Paper to be presented at the DRUID Academy Conference 2017 at University of Southern Denmark, Odense, Denmark on January 18-20, 2017

The Impact of Intellectual Property Rights on Cumulative Innovation: Evidence from Biotechnology Industry

Sina Khoshsokhan
Boston University
Strategy and Innovation
sina@bu.edu

Abstract

The Impact of Intellectual Property Rights on Cumulative Innovation: Evidence from Biotechnology Industry Sina Khoshsokhan Questrom School of Business, Boston University Year of enrollment: 2013 Expected final date: May 2018 Email: sina@bu.edu Whether intellectual property rights encourage or hinder cumulative innovation has been a subject of debate. On one hand, proponents of the positive impact claim that intellectual property rights provide incentives for researchers and private investors and direct their efforts. On the other hand, proponents of the negative impact (or so-called anti-commons effect) argue that intellectual property rights limit the ability of researchers to build upon existing innovations and technologies. Although several recent empirical works find evidence in support of anti-commons effect, the focus of these studies has been primarily on the "existing disclosed technologies". In spite of their prominent role in deepening our understanding of the relationship between intellectual property rights and cumulative innovation, such research designs merely allow for studying the "accessibility" aspect of innovations and fall short of examining other dimensions (e.g. reward and disclosure). In order to address this gap, in this study, I exploit a natural experiment: The US Supreme Court's ruling on "Association of Molecular Pathology v. Myriad Genetics, Inc." case which led to the termination of gene patenting practice in the United States. I examine how the abolition of intellectual property rights on isolated genes has affected not only the subsequent innovations based on existing isolated genes, but also the incentives to develop new gene-based diagnostic tests and pushing the frontiers in biotechnology industry. Collecting data on patented genes from multiple sources (e.g. National Center for Biotechnology Research (NCBI), PubMed, genetests.org, etc.), I employ a difference-in-difference method to estimate the impact of intellectual property rights not only on follow-on innovations (which has been the subject of prior empirical studies), but also on new-to-the-world innovations (which is the main contribution of this paper). While I find no evidence on removing intellectual property rights affecting scientific research progress (measured by publications), its impact on commercialized innovations (measured by gene-based diagnostic tests) is significant. I find that while removing intellectual property rights increases the proliferation of diagnostic tests (follow-on innovations) by %11-%21 in different specifications, it leaves negative impacts on the development of new-to-the-world diagnostic tests (novel innovations), by %3-%17 in different models.

Jelcodes:M10,-

The Impact of Intellectual Property Rights on Cumulative Innovation: Evidence from Biotechnology Industry

Sina Khoshsokhan
Questrom School of Business, Boston University

October 2016

Abstract

Whether intellectual property rights encourage or hinder cumulative innovation has been a subject of debate. On one hand, proponents of the positive impact claim that intellectual property rights provide incentives for researchers and private investors and direct their efforts. On the other hand, proponents of the negative impact (or so-called anti-commons effect) argue that intellectual property rights limit the ability of researchers to build upon existing innovations and technologies. Although several recent empirical works find evidence in support of anti-commons effect, the focus of these studies has been primarily on the "existing disclosed technologies". In spite of their prominent role in deepening our understanding of the relationship between intellectual property rights and cumulative innovation, such research designs merely allow for studying the accessibility aspect of innovations and fall short of examining other dimensions (e.g. reward and disclosure). In order to address this gap, in this study, I exploit a natural experiment: The US Supreme Courts ruling on "Association of Molecular Pathology v. Myriad Genetics, Inc." case which led to the termination of gene patenting practice in the United States. I examine how the abolition of intellectual property rights on isolated genes has affected not only the subsequent innovations based on existing isolated genes, but also the incentives to develop new gene-based diagnostic tests and pushing the frontiers in biotechnology industry.

1 Introduction

How do intellectual property rights affect cumulative innovation? Theoretical answers to this question come from, at least, two camps. On one hand, the proponents of the positive impact argue that intellectual property right systems provide incentives to motivate innovation. While due to non-rivalrousness and non-excludablity features of ideas, competitive markets are not able to effectively incentivize innovation, patent systems (as a form of intellectual property rights regime) motivates innovators through granting them a short-term monopoly on the innovation. By excluding others from commercializing their innovations, innovators are able to earn economic returns on their ideas, either by commercializing or through licensing them. In this regard, Arora et al. (2001) highlight the prominent role that patent systems play in facilitating the functioning of markets for technologies. They maintain that in a market where intellectual property rights are well-recognized, patent assignees are able to sell or license their patents and, in exchange for a license fee, allow other innovators to utilize or build upon their technologies and ideas. Such an environment allows innovators to earn returns even in cases where they are not able to commercialize their innovation themselves. Gans and Stern (2003) share the same idea and drawing on literature on profiting from innovation (Teece 1986) introduce a set of strategies to innovators to choose based on their access to complementary assets required for commercializing their innovations and the strength of intellectual property rights regime. In their framework, individual innovators and small entrepreneurial firms, as two of the main engines of innovation, usually lack the complementary assets (e.g. manufacturing processes, marketing capabilities, etc.) required for producing a product (or provide a service). The existence of intellectual property rights regimes allows these small entities to engage in negotiations with established firms who own the needed complementary assets on commercialization of the innovation, without fearing that their ideas would be stolen. This also enables small firms to become so-called "*idea factories*" - focusing only on developing ideas, licensing (or selling) them, and leaving the commercialization to established firms. An environment with such characteristics is prone to experience significant growth in innovation.

On the other hand, advocates of the so-called "*anti-commons effect*" propose a negative relationship between patent rights and cumulative innovation (Heller and Eisenberg 1998). According to this perspective, the existence of concurrent and fragmented intellectual property rights can deter cumulative innovation. This theory maintains that innovation does not happen in isolation, but by building upon prior innovations (Scotchmer 1991) and in environments where such fragmented and concurrent intellectual property rights exist, innovators have to negotiate with several different patent holders in order to be able to build on their innovations. This would increase the innovation cost and make the entire process susceptible to hold-up problem. This perspective sheds light on the flip side of intellectual property rights regime which enables

patent holders to aggressively enforce their intellectual property rights and bar others from building upon their innovations.

While the theoretical arguments in this debate are well-established, the empirical tests for these arguments are not more than a handful. These empirical studies, which will be discussed in more details in section 2, provide a variety of (sometimes even opposite) findings. As we will discuss later, this might be due to the heterogeneity in the existence and magnitude of anti-commons effect among different industries and settings. Another reason may be the changing behavior of mechanisms through which these effects are at work. Although the magnitudes of anti-commons effect have been measured in different settings, we know little about the processes that eventually lead to these impacts. Such processes may result in different outcomes in different settings. Finally, a large part of empirical studies have focused on removing intellectual property rights (or granting them) from (to) "*existing innovation*" that are already "*disclosed*" to public. In fact, the counterfactual in these studies is the disclosed innovations with (or without, depending on the research design) intellectual property rights. These settings, however, falls short of allowing the researcher to test some functions of intellectual property right, such as the reward that they provide in return to disclosing the innovation to public. In this study, I aim to address this final concern. I hypothesize that although removing intellectual property rights from innovations provides free access to their underlying ideas and consequently facilitates building on these innovations, nonetheless, it motivates innovators to not disclose their innovations. It also does not incentivize them to innovate, and these may eventually leave negative impacts on cumulative innovation.

In order to empirically test these impacts, I examine the United States Supreme Court's ruling on Association of Molecular Pathology v. Myriad Genetics, Inc. The decision on this case which was announced in June 2013 put an end to gene patenting practice. In this study, I take this ruling as an exogenous shock to biotechnology industry and take the whole setting as a quasi-experiment. Using a difference-in-difference estimation method, then, I estimate the impact of lifting patenting rights from isolated genes on cumulative innovation in this industry. This paper joins the set of empirical papers exploiting natural experiments to examine the relationship between intellectual property rights and cumulative innovation. A more detailed review of these studies is provided in section 2.

The organization of the paper is as follows. In section 2, a review of previous studies in the field, together with the hypotheses are provided. The context of the natural experiment is described in Section 3. Data construction and empirical methods are discussed in Section 4. In section 5, the results are provided. Section 6 concludes.

2 Theory and Hypotheses

To the best of my knowledge, the first empirical examination of the anti-commons theory dates back to 2003 when Walsh et al. (2003) interviewed 70 scientists, IP attorneys and business managers in biotechnology industry about the impact of intellectual property rights on the progress of their projects. In most cases, they did not find any significant evidence for intellectual property rights to create obstacles for progress of projects, particularly in university research, where companies are much more lenient and rarely enforce their patent rights. They found, however, an important exception to this rule that is patents on "targets" which, as they describe it "refers to any cell receptor, enzyme, or other protein implicated in a disease, thus representing a promising locus for drug intervention" (e.g. gene patents). Walsh and coauthors document several complaints regarding gene patent holders refusing to allow access to patented genes for research purposes, arguing that this potentially leads to commercial gains.

Using quantitative empirical methods, Fionna Murray, Scott Stern, and colleagues started a stream of works to examine the salience of anti-commons effect in a variety of settings. Murray and Stern (2007) take patent-paper pairs as unit of analysis and design a quasi-experiment to assess the impact of granting intellectual property rights on cumulative innovation. They exploit the *dual-disclosure method* in science, through both journal publications and patents, and select a sample of articles in the journal of "Nature Biotechnology" over the period of 1997-1999 to which a patent is also granted. By comparing the citations these papers receive before and after patents are granted, Murray and Stern find that despite the initial boost following the grant of patents, there is a 10 to 19 percent decrease in citations that articles receive after patents are granted. These results support the existence of anti-commons effect in biotechnology industry.

In a similar vein, but taking an opposite approach, Murray et al. (2009) examine the impact of "openness" on cumulative innovation. Their study sheds light on the other side of the effects of intellectual property rights: what happens to cumulative innovation when intellectual property rights are lifted. Studying the outcomes of a NIH-imposed openness shock, namely a set of agreements between NIH and DuPont to give university researchers access to a large number of genetically-engineered mice, Murray and colleagues find that such a shock leads not only to a significant increase in the follow-on research, but also opens new avenues to increase the diversity of subsequent research streams. In more precise terms, overall, the openness shock leads to 21 percent increase in annual citations that the "mouse-articles" (those affected by the shock) receive. This boost is particularly salient where novel realms of research are addressed (articles that include new keywords). This article documents more evidence supporting the anti-commons effect.

Although not directly addressing the anti-commons effect, Furman and Stern's (2011) findings also give insight regarding the role that making research materials accessible to researchers plays in boosting

follow-on knowledge creations. They underline the role of institutions in making research materials available for university researchers and its positive impact on the value of the subsequent research (captured by forward citations). Furman and Stern document a 57-135 percent increase in forward citations in the published articles that have used research materials provided by biological research centers (i.e. institutions that "collect, certify, and distribute biological organisms, such as cell lines, microorganisms, and DNA material"). These results indicate that when institutions make research inputs available for researchers, the value of research sharply increases. A similar case could be made for research inputs that are not-accessible for researchers due to intellectual property rights: giving access to such research inputs (e.g. gene patents) would lead to increase in the value of research outputs.

Williams (2013) also adds to the discussion by linking the impact of lifting intellectual property rights from a certain set of genes (those owned by a private company named Celera) on follow-on gene-based diagnostic tests. This paper is particularly important in this debate, because the outcome variable is not merely forward citations to a publication, which could be a noisy measure for estimating the importance of follow-on research (Cole 2000), but it also includes actual commercialized products of innovation (gene-based diagnostic tests). She finds a lingering decrease in both publications and product developments (between 20 to 30 percent) that build upon genes that were protected by intellectual property rights, even for short periods of time.

Galasso and Schankerman (2015) broaden the spectrum of settings by testing anti-commons theory in several different industries. They take advantage of the random assignment of judges at the United States Court of Appeals for the Federal Circuit to develop an instrumental variable called Judges Invalidity Propensity (JIP), measuring the likeliness of a judge to invalidate a patent in response to litigation. Basing their identification strategy on this instrumental variable, they find an average of 50 percent increase in forward citations to a patent after its invalidation. The more interesting finding of their study, however, is the heterogeneity of this impact among different industries. While they find no evidence of significant increase in citation rates following patent invalidations in chemical, mechanical, and drugs fields, patent invalidation are estimated to raise those citation rates by 178 percent in computers, 203 percent in electronics, and 320 percent increase in medical instruments and biotechnology.

Galasso and Schankerman's findings in electronics and computer industry is supported also by Watzinger et al. (2016). They take an antitrust lawsuit that forced Bell Labs to license all its patents that were published before 1957 for free as a natural experiment to estimate its impact on follow-on innovation. Their analysis show that lifting intellectual property rights from these patents has increased their citation rates by 9 percent. It has also led to 23 percent boost in patent subcalsses that were affected by this lawsuit.

Finally, and more interestingly, Sampat and Williams (2015) address the question of anti-commons

effect by comparing subsequent research and product developments on isolated genes that are protected by patent rights with those whose patent applications are rejected by USPTO. Their initial analysis indicates that these genes are similar in research investments, in terms of scientific publications, and therefore comparable. Further analyses show that product developments, both in terms of gene diagnostic tests and drug development based on these genes do not significantly differ between these two groups, meaning that intellectual property rights do not affect subsequent innovations. This is in sharp contrast with the findings of previous empirical studies, most notably with those of one of the coauthors: Williams's (2013) study. They argue this difference is the result of difference in settings: While Williams (2013) studies lifting property rights from genes that were patented by only one private firm (Celera), Sampat and Williams (2015) compare genes that are patent protected with those that are not (although for which a patent application has been filed). Based on their findings, Sampat and Williams particularly question the arguments of the proponents of terminating gene patenting practice on the grounds of its role in impeding follow-on research and innovation which led to the Supreme Court's ruling on the AMP v. Myriad (the setting of the current study).

Based on the discussions above, therefore, I propose:

Hypothesis 1: Removing intellectual property rights from an innovation leads to higher rates of follow-on innovations based on that innovation.

All the articles discussed above are based on clear identification strategies, taking advantage of natural experiments which enable the authors to interpret causal relationships between explanatory and outcome variables. However, as briefly discussed in section 1, their research settings allow only to compare existing disclosed innovations that are patent-protected with existing disclosed innovations that are not. Such a research design is instrumental in examining the extent to which "accessibility" of prior innovations help in developing new innovations. It falls short, nevertheless, of testing for other dimensions of cumulative innovation.

As Murray and O'Mahony (2007) describe in their conceptual model, for the processes of cumulative innovation to work properly, at least three conditions should be met:

1. *Disclosure* of prior (generations of) knowledge: making the knowledge available for innovators who need it as an important input for their innovation,

2. *Accessibility*: because availability of knowledge does not necessarily confer its accessibility to reuse. Innovators need to be able to reuse or apply the prior innovations in order to innovate cumulatively.

And finally,

3. *Rewards* for both disclosure and providing access: Innovators should be incentivized to make the underlying knowledge in their innovation available and accessible to other innovators.

From these three conditions, the research designs of the aforementioned empirical studies, merely allow for testing the accessibility condition through exogenously removing or imposing patent rights. However, when intellectual property rights regimes are removed in industry-levels, the other aspects like disclosure and reward are also affected. In such settings, where innovations are not protected by intellectual property rights, innovators tend to keep their ideas as secrets and are less motivated to disclose them. This would interrupt knowledge flows and may leave negative impacts on cumulative innovation. Moreover, in these contexts, entrepreneurial firms lose one of their most valuable resources, patents, and it may drastically affect their leverage in competition with other firms. Putting small and entrepreneurial innovative firms at risk of losing the competition, can significantly mitigate innovation rates in an industry. Although examining each of these mechanisms and how they impact cumulative innovation is out of the scope of this study, I test for the extent to which removing patent rights would affect the development of "new-to-the-world" innovations. I expect this impact to be negative

Based on the discussions above, therefore, I hypothesize:

Hypothesis 2: Removing intellectual property rights from an industry leads to the emergence of fewer novel innovations in that industry.

3 Empirical Context

On June 13, 2013, the Supreme Court of the United States announced its decision on the "Association of Molecular Pathology v. Myriad Genetics, Inc." case. According to this decision, isolated genes were considered to be products of nature and therefore, based on Section 101 of Title 35 of the United States Code, they were not patent-eligible subject matter. This ruling invalidated the claims of Myriad Genetics, Inc. over BRCA1 and BRCA2 genes, together with the patent claims of other entities over isolated genes. The United States Patents and Trademarks Office promptly issued a new guideline, maintaining that they no longer issue patents on isolated nucleic acids. The ruling excluded, however, the claims on complementary DNAs (cDNAs) from invalidation on the basis that cDNAs are human-made and are not naturally occurring.

This decision put an end to gene patenting practice which claimed to be one of the pillars of growth in biotechnology industry.

Gene patenting practice started in 1982, when Howard Goodman, John Shine, and Peter Seeburg of University of California, Berkeley, filed a patent application for "Microorganism Containing Gene for Human Chorionic Somatomammotropin". The patent (Patent No.: 4,447,538) was issued on May 8, 1984. The Supreme Court of the United States' decision on Diamond v. Chakrabarty case in 1980, had made it possible for inventors to patent "genetically modified organisms" and it paved the path for biotechnology researchers to patent isolated genes. By 2005, reportedly more than 20 percent of human genome was patented (Jensen and Murray 2005). Granting patents on isolated genes provided legal rights for the assignee to exclude others from performing scientific activities with the gene, including developing diagnostic tests, producing therapeutic drugs, etc. (Park 2013).

Among the firms that possessed gene patents and aggressively enforced their patent rights, Myriad Genetics, Inc. was of particular prominence. Myriad owned patents over BRCA1 and BRCA2 genes. Certain mutations of these genes could make the human body susceptible to breast and ovarian cancer. Enforcing its intellectual property rights, Myriad did not allow other institutions to develop *BROCA* (breast and ovarian cancer) tests, which had led to a monopoly, allowing Myriad to charge high prices (about $4,000) for its tests. Myriad also refused to update its tests that produced a considerable number of false negative and false positive results.

In response to this aggressive patent enforcement, the American Civil Liberties Union (ACLU) started a campaign against Myriad and, in a broader sense, against gene patenting practice. They litigated Myriad's claims over BRCA1 and BRCA2 and in March 2010, the United States District Court for the Southern District of New York ruled in favor of ACLU, declaring a number of Myriad's claims over BRCA1 and BRCA2 genes invalidated. Myriad appealed this decision in the United States Court of Appeals for the Federal Circuit which led to the overturn of the District Court's decision and Myriad regained its patent rights over the two genes.

In response, The Association for Molecular Pathology petitioned the case to the Supreme Court with a simple question: "Are human genes patentable?" SCOTUS agreed to hear the case and their meeting was held in April 2013. On June 13, 2013 the Supreme Court announced its decision on the case, ruling against gene patenting practice, only with the exception of cDNAs. The Supreme Court's decision holds that "A naturally occurring DNA segment is a product of nature and not patent eligible merely because it has been isolated, but cDNA is patent eligible because it is not naturally occurring".

The Supreme Court's ruling on AMP v. Myriad case, provides a setting for quasi-experiment analysis. First, by lifting intellectual property rights from isolated genes and preserving those rights for cDNAs, this

ruling creates control and treatment groups. In this setting, on one hand, genes that have been already protected by patent rights over both genomic DNA and complementary DNA act as a control group. These genes lose their claims on genomic DNAs after the ruling, but retain their intellectual property rights over their cDNAs and at least to some extents preserve their intellectual property rights. On the other hand, genes for which a cDNA patent has never been obtained and according to the ruling, could not be protected by any patent rights, constitute the treatment group.

One important feature of this ruling is that the main logic behind dividing gene patents into two groups of genomic DNAs (treatment group) and cDNAs (control group) is merely their "patent-eligibility" which is not related to outcome variables such as the importance of patent in the commercialization processes, the citations that it has received, etc. This feature is of high importance in arguing for the shock to be strictly exogenous.

One other important concern, here, is that whether the exogenous shock was predicted by the subjects of experiment. This concern is particularly pronounced due to the fact that the decision of Supreme Court on the case has been unanimous. Supreme Court Justices, with ideas closer to either party, all have agreed on this decision and such a comprehensive agreement over nonpatentability of genes raises the concerns that the decision could be predicted to be in favor of ACLU, even before the announcement. Both qualitative and financial data negate this assumption.

A quick look at reports published after the case hearing shows that the Supreme Court justices seemed to be taking different positions on the issue of patent-eligibility of human genes. On May 14, 2013 Los Angeles Times reports: "[...]justices seemed split. Some wondered about the implications of allowing patents for items extracted from nature. Would that logic apply to a plant found in the Amazon jungle? Or to an entire chromosome extracted from a cell? Others feared that companies wouldn't be willing to make the investment needed to isolate sequences if they couldn't patent the results." [1] CQ Researcher publishes a similar note about the division among justices on the day of hearing: "Justices grappled with the science and the law during the hour-long argument on April 15, but several seemed skeptical of allowing a patent for a gene naturally found in the human body even if scientists had to work to isolate it. Other justices, however, voiced concerns that researchers needed the prospect of patent protections to provide economic incentives for their work." [2] Finally, Wall Street Journal reports that while Justice Sotomayor uses chocolate-chip cookies as an analogy to argue that its natural elements like flour and eggs are not patentable "simply because I've created a new use or a new product from those ingredients", Justice Kennedy highlights the concerns about invalidating patents and its implications: "I just don't think we can decide the case on the ground, 'Oh,

[1] http://articles.latimes.com/2013/may/14/news/la-ol-angelina-jolie-gene-patents-20130514
[2] Jost, K. (2013, May 31). Patenting human genes. CQ Researcher, 23, 473-496

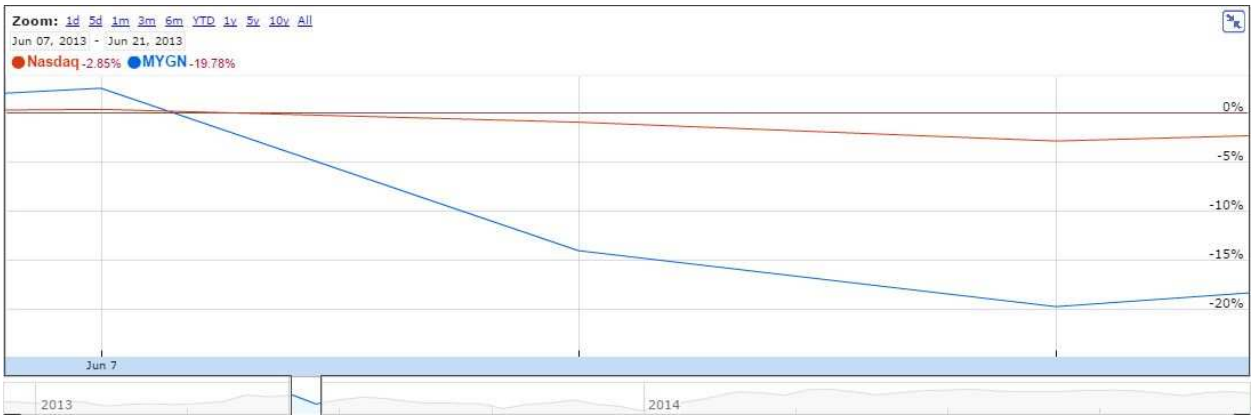


Figure 1: Myriad Genetics share price. The graph is adopted from google.com/finance

don't worry about investment". [3] Such disagreements in the hearing session implies that it is very unlikely that the decision was predicted before the announcement.

On the financial side, also, figures indicate that such a decision was not anticipated by the companies that owned gene patents. Figure 1 shows the share value of Myriad Genetics over the period between Jun 6, 2013 (one week before the announcement) and June 20, 2013 (one week after the announcement). The graph clearly indicates a significant drop in share value of the company (more than 20 percent), showing that Myriad, as the major stakeholder in this case, did not anticipate such a ruling to get prepared for the consequences.

On this basis, in this study I consider the case as a strictly exogenous shock and exploit it as a quasi-experiment. The details on empirical methods and data structure are provided in the next section.

4 Data and Methods

4.1 Empirical Models

I use a difference-in-difference model to estimate the impact of removing intellectual property rights on cumulative innovation in biotechnology industry. In the first model, I use publications associated with each gene as a measure of scientific innovation. This is not a perfect measure for capturing innovation, at least for two reasons: first, not all scientific discoveries are published in journals, especially in a thriving and competitive industry like biotechnology. Companies have strong incentives to keep their novel findings, in particular those that are not legally protected, within the firm and do not engage in activities that would result in knowledge spillover from the firm. Second, not all publications fall into the category of innovation.

[3] http://www.wsj.com/articles/SB10001424127887324485004578424782830965300

While some articles introduce breakthrough ideas, others are not as innovative. Usually this asymmetry is dealt with by using forward citations as a measure of article's quality or its importance in the field. This measure could not be used in my research design due to the fact that at the time I collect data only 3 years have been passed from announcement of the Supreme Court's ruling. In such a short period of time after the shock, the citation data are subject to truncation. In spite of these issues, however, publication rates associated with each gene can be taken as a noisy, but useful, measure of innovation. In this study, I use articles published in PubMed in each quarter, and use it both as outcome variable (in the first model) and control variable (in the second model).

In the first model, I compare the publication rates for isolated genes that were protected by patent rights before the AMP v. Myriad case ruling with the genes that had patents on both isolated gene and cDNA, before and after the decision on the case is announced. The model is:

$$Gene\_Articles_{it} = f(\epsilon_{it}; \beta_i + \lambda_t + \gamma Treated_{it} + X_{it})$$

Where β_i captures the fixed-effect for each gene, λ represents time dummies for each period, $Treated$ is a dummy variable equal to one for isolated genes after the shock (when their patent protection is removed) and zero otherwise, and finally X is the vector of control variables. Since $Gene\_Articles$ is a count variable, its distribution is highly skewed. Therefore, I use a Poisson model in this analysis.

The second model examines the development of gene-based diagnostic tests as a measure of innovation. I examine two outcome variables to test the two hypotheses. First, I study the change in number of diagnostic tests offered in the U.S. after the shock. This variable measures the proliferation of diagnostic tests after the removal of patents from genes and tests Hypothesis 1. Second, I study the novel diagnostic tests that have been developed after the shock and did not exist before the shock. The research design allows me to compare the two groups based on this measure and examine how removing intellectual property rights has affected the development of these novel innovation. As I will describe later in the Data Construction section, one limitation of "genetests.org", from which I scraped data for gene-based diagnostic tests, is that it only offers data in one cross-section. Although, using "archive.org", I was able to collect data from genetests.org in 2014 too and use it as a baseline for comparison of the new versus old diagnostic tests, I was not able to create a panel dataset for diagnostic tests before and after the shock. Therefore, my analyses here are cross-sectional. The models that I use for testing the hypotheses are:

$$Diagnostic\_Test_i = f(\epsilon_i; \gamma Treated_i + X_i)$$

$$NEW\_Diagnostic\_Test_i = f(\epsilon_i; \gamma Treated_i + X_i)$$

Since the outcome variables are dummies, I use both OLS and Logit regressions for this set of analyses. The control variables the I use in gene-level include patent age, patent age squared, number of diseases predicted by mutations of the gene, the importance of these diseases (one for genes that help in predict severe diseases like HIV AIDS, Alzheimer, different types of cancer, etc.) and chromosomal location of the gene.

4.2 Data Construction

To identify patented genes, I followed Jensen and Murray (2005) and took advantage of the nucleotide sequences that "DNA-patent" applications are required to include in their texts since 1990. In patent claims section, these nucleotide sequences are referred to by their ID as "SEQ ID". Information on these patents together with the set of their nucleotide sequences is sent to GenBank on a weekly basis and is publicly available. Using keywords such as "PATB" and "SEQ ID", I searched for these patents and downloaded an aggregate of more than 7,000,000 sequences (not necessarily unique) in more than 45,000 patents in September 2016. In order to identify human genes, I selected nucleotide sequences that consisted of at least 150 nucleotides, because that is the average length of a human exon (Jensen and Murray 2005, Lander 2001). I then aligned these nucleotide sequences using an algorithm that National Center for biotechnology Information (NCBI) provides (Zhang et al. 2000) to identify the gene symbols associated with each sequence. I then selected the genes that included "Human", "Homo Sapien", or "H. Sapien" in their title to limit my sample only to the human genome. In these titles, I searched for the keyword "cDNA" and categorized these genes as my control group. This group includes 207 unique genes from the 870 genes that were found. I compare this control group with a treatment group of patented human genes, including 420 unique genes.

As my outcome variables, I first studied the publication counts associated with each gene in each quarter and second I examine diagnostic tests. The latter measure has been used in a number of studies (e.g. Sampat and Williams 2015 and Williams 2013) as one of the main ways of commercializing innovation in biotechnology industry.

Publications were collected from searching in the "PubMed" database in NCBI. I wrote a script to search for each gene and scrape the data on the number of associated articles published in PubMed in each quarter from Spring 2009 through Summer 2016. This time period covers publication data from 3 years before and 3 years after the decision on AMP v. Myriad is announced.

The data on gene-based diagnostic tests were collected from the website "genetests.org". This website gathers a large database including diagnostic tests, the clinics that provide these tests, their locations, the method they use for testing, and an approximate time for receiving the test results. Although not a complete database of all gene tests, "genetests.org" provides a large dataset on diagnostic gene tests, including more than 63,000 diagnostic tests for more than 5,600 genes by Summer 2016. As mentioned before, a major limitation of this database is that it only provides cross-sectional data and not a panel dataset. Researchers that exploit this database usually define a dummy variable equal to one in cases where a diagnostic test exists for the gene and zero otherwise, and use it as a binary outcome variable in cross-sectional analyses (e.g. Sampat and Williams 2015, Williams 2013). In this study, however, I take advantage of the website "www.archive.org" to collect data on a date closer to the Supreme Court's ruling announcement. *archive.org* collects snapshots from a large set of websites to create a library of front pages of these websites over years. According to its reports, *archive.org* has stored data for more than 237 billion webpages on the internet. I exploit this feature to go back in the history of "genetests.org" and collect data on diagnostic tests on a date closer to the shock. The earliest date that I could collect data for was July 2, 2014. It is one year after the shock, but given the time it takes to develop a diagnostic test, it can be used as the baseline for identifying newly developed gene tests and proliferation analyses. It allows me to define a novel variable, "NEW diagnostic test", that captures genes for which a diagnostic test did not exist in July 2014 and they appeared by September 2016. I use this variable to measure the rate of innovation in terms of developing a new gene-based diagnostic test in the industry.

5 Results

Table 1 presents the results of publication rate analyses. Model 1 shows the estimates for OLS regression while Model 2 presents the estimates for Poisson regression. The standard errors in Model 2 are clustered in gene level. As shown in Table 1, the estimates in neither model is statically significant at 90 percent level and therefore I am not able to reject the null hypothesis that removing intellectual property rights does not impact publishing behavior.

These results, also, can be visually deduced from Figure 2, where the density distributions of the articles are shown for control and treatment groups, both for pre- and post-shock periods. As shown in Figure 2, the density distributions for the two groups have similar shapes both before and after Supreme Court's decision is announced. These plots imply that the shock has not affected journal publication rates.

A similar pattern also can be seen in Figure 3, where the average numbers of articles in the two groups over years are plotted. Although the levels differ for the two groups, but the trends are very similar, even

Table 1: The impact of removing intellectual property rights on publication rates

VARIABLE	(1) OLS Gene articles	(2) Poisson Gene articles
Treated	0.00507	0.00244
	(0.0152)	(0.0133)
Controls	Yes	Yes
Gene fixed-effects	Yes	Yes
Time dummies	Yes	Yes
Observations	16,302	16,302
R-squared	0.132	
Number of genes	627	627

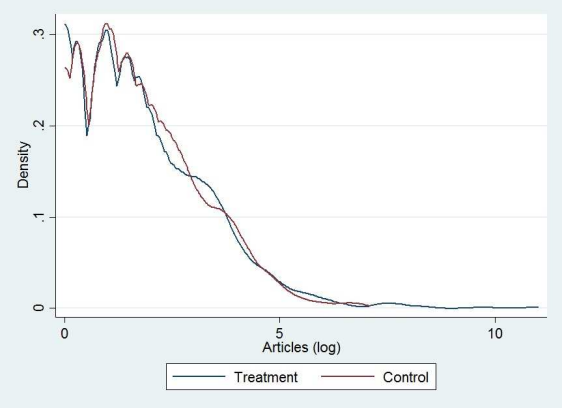
Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

after the Supreme Court's decision is announced (showing by a green line in the figure), meaning that this ruling has not drastically affected the publishing behavior in the sample.

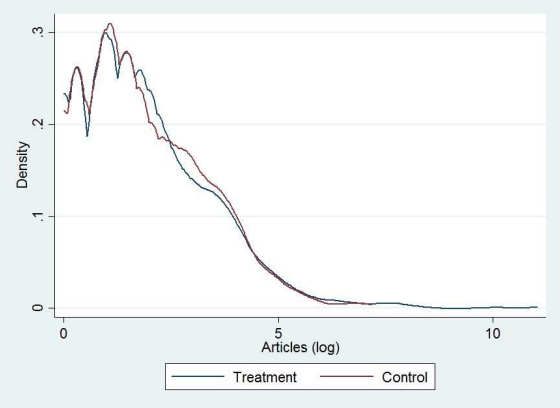
These results are in line with the findings of Sampat and Williams (2015) that the existence of intellectual property rights over genes does not affect publication rates in the biotechnology industry. It, however, contradicts with the findings of Walsh et al. (2003) who report gene patents to impede scientific research progress.

In the second set of analyses, I take gene-based diagnostic tests as the outcome variable and test the two proposed hypotheses using these variables. For testing Hypothesis 1, I examine the proliferation of diagnostic tests for each gene after the shock. As shown in Table 2, I study the difference between the number of diagnostic tests offered in the U.S. between July 2014 (the earliest date for which I could collect data) and October 2016.

Model 1 presents the results for OLS regression. The estimate for the "Treated" variable has positive sign and is statically significant at 95 percent significance level, showing that removing intellectual property rights has led to increase in the number of diagnostic tests offered in the U.S. According to this model, removing intellectual property rights from isolated genes has led to between 4 to 5 more diagnostic tests for each gene in the U.S. on average. Since the outcome variable is a count variable and its distribution is disperse, I repeat the same analysis , this time using a Poisson regression model. The estimates for this



(a) Pre-shock



(b) Post-shock

Figure 2: Density distribution for article counts (log)

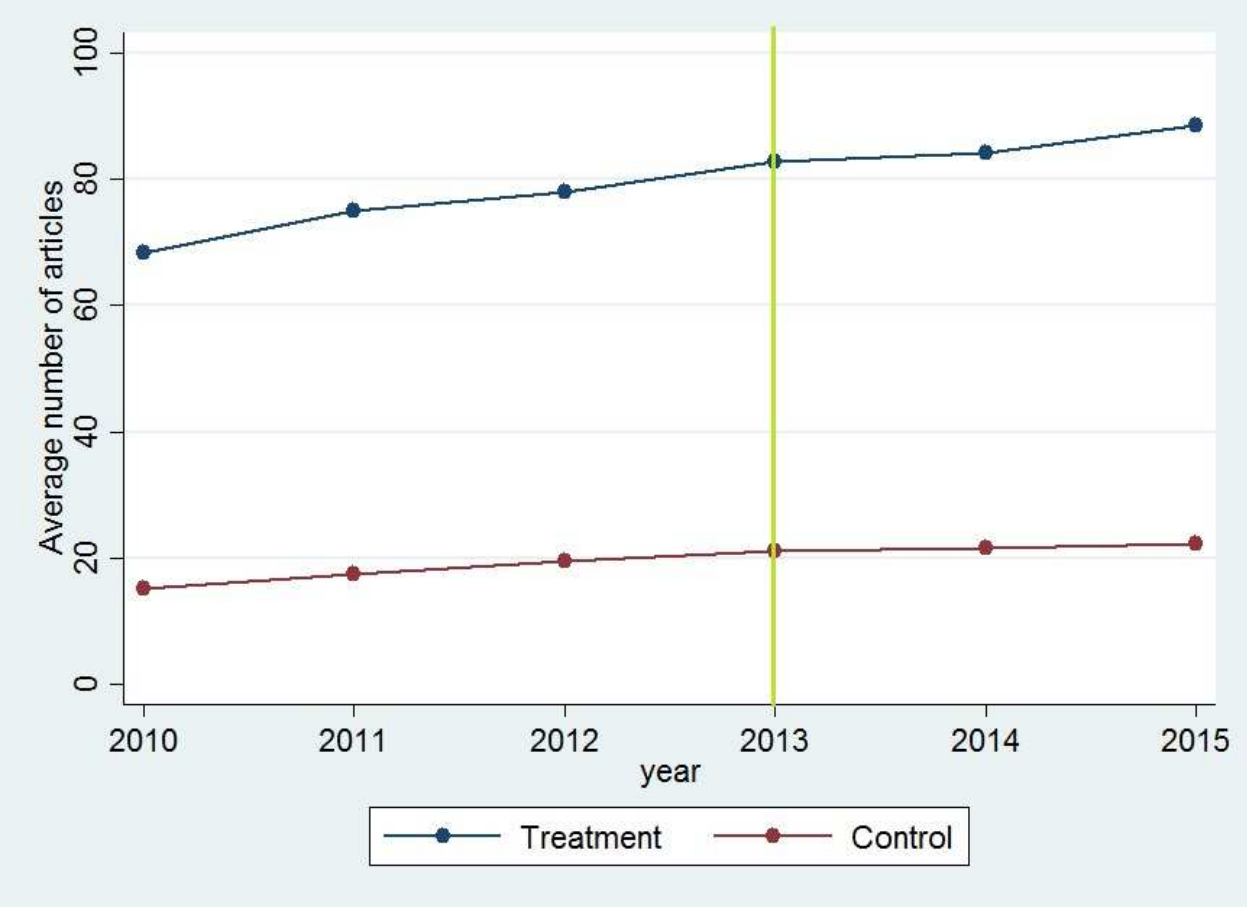


Figure 3: Average publication counts from 2010 to 2016

model are shown in Table 2 Model 2. This model shows that removing intellectual property rights from genes, on average, has led to 21 percent increase in the number of offered diagnostic tests in the U.S.

Because the difference between two numbers is a raw measure and may imply different growth rates in different levels (the difference between 7 and 9 is the same as the difference between 1007 and 1009, but the growth rates are different), I complement this analysis with another outcome measure: the ratio of number of diagnostic tests in 2016 over those in 2014. The results of these analyses are presented in Table 2 Models 3 and 4. While the estimate of OLS regression is not statically significant, the estimate of Poisson regression (Model 4) shows an 11 percent growth in gene-based diagnostic tests offered in the U.S. Models 1, 2, and 4 therefore support Hypothesis 1, and show removing intellectual property rights from genes has led to the considerable proliferation of diagnostic tests.

Finally we test Hypothesis 2 to examine the impact of removing patent rights on the development of "new-to-the-world" diagnostic tests. For this purpose we limit our sample to the genes that did not have diagnostic test in 2014 and compare the two groups of treatment and control to see whether there is a systematic difference between the two in terms of development of "new" diagnostic tests by 2016. The results are presented in Table 3. The estimates in these models are negative and statically significant, showing that removing intellectual property rights has left negative impacts on the development of novel diagnostic tests. In particular, the results for the Logit model (presented in Table 3 Model 2) show that for "non-patent-protected" genes the number of newly developed genes is 17 percent less than that of the genes that experience some sort of intellectual property protection. These results support Hypothesis 2.

6 Conclusion

In this study, I examined the impact of intellectual property rights on cumulative innovation. Drawing on the literature on anti-commons effect and building on previous empirical works, I studied the extent to which removing patent rights from isolated genes, following Supreme Court's decision on AMP v. Myriad case, has impacted both follow-on and novel innovations in biotechnology industry. I found no evidence for intellectual property rights to affect publishing behavior in the field. For gene-based diagnostic tests, however, I found that removing patent rights from isolated genes has positively affected the number of diagnostic tests for those genes. This result supports the premises of anti-commons effect, namely intellectual property rights hinder the follow-on innovation. Nonetheless, as my analyses for the novel diagnostic tests suggest, lack of intellectual property rights has led to the development of fewer new diagnostic tests for the genes affected by the Supreme Court's decision. These results together shows that although removing intellectual property rights improves accessibility of knowledge for follow-on innovations, it decreases the

Table 2: Proliferation of gene-based diagnostic tests

VARIABLE	(1) OLS Diagnostic gene test (difference)	(2) Poisson Diagnostic gene test (difference)	(3) OLS Diagnostic gene test (ratio)	(4) Poisson Diagnostic gene test (ratio)
Treated	4.287**	0.211***	0.375	0.110***
	(1.830)	(0.0111)	(0.308)	(0.0263)
Articles (log)	1.115*	0.0645***	0.337***	0.0973***
	(0.672)	(0.00405)	(0.113)	(0.00963)
Number of Diseases	1.126	0.0826***	0.0242	0.0138
	(1.000)	(0.00606)	(0.168)	(0.0139)
Important Disease	10.97***	0.488***	0.845**	0.211***
	(2.022)	(0.0119)	(0.340)	(0.0275)
Patent Age	-0.591**	-0.0361***	-0.101**	-0.0302***
	(0.238)	(0.00154)	(0.0400)	(0.00357)
Patent Age2	0.000292**	1.79e-05***	5.06e-05**	1.51e-05***
	(0.000118)	(7.64e-07)	(1.98e-05)	(1.77e-06)
Chromosomal Location	Yes	Yes	Yes	Yes
Observations	2,496	2,496	2,496	2,496
R-squared	0.128		0.135	

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

chances for novel innovations to emerge. This might be simply due to lack of incentives for innovators or due to failures in market for ideas. An avenue for future research, therefore, can be to study the mechanisms that lead to these impacts.

Table 3: Development of new diagnostic tests

	(1) OLS	(2) Logit
VARIABLE	NEW Diagnostic gene test	NEW Diagnostic gene test
Treated	-0.0327***	-0.170***
	(0.00891)	(0.0476)
Articles (log)	0.0530***	0.252***
	(0.00324)	(0.0173)
Number of Diseases	0.0409***	0.226***
	(0.00333)	(0.0185)
Important Disease	-0.0217**	-0.123**
	(0.00916)	(0.0494)
Patent Age	0.00433***	0.0249***
	(0.00111)	(0.00595)
Patent Age2	-2.17e-06***	-1.25e-05***
	(5.50e-07)	(2.94e-06)
Chromosomal Location	Yes	Yes
Observations	13,754	13,546
R-squared	0.134	

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

References

[1] Arora, A., Fosfuri, A., & Gambardella, A. (2001). Markets for technology and their implications for corporate strategy. *Industrial and corporate change*, 10(2), 419-451.

[2] Cole, J. R. (2000). A short history of the use of citations as a measure of the impact of scientific and scholarly work. The Web of Knowledge: A Festschri in Honor of Eugene Garfield, 281-300.

[3] Furman, J. L., & Stern, S. (2011). Climbing atop the shoulders of giants: The impact of institutions on cumulative research. *The American Economic Review*, 101(5), 1933-1963.

[4] Galasso, A., & Schankerman, M. (2014). Patents and cumulative innovation: Causal evidence from the courts (No. w20269). National Bureau of Economic Research.

[5] Gans, J. S., & Stern, S. (2003). The product market and the market for ideas: commercialization strategies for technology entrepreneurs. *Research policy*, 32(2), 333-350. Chicago

[6] Heller, M. A., & Eisenberg, R. S. (1998). Can patents deter innovation? The anticommons in biomedical research. *Science*, 280(5364), 698-701

[7] Jensen, K., & Murray, F. (2005). Intellectual property landscape of the human genome. *Science*, 310(5746), 239-240.

[8] Jost, K. (2013). Patenting human genes. *CQ Researcher*, 23, 473-496

[9] Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., ... & Funke, R. (2001). Initial sequencing and analysis of the human genome. *Nature*, 409(6822), 860-921.

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

[11] Murray, F., & O'Mahony, S. (2007). Exploring the foundations of cumulative innovation: Implications for organization science. *Organization Science*, 18(6), 1006-1021.

[12] Murray, F., Aghion, P., Dewatripont, M., Kolev, J., & Stern, S. (2009). Of mice and academics: Examining the effect of openness on innovation (No. w14819). National Bureau of Economic Research.

[13] Patent author, *Patent Office*, 2013, University Press

[14] Sampat, B., & Williams, H. L. (2015). How do patents affect follow-on innovation? Evidence from the human genome (No. w21666). National Bureau of Economic Research.

[15] Scotchmer, S. (1991). Standing on the shoulders of giants: cumulative research and the patent law. *The Journal of Economic Perspectives*, 5(1), 29-41.

[16] Teece, D. J. (1986). Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research policy*, 15(6), 285-305.

[17] Walsh, J. P., Arora, A., & Cohen, W. M. (2003). Effects of research tool patents and licensing on biomedical innovation. *Patents in the Knowledge-based Economy*, 285, 286.

[18] Watzinger, M., Fackler, T., Nagler, M., & Schnitzer, M. (2016). Antitrust, Patents, and Cumulative Innovation: Evidence from Bell Labs.

[19] Williams, H. L. (2010). Intellectual property rights and innovation: Evidence from the human genome (No. w16213). National Bureau of Economic Research.

[20] Zheng Zhang, Scott Schwartz, Lukas Wagner, & Webb Miller (2000), "A greedy algorithm for aligning DNA sequences", *J. Comput. Biol.* 2000; 7(1-2):203-14.