetically engineered mammal (the Oncomouse), human embryonic stem cells and recombinant DNA were patented and published. Academic scholars are not alone in pursuing these strategies. For example, while doing corporate research at Bell Labs, William Shockley developed the foundations for early semiconductors with his development of the transistor. The experiments he undertook in January 1948 are described in his Bell Labs lab notebook (No. 20455 pp: 128-32 (January 1948). Less than six months later, in June 1948, he filed for a US patent on the solid state transistor: Shockley, W. "Circuit Element Utilizing Semiconductive Material," (*U. S. Patent 2,569,347* issued September 25, 1951). In 1949, he published the theory of P-N junctions underlying the transistor inventor in the widely circulated Bell Labs journal (Shockley, 1949).

An emerging body of evidence suggests that across organizational settings, there is significant heterogeneity in disclosure even for pieces of knowledge equally "distant" to the market. Consider our knowledge of the human genome. The Human Genome Project, initiated by the Department of Energy in the late 1980s, was a publicly funded effort to provide a template/map of the entire human genome, essentially creating a "book of life" that might serve as a reference and guide to researchers linking genes to a wide variety of human diseases. Although researchers had already implicated a small number of genes in diseases such as cystic fibrosis prior to the mapping exercise, in the absence of any additional knowledge, the decision to map the full genome was based on the notion that one gene is as likely as another to provide important insights into the genetic basis of disease. By 2001, when (draft) sequencing was completed, data from the publically funded HGP, together with the competitive mapping efforts by privately funded start-up Celera, formed the basis of the full-genome database in the National Center for Biotechnology Information (NCBI) (Nature 2001, Science 2001). An entirely open, resource, the NCBI links each of the 20,000 individual genes scattered across 23 chromosomes to its position on the genome and any additional information about the gene including its documented role in disease, mouse models incorporating the gene, links to other species with similar genes etc. In the period since the map's release, it has guided considerable cumulative knowledge production on individual genes including over 4,000 patents on DNA sequences coding for human genes (Jensen & Murray 2005). What distribution of knowledge accumulation might be expect across the complement of 20,000 human genes? Romer's perspective, that knowledge maps are critical in shaping and enhancing the production of non-rival knowledge, anticipates that following basic mapping of the genome, knowledge accumulation would likely proceed at a similar rate across the entire landscape. Moreover, whatever the level of accumulation, we would expect that such substantially similar types of knowledge would be subject to one dominant disclosure strategy.

The evidence, however, is far more complex as illustrated by an analysis of knowledge disclosure associated with Chromosome 10 presented in Figure 1. While the full genetic sequence and location information of all 1005 protein coding human genes on the Chromosome have been known for almost a decade, 40% of all genes (400) are currently associated with no additional disclosed knowledge beyond the baseline HGP mapping. Another 453 genes (45%) of all Chromosome 10's mapped genes are associated with public knowledge (i.e. peer-reviewed publications) but lack any formal intellectual property claims. Essentially, disclosure of knowledge about these genes follows the logic of Open Science. Conversely, only 3 genes are disclosed following only the logic of Commercial Science. It is important to emphasize that while their disclosure patterns are striking different, an *ex ante* perspective there are few differences between these two groups of genes. Finally, about 149 genes are associated with knowledge disclosed as patent-paper pairs. In other words for 15% of Chromosome 10 the disclosed body of research on the salient human genes includes publications in the scientific literature, which at the same time are associated with formal IP rights over some application of that knowledge (Huang & Murray 2008).

Gene Patent	Yes	3	149
	No	400	453
		No	Yes
		Gene Publication	

Figure 1: Draft Disclosure Map of the 1005 genes on Chromosome 10

These patterns are provocative and according to our preliminary estimates reflective of the entire human genome landscape. Most important, the lack of publicly disclosed knowledge for a significant share of the genome is noteworthy. One might expect that with only 20,000 genes and with nearly hundreds of thousands of life science researchers around the world interested in these genes, we would find at least some documented, peer-reviewed evidence about each human gene. After all, one potential strategy for a new entrant into research would simply be to provide information about an otherwise unexplored gene. Of course, the lack of disclosure is at best inconclusive evidence that these genes are not the subject of research or that no knowledge about them exists. We cannot rule out the possibility that private researchers (or even public ones) have investigated a given gene but chosen to keep their knowledge secret.

Secrecy might be the chosen (non) disclosure strategy in cases where genetic knowledge is more commercially valuable in the absence of any disclosure. For example, when a firm wants to use a particular gene mutation as part of a complex genetic screening test and can certify that their results work without revealing exactly which mutations they are measuring. Alternatively, firms may keep their genetic knowledge a secret hoping that later innovators with generate complementary knowledge that relies upon their genetic secrets. Another possible source of secrecy is that, while revelation of ‘negative” knowledge accumulation on a particular gene may provide a significant public good (i.e. it reduces the marginal search cost for others), few opportunities for such “negative disclosures” exist. A researcher cannot secure IP rights for “failures” and publication opportunities continue to be extremely limited.

The second notable feature of our data is the diversity in disclosure across the relatively homogeneous knowledge setting; a finding that is suggestive of the notion that predictions of likely disclosure strategies are not simply grounded in the type of knowledge generated by a given research project. Instead, the logics of Open Science and Commercial Science can coexist for knowledge that is essentially of a fixed “type.” This is possible only when institutional arrangements are sufficiently flexible that they can incorporate the same type of knowledge. As the human genetics example illustrates, such flexibility arises when a single research project generates knowledge that not only meets the requirements of peer-reviewed journal disclosure but also the inventiveness requirements defined by intellectual property law, particularly patents. While knowledge of the human genome has been controversial in this regard – with widespread debate over the degree of disclosure and inventiveness properly associated with gene sequence patents – the two disclosure regimes provide researchers (and their funders) with the flexibility to make strategic choices between publication and patenting. The potential for a single project to generate knowledge that can either be published or patented speaks to the third feature of our data; the prevalence of patent-paper pairs. The fact that almost 20% of all protein coding human genes with any form of disclosure are associated with patent-paper pairs, underscores their importance as a disclosure strategy in this arena and throughout the modern knowledge economy. Each patent-paper pair (PPP) involves the disclosure of related (although not necessarily identical) knowledge in two distinct forms, providing very different forms of credit and claims over future cumulative knowledge. Disclosure in PPPs emphasizes that, at the discretion of researchers, a given investment in knowledge can simultaneously garner protection through the patent system while also being disclosed into the republic of science. Moreover, as emphasized in Murray (2002), such dual disclosures lead to diffusion into different communities with strong qualitative evidence that the community relying on publications for their follow-on knowledge accumulation is largely distinctive from (even though it may informally overlap with) the community building on the patent disclosure. In other words, PPPs provide a direct mechanism for and instantiation of the type of protection and diffusion of knowledge that lies at the core of the Romer model.

It is useful to emphasize that the amount and type of knowledge, produced in a particular project, and disclosed through patents and publications is closely related but important distinctions can arise. The most simple form of patent-paper pair – what we refer to as a *congruent pair* – is the case when the knowledge disclosed in the patent and in the paper is effectively identical. The formatting requirements of the two documents differ – patents structured around a set of claims, invention background and examples, publications around the articulation of a question, prior literature, methods and results. However, from the perspective of a follow-on innovator *congruent pairs* provide identical learning, as illustrated in the patent-paper pair on BLNK (pronounced "blink") on Chromosome 10. The BLNK gene encodes a B cell linker protein. In simple terms, proteins of this type sit on the surface of B cells and provide bridges between receptors and other proteins, thus regulating the biologic outcomes of B cell function and development. In 1999, scientists published discovered and described the sequence and function of the BLNK gene in a paper in the journal Immunity. As their abstract highlights: "*We describe here the identification of a novel B cell linker protein, termed BLNK, that interfaces the B cell receptor-associated Syk tyrosine kinase with PLCgamma, the Vav guanine nucleotide exchange factor, and the Grb2 and Nck adapter proteins*" (Fu et al, 1999). The researchers also filed patent # 5994522 on BLNK proteins describing in the patent background how "the discovery of molecules which interact with either Grb2 or PLC-γ … play a role in the regulation of … signaling pathways are desired. Accordingly, it is an object of the present invention to provide such molecules, termed 'BLNK' proteins, and to provide methods of using such molecules in screening assays." The paper provides no additional disclosure relative to the patent.

Two further types of patent-paper pairs arise when knowledge disclosed in the paper strictly exceeds the knowledge disclosed in the patent. In the case of what we refer to as a *landscape patent-paper pair*, the additional knowledge disclosed in the paper provides a "map" for additional search and future discoveries (Fleming & Sorenson, 2004). Due to the requirements in the scientific literature that a paper speak to the broader relationship between an invention and the scientific field, or to the underlying relationship between a set of possible invention variants, scientists must provide strictly greater disclosure in the paper over and above the patent. As such, landscape patent-paper pairs allow potential rivals to lower their cost of search for a distinct alternative to the knowledge covered by a piece of formal IP. The simplest illustration of this case can be found in chemical patents: many patents on a chemical substance will disclose a wide variety of chemical structure variants as a strategy for expanding the scope of IP protection. Most peer-reviewed chemistry journals would not consider such a list of variants to be "publishable." Instead, a researcher wishing to submit a publication would have to disclose additional insights into the relationship between the structures and the variation in their functions in the application of interest or some related property (for example the variation in tensile strength). In doing so, the scientist provides deeper insights into the scientific landscape guiding his own search for compounds and therefore the general principles for future researchers. Take as an example from Chromosome 10, the patent-paper pair on the SARP gene – the secreted frizzled-related protein 5. In 1997, researchers at LXR Biotechnology in California filed a patent application on a "Family of genes encoding apoptosis-related peptides, peptides encoded thereby and methods of use thereof" (Patent # 6433155). The patent disclosed and claimed the gene "sequence encoding the SARP polypeptide; vectors containing the sequence that encode all or part of the SARP protein, antibodies specific for SARP and their in diagnostic and therapeutic methods for treating diseases related to the regulation of SARP expression in tissue and bodily fluid samples, including cancers." In the same year, they published a paper entitled "SARPs: a family of secreted apoptosis-related proteins" in the leading journal, Proceedings of the National Academy of Sciences. Like the patent, the paper provided detailed gene sequence information, but went well beyond its disclosure in describing the potential role of the protein in cell death (apoptosis) and elaborating how "expression of SARPs modifies the intracellular levels of beta-catenin, suggesting that SARPs interfere with the Wnt-frizzled proteins signaling pathway." These important additional analyses make key contributions to the broader understanding of this class of genes but of course also provide follow-on researchers with a more detailed scientific landscape to further their research.

The final type of patent-paper pairs we identify as e*nabling patent-paper pairs*, characterized by paired papers providing future researchers with additional information on how to usefully replicate or use the invention described in the patent. Pairs of this type arise because of differences in the level of "enabling" disclosure required in the paper as compared to the patent rather requirements for disclosure that is of a quite distinctive nature. We again take an example from Chromosome 10, this time the patent-paper pair on the ERG gene encoding "early growth regulatory" proteins. The patent (# 5866325), filed in 1995 discloses "DNA sequences encoding novel DNA binding proteins implicated in regulation of early stages of cell growth. Illustratively provided are human and mouse origin DNA sequences encoding early growth regulatory ('Egr') proteins ... Also disclosed is a detailed analysis of the structure and function of the early growth regulatory protein, Egr-1, delineating independent and modular activation, repression, DNA-binding, and nuclear localization activities …." The researchers published a paper the same year disclosing the "Growth of LLC-PK1 renal cells is mediated by EGR-1" (Kinane et al., 1994) provides additional enabling methodological details over and above the patent disclosure. In this and other cases of enabling pairs, while the paper lowers the marginal cost of replication, in the absence of formal IP rights, the patent itself may blockade imitation. Whether other innovators are able to build on the patent-paper pair depends crucially on both the enforceability of the patent and whether the patent owner is willing to license at a cost allowing others to take advantage of the invention.

It is the contention of this paper that the complex disclosure decision over all three types of patent-paper pairs and the other forms of disclosure – patents only, papers only and, of course, secrecy – turns on a conflict in the disclosure incentives faced by researchers and their funders. Given conflicting objectives, disclosure choices arise from the strategic negotiation between the two parties. By assuming that for most research projects, part of the knowledge generated could be disclosed through patenting or publication, we can usefully separate disclosure choices from the type of knowledge. Second, while the organization of knowledge production (public or private) conditions the objectives of funders, they do not explicitly and simply drive disclosure outcomes. Instead, the micro-foundations of disclosure lie in a more subtle strategic interaction between researchers and the organizations that fund them.

As background to this strategic interaction, start by considering the range of motivations of researchers engaging in knowledge production. We build on Rosenberg (1990) who argues:

> Many scientists in private industry could honestly say that they are attempting to advance the frontiers of basic scientific knowledge, without any interest in possible applications. At the same time, the motivation of the research managers who decide to finance research in some basic field of science, may be strongly motivated by expectations of eventually useful findings. (p.169)

This suggests that even those scientists employed in the private sector are motivated by concerns beyond purely monetary rewards. Their career incentives, when coupled with the socialization that arises during their research training, led them to value their contributions to scientific knowledge and the recognition they garner in this activity (Merton, 1957). There are two key implications of this observation. First, as explored by Aghion, Dewatrapoint, & Stein (2005), scientists value the autonomy and freedom to control their own research direction and develop project choices that meet their individual interests. Regardless of whether or not the projects that they undertake produce basic or applied knowledge, this remains an important element in their career choices. Second, scientists place considerable value on the ability to disclose the knowledge that they produce through publications. While the precise mechanism by which profession-specific values are internalized is open to debate (e.g., a hallmark of graduate-level scientific training is the emphasis placed on the contributions to scientific research by senior academics), researcher sensitivity to the scientific disclosure and reward system may be important for understanding the economic behavior of scientists. Simply put, scientists may have a "taste" for Science (Merton, 1973; David and Dasgupta, 1994). The researchers reveal their utility for "being a scientist" in their willingness to accept lower wages in return for autonomy in agenda setting. This notion is supported by evidence of a significant wage differential between academia and industry (Murray & Hsi, 2007) and in the lower wages industry researchers will accept when a private-sector position is combined with opportunities to disclose through publication (Stern, 2004). As will be seen, this enriches the comparative statics results that follow.

In contrast to researcher preferences, those who fund research have objectives that are more complex. We assume that the need provide barriers to entry to appropriate the rents from knowledge production are central in shaping the disclosure preferences of private funders. As a result, secrecy or the property rights provided by patenting are preferred alternatives over publication alone. Taken together, the disclosure preferences of researchers and funders diverge, with important consequences for the two-party negotiation. We model the negotiation to predict the conditions under which we expect to observe different regimes of knowledge disclosure. The remainder of this paper considers a simple yet instructive model to identify the key economic and strategic forces under which different disclosure strategies, particularly patent-paper pairs, may arise.

3. Baseline Model

We begin with a baseline model that endogenizes the disclosure strategy that is chosen between a single scientist and a firm and, in particular, the types of regimes that might arise.

Our focus here is on the nature of the bilateral relationship between the scientist and the firm so as to highlight how changes in the external environment not only impact on the direct incentives for disclosure but also on the choice of disclosure strategy.

Regimes

Our starting point is a single scientist who is matched with a single firm. They engage in negotiations over the disclosure outcomes from a project. Disclosure, if it occurs, can be achieved through both patenting and publication. Specifically, the scientist and firm agree to whether a project is patented or not – that is, a choice of $i \in \{0,1\}$ with $i = 1$ denoting a choice of patenting – and how much (if any) of the outcomes of the project are disclosed through publication. Choosing to patent involves a quid pro quo of certain disclosures the level of which is represented by a parameter, d_{PAT}.[2] Disclosure through publication is a choice $d \in [0, D]$, which is negotiated between the firm and scientist. A choice of $d = D$ represents full disclosure of the project's outcomes. When the firm and scientist choose $d = 0$ we term this 'no publication' whereas any agreed $d > 0$ is considered a publication.[3] As will be discussed in more detail below, both d and d_{PAT} will be regarded as probabilities that knowledge is disclosed in a way that is useful to others rather than a description of what is disclosed per se.

Importantly, the knowledge in a patent disclosure may concord with knowledge disclosed through publication. We demonstrate that the degree of *concordance* has important implications for the nature of disclosure regimes that may be agreed upon by a scientist and a

[2] Patenting, of course, involves other costs such as those associated with filing and enforcement that can impact on firms and scientists. We assume here that such costs are zero but not that their inclusion would serve merely to reduce the returns to patenting while not changing any of the results obtained below.

[3] It would be possible to imagine a situation where publication also required a minimum level of disclosure. However, because, as will be argued below, scientists and firms disagree over the degree of publication disclosures but not over patent disclosures, we focus on that as a continuous and free variable. Imposing a lower constraint on the level of publication would have the effect of reducing the parameters for which publication might be observed.

firm.

The combinations of choices between patenting and publication give rise to four possible broad disclosure regimes. First, if there is no patent or publication ($i = d = 0$), what results is a regime of *secrecy*. We use this term because, as will be demonstrated, no disclosure occurs in equilibrium. Second, if there is a patent but no publication ($i = 1,\ d = 0$), there is a *commercial science* regime. Third, if there is a publication but no patent ($i = 0,\ d > 0$), there is a *open science* regime. Finally, if there is both patenting and publication ($i = 1, d > 0$), there is a *patent-paper pairs* regime.

Scientists and Firms

Scientists are motivated by money (in the form of a wage, w) and scientific kudos that results from publication and citation by other scientists.[4] They are also presumed to be liquidity constrained and so cannot agree to $w < 0$.[5] We represent their utility by $U = w + b_S d$ where b_S is the marginal benefit of disclosure in terms of kudos. Scientists receive this utility if the project goes ahead, otherwise they receive 0.[6] Note that this specification presumes that only disclosures through publication generate scientific kudos and patent disclosures do not matter for scientific prestige and rewards. However, in many respects this is for notational simplicity. As we explain below, if there is knowledge that is disclosed through patenting that is also valuable for the generation of scientific kudos, a firm and scientist have a common interest in agreeing to the publication of that knowledge. Hence, without loss in generality, we can

[4] This could also be paid in equity but with no change to the results that follow.
[5] This is a natural assumption given our specification here that focuses on the firm as the provider of capital for the project. If scientists had independent wealth, they could simply choose to provide that capital themselves.
[6] This is a standard normalization that is reasonable in this context where negotiations take place over individual projects rather than say, career research paths. Of course, an interesting issue for future analysis is how the nature of outside options and their terms – disclosure and wage mix – impact on the negotiations between the scientist and a given funder.

represent scientist utility as depending upon publication disclosures only.

Firms provide capital for the scientist's research, *k*, and pay the scientist's wages (if any). Their profits net of these costs depend upon whether they have a monopoly, earning Π, or face imitative competition, in which case they earn π (< ½Π).

Competition occurs if there is entry. The probability that a firm faces entry is a function, $f(d, id_{PAT}) \in (0,1)$ which is non-decreasing in d and id_{PAT} (disclosures assist imitative entry) Thus, a firm's expected payoff from agreeing to fund a scientist's research is: $\Pi - k - w - f(d, id_{PAT})(\Pi - \pi)$. The firm, therefore, has a conflicting preference over the level of disclosures through publication to that of the scientist.

For ease of exposition, in what follows, we adopt a specific model of entry that allows us to generate closed form solutions. We stress, however, that the qualitative predictions in what follows could all be established in a more general specification. The specific model is depicted in Figures 2 and 3 for the cases where a patent does or does not exist. Descriptively, it is as follows. There is room in the market for a single potential entrant who can earn π in gross profits should they enter. There is uncertainty as to the nature of fixed entry costs, θ, faced by the entrant. We assume that those costs are drawn from a uniform distribution on [0,1].

Figure 2: Entry when Patent Exists

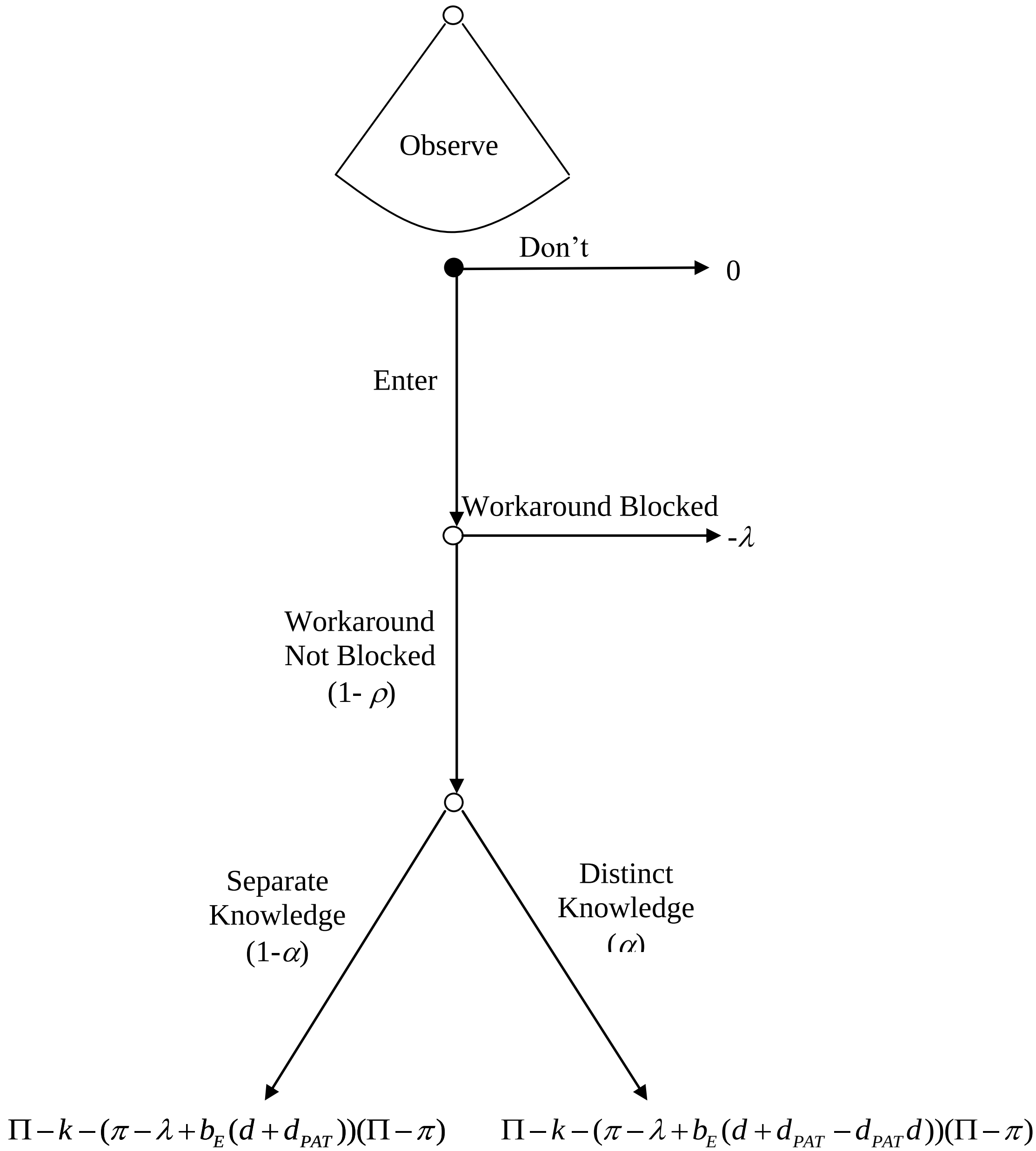
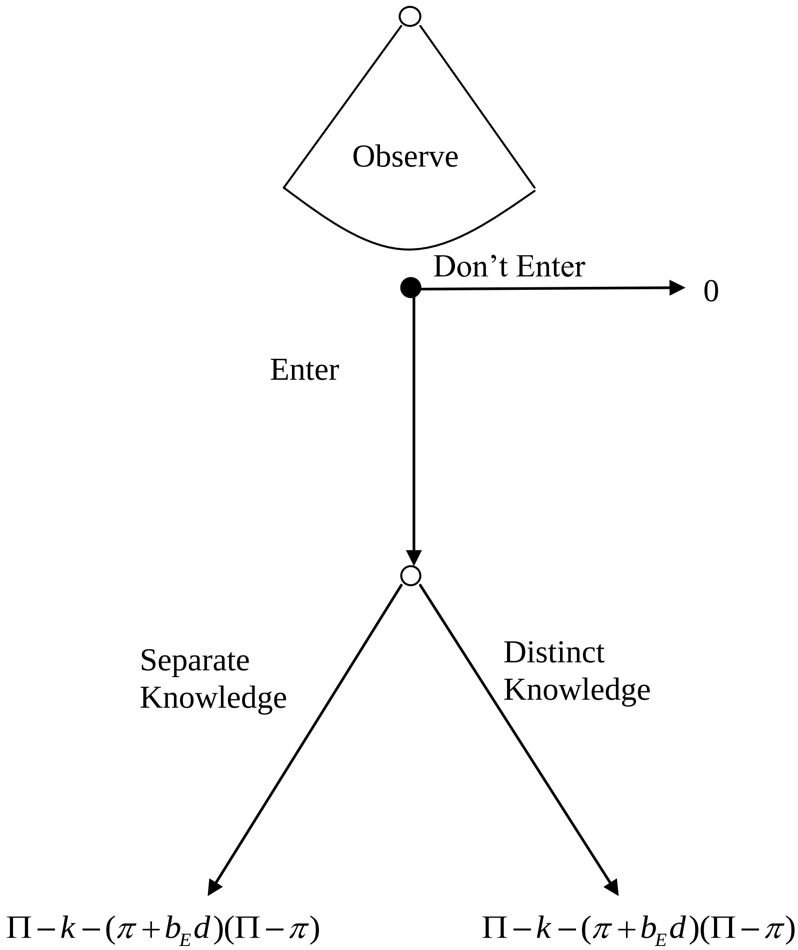


Figure 3: Entry when there is no Patent



In addition, if there is a patent, this impacts on the entry decision: specifically, upon entry costs is the strength of patent protection and the nature of disclosures made. Patent protection is modeled by assuming that at cost, λ, the entrant can develop a work-around any patent that might exist. If there is no patent, these costs are not incurred. However, just sinking costs in a workaround does not guarantee entry. Ex post there is a probability, ρ, that the work-around will be blocked and no entry will be possible even though workaround development costs have been incurred. If $\rho = 1$, then no entry would occur as patent protection would be perfect. With $\rho < 1$, patent protection is imperfect although both ρ and λ govern the strength of patent protection.

If entry is not blocked, the entrant can potentially take advantage of disclosures that have been made through patents and publications. As noted earlier, disclosures through a patent can provide technical information that assists in developing a competing product. We model this by assuming that there is a probability, d_{PAT}, that patent disclosure, on its own, mitigate development costs by b_E. d_{PAT} is literally a measure of disclosures through patent requirements and is a fixed exogenous level in what follows. Similarly, publication offers the opportunity to learn key information. Specifically, an enabling paper can allow an entrant to defray development costs in a similar manner to patent disclosures. We assume that the probability that an entrant learns the information through publication alone is d (where now, by definition, $D \leq 1$) and this can yield the entrant a benefit, b_E.

However, the information learned through publication might be similar or distinct to that information that could be earned through a patent. To consider this possibility, we assume that with probability, α, patent and publication disclosures inform on the same type of information. Consequently, the probability that b_E is saved is $d_{PAT} + d - d_{PAT}d$. On the other hand, with probability, $1 - \alpha$, there are separate pieces of information that could be learned from patent and publication disclosures. In this case, the probability that b_E is saved is $d_{PAT} + d$.[7] In expectation and at the time of the entry decision, the probability of learning information that saves b_E is $d_{PAT} + d - \alpha d_{PAT} d$. Thus, α parameterizes the degree of concordance between patent and publication disclosures in terms of their usefulness in assisting entry.

Using this we can derive the specific probability of entry. Entry will only occur if this

[7] There is a sense in which this outcome involves more opportunities for learning from disclosures than the similar information case. We could adjust for this by choosing the benefit from learning through any one channel as $b_E/2$. However, this would have no impact on qualitative results at the expense of additional complication.

expected profit is positive. Given the uniform distribution of θ, the probability of entry is simply, $(1-i\rho)\left((1-i\rho)\pi - i\lambda + b_E(d + id_{PAT} - \alpha id_{PAT} d)\right)$. We assume that $\pi + b_E < 1$ that ensures that even when entry conditions are most favorable, it still may not occur.[8]

Negotiations

The scientist and firm negotiate over a wage, *w*, and the disclosure strategy (i,d).[9] We assume that the scientist has no financial capital and so the wage must be non-negative. Initially, we restrict attention to the case where it is strictly positive before turning to look at what happens when it may be 0. We use the Nash bargaining solution to describe the outcomes of this negotiation. Specifically, under that solution, variables are chosen to solve the following:

$$\max_{w,i,d} \left(w + b_S d\right)\left(\Pi - k - w - (1-i\rho)\left((1-i\rho)\pi - i\lambda + b_E(d + id_{PAT} - \alpha id_{PAT} d)\right)(\Pi - \pi)\right) \quad (1)$$

Note that we assume that the scientist and firm have equal bargaining power.[10]

Holding the disclosure strategy constant, the wage is given by:

$$\max\left[0, \tfrac{1}{2}\left(\Pi - k - (1-i\rho)\left((1-i\rho)\pi - i\lambda + b_E(d + id_{PAT} - \alpha id_{PAT} d)\right)(\Pi - \pi) - b_S d\right)\right] \quad (2)$$

This wage will be strictly positive if $\Pi - k - (1-i\rho)\left((1-i\rho)\pi - i\lambda + b_E(d + id_{PAT} - \alpha id_{PAT} d)\right)(\Pi - \pi) > b_S d$. In this case, notice that the wage payment is reduced by the fact that the scientist receives kudos if a publication occurs. This is consistent with the findings of Stern (2004) that scientist wages within firms are lower if they are allowed to freely publish their research results. Also, note that if disclosure through publication is sufficiently high, wages will fall to zero. The critical level of disclosure that

[8] Relaxing this would not change any of the qualitative results below but the assumption is designed to restrict our analysis to the cases of most interest.

[9] Below we consider what happens when the level of disclosure from various options is unconstrained by required thresholds and so might itself be a choice variable.

[10] The model could easily be extended to parameterize the degree of bargaining power. However, doing so does not change the qualitative results below.

would result in this is given by:

$$\underline{d} = \frac{\Pi - k - (1 - i\rho)((1 - i\rho)\pi - i\lambda + b_E i d_{PAT})(\Pi - \pi)}{b_S + b_E (1 - i\rho)(1 - \alpha i d_{PAT})(\Pi - \pi)} \tag{3}$$

Expositionally, as well as analytically for some results, it is useful to distinguish between the cases where $\underline{d} > D$ and $\underline{d} \leq D$. The former case will arise for projects where the potential for earning scientific kudos is low relative to commercial returns and consequently, in equilibrium, wage payments are positive. The latter case will arise where the potential for earning scientific kudos is high relative to commercial returns and will correspond to a situation where scientists accept funding but no wages from the firm. In many respects, this latter case corresponds to the type of research projects typically undertaken within universities as opposed to firms (even when commercially funded). Because it is the more straightforward case, we will consider restrict wages to be positive (i.e., assume $\underline{d} > D$) for the remainder of this section as a baseline case before turning to consider what arises when wages can fall to zero and any differing results in Section 5.

Complete Congruence

It is useful at this point to highlight a situation that might be observationally important that we are ruling out in this specification of how disclosures impact upon entry. Suppose that the knowledge disclosed through a patent was always valuable in terms of scientific kudos – say, through a 'marketing paper.' In this case, whenever a scientist and firm agreed to patent their results, there would be an associated publication. Technically, a situation of commercial science would never be observed if this were possible. Similarly, suppose that any information that might be disclosed through a publication would be sufficient to allow a patent to be granted, then open science would never arise. Consequently, to allow for a full range of equilibrium outcomes beyond secrecy and patent-paper pairs, our specification assumes that such *complete congruence* is not possible.

Total Surplus Maximization

When $\underline{d} > D$, wages will be positive in equilibrium. In this case, (*i*, *d*) is chosen to maximize joint surplus:

$$b_S d + \Pi - k - (1-i\rho)\big((1-i\rho)\pi - i\lambda + b_E(d + id_{PAT} - \alpha i d_{PAT} d)\big)(\Pi - \pi) \tag{4}$$

Importantly, this means that a single condition determines whether there will be a publication or not. Let $\Delta_d \equiv \frac{b_S}{b_E(\Pi-\pi)}$ be the ratio of marginal benefits to marginal costs of disclosure through publication when there is no patent. Then, in terms of maximizing total surplus, $d > 0$ if and only if $\Delta_d > (1-i\rho)(1-id_{PAT})$. Publication increases joint surplus when the benefits of disclosure in terms of kudos outweigh the losses associated with potentially greater entry.

A patent impacts an entrant in two ways: it reduces its profits directly but also lowers its entry costs. Thus, a firm will only find it worthwhile to patent if it forestalls rather than encourages entry. Let $\Delta_i \equiv \frac{\Pi-k-(1-\rho)((1-\rho)\pi-\lambda+b_E d_{PAT})(\Pi-\pi)}{\Pi-k-\pi(\Pi-\pi)}$ be the ratio of firm profits (net of publication disclosure costs) when they do and do not obtain a patent and there is no publication. Note that $\Delta_i > 1 \Leftrightarrow b_E d_{PAT} - \lambda < \pi \frac{\rho(2-\rho)}{1-\rho}$; a patent raises firm profits if the losses from disclosure are outweighed by the forestalling effect of the patent's potential excludability. Like publication, when maximizing total surplus, whether a patent is chosen or not comes down to a single condition; namely, $i = 1$ only if $\Delta_i \geq 1$.

Proposition 1 summarizes the resulting equilibrium outcomes:

Proposition 1. *Suppose that* $\underline{d} > D$. *Then the equilibrium outcome is:*

(i) Secrecy $(i = d = 0)$ *if* $\Delta_i < 1$, $\Delta_d < 1$ *and*

$\Delta_i + \left(\Delta_d - (1-\rho)(1-\alpha d_{PAT})\right)\frac{b_E(\Pi-\pi)}{\Pi-k-\pi(\Pi-\pi)} < 1$.

(ii) *Commercial Science* ($i=1,\ d=0$) *if* $\Delta_d < (1-\rho)(1-\alpha d_{PAT})$ *and* $\Delta_i > 1$ *;*

(iii) *Open Science* ($i=0,\ d>0$) *if* $\Delta_d > 1$ *and* $\Delta_i < 1 + \frac{b_E(\Pi-\pi)}{\Pi-k-\pi(\Pi-\pi)}$

(iv) *Patent-Paper Pairs* ($i=1,\ d>0$) *if* $\Delta_i + \left(\Delta_d - (1-\rho)(1-\alpha d_{PAT})\right)\frac{b_E(\Pi-\pi)}{\Pi-k-\pi(\Pi-\pi)} > 1$, $\Delta_d > (1-\rho)(1-\alpha d_{PAT})$ *and* $\Delta_i > 1 + \frac{b_E(\Pi-\pi)}{\Pi-k-\pi(\Pi-\pi)}$

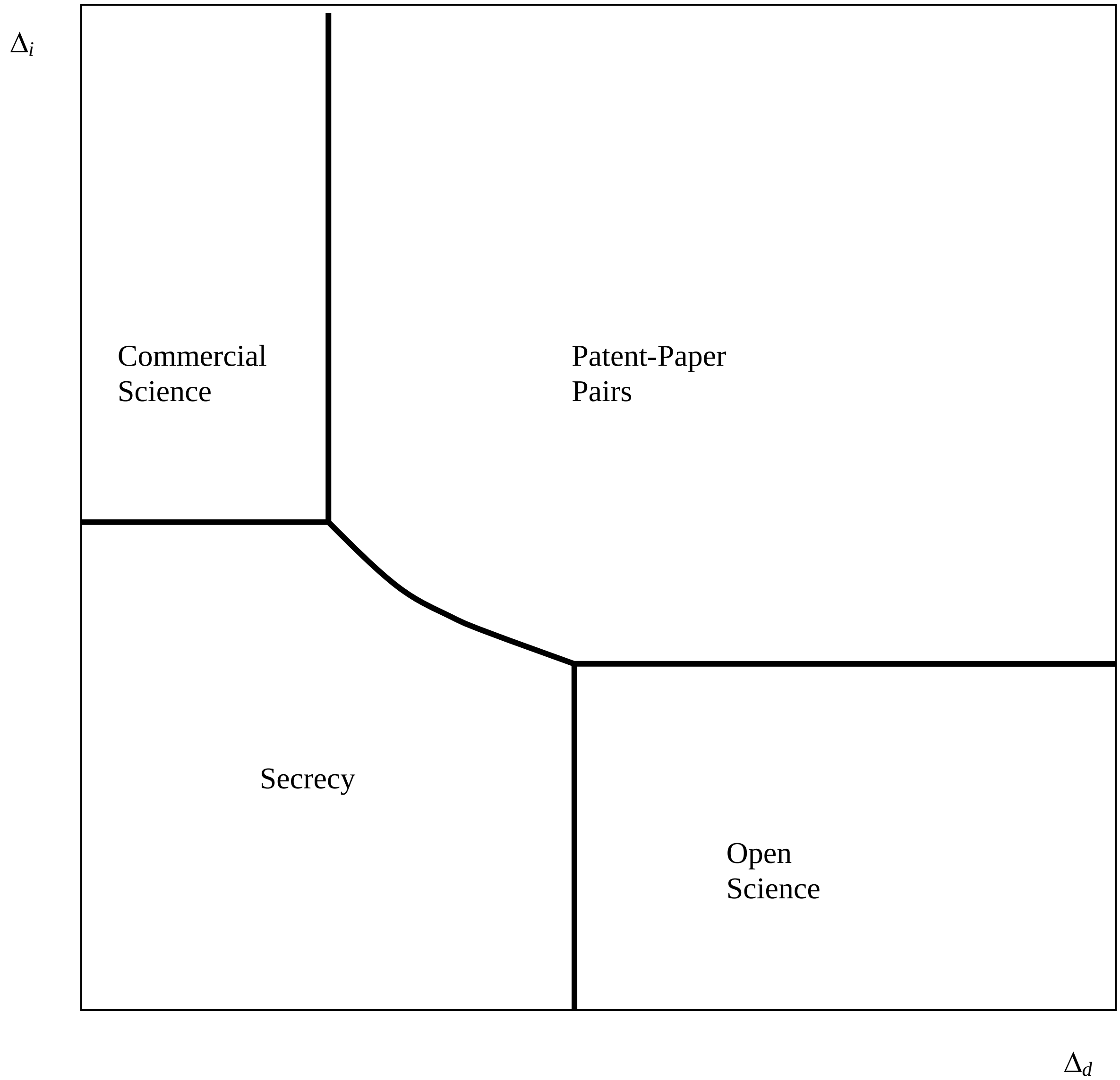
All proofs are in the appendix.

The outcome of the proposition is depicted in Figure 4. There are a number of features to note about the characterization of equilibrium outcomes in Proposition 1. The most important feature is that there is a complementarity between patents and publications. Specifically, the marginal cost to publishing is:

$$b_E(1-i\rho)(1-\alpha i d_{PAT})(\Pi-\pi) \quad (5)$$

Note that this is decreasing in *i*. This occurs for two reasons. First, for $\alpha > 0$, there is some concordance between the knowledge disclosed through patents and that disclosed through publications. Consequently, if knowledge is being disclosed through one channel, it is less likely knowledge that will promote entry will be disclosed through the other, reducing the costs to such disclosure. Second, for $\rho > 0$, patent protection is imperfect but has a chance of blocking entry and hence, the ability of an entrant to take advantage of any disclosures (through publication or patent). Intuitively, the greater the level of patent protection, the less likely are disclosures going to boost the probability of successful entry. Conversely, the marginal return to patenting is increasing in the level of publication. That return is $(\rho\pi + \lambda - b_E d_{PAT}(1-\alpha d))(\Pi-\pi)$. Note that the greater the level of *d*, the lower is the expected marginal impact of patent disclosures. Thus, if the firm decides to obtain a patent, the expected cost from increased disclosures falls from $b_E(\Pi-\pi)$ to $b_E(1-\alpha d)(\Pi-\pi)$ and so the firm and scientist negotiate some level of disclosure.

Figure 4: Baseline Model



The strength of intellectual property protection

It is useful to note that under both open science and patent-paper pairs, the level of disclosure is full disclosure, *D*. This is because in this model the marginal benefit and marginal cost of disclosure are constants. However, if total surplus were concave in publication disclosure, the complementarity between patents and publications will mean that a higher level of disclosure will be negotiated under patent-paper pairs than under open science.

However, we can still utilize our baseline model is equipped to consider the debate regarding whether stronger intellectual property protection would limit openness in science performed by firms. In our baseline model, intellectual property protection is strengthened by increasing λ, increasing ρ or decreasing d_{PAT}.

First, strengthening patent protection will raise the returns to patenting over not patenting (encouraging both commercial science and patent-paper pairs relative to other regimes). In addition raising ρ and reducing d_{PAT} will impact on the publish/not publish margin. Hence, it will encourage patent-paper pairs relative to other regimes; specifically, commercial science.

Of course, reducing d_{PAT}, makes patenting more likely to be chosen (and so leads greater levels of disclosure through that pathway) and also, within patent-paper pairs, allows the scientist to negotiate greater levels of disclosure through publication. Thus, this change in IP protection alters the *type* of disclosure that we observe.

Independence

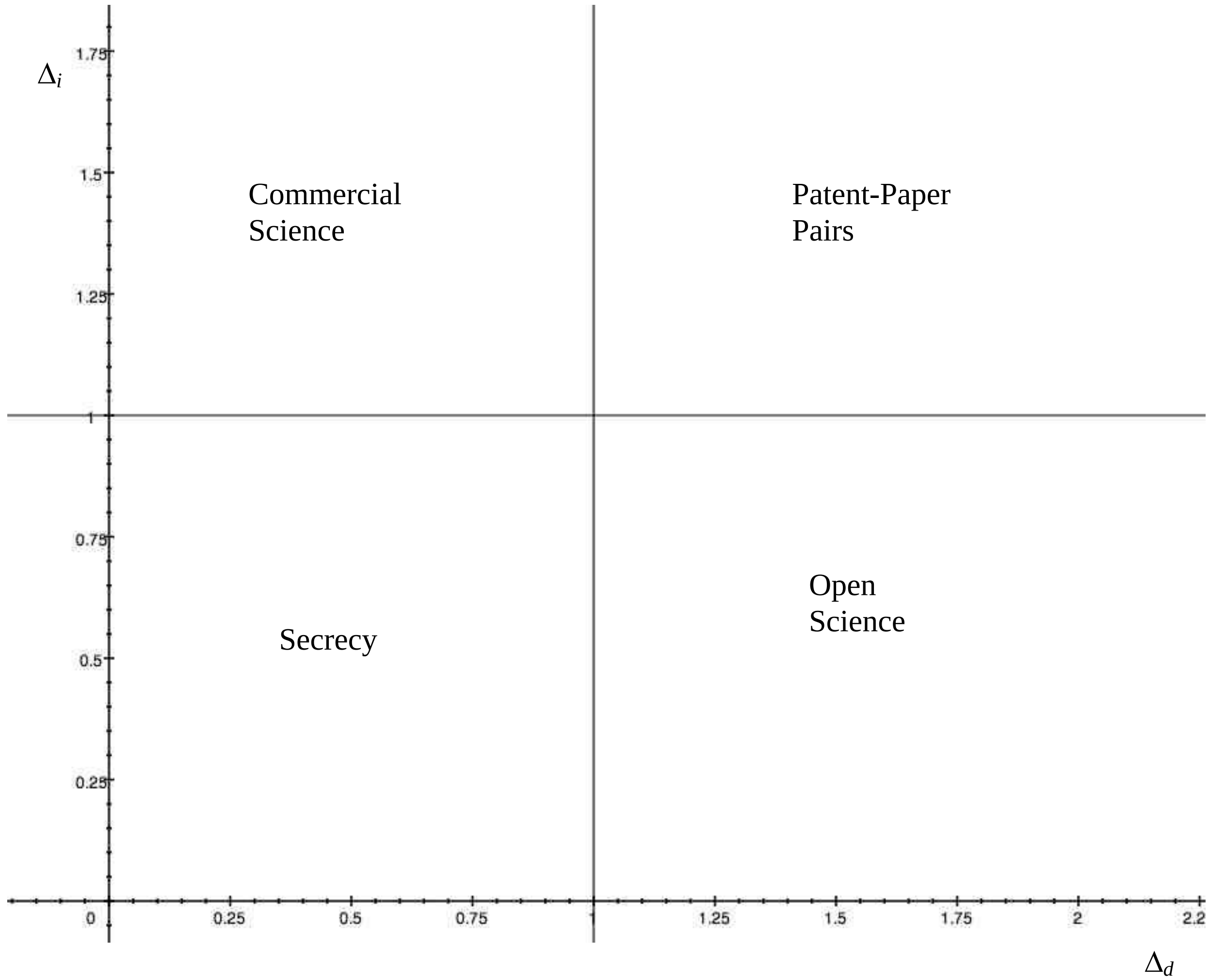
It is instructive to demonstrate the sources of complementarity (α, $\rho > 0$) by considering what happens in a limiting case. For instance, if $\alpha = 0$, there is no concordance in knowledge disclosed through publication and patents but there still exists a complementarity because a patent may block entry completely. However, if $\rho = 0$, a patent may deter but not possibly block entry. In this case, we can demonstrate the following:

Proposition 2. *Suppose that $\alpha = \rho = 0$. The equilibrium outcome is:*

- *(i) Secrecy ($i = d = 0$) if $\Delta_i < 1$ and $\Delta_d < 1$;*
- *(ii) Commercial Science ($i = 1,\ d = 0$) if $\Delta_i > 1$ and $\Delta_d < 1$;*
- *(iii) Open Science ($i = 0,\ d > 0$) if $\Delta_i < 1$ and $\Delta_d > 1$;*
- *(iv) Patent-Paper Pairs ($i = 1,\ d > 0$) if $\Delta_i > 1$ and $\Delta_d > 1$.*

This outcome is depicted in Figure 5. In what follows, we may consider this independent case (where $\alpha = \rho = 0$) in order to highlight alternative factors that drive the interaction between patents and publications.

Figure 5: Independence



What this demonstrates is that as α increases, patent-paper pairs become more likely. Consequently, this appears to indicate that if patent disclosures were required to be more like publication disclosures, more disclosure might result. Of course, in actuality, this change would likely cause pairs to negotiate regimes other than commercial and open science rather than more disclosure per se.

Allowing for Zero Wages

We now turn to consider the case where $\underline{d} \leq D$, where the potential value of kudos is high relative to commercial returns. While the research is commercially funded, and the firm can earn profits, it is possible that they might negotiate a level of disclosures through publication such that wages hit their constraint of zero. In this situation, we cannot simply apply an intuition based on maximizing total surplus – something that only occurs if wages are positive.

However, the same basic drivers of the decisions to patent and publish (namely, Δ_d and Δ_i) continue to play a significant and similar qualitative role when $\underline{d} \leq D$. The outcome is summarized in Proposition 3.

Proposition 3. *Suppose that* $\underline{d} \leq D$*. Then the equilibrium outcome is:*

(i) *Secrecy (* $i = d = 0$ *) if* $\Delta_i < 1$, $\Delta_d < 1$ *and* $\Delta_d \Delta_i^2 < (1-\rho)(1-\alpha d_{PAT})$

(ii) *Commercial Science (* $i = 1,\ d = 0$ *) if* $\Delta_i > 1$ *and* $\Delta_d < (1-\rho)(1-\alpha d_{PAT})$;

(iii) *Open Science (* $i = 0,\ d > 0$ *) if* $\Delta_d > 1$ *and* $\Delta_i^2 < (1-\rho)(1-\alpha d_{PAT})$

(iv) *Patent-Paper Pairs (* $i = 1,\ d > 0$ *) if* $\Delta_d \Delta_i^2 > (1-\rho)(1-\alpha d_{PAT})$, $\Delta_d > (1-\rho)(1-\alpha d_{PAT})$ *and* $\Delta_i^2 > (1-\rho)(1-\alpha d_{PAT})$

The proof of the proposition shows that even if, under publication, disclosure is so high that it causes wages to hit their non-negative constraint, the choice between open science and patent-paper pairs is driven by whether patents deter or promote entry. In addition, the two drivers of complementarity between patents and publications are still present in this case. And, for the special case where $\alpha = \rho = 0$, the conditions that drive each regime are identical to those when wages are restricted to be positive.

There are, however, a number of significant differences between the baseline case and that in this section. First, when there is publication, wages are zero whereas under secrecy and commercial science, wages are positive (Lemma 1). This simplifies the analysis considerably. It says that if greater publications maximize total surplus, the scientist and the firm will choose that as a means of compensating scientists rather than the use of wages which are a simple transfer between them. So long as there is the potential for disclosures to serve this function, that is, $\underline{d} < D$, zero wages will be negotiated.

Second, when there is publication, the level of disclosure negotiated is:

$$d_i^* = \frac{\Pi - k - (1 - i\rho)((1 - i\rho)\pi - i\lambda + b_E i d_{PAT})(\Pi - \pi)}{2b_E(1 - i\rho)(1 - \alpha i d_{PAT})(\Pi - \pi)} \qquad (6)$$

Notice that, this is independent of the marginal return to disclosure from scientific kudos (b_S) but positively related to firm profits under secrecy and commercial science. It also interacts with the patenting choice of the firm. If patenting is commercially desirable, the level of disclosure is higher under patenting. If it is not commercially desirable, the level of disclosure is lower under patenting. Thus, the firm and scientist share a common interest in the patenting decision.

Third, this highlights an additional complementarity between patents, and in general, commercial returns and publication that does not exist in the baseline case. If a patent is granted, a stronger patent increases commercial returns. As noted earlier, as wages are zero when there is a publication, the only way this additional surplus can be transferred to the scientist is by allowing more disclosure. Similarly, any strengthening of patent protection, will through all three paths, not only make patent-paper pairs more desirable as a regime but also increase the level of disclosure within that regime. Thus, in contrast to concerns that changes (such as the Bayh-Dole Act) would diminish levels of openness and publications, there are stronger incentives to generate published disclosures for commercially funded projects with high potential kudos as as patent protection becomes stronger.

It is instructive to note, however, that the greater level of disclosure through publication does not offset the reduction of disclosure through patenting (through a reduction in d_{PAT}). A 'unit' reduction in knowledge disclosed through patenting, from (6), only results in a half 'unit' increase in knowledge disclosed through publication.

In interpreting the distinction between the cases where wages with publication may or may not be zero, it is tempting to characterize the positive wage case as being performed by scientists employed within firms while the zero wage case involves research performed in universities. Indeed, it is the case that it is only in universities (and similar organizations) where you see scientists being paid from other means and not for outcomes directly related to a particular project. However, it is equally possible that, for low kudos (or high commercial return) projects, university scientists will be directly paid for research (usually termed consulting) even when they have publication rights. Nonetheless, there is a sense in which an extension of this model – say to include the role of publication as a means of monitoring research quality [REFS HERE] – would lead to a predictive distinction as to the employment choices of scientists. That possibility is left for future research.

4. Licensing and Landscape Papers

In the baseline model, we focused on publications as 'enabling papers' that were more likely to contain knowledge that would assist current competitors and impact on firm profits in the short-term. However, many papers do not fall into this class and are better assessed as 'landscape' papers. These papers involve the discovery and dissemination of more fundamental knowledge. In particular, that knowledge, while resulting in some immediate kudos for scientists, can also generate kudos as it is built upon and cited by future researchers.

From the perspective of commercial funders, however, landscape papers are not as risky in terms of fostering immediate entry. Instead, if an innovation is patented, they can earn future revenues from licensing the rights to utilise that patent in future research. However, it may be that publication of a landscape paper may allow future scientists and their funders to avoid having to pay license fees. Moreover, their incentive to avoid such payments might themselves lead to those papers not being cited by future scientists, thereby causing a loss in potential kudos. In the end, if future researchers are impeded by the need to license prior intellectual property from commercial funders, this might cause them to divert their research agendas away from paths that would lead to them citing the scientist's publications. This will generate a new source of conflict between scientists and their funders regarding current disclosure through publication.

This possibility has received attention recently amongst scientists. Concerns have been raised that patenting of scientific knowledge gives rise to an *anti-commons effect* (Heller & Eisenberg, 1998). This occurs when future scientists fear that patent protection will generate a thicket of licenses, permissions and other transaction costs and so avoid researching in areas or building on research where formal intellectual property exits. Murray and Stern (2007) demonstrated that when a paper becomes associated with a patent, its citations drop significantly; potentially, an indicator that scientists are avoiding research areas with potential intellectual property issues. Moreover, there is some concern that, either as a result of these measurable effects on future citations, or as a result of the decrease in kudos, scientists may face a reputation discount if they patent as well as publish their research.

In this section, we construct a model designed to evaluate this concern. In contrast to the baseline model, firms and scientist are assumed to live for two periods and in each time period a new firm/scientist pair is born. In the first period, they negotiate over various terms, research is conducted and immediate outcomes (profits, patents and publications) are realized. This corresponds to the static model already examined. In the second period, the next generation of scientist and firm do the same thing. Their choice as to whether there is a publication or not determines the kudos achieved by the previous scientific generation with that kudos only emerging if there is a citation to that previous work. However, in this situation, if there exists a patent in the previous generation, the firm and scientist may be required to pay a license fee to the previous generation funder. In effect, negotiations are now three-party; between a firm, a scientist and the previous patent rights holder. Thus, our static model is turned into an overlapping generations, dynamic model.

To this end, we make several substantive changes to the static model. First, for simplicity we assume that $\alpha = \rho = 0$ so as to eliminate static drivers of interactions between patents and papers. Second, we assume that there is an additional level of kudos from publication, B_S, that is only realised if the scientist publishes and the publication is cited by the next generation. We assume that firms and scientists share a common discount factor, δ. Thus, a scientist earns $(b_S + \delta B_S)d$ if they have a publication with disclosure d that is cited.

Third, if a patent is granted, then the patent holder always negotiates with the next generation research team. If those negotiations succeed, there is a transfer of knowledge between generations meaning that no capital costs of research are incurred and a lump sum license fee, τ_t, is paid. If those negotiations fail, then with probability γ, the new generation research team must choose a different research path. Thus, γ is the probability that a patent blocks future research based on the past knowledge. There is no knowledge transfer (they incur the full capital costs, k). On the other hand, with probability $1-\gamma$, the new generation at time t can, if there is a previous publication with disclosure level, d_t, exploit the knowledge from that publication in another way and incur expected capital costs, $(1-d_t)k$. In either of these cases, there is no citation to the previous generation as the next generation research team avoids citation as part of its strategy of working-around the previous patent rights. In this respect, γ corresponds to a measure of the strength of future IP protection.[11]

Fourth, if no patent is granted, then the previous funder is not involved in future negotiations. If there is no previous publication, there is no transfer of knowledge and the future team incurs the full capital costs, *k*. On the other hand, if there is a previous publication, those capital costs are eliminated and the future scientist cites the past one who can grants them kudos.[12]

To model the three-party bargaining game, should it emerge, we utilise a multi-lateral variant of the Nash bargaining solution. That solution is found by solving the following:

$$\max_{\tau_t, w_t, i_t, d_t} \tau_t (w_t + b_S d_t + I_{d_{t+1}>0} \delta B_S d_t - u_t)(V_t - v_t) \quad (7)$$

where the subscript, *t*, in the choice variables corresponds to the disclosure strategy and payments of generation *t*, *I* is an indicator function taking a value of 1 if a publication occurs in *t*+1 and 0 otherwise, u_t is the scientist's expected utility and v_t is the firm's expected payoff should no licensing agreement be reached with the firm in period *t*-1. Those latter outside options are determined as the result of a bilateral negotiation between the scientist and firm of generation *t* in the same manner as those in the baseline model. We will assume that *D* is sufficiently low that wages in equilibrium are always positive. Hence, we can focus on

---

[11] An alternative way of modeling the imperfection of future IP rights would be to assume that the uncertainty regarding this was released prior to any negotiations. This alternative timing is more complex than the one we have chosen and does not appear to result in any significant alternative qualitative conclusions.

[12] It would be possible to model the savings on capital cost being a function of the level of disclosure. This, however, would not change the model's predictions and so we opt for a simpler, equivalence between unimpeded transfer through publication and a transfer alongside a patent.

negotiation outcomes that maximize total surplus.[13]

We focus on symmetric dynamic equilibria whereby each generation chooses the same disclosure strategy.[14] Consequently, it is useful to discuss what this means for each regime in turn before formally characterizing the symmetric equilibria.

Secrecy. Suppose that each generation chooses $i=0$ and $d=0$. Then research capital costs of k are realised in every period and each scientist and firm generation earn their static outcomes; each has a payoff of $\frac{1}{2}\left(\Pi-k-\pi(\Pi-\pi)\right)$ for the first period only. Note that there is no additional incentive for a single generation to deviate and publish beyond the drivers in the static case (Δ_d) as this does not earn the scientist any kudos because, as we will show formally below, the future generation has no additional incentive to publish as a result of this.

On the other hand, patenting allows the firm to obtain license fees in the future. As there is no publication, by using the license and corresponding knowledge transfer, the future generation saves capital costs, k. In this case, they solve:

$$\max_{\tau_t, w_t, d_t, i_t} \tau_t (w_t - u_t)\left(\Pi - (\pi - i_t\lambda + i_t b_E d_{PAT})(\Pi-\pi) + i_t\delta\tau_{t+1} - w_t - \tau_t - v_t\right) \qquad (8)$$

Of course, in a symmetric equilibrium, $u_t = v_t = \frac{1}{2}\left(\Pi-k-\pi(\Pi-\pi)\right)$. Consequently, assuming no other changes (that is, $d_t = i_t = 0$), the license fee, τ_t, equals $k/3$. In this case, by choosing to patent, current expected surplus changes by $(\lambda - b_E d_{PAT})(\Pi-\pi)+\delta\frac{1}{3}k$. As we did with the baseline model, this incentive is naturally expressed using the ratio of profits with and without a

[13] As with the static model, considering the case where D is relatively high would do little to change the qualitative drivers and sources of interaction between publications and patents and where there is a difference, it is of the form already discussed in the static case.

[14] There may be asymmetric equilibria whereby firms and scientists alternate in their disclosure strategies from generation to generation or, alternatively, chose their disclosure strategy contingent upon disclosure strategy of the immediate past generation.

patent; that is, $\Delta_i \equiv \frac{\Pi+\delta\frac{1}{3}k-(\pi-\lambda+b_E d_{PAT})(\Pi-\pi)}{\Pi-\pi(\Pi-\pi)}$.[15] If this is less than 1, secrecy will be preferred to commercial science. Note that the potential to earn future license fees increases the returns to patenting.

Commercial Science. Like secrecy, commercial science involves no publication and so if future scientists are not expected to publish, there is no additional incentive for the current generation to publish and earn kudos. However, in contrast to the static model, a license fee is earned but also paid. Consequently, the scientist and firm each expect a payoff of, $\frac{1}{2}\left(\Pi-\frac{1}{3}k-(\pi-\lambda+b_E d_{PAT})(\Pi-\pi)+\delta\frac{1}{3}k\right)$. Thus, if $\Delta_i \geq 1$, the firm and scientist will have an on-going incentive to continue patenting.

Open Science. Like commercial science, open science results in no research capital costs, *k*, being incurred by the next generation. In addition, the scientist receives kudos as they publish and are cited. If they agreed to a publication disclosure, *d*, forgoing this would reduce total surplus by $(b_S+\delta B_S)d$. Notice that the returns to publication are lower than the static case but this is an artifact of the assumed delay in kudos.

What about the incentives to deviate and patent? This would allow the pair to earn, $(\lambda-b_E d_{PAT})(\Pi-\pi)$ immediately and also to earn a future license fee. Because of the publication that fee would be $\tau=\frac{1}{3}k(1-d(1-\rho))$. Thus, compared with a move from secrecy to commercial science, a move from open science to patent paper pairs involves a lower license fee and hence, will not occur unless $\frac{1}{3}kd(1-\rho)<\Delta_i$. This is in contrast to the static case where the returns to patenting were unambiguously higher when a publication existed.

Patent-Paper Pairs. In patent-paper pairs, negotiations are determined by:

[15] While this parameter has the same interpretation in the baseline model, it is slightly different is that there is a transfer of knowledge across generations when patent licenses are negotiated.

$$\max_{\tau_t, w_t, d_t, i_t} \tau_t (w_t + (b_S + \delta B_S) d_t - u_t)\left(\Pi - (\pi - i_t\lambda + i_t b_E (d_{PAT} + d_t))(\Pi - \pi) + i_t \delta \tau_{t+1} - w_t - \tau_t - v_t\right) \quad (9)$$

What is interesting here is that the scientist and firm outside options are influenced by the publication disclosures of the previous generation. The greater these are, the lower are the gains from trade from reaching a licensing agreement over the use of past IP.

Note that in the dynamic game,

$$\underline{d} = \frac{\Pi - (\pi - i\lambda + ib_E d_{PAT})(\Pi - \pi) + i\,\tfrac{1}{3}\delta k}{b_S + \delta B_S + b_E(\Pi - \pi) + i\,\tfrac{1}{3}\delta k(1-\gamma)} \quad (10)$$

which we assume to be greater than *D*.[16] Suppose that $\Delta_d^f \equiv \frac{B_S}{b_E(\Pi-\pi)}$. The following proposition fully characterizes the symmetric equilibrium outcomes for the dynamic game.

Proposition 4. *Assume that* $D < \underline{d}$ *. The following represent symmetric equilibria in the dynamic game:*

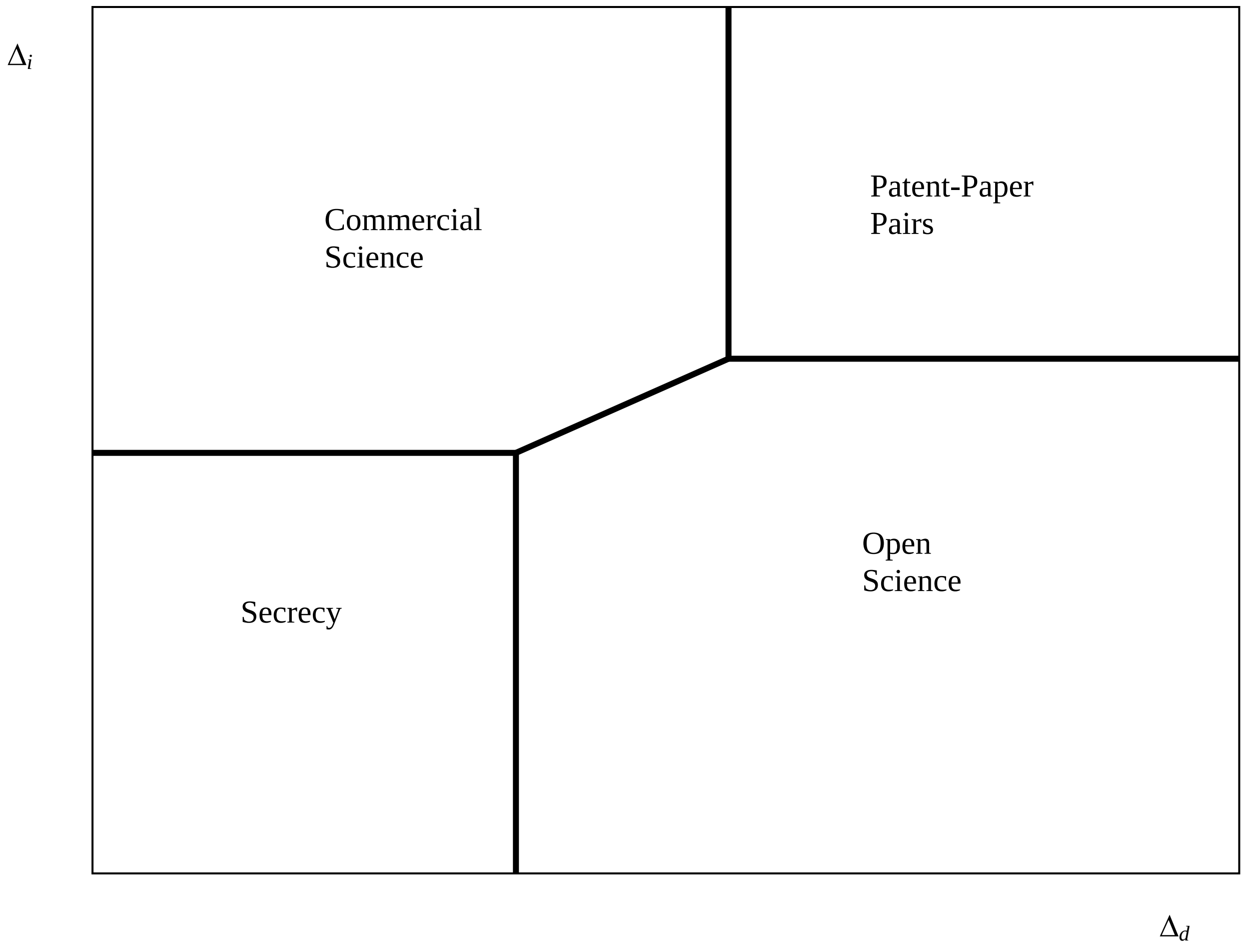
- *(i)* *Secrecy (* $i = d = 0$ *) if* $\Delta_d \leq 1,\ \Delta_i \leq 1$;
- *(ii)* *Commercial Science (* $i = 1,\ d = 0$ *) if* $\Delta_d \leq 1,\ \Delta_i \geq 1$;
- *(iii)* *Open Science (* $i = 0,\ d > 0$ *) if* $\Delta_d + \delta\Delta_d^f \geq 1$, $\Delta_i \leq 1 + \delta\,\tfrac{1}{3}\frac{kD(1-\gamma)}{\Pi - \pi(\Pi-\pi)}$ *and* $\frac{\Delta_d + \delta\Delta_d^f - 1}{\Pi - \pi(\Pi-\pi)} D \geq \frac{\Delta_i - 1}{b_E(\Pi-\pi)}$;
- *(iv)* *Patent-Paper Pairs (* $i = 1,\ d > 0$ *) if* $\Delta_i \geq 1 + \delta\,\tfrac{1}{3}\frac{kD(1-\gamma)}{\Pi - \pi(\Pi-\pi)}$ *and* $\Delta_d + \delta\Delta_d^f \geq \delta\,\tfrac{1}{3}\frac{k(1-\gamma)}{b_E(\Pi-\pi)}$.

The proof is in the appendix. An example of the equilibrium outcomes are depicted in Figure 6. Interestingly, there can be multiple equilibria. This emerges because there is an inter-generational complementarity in publication decisions. Specifically, publication today is only valuable if there is publication and citation tomorrow. Consequently, when $\Delta_d + \delta\Delta_d^f \geq 1 \geq \Delta_d$, there always exists a 'non-communication' equilibrium involving *d* = 0 as the future expectation of this, creates no additional incentives for publication today.

---

[16] Akin to the baseline model.

Figure 6: License Fees and Publication



Critically, there is now substitutability between patents and publications. This is because the future licensing revenues from patenting are reduced if there is publication. This means that the domains of open and commercial science are wider than in the static case.

The strength of IP protection

In contrast to the previous static models, the dynamic model with licensing appears to suggest that the IP protection does indeed impede publication as it might make these substitutes rather than complements.

However, in many respects, this appearance belies the true impact of IP protection. Firstly, the domain of open science is expanded as a result of the substitutability between patents and papers. Consequently, scientists and firms opt for open science that involves a higher level of disclosure than patent-paper pairs as the returns for securing intellectual property protection are too low. Thus, licensing expands rather than contracts the domain for open science.

That said, stronger intellectual property protection does increase the level of disclosure within the regime of patent-paper pairs. This happens for reasons similar to those identified earlier, namely, that (i) commercial returns rise and so negotiated disclosure rises so as to transfer some of those returns to scientists and (ii) as the cost of publication was a reduction in license fees as a result of imperfect IP protection, strengthening this reduces those costs and so increases the incentive to disclose more through publication. Indeed, as γ increases, this increases the returns to patenting but also the negotiated disclosure level under patent-paper pairs and so a greater range of parameter values will support this as an equilibrium at the expense of both commercial and open science. Indeed, in the limit as γ approaches 1, the substitutability between papers and patents disappears and the regimes are divided in a manner akin to the baseline case.

Ownership of IP rights

In the above discussion, firms are assumed to be the owners of patents and hence, are the only party that participates in future negotiations over license fees. This means that firms have no interest in the disclosure decisions of future generations and, indeed, given their hold-up power, only care about future capital costs.

But what if scientists, instead of firms, owned IP rights? In this respect, the impact on immediate entry is unchanged but in licensing negotiations, the scientist, if they had a publication, would care about whether the next generation published or not. More importantly, they would be concerned that, in the event of a breakdown, they would not be cited and would lose their kudos value.

What this means is that, in patent-paper pairs, where this is relevant, the scientist would discount the license fee to take into account the ability of the next generation to hold up their kudos value. Indeed, that license fee would become: $\tau_{t+1} = \frac{1}{3}\left(k(1-d_t(1-\rho)) - B_S d_t(1-\rho)\right)$. In this situation, total surplus becomes:

$$\begin{aligned}&(b_S + \delta B_S)d_t + \delta \tfrac{1}{3}\left(k(1-d_t(1-\rho)) - B_S d_t(1-\rho)\right)\\&+\Pi - (\pi - i_t\lambda + i_t b_E(d_{PAT} + d_t))(\Pi - \pi) - \tfrac{1}{3}\left(k(1-d_{t-1}(1-\rho)) - B_S d_{t-1}(1-\rho)\right)\end{aligned} \tag{11}$$

This means that the pair jointly appropriate future kudos of $\delta B_S d_t\left(1-\frac{1}{3}(1-\rho)\right)$. Thus, the threshold for publication to be chosen is higher and consequently, the domains for open science and patent-paper pairs would shrink relative to the case where only the firm owned patent rights. Intuitively, by bringing future kudos to the table as a valuable item, scientist ownership rights means that some of that value will be seeded to future generations. As licensing agreements would be accepted anyway, this shift in value is a mere transfer. However, the lower appropriation of kudos biases against agreements that involve publication.

Pubs reduce future capital costs without licensing. What if reduced future capital costs with licensing too. Pre-emptive or exploratory work makes next research project using licensed tech more likely.

5. Conclusions

To be done.

6. Appendix

Proof of Proposition 1

Total surplus from each regime is:

- Secrecy: $\Pi - k - \pi(\Pi - \pi)$
- Commercial Science: $\Pi - k - (1-\rho)\left((1-\rho)\pi - \lambda + b_E d_{PAT}\right)(\Pi - \pi)$
- Open Science: $b_S D + \Pi - k - \left(\pi + b_E D\right)(\Pi - \pi)$
- Patent-Paper Pairs: $b_S D + \Pi - k - (1-\rho)\left((1-\rho)\pi - \lambda + b_E(D + d_{PAT} - \alpha d_{PAT} D)\right)(\Pi - \pi)$

Comparing these with one another gives the conditions in the proposition. However, note that when comparing commercial science and open science, commercial science will be preferred if: $D(\Delta_d - (1-\rho)(1-\alpha d_{PAT}))\frac{b_E(\Pi-\pi)}{\Pi-k-\pi(\Pi-\pi)} < \Delta_i - 1$. Note that the other two conditions for commercial science are sufficient to guarantee that this will hold. Hence, a condition comparing these two is redundant.

Proof of Proposition 3

We begin by proving Lemma 1.

> ***Lemma 1.*** *Suppose that full disclosure maximizes total surplus (ie., $b_S \geq b_E(1-i\rho)(1-\alpha i d_{PAT})(\Pi - \pi)$. Then negotiated wages equal zero.*

PROOF: Suppose that *w* = 0. Then the choice of *d* that solves (1) with wages set to zero involves:

$$\begin{aligned}&\Pi - k - (1-i\rho)\left((1-i\rho)\pi - i\lambda + b_E(d + id_{PAT} - \alpha i d_{PAT} d)\right)(\Pi - \pi)\\&= db_E(1-i\rho)(1-\alpha i d_{PAT})(\Pi - \pi)\end{aligned} \quad (12)$$

which yields:

$$d^* = \frac{\Pi - k - (1-i\rho)((1-i\rho)\pi - i\lambda + b_E i d_{PAT})(\Pi - \pi)}{2b_E(1-i\rho)(1-\alpha i d_{PAT})(\Pi - \pi)} \quad (13)$$

It is easy to see that $d^* \geq \underline{d}$ if and only if $b_S \geq b_E(1-i\rho)(1-\alpha i d_{PAT})(\Pi - \pi)$.

Given Lemma 1, we can write the various Nash bargaining objectives as follows (assuming $d^* < 1$):

- Secrecy: $\frac{1}{4}\left(\Pi - k - \pi(\Pi - \pi)\right)^2$
- Commercial Science: $\frac{1}{4}\left(\Pi - k - (1-\rho)((1-\rho)\pi - \lambda + b_E d_{PAT})(\Pi - \pi)\right)^2$

- Open Science: $\frac{b_S}{4b_E(\Pi-\pi)}\left(\Pi-k-\pi(\Pi-\pi)\right)^2$
- Patent-Paper Pairs: $\frac{b_S}{4b_E(1-\rho)(1-\alpha d_{PAT})(\Pi-\pi)}\left(\Pi-k-(1-\rho)((1-\rho)\pi-\lambda+b_E d_{PAT})(\Pi-\pi)\right)^2$

Comparing these yields the conditions in the proposition. Note that for commercial science, the condition for it to be preferred to open science is: $\Delta_i^2 > \Delta_d$. However, if it is preferred to Patent-Paper Pairs and Secrecy then this condition is automatically satisfied. A similar logic holds when evaluating open science.

Proof of Proposition 4

First, consider the conditions for secrecy. If a scientist/firm pair deviated and patented their research, this would result in a change in surplus of $(\lambda-b_E d_{PAT})(\Pi-\pi)+\delta\frac{1}{3}k$ so long as the next generation (a) agreed to this license fee and (b) did not themselves choose to patent or publish. Note that, in a secrecy equilibrium, $u_t = v_t = \frac{1}{2}\left(\Pi-k-\pi(\Pi-\pi)\right)$ and $\tau_{t+1}=\frac{1}{3}k$ Substituting this into (8) yields:

$$\max_{\tau_t,w_t,d_t,i_t} \tau_t\left(w_t-\tfrac{1}{2}\left(\Pi-k-\pi(\Pi-\pi)\right)\right)\left(\tfrac{1}{2}(\Pi-k)-(\tfrac{1}{2}\pi-i_t\lambda+i_t b_E d_{PAT})(\Pi-\pi)+i_t\delta\tfrac{1}{3}k-w_t-\tau_t\right) \quad (14)$$

As τ_t and w_t are positive, the choice of i_t and d_t is determined by total surplus. Under the conditions for secrecy, increasing either of these reduces total surplus and hence, even if a licensing agreement is signed upon deviation, these choices will not change. Similarly, deviating to publication will reduce total surplus as it assists entry, does not result in kudos (as there is no future publication) and reduces future research capital costs (something not captured by the current generation). Finally, deviating to both patents and publications, increases u_t and v_t but does not otherwise change the drivers of licensing. Hence, it results in a lower return than a move to commercial science alone and is not profitable.

Second, consider the conditions for commercial science. A deviation to secrecy will not change the incentives for the next generation to engage in commercial science (although it will change their costs) and so the condition supporting commercial science is the mirror of the condition supporting secrecy. A deviation to patent-paper pairs would reduce commercial returns – both immediately and in the future – and not offer the benefit of additional scientific kudos as it would still maximize total surplus of future pairs not to publish. This is also true for a move to open science. Hence, the existence of an equilibrium with commercial science is solely driven by whether $\Delta_i \geq 1$ and $\Delta_d \leq 1$.

Third, open science results in disclosure equal to D with scientists and firms earning an expected payoff of $\frac{1}{2}\left((b_S+\delta b_S)D+\Pi-(\pi+b_E D)(\Pi-\pi)\right)$. Secrecy would reduce total surplus by $\left(b_S+\delta B_S-b_E(\Pi-\pi)\right)D$ which is positive if $\Delta_d+\delta\Delta_d^f \geq 1$ where $\Delta_d^f \equiv \frac{B_S}{b_E(\Pi-\pi)}$. In contrast, obtaining a patent increases total surplus by $(\lambda-b_E d_{PAT})(\Pi-\pi)+\delta\frac{1}{3}k(1-D(1-\gamma))$. This is negative if: $\Delta_i \leq \delta\frac{1}{3}\frac{kD(1-\gamma)}{\Pi-\pi(\Pi-\pi)}$. Finally, by reverting to commercial science, total surplus falls by

$\left(b_S + \delta B_S - b_E(\Pi - \pi)\right) D - (\lambda - b_E d_{PAT})(\Pi - \pi) - \delta \tfrac{1}{3} k$ or $\frac{\Delta_d + \delta\Delta_d^f - 1}{\Pi - \pi(\Pi - \pi)} D \geq \frac{\Delta_i - 1}{b_E(\Pi - \pi)}$.

Finally, patent-paper pairs results in disclosure equal to *D*. This will be preferred to open science if $\Delta_i \geq \delta \frac{1}{3} \frac{kD(1-\gamma)}{\Pi - \pi(\Pi - \pi)}$ and to commercial science if $b_S + \delta b_S \geq \delta \frac{1}{3} k(1-\gamma) \Rightarrow \Delta_d + \delta\Delta_d^f \geq \delta \frac{1}{3} \frac{k(1-\gamma)}{b_E(\Pi - \pi)}$. Compared to secrecy, patent-paper pairs results in higher total surplus if: $\frac{\Delta_d + \delta\Delta_d^f - 1}{\Pi - \pi(\Pi - \pi)} D + \frac{\Delta_i - 1}{b_E(\Pi - \pi)} \geq \delta k D \frac{1-\gamma}{3 b_E(\Pi - \pi)(\Pi - \pi(\Pi - \pi))}$. Note, however, that if the other two conditions hold, then this holds as well and so is redundant.

7. References

Arrow, K. 1962. Economic Welfare and the Allocation of Resources for Invention. R. Nelson, ed. *The Rate and Direction of Inventive Activity*. Princeton University Press, Princeton, NJ, 609-25.

Aghion, P., Dewatripont, M., & Stein, J. 2005. Academic freedom, private-sector focus and the process of innovation**.** Working paper no. 11542, *National Bureau of Economic Research*, Cambridge, MA.

Azoulay, P., W. Ding, T. E. Stuart. 2007. The determinants of faculty patenting behavior: Demographics or opportunities. *J. of Economic Behavior and Organ.* (forthcoming).

Bar-Gill, O., G. Parchomovsky. 2003. The Value of Giving Away Secrets. Harvard Law and Econ Disc Paper 417; Harvard Public Law Working Paper 64.

Blumenthal, D. 1997. Withholding research results in academic life science: Evidence from a national survey of faculty. *Journal of the American Medical Association***,** 277: 1224-8.

Bush, V. 1945. *Science: The endless frontier***.** Washington, D.C.: United States Government Printing Office.

Sells SF, Muthukumar S, Sukhatme VP, Crist SA, Rangnekar VM. 1995. The zinc finger transcription factor EGR-1 impedes interleukin-1-inducible tumor growth arrest. *Mol. Cell Biol*. Feb;15(2):682-92.

Chesbrough, H. 2003. *Open Innovation.* Harvard Business School Press, Boston, MA.

Dasgupta, P., & David, P. A. 1994. Towards a new economics of science. *Research Policy***,** 23: 487-521.

Ding, W., F. Murray, T. Stuart. 2007. An Empirical Study of Gender Differences in Patenting in the Academic Life Sciences. *Science* (5787) 665-667.

Ducor, P. 2000. Intellectual property: Coauthorship and coinventorship. *Science***,** 289: 873-875.

Fleming, L., & Sorenson, O. 2004. Science as a map in technological search. *Strategic Management Journal*, 25: 909-928

Fu C, Turck CW, Kurosaki T, Chan AC. 1998. BLNK: a central linker protein in B cell activation. *Immunity*. Jul;9(1):93-103.

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

Heller, Michael. 2008. *Gridlock Economy: How too much ownership wrecks markets, stops innovation and costs lives*. Philadelphia PA: Basic Books.

Huang, Kenneth, F. Murray. 2008. Does Patent Strategy Shape The Long-Run Supply Of Public Knowledge? Evidence From The Human Genome. Forthcoming in *Academy of Management Journal*.

Jensen, K., & Murray, F. 2005. The intellectual property landscape of the human genome. *Science*, 310: 239-240.

Kinane TB, Finder JD, Kawashima A, Brown D, Abbate M, Shang C, Fredericks WJ, Rauscher FJ 3rd, Sukhatme VP, Ercolani L. 1994. Growth of LLC-PK1 renal cells is mediated by EGR-1 up-regulation of G protein alpha i-2 protooncogene transcription. *Journal of Biological Chemistry* Vol. 269, Issue 44, 27503-27509, Nov, 1994

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

Lemley, Mark, A. 2007. The Surprising Virtues of Treating Trade Secrets as IP Rights. Stanford Law and Economics Olin Working Paper, No. 358.

Levin, R., Klevorick, A., Nelson, R., and S. Winter.1987. Appropriating the Returns from Industrial Research and Development. *Brookings Papers on Economic Activity* 3 pp. 783-831.

Lim, K. 1999. Do Firms Have to Perform Basic Research in order to Capture Spillovers, mimeo, MIT Sloan School.

Lessig, L. 2004. *Free Culture: How Big Media Uses Technology and the Law to Lock Down Culture and Control Creativity*. The Penguin Press.

Merton, R.K. 1957. Priorities in Scientific Discovery: A Chapter in the Sociology of Science. *American Sociological Review* 22(6) pp. 635-59.

Merton, R.K. 1973. The normative structure of science. In *The sociology of science: Theoretical and empirical investigations:* 267-280 (Ch. 13). Chicago, IL: The University of Chicago Press.

Mukherjee, A. & Stern, S. 2008. "Disclosure or Secrecy: The Dynamics of Open Science," *mimeo*., Northwestern.

Murray, F., & Stern, S. 2007. Do formal intellectual property rights hinder the free flow of scientific knowledge? An empirical test of the anti-commons hypothesis. *Journal of Economic Behavior and Organization*, 63(4): 648-687.

Murray, F. 2002. Innovation as co-evolution of scientific and technological networks: Exploring tissue engineering. *Research Policy*, 31 (8-9): 1389-1403

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

Romer, P. M. 1990. Endogenous Technological Change. *J. of Political Economy* **98**(5), S71-102, University of Chicago Press, October.

Rosenberg, N. 1990. Why Do Firms Do Basic Research (with Their Own Money)? *Research Policy* 19(2): 165-174.

Scotchmer, Suzanne. 1996. "Protecting Early Innovators: Should Second-Generation Products be Patentable?" *The Rand Journal of Economics* 27 Summer 1996: 322-331.

Shockley, William. 1949. "The Theory of P-N Junctions in Semiconductors and P-N Junction Transistors," *Bell System Technical Journal Vol. 28* pp. 435–89.

Smith, C. and M. N. Wise. 1989. *Energy and Empire: A Biographical Study of Lord Kelvin*. Cambridge: Cambridge University Press.

Stern, Scott. 2004. Do scientists pay to be scientists? *Management Science*. **50**(6) 835-853.

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