11-2023

Measuring follow-on innovation

Janet Freilich

Sepehr Shahshahani

Measuring Follow-On Innovation

Janet Freilich* Sepehr Shahshahani†‡

November 4, 2022

Abstract

How patents affect follow-on innovation is a key question for the patent system. We disaggregate follow-on innovation into activities that infringe patents and others that do not infringe but can be indirectly affected by patents. Replicating an important study using our disaggregated measure, we find that 87 percent of follow-on innovation is not patent infringement. Supplementing the study's empirical strategy with data on patent expiration dates, we find that gene patents which are not close to expiration cause an increase in noninfringing follow-on research, but the effect disappears for patents close to expiration. Our nuanced measure helps better identify the mechanisms of patents' effect, reconcile disparate results in the literature, and evaluate policy reform. (JEL O31, O34, K20)

*Fordham Law School. jfreilich1@fordham.edu.

†Fordham Law School. sshahshahani@fordham.edu.

‡We thank Bhaven Sampat and Heidi Williams for generously sharing their data. We also thank Sharon Belenzon, Ben Hammond, Deepak Hegde, Daniel Hemel, Jeffrey Kuhn, Michael Meurer, Lisa Larrimore Ouellette, Leah Rosenstiel, Bhaven Sampat, David Schwartz, Heidi Williams, Patricia Zettler, Samantha Zyontz, and participants in the 2021 Intellectual Property Scholars Conference, the 2022 NYU Tri-State IP Workshop, the 2022 American Law and Economics Associations Annual Meeting, and the 2022 Law & STEM Junior Faculty Forum for their helpful comments. Finally, we thank JC Martinez for excellent research assistance.

The patent system is a primary regulatory mechanism to incentivize scientific and technological innovation, a function it performs by awarding exclusive rights over an invention for a limited time. But exclusivity comes at a price because it makes access to the patented invention costlier. The central tradeoff of the patent regime is thus to balance innovator incentives against access costs. The access side of the tradeoff applies not only to end use of patented inventions but also to follow-on innovation, meaning innovation building on earlier innovations. The latter aspect is of critical importance in assessing patent policy because all innovation is really follow-on innovation.

Scholars have long been concerned with patent law's tradeoff between initial and follow-on innovation, particularly because a great deal of follow-on innovation appears to fall within the scope of a patent (Scotchmer (1991)). Recent influential empirical work has used creative identification strategies to ascertain the effects of patents on follow-on innovation (e.g., Murray and Stern (2007); Williams (2013); Galasso and Schankerman (2015); Murray et al. (2016); Sampat and Williams (2019)).

The most straightforward effect of patents on follow-on innovation occurs when the follow-on activity itself infringes the patent (what we call "direct" effects). But this is not the only mechanism for patents to affect follow-on innovation. For instance, patents may spur development of substitute, noninfringing technologies; disclosure provided in the patent document might encourage innovation; the possibility that a commercial embodiment of noninfringing research may infringe a patent could discourage the research; or patent holders may invest resources in the patented field to increase follow-on innovation (what we call "indirect" effects).[1] Existing empirical studies provide insight into the overall impact of patents on follow-on innovation but do not distinguish between follow-on activities that infringe the patent and those that may be affected by indirect mechanisms. Previous work generally cannot, therefore, isolate different mechanisms by which patents affect follow-on innovation.

Our contribution is to disaggregate follow-on innovation into activities that are (1) directly affected by a patent, meaning they may constitute patent infringement, and (2) indirectly affected by a patent, meaning that while they are not patent infringement, they

[1] Most of these mechanisms can also affect innovative activities that infringe the patent.

may plausibly be affected by upstream patents. Further, we explore different pathways through which patents may indirectly affect follow-on innovation. This approach provides a more nuanced understanding of follow-on innovation, sheds greater light on mechanisms by which patents affect follow-on innovation, and allows novel evaluation of various policy proposals.

We distinguish between direct and indirect follow-on innovation by exploiting legal doctrines that determine whether an activity is infringing depending on where, by whom, and for what purpose it was conducted. Previous empirical studies have generally defined follow-on innovation as the set of activities related to a patented technology. However, this approach includes many activities that clearly do not constitute patent infringement. The most important categories of clearly noninfringing activities (with respect to US patents) are those by non-US persons and by state-affiliated persons such as state universities.[2] Other noninfringing activities that are commonly counted as follow-on innovation include activities generating information for submission to the Food and Drug Administration. Previous empirical studies group these types of follow-on research together; we provide new insight by disaggregating them.

First, we show that surprisingly little of what is traditionally considered follow-on research actually infringes a patent. We replicate the important recent study by Sampat and Williams (2019) which investigates the effect of gene patents on follow-on research on those genes and find both that 87 percent of what was conceptualized as follow-on innovation is in fact not patent infringement and that patents have little effect on the remaining follow-on innovation. The fact that much, if not most, of the activities typically conceptualized as follow-on innovation are in fact not patent infringement shows that the direct effect of patents operates through a much narrower set of activities than previously recognized in the economics literature. This suggests that, as it relates to direct effects, the tradeoff between incentives for early technological development and later follow-on innovation may not be as stark as theorized. To our knowledge, ours is the first study to use a targeted measure of follow-on innovation and thus the first to quantify the scale of infringing follow-on innovation, the extent to which it is affected by patents, and the

[2] As is common in law, we use "person" to denote both individuals and entities.

degree to which direct effects are a plausible mechanism for patents to affect follow-on innovation.

Our approach thus sheds new light on one mechanism by which patents may affect follow-on innovation: infringement. In addition, we can better study how patents may *indirectly* affect follow-on innovation by isolating and removing direct effects. There are many indirect effects and our empirical context is not suited to studying them all, nor can we provide definitive answers about indirect mechanisms. Our goal here, consequently, is to illustrate how disaggregating follow-on innovation helps in exploring different causal mechanisms rather than to provide definitive answers or a thorough survey of indirect effects.

We explore two different pathways to indirect effects. Using the set of *non*infringing follow-on publications in Sampat and Williams's dataset, we (1) ask whether the effect of a patent on noninfringing follow-on research is attenuated as the patent gets closer to its expiration date, and (2) investigate whether patents have a greater effect on noninfringing follow-on research that is closer to commercialization by exploring the difference in effects on human, animal, and *in vitro* research. On the first question, we find that gene patents that are not expiring soon have a positive and statistically significant effect on noninfringing follow-on research, but the effect disappears for patents that are expiring soon. That is, genes protected by patents that are not close to expiration have statistically significantly greater follow-on publications than genes that are not protected by any patents, but the difference in follow-on publications between genes with soon-expiring patents and genes with not-soon-expiring patents is negative (and statistically significant in most specifications). These results support the firm theoretical prediction that the indirect effect of a patent should be greater when it is farther from expiration; they also show that the net indirect effect of not-soon-expiring patents on follow-on research is positive. On the second question, we do not find a statistically significant indirect effect for any type of research, nor a statistically significant difference in indirect effects between different types of research.

An added benefit of our nuanced approach to measuring follow-on innovation is that it can help evaluate how changes to the patent system might affect follow-on research.

For example, in recent decades, US courts eliminated an exemption from infringement for academic researchers (*Madey v. Duke University*, 307 F.3d 1351 (Fed. Cir. 2002)). This decision has been criticized on the ground that it negatively affects follow-on innovation (see Malakoff (2003)). But whether or not that is true depends on the mechanism by which patents affect follow-on innovation. Eliminating (or restoring) an exception for academic researchers targets only the direct effect of patents and is impactful only if (1) there is a meaningful amount of infringement by research, and (2) patents affect such infringement. Our results in the gene patent context suggest that neither is true (although, on the second point, we cannot rule out the possibility that patents have both positive and negative effects that counteract each other), which is consistent with other work finding that changing the research exception did not greatly affect follow-on research (Walsh, Arora and Cohen (2003)). We use our disaggregated measure of follow-on innovation to explore this and other policy questions.

Section 1 connects our work to the literature. Section 2 explains the legal basis for our claim that certain types of follow-on innovation are not patent infringement. We support our claims by reference to primary sources (statutes and court decisions) to make them transparent and verifiable. Section 3 explores the direct effect of patents on follow-on innovation. We replicate Sampat and Williams (2019) using a refined measure of follow-on innovation, in the process showing how researchers may construct the set of follow-on innovations to better target the hypothesized mechanism for how an upstream patent may affect downstream activities. Our reanalysis does not overturn Sampat and Williams's results, but it shows that the net direct effect of patenting on follow-on innovation is smaller than the net overall effect they found. It also shows that surprisingly little follow-on innovation is directly affected by the patent, with implications for theorizing the tradeoff between initial and follow-on innovation. Section 4 then explores the indirect effect of patents, including our finding that gene patents far from expiration increase noninfringing follow-on innovation. Section 5 discusses the implications of our analysis, including how it applies to several important policy questions on design of the patent system and how it helps explain discrepancies among several prominent empirical studies.

1 Literature

The fundamental incentive-access tradeoff of intellectual property has been theoretically appreciated in economics for a long time (e.g., Plant (1934); Nordhaus (1969)) and in law for an even longer time (e.g., *Wheaton v. Peters*, 33 U.S. 591 (1834); *Kendall v. Winsor*, 62 U.S. 322 (1858)). Also theoretically well-known are the challenges that cumulative innovation pose for the optimal design of the patent system (e.g., Scotchmer (1991); Green and Scotchmer (1995); Hopenhayn, Llobet and Mitchell (2006); Bessen and Maskin (2009)).[3] More recently, empirical work has begun to reveal in greater depth the workings of the patent system,[4] including by studying the effect of patents on follow-on innovation.

Ascertaining the effects of patents on follow-on innovation is beset by many challenges, not the least of which is the selection of more valuable innovations into patenting. Sampat and Williams (2019) find that although patented human genes (i.e., genes containing patented DNA sequences) were subject to greater follow-on innovation than unpatented genes prior to patenting, there is no meaningful difference in pre-patent follow-on innovation between patented and unpatented genes conditional on being included in a patent application. The authors therefore compare post-patent follow-on innovation between patented and unpatented genes that were included in a patent application, finding no meaningful effect of patenting on follow-on innovation. They supplement this analysis by using the leniency of patent examiners as an instrument for patent grant. The instrumental-variable analysis shows a negative but very small effect of patenting on follow-on innovation. Sampat and Williams use three measures of follow-on innovation: articles published in scientific journals that discuss research on the patented genes, diag-

[3] Other theoretical work on intellectual property protection includes, for example, Loury (1979) and Lee and Wilde (1980) (patent races); Reinganum (1982) (firms' dynamic optimal resource allocation to R&D in the presence of patents); Katz and Shapiro (1986) (IP licensing); Meurer (1989) (settlement between patentee and potential challenger of patent); Anton and Yao (1994) (how poor inventors may appropriate value when IP rights are weak); Arora (1995) (transferring know-how); Kortum (1997) (why research inputs have grown rapidly but patents per researcher have fallen and productivity growth has not increased); Henry and Ponce (2011) (knowledge trading versus IP); Weyl and Tirole (2012) (prizes versus IP); Shahshahani (2018) (the effect of court decisions in close IP cases on subsequent legislation).

[4] E.g., Jaffe, Trajtenberg and Fogarty (2000); Hall and Ziedonis (2001); Mowery et al. (2001); Lanjouw and Schankerman (2001); Harhoff, Scherer and Vopel (2003); Moser (2005); Hall, Jaffe and Trajtenberg (2005); Murray and Stern (2007); Gans, Hsu and Stern (2008); Bessen (2008); Lemley and Sampat (2008); Furman and Stern (2011); Moon (2011); Lemley and Sampat (2012); Moser (2012); Williams (2013); Galasso and Schankerman (2015); Murray et al. (2016); Frakes and Wasserman (2017); Kogan et al. (2017); Li, Azoulay and Sampat (2017); Gaulé (2018); Kline et al. (2019); Sampat and Williams (2019); Farre-Mensa, Hegde and Ljunqvist (2020); Feng and Jaravel (2020).

nostic tests on those genes, and clinical trials relating to those genes.

The use of scientific publications to study follow-on innovation is common. The measure is also used by Murray and Stern (2007) and Galasso and Schankerman (2015). Galasso and Schankerman use random assignment to patent cases of Federal Circuit judges with different propensities to invalidate patents to construct an instrument for patent validity. Using downstream patent citations to upstream focal patents before and after the upstream patent was invalidated as their main measure of follow-on innovation, supplemented by data on medical devices and clinical trials for pharmaceuticals, they find that patent invalidation leads to a large (50 percent) increase in citations to a patent in later patents. However, the effect is concentrated in the fields of computers, electronics, and medical devices but absent in drugs and chemical and medical technologies. Focusing on the biomedical field, Murray and Stern (2007) exploit the lag between publication of scientific papers and patents with the same information (patent-paper pairs) to study how citations to a scientific publication change after publication of a granted patent. They find a 10-20 percent drop after patent grant.

By using careful research designs, these studies have contributed greatly to our understanding of the overall effect of patents on follow-on innovation. As explained below, the follow-on innovation in these studies includes both activities that infringe the upstream patent and many that do not infringe. We show that it can be informative to disaggregate these categories.

2 Disaggregating Follow-On Innovation

There are two challenges in studying the effect of patents on follow-on innovation. The first arises from ambiguity in the language of a patent "claim," the part of the patent application that defines the boundaries of the patent right. Because claims use words (rather than, say, precise numerical or formulaic identifiers), it is difficult to determine what is covered by a claim, particularly in a largescale study encompassing many patents. Citations, which are often used as a proxy for follow-on innovation, are recognized to be imperfect for this reason (Galasso and Schankerman (2015)). Sampat and Williams

(2019), following Jensen and Murray (2005), use an ingenious approach to avoid this problem: patent claims on genetic sequences are standardized (the claim is a particular DNA sequence) and thus free of the difficulties of interpreting claim language.[5] We use data from Sampat and Williams (2019) because it allows us to avoid the noise that non-standard language creates in other studies of follow-on innovation.

The second challenge in studying follow-on innovation is determining which follow-on innovations to investigate. Studies of the overall effects of patents on follow-on innovation may look at all subsequent innovative activity related to a patent. But, as we show below, many hypotheses about the effect of patents are usefully tested on a more targeted group of follow-on activities. Below, we set out several such groups that may be differently affected by patents. We do so using the body of US legal doctrines defining infringement.

At the outset, infringement in the United States is defined as making, using, selling, offering to sell, or importing any patented invention (35 U.S.C. § 271(a)).[6] From this baseline, patent law is riddled with doctrines that narrow the range of infringing activities. It is impossible to discuss all such exceptions here, but we set out several doctrines that either (1) create brightline rules about whether or not an activity infringes or (2) affect a large swath of activities (see Freilich (2022), on which much of the following is based, for deeper discussion). In all cases where the borders of a doctrine may reasonably be considered unclear, our robustness checks in Appendix 1 consider alternative judgments about (non)infringement (see Section 3.3). We also note that our analysis considers only *de jure* exceptions to infringement, but there are also situations where entities that infringe patents as a legal matter will not be sued as a practical matter—a *de facto* exemption from infringement (certain university research, for example; see p. 16). Table 1 summarizes the relevant doctrines. They are US-specific, though many countries have similar laws.

[5] However, there are other challenges to determining the scope of gene patent claims (Holman (2012)).

[6] In the context of publications relating to patented genes, a measure of follow-on innovation in Sampat and Williams (2019), this could mean sequencing or detecting a gene using a method that requires copying or amplifying the patented genetic sequence, using a primer that includes the sequence, or creating or using a vector that incorporates the sequence (Sherkow and Abbott (2018)). Each of these activities requires making or using the patented gene. These activities are quite common in research that, for instance, seeks to determine whether a gene sequence is present in a particular population or attempts to create a mouse model with a particular allele. Researchers may import a patented technology if they buy a probe, vector, or primer incorporating a patented genetic sequence abroad and import it into the United States.

Table 1: Important US patent infringement doctrines

Infringement	**Not Infringement**
Activities conducted in the US	Activities conducted outside the US
Activities conducted by non-state-affiliated institutions	Activities conducted by state governments and state-affiliated institutions
Non-applied life sciences research; non-life sciences research	Activities related to submission of information to the FDA
Using a patented technique to generate information	Using information generated by a patented technique

Extraterritorial activities. All patents are limited in geographic scope. In the US, the Patent Act restricts liability to infringing activities conducted "within the United States" or imports "into the United States" (35 U.S.C. § 271(a)). A patent filed in the US therefore allows its holder to prevent use of the patented technology in the US alone, not in other countries.[7] There is anecdotal evidence that researchers move their work abroad in order to avoid US patents. A survey exploring whether patents affect follow-on research found that firms avoided "patents by using the patented technology offshore" (Walsh, Arora and Cohen (2003)). Another study found that firms "innovate where their exposure to patent enforcement and litigation is minimized" (Day and Udick (2019)).

The 271(e)(1) safe harbor. In the life sciences, a great deal of research falls into an exception to patent infringement called alternatively the 271(e)(1) safe harbor, the Bolar exception, or the statutory research exception. This exception provides that it is "not an act of infringement" to make, use, offer to sell, sell, or import a patented invention as long as it is done "solely for uses reasonably related to the development and submission of information under a Federal law which regulates the manufacture, use, or sale of drugs or veterinary biological products" (35 U.S.C. § 271(e)(1)). In practice, this exception almost always applies to information generated for submission to the FDA. The Supreme Court has interpreted the words "reasonably related to the development and submission of information" broadly to include clinical trials, some post-approval drug development, and even early stage research into compounds as long as there is a "reasonable basis for believing that a patented compound may work ... to produce a particular physiological effect, and ... if successful, would be appropriate to include in a submission to the FDA" (*Merck KGaA v. Integra Lifesciences I, Ltd.*, 545 U.S. 193

[7] Note that importing or selling a product made abroad using a process patented in the United States is patent infringement (35 U.S.C. § 271(g)).

(2005)). However, the exception does not exempt "[b]asic scientific research" (*id.*). Since a great deal of life sciences research is done with an eye toward practical applications, this law exempts a considerable amount of life sciences research from patent infringement. However, researchers working on early stage or non-applied work must still be cautious of patent infringement. The boundaries of the 271(e)(1) safe harbor are ambiguous.

Using or importing results generated from a patented technique. While using a patented technique to generate useful information (for instance, using a patented gene to diagnose a patient with a disease) is infringement, a researcher who obtains information from someone else's use of a patented technique does not infringe. For instance, many clinical trials on genetic disorders enroll patients with a particular mutation but don't do the sequencing themselves—instead including only patients previously diagnosed with the mutation, thereby avoiding infringement. Likewise, while importing a patented product is infringement, importing information generated by a patented process is not.[8]

State activities. States, and state-affiliated entities such as state universities, cannot be sued for patent infringement (*Mt. Healthy City Sch. Dist. Bd. of Educ. v. Doyle*, 429 U.S. 274 (1977)). The Eleventh Amendment to the Constitution includes a provision that "[t]he Judicial power of the United States shall not be construed to extend to any suit ... prosecuted against one of the United States," and the Supreme Court has affirmed that states, as sovereign entities, may not be sued by private parties without their consent (*Seminole Tribe v. Fla.*, 517 U.S. 44 (1996)). Because states have not consented to be sued for patent infringement and the circumstances under which the federal government can abrogate state immunity have not applied, states have immunity from patent infringement suits (*Florida Prepaid Postsecondary Educ. Expense Bd. v. College Savings Bank*, 527 U.S. 627 (1999)). For example, when a patentee alleged that the University of Idaho's experimental yellow mustard seed breeding program infringed, the court dismissed the case on sovereign immunity grounds (*Sharafabadi v. Univ. of Idaho*, 2009 WL 10725549 (W.D. Wash. 2009)). When non-state universities are sued, by contrast, this defense is not available (*Madey v. Duke University*, 307 F.3d 1351 (Fed. Cir. 2002)).

[8] For instance, a company that owned a patent on a method of screening compounds that might be useful as drugs was not able to enforce its patent against a pharmaceutical company that used the patented method outside the U.S. and then imported the data (*Bayer AG v. Housey Pharms., Inc.*, 340 F.3d 1367 (Fed. Cir. 2003)).

However, the Supreme Court left open the possibility of certain limited state remedies for patent infringement. In addition, in some circumstances patentees may be able to obtain injunctions (*Ex parte Young*, 209 U.S. 123 (1908)). State activities may therefore also be thought of as more difficult to remedy, rather than entirely noninfringing.

Federal government activities. Like states, the federal government is immune from certain suits. In the context of patent infringement, the federal government has partially waived that immunity, meaning that patent holders can recover "reasonable" compensation for infringement (28 U.S.C. § 1498). But the federal government does not need the patentee's permission before working with the patented invention, and patentees cannot obtain an injunction to stop the federal government's use of a patented invention. Thus, some of the transaction costs that motivate concerns about patents' effect on follow-on innovation do not apply in the federal government context. As with state activities, federal government activities may be difficult for patentees to remedy.

3 Exploring Direct Effects

To explore the direct effect of patents on follow-on innovation, we replicate Sampat and Williams (2019) but redefine categories of follow-on innovations in accordance with our discussion in Section 2. The reanalysis serves both to critically examine Sampat and Williams's important findings and to concretely illustrate how to construct different sets of follow-on innovations. To focus the discussion on our thesis, we follow Sampat and Williams's analysis in everything except for the measure of relevant follow-on innovation.

Sampat and Williams (2019) have two analyses using three datasets of follow-on innovation. The first analysis is a comparison, restricted to genes that were claimed in at least one patent application, between genes that were and were not granted a patent; the second analysis uses patent examiner leniency as an instrument for a patent grant (see Section 1 for more details). The three measures of follow-on innovation are gene-related scientific publications (collected from the Online Mendelian Inheritance in Man (OMIM) database), pharmaceutical clinical trials (from the Citeline Pharmaprojects database), and diagnostic tests (from GeneTests.org). Of these three datasets, only the one on scientific publications is publicly available in a way that enables us to determine whether

a given follow-on innovation could be directly affected by the upstream patent.[9] But, as we explain in Appendix 2, the other two datasets likely contain both infringing and noninfringing follow-on innovation, similarly to the categories we identify below for the first dataset.

To collect follow-on scientific publications, Sampat and Williams extracted publications from the OMIM database and linked them to genes via two sets of gene identifiers, MIM numbers and Entrez geneIDs.[10] Sampat and Williams generated 4,014 unique MIM number-geneID-publication triads. As explained in Section 2, this procedure results in a broad definition of the universe of relevant downstream research because it includes research that is noninfringing, including research produced by non-US and state-affiliated persons. We instead measure the direct effect of patents on of follow-on research by replicating the procedure used by Sampat and Williams and excluding publications that are not infringing for one of the reasons described above. Our procedure is as follows.

3.1 Methodology

In order to ascertain whether a publication was authored by an individual outside the US or associated with federal or state governments, we sought information on the institutional affiliation of the follow-on researcher. We linked each OMIM record to a PubMed record because PubMed includes information on authors' institutional affiliation. To link OMIM and PubMed records, we extracted the title, first author, and year of publication from the OMIM record and then searched for that information in the PubMed database. We excluded OMIM records that were not in the PubMed database and also excluded PubMed records that did not list an institutional affiliation. To ensure accuracy, we excluded OMIM-PubMed matches whose titles were less than 90 percent similar

[9] Data linking pharmaceutical clinical trials to genes are proprietary and not available to us. Data on diagnostic tests are available from the NIH (https://ftp.ncbi.nih.gov/pub/GeneTests/), but the most recently available file was updated in 2014, one year after gene patents were invalidated in the US (*Assoc. for Molecular Pathology v. Myriad Genetics, Inc.*, 569 U.S. 576 (2013)). Because the file does not include the date on which the test first became available, it cannot be used for our purposes. An earlier version of the data, which was used by Sampat and Williams, does not identify whether the tests were available in or outside the US, and therefore cannot be categorized according to our methodology.

[10] The former is used by the OMIM database, the latter by other databases, including the one used by Sampat and Williams to link genes to patents. A file converting between MIM numbers and geneIDs can be found at https://www.omim.org/downloads (genemap2.txt). GeneIDs may be associated with more than one MIM number and MIM numbers may be associated with more than one geneID.

as defined using Levenshtein distance. After these exclusions, our sample included 3,370 publication-MIM number-gene ID triads, which is 644 entries (16 percent) smaller than Sampat and Williams's sample.

For each publication, we manually classified whether or not the first author was affiliated with a US address.[11] If the first author was affiliated with more than one institution, we classified the publication as US if at least one affiliation was in the US. For US institutions, we additionally manually classified whether the institution was (1) for-profit or (2) affiliated with a state government (including state universities and affiliated hospitals) or (3) affiliated with the federal government.

In addition, we manually classified whether a publication fell into the 271(e)(1) exception. This cannot be done precisely at a large scale, nor did we conduct a legal analysis of each publication. Instead, we used a proxy: whether the publication discussed a potential therapeutic application. This is a reasonable proxy because § 271(e)(1) exempts research from infringement where there is a "reasonable basis for believing that a patented compound [or gene] may work ... to produce a particular physiological effect" which would lead to FDA submission (*Merck KGaA v. Integra Lifesciences I, Ltd.*, 545 U.S. 193 (2005)). However, the proxy likely underestimates the true number of publications that fall into the 271(e)(1) exception because a great deal of research is done with an eye towards eventual practical applications, even if those applications are not spelled out in the paper itself. As noted, the precise contours of 271(e)(1) are not clear, particularly as applied to research tools, so this classification is imprecise.

Finally, we manually classified whether a publication's authors made or used genetic sequences themselves or whether they relied on genetic information obtained elsewhere, based on the publication's description of their methodology.

3.2 Results

Figure 1 shows the changes resulting from our refinement. Remarkably, only 369 out of the 2,771 publications identified by Sampat and Williams—a mere 13 percent—might

[11] PubMed's API, Entrez, provides institutional affiliation only for first authors for papers published before 2014. This leaves open the possibility that publications classified as not affiliated with a US address will contain some work done in the US by non-first authors.

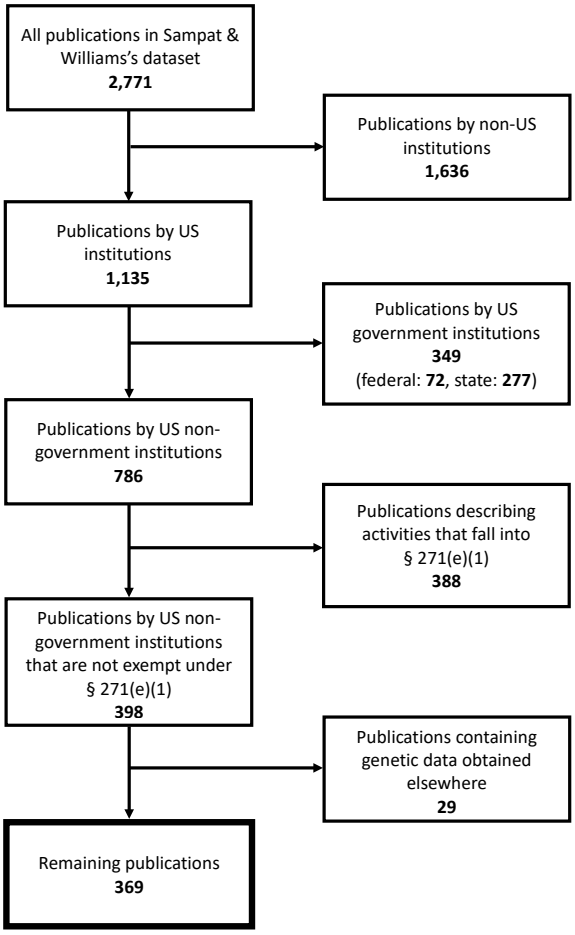

Figure 1: Plausibly Infringing Follow-On Publications in Sampat and Williams (2019)

constitute patent infringement. The rest clearly do not constitute patent infringement because they are authored by non-US or government-affiliated persons, fall under the 271(e)(1) exemption, or obtain their genetic data from others.[12]

We now carry out Sampat and Williams's first analysis on our narrowed list of follow-on publications. Table 2 reports the results. Columns (1) and (2) use the entire set of follow-on publications, as Sampat and Williams did. The results replicate those in Table 2, Panel A of Sampat and Williams (2019). Columns (3) and (4) instead use only publications that could plausibly be considered patent infringement, implementing our disaggregation. Specifically, we included only publications authored by non-government-

[12] Some patents are filed in more than one country, so non-US authors could still be affected by a gene patent that was filed both in the US and in their country. However, of the 499 patent families (including granted and never granted patents) in our sample, less than a third (152) have a family member filed outside the US, let alone one filed in the author's country. Further, many countries have robust research exceptions that exempt nonprofit research from patent infringement, and also have an equivalent to the 271(e)(1) exception. In addition, foreign patents are not subject to the examiner leniency instrument used by Sampat and Williams. Data on patent families and countries was obtained from Google Patents.

Table 2: Patents and Follow-On Publications on Human Genes Claimed in Granted and Not Granted Patent Applications: Regression Estimates

	Dependent variable:			
	log of 2011 pub's	any 2011 pub's	log of potentially infringing 2011 pub's	any potentially infringing 2011 pub's
	(1)	(2)	(3)	(4)
Patent granted	0.0019	−0.0014	−0.0016	−0.0026
	(0.0059)	(0.0054)	(0.0020)	(0.0024)
Mean of dep. var.	0.1104	0.1094	0.0155	0.0197
Observations (genes)	15,524	15,524	15,188	15,188

Note: *p<0.1; **p<0.05; ***p<0.01

This table shows estimates of the difference in follow-on publications in 2011 for genes that were granted a patent versus genes that were claimed in at least one patent application but not granted a patent by 2010. Columns (1) and (2) use all follow-on publications as the dependent variable, replicating Table 2, Panel A in Sampat and Williams (2019). Columns (3) and (4) instead use only follow-on publications that could plausibly be considered patent infringement, implementing our disaggregation. Namely, we consider only publications authored by non-government-affiliated US persons where the 271(e)(1) exemption was not applicable and the author did not obtain the genetic information elsewhere. Each column reports estimates from a separate OLS regression with heteroskedasticity-robust standard errors.

affiliated US persons where the 271(e)(1) exemption was not applicable and the author did not obtain the genetic information elsewhere. The main results are unaffected. Like Sampat and Williams, we fail to reject the null hypothesis that patenting had no effect on follow-on scientific research as measured by publications. But our correlation coefficients are more precisely estimated. Whereas their "95 percent confidence intervals can reject declines or increases of more than 2 percent" (Sampat and Williams (2019)), our 95 percent confidence intervals reject declines of more than half a percent and increases of more than a quarter of a percent (model 3).

Table 3: Patents and Follow-On Publications on Human Genes by Examiner Leniency: Instrumental Variable Estimates

	Dependent variable:			
	log of 2011 pub's	any 2011 pub's	log of potentially infringing 2011 pub's	any potentially infringing 2011 pub's
	(1)	(2)	(3)	(4)
Patent granted (instrumented)	−0.0230**	−0.0187**	−0.0090***	−0.0119***
	(0.0102)	(0.0089)	(0.0027)	(0.0036)
Mean of dep. var.	0.0798	0.0888	0.0060	0.0080
Observations (gene-application pairs)	293,652	293,652	288,021	288,021

Note: *p<0.1; **p<0.05; ***p<0.01

This table shows instrumental-variable estimates of differences in follow-on publications in 2011 for genes that were granted a patent versus genes that were claimed in at least one patent application but not granted a patent by 2010, using the patent examiner's leniency as an instrument for patent grant. Columns (1) and (2) use all follow-on publications as the dependent variable, replicating Table 3, Panel A in Sampat and Williams (2019). Columns (3) and (4) instead use only follow-on publications that could plausibly infringe a patent. Namely, we consider only publications authored by non-government-affiliated US persons where the 271(e)(1) exemption did not apply and the author did not obtain the genetic information elsewhere. Each column is from a separate OLS regression with Art-Unit-by-application-year fixed effects and standard errors clustered by patent application.

Next we proceed to the instrumental-variable analysis, using the patent examiner's leniency as an instrument for the probability that the patent would be granted. After successfully replicating Sampat and Williams's first-stage results (not shown here), we implemented their second-stage analysis using our infringing set of follow-on innovations. Table 3 reports the results. Columns (1) and (2) use the entire set of follow-on scientific publications, as in Table 3, Panel A in Sampat and Williams (2019). Columns (3) and (4) instead use only publications that could plausibly be deemed patent infringement, namely those authored by non-government-affiliated US persons where the 271(e)(1) exemption was not applicable and the author did not obtain the genetic information elsewhere. Sampat and Williams's IV analysis shows a negative effect of patenting on follow-on scientific research, though the magnitude of the effect is small. Our disaggregated analysis replicates qualitatively the same result. But the magnitude of the negative effect we find is smaller. And, as in the basic analysis (Table 2), our point estimates are more precise. Whereas Sampat and Williams's 95 percent confidence intervals "can reject declines of

more than 9 percent" (Sampat and Williams (2019)), we can reject declines of greater than 1.45 percent or of less than 0.3 percent (model 3).

In all, our reanalysis of Sampat and Williams (2019) on the set of infringing follow-on publications has shown two things. First, very few follow-on publications relating to gene patents (a mere 13 percent of Sampat and Williams's sample) can plausibly be thought to infringe the patents. Second, whether a gene was protected by a patent had little or no effect on follow-on scientific publications that might infringe the patent (a small or nonexistent "direct" effect, in our terminology). For both the basic and IV analyses, our estimates of a negative effect were more precise and smaller in magnitude than Sampat and Williams's estimates.

One might have expected our analysis to overturn Sampat and Williams's null result. If Sampat and Williams's set of follow-on innovations includes activities that clearly do not constitute patent infringement and as such would not be expected to be affected by an upstream patent (at least not directly), and if patents are expected to affect follow-on innovation, then by eliminating clearly noninfringing activities we might be expected to find a greater effect than what Sampat and Williams found. That did not happen. Why?

We can think of a context-specific and a more general explanation. First, the follow-on innovations in Sampat and Williams (2019), when narrowed to infringing activities, consist mainly of innovations by university researchers—where survey evidence has shown that innovators tend to ignore patents and patentees tend to tolerate infringement (Walsh, Arora and Cohen (2003)). If that is the explanation, these results would not be expected to generalize to non-research contexts (see also our discussion in Section 5 on reconciling different studies). Second, and more generally, while patents (like any monopoly) are expected to increase prices and restrict access, the theoretical expectation that patenting restricts follow-on innovation is not unchallenged. Theoretical and empirical studies have argued that the exclusivity provided by patent ownership encourages investing in follow-on innovation and reduces the transaction costs of doing so (see Kitch (1977); Arora (1995); Kieff (2001); Gans, Hsu and Stern (2002); Gans, Hsu and Stern (2008)). These forces might counteract the restriction on access, netting out in our context to a neutral effect on follow-on innovation. This explanation, by contrast to the first one, is not

specific to university research. We cannot tell whether the specific or general explanation is driving the null result because the available data only allowed us to apply our refined measure of follow-on innovation in the research context. Implementing our disaggregated measure in other settings would help adjudicate between contending explanations, as we further discuss in Section 5.

3.3 Dealing with Uncertainty

Patent scope, infringement, and validity are famously characterized by uncertainty (Lemley and Shapiro (2005)) and, while the doctrines discussed herein are more certain than most, we cannot guarantee that all publications falling into a particular category are noninfringing. To address potential uncertainties, in Appendix 1 we present robustness checks replicating our analyses with alternative classifications of clearly noninfringing activities. First, as explained in Section 2, activities by the federal government are not technically immune from infringement, though they do not require a patentee's prior license. Tables 7 and 8 in Appendix 1 accordingly present alternative versions of Tables 2 and 3 that exclude papers affiliated with the federal government from our noninfringing list. Second, a possible objection to our analysis is that some of the activities we classify as noninfringing may not be known by scientists to be noninfringing, so they might still be directly affected by patents. This might particularly apply to activities conducted by state universities because sovereign immunity is a lesser-known legal doctrine and to the 271(e)(1) exemption because its boundaries are not precise. Moreover, the exemption of state activities from infringement may be complicated where scientists at state-affiliated institutions collaborate with scientists at other institutions. To address these concerns, Tables 9 and 10 in Appendix 1 present alternative versions of Tables 2 and 3 that exclude from the set of infringing follow-on innovations only activities conducted outside the US. Finally, to address potential uncertainty about the classification of US authors, Tables 11 and 12 in Appendix 1 classify an author with multiple institutional affiliations as US-based only if *all* of the author's affiliations are in the US (as opposed to Tables 2 and 3, which classify an author with multiple affiliations as US-based if *any* of the author's affiliations are in the US). Results are similar in all robustness checks.

4 Exploring Indirect Effects

Disaggregating follow-on innovation into direct and indirect effects permits a more nuanced look at indirect effects—ways in which patents can affect follow-on innovation that does not infringe. For instance, a foreign innovator or state university may be able to conduct research without fear of infringement but would not be able to sell the fruits of that research in the United States without infringement. The diminished prospects for commercialization may negatively impact incentives for noninfringing research, just as other effects of patents—disclosure, for instance—may positively affect it. Although many mechanisms of patents' indirect effect on follow-on innovation can also apply to infringing follow-on innovation, they are more difficult to isolate in that context, so it is useful to study them in the context of noninfringing activities. There are many different indirect effects and we are unable to investigate them all in our specific empirical context.[13] Instead, we show how disaggregating noninfringing follow-on innovation can improve studies of indirect effects by looking at whether and how (1) the closeness of a patent to expiration and (2) the closeness of follow-on research to commercialization or clinical application moderate the effect of patents on follow-on research.[14]

4.1 Closeness of Patent to Expiration

In the US, patent protection typically expires 20 years after the application was filed (35 U.S.C. § 154(a)). We explore whether the effect of patents on noninfringing research varies depending on how close the patent is to expiration. There is good reason to think that it might, because the (positive or negative) effect of a patent on activities that will bear fruit in the future is likely to be different depending on whether the patent will still be in effect in that future. The direction of the difference is not obvious as a theoretical

[13] In aggregate, we find that the effect of gene patents on noninfringing follow-on publications is the same as on infringing follow-on publications—that is, not statistically distinguishable from zero (Appendix 1, Table 13).

[14] Initially, we also attempted to investigate one subset of innovations that might be strongly affected by the indirect effect through commercialization prospects—namely, publications by for-profit entities. This does not exhaust the set of relevant follow-on innovations for estimating indirect effects, but it is a relevant subset because the probability that commercialization prospects are motivating research is greater for for-profits than for nonprofits. However, we found that only 33 publications in the Sampat and Williams (2019) dataset are from for-profit entities—too small a sample and with too little variation in the independent variable (patent grant) to allow for meaningful statistical inference.

matter. As discussed in Section 3.2, there are sound reasons for believing that patents discourage follow-on innovation by raising prices and barriers to access, or alternatively that they encourage investment in follow-on innovation by promising exclusive rewards, reducing transaction costs, or signaling that a field is worth exploring. But there is good reason to think that, whatever its direction, the magnitude of the effect should be greater when the patent is farther from expiration. This applies for both negative and positive effects: a researcher is less likely to be deterred from research for fear that its commercialization will be infringing when the patent will have expired by the time the research results get to market, and a patentee or licensee is less likely to invest in a field when the patent will have expired by the time the investment bears fruit. A similar prediction about magnitude would hold for indirect-effect mechanisms that are based not on commercialization but on signaling. It is plausible to think, for example, that patents have a positive effect on follow-on research by signaling that a certain field is fruitful for research, such that there is an uptick in research in the field after the patent is granted and the topic becomes hot, but interest dampens or potential is exhausted as time passes and the patent nears expiration.

To get an empirical handle on the question, we gathered data on gene patent expiration dates from Google Patents.[15] Table 4 shows the distribution of expiration dates for gene patents in the Sampat and Williams dataset.[16]

Table 4: Number of genes covered by patents with different expiration dates

not patented	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027
10,432	68	86	402	4	300	1,404	112	125	1,008	600	21

We estimate the linear model

$$y_i = \beta_0 + \beta_1 X_i + \beta_2 Z_i + \epsilon_i \tag{1}$$

[15] Patent expiration data are, of course, patent-level, but the Sampat and Williams publications dataset is gene-level. To account for the fact that some genes are covered by more than one patent, we ascribed the *latest* expiration date of the patents covering the gene to the gene. This is the most appropriate approach to account for lack of a one-to-one correspondence between genes and gene patents because it chooses the date as of which a gene is no longer covered by *any* patents, thereby facilitating the proper comparison between the patent and no-patent counterfactuals.

[16] We were unable to find expiration dates for 626 gene patents in the Sampat and Williams dataset, which is why the number of observations in Table 5 to come is lower than in previous tables. We replicated previous analyses on the smaller dataset to confirm that the relatively slight (4 percent) decrease in the number of genes does not meaningfully affect the results.

where i indexes genes, y_i measures follow-on publications in 2011 (like Sampat and Williams, we use both the logged number of publications and a binary measure for whether there were any publications), X_i is an indicator for whether the gene was protected by a patent, and Z_i is an indicator for whether the patent was close to expiration. In this specification, the estimate of β_1 is interpretable as the effect of a patent that is *not* expiring soon on follow-on publications, and the estimate of β_2 is interpretable as the difference between the effect of a patent that is expiring soon and the effect of a patent that is not expiring soon on follow-on publications.[17] Thus we follow the same empirical strategy as Sampat and Williams but we restrict the analysis to noninfringing publications and we add a measure of whether the patent is expiring soon.

Table 5 reports the results where we defined "expiring soon" to mean patents expiring before 2023 (and all other patents are considered "not expiring soon"). This divides expiration years and the number of associated gene patents roughly in half. We see that not-expiring-soon gene patents have a positive and statistically significant effect on follow-on publications. Being protected by a patent that is not expiring soon is associated with a 4 percent increase in the number of follow-on publications relative to not being protected by a patent. Moreover, the difference between expiring-soon and not-expiring-soon patents is negative and statistically significant. A gene protected by a patent that is expiring soon has statistically significantly fewer follow-on publications, and a statistically significantly lower probability of having any follow-on publications, than a gene protected by a patent that is not expiring soon.

We check the robustness of these results to different definitions of what it means for a patent to be expiring "soon" or "late." In Appendix 1, Tables 14-18, we experiment with all possible symmetric ways to categorize soon-expiring and late-expiring patents: (1) expiring soon means 2017 and expiring late means 2027 (first versus last expiration years), (2) expiring soon means 2017-18 and expiring late means 2026-27 (first two versus last two expiration years), (3) expiring soon means 2017-19 and expiring late means

[17] The specification is equivalent to the following specification with an interaction term,

$$y_i = \beta_0 + \beta_1 X_i + \beta_2 (X_i \times Z_i) + \epsilon_i$$

because $Z_i = 1$ only if $X_i = 1$ (a gene has an expiring-soon patent only if it has a patent to begin with).

Table 5: Soon-Expiring and Late-Expiring Patents and Noninfringing Follow-On Publications: Regression Estimates

	Dependent variable:	
	log of 2011 publications	any 2011 publications
	(1)	(2)
Patent granted	0.0364***	0.0165***
	(0.0074)	(0.0053)
Expiring soon	−0.0558***	−0.0182***
	(0.0092)	(0.0066)
Mean of dep. var.	0.0831	0.0464
Observations (genes)	14,562	14,562
Note:		*p<0.1; **p<0.05; ***p<0.01

This table shows estimates of the effect of gene patents that are not expiring soon on follow-on publications in 2011, as well as estimates of the difference in effects between patents that are expiring soon and patents that are not expiring soon. A patent is defined as expiring soon if it expires before 2023.

2025-27 (first three versus last three expiration years), (4) expiring soon means before 2021 and expiring late means after 2023 (first four versus last four expiration years), and (5) expiring soon means before 2022 and expiring late means after 2022 (first five versus last five expiration years).[18] Results are largely consistent with Table 5. The not-expiring-soon coefficient is positive and statistically significant for the logged publications outcome in all specifications; it is also positive for the binary publications outcome in all specifications, and statistically significant for the binary publications outcome in all but one of the specifications. We can therefore be confident that the positive effect of not-expiring-soon patents on follow-on research is not an artifact of a particular definition of "soon." The difference between the expiring-soon and not-expiring-soon effects is also negative in all specifications but statistically significant in only some specifications. In short, we have a robust finding that being protected by a patent that is not expiring soon is associated with greater follow-on research on a gene relative to not being patented.

[18] In Tables 14-18 in Appendix 1, we drop other patented genes from the dataset when we compare soon-expiring and late-expiring patents (i.e., we drop "middle-expiring" patents). Alternatively we could keep all genes, specify a model with more covariates (an additional indicator for expiring late), and do a hypothesis test on the difference between the expiring-soon and expiring-late coefficients. We have performed that analysis for all tables, and it yields the same results. We choose the specification in Tables 14-18 for ease of interpretation.

Our finding confirms the theoretical prediction that the magnitude of a patent's effect on noninfringing follow-on innovation should be greater for patents that are farther away from expiration. It also shows that the net effect on noninfringing follow-on research is positive for genes that are protected by a patent that is not close to expiration (relative to genes that are not protected by a patent). These results are consistent with a mechanism whereby the patentee (or its licensee) is encouraged to invest in research in a patented field by the prospect of exclusive returns and reduced transaction costs (or researchers are encouraged to boost their work in the field in hopes of marketing it to the patentee), but the effect diminishes as the patent gets closer to expiration and the patentee thus becomes less certain of its ability to reap rewards from further investment. They are also consistent with a signaling mechanism whereby a patent fuels follow-on research on a topic by signaling its fruitfulness, but the new interest or the field's potential for new research diminishes as time passes.

We caution, though, that one limitation of the data is that it has no patents with a very early expiration date relative to the outcome variable date. The earliest expiration date is 2017, six years after the year in which follow-on publications were measured. Future research will hopefully shed light on how the indirect effect of a patent on follow-on innovation is moderated by the patent's closeness to expiration in contexts where some patents are expiring very soon.

4.2 Closeness of Follow-On Research to Commercialization

Whereas the previous section focused on how a patent's indirect effect on follow-on research is moderated by a characteristic of the patent, this section explores how the effect is moderated by a characteristic of the research. We ask whether the effect of a patent on noninfringing follow-on research varies depending on whether the research is close to commercialization or clinical application.

This question speaks to the possibility that follow-on research that does not infringe a patent may nonetheless be affected in situations where commercializing the research constitutes patent infringement. This indirect inhibitory mechanism could apply to certain activities explored in this paper—for example, extraterritorial research potentially

leading to sales in the US. Similarly, many research activities are not infringing because they fall into the 271(e)(1) safe harbor (e.g., clinical trials), but that safe harbor does not apply to sales of the researched product (e.g., sales of a drug after FDA approval). As in the previous section, we have a firmer theoretical expectation about the magnitude than about the direction of the difference in effects. The patent's impact could be negative if the inability to monetize the results of research dampens incentives to conduct the research; the effect could be positive if the patentee invests in the field.[19] Whatever the direction, we hypothesize that because patents have a limited term, the effect of a patent on future commercialization opportunities for noninfringing research will be stronger for research that is closer to commercialization. This applies for both negative and positive effects: a researcher who is decades away from commercialization will not be deterred by a patent that will have likely expired by the time the research gets to market; a patentee investing in a field in order to encourage development of technologies that will pay it licensing fees will not invest in technologies likely to come to market after the patent has expired.[20]

We do not have direct information on how close our instances of follow-on innovation are to commercialization or clinical application, so we use a proxy: we assume that research on humans is closer to commercialization than research *in vitro*. This is an imprecise assumption but holds true generally. To implement our proxy measure, we read the abstracts of all 2,402 noninfringing publications in the Sampat-Williams dataset (with the help of research assistants) and classified each publication as either human, animal, or *in vitro*.[21] We then ran regressions to see whether the effect of patents on follow-on research is more pronounced in research that is closer to commercialization, as approximated by these categories. That is, we reran Sampat and Williams's regression analysis on different categories of noninfringing publications defined by closeness to commercialization.

[19] The effect could also be positive if the patent encourages the development of substitute technologies or spurs activity by disclosing the initial technology, but our data is not set up to measure these effects.

[20] This is complicated in situations where patentees have portfolios of patents with staggered expiration dates.

[21] If a paper involved different stages of research, we classified it according to the latest stage. So papers containing both human and animal research were classified as "human," and those containing both animal and *in vitro* research were classified as "animal."

Table 6 reports the results when we divide all noninfringing studies into those done on humans and those not done on humans, meaning animal or *in vitro* studies. (Table 19 in Appendix 1 uses an alternative comparison that groups human and animal research together and compares it to *in vitro* research, with similar results.) The correlation coefficient estimates are not statistically significant at conventional levels for either category. But the correlation is negative for human research and positive for nonhuman research, which is consistent with the conjecture that research that is closer to commercialization is more likely to be deterred by the existence of a gene patent when the research itself is not infringing but the commercialized fruit of it might be. The lack of a statistically significant effect is not particularly surprising given that there were no significant effects in the entire population of publications nor for all infringing or all noninfringing publications (see Tables 2 and 13). But the analysis hopefully shows the way for future studies that can exploit the idea of closeness to commercialization.

Table 6: Patents and Noninfringing Follow-On Publications Classified by Closeness of Research to Commercialization Based on Type of Research: Regression Estimates

	Dependent variable:			
	log of 2011 human pub's	log of 2011 nonhuman pub's	any 2011 human pub's	any 2011 nonhuman pub's
	(1)	(2)	(3)	(4)
Patent granted	−0.0036	0.0041	−0.0023	0.0037
	(0.0034)	(0.0038)	(0.0033)	(0.0042)
Mean of dep. var.	0.0358	0.0530	0.0369	0.0612
Observations (genes)	15,188	15,188	15,188	15,188

Note: *p<0.1; **p<0.05; ***p<0.01

This table shows estimates of the difference in noninfringing follow-on publications in 2011 for genes that were granted a patent versus genes that were claimed in at least one patent application but not granted a patent by 2010, with separate estimates of the difference for studies done on humans (closer to commercialization) and studies not done on humans (farther from commercialization).

5 Discussion

Understanding how well the patent system works requires understanding how patents affect innovations that build on patented inventions. Economists have accordingly devoted considerable energy to examining the effect of patents on follow-on innovation. But prior studies have grouped direct and indirect effects together, which allows observation of only the overall effect of patents on follow-on innovation. Our disaggregated measure provides insight into the mechanisms by which patents can affect follow-on innovation. We discuss some implications below.

5.1 Applying the Disaggregated Measure to Policy

Isolating the direct effect of patents on follow-on innovation is useful to assess a number of policy questions. We do so here with the caveat that this study measured effects in the narrow context of gene patents and conclusions about policy may not generalize.

Research Exceptions. Many countries have some form of research exception to patent infringement. The specifics vary, but these exceptions generally preclude infringement liability for activities conducted on patented technologies for purposes of research. The United States historically had a research exception, but it was effectively eliminated in 2002 (*Madey*). There is an extensive debate about the optimal contours of a research exception, with arguments both for and against a robust exception focused on the effect of such an exception on follow-on research (Eisenberg (2008)).

As a policy intervention, the research exception targets the direct effect of patents on follow-on research. It has no influence on indirect effects: research permitted by such an exception could still lead to infringement if the research were sold commercially, and research would still be promoted by disclosure in patents and by any investment the patentee made in the field. The research exception should therefore be assessed only in terms of direct effects. Our findings show that the number of infringing publications is small (and not because such research is suppressed by patents), suggesting that a research exception would have a limited impact. Further, we find no evidence that patents have a direct effect on follow-on innovation. If the null result indicates a lack of effect (as opposed

to a negative effect counteracted by a positive effect, which we cannot rule out), then instituting a research exception in the US would be unlikely to affect follow-on innovation. However, because studies outside the gene context (Murray and Stern (2007); Galasso and Schankerman (2015)) have found that patents reduce overall follow-on innovation, direct effects—and the impact of a research exception—may be nontrivial in those fields.

Categorical Patent Exclusions. There are a variety of categorical patent exclusions, fields where innovation is not eligible for a patent. These exclusions are often justified in part on the grounds that patents would restrict follow-on research. For example, in *Lab. Corp. of America Holdings v. Metabolite Labs., Inc.*, 548 U.S. 124 (2006), Justice Breyer explained that laws of nature, natural phenomena, and abstract ideas cannot be patented because patents "can discourage research by impeding the free exchange of information, for example by forcing researchers to avoid the use of potentially patented ideas, by leading them to conduct costly and time-consuming searches of existing or pending patents, by requiring complex licensing arrangements, and by raising the costs of using the patented information, sometimes prohibitively so."[22] If patents hinder follow-on innovation, banning patents in a field would assuredly preclude that effect, whether the effect occurred through direct or indirect mechanisms. However, subject-matter exclusions are controversial in part because a categorical ban on patents has many additional effects, such as potentially disincentivizing basic scientific research that might fall into the "abstracts ideas" or "laws of nature" categories. Policymakers must therefore determine whether a broad, categorical policy—one that affects both follow-on innovation and other aspects of the patent system—is necessary. If the effect of patents on follow-on innovation is mostly from the direct mechanism, then a more targeted policy, such as a research exception, might serve. But if the effect of patents on follow-on innovation is from indirect mechanisms, then a broader policy, such as subject-matter exclusions, might be necessary. Thus, to assess whether subject-matter exclusions are necessary to sustain follow-on research, one must separate direct and indirect effects and measure the indirect effect of patents on follow-on research.

International Arbitrage. Institutions with facilities in multiple countries can shift

[22] Breyer was writing in dissent from the Supreme Court's decision not to hear an appeal in the case (dismiss the writ of certiorari).

research activities to a jurisdiction with no relevant patent, which is a complete bar to liability. This presents arbitrage opportunities (Samuelson (2004)) and has been reported anecdotally (see p. 8). There are, however, governmental pushes to keep research domestic. One factor in keeping research domestic is the extent to which researchers worry about the prospect of patent liability. The extent to which international arbitrage occurs, and the extent to which it can be prevented by policy, depends on whether patents directly affect follow-on research. Mechanisms by which patents can indirectly affect follow-on research, including disclosure and the potential for the commercialized patents to infringe, are not country-dependent.[23] Our results can shed some light on the likelihood of international arbitrage. Most notably, given that we find patents do not directly affect follow-on research, the presence of patents is unlikely to be a frequent motivation for researchers to move their work abroad in the gene context.

5.2 Applying Our Disaggregated Measure to Other Studies

To illustrate the importance of disaggregating categories of follow-on innovation, we show how our refinement applies to other studies of follow-on innovation and how it may help explain discrepancies between several key studies. Murray and Stern (2007) use citations as their measure of follow-on innovation and include citations by entities outside the United States. This extraterritorial research is not patent infringement. Further, Murray and Stern observe citations from 2002 and earlier, and for most of that period the United States had a common-law research exception which was thought to exempt university research from patent infringement.[24] Murray and Stern found a 10-20 percent decrease in follow-on innovation after the patent grant; had they narrowed their definition of follow-on innovation to citations that were infringement, the effect may have been different.

Galasso and Schankerman (2015) also use citations to measure follow-on innovation. They appear to count only citations by US patents, but US patents are often filed by

[23] Barring language and access barriers for disclosure and assuming that the follow-on researcher desires commercial opportunities internationally.

[24] The precise contours of the exception were never clear. The exception essentially ceased to exist after the Federal Circuit's decision in *Madey v. Duke University*, 307 F.3d 1351 (Fed. Cir. 2002). By contrast, the 271(e)(1) safe harbor, although it would have applied to much of Murray and Stern's sample, did not exist as such in the period they studied (the statute itself was on the books, but it was not interpreted to apply broadly until 2005).

foreign inventors,[25] meaning they can represent research that was done outside the US and therefore could not infringe a US patent. Further, citations in their sample could reflect work done at state-based institutions such as universities. And, particularly in the life sciences, many of the citations likely disclose work falling into the 271(e)(1) exception. This last exception—which applies strongly in the life sciences but not in other areas—may explain why Galasso and Schankerman found no effect in the biomedical sciences but found a substantial effect in other industries.

Greater attention to the nuances of measuring follow-on innovation may also explain why Galasso and Schankerman (2015) found a large (50 percent) effect of patents on follow-on innovation in some fields but Murray and Stern (2007) found a much smaller effect and Sampat and Williams (2019) found no effect. Galasso and Schankerman used patents as their measure of follow-on innovation, whereas the other studies used publications. It is likely that much of the follow-on research in these studies, whether patents or publications, describes research that is not infringing, so an upstream patent might not be expected to have a direct effect. However, there could be a difference in indirect effects because follow-on patents are much more likely than publications to reflect research from for-profit companies.[26] Follow-on patents may reflect a desire to monetize research, which would be patent infringement and would not be subject to the noninfringement doctrines discussed in Section 2.

Finally, Williams (2013) studied publications relating to genes that were available publicly and genes that were held in secret by Celera and disclosed only under a restrictive contract. Williams found that Celera's IP (trade secret) protections decreased follow-on innovation by 20-30 percent. This is a notable difference from the later finding in Sampat and Williams (2019) that gene patents had no effect on follow-on innovation. As Sampat and Williams (2019) point out, this discrepancy may arise in part because Celera's IP protection was not patent-based and therefore not subject to any of the doctrines described above.

As a generalizable point, the effect of patents on follow-on innovation will differ sub-

[25] US Patent Office statistics for 2019 show that 55 percent of US patents were of foreign origin (USPTO (2020)).

[26] Patents filed by universities make up a small portion of all patents—approximately 2 percent, according to a 2013 US Patent Office report (USPTO (2013)).

stantially in different industries, at different stages of follow-on innovation (e.g., research versus commercialization), and in different countries. Studies may thus reach divergent conclusions based on the specifics of the studies' design. An important takeaway from our analysis is that the proper measure of follow-on innovation depends on the hypothesized mechanism through which a patent might affect downstream innovation. By showing how more specific measures can be constructed, we hope to have contributed to a better understanding of the effect of patents and, ultimately, to more informed patent policy.

Finally, our work is part of a larger trend towards refined measures in studies of innovation. In the context of patent citations, for example, researchers initially used simple counts but over time adopted increasingly fine-tuned measures to improve the correspondence between citation counts and the question of interest (Jaffe and de Rassenfosse (2019)). We hope that our contribution to the measurement of follow-on innovation will likewise help advance understanding of the effect of patents and, ultimately, inform patent policy.

Appendix 1: Robustness Checks

Table 7: Patents and Follow-On Publications on Human Genes Claimed in Granted and Not Granted Patent Applications: Regression Estimates (Alternative Version)

	Dependent variable:	
	log of potentially infringing 2011 pub's	any potentially infringing 2011 pub's
	(1)	(2)
Patent granted	−0.0020	−0.0032
	(0.0021)	(0.0025)
Mean of dep. var.	0.0172	0.0215
Observations (genes)	15,188	15,188
Note:		*p<0.1; **p<0.05; ***p<0.01

This table presents an alternative version of columns (3)-(4) in Table 2. Whereas Table 2 considered publications authored by non-*government*-affiliated US persons where the 271(e)(1) exemption did not apply and the author did not obtain the genetic information elsewhere, this table considers publications authored by non-*state*-affiliated US persons where the 271(e)(1) exemption did not apply and the author did not obtain the genetic information elsewhere. That is, publications authored by federal-government-affiliated authors were not considered potentially infringing in Table 2 but are considered so here.

Table 8: Patents and Follow-On Publications on Human Genes by Examiner Leniency: Instrumental Variable Estimates (Alternative Version)

	Dependent variable:	
	log of potentially infringing 2011 pub's	any potentially infringing 2011 pub's
	(1)	(2)
Patent granted (instrumented)	−0.0098***	−0.0122***
	(0.0029)	(0.0038)
Mean of dep. var.	0.0072	0.0092
Observations (gene-application pairs)	288,021	288,021
Note:		*p<0.1; **p<0.05; ***p<0.01

This table presents an alternative version of columns (3)-(4) in Table 3. Whereas Table 3 considered publications authored by non-*government*-affiliated US persons where the 271(e)(1) exemption did not apply and the author did not obtain the genetic information elsewhere, this table considers publications authored by non-*state*-affiliated US persons where the 271(e)(1) exemption did not apply and the author did not obtain the genetic information elsewhere. That is, publications authored by federal-government-affiliated authors were not considered potentially infringing in Table 3 but are considered so here.

Table 9: Patents and Follow-On US Publications on Human Genes Claimed in Granted and Not Granted Patent Applications: Regression Estimates

	Dependent variable:	
	log of 2011 US pub's	any 2011 US pub's
	(1)	(2)
Patent granted	−0.0016	−0.0030
	(0.0035)	(0.0037)
Mean of dep. var.	0.0414	0.0462
Observations (genes)	15,188	15,188
Note:		*p<0.1; **p<0.05; ***p<0.01

This table presents an alternative version of columns (3)-(4) in Table 2. Whereas Table 2 considered publications authored by non-government-affiliated US persons where the 271(e)(1) exemption did not apply and the author did not obtain the genetic information elsewhere, this table considers all publications authored by US persons.

Table 10: Patents and Follow-On US Publications on Human Genes by Examiner Leniency: Instrumental Variable Estimates

	Dependent variable:	
	log of 2011 US pub's	any 2011 US pub's
	(1)	(2)
Patent granted (instrumented)	−0.0170***	−0.0157***
	(0.0061)	(0.0061)
Mean of dep. var.	0.0276	0.0319
Observations (gene-application pairs)	288,021	288,021
Note:		*p<0.1; **p<0.05; ***p<0.01

This table presents an alternative version of columns (3)-(4) in Table 3. Whereas Table 3 considered publications authored by non-government-affiliated US persons where the 271(e)(1) exemption did not apply and the author did not obtain the genetic information elsewhere, this table considers all publications authored by US persons.

Table 11: Patents and Follow-On Publications on Human Genes Claimed in Granted and Not Granted Patent Applications: Regression Estimates (Alternative Classification of US Authorship)

	Dependent variable:	
	log of potentially infringing 2011 pub's	any potentially infringing 2011 pub's
	(1)	(2)
Patent granted	−0.0003	−0.0029
	(0.0028)	(0.0031)
Mean of dep. var.	0.0273	0.0314
Observations (genes)	15,188	15,188
Note:		*p<0.1; **p<0.05; ***p<0.01

This table presents an alternative version of columns (3)-(4) in Table 2. The difference is in the classification of US authors. Whereas Table 2 classifies an author with multiple institutional affiliations as US-based if *any* of the author's affiliations are in the US, this table classifies an author with multiple institutional affiliations as US-based if *all* of the author's affiliations are in the US.

Table 12: Patents and Follow-On Publications on Human Genes by Examiner Leniency: Instrumental Variable Estimates (Alternative Classification of US Authorship)

	Dependent variable:	
	log of potentially infringing 2011 pub's	any potentially infringing 2011 pub's
	(1)	(2)
Patent granted (instrumented)	−0.0129***	−0.0143***
	(0.0045)	(0.0047)
Mean of dep. var.	0.0159	0.0188
Observations (gene-application pairs)	288,021	288,021
Note:		*p<0.1; **p<0.05; ***p<0.01

This table presents an alternative version of columns (3)-(4) in Table 3. The difference is in the classification of US authors. Whereas Table 3 classifies an author with multiple institutional affiliations as US-based if *any* of the author's affiliations are in the US, this table classifies an author with multiple institutional affiliations as US-based if *all* of the author's affiliations are in the US.

Table 13: Direct and Indirect Effect of Patents on Follow-On Publications: Regression Estimates

	Dependent variable:			
	log of potentially infringing 2011 pub's	any potentially infringing 2011 pub's	log of noninfringing 2011 pub's	any noninfringing 2011 pub's
	(1)	(2)	(3)	(4)
Patent granted	−0.0016	−0.0026	0.0017	0.0050
	(0.0020)	(0.0024)	(0.0051)	(0.0037)
Mean of dep. var.	0.0155	0.0197	0.0820	0.0462
Observations (genes)	15,188	15,188	15,188	15,188
Note:				*p<0.1; **p<0.05; ***p<0.01

This table compares the effect of patents on potentially infringing follow-on research (columns 1-2, replicating columns 3-4 in Table 2) with their effect on noninfringing follow-on research (columns 3-4). Both direct and indirect effects are statistically indistinguishable from zero.

Table 14: Soon-Expiring and Late-Expiring Patents and Noninfringing Follow-On Publications: Alternative Regression Estimates

	Dependent variable:	
	log of 2011 publications	any 2011 publications
	(1)	(2)
Patent granted	0.3712***	0.0983**
	(0.0642)	(0.0452)
Expiring soon	−0.3652***	−0.0840
	(0.0734)	(0.0517)
Observations (genes)	10,521	10,521
Note:		*p<0.1; **p<0.05; ***p<0.01

This is an alternative version of Table 5. It shows estimates of the effect of gene patents that are expiring late on follow-on publications in 2011, as well as estimates of the difference in effects between patents that are expiring soon and patents that are expiring late. A patent is defined as expiring soon if it expires in 2017 and as expiring late if it expires in 2027.

Table 15: Soon-Expiring and Late-Expiring Patents and Noninfringing Follow-On Publications: Alternative Regression Estimates

	Dependent variable:	
	log of 2011 publications	any 2011 publications
	(1)	(2)
Patent granted	0.0250**	0.0005
	(0.0122)	(0.0085)
Expiring soon	−0.0458*	−0.0126
	(0.0265)	(0.0186)
Observations (genes)	11,207	11,207
Note:		*p<0.1; **p<0.05; ***p<0.01

This is an alternative version of Table 5. It shows estimates of the effect of gene patents that are expiring late on follow-on publications in 2011, as well as estimates of the difference in effects between patents that are expiring soon and patents that are expiring late. A patent is defined as expiring soon if it expires in 2017-18 and as expiring late if it expires in 2026-27.

Table 16: Soon-Expiring and Late-Expiring Patents and Noninfringing Follow-On Publications: Alternative Regression Estimates

	Dependent variable:	
	log of 2011 publications	any 2011 publications
	(1)	(2)
Patent granted	0.0381***	0.0193***
	(0.0080)	(0.0057)
Expiring soon	−0.0280*	−0.0081
	(0.0147)	(0.0104)
Observations (genes)	12,617	12,617
Note:		*p<0.1; **p<0.05; ***p<0.01

This is an alternative version of Table 5. It shows estimates of the effect of gene patents that are expiring late on follow-on publications in 2011, as well as estimates of the difference in effects between patents that are expiring soon and patents that are expiring late. A patent is defined as expiring soon if it expires in 2017-19 and as expiring late if it expires in 2025-27.

Table 17: Soon-Expiring and Late-Expiring Patents and Noninfringing Follow-On Publications: Alternative Regression Estimates

	Dependent variable:	
	log of 2011 publications	any 2011 publications
	(1)	(2)
Patent granted	0.0354***	0.0170***
	(0.0077)	(0.0055)
Expiring soon	−0.0234	−0.0062
	(0.0146)	(0.0103)
Observations (genes)	12,746	12,746
Note:		*p<0.1; **p<0.05; ***p<0.01

This is an alternative version of Table 5. It shows estimates of the effect of gene patents that are expiring late on follow-on publications in 2011, as well as estimates of the difference in effects between patents that are expiring soon and patents that are expiring late. A patent is defined as expiring soon if it expires in 2017-20 and as expiring late if it expires in 2024-27.

Table 18: Soon-Expiring and Late-Expiring Patents and Noninfringing Follow-On Publications: Alternative Regression Estimates

	Dependent variable:	
	log of 2011 publications	any 2011 publications
	(1)	(2)
Patent granted	0.0364***	0.0165***
	(0.0075)	(0.0054)
Expiring soon	−0.0354***	−0.0018
	(0.0123)	(0.0088)
Observations (genes)	13,158	13,158
Note:		*p<0.1; **p<0.05; ***p<0.01

This is an alternative version of Table 5. It shows estimates of the effect of gene patents that are expiring late on follow-on publications in 2011, as well as estimates of the difference in effects between patents that are expiring soon and patents that are expiring late. A patent is defined as expiring soon if it expires in 2017-21 and as expiring late if it expires in 2023-27.

Table 19: Patents and Noninfringing Follow-On Publications Classified by Closeness of Research to Commercialization Based on Type of Research (Alternate Classification)

	Dependent variable:			
	log of 2011 human or animal pub's	log of 2011 *in vitro* pub's	any 2011 human or animal pub's	any 2011 *in vitro* pub's
	(1)	(2)	(3)	(4)
Patent granted	−0.0019	0.0023	−0.0001	0.0009
	(0.0041)	(0.0031)	(0.0041)	(0.0035)
Mean of dep. var.	0.0538	0.0350	0.0568	0.0410
Observations (genes)	15,188	15,188	15,188	15,188

Note: *p<0.1; **p<0.05; ***p<0.01

This table presents an alternative version of Table 6. Whereas Table 6 showed whether the effect of a gene patent on noninfringing follow-on research is different depending on whether the research was done on human or nonhuman subjects, this table presents a comparison between noninfringing human or animal research on one hand and noninfringing *in vitro* research on the other hand.

Appendix 2: Other Measures of Follow-On Innovation

Sampat and Williams (2019) include not only publications but also two other measures of follow-on innovation: diagnostic tests and clinical trials. These alternative measures avoid the possible idiosyncrasies of scientific publications—namely that they are often authored by academics who have different incentives with respect to patents than non-academics. While we unfortunately do not have the data to analyze these additional measures, we note that they are likely subject to the same issues as the publication measure.

First, both diagnostic tests and clinical trials can be run outside the United States. To get a rough sense of how common this might be, and how it might compare with our results on extraterritorial publications, we searched https://clinicaltrials.gov/ for "BRCA1" (a gene involved in breast, ovarian, and prostate cancers), which yielded 288 trials (for a search conducted in June 2021). Of these, 122 (42 percent) were based outside the United States. For diagnostic tests, we used the most recently available data from GeneTests (2014), which lists genes and corresponding tests. 5,178 genes had diagnostic tests available outside the US and 4,644 had diagnostic tests available in the US. So 53 percent of diagnostic tests were extraterritorial, roughly similar to 59 percent of publications that were extraterritorial. Overall, then, the data suggest that extraterritoriality has a comparable impact on all three measures of follow-on innovation (see Table 20). If anything, the data on genetic tests and clinical trials likely overestimate the amount of US-based follow-on research occurring in 2011 (the year in which publications were counted) because data on genetic tests and clinical trials come from after the US Supreme Court invalidated gene patents in 2013 (*Assoc. for Molecular Pathology v. Myriad Genetics, Inc.*, 569 U.S. 576), which presumably removed any inhibitory effect of patents on follow-on genetic research.

Table 20: Impact of extraterritoriality on measures of follow-on innovation

	Number of publications	Number of clinical trials for BRCA1	Number of genes with diagnostic tests
US	1,135	166	4,644
Non-US	1,636	122	5,178

Sovereign immunity also applies to both clinical trials and diagnostic tests. Of the

US clinical trials involving BRCA1, 53 (32 percent) were based out of state- or federal-government affiliated hospitals. Diagnostic tests are also offered at government affiliated institutions,[27] although it is more difficult to get precise numbers on this.

Finally, the 271(e)(1) exception likely plays a significant role in exempting clinical trials from patent infringement. Section 271(e)(1) applies to research intended to gather information for FDA submission and almost all clinical trials are so intended. Genetic tests are less likely to fall into 271(e)(1), but many genetic tests are offered for research purposes—for instance as part of clinical trials.[28]

[27] For instance, BRCA1 diagnostic tests are offered at laboratories affiliated with the University of Michigan, the University of Oklahoma, and the University of Minnesota.

[28] There is debate about whether use of research tools falls into the 271(e)(1) exception. Compare *Proveris Sci. Corp. v. Innovasystems, Inc.*, 536 F.3d 1256, 1265 (Fed. Cir. 2008), with *Classen Immunotherapies Inc. v. Elan Pharms., Inc.*, 786 F.3d 892, 897 (Fed. Cir. 2015).

References

Anton, James J. and Dennis A. Yao. 1994. "Expropriation and Inventions: Appropriable Rents in the Absence of Property Rights." *American Economic Review* 84(1):190–209.

Arora, Ashish. 1995. "Licensing Tacit Knowledge: Intellectual Property Rights And The Market For Know-How." *Economics of Innovation and New Technology* 4(1):41–60.

Bessen, James. 2008. "The Value of U.S. Patents by Owner and Patent Characteristics." *Research Policy* 37(5):932–945.

Bessen, James and Eric Maskin. 2009. "Sequential Innovation, Patents, and Imitation." *RAND Journal of Economics* 40(4):611–635.

Day, Gregory and Steven Udick. 2019. "Patent Law and the Emigration of Innovation." *Washington Law Review* 94:119–170.

Eisenberg, Rebecca S. 2008. "Noncompliance, Nonenforcement, Nonproblem? Rethinking the Anticommons in Biomedical Research." *Houston Law Review* 45:1059–1099.

Farre-Mensa, Joan, Deepak Hegde and Alexander Ljunqvist. 2020. "What Is a Patent Worth? Evidence from the U.S. Patent "Lottery"." *Journal of Finance* 75(2):639–682.

Feng, Josh and Xavier Jaravel. 2020. "Crafting Intellectual Property Rights: Implications for Patent Assertion Entities, Litigation, and Innovation." *American Economic Journal: Applied Economics* 12(1):140–81.

Frakes, Michael D. and Melissa F. Wasserman. 2017. "Is the Time Allocated to Review Patent Applications Inducing Examiners to Grant Invalid Patents? Evidence from Microlevel Application Data." *Review of Economics and Statistics* 99(3):550–563.

Freilich, Janet. 2022. "Paths to Follow-on Innovation." *U.C. Davis Law Review* (forthcoming).

Furman, Jeffery L. and Scott Stern. 2011. "Climbing atop the Shoulders of Giants: The Impact of Institutions on Cumulative Research." *American Economic Review* 101(5):1933–1963.

Galasso, Alberto and Mark Schankerman. 2015. "Patents and Cumulative Innovation: Causal Evidence from the Courts." *Quarterly Journal of Economics* 130(1):317–369.

Gans, Joshua, David Hsu and Scott Stern. 2002. "When Does Start-Up Innovation Spur the Gale of Creative Destruction?" *RAND Journal of Economics* 33(4):571–586.

Gans, Joshua S., David H. Hsu and Scott Stern. 2008. "The Impact of Uncertain Intellectual Property Rights on the Market for Ideas: Evidence from Patent Grant Delays." *Management Science* 54(5):982–997.

Gaulé, Patrick. 2018. "Patents and the Success of Venture-Capital Backed Startups: Using Examiner Assignment to Estimate Causal Effects." *Journal of Industrial Economics* 66(2):350–376.

Green, Jerry R. and Suzanne Scotchmer. 1995. "On the Division of Profit in Sequential Innovation." *RAND Journal of Economics* 26(1):20–33.

Hall, Bronwyn H., Adam Jaffe and Manuel Trajtenberg. 2005. "Market Value and Patent Citations." *RAND Journal of Economics* 36(1):16–38.

Hall, Bronwyn H. and Rosemarie Ham Ziedonis. 2001. "The Patent Paradox Revisited: An Empirical Study of Patenting in the U.S. Semiconductor Industry, 1979-1995." *RAND Journal of Economics* 32(1):101–128.

Harhoff, Dietmar, Frederic M Scherer and Katrin Vopel. 2003. "Citations, Family Size, Opposition and the Value of Patent Rights." *Research Policy* 32(8):1343–1363.

Henry, Emeric and Carlos J. Ponce. 2011. "Waiting to Imitate: On the Dynamic Pricing of Knowledge." *Journal of Political Economy* 119(5):959–981.

Holman, Christopher M. 2012. "Will Gene Patents Derail the Next-Generation of Genetic Technologies?: A Reassessment of the Evidence Suggests Not." *UMKC Law Review* 80:563–607.

Hopenhayn, Hugo, Gerard Llobet and Matthew Mitchell. 2006. "Rewarding Sequential Innovators: Prizes, Patents, and Buyouts." *Journal of Political Economy* 114(6):1041–1068.

Jaffe, Adam B. and Gaétan de Rassenfosse. 2019. Patent Citation Data in Social Science Research: Overview and Best Practices. In *Research Handbook on the Economics of Intellectual Property*. Elgar.

Jaffe, Adam B., Manuel Trajtenberg and Michael S. Fogarty. 2000. "Knowledge Spillovers and Patent Citations: Evidence from a Survey of Inventors." *American Economic Review* 90(2):215–218.

Jensen, Kyle and Fiona Murray. 2005. "Intellectual Property Landscape of the Human Genome." *Science* 310(5746):239–240.

Kieff, F. Scott. 2001. "Property Rights and Property Rules for Commercializing Inventions." *Minnesota Law Review* 85:697–754.

Kitch, Edmund. 1977. "The Nature and Function of the Patent System." *Journal of Law and Economics* 20(2):265–290.

Kline, Patrick, Neviana Petkova, Heidi Williams and Owen Zidar. 2019. "Who Profits from Patents? Rent-Sharing at Innovative Firms." *Quarterly Journal of Economics* 134(3):1343–1404.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru and Noah Stoffman. 2017. "Technological Innovation, Resource Allocation, and Growth." *Quarterly Journal of Economics* 132(2):665–712.

Kortum, Samuel S. 1997. "Research, Patenting, and Technological Change." *Econometrica* 65(6):1389–1419.

Lanjouw, Jean O. and Mark Schankerman. 2001. "Characteristics of Patent Litigation: A Window on Competition." *RAND Journal of Economics* 32(1):129–151.

Lee, Tom and Louis L. Wilde. 1980. "Market Structure and Innovation: A Reformulation." *Quarterly Journal of Economics* 94(2):429–436.

Lemley, Mark A. and Carl Shapiro. 2005. "Probabilistic Patents." *Journal of Economic Perspectives* 19(2):75–98.

Lemley, Mark and Bhaven Sampat. 2008. "Is the Patent Office a Rubber Stamp?" *Emory Law Journal* 58:181–203.

Lemley, Mark and Bhaven Sampat. 2012. "Examiner Characteristics and Patent Office Outcomes." *Review of Economics and Statistics* 94(3):817–827.

Li, Danielle, Pierre Azoulay and Bhaven Sampat. 2017. "The Applied Value of Public Investments in Biomedical Research." *Science* 356(6333):78–81.

Loury, Glenn C. 1979. "Market Structure and Innovation." *Quarterly Journal of Economics* 93(3):395–410.

Malakoff, David. 2003. "Universities Ask Supreme Court to Reverse Patent Ruling." *Science* 299:26–27.

Meurer, Michael J. 1989. "The Settlement of Patent Litigation." *RAND Journal of Economics* 20(1):77–91.

Moon, Seongwuk. 2011. "How Does the Management of Research Impact the Disclosure of Knowledge? Evidence from Scientific Publications and Patenting Behavior." *Economics of Innovation and New Technology* 20(1):1–32.

Moser, Petra. 2005. "How Do Patent Laws Influence Innovation? Evidence from Nineteenth-Century World's Fairs." *American Economic Review* 95(4):1214–1236.

Moser, Petra. 2012. "Innovation without Patents: Evidence from World's Fairs." *Journal of Law and Economics* 55(1):43–74.

Mowery, David, Richard Nelson, Bhaven Sampat and Arvids Ziedonis. 2001. "The Growth of Patenting and Licensing by U.S. Universities: An Assessment of the Effects of the Bayh–Dole Act of 1980." *Research Policy* 30(1):99–119.

Murray, Fiona, Philippe Aghion, Mathias Dewatripont, Julian Kolev and Scott Stern. 2016. "Of Mice and Academics: Examining the Effect of Openness on Innovation." *American Economic Journal: Economic Policy* 8(1):212–52.

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 & Organization* 63(4):648–687.

Nordhaus, William. 1969. *Invention, Growth, and Welfare: A Thoeretical Treatment of Technological Change*. MIT Press.

Plant, Arnold. 1934. "The Economic Theory Concerning Patents for Inventions." *Economica* 1(1):30–51.

Reinganum, Jennifer F. 1982. "A Dynamic Game of R and D: Patent Protection and Competitive Behavior." *Econometrica* 50(3):671–688.

Sampat, Bhaven and Heidi L. Williams. 2019. "How Do Patents Affect Follow-On Innovation? Evidence from the Human Genome." *American Economic Review* 109(1):203–36.

Samuelson, Pamela. 2004. "Intellectual Property Arbitrage: How Foreign Rules Can Affect Domestic Protections." *University of Chicago Law Review* 71(1):223–239.

Scotchmer, Suzanne. 1991. "Standing on the Shoulders of Giants: Cumulative Research and the Patent Law." *Journal of Economic Perspectives* 5(1):29–41.

Shahshahani, Sepehr. 2018. "The Role of Courts in Technology Policy." *Journal of Law and Economics* 61(1):37–61.

Sherkow, Jacob and Ryan Abbott. 2018. "Fortune and Hindsight: Gene Patents' Muted Effect on Medical Practice." *British Medical Bulletin* 126(1):37–45.

USPTO. 2013. "U.S. Colleges and Universities - Utility Patent Grants 1969-2012." `https://www.uspto.gov/web/offices/ac/ido/oeip/taf/univ/doc/doc_info_2012.htm`.

USPTO. 2020. "U.S. Patent Statistics Chart Calendar Years 1963 - 2019." `https://www.uspto.gov/web/offices/ac/ido/oeip/taf/us_stat.htm`.

Walsh, John, Ashish Arora and Wesley Cohen. 2003. "Working Through the Patent Problem." *Science* 299(5609):1021.

Weyl, E. Glen and Jean Tirole. 2012. "Market Power Screens Willingness-to-Pay." *Quarterly Journal of Economics* 127(4):1971–2003.

Williams, Heidi L. 2013. "Intellectual Property Rights and Innovation: Evidence from the Human Genome." *Journal of Political Economy* 121(1):1–27.