



MIT Open Access Articles

Intellectual Property Rights and Innovation: Evidence from Health Care Markets

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation	Williams, Heidi L. "Intellectual Property Rights and Innovation: Evidence from Health Care Markets." <i>Innovation Policy and the Economy</i> 16, no. 1 (January 2016): 53–87. © 2016 the National Bureau of Economic Research
As Published	http://dx.doi.org/10.1086/684986
Publisher	University of Chicago Press
Version	Final published version
Citable link	http://hdl.handle.net/1721.1/114270
Terms of Use	Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.

Intellectual Property Rights and Innovation: Evidence from Health Care Markets

Heidi L. Williams, *Massachusetts Institute of Technology and NBER*

Executive Summary

A long theoretical literature has analyzed optimal patent policy design, yet there is very little empirical evidence on a key parameter needed to apply these models in practice: the relationship between patent strength and research investments. I argue that the dearth of empirical evidence on this question reflects two key challenges: the difficulty of measuring specific research investments and the fact that finding variation in patent protection is difficult. I then summarize the findings of two recent studies that have made progress in starting to overcome these empirical challenges by combining new data sets measuring biomedical research investments with novel sources of variation in the effective intellectual property protection provided to different inventions. The first study, Budish, Roin, and Williams (2015a), documents evidence consistent with patents affecting the rate and direction of research investments in the context of cancer drug development. The second study, Williams (2013), documents evidence that one form of intellectual property rights on the human genome had quantitatively important impacts on follow-on scientific research and commercial development. I discuss the relevance of both studies for patent policy and discuss directions for future research.

I. Introduction

Academics and policymakers have long recognized that—in the absence of government intervention—competitive markets may provide too little incentive for research investments in new technologies. If a

market failure arises due to the public good nature of ideas, then there exists the potential for government intervention to improve welfare. In theory, many different policy instruments could be used to address this market failure. In practice, perhaps the most widely used policy lever both historically and globally has been intellectual property rights such as the patent system.

Intellectual property rights aim to increase private research investments in new technologies by allowing inventors to capture a higher share of the social returns to their inventions. Patents, for example, award inventors a temporary right to exclude competitors from marketing their invention. By allowing inventors to earn supracompetitive profits temporarily—during the term of the patent—patents aim to allow inventors to recoup their research and development expenditures and, hence, to provide dynamic incentives for investments in new technologies.

Dating back at least to William Nordhaus's dissertation (Nordhaus 1969), the optimal patent policy design problem has traditionally been framed as a trade-off between this benefit of providing incentives for the development of new technologies and the cost of deadweight loss from higher prices during the life of the patent.¹ However, there is very little evidence on a key parameter that is needed in order to apply such a framework to estimate the optimal patent term in practice: specifically, how does patent strength affect how much is spent on research investments? As we will discuss, the two most carefully conducted empirical studies that have attempted to investigate this question directly have uncovered little evidence that stronger (longer or broader) patents induce additional research investments (Sakakibara and Branstetter 2001; Lerner 2009).

Nordhaus-style models of optimal patent policy design have traditionally modeled innovations as isolated discoveries and predict an unambiguously positive relationship between patent strength and the rate of innovation. Most public policies are designed with this type of assumption in mind. However, as stressed in an influential line of work by Suzanne Scotchmer (Scotchmer 1991, 1996, 2004), many or most innovations are cumulative—in the sense that any given discovery is also an input into later follow-on discoveries. In markets where innovation is cumulative in this sense, optimal patent policy design also depends on how patents on existing technologies affect follow-on innovation. Until a few years ago, little empirical evidence was available on how patents affect follow-on innovation in practice, but many had argued

both qualitatively (Heller and Eisenberg 1998) and theoretically (Bessen and Maskin 2009) that patents may hinder follow-on innovation.

Taken at face value, a literal interpretation of the current set of available empirical evidence—that patents may not induce additional research investments and that patents may hinder follow-on innovation—would be that the patent system generates little social value. Indeed, such a position has been argued by Boldrin and Levine (2008) among others. I argue that drawing such a conclusion would be premature. In my view, there is a dearth of empirical evidence on how patents shape research investments and too little evidence available to draw strong conclusions. Two challenges have hindered empirical work in this area. First, finding the “right” data is challenging. Measuring specific research investments is usually quite difficult. A long literature in economics and related fields has used patent data as a proxy for research investments, but, by construction, patent data is ill-suited to investigating how variation in patent protection affects innovation given that changes in patent laws that grant longer patent terms will change inventors’ incentives to file for patent protection on existing research investments. Second, finding any variation (much less “clean” or quasi-experimental variation) in patent protection is difficult. There are few or no equivalents of clinical trials that have randomized patent protection across technological areas.² On paper, the patent system is uniform, providing a twenty-year term for all inventions. Although historically (and, more recently, for developing countries) there exists some cross-country variation in patent protection, this type of cross-country variation seems likely to be underpowered to test how patents affect research investments given that many inventions are created for a global market. That is, a relatively small market like Austria extending its patent term from 17 years to 20 years may reasonably be expected to induce a relatively small change in global research incentives.

In this paper, I summarize evidence from two of my recent investigations, which have attempted to overcome these empirical challenges in order to shed light on how patents affect research investments (Budish, Roin, and Williams 2015a; Williams 2013). First, these papers have taken advantage of new sources of data disclosing very specific (nonpatent) metrics of research investments that were able to be constructed by focusing in on single-market case studies: one paper focused on measuring research investments in cancer drug development; the second paper focused on measuring scientific research and product development investments related to the human genome. Second, each of these papers

identified and took advantage of new sources of variation in the *effective* intellectual property protection provided to different inventions, even if on paper the protection provided to these inventions was uniform. While both papers have important limitations, I argue that each investigation makes progress toward constructing the right “match” of data and variation, which can provide a framework for estimating the volume, type, and value of “missing” research and development (R&D) investments—exactly the parameters that are needed to inform a Nordhaus-style or Scotchmer-style calculation of optimal intellectual property policy design.

The goal of generating such empirical evidence is, of course, to inform policy design. The more effective are patents in inducing research investments, the stronger the case for longer or broader patents; on the other hand, the larger the costs of patents in terms of hindering follow-on innovation, the weaker is this case. While patent reform on this scale may seem like a relatively distant possibility, economically important patent reforms have been taking place incrementally in the form of a recent set of important US Supreme Court decisions. Specifically, the assumption that patents hinder follow-on innovation has informed at least four recent US Supreme Court decisions that have restricted the set of discoveries eligible for patent protection (Kesselheim et al. 2013). First, in *Bilski v. Kappos*, the Court invalidated patent claims on an investment strategy, announcing it supported a “high enough bar” on patenting abstract ideas that it would not “put a chill on creative endeavor and dynamic change.” Second, in *Mayo Collaborative Services v. Prometheus Laboratories, Inc.*, the Court invalidated patent claims on methods of using genetic variation to guide pharmaceutical dosing, expressing concern that “patent law not inhibit further discovery by improperly tying up the future of laws of nature.” Third, in *Association for Molecular Pathology v. Myriad Genetics, Inc.*, the Court invalidated patent claims on genomic DNA, concluding that “[g]roundbreaking, innovative, or even brilliant” discoveries of natural phenomena are not patentable given that such patents “would ‘tie up’ the use of such tools and thereby inhibit future innovation premised upon them.” Finally, since the publication of Kesselheim et al. (2013), in *Alice Corp v. CLS Bank International*, the Court invalidated patent claims on software based on similar arguments. These US Supreme Court decisions have sharply reduced patent protection in several economically important sectors of the economy, and yet economists have little to contribute in terms of empirical evidence on what to expect the welfare consequences

of these decisions to be. In particular, the elimination of patent protection in these sectors could encourage firms to rely on other methods for protecting their intellectual property—such as trade secrets—and an understanding of the *relative* costs and benefits of patent protection compared to these outside options is needed in order to evaluate the welfare consequences of these court decisions.

Section II clarifies the challenges that I see as having hindered empirical research on the key research questions needed to inform optimal intellectual property policy design. Sections III and IV then summarize evidence from two of my recent papers, which have attempted to overcome these challenges: Budish, Roin, and Williams (2015a) and Williams (2013). Section V concludes.

II. Challenges in Quantifying “Missing” R&D Investments

In many markets, the costs of poorly designed policies are conceptually straightforward to measure. For example, many government old-age support programs create incentives that hasten elderly individuals’ transitions from the labor force into retirement. By estimating how labor supply and retirements change in response to changes in incentives, one can begin to evaluate the welfare consequences of such policies.

In contrast, in markets for innovation, the costs of poorly designed policies are much more difficult to detect and quantify. If the patent system as currently designed contains flaws relative to an optimally designed system, the cost of those flaws in policy design is that there are “missing” R&D investments—in the sense that there are scientifically feasible R&D investments that are not undertaken due to misaligned incentives in the patent system. Unfortunately, measuring such missing R&D investments directly is usually quite difficult. For example, pharmaceutical firms do not keep systematic records of—or share publicly—scientifically promising drug compounds that were abandoned due to misaligned incentives in the patent system. Estimating the quantity and type of such missing R&D investments indirectly is also usually quite difficult because missing private R&D due to misaligned incentives in the patent system is difficult to distinguish from private R&D investments that are “efficiently” missing due to alternative explanations such as a lack of market demand or a lack of scientific opportunities.

In my view, estimating the volume, type, and value of missing R&D investments requires overcoming two challenges: first, finding the right

data, and, second, identifying variation in patent protection across otherwise similar potential inventions. In this section, I briefly discuss each of these challenges in more detail.

A. *Measurement Challenges*

The first empirical challenge facing researchers seeking to estimate the relationship between patent strength and research investments is that measuring specific research investments is usually quite difficult. Large literatures in economics, law, management, finance, and related fields have used patent-based outcome variables as measures of research investments. However, I argue that patent-based measures are usually *not* a useful outcome variable in studies where researchers wish to analyze how variation in patent protection affects research investments.

Take as an example a recent paper by Abrams (2009). Abrams investigates the 1995 Agreement on Trade-Related Aspects of Intellectual Property Rights (TRIPS), which induced patent term extensions that were heterogeneous across technology classes (as a function of the United States Patent and Trademark Office [USPTO] processing times in different technology areas). He uses patent counts and citation-weighted patent counts as outcome variables to test whether technology classes that experienced longer patent term extensions had larger increases in innovation, as measured by these patent-based outcomes. The concern that arises with this type of study is that such changes in patent policy change inventors' incentives to file for patent protection on existing research investments. In the absence of some outside (nonpatent based) measure of research investments, seeing that patent counts increased in response to TRIPS-induced patent term extensions is difficult to interpret.

That said, there are at least two clever exceptions in the literature of how researchers have circumvented this concern. First, Lerner (2009) undertook an ambitious data collection effort in order to investigate how major patent policy changes in 60 nations over 150 years affected research investments.³ Lerner's key measurement insight is that while looking at changes in patent filings, for example, in Austria when Austria's patent term lengthened may be difficult to interpret, looking at changes in patents filed in Great Britain by residents of Austria—where Great Britain was chosen as a country that had relatively stable patent policy over the time period of study—can provide a “clean” measure of research investments. Table 1 reproduces table 1 from Lerner (2009).

Table 1
Change in Patenting Activity around Policy Shifts (Year -2 to +2)

	Residents' Patenting in Great Britain	Residents' Patenting in Country	Foreign Patenting
<i>Panel A: Unadjusted Changes in Patenting around Policy Changes</i>			
Positive patent policy changes	-27	+2,424	+8,662
Ambiguous/negative changes	+210	+529	+1,401
<i>Panel B: Changes in Patenting, Adjusted by Equal-Weighted Index</i>			
Positive patent policy changes	-101	-1,617	+4,979
	-(4.61)***	(1.86)*	(2.41)**
Ambiguous/negative changes	-217	-525	+390
	-(3.19)***	(0.34)	(1.28)
<i>Panel C: Changes in Patenting, Adjusted by Value-Weighted Index</i>			
Positive patent policy changes	-100	-932	+5,617
	-(4.52)***	(1.69)*	(2.85)***
Ambiguous/negative changes	-137	-408	+501
	(2.40)**	(0.07)	(1.65)

Note: This reproduces table 1 from Lerner (2009).

* Significant at the 10% level.

** Significant at the 5% level.

*** Significant at the 1% level.

Panel A documents the changes in patents filed from two years before to two years after the policy shift, separately for policy changes that unambiguously increased patent protection and for policies that either reduced or ambiguously changed the strength of patent protection. The estimates in columns (3) and (4) suggest that both domestic and foreign patent applications increased in countries undertaking increases in patent protection; in contrast, the estimates in column (2) provide no evidence of an increase in British patent applications by residents of the country where the policy change occurred.⁴ Lerner documents and reports results from a number of additional specifications, none of which suggest that stronger patent laws induced more R&D investments as measured by British patent applications filed by residents of the country where the policy change occurred.

Second, Galasso and Schankerman (2015) propose a novel approach for investigating how patents on existing technologies affect follow-on innovation. Social scientists have typically focused on patent citations as a measure of follow-on innovation (e.g., Jaffe, Trajtenberg and Henderson 1993) because patents are required to cite previous patents and other so-called prior art on which their work builds. But, by construc-

tion, patent citations cannot be used to measure cumulative innovation on nonpatented technologies. While most economists would, I think, agree that much of the innovation we observe in real-world markets is cumulative in the sense that any given discovery builds on past discoveries, finding ways of measuring the cumulative nature of innovation in practice is quite challenging. The key measurement insight of Galasso and Schankerman (2015) is that by focusing on patent *invalidations*, because later patent applicants are still required to cite patents when relevant even if they have been invalidated and thus put into the public domain, the authors can “rescue” patent citations as a follow-on innovation measure in order to compare follow-on citations across patented and nonpatented (invalidated) technologies.

Hence, while patent-related outcomes are difficult to interpret in general, thoughtful exceptions exist that have enabled researchers to use patent-based data to measure research investments and follow-on innovation. However, in general, my view is that an important direction for future work is to develop nonpatent based measures of both research measures and follow-on innovation. An early contribution on this measurement challenge was the work of Sakakibara and Branstetter (2001)—an analysis we will discuss in more detail in the following—who use reported R&D investments as an outcome to measure whether stronger patent laws affected research investments. Either as a marker for general research investments or, more specifically, as a measure of basic scientific research investments, researchers have used either counts of scientific papers (Williams 2013) or citations to scientific papers (Murray and Stern 2007; Murray et al. 2008) as outcomes. A relatively larger and growing literature—starting with Acemoglu and Linn (2004) and Finkelstein (2004)—has used clinical trials and drug approvals as measures of research investments.⁵ Relatedly, two of my papers—Williams (2013) and Sampat and Williams (2015)—have used diagnostic test availability as an outcome variable. Finally, Moser, Ohmstedt, and Rhode (2014) collect data on plant yields as a measure of technological progress in agriculture.

While the preceding list is certainly not comprehensive, I hope that it gives some sense of the types of measures that have been productively used in the past literature as metrics of research investments. While much progress has been made in this area relatively recently, it is worth stressing that ideally measurement progress would continue to be made on at least two fronts. First, progress on nonpatent measures of research investments has been disproportionately made in

some markets—particularly health care—and not others. Second, even in health care markets, ideally, we would continue to make progress toward measuring more welfare-relevant outcomes. That is, if a paper was able to document that longer patent terms induced more research investments, that would be very useful, but, ideally, we would like to trace out how those research investments in turn affect welfare-relevant outcomes such as prices and health outcomes.⁶

B. Inference Challenges

A second empirical challenge facing researchers seeking to estimate the relationship between patent strength and research investments is that finding any variation in patent protection is difficult. On paper, the patent system provides uniform protection—a twenty-year patent term—to all inventions. Indeed, under the TRIPS agreement, administered by the World Trade Organization (WTO), WTO members (almost all countries) are required to make twenty-year patent terms available for all technology fields.

Historically—and more recently, prior to the TRIPS agreement, in developing countries—there has existed some cross-country variation in patent protection. In fields such as labor economics and public finance, this type of legal variation across geographic markets can often be quite useful in estimating the causal impacts of public policies, and, indeed, this type of cross-country variation in patent laws is very useful in answering many research questions: for example, testing how country-level patent protection affects country-specific drug launch decisions (as in Cockburn, Lanjouw, and Schankerman 2014; Duggan, Garthwaite, and Goyal 2014; and Kyle and Qian 2014) or testing how patent laws affect the direction of technological change within a country (as in Moser 2005). However, in my view, country-level changes in patent laws are not well-suited to investigating the question of how patent strength affects research investments because any given market extending its patent term from, for example, 17 years to 20 years will induce a relatively small change in global incentives for research investments. That is, because many innovations are created for a global market, changes in any given country's patent law may be expected to be underpowered to test how patents affect (global) research investments.

This concern about country-level changes in patent laws being potentially underpowered to empirically test how patents affect research investments is one interpretation of the Lerner (2009) results, which failed

to uncover any evidence that stronger patent terms increased research investments. This also provides a potential explanation for the results of Sakakibara and Branstetter (2001), who investigate a set of Japanese patent reforms enacted in 1988 that strengthened patent protection in Japan. While this package of reforms involved a number of changes, a key feature was an increase in patent breadth that arose from a switch from the Japanese single-claim (“sashimi”) system to the type of multiclaim system common in the United States and elsewhere.⁷ Figure 1 reproduces figure 2 from Sakakibara and Branstetter (2001), which confirms that there was a substantial behavioral response to this reform, namely—as expected given the switch from a single-claim system to a multiclaim system—we observe a discrete and sharp increase in 1988 in the number of claims per patent application.⁸ The authors then use micro-level data on the R&D investments of 307 publicly traded Japanese manufacturing firms in various industries to test whether this

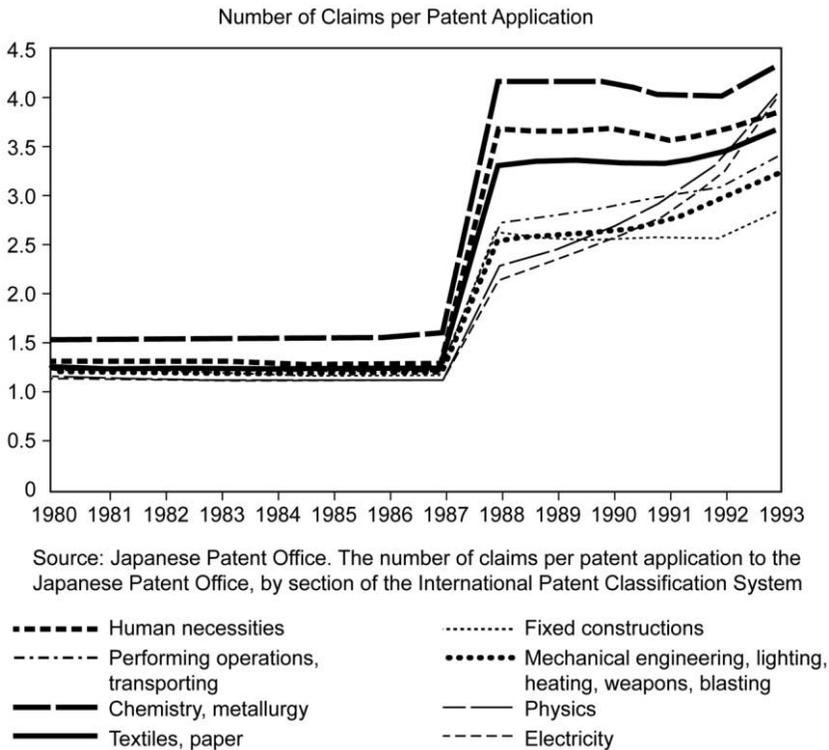


Fig. 1. Number of claims per patent application

Note: This reproduces figure 2 from Sakakibara and Branstetter (2001).

strengthening of patent protection induced additional research investments. Aggregate time series evidence indicates that there was an increase over time in R&D spending by Japanese firms which *predated* the 1988 reform, that the years 1988 and 1989 were marked by a relative *decline* in R&D spending, and that the prereform trend of increasing R&D resumed in 1990 but that there is little evidence of a deviation from the earlier trend.⁹ The authors document a number of additional specifications, none of which suggest that stronger patents induced more R&D investments, and conclude that, while the policy change induced a behavioral response by Japanese firms, it failed to induce discernible increases in R&D.

III. How Do Intellectual Property Rights Affect Research Investments?

Both the Lerner (2009) and Sakakibara and Branstetter (2001) investigations uncover little evidence that stronger (longer or broader) patent terms induce additional research investments, but, because both rely on country-specific law changes as variation, they may be underpowered to detect such effects. Beyond these papers, there is very little direct evidence on the question of how patent strength affects research investments. The ideal setup for investigating this question empirically would start with (quasi-random) variation in the strength of patent protection available across different types of potential inventions. For example, given a large number of technology markets, if some markets were eligible for 10 years of patent protection and others were eligible for 20 years of patent protection, then one could then collect data measuring research investments across these different markets and estimate whether markets with longer patent protection tended to realize higher levels of private research investments.

In practice, the relative uniformity of the patent system stymies researchers' abilities to find such empirical variation. Within a country, the incentives provided by the patent system are uniform: for example, in the United States, all inventions are awarded a 20-year period of market exclusivity. In Budish, Roin, and Williams (2015a), we start with the observation that although the incentives provided by the patent system are uniform in theory, in practice, the patent system can provide remarkably uneven protection across different classes of potential inventions. We investigate this idea in the context of cancer drug development. As we will discuss, while our estimates are not able to nail down

the relationship between patents and research investments, our analysis identifies—both theoretically and empirically—a distortion of private research investments away from certain types of research projects and analyzes a number of policy levers that could be used to address this distortion.¹⁰

A. Empirical Context: Cancer Drug Development

Most new cancer drugs are approved to treat patients with relatively advanced forms of late-stage cancer, as opposed to treating patients with early stage cancer or preventing cancer. While this pattern could, in theory, be explained entirely by market demand (patients have higher demand for treatments for late-stage cancers relative to early stage cancers) or by scientific challenges (the “science” of treating late-stage cancers is easier than the science of treating early stage cancers), we investigate an alternative hypothesis: private firms may invest more in late-stage cancer treatments—and too little in early stage cancer treatments or cancer prevention drugs—because late-stage cancer drugs can be brought to market comparatively quickly, whereas drugs to treat early stage cancer or to prevent cancer require a much longer time to bring to market.

These differences in time-to-market (which we refer to as “commercialization lags”) arise because, prior to selling their inventions to consumers, firms developing new pharmaceutical drugs must complete US Food and Drug Administration (FDA)-required clinical trials documenting evidence that their drugs are safe and effective. “Effective” in this context is usually interpreted as improving patient survival rates relative to a placebo or relative to another available drug treatment in a randomized control trial.¹¹ Standard power calculations suggest that a statistically significant difference in survival outcomes between the treatment and control groups of a randomized trial can be observed more quickly in patient populations with a higher mortality rate, whereas such a difference will take longer to observe if patients have a relatively longer life expectancy. This implies that clinical trials must be longer in duration when evaluating treatments for early stage cancer patients relative to treatments for late-stage cancer patients.

We present a simple theoretical model to clarify that these longer required clinical trials—and, hence, longer commercialization lags—could distort private research investments away from long-term projects for either of two reasons. First, private firms may be more impa-

tient than the social planner. Second, the fixed-term structure of the patent system may also induce a distortion. Because pharmaceutical firms face strong incentives to file for patents at the time of discovery (“invention”) rather than the time of first sale (“commercialization”), the effective patent term that a firm receives decreases with the commercialization lag of the project. That is, while a firm that developed, patented, and sold an invention near the time of discovery would receive a 20-year patent term, firms facing long commercialization lags can receive much shorter effective patent terms. In the extreme, research projects requiring clinical trials that take longer than 20 years would receive zero effective patent protection, implying that patents provide very little incentive for private firms to engage in long-term research.

B. Measurement Challenges

The key prediction we want to empirically test is that there are “missing” research investments on long-term projects that are scientifically feasible but that are not pursued due to this distortion. Testing this prediction encounters two measurement challenges. First, we do not observe commercialization lags for research projects that are never pursued. Second, we need to find a way of measuring research investments across different types of cancers for which firms face different expected commercialization lags.¹²

Two features of cancer markets allow us to make progress on these measurement challenges. First, the treatment of cancer patients is organized around the organ (e.g., lung, breast, prostate) and stage (localized, regional, metastatic) of disease. Because cancer treatments tend to be specific to the organ and stage of the primary tumor, these classifications provide a natural categorization of observed and potential research investments. To be specific, because the treatment of patients is organized around these organ-stage designations (for example, metastatic prostate cancer patients are treated differently than localized breast cancer patients), the FDA’s approval of new cancer drugs is organized around providing evidence on the safety and efficacy of a new drug for patients of a given organ-stage designation. For example, Genentech’s drug Bevacizumab was approved by the FDA in 2004 for the treatment of patients with metastatic carcinoma of the colon and rectum.

Second, for each cancer-stage group, we observe a good predictor of how long it would take to commercialize a new drug even if no drug

has ever been developed for that group of patients: namely, the survival time of that patient group. As discussed, survival time predicts commercialization lags because patient groups with longer survival times will require longer clinical trials. While this idea is quite simple, the fact that we can use survival time as an observable variable that predicts how long commercialization lags would be for developing a drug for a patient group with no available treatments is quite valuable.

Given this framework of cancer-stage types and our proxy variable (survival rates) for commercialization lags, we then need to construct a measure of the research investments made investigating potential treatments for each cancer-stage group of patients. The usual measures of research investments—patent counts and scientific publication counts—are difficult to construct in this context because it is technically challenging to “read” the text of a patent grant or the text of a scientific publication and discern whether the observation is related to, for example, a treatment for metastatic breast cancer or a treatment for regional breast cancer. Instead, we take advantage of the fact that private firms in this market have an incentive to reveal exactly the patient groups that they are attempting to develop drugs for in a particular form: namely, by choosing which patient groups are eligible to enroll in their clinical trials. As noted in the preceding, because drugs are FDA approved for specific cancer-stage groups, a private firm aiming to develop a treatment for metastatic breast cancer patients must show the FDA evidence that its drug is safe and efficacious in a group of metastatic breast cancer patients. We, thus, construct our data by compiling a new data set of advertisements recruiting patients to participate in clinical trials and parse information on which patient groups are eligible to enroll in each trial. This aspect of our data construction enables a relatively precise link between our measures of expected commercialization lags (survival time) and our measure of R&D activity (clinical trials) across cancer types and stages of disease.

C. Inference Challenges

Using this data, we first document that—consistent with our conjectured distortion—patient groups with longer commercialization lags (as proxied by higher survival rates) tend to have lower levels of R&D investments. Figure 2—which reproduces figure 1(a) from Budish, Roin, and Williams (2015a)—gives a sense of this pattern in aggregated stage level data. Patients with metastatic cancer have a five-year survival rate

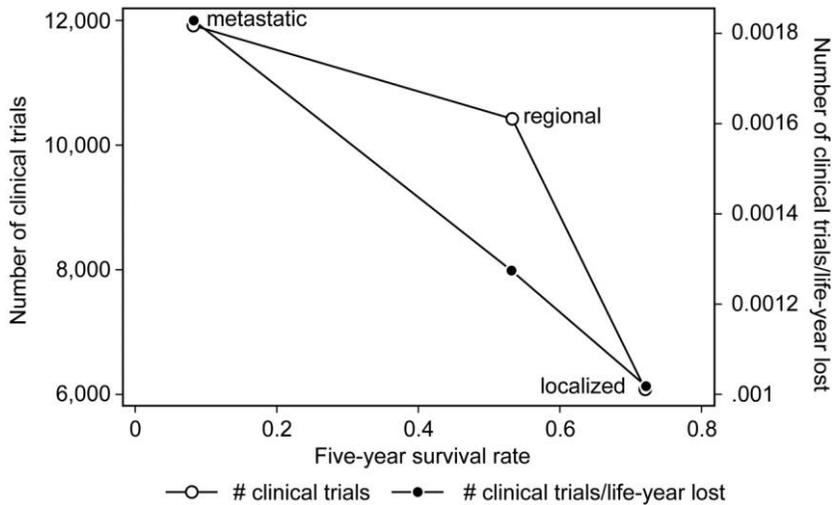


Fig. 2. Survival time and research investments

Note: This reproduces figure 1a from Budish, Roin, and Williams (2015a).

of around 10% and are eligible to enroll in around 12,000 clinical trials in our data. In contrast, patients with localized cancer have a five-year survival rate of around 70% and are eligible to enroll in just over 6,000 clinical trials. A rough adjustment for market size (shown on the right-hand-side y axis, looking at the number of clinical trials per life-year lost from cancer) does little to change this basic pattern.

By itself, this fact is difficult to interpret for two reasons. First, this correlation need not reflect a causal relationship between commercialization lags and R&D investments. If patient demand is relatively lower or scientific opportunities are relatively less available for early stage cancers, then a policy that shortened commercialization lags may have no effect on R&D investments. Second, even if this fact did reflect a causal effect of commercialization lags on R&D investments, on its own, this fact need not be evidence of a distortion because the social planner—due, for example, to the social discount factor—would also prefer to pursue research projects that can be completed more quickly. To address these two concerns, we document evidence from two complementary empirical tests.

First, we document causal evidence that shortening commercialization lags increases R&D investments. The key idea behind this empirical test is to take advantage of the fact that some types of cancers are

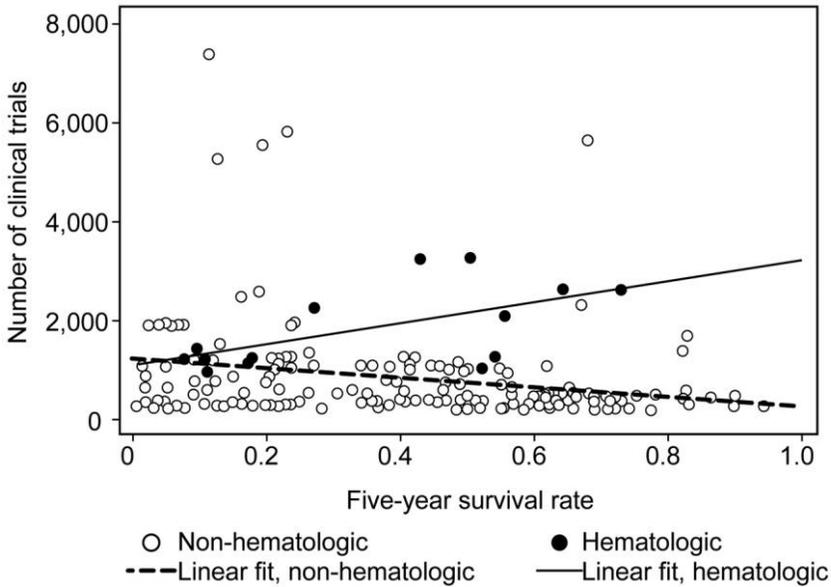


Fig. 3. Surrogate endpoints, survival time, and R&D investments
 Note: This reproduces figure 4 from Budish, Roin, and Williams (2015a).

allowed to use surrogate endpoints (nonmortality endpoints), which break the link between patient survival times and clinical trial lengths. Perhaps the most clearly established nonmortality related endpoint is “complete response” for leukemias, which is measured based on blood cell counts and related bone marrow measures. Because improvements in these outcomes have been documented to predict subsequent improvements in survival outcomes, the FDA has long approved new drugs for leukemias and lymphomas on the basis of the fact that drugs improve these complete response measures. This variation is useful because it means that for some cancers—namely, leukemias and lymphomas (known together as hematologic cancers)—the survival time of a given patient group should *not* predict commercialization lags, and, hence, we should see a different (less negative) relationship between survival times and research investments for those cancer types. As documented in figure 3—which reproduces figure 4 from Budish, Roin, and Williams (2015a)—indeed there is *not* a negative relationship between survival time and R&D in this sample of cancers allowed to use surrogate endpoints. This suggests a causal relationship between commercialization lags and research investments: if commercialization lags

were shortened, there are scientific research opportunities that would be beneficial to pursue, and patients would have demand for those treatments. However, this test leaves open the possibility that the social planner and private firms symmetrically respond to commercialization lags and, thus, does not provide direct evidence of a distortion of private research dollars away from long-term projects.

Second, to address this gap, we contrast public and private research investments. While—consistent with our theoretical model—commercialization lags reduce both public and private research investments (results not shown), figure 4—which reproduces figure 5(b) from Budish, Roin, and Williams (2015a)—documents that the commercialization lag-R&D correlation is more negative for privately financed R&D than for publicly financed R&D. Taken at face value, this suggests that opportunities exist to conduct clinical trials that the public sector is willing to fund and that patients are willing to enroll in but which the private sector is unwilling to finance. Taken together with the surrogate endpoint evidence, the evidence from these two empirical tests provides support for the idea that commercialization lags distort private research investments away from long-term projects.

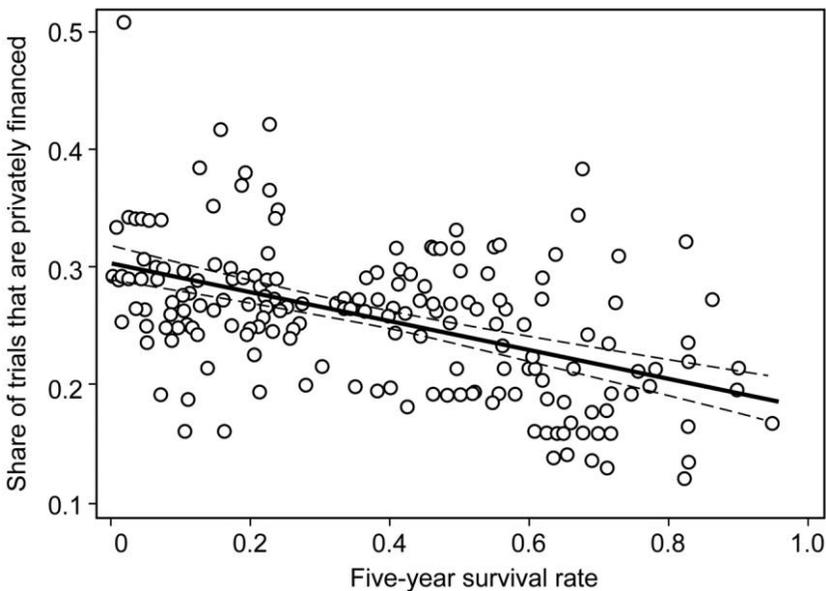


Fig. 4. Survival time and financing of clinical trials

Note: This reproduces figure 5b from Budish, Roin, and Williams (2015a).

D. Interpretation

These empirical estimates are directly relevant to analyzing two policy levers that could address the distortion of private research dollars away from long-term projects: namely, allowing firms to rely on (valid) surrogate endpoints that would enable shorter clinical trials and targeting publicly funded R&D subsidies toward projects with long commercialization lags. However, these estimates *cannot* directly speak to the importance of patents as a mechanism for these results, given that our empirical estimates do not speak to which mechanism—patents or corporate impatience—are driving our results and, thus, cannot be used to estimate the value of patent design changes that would address only the patent distortion (such as starting the patent clock at commercialization). In ongoing work, Budish, Roin, and I are pursuing a second joint project to more directly investigate how patents impact research investments by examining how research investments into new uses of old drugs vary with patent protection on initial pharmaceutical innovations (Budish, Roin, and Williams 2015b).

Despite being unable to pin down the precise mechanism behind the distortion of private research dollars away from long-term research investments, these results are directly policy relevant in estimating the potential value of reducing commercialization lags. Taking advantage of our surrogate endpoint variation, we estimate counterfactual R&D allocations and induced improvements in cancer survival rates that would have been observed if commercialization lags were reduced. Our back-of-the-envelope calculation suggests that the distortion of private research dollars away from long-term projects has quantitatively important implications for the survival outcomes of US cancer patients: we estimate that among one cohort of patients—US cancer patients diagnosed in 2003—longer commercialization lags generated around 890,000 lost life-years. Valued at \$100,000 per life-year lost (Cutler 2004), the estimated value of these lost life-years is on the order of \$89 billion for this single cohort of patients.

While we are not aware of previous attempts to estimate the value of reductions in commercialization lags, this quantitative evidence is consistent with prior qualitative evidence for the case of heart disease. Heart disease is the leading cause of death in the United States, but the age-adjusted rate of death from heart disease has dropped by 50% since 1968. While factors such as the decline in smoking have contributed to this trend, much of the decline in heart disease has been attributed to

improved pharmacological preventives and treatments for cardiovascular disease, including the development of beta-blockers, angiotensin-converting enzyme (ACE)-inhibitors, and statins (Weisfelt and Zieman 2007). All of these drugs were approved on the basis of evidence that they reduced blood pressure and cholesterol levels—outcomes that can be measured much more quickly than morbidity and mortality (Psaty et al. 1999). Some researchers have argued that without these surrogate endpoints, it is unclear whether these drugs would have reached the market as treatments for heart disease (Lathia et al. 2009; Meyskens et al. 2011). Importantly, from a policy perspective, these surrogate endpoints for heart disease were first identified by the Framingham Heart Study, a large-scale, multidecade, federally funded observational study that documented evidence that high blood pressure and low-density lipoprotein (LDL) cholesterol are critical risk factors linked to cardiovascular disease. Subsequent clinical trials helped to validate these prognostic factors, which led the FDA to accept them as surrogate endpoints in cardiovascular trials (Meyskens et al. 2011). Note that public subsidies—such as federal support for the Framingham study—were likely important in this context because any individual firm's investment in discovering and validating surrogate endpoints would generate benefits that largely spill over to other firms. Both our empirical evidence on the effects of surrogate endpoints for hematologic cancers and this historical case study for heart disease suggest that public research investments aimed at establishing and validating surrogate endpoints may have a large social return.

IV. How Do Intellectual Property Rights on Existing Technologies Affect Follow-on Innovation?

As discussed in the introduction, Nordhaus-style models of optimal patent policy design have traditionally modeled innovations as isolated discoveries and predict an unambiguously positive relationship between patent strength and the rate of innovation. In the spirit of such theoretical models, researchers' attempts to empirically evaluate the effectiveness of intellectual property rights in spurring innovation have traditionally focused on the question of whether the prospect of being awarded a patent induces additional research investments.

However, in practice, many or most innovations are cumulative in the sense that any given discovery is also an input into later follow-on discoveries. In such markets, the overall effectiveness of intellectual

property rights in spurring innovation also depends on how patents on existing technologies affect follow-on innovation. Consider an example that has been prominent in recent policy debates the case of patents on sequenced human genes. The firm Myriad Genetics was granted patents on human genes correlated with risks of breast and ovarian cancer. These genes have been the focus of a large amount of basic scientific research—both before and after these patents were granted—and the evidence that variation on these genes is linked with breast and ovarian cancer risks has been applied to develop a diagnostic test, BRACAnalysis, which assesses an individual's risk profile for these diseases. Of course, the prospect of obtaining these gene patents may have shaped Myriad's willingness to invest from an *ex ante* perspective. But the question of interest in this section is whether—from an *ex post* perspective—Myriad's gene patents discouraged follow-on scientific research or commercial product development. While there is a well-developed theoretical literature on this question, empirical evidence is scarce. In Williams (2013), I attempt to shed light on this question by analyzing how one nonpatent form of intellectual property on the human genome affected follow-on innovation.

A. *Empirical Context: Sequencing the Human Genome*

Understanding my empirical analysis of this question requires some background details on the public and private efforts to sequence the human genome. The publicly funded Human Genome Project was launched in 1990 and originally planned to finish sequencing the human genome by 2005 (Collins and Galas 1993). In 1998, a private firm—Celera, led by former NIH scientist Craig Venter—was launched, with the intention to finish sequencing the human genome by 2001; the public effort subsequently announced a revised plan to complete its sequencing effort by 2003 and to release a “draft” sequence of the human genome even earlier (Collins et al. 1998; Venter et al. 1998; Pennisi 1999). The two efforts jointly published draft versions of their sequenced human genomes in February 2001, Celera in *Science* (Venter et al. 2001) and the Human Genome Project in *Nature* (Lander et al. 2001). Importantly, both of these published versions were drafts, in the sense that they were incomplete. Celera's sequencing efforts stopped with this draft human genome publication, whereas the Human Genome Project's sequencing effort continued and was completed in 2003 (Wade 2003).

This “race” to sequence the human genome induced variation across

human genes in whether they were held with intellectual property rights or whether they were always held in the public domain. As of 1996, the publicly funded human genome effort was regulated by the so-called Bermuda rules, which required researchers to post sequenced data on an open-access website within 24 hours of sequencing. The stated goal was “to encourage subsequent research and development and to maximize [the data’s] benefit to society.” In contrast, Celera—as a private firm—attempted to cover its sequenced genes with intellectual property rights so as to realize a higher private return on their investments. Celera filed a large number of patent applications on its sequenced genes, but these patent applications were largely rejected (likely because—at the time—only knowing the sequence of a gene and not something about its medical usefulness was insufficient to meet the USPTO’s “utility” standard).

In the absence of being able to obtain patent protection, Celera designed a contract law-based form of intellectual property that has several important features. First, Celera’s intellectual property enabled the firm to disclose their data in the sense that any individual could view data on Celera’s draft genome through Celera’s website or could obtain a free data DVD from the firm (Science Online 2001). Academic researchers faced no restrictions in using Celera’s data for academic publications. Second, Celera’s intellectual property—by placing restrictions on redistribution of their data—enabled the firm to price discriminate across different users of their data. While the data was free to individual researchers interested in using the data for academic research, larger academic labs were charged fees, and private biotechnology and pharmaceutical firms were charged larger fees (Service 2001). Third, any downstream users interested in using the data for commercial purposes needed to negotiate a licensing agreement with Celera.

While this package of intellectual property was successful—from Celera’s perspective—on several margins, it had an important weakness. Patents offer protections to inventors that are durable to reinvention: if one firm develops and patents an idea, and another firm rediscovers the same idea, the patent enables the first firm to exclude the second firm from entering the market. In contrast, Celera’s intellectual property was *not* durable with respect to reinvention. From a practical perspective, what this meant was the following. When Celera and the Human Genome Project published their draft genomes in 2001, neither draft was complete: some genes had been sequenced only by Celera, others had been sequenced only by the Human Genome Project, some

genes had been sequenced by both efforts, and other genes had been sequenced by neither effort. Celera's intellectual property applied to genes sequenced by Celera as of 2001 but not yet sequenced by the Human Genome Project. As the Human Genome Project continued its sequencing effort post-2001, it resequenced genes that had initially been held with Celera's intellectual property, an action that moved these genes into the public domain. By the time the Human Genome Project's sequencing effort was completed in 2003, all of Celera's genes had been placed in the public domain. The institutional context suggests that market participants were well aware of this feature of Celera's intellectual property: Shreeve (2005) quotes Craig Venter as saying, "Amgen, Novartis, and now Pharmacia Upjohn have signed up [for Celera data purchases] knowing damn well the data was going to be in the public domain in two years anyways. They didn't want to wait for it" (pg. 205). In Williams (2013), I analyze how Celera's nonpatent form of intellectual property rights affected follow-on scientific research and product development, relative to a counterfactual of Celera's genes having always been in the public domain.

B. Measurement Challenges

A first challenge in investigating this question of how intellectual property rights affect follow-on innovation is a set of measurement issues: which human genes were covered by Celera's intellectual property and how can we measure follow-on innovation on those genes relative to human genes which were always in the public domain (by nature of having first been sequenced by the Human Genome Project)?

Identifying Which Technologies Are Covered by Intellectual Property Rights

I construct my data based on the census of human genes (around 30,000 genes). For each gene, I was able to track the timing of when the gene was sequenced by the Human Genome Project by querying an online NIH database to identify the date that gene sequence first appeared in that online database. Because the Bermuda rules—as described in the preceding—required publicly sequenced data to be posted online within 24 hours of sequencing, the date of appearance in this online database is a reasonable proxy for when each gene was sequenced by the publicly funded sequencing effort.

The key input enabling the measurement of which genes were covered with Celera's intellectual property was an academic paper by Istrail et al. (2004), which compared Celera's 2001 draft genome with a snapshot of the publicly sequenced data at a given point in time. By combining this cross-sectional comparison with the data described in the preceding on when each gene was sequenced by the publicly funded effort, I am able to identify which genes were "Celera" genes in 2001 and the timing of when each of those genes was resequenced by the Human Genome Project between 2001 and 2003.

This measurement challenge—while it may at first seem unique to this particular investigation—is, at least in my view, a major stumbling block to research in general on how intellectual property rights affect follow-on innovation. Any analysis of that question requires cataloging which discoveries are and are not covered by intellectual property rights. Unfortunately, this type of linkage is very rarely straightforward to construct. It is difficult—even for an expert—to look at the text of a patent grant, for example, and discern what product in the real world is covered by that patent. In some cases, regulations require disclosures of these linkages, such as in the FDA's *Orange Book* (as used by Hemphill and Sampat 2011), which lists patents that are self-reported by pharmaceutical firms to be relevant to FDA approved drugs, but such linkages are quite rare. As I will discuss in the conclusions of this paper, uncovering new ways of identifying such linkages could greatly ease researchers' ability to investigate the question of how patents on existing technologies affect follow-on research investments.

Tracing Follow-On Innovation

Measuring the cumulative nature of innovation is quite challenging in most markets. For this investigation, I constructed measures of cumulative innovation based on linkages between genes and phenotypes. For example, the link between the breast cancer (BRCA) gene and breast or ovarian cancer risk represents a genotype-phenotype linkage. For each gene, I collected data on scientific publications documenting evidence for genotype-phenotype linkages from the Online Mendelian Inheritance in Man (OMIM) database, which curates biomedical publications with identifiers for the genes and diseases relevant to each publication. In order to measure commercial development related to each gene, I collected data from GeneTests.org, a voluntary listing of US and international laboratories offering genetic testing. While not comprehensive,

over the time period of my study, private firms and academic medical centers had an incentive to list their tests in GeneTests.org as a way of advertising the availability of their tests to both patients and doctors. Together, these linked measures of follow-on innovation provide a window into how genes transition from being the subject of basic scientific research (as documented by the publication measure) to being incorporated into commercial products that are available to consumers (as measured by their inclusion in genetic diagnostic tests) and measure these aspects of follow-on innovation in a relatively precise way.

C. Inference Challenges

A second challenge in investigating this question of how intellectual property rights affect follow-on innovation is an inference challenge: how can we construct a counterfactual for what follow-on innovation on Celera genes would have been if these genes would have always been in the public domain?

I start by documenting the simplest possible comparison: namely, comparing follow-on innovation on Celera genes relative to follow-on innovation on non-Celera genes that were sequenced in the same year (2001). Table 2—which reproduces table 1 from Williams (2013)—documents this comparison. Taken at face value, these estimates suggest that Celera’s intellectual property led to economically and statistically significant reductions in follow-on R&D on the order of 20–30% of the sample means. Celera genes had an average of 1.2 scientific publications between 2001 and 2009, relative to 2.1 scientific publications for non-Celera genes published in the same year. Around 3% of Celera genes were used in a gene-based diagnostic test as of 2009, compared to 5.4% of non-Celera genes sequenced in the same year.

Of course, the concern that arises with this simple cross-tabulation is that it could reflect either a negative effect of Celera’s intellectual property on follow-on research or could reflect that Celera’s genes had lower inherent potential for follow-on research. That is, table 2 reflects a causal effect of Celera’s intellectual property on follow-on innovation only to the extent that Celera’s intellectual property was as good as randomly assigned across genes sequenced in 2001. Both qualitative evidence that I summarize and new quantitative evidence that I document in Williams (2013) suggest that the second factor—a type of selection bias—could be a concern, which motivates the construction of two additional empirical tests that directly address selection. First, I restrict

Table 2
Innovation Outcomes for Celera and non-Celera Genes Sequenced in 2001

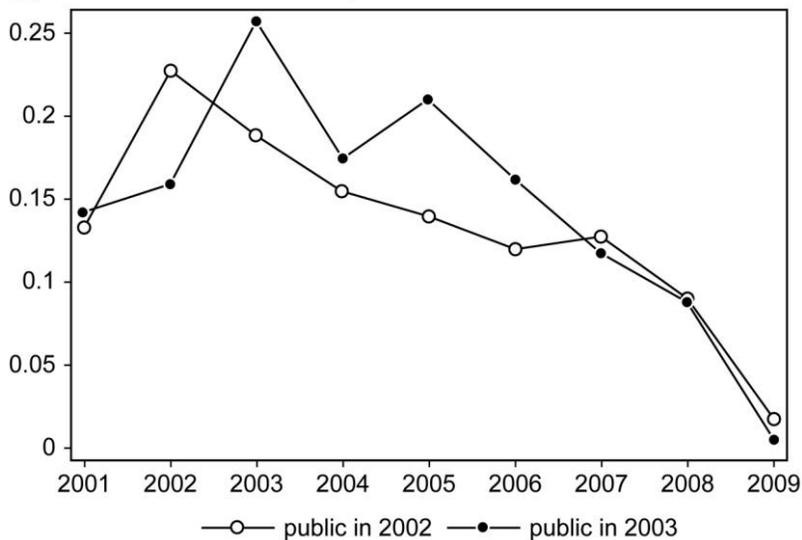
	Celera Mean (1)	Non-Celera Mean (2)	Difference [(1) – (2)] (3)	<i>p</i> -Value of Difference (4)
Publications in 2001–9	1.239	2.116	–.877	[.000]
1(known, uncertain phenotype)	.401	.563	–.162	[.000]
1(known, certain phenotype)	.046	.073	–.027	[.000]
1(used in any diagnostic test)	.030	.054	–.024	[.000]
Observations	1,682	2,851		

Notes: This reproduces table 1 from Williams (2013). This table compares subsequent innovation outcomes for Celera genes relative to non-Celera genes sequenced in the same year. Gene-level observations. The sample in col. 1 includes all Celera genes; the sample in col. 2 includes all non-Celera genes sequenced in 2001. The *p*-value reported in col. 4 is from a *t*-test for a difference in mean outcomes across cols. 1 and 2. See the text and online app. A for more detailed data and variable descriptions.

attention to within-gene variation in Celera's intellectual property and test whether the removal of Celera's intellectual property increased follow-on innovation on a given gene. Second, I limit the sample to Celera genes and test for a link between the amount of time a gene was held with Celera's intellectual property and follow-on innovation. Quantitative evidence in Williams (2013) suggests that both of these supplementary empirical tests appear to eliminate selection bias, and the estimates from these empirical tests provide evidence that is consistent with the estimates from the simple cross-tabulation presented in table 2.

Figure 5—which reproduces figure 5 from Williams (2013)—documents evidence from the second of these two supplementary analyses, limiting the sample to Celera genes and testing for a link between the amount of time a gene was held with Celera's intellectual property and follow-on innovation. The solid black lines plot mean follow-on innovation outcomes for Celera genes that were resequenced by the Human Genome Project in 2002 ($N = 1,047$), while the dashed lines plot mean follow-on innovation outcomes for Celera genes that were held with Celera's intellectual property for one additional year by nature of having been resequenced by the Human Genome Project in 2003 ($N = 635$). Panel A plots the average count of scientific publications linked to each gene by year. Panel B plots the mean of an indicator variable for whether genes had any conjectured phenotype relationship as of that year. An important distinction between the two panels is, hence, that panel A measures the flow of research investments by year, whereas panel B measures the stock of knowledge about a gene as of a given year.

(a) Outcome variable: Scientific publications



(b) Outcome variable: Conjectured phenotype

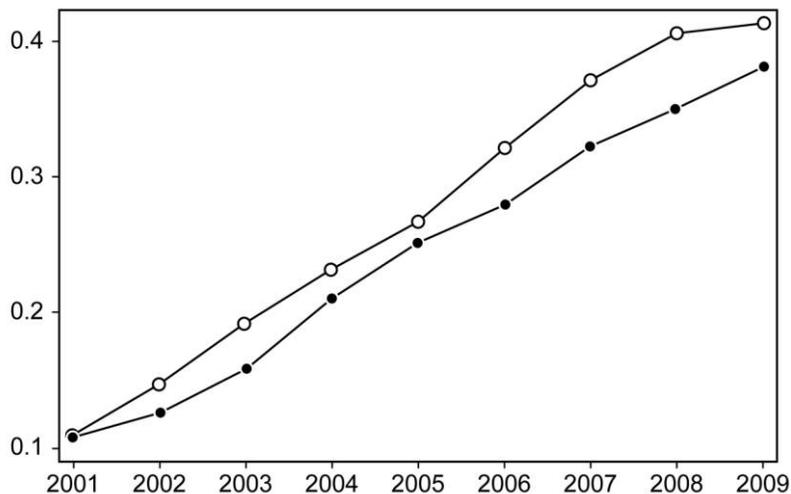


Fig. 5. Follow-on innovation for Celera genes by year and year of resequencing by the public effort

Notes: This reproduces figure 5 from Williams (2013). The outcome variable in panel A is the count of gene-level scientific publications in that year; the outcome variable in panel B is an indicator variable for whether the gene had at least one conjectured phenotype link as of that year. Means of these variables are shown separately for Celera genes that were resequenced by the public effort in 2002 ($N = 1,047$) and for Celera genes that were resequenced by the public effort in 2003 ($N = 635$).

Consistent with this exercise providing a clean “experiment,” the mean levels of follow-on innovation across these two cohorts of genes are similar in 2001 when both cohorts were held with Celera’s intellectual property. In panel A, we see that scientific publications on genes show a relative uptick in the year that they enter the public domain—2002 for the 2002 cohort and 2003 for the 2003 cohort. However, the flow of scientific effort into these two cohorts of genes seems to have converged over time: although the difference in means in 2002 is statistically significant at the 10% level, mean differences in other years are not statistically significant.

In panel B, we see that the stock of scientific knowledge also shows a relative uptick in 2002 for the 2002 cohort. However, the 2003 cohort shows persistently lower levels of this knowledge stock variable through the end of my data (the difference in means is statistically significant in 2003, 2006, 2007, and 2008). Hence, while the flow of scientific effort into these two cohorts of genes appears to have converged over time, the stock of scientific knowledge about these two cohorts of genes has not converged. The data cannot reject a model in which being held with Celera’s intellectual property for one additional year induces a permanent loss of one year of follow-on research, suggesting that even very temporary forms of intellectual property can have persistent effects on follow-on innovation.

D. Interpretation

Importantly, this analysis does not attempt to estimate the overall welfare effects of Celera’s intellectual property. Celera’s entry may have spurred faster sequencing of the human genome, and those benefits may have been much larger than the negative effects that I estimate of how Celera’s intellectual property affected follow-on innovation. Instead, my empirical estimates speak to a counterfactual in which Celera’s entry was held constant, in which case an alternative lump-sum reward mechanism may have had social benefits relative to Celera’s chosen package of intellectual property.

From a policy perspective, it is also important to relate these estimates to the ongoing legal controversy regarding patents on human genes. Despite not directly investigating the effects of gene patents, this Williams (2013) evidence was cited in several briefs submitted to the recent *Association for Molecular Pathology v. Myriad Genetics, Inc.* case in which the US Supreme Court invalidated patent claims on genomic de-

oxyribonucleic acid (DNA), concluding that “[g]roundbreaking, innovative, or even brilliant” discoveries of natural phenomena are not patentable given that such patents “would ‘tie up’ the use of such tools and thereby inhibit future innovation premised upon them.”¹³ That is, the US Supreme Court decision that human genes should be ineligible for patent protection rested on an argument that patents on human genes would deter follow-on innovation.

However, survey evidence such as Walsh, Arora, and Cohen (2003) suggests that unlike the nonpatent form of intellectual property that I studied in Williams (2013), patents on human genes may *not* discourage follow-on innovation, largely because patents (unlike Celera’s intellectual property) generally preserve open access to materials for academic scientists at the exploratory research stage. In ongoing joint work with Bhaven Sampat, we are investigating how gene patents impact follow-on innovation in order to speak directly to patent policy and to this US Supreme Court case (Sampat and Williams 2015). Specifically, we use newly collected data and two novel quasi-experimental approaches to investigate how gene patents affect follow-on innovation. First, we compare genes claimed in accepted and rejected patent applications. Second, we construct a new instrumental variable for which patent applications are granted patents, based on the “leniency” of the (conditionally randomly assigned) patent examiner. Both approaches suggest gene patents have *not* reduced follow-on innovation; in particular, we can reject declines in follow-on innovation on the order of my earlier estimates.

Taken at face value, the evidence from these two papers suggests that, at least in the context of human genes, the traditional patent trade-off (ex ante incentives versus deadweight loss) may be sufficient to analyze optimal patent policy design but that nonpatent policies governing access to materials may have important effects on follow-on innovation (consistent with the work of Murray et al. 2008; and Furman and Stern 2011). While I am, of course, hesitant to extrapolate beyond the specific case study of human genes, the results of these two papers also offer reason for concern about the recent *Association for Molecular Pathology v. Myriad Genetics, Inc.* ruling. Celera tried (but largely failed) to obtain patent protection on its sequenced genetic data, and in the absence of patent protection Celera used the nonpatent form of intellectual property described in the preceding as the next best available alternative means of capturing returns to their investment in sequencing the hu-

man genome. Ex post, it appears likely that both Celera and the social planner may have preferred for Celera to have been granted patents if those patents would have enabled open disclosure of Celera's data in the public domain, as it appears that such open disclosure would have enabled additional follow-on scientific research and commercial investments. Consistent with this case study, theoretical models such as Aghion, Dewatripont, and Stein (2008) and Murray et al. (2008) predict that private firms should prefer patents to other forms of intellectual property that restrict basic scientists' access to research materials. While the case study of human genes is at best suggestive, these results suggest that any welfare evaluation of rulings like the *Association for Molecular Pathology v. Myriad Genetics, Inc.* decision must carefully consider whether the elimination of patent protection may encourage firms to rely on other methods for protecting their intellectual property and to measure the *relative* costs and benefits of patent protection compared to these outside options.

V. Conclusion

A long theoretical literature has analyzed optimal patent policy design, yet there is very little empirical evidence on a key empirical parameter needed to apply these models in practice: namely, the relationship between patent strength and research investments. In my view, this dearth of empirical evidence reflects two challenges: the difficulty of measuring specific research investments and the fact that finding any variation (much less "clean" or quasi-experimental variation) in patent protection is difficult.

In this paper, I have summarized two of my recent investigations that have attempted to make progress in starting to overcome these empirical challenges.¹⁴ Methodologically, both papers make progress toward quantifying R&D investments that are missing due to misaligned incentives in the patent system by identifying a "risk set" of potential technologies that could have been developed. The key insight of this approach is that starting with data that only measures the observed set of technologies that have been developed will often be insufficient in the absence of some type of framework in which to interpret this data. By comparing the set of technologies that have been observed to be developed against a risk set of potential technologies that could have been developed, researchers can develop a framework for starting to

evaluate potential explanations for the gap between the set of observed and set of potential inventions. In the first paper, the census of cancer types was a useful risk set (Budish, Roin, and Williams 2015a); in the second paper, the census of human genes was a useful risk set (Williams 2013). Importantly, having a risk set does *not* address either of the two empirical challenges discussed in this paper—that is, measuring R&D investments and inferring the causal role of patent protection relative to other factors—but a risk set can, nonetheless, be a very important first step in empirically tackling the question of how patents affect research investments.

While I view both papers as making progress on the relevant policy questions in this area, each investigation faces important limitations in its ability to speak directly to patent policy. In the case of Budish, Roin, and Williams (2015a), we are able to provide evidence that private research investments are distorted away from long-term research projects but are unable to conclude that this distortion is generated by the patent system. However, our estimates are able to speak directly to the potential value of reducing commercialization lags for drug development. In the case of Williams (2013), I am able to document evidence that one nonpatent form of intellectual property on human genes (which a firm used when its patent applications were rejected by the USPTO) induced substantial and persistent decreases in follow-on scientific research and commercial product development. However, preliminary results from a follow-on project (Sampat and Williams 2015) suggest that this conclusion may not generalize to gene patents. Taken at face value, this pair of papers on gene patents is useful in highlighting that any welfare evaluation of rulings like the *Association for Molecular Pathology v. Myriad Genetics, Inc.* decision must take into account the *relative* costs and benefits of patent protection compared to the outside options that firms have available for protecting their intellectual property in the absence of patents. My hope is that future work in this area will be able to provide evidence to inform both applications of the Nordhaus-style framework to estimating optimal patent terms in practice and to inform specific policy decisions such as the recent set of US Supreme Court rulings relating to follow-on innovation.

Endnotes

I am very grateful to Navid Bazzazian, Dan Fetter, Josh Lerner, Jim Poterba, and, especially, Scott Stern for comments. Research reported in this publication was supported

by the National Institute on Aging (NIA) and the National Institutes of Health (NIH) Common Fund, Office of the NIH Director, through Grant U01-AG046708 to the National Bureau of Economic Research (NBER); the content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or NBER. Financial support from NIA Grant Number T32-AG000186 to the NBER, National Science Foundation (NSF) Grant Number 1151497, the Center for American Political Studies at Harvard, the Chicago Booth Initiative on Global Markets, the NBER Innovation Policy and the Economy program, and the Petrie-Flom Center at Harvard Law School is also gratefully acknowledged. For acknowledgments, sources of research support, and disclosure of the author's material financial relationships, if any, please see <http://www.nber.org/chapters/c13586.ack>.

1. On theoretical models of optimal patent breadth and length, see also Scherer (1972), Kaplow (1984), Klemperer (1990), and Gilbert and Shapiro (1990).

2. Ouellette (2015) argues that there could exist feasible opportunities to implement randomized policy experiments to inform patent policy.

3. His sample of 60 countries was chosen based on those listed in the International Monetary Fund's *International Financial Statistics* with the 60 highest total gross domestic products in 1997. For each of these countries, he then undertook archival research using sources including guidebooks to the world patent systems, publications of the world's patent offices, and legal monographs in order to document the following characteristics of patent policy regimes by country over time: whether the country offered comprehensive patent protection, the length of patents, the cost of awards, and provisions for patent revocation. This work identified 177 policy changes in 51 of the 60 countries in the sample.

4. Panels B and C of table 1 construct "adjusted" versions of these comparisons that control for changes in patent filings that are common across countries within a year, which only strengthens the lack of evidence of an increase in British patent applications by residents of the country where the policy change occurred.

5. Subsequent papers using this type of data include Yin (2008), Yin (2009), Kyle and McGahan (2012), Blume-Kohout and Sood (2013), Sampat and Williams (2015), Budish, Roin, and Williams (2015a), and Dubois et al. (2015).

6. As we will discuss, Budish, Roin, and Williams (2015a) offer one—albeit quite speculative—attempt to undertake this type of exercise for health outcomes.

7. This 1988 Japanese package of patent reforms also provided patent term restoration of up to five years to compensate for safety and efficacy delays required in the pharmaceutical industry, which the authors argue amounted to an effective extension of the length of patent rights in the pharmaceutical industry in Japan.

8. Consistent with this figure, the authors note that interviews with general managers in the intellectual property divisions of several pharmaceutical and chemical companies confirmed that they were quick to respond to the reform.

9. Specifically, figure 5 from Sakakibara and Branstetter (2001) plots the coefficients on year fixed effects from a regression relating log R&D spending at the firm-year level to Tobin's Q, the log of contemporaneous sales (as a measure of firm size), firm fixed effects, and year fixed effects.

10. In ongoing follow-up research, Budish, Roin, and I are pursuing a second joint project to more directly investigate the question of how patents impact research investments (Budish, Roin, and Williams 2015b).

11. There are some exceptions to this overly simplified definition of efficacy, which our empirical work takes advantage of.

12. Of course, a third challenge is that the assertion that there are "missing" research investments is difficult to distinguish from alternative explanations such as a lack of market demand or a lack of scientific opportunities; we discuss this inference challenge in detail in Section III.

13. See http://www.americanbar.org/content/dam/aba/publications/supreme_court_preview/briefs-v2/12-398_pet.authcheckdam.pdf and <http://isp.yale.edu/wp-content/uploads/2012/01/31253-BRCA-AMICUS-BRIEF-FINAL.pdf>, as well as Marshall (2013) for a recent discussion.

14. While this paper has focused on patents, it is worth noting that a growing literature is investigating analogous issues for copyright, including Lerner and Rafert (2014), Giorelli and Moser (2015), and Waldfogel (forthcoming).

References

- Abrams, David. 2009. "Did TRIPS Spur Innovation? An Analysis of Patent Duration and Incentives to Innovate." *University of Pennsylvania Law Review* 157 (6): 1613–47.
- Acemoglu, Daron, and Joshua Linn. 2004. "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry." *Quarterly Journal of Economics* 119 (3): 1049–90.
- Aghion, Philippe, Mathias Dewatripont, and Jeremy Stein. 2008. "Academic Freedom, Private-Sector Focus, and the Process of Innovation." *RAND Journal of Economics* 39 (3): 617–35.
- Bessen, James, and Eric Maskin. 2009. "Sequential Innovation, Patents, and Imitation." *RAND Journal of Economics* 40 (4): 611–35.
- Blume-Kohout, Margaret, and Neeraj Sood. 2013. "Market Size and Innovation: Effects of Medicare Part D on Pharmaceutical Research and Development." *Journal of Public Economics* 97 (1): 327–36.
- Boldrin, Michele, and David K. Levine. 2008. *Against Intellectual Monopoly*. New York: Cambridge University Press.
- Budish, Eric, Benjamin Roin, and Heidi Williams. 2015a. "Do Firms Underinvest in Long-Term Research? Evidence from Cancer Clinical Trials." *American Economic Review* 105 (7): 2044–85.
- . 2015b. "Missing Markets for Innovation: Evidence from New Uses of Old Drugs." Manuscript, MIT.
- Cockburn, Iain, Jean Lanjouw, and Mark Schankerman. 2014. "Patents and the Global Diffusion of New Drugs." Working Paper no. 20492, NBER, Cambridge, MA.
- Collins, Francis, and David Galas. 1993. "A New Five-Year Plan for the U.S. Human Genome Project." *Science* 262 (5130): 43–46.
- Collins, Francis, Ari Patrinos, Elke Jordan, Aravinda Chakravarti, Raymond Gesteland, LeRoy Walters, and the members of the DOE and NIH planning groups. 1998. "New Goals for the US Human Genome Project: 1998–2003." *Science* 282 (5389): 682–89.
- Cutler, David. 2004. *Your Money or Your Life: Strong Medicine for America's Health Care System*. Oxford: Oxford University Press.
- Dubois, Pierre, Olivier de Mouzon, Fiona Scott Morton, and Paul Seabright. 2015. "Market Size and Pharmaceutical Innovation." *RAND Journal of Economics* 46 (4): 844–71.
- Duggan, Mark, Craig Garthwaite, and Aparajita Goyal. 2014. "The Market Impacts of Pharmaceutical Product Patents in Developing Countries: Evidence from India." Working Paper no. 20548, NBER, Cambridge, MA.
- Finkelstein, Amy. 2004. "Static and Dynamic Effects of Health Policy." *Quarterly Journal of Economics* 119 (2): 527–67.
- Furman, Jeffrey, and Scott Stern. 2011. "Climbing Atop the Shoulders of Giants: The Impact of Institutions on Cumulative Research." *American Economic Review* 101 (5): 1933–63.
- Galasso, Alberto, and Mark Schankerman. 2015. "Patents and Cumulative In-

- novation: Causal Evidence from the Courts." *Quarterly Journal of Economics* 130 (1): 317–69.
- Gilbert, Richard, and Carl Shapiro. 1990. "Optimal Patent Length and Breadth." *RAND Journal of Economics* 21:106–12.
- Giorcelli, Michela, and Petra Moser. 2015. "Copyrights and Creativity: Evidence from Italian Operas." 2015. Manuscript, Stanford University.
- Heller, Michael, and Rebecca Eisenberg. 1998. "Can Patents Deter Innovation? The Anticommons in Biomedical Research." *Science* 280 (5364): 698–701.
- Hemphill, C. Scott, and Bhaven Sampat. 2011. "Evergreening, Patent Challenges, and Effective Market Life in Pharmaceuticals." *Journal of Health Economics* 31 (2): 327–39.
- Istrail, Sorin, Granger G. Sutton, Liliana Florea, Aaron L. Halpern, Clark M. Mobarry, Ross Lippert, Brian Walenz, et al. 2004. "Whole-Genome Shotgun Assembly and Comparison of Human Genome Assemblies." *Proceedings of the National Academy of Sciences* 101 (7): 1916–21.
- Jaffe, Adam, Manuel Trajtenberg, and Rebecca Henderson. 1993. "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." *Quarterly Journal of Economics* 108 (3): 577–98.
- Kaplow, Louis. 1984. "The Patent-Antitrust Intersection: A Reappraisal." *Harvard Law Review* 97:1815–92.
- Kesselheim, Aaron, Robert Cook-Deegan, David Winickoff, and Michelle Mello. 2013. "Gene Patenting—The Supreme Court Finally Speaks." *New England Journal of Medicine* 369 (9): 869–75.
- Klemperer, Paul. 1990. "How Broad Should the Scope of Patent Protection Be?" *RAND Journal of Economics* 21:113–30.
- Kyle, Margaret, and Anita McGahan. 2012. "Investments in Pharmaceuticals before and after TRIPS." *Review of Economics and Statistics* 94 (4): 1157–72.
- Kyle, Margaret, and Yi Qian. 2014. "Intellectual Property Rights and Access to Innovation: Evidence from TRIPS." Working Paper no. 20799, NBER, Cambridge, MA.
- Lander, Eric, L. M. Linton, B. Birren, C. Nusbaum, M. C. Zody, J. Baldwin, K. Devon, et al. 2001. "Initial Sequencing and Analysis of the Human Genome." *Nature* 409 (6822): 860–921.
- Lathia, C. D., D. Amakye, W. Dai, C. Girman, S. Madani, J. Mayne, P. MacCarthy, et al. 2009. "The Value, Qualification, and Regulatory Use of Surrogate End Points in Drug Development." *Clinical Pharmacology & Therapeutics* 86 (1): 32–43.
- Lerner, Joshua. 2009. "The Empirical Impact of Intellectual Property Rights on Innovation: Puzzles and Clues." *American Economic Review Papers and Proceedings* 99 (2): 343–48.
- Lerner, Joshua, and Greg Rafert. 2014. "Lost in the Clouds: The Impact of Changing Property Rights on Investment in Cloud Computing Ventures." Manuscript, Harvard Business School. <http://www.people.hbs.edu/jlerner/lost-in-the-clouds-05042014.pdf>.
- Marshall, Eliot. 2013. "Lock up the Genome, Lock down Research." *Science* 342:72–73.
- Meyskens, Frank, Gregory Curt, Dean Brenner, Gary Gordon, Ronald Herberman, Olivera Finn, Gary Kelloff, Samir Khleif, Caroline Sigman, and Eva Szabo. 2011. "Regulatory Approval of Cancer Risk-Reducing (Chemopreventive) Drugs: Moving What We Have Learned into the Clinic." *Cancer Prevention Research* 4 (3): 311–23.

- Moser, Petra. 2005. "How Do Patent Laws Influence Innovation? Evidence from Nineteenth-Century World's Fairs." *American Economic Review* 95 (4): 1214–36.
- Moser, Petra, Joerg Ohmstedt, and Paul Rhode. 2014. "Patent Citations and the Size of Innovation: Evidence from Hybrid Corn." Manuscript, Stanford University.
- Murray, Fiona, Philippe Aghion, Mathias Dewatripont, Julian Kolev, and Scott Stern. 2008. "Of Mice and Academics: Examining the Effect of Openness on Innovation." Manuscript, MIT.
- Murray, Fiona, and Scott Stern. 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): 648–87.
- Nordhaus, William. 1969. *Invention, Growth, and Welfare: A Theoretical Treatment of Technological Change*. Cambridge, MA: MIT Press.
- Ouellette, Lisa Larrimore. 2015. "Patent Experimentalism." *Virginia Law Review* 101:65.
- Pennisi, Elizabeth. 1999. "Human Genome: Academic Sequencers Challenge Celera in a Sprint to the Finish." *Science* 283 (5409): 1822–23.
- Psaty, Bruce, Noel Weiss, Curt Furberg, Thomas Koepsell, David Siscovick, Frits Rosendaal, Nicholas Smith, et al. 1999. "Surrogate End Points, Health Outcomes, and the Drug-Approval Process for the Treatment of Risk Factors for Cardiovascular Disease." *Journal of the American Medical Association* 282 (8): 786–90.
- Sakakibara, Mariko, and Lee Branstetter. 2001. "Do Stronger Patents Induce More Innovation? Evidence from the 1988 Japanese Patent Law Reforms." *RAND Journal of Economics* 32 (1): 77–100.
- Sampat, Bhaven, and Heidi Williams. 2015. "How Do Patents Affect Follow-On Innovation? Evidence from the Human Genome." Manuscript, MIT.
- Scherer, F. M. 1972. "Nordhaus' Theory of Optimal Patent Life: A Geometric Reinterpretation." *American Economic Review* 62 (3): 422–27.
- Science Online. 2001. "Accessing the Celera Human Genome Sequence Data." <http://www.sciencemag.org/feature/data/announcement/gsp.dtl>.
- Scotchmer, Suzanne. 1991. "Standing on the Shoulders of Giants: Cumulative Research and the Patent Law." *Journal of Economic Perspectives* 5 (1): 29–41.
- . 1996. "Protecting Early Innovators: Should Second-Generation Products be Patentable?" *RAND Journal of Economics* 27 (2): 322–31.
- . 2004. *Innovation and Incentives*. Cambridge, MA: MIT Press.
- Service, Robert. 2001. "Can Data Banks Tally Profits?" *Science* 291 (5507): 1203.
- Shreeve, James. 2005. *The Genome War: How Craig Venter Tried to Capture the Code of Life and Save the World*. New York: Ballantine Books.
- Venter, J. Craig, M. D. Adams, E. W. Myers, P. W. Li, R. J. Mural, G. G. Sutton, H. O. Smith, et al. 2001. "The Sequence of the Human Genome." *Science* 291 (5507): 1304–51.
- Venter, J. Craig, Mark Adams, Granger Sutton, Anthony Kerlavage, Hamilton Smith, and Michael Hunkapiller. 1998. "Shotgun Sequencing of the Human Genome." *Science* 280 (5369): 1540–42.
- Wade, Nicholas. 2003. "Once Again, Scientists Say Human Genome is Complete." *New York Times*, April 15.
- Waldfoegel, Joel. Forthcoming. "Copyright Protection, Technological Change, and the Quality of New Products: Evidence from Recorded Music since Napster." *Journal of Law and Economics*.

- Walsh, John, Ashish Arora, and Wesley Cohen. 2003. "Working through the Patent Problem." *Science* 299 (5609): 1021.
- Weisfelt, Myron, and Susan Ziemann. 2007. "Advances in the Prevention and Treatment of Cardiovascular Disease." *Health Affairs* 26 (1): 25–37.
- Williams, Heidi. 2013. "Intellectual Property Rights and Innovation: Evidence from the Human Genome." *Journal of Political Economy* 121 (1): 1–27.
- Yin, Wesley. 2008. "Market Incentives and Pharmaceutical Innovation." *Journal of Health Economics* 27 (4): 1060–77.
- . 2009. "R&D Policy, Agency Costs and Innovation in Personalized Medicine." *Journal of Health Economics* 28 (5): 950–62.