

Williams, Heidi

Article

What inventions are we missing?

NBER Reporter

Provided in Cooperation with:

National Bureau of Economic Research (NBER), Cambridge, Mass.

Suggested Citation: Williams, Heidi (2019) : What inventions are we missing?, NBER Reporter, National Bureau of Economic Research (NBER), Cambridge, MA, Iss. 1, pp. 18-21

This Version is available at:

<https://hdl.handle.net/10419/219432>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



Heidi Williams is a research associate in the NBER's Aging, Health Care, and Productivity Programs. She is an associate professor of economics at MIT, and is currently Trione Visiting Associate Professor at the Stanford Institute for Economic Policy Research. She is co-editor of the *Journal of Economic Perspectives*.

Her research focuses on investigating the causes and consequences of technological change, with a particular focus on health care markets. Williams received her AB in mathematics from Dartmouth College in 2003, her MSc in development economics from Oxford University in 2004, and her PhD in economics from Harvard in 2010. Her research has received support from the National Institutes of Health and an NSF CAREER grant, among others, and she has received a Sloan Research Fellowship and a MacArthur Foundation Fellowship.

Williams grew up in North Dakota, but has most recently been living in Cambridge and Palo Alto with her husband and two young sons.

What Inventions Are We Missing?

Heidi Williams

Academics and policymakers have long recognized that competitive markets may under-incentivize innovation. This concern has motivated the design of public policies such as the patent system, which aims to encourage research investments into new technologies by allowing inventors to capture a higher share of the social returns to their research investments.

A well-developed theoretical literature has analyzed optimal patent policy, with a focus on the trade-off between providing incentives for the development of new technologies and tolerating higher prices during the life of the patent. Although such theoretical models—and, importantly, public policies—typically assume that stronger (e.g. longer or broader) patents will induce additional research investments, there is remarkably little empirical evidence on how patents affect research investments in practice.

This question has been difficult to tackle empirically for at least two reasons. First, measuring research investments can be quite challenging. Second, finding variation in patent protection that can be leveraged in an empirical study is difficult. On paper, the U.S. patent system is uniform, providing a 20-year period of protection for all inventions. While historically some cross-country variation in patent laws has existed, because innovations are generally developed for a global market, country-specific patent law changes may often induce only a relatively small change in global research incentives.

My research agenda attempts to overcome both of these challenges in order to develop empirical estimates of the key parameters needed to inform optimal patent policy. By combining detailed new measures of research investments with novel sources of variation in the *effective* patent terms provided to otherwise similar inventions, my work aims to construct frameworks within which we can infer the volume, type, and value of “missing” research investments that

would have occurred under counterfactual patent policies. In this piece, I summarize some of the main findings that have emerged from my research in this area.

Measuring Innovation

Traditionally, economists who study innovation have relied on patent counts (or citation-weighted patent counts) as a measure of innovation, often leveraging the data constructed by Bronwyn Hall, Adam Jaffe, and Manuel Trajtenberg.¹ Although this approach has been useful in many settings, it encounters two major limitations. First, in many cases it is difficult or impossible to match patents with the specific products they protect, or to identify specific groups of consumers that might benefit from those products. For example, the text in a patent protecting a delivery method for a breast cancer drug may have no information suggesting the patent is relevant to breast cancer patients. Hence, it can be very difficult to use patents to measure research investments in a way that can be linked to product-market or consumer-level outcomes. Second, by construction, patent data can only measure patented inventions. Because many technologies are not patented, changes in patent counts may in some settings reflect changing levels of inventors' willingness to file for patents on their research investments, rather than changes in the underlying research investments themselves.

A major focus of my research agenda has been to attempt to overcome these two challenges by compiling “real,” non-patent measures of innovation. For example, Eric Budish, Ben Roin, and I aimed to develop measures of research investments in cancer drugs.² The core of our data construction was to take advantage of a National Cancer Institute (NCI) clinical trial registry that includes an explicit listing of the patient groups eligible to enroll in each

clinical trial. Cancer treatment tends to be specific to an organ of origin, such as prostate, and stage of disease, for example, metastatic. As such, the organ-stage classification tends to be used both to label clinical trial-eligible groups (as in the NCI data, where such classifications are used to describe which patients can enroll in any given clinical trial) and to label patients in standard clinical datasets (e.g. the Surveillance, Epidemiology, and End Results cancer registry data, which provides data on the survival outcomes of U.S. cancer patients). Leveraging this organ-stage classification hence allowed us to use clinical trials as a real measure of research investments in new cancer drugs, and to link those research investments to measures of patient health outcomes—namely, survival outcomes for different groups of cancer patients over time.

As a second example, two of my papers—one joint with Bhaven Sampat—have constructed data on non-patent measures of innovation related to the human genome.³ The set of human genes is curated by scientists in a way that assigns the equivalent of Social Security numbers—unique identifiers for each gene—called Entrez Gene IDs. These Entrez Gene IDs are in turn linkable to various databases which catalog scientific papers that have documented evidence for links between genes and diseases; this “paper trail” provides a consistent measure of scientific publications related to a given gene. Entrez Gene IDs can also be linked to product market databases such as the GeneTests.org database, which provides information on the use of genes in gene-based diagnostic tests, and the Pharmaprojects database, which provides information on drug compounds in clinical

trials that relate to specific genetic variations. These types of curated scientific identifiers can thus provide unique opportunities to trace out meaningful links between basic scientific discoveries and commercialized products.

Our goal in constructing these data is to apply them in order to test and evaluate theories about economic factors that may be encouraging or hindering innovation. Let me now summarize some of the substantive findings of our research.

Incentives to Develop New Drugs

Why don't we have a cure for cancer? Informal discussions with doctors and scientists usually provide a variety of answers to this question, many of which boil down to some version of “the basic science is hard.” But from an economic perspective, we are of course inclined to think that how “hard” any given scientific problem seems might reflect not only the innate difficulty of the problem, but also the level of past research efforts to solve it.

Budish, Roin, and I investigated one aspect of this question—namely, whether private firms underinvest in cancer drug development projects that require a long time to complete.⁴ The basic idea of our study was this:

Pharmaceutical firms face strong incentives to file patent applications at the time of invention. However, prior to marketing a new drug, firms must submit clinical trial results to the U.S. Food and Drug Administration (FDA) documenting that their product meets a set of safety and efficacy standards. These clinical trials generate a lag between the time of patenting and the time of commercialization, which reduces an invention's effective patent life—more so for drugs that require longer clinical trials. A key determinant of clinical trial length is patient survival time: Because clinical trials generally must show evidence that the treatment improves mortality-related outcomes, clinical trials tend to be longer—and hence, effective patent terms shorter—when enrolling patients with longer survival times. All else being equal, firms are thus awarded longer terms of market exclusivity for successfully developing drugs to treat patients with shorter survival times (such as patients with late stage, metastatic cancers) relative to drugs to treat patients with longer survival times (such as patients with early stage, localized cancers). This motivates the following empirical test: If longer patent terms encourage more research investments, we should see higher levels of private research investments on treatments for patient groups that require shorter clinical trials.

Using the cancer clinical trial registry data described above, we document several sources of empirical evidence that together are consistent with private research investments being distorted away from long-term projects such as drugs to prevent or treat early-stage cancers. For example, Figure 1, reproduced from our paper, documents that research on cancers which require longer

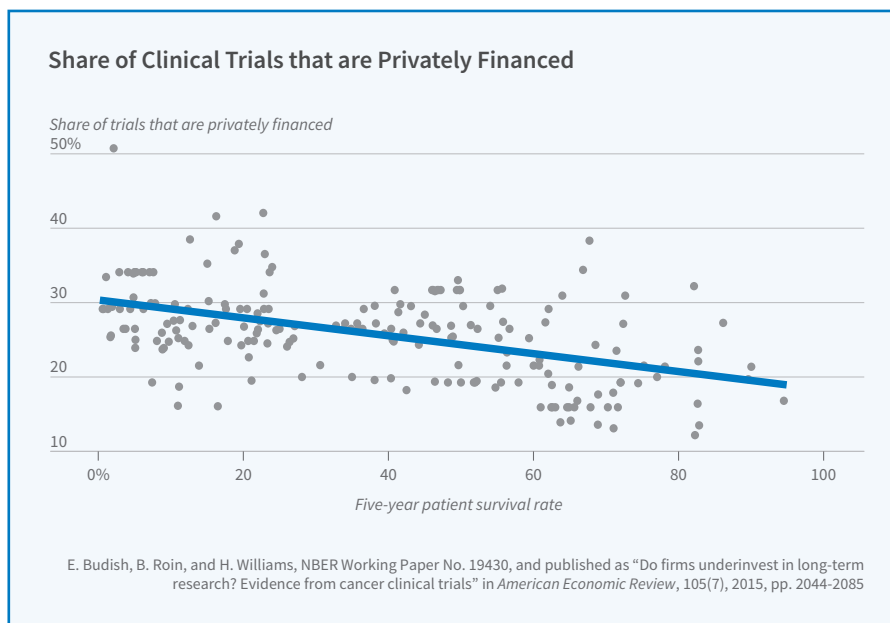


Figure 1

clinical trials because they have higher five-year survival rates is more likely to be publicly funded. A back-of-the-envelope calculation we present suggests that this distortion has quantitatively important implications for the survival outcomes of U.S. cancer patients.

Unfortunately, our estimates cannot speak directly to the potential role of patents, since excess impatience of private firms may also under-incentivize long-term research.⁵ Our empirical evidence is nonetheless directly relevant to at least two policy levers. First, to the extent that valid surrogate endpoints—non-survival endpoints that are known to be causally linked with subsequent survival improvements but which can generally be observed more quickly than can changes in survival outcomes—are available, our evidence suggests that allowing clinical trials to rely on valid surrogate endpoints can meaningfully increase research investments and substantially improve patient health outcomes. Second, our results suggest a rationale for targeting public research subsidies toward projects that are expected to have long commercialization lags, such as many Alzheimer’s drugs.

Human Genomes: Public and Private

The prediction that stronger patent protection induces additional research investments emerges unambiguously from a class of theoretical models that treat innovations as isolated discoveries. In practice, however, innovation is often “cumulative,” that is, any given discovery is also an input into later, follow-on discoveries. In such cases, optimal patent policy will depend in part on how patents on existing technologies affect

follow-on innovation. A well-developed literature has documented theoretically ambiguous predictions on how patents affect follow-on innovation, but there is little available empirical evidence.⁶ I here describe two papers that aim to shed light on this question in the context of the human genome.

In 2001, the journals *Nature* and *Science* published the initial sequences of the human genome: *Nature* published a version completed by the publicly funded Human Genome Project, and *Science* published a version completed by a private firm, Celera.⁷ As

rounded the question of whether Celera’s data not being open access would hinder subsequent scientific research and product development.

I collected records of when each human gene was sequenced by Celera and by the Human Genome Project,⁹ which indicates whether a gene was ever held with Celera’s intellectual property. Genes sequenced first by Celera were held with Celera’s intellectual property until the gene was re-sequenced by the Human Genome Project, at which time it moved into the public domain. I linked these records to the types of gene-level innovation measures described above. Using these data, I documented that Celera’s property rights reduced subsequent scientific research and product development by approximately 30 percent.

Taken at face value, these results clearly suggest that data access rules can have quantitatively important effects on the rate of innovation. Because the timing of that study coincided with a high-profile U.S. Supreme Court case (*AMP v.*

Myriad) that ruled on whether human genes should qualify as patentable subject matter, a question naturally arose of whether the negative effects of Celera’s non-patent intellectual property would be similar to the effects of gene patents. Sampat and I directly investigated this question.¹⁰

We proposed two quasi-experimental approaches for investigating how gene patents affect follow-on innovation. First, we compared follow-on innovation on genes claimed in accepted versus rejected patent applications. Figure 2 above, reproduced from our paper, shows that genes included in accepted and rejected patent applications were the

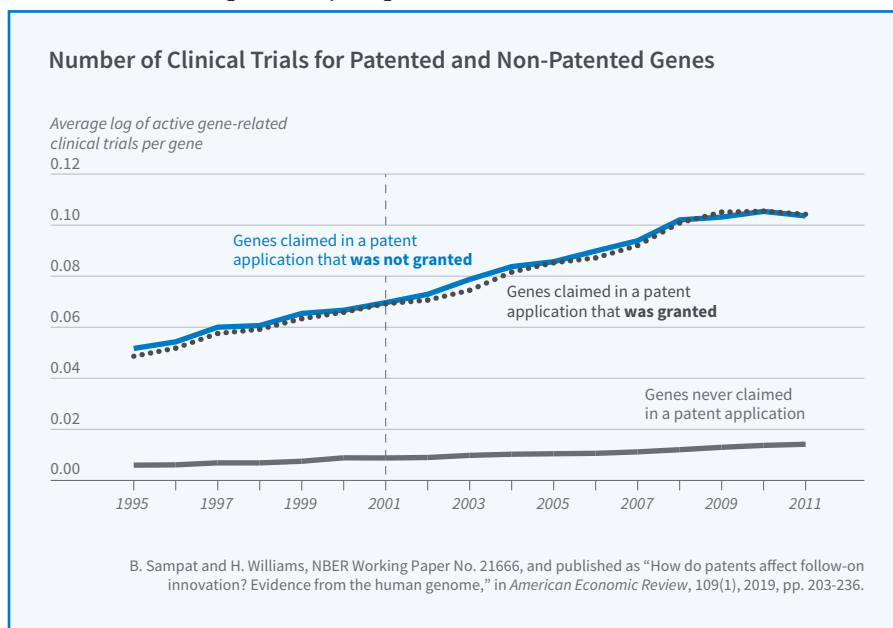


Figure 2

was described in an accompanying editorial in *Nature*, the rules surrounding data access differed across the two efforts.⁸ The Human Genome Project placed all of its sequenced data in an open-access database within 24 hours of being sequenced, with the stated goal of maximizing the data’s benefit to society. Celera instead chose to protect its data with a contract law-based form of intellectual property, which allowed free use of the data for academic research, but placed restrictions on redistribution of the data and required licensing agreements to be negotiated for any downstream discoveries. As emphasized in the *Nature* editorial, much public debate sur-

subject of a similar number of clinical trials both before and after the relevant patent applications were filed. The earliest patent application filing date in our sample is denoted by the vertical line.¹¹

Second, we constructed an instrumental variable for whether a patent application is granted a patent, based on the “leniency” of the conditionally randomly assigned patent examiner. In contrast with the effects I observed with Celera’s intellectual property, both of these approaches suggest that gene patents have *not* reduced follow-on innovation; in particular, we can statistically reject declines in follow-on innovation on the order of my earlier estimates for Celera’s intellectual property.

Both of the quasi-experimental approaches developed in this paper have already been re-applied in a number of other papers, by myself as well as other authors, and I have been pleased to see refinements and improvements of them being developed as well, as in the work of Josh Feng and Xavier Jaravel.¹²

Taken together, the evidence from this set of papers suggests that the traditional patent trade-off (*ex ante* incentives versus deadweight loss) may be sufficient to analyze optimal patent policy design at least in some markets, but that non-patent policies governing access to materials — such as data exclusivity — may have important effects on follow-on innovation.

¹ B. Hall, A. Jaffe, and M. Trajtenberg, “The NBER Patent Citation Data File: Lessons, Insights, and Methodological Tools,” NBER Working Paper No. 8498, October 2001, and in A. Jaffe and M. Trajtenberg, eds., *Patents, Citations, and Innovations*, Cambridge, Massachusetts, The MIT Press, 2002.
[Return to Text](#)

² E. Budish, B. Roin, and H. Williams, “Do Firms Underinvest in Long-Term Research? Evidence From Cancer Clinical Trials,” NBER Working Paper No. 19430, September 2013, and *American Economic Review*, 105(7), 2015, pp. 2044–85.
[Return to Text](#)

³ H. Williams, “Intellectual Property Rights and Innovation: Evidence From the Human Genome,” NBER Working Paper No. 16213, July 2010, and *Journal of Political Economy*, 121(1), 2013, pp. 1–27; B. Sampat and H. Williams, “How Do Patents Affect Follow-On Innovation? Evidence From the Human Genome,” NBER Working Paper No. 21666, October 2015, and *American Economic Review*, 109(1), 2019, pp. 203–36.
[Return to Text](#)

⁴ *Ibid*, NBER Working Paper No. 19430.
[Return to Text](#)

⁵ See, e.g., J. Stein, “Agency, Information, and Corporate Investment,” NBER Working Paper No. 8342, June 2001, and in G. Constantinides, M. Harris, and R. Stulz, eds., *Handbook of the Economics of Finance*, Amsterdam, The Netherlands: Elsevier B.V., (1A), 2003, pp. 111–65; J. Stein, *Quarterly Journal of Economics*, 104(4), 1989, pp. 655–69.
[Return to Text](#)

⁶ For one recent paper that is an exception to this statement, see A. Galasso and M. Schankerman, “Patents and Cumulative Innovation: Causal Evidence From the Courts,” NBER Working Paper No. 20269, July 2014, and *Quarterly Journal of Economics*, 130(1), 2015, pp. 317–69.
[Return to Text](#)

⁷ International Human Genome Sequencing Consortium, “Initial

Sequencing and Analysis of the Human Genome,” *Nature*, 409(6822), 2001, pp. 860–21; J. Venter, et al., “The Sequence of the Human Genome,” *Science*, 291(5507), 2001, pp. 1304–51.
[Return to Text](#)

⁸ “Human Genomes, Public and Private,” *Nature*, 409(6822), 2001, p. 745.
[Return to Text](#)

⁹ *Ibid*, NBER Working Paper No. 16213.
[Return to Text](#)

¹⁰ *Ibid*, NBER Working Paper No. 21666, and *American Economic Review*, 109(1), 2019, pp. 203–36.
[Return to Text](#)

¹¹ The lower line in Figure 2 shows that genes that were never included in any patent applications were the subject of many fewer clinical trials over this entire period, suggesting that — perhaps unsurprisingly — there was positive selection of genes into patent applications.
[Return to Text](#)

¹² J. Farre-Mensa, D. Hegde, and A. Ljungqvist, “What Is a Patent Worth? Evidence From the U.S. Patent ‘Lottery,’” NBER Working Paper No. 23268, December 2018; P. Kline, N. Petkova, H. Williams, and O. Zidar, “Who Profits from Patents? Rent-Sharing at Innovative Firms,” NBER Working Paper No. 25245, November 2018, and forthcoming in the *Quarterly Journal of Economics*; P. Gaule, “Patents and the Success of Venture-Capital Backed Startups: Using Examiner Assignment to Estimate Causal Effects,” 2015, unpublished CERGE-EI Working Paper, Series No. 546; J. Feng and X. Jaravel, “Crafting Intellectual Property Rights: Implications for Patent Assertion Entities, Litigation, and Innovation,” 2016, unpublished Harvard mimeo.
[Return to Text](#)