Julian Boulanger

THE IMPACT OF THE PATENT SYSTEM ON INNOVATION

STOCKHOLM SCHOOL OF ECONOMICS

THE IMPACT OF THE PATENT SYSTEM ON INNOVATION

This Ph.D. thesis contains the following three independent chapters.

"Patents and Follow-On Innovation: Evidence From Patent Renewal Decisions" examines whether patent protection on existing technologies blocks or facilitates technical advances that spur from those existing technologies.

"The Examination of Continuation Applications and the Problem of Invalid Patents in the U.S." explores how features of the examination process affects examiners' behaviour and, in particular their granting decisions, in the context of continuation applications in the U.S.

"Patent Length and Innovation Incentives for Industrial Designs" studies the reaction of inventors to an increase in the term of protection of design patents in the U.S.

JULIAN BOULANGER holds a Bachelor's degree in Economics and Management from Université Saint-Louis, a joint M.Sc. in Economics from Bocconi University and Université Catholique de Louvain and an M.Sc. in Economic Theory and Econometrics from the Toulouse School of Economics. His main research fields are the economics of intellectual property rights and innovation.

ISBN 978-91-7731-137-9

DOCTORAL DISSERTATION IN ECONOMICS
STOCKHOLM SCHOOL OF ECONOMICS, SWEDEN 2019

The Impact of the Patent System on Innovation

Julian Boulanger

Akademisk avhandling

som för avläggande av ekonomie doktorsexamen
vid Handelshögskolan i Stockholm
framläggs för offentlig granskning
måndagen den 17 juni 2019, kl 13.15,
rum 550, Handelshögskolan,
Sveavägen 65, Stockholm

The Impact of the Patent System on Innovation

The Impact of the Patent System on Innovation

Julian Boulanger

Dissertation for the Degree of Doctor of Philosophy, Ph.D., in Economics
Stockholm School of Economics, 2019

The Impact of the Patent System on Innovation

ISBN 978-91-7731-137-9 (printed)
ISBN 978-91-7731-138-6 (pdf)

This book was typeset by the author using LaTeX.

Front cover photo:
© Julian Boulanger

Back cover photo:
© Zabrina E. Kunkel

Printed by:
BrandFactory, Gothenburg, 2019

Keywords:
Follow-on innovation, patent renewal, patent examination, patent office, invalid patent, continuation application, examiner behaviour, innovation incentives, patent term, patent length, design patent.

Foreword

This volume is the result of a research project carried out at the Department of Economics at the Stockholm School of Economics (SSE).

This volume is submitted as a doctor's thesis at SSE. In keeping with the policies of SSE, the author has been entirely free to conduct and present his research in the manner of his choosing as an expression of his own ideas.

SSE is grateful for the financial support provided by the Jan Wallander and Tom Hedelius Foundation which has made it possible to fulfill the project.

Göran Lindqvist
Director of Research
Stockholm School of Economics

Tore Ellingsen
Professor and Head of the
Department of Economics
Stockholm School of Economics

Acknowledgements

I thank my supervisor, Tore Ellingsen, for his support and guidance throughout my Ph.D. studies. Early on, Tore was available to discuss ideas and gave me directions for my work. His broad knowledge and sharp mind have helped me develop new insights into my research and make progress in my dissertation. Over the years, Tore gave me ample freedom to develop my own research, allowing me to foster my independent thinking. Learning how to conduct research independently was the main reason I entered the Ph.D. programme at the Stockholm School of Economics and I am very grateful to Tore for trusting me and providing me with resources to achieve that goal.

I also thank Jörgen Weibull, Mark Voorneveld and Richard Friberg. Jörgen was always generous with his time and I benefited a lot from our discussions and the papers that he sent me for further reading. Mark, on top of being an exemplary instructor, has been a great collaborator. I enjoyed tutoring on two separate occasions for his Microeconomics Ph.D. course and learned a lot from this experience from a pedagogical point of view. Finally, I am grateful to Richard who came forth following my seminar presentations with helpful comments and suggestions on how to improve my papers and presentations.

I am grateful to the Jan Wallander and Tom Hedelius Foundation for financial support. In particular, thanks to a travel scholarship by this Foundation, I was able to spend the first semester of my fourth year at ECARES at the Université libre de Bruxelles and the subsequent one at the Department of Economics at the University of California, Berkeley. I thank Bruno van Pottelsberghe for his invitation at ECARES and for the many stimulating discussions that we had during my stay. I thank Bronwyn H. Hall and Brian Wright at UC Berkeley for their availability, encouragements and valuable feedback on my work.

Finally, I thank my fellow Ph.D. and post-doctoral students at SSE, with whom over the years I have enjoyed nice lunches, after works, poorly refereed football games, sosta coffee breaks and animated dinners. I also thank the administrative staff at SSE for taking care of our scholarship payments, course registrations and many other important tasks. Thanks to my friends and family back in Belgium, home has remained a special place and I am really looking forward to spending the next years by their side. Last but not least, I thank Zabrina for travelling halfway across the world to come live with me in Stockholm, for her kindness, love and invaluable support, and for helping me keep a sense of proportion.

Stockholm, April 26, 2019

Julian Boulanger

Contents

Introduction

This Ph.D. thesis is a collection of three self-contained chapters. These three chapters share a common theme, the empirical analysis of the patent system, and they address a fundamental question in the economics of innovation: How does the patent system affect innovation?

Each chapter focuses on a unique aspect of this complex question. Chapter 1 explores empirically how patent protection on existing technologies affects "follow-on" innovation, that is, innovation that spurs from existing technologies or research. Chapter 2 offers an empirical analysis of the relationship between features of the patent examination process and the granting of invalid patents, which create a climate of uncertainty for innovators and may impede future innovation. Finally, Chapter 3 examines empirically how inventors respond to an increase in the term of patent protection for industrial designs in the U.S.

The three chapters are now briefly summarised. More detailed overviews can be found in the chapters' introductions.

Patents and Follow-On Innovation: Evidence From Patent Renewal Decisions

Patents motivate innovation by rewarding inventors with exclusivity rights over their inventions. But scholars have argued that this reward theory of patents is not sufficient to justify the patent system. When innovation is cumulative — that is, when existing innovations are used as inputs to the development of future innovations — patents can either promote or block follow-on innovation and knowing the conditions under which these effects prevail is key to understanding whether patents are, overall, desirable or not.

In this chapter, I study how patent protection on existing technologies affect follow-on innovation using data on patent renewal decisions in the U.S. To address the endogeneity of renewal decisions, I construct an instrument using the large increases in maintenance fees that took place in 2013 following the enactment of the America Invents Act.

My findings show that, on average, patents have a strong blocking effect on follow-on innovation. But this average effect masks considerable heterogeneity. First, I explore variation across the three stages at which patents must be renewed over their lifetime and find that the blocking effect is entirely driven by patents in the first and second stages. At the third maintenance stage, patents appear, instead, to promote follow-on innovation. Second, I assess potential heteregeneity across different technology areas and find that patents block follow-on innovation more in discrete technology areas and when patent ownership is highly fragmented.

These findings suggest that patents harm follow-on innovation only in specific contexts. Targeted policies aimed at facilitating the licensing of certain patents, such as those at early stages of their lifespans, may thus be more suitable than policies which aim at removing patent rights all together.

The Examination of Continuation Applications and the Problem of Invalid Patents in the U.S.

It has long been recognised that the U.S. patent office routinely grants a large number of invalid patents. Scholars have recently started shedding light on the root causes of this issue and have emphasised the key role played by the patent examination process itself.

In this chapter, I study how features of the examination process affect examiner behaviour and the problem of invalid patents in the context of continuation applications. These applications emanate from earlier patent applications filed at the patent office and allow applicants to submit new claims for a given invention. In the U.S., continuations are generally examined by the same examiner who was assigned to the earlier application. Using application-level data, I explore how this feature, which I call "relatedness," affects examiners' grant decisions and other examination practices.

I find that relatedness increases the grant rate, decreases examiners' efforts to narrow down the scope of protection claimed by patent applicants and decreases the examiners' search efforts for "prior art," which constitutes any evidence that the claimed invention was already known to the public. I also show that the effects of relatedness do not seem to be driven by a "wearing down" effect on examiners and differ only slightly across different technology areas. Finally, I find evidence that relatedness leads to the granting of patents of more dubious validity.

A key implication of these results is that the way the U.S. continuation application system is designed causes examiners to adopt softer examination practices, which in turn contributes to the problem of invalid patents.

Patent Length and Innovation Incentives for Industrial Designs

An optimal patent system balances the costs and benefits of patents at the margin. A key parameter needed to achieve optimality is the sensitivity of innovation to the length of patent protection.

In this paper, I study the impact of patent length on innovation incentives in the context of design patents. I exploit the exogenous increase in the term of design patents, from fourteen to fifteen years, that came into effect in 2015 when the U.S. implemented the Hague Agreement Concerning the International Registration of Industrial Designs.

The term extension increased patent filings by 1.5-8%, though not in a statistically significant way. In addition, there is little evidence of variation in the term extension's effect across different industries or types of patent applicants.

These findings add to the existing body of empirical studies that document inventors' low sensitivity to the strengthening of patent protection.

Chapter 1

Patents and Follow-On Innovation: Evidence From Patent Renewal Decisions[1]

[1] I thank Markus Nagler, Bronwyn H. Hall, Tore Ellingsen, Richard Friberg and Brian Wright for very helpful comments and feedback. I also thank participants at the Stockholm School of Economics lunch seminar and at the Norwegian School of Economics course on Productivity for useful remarks. Financial support from the Jan Wallander and Tom Hedelius Foundation is gratefully acknowledged. All errors are my own.

1.1 Introduction

A common justification for patents is that they reward inventors and, thereby, boost innovative activity. But researchers have questioned this view, noting that because innovation is a cumulative process, with new technologies often building on earlier ones, the incentive effects of patent protection on follow-on innovation must also be factored in.[2] And while some scholars argue that patents can promote follow-on innovation by solving coordination issues among downstream innovators and facilitating the commercialisation of ideas (Spulber, 2015; Kitch, 1977), others contend that patents may block follow-on innovation if bargaining failures prevent the efficient licensing of patented technologies between upstream and downstream innovators (Bessen and Maskin, 2009; Bessen, 2004). Hence, theory alone cannot rule out that, overall, patents deter rather than motivate innovation.

A growing body of research investigates empirically the effects of patent protection on follow-on innovation. The major challenge for empirical studies is to identify credible sources of variation in patent protection. First, patent laws have been in force for decades in most countries, leaving little variation in patent protection at the aggregate level that can be exploited by researchers. Second, simple comparisons of patented and unpatented technologies are hard to make as unpatented technologies are typically not observable, one of the main alternatives to patenting being secrecy (Hall et al., 2014). In addition, assuming such comparisons could be made, they would likely be misleading due to selection into patenting (Sampat and Williams, 2019). Existing studies have thus focused on specific technology areas or patents in the tail of the value distribution for which clear natural experiments exist. As a result, while we now know a lot about the effect of patents on follow-on innovation in those specific contexts, still very little is known about this effect for more representative patents.

In this paper, I investigate how patent protection on existing technologies affects follow-on innovation, as measured by citations made by later patents, using U.S. data on patent renewal decisions. In the U.S., patents have to be renewed three times to make it to the full legal term of twenty years: at years

[2]This paper follows the literature on cumulative innovation and focuses on follow-on innovation by other inventors than the inventor of the earlier innovation.

3.5, 7.5 and 11.5 counted from the date of issue. As maintenance fees must be paid at each renewal stage, some patent holders choose to let their patents lapse, in which case the underlying technology enters the public domain and can be used without needing a license at first. Renewal decisions thus induce variation in patent protection that can potentially be used to estimate the effects of patents on follow-on innovation. Importantly, because every patent must be renewed, the renewal decision allows me to discuss these effects for patents across all major technology areas and whose values are not limited to the tail of the value distribution.

The main concern with this empirical approach is that renewal decisions are endogenous. Patent holders are more inclined to renew valuable patents, which are more likely to generate more follow-on innovation. Due to the influence of unobservable characteristics, such as patent value, a simple comparison of renewed and lapsed patents would certainly not give a credible estimate of the true causal effect of patent protection on follow-on innovation.

To address this endogeneity issue, I construct an instrument for renewal decisions using the maintenance fee increases that took place in 2013 when, following the enactement of the America Invent Act in 2011, the U.S. Patent and Trademark Office (PTO) gained fee-setting authority. Patent holders that had to renew their patents after March 19, 2013, faced maintenance fees that were, on average, 39% higher across the three maintenance stages. The basic idea behind my identification strategy is that, while higher fees should lower renewal rates, there is no apparent connection between whether a patent had to be renewed before or after the fee increases and its value or potential to generate more or less follow-on innovation. The fee increases can then be used to generate quasi-experimental variation in renewal decisions.

To execute this empirical strategy, I first rely on the fact that a patent's maintenance fee payments are scheduled on the basis of its date of issue to separate patents into two groups, a "low-fee" group and a "high-fee" group. I then use a two-step Instrumental Variables (IV) method in which I first estimate how the belonging to the low-fee and high-fee groups affects renewal decisions using a Probit model and then use the predicted probabilities from that model to instrument for the renewal decisions. This estimation method, which is discussed in more detail in Wooldridge (2010) in Section 21.4, has been commonly used in the literature and has good efficiency properties.

In addition to estimating the average effect of patents on follow-on innovation, this paper explores the heterogeneity of this effect along some key dimensions. First, it assesses potential variation across the different maintenance stages. Although patentees tend to be initally uncertain about the potential returns to their patents (possibly in the form of licensing agreements with downstream innovators), they quickly learn over time about these returns and can then form a better judgement of the value of their patents (Lanjouw et al., 1998). As a result, upstream and downstream innovators' ability to negotiate mutually beneficial licensing agreements improves over time and the blocking effect of patents that might initially prevail due to bargaining failures should dissipate progressively. Second, it studies how the effect of patents on follow-on innovation varies across different types of technologies and focuses, in particular, on the roles played by technological complexity and the fragmentation of patent ownership, which have received a lot of attention in the literature but for which mixed evidence is currently available.

My findings show that, on average, patents have a strong blocking effect on follow-on innovation. But this average effect masks considerable variation across the three maintenance stages and is entirely driven by patents in the first and second stages. I find that the lapse of a patent at the first maintenance stage results in about 35% more citations by later patents. This figure drops to about 15% for patents at the second maintenance stage. For patents at the third maintenance stage, the lapse of patent protection results in about 20% fewer citations by later patents. Thus, patents appear to block follow-on innovation at early stages, but the blocking effect dissipates and gives place to a facilitating effect at later stages. This result is consistent with the learning mechanism discussed earlier and also supports the presence of coordination failures between downstream innovators, as I later clarify. Finally, I find that patents block follow-on innovation more in discrete technology areas and when patent ownership is highly fragmented.

This paper is mainly related to the growing empirical literature that investigates how patents affect follow-on innovation. Most closely related to this paper are the studies by Galasso and Schankerman (2015) and Gaessler et al. (2017), which both use patent invalidations as a source of variation in patent protection. Galasso and Schankerman (2015) focus on patent invalidation cases brought in front of the U.S. Court of Appeals for the Federal Circuit and exploit the random allocation of judges, along with their varying propensities to invalidate patents, to instrument for the invalidation decisions. Gaessler et al.

(2017) focus on post-grant oppositions at the European Patent Office and use the participation of an opposed patent's examiner in the opposition division as an instrument for patent invalidation. Both papers find that, on average, patents have a strong and significant blocking effect on follow-on innovation. However, whereas Galasso and Schankerman (2015) find that the effect is more pronounced in complex technology fields and when patent rights are highly fragmented, Gaessler et al. (2017) find that the effect is driven primarily by patents in discrete technologies and when patent fences and thickets are absent. As these papers focus on litigation and opposition cases, they are naturally based on selected samples of highly valuable patents. My paper complement these studies by providing evidence on the effect of patents on follow-on innovation for patents of more representative values. In addition, my paper shows that neither study's results on the drivers of the effect of patents can be fully replicated.

Another strand of the literature uses patent grants as a source of variation in patent protection. Sampat and Williams (2019) consider human genes patenting and show that patents do not seem to affect follow-on innovation in that particular market. An earlier study by Murray and Stern (2007) uses patent-paper pairs in human genetics and shows that the grant of a patent decreases the number of citations made to the paired scientific paper. Scholars have also considered exogenous events of compulsory licensing as a source of variation in patent protection. With compulsory licensing of a given upstream technology, downstream innovators can use the technology without needing a license from the owner of the upstream technology. Moser and Voena (2012) study the effect of an exogenous episode of compulsory licensing induced by World War I on domestic innovation in the U.S. in the chemical technology area and find that compulsory licensing increased domestic innovation in that area by 20%. In a similar study, Watzinger et al. (2017) study the effect of the compulsory licensing of Bell's existing patents in 1956 and show that it had a significant effect on follow-on innovation, but only in technology areas where Bell could not foreclose competition.

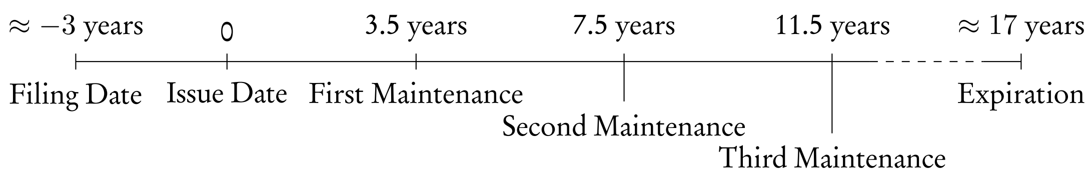
This paper is structured as follows. In Section 1.2, I explain how patents are renewed in the U.S., how follow-on innovation can be measured and what existing theories tell us about the potential effects of patents on follow-on innovation. Section 1.3 describes the data and reports summary statistics. In Section 1.4, I introduce my empirical model and, in Section 1.5, I present its results. Section 1.6 concludes the paper.

1.2 Empirical Setting

1.2.1 Patent Renewal

The current U.S. renewal system was introduced in the early 1980s and has since then consisted in three maintenance stages at which patent holders must pay fees. Figure 1.1 illustrates the renewal process. The first, second and third maintenance fees are due, respectively, 3.5, 7.5 and 11.5 years after the patent's date of issue. The PTO allows patentees to pay during a one-year window around the due date. For instance, U.S. patent 6,161,220, which issued on December 19, 2000, had its third maintenance fee due on June 19, 2012, but could be paid during the period that started on December 19, 2011, and ended on December 19, 2012.[3] Only patents that have been renewed three times stay in force until the full legal term, which equals 20 years from the patent's date of filing. For a typical patent application, the patent issues about three years after being filed, hence patent protection typically expires about 17 years as counted from the date of issue if all three maintenance fees have been paid.[4]

Figure 1.1: The Patent Renewal Process



Notes: This figure depicts the patent renewal process for a typical patent.

The renewal decision by patentees has been the focus of an extensive literature, which poses the problem as a comparison between the costs of re-

[3]The second half of the payment window is known as the "grace period." During this period, the payment of additional surcharge fees by the patentee are required to maintain the patent in force.

[4]The PTO applies patent term adjustments for any delays in the examination process that it is responsible for. As a result, individual patents have different lengths.

newal (i.e. the maintenance fees) with the current and potential future returns from patent protection. In its simplest form, the model of patent renewal assumes that the costs and benefits are known to patentees and are deterministic (Pakes and Schankerman, 1984). In a more intricate "option value" version of the model, patentees are allowed to learn about their patents' stochastic returns over time (Pakes, 1986). The learning process shapes renewal decisions, as patentees may decide to hold on to patents with low current value if they believe these patents will yield greater returns in the future. A consistent finding about the learning process is that most learning occurs early in a patent's life, within the first five to seven years (Lanjouw et al., 1998). In other words, by the time a patent reaches the second maintenance stage, its owner should be quite confident as to whether it will generate low or high future returns. This learning, together with the fact that maintenance fees increase from one stage to the next, can be used to explain the observed decrease in the fraction of patents renewed across the various maintenance stages. In recent years, respectively 85%, 70% and 50% of all first maintenance stage, second maintenance stage and third maintenance stage patents were renewed.

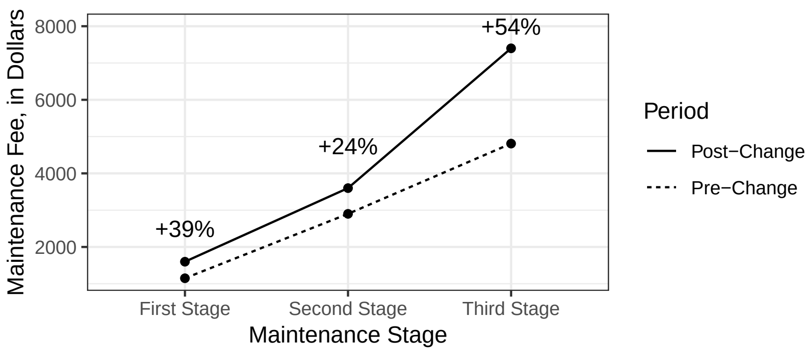
The U.S. has kept its maintenance fees relatively low over the years. From the early nineties, maintenance fees were actually adjusted annually only to reflect changes in the consumer price index. But the America Invents Act, a substantial patent reform signed into law in 2011, gave the PTO temporary fee-setting authority, which later resulted in a number of important adjustements made to patent fees.[5] Maintenance fees were increased significantly as a result of these changes, which came into force on March 19, 2013. Figure 1.2 shows the fee schedules before and after March 19, 2013. The first maintenance fee increased by 39%, from $1,150 to $1,600. The second fee increased by 24%, from $2,900 to $3,600. Finally, the third fee increased by 54%, from $4,810 to $7,400.

1.2.2 Follow-On Innovation

In theory, the idea of follow-on innovation is clear and is typically understood as one or several new technologies that build upon an existing technology. Absent the earlier technology, later technological advances would not have taken place. But measuring follow-on innovation has proved to be challenging in

[5]The fee-setting authority was recently renewed and the PTO will retain it until 2026.

Figure 1.2: Maintenance Fees Before and After the 2013 Fee Increases



Notes: This plot shows the maintenance fee schedules applicable before and after the fee increases that came into force on March 19, 2013. For detailed information on the 2013 fees increases, see United States Patent and Trademark Office (2013).

practice. The reason is that there is no ultimate way to trace back all technological advances to the innovations that they spurred from.

While a few measures have been proposed in the literature, as argued by Galasso and Schankerman (2015) the only one that can be used in empirical studies that cover all major technology areas are citations made by later patents, commonly called "forward citations." There are more direct measures, like subsequent product developments, but they are only available for specific markets. Sampat and Williams (2019), for instance, measure follow-on innovation in the context of human genes by linking individual genes to the number of related pharmaceutical clinical trials and diagnostic tests. But even if we could measure follow-on innovation across a large number of markets using product developments, it is unclear how we would interpret empirical results based on this measure, as there would not be a single unit for comparison across the various markets. Citations, on the other hand, are both readily interpretable and available across a wide range of technologies.

Forward citations are by no means a perfect measure of follow-on innovation. As Galasso and Schankerman (2015) explain, they can both overestimate and underestimate the amount of follow-on innovation. They will underestimate it when subsequent developments are not patented (for instance, when they are kept secret instead) and they will overestimate it when the earlier patented innovation did not actually spur the subsequent patented development but must still be cited as it is part of the "prior art."[6]

Nonetheless, I follow the common practice found in previous studies that cover patents in a broad range of different technology fields and measure follow-on innovation using forward citations. Importantly, when a patent lapses, although it enters the public domain, the obligation to cite it in future patents that rely on the underlying technology remains in place. This ensures that I can equally measure follow-on innovation for both renewed and lapsed patents.

1.2.3 Theoretical Framework

The interest of economists in the effect of patents on follow-on innovation is not new. In an early contribution, Kitch (1977) noted that patents on upstream technologies may prevent inefficient races between downstream innovators, resulting in increased social welfare. However, scholars have also argued that bargaining failures imply that patents on existing technologies may prevent the development of follow-on technologies if licensing agreements cannot be reached. Building on the different ideas proposed in the literature, Galasso and Schankerman (2015) develop a unified framework to study the effect of patents on follow-on innovation. The key mechanism in their model is a trade-off between coordination failures between downstream innovators and bargaining failures between upstream and downstream innovators that are caused by the presence of asymmetric information.

Their model assumes that there is one upstream technology from which one downstream technology can be developed by two identical potential downstream innovators. The downstream technology's value can be either high or low, but its cost of development is fixed. Asymmetric information is introduced in the model by assuming that downstream innovators know the value of the potential follow-on technology, while the upstream innovator only knows that it is of high value with some positive probability. Asymmetric information drives bargaining failures as it prevents the upstream and downstream innovators to agree on mutually beneficial licensing terms. Coordination issues are introduced in the model by assuming that the development of the follow-on innovation is profitable only when a single downstream innovator develops

[6]The prior art constitutes the public information that cannot be claimed in a given patent as it existed before the patent application was filed. It generally consists of prior patents and scientific publications.

it.

The main result of the model is that patents on existing upstream technologies can either block or promote follow-on innovation depending on the relative intensities of coordination and bargaining failures. Holding coordination failures constant, when information is strongly asymmetric, patents block follow-on innovation, whereas when it is weakly asymmetric, patents promote follow-on innovation. The intuition behind this result is the following. Patent protection on the upstream innovation allows the patentee to exclude one of the two downstream innovators, solving coordination issues in the downstream market. But a patent will only increase follow-on innovation if the upstream and downstream innovators ultimately reach a licensing agreement. This is more likely to be the case when uncertainty on the follow-on innovation's value is low, as then the upstream innovator is more informed on this value and better able to offer licensing terms that the downstream innovator will accept.

An implication of this result is that, other things constant, higher uncertainty results in a stronger blocking effect. When patentees learn about their patents' returns, effectively they learn about the value of potential follow-on developments of their technologies. This learning lowers uncertainty and, in turn, the blocking effect of patent protection. The conclusion that emerges is that patent protection should have a stronger blocking effect for patents in the early maintenance stages. As noted previously, most learning occurs within the first seven years of a patent's life and, thus, we expect patents to have a noticeable effect on follow-on innovation mainly at the first two maintenance stages.

1.3 Data

This paper uses two types of data: (1) data on maintenance fee events, collected from Reed Tech, and (2) data on a range of patent and owner characteristics, collected from PatentsView.[7] Every time a patent is renewed or is allowed to lapse by its owner, this information is recorded internally by the

[7]The Reed Tech data is available at https://patents.reedtech.com/maintfee.php. This paper makes use of the May 28, 2018, release of the PatentsView data. Bulk downloads can be made at http://www.patentsview.org/download/.

PTO as a maintenance fee event. The data made available by Reed Tech contains all recorded events for patents granted since September 1, 1981, and is updated weekly. I match this data with the information on patent and owner characteristics collected from PatentsView using the unique patent number assigned by the PTO when the patent issues. PatentsView allows researchers to download bulk data organised in separate "themes." One can then easily access information on granted patent applications, the number of citations a given patent made or received, information on assignees and inventors, among other things.[8]

My empirical strategy, detailed in Section 1.4, uses the large increases in maintenance fees that came into effect on March 19, 2013, to generate quasi-experimental variation in renewal decisions. For this strategy to work, two conditions must be met and this places certain restrictions on the sample. First, patent holders should not be able to manipulate the applicable maintenance fees. Because the PTO allows patentees to pay their maintenance fees over a one-year window around the due date, any patent whose payment window includes March 19, 2013, must be excluded.[9] My sample thus contains only two types of patents: (1) patents whose latest possible dates of maintenance fee payment were before March 19, 2013 and (2) patents whose earliest possible dates of maintenance fee payment were on or after March 19, 2013.

Second, the impact that the increase in maintenance fees has had on renewal decisions should be estimated as precisely as possible. In particular, we must capture the effect of the shock itself rather than the influence of alternative factors. We can avoid potential confounders by restricting the sample over a relatively narrow window around March 19, 2013. I restrict my sample to all patents whose latest dates of payment fall within a thirteen weeks period prior to March 19, 2013 and to all patents whose earliest dates of payment fall within a thirteen weeks period on or after March 19, 2013. This particular choice of

[8]This paper makes use of the following files: application.tsv, assignee.tsv, foreign_priority.tsv, ipcr.tsv, nber.tsv, patent.tsv, patent_assignee.tsv, uspatentcitation.tsv and uspc.tsv. All these files can be linked together through an internal identification system.

[9]Suppose we did not exclude patents whose payment window includes March 19, 2013. Then our sample would likely include many cases in which the patent holder paid the maintenance fees early in order to avoid paying the higher fees. This would introduce a sorting issue and a bias in the estimates.

window being somewhat arbitrary, I verify in a robustness check in Section 1.5 that the main findings are unaltered by the choice of a larger window of twenty-six weeks.

My main dependent variable, *PostCites*, is the number of forward citations made by patents that have different assignees (i.e. owners) than the focal patent in the three years that followed the maintenance event.[10] My choice of a three-year window is motivated by the time intervals between the various maintenance events. It is shorter than in previous studies, which have typically relied on five-year windows, because a longer window than three years would risk confounding the effects of a given maintenance fee event by a subsequent one.[11]

The construction of my outcome variable raised a couple of issues. First, a small number of citing and cited patents could not be matched to any assignee. As the distinction between citations made by other assignees and the same assignees as focal patents is needed to construct my outcome variable, citations for which assignee information was missing were discarded.[12] Second, the citations made by a given patent to earlier patents are only revealed once that patent has been granted. Due to the significant time gap between the time of filing and of grant (three years on average), many citations are not recorded for several years until the citing patents are granted, leading to truncation of the data. Truncation may be an issue in my sample because forward citations are counted over the three-year period that follows maintenance fee events, some of which occur as late as June 2014. To correct for truncation, I adjust forward citations using the fixed-effects method proposed by Hall et al. (2001). This approach works as follows: for each patent, divide its number of forward citations by the average number of forward citations received by all patents granted in a given pre-defined cohort to which it belongs (for instance, technology areas). I apply a conservative adjustment by defining a cohort as a given year of

[10]For many of my variables, I follow the notation in Galasso and Schankerman (2015) as my empirical model is very close in spirit to theirs.

[11]Exactly four years separate any two maintenance fee due dates. However, since patent holders can pay maintenance fees up to six months in advance or with delay, the shortest possible time between any two payments will be three years.

[12]About 2% of all citing patents could not be matched to an assignee and about 4% of cited patents could not be matched to an assignee.

issue and one of the six NBER technological categories defined in Hall et al. (2001).[13] This corrects for truncation by removing all year, technology field and year-field effects and improves comparability of forward citations across patents.

My main independent variable of interest is $Lapsed$, an indicator variable that takes the value 1 if a patent was allowed to lapse by its owner and 0 if it was renewed. A number of additional independent variables play an important role in my empirical model. $PreCites$ is the number of forward citations made by patents that have different assignees than the focal patent, counted from the date of issue of the focal patent and until the maintenance fee event. $PreSelfCites$ is the number of forward citations made by patents that have the same assignee as the focal patent, counted from the date of issue of the focal patent and until the maintenance fee event. $Claims$ is the total number of claims in the given patent. Finally, I always include technology effects based on the 37 NBER subcategories defined in Hall et al. (2001), as well as maintenance effects indicating whether a patent had to be renewed for the first, second or third time.

My sample contains 221,077 patents, each of which is associated with a single maintenance fee event.[14] Among these, 91,677 (41%) had a first maintenance fee to be paid, 69,888 (32%) had a second maintenance fee to be paid and the remaining 59,512 (27%) had a third maintenance fee to be paid. The dates of recorded maintenance events range between December 21, 2011 and June 11, 2014. Sample patents were issued between 2000 and 2002 (third maintenance stage patents), 2004 and 2006 (second maintenance stage patents), and 2008 and 2010 (first maintenance stage patents). The average patent in the sample is 10 years old as counted from its filing date. Of the 221,077 patents, 30% are in the Computers and Communications category (NBER category 2), 24% are in the Electrical and Electronic category (NBER category 4), 14% are in the Mechanical category (NBER category 5), 12% are in the Other category (NBER

[13]The NBER categories are: (1) Chemical (NBER category 1); (2) Computers and Communications (NBER category 2); (3) Drugs and Medical (NBER category 3); (4) Electrical and Electronics (NBER category 4); (5) Mechanical (NBER category 5) and (6) Other (NBER category 6).

[14]Some additional cleaning was necessary before the sample was ready to be used for my empirical analysis. Details are given in Appendix 1.B.

category 6), 11% are in the Chemical category (NBER category 1) and, finally, 9% are in the Drugs and Medical category (NBER category 3).

Table 1.1 reports summary statistics for the key variables used in the empirical analysis. About one patent in five in the sample lapsed, but there is variation in renewal rates across the three maintenance stages. Among patents that had a first maintenance fee to be paid, only 13% lapsed. This figure jumps to 21% for patents at the second maintenance stage and to 27% for patents at the third maintenance stage. A little more than half of the sample patents faced the higher maintenance fees. The average patent received about 2.5 forward citations by other assignees within three years after the maintenance event and about 9.5 forward citations made by other assignees prior to the maintenance event. The average patent in my sample contains about 18 claims.

Table 1.1: Summary Statistics

Variable	Mean	Std. dev.	Min	Max
PostCites	2.44	9.91	0	1316
Lapsed	0.20	0.40	0	1
PreCites	9.55	25.4	0	1386
PreSelfCites	1.46	7.53	0	1397
Claims	17.9	13.9	1	520

Notes: The sample contains 221,077 individual patents. PostCites is the number of forward citations made by different assignees than the given patent's assignee in the three years that followed the patent's maintenance event. Lapsed is a binary variable indicating if the given patent was allowed to lapse by the patentee or not. PreCites is the number of forward citations made by other assignees than the given patent's assignee before the maintenance event. PreSelfCites are the forward citations made by the same assignee as the given patent's assignee before the maintenance event. Claims is the number of claims in the given patent.

1.4 Empirical Model

To investigate how patent protection affects follow-on innovation, I estimate the following baseline specification:

$$\begin{aligned}\log(PostCites_i + 1) = {} & \beta Lapsed_i + \lambda_1 \log(PreCites_i + 1) \\ & + \lambda_2 \log(PreSelfCites_i + 1) + \lambda_3 \log(Claims_i) \\ & + Tech_i + Maintenance_i + x_i'\zeta + \epsilon_i, \end{aligned} \tag{1.1}$$

where the unit of observation is a given patent i. The dummy $Lapsed_i = \{0, 1\}$ indicates whether patent i was allowed to lapse by its owner. The outcome variable is defined as the log of one plus $PostCites$. This log transformation is commonly employed in the literature to tackle the high degree of skewness and the large number of zeros in patent citations data. Technology and maintenance stage effects are included in all specifications, as well as the log of one plus $PreCites$, the log of one plus $PreSelfCites$ and the log of $Claims$.[15] The additional controls that are included in x_i in separate specifications are described in Appendix 1.A. These controls include, for instance, dummies for filing years and entity size dummies.

The coefficient of interest, β, captures the effect of the lapse of patent protection on forward citations made by other inventors. A positive β would show that lapsed patents have received more forward citations than renewed patents, which would suggest that patents have a blocking effect on follow-on innovation. A negative β, to the contrary, would suggest that the lapse of patents decreases follow-on innovation. Finally, if β equals zero, we should conclude that patents do not seem to have any quantitatively important effects on follow-on innovation.

The main empirical challenge is that the renewal decision is endogenous. Patents that protect more valuable technologies are more likely to both be renewed and generate high levels of follow-on innovation. Indeed, the prospects of large future revenues, possibly in the form of licensing agreements with downstream innovators, is necessarily taken into account when the decision to renew is made. Without a variable that can accurately capture patent value, Ordinary Least Square (OLS) estimates of regression (1.1) may be biased. Al-

[15]The distribution of claims is also highly skewed in the data, which is why I also use the log transformation.

though some patent value indicators can be included as controls, it is unlikely that they would alone suffice to eliminate the bias.[16] We can form a guess on the sign of the bias using the omitted variable bias formula discussed in Angrist and Pischke (2008). If more valuable patents are more likely to be renewed, the correlation between $Lapsed$ and patent value is negative.[17] As a result, β would be downward-biased if more valuable patents also tend to generate more follow-on innovation.

To identify β, we need renewed and lapsed patents to be comparable. The basic idea behind my identification stragegy is to use the 2013 maintenance fee increases to generate quasi-experimental variation in $Lapsed$ and solve the endogeneity issue. Intuitively, while higher fees should lower renewal rates, there is no apparent connection between whether a patent had to be renewed before or after the fee increases and its value or potential to generate more or less follow-on innovation. I show in Section 1.5.1 that the fee increases did affect patentees' renewal decisions, so that this natural experiment generates meaningful variation in $Lapsed$ that can be exploited to estimate an unbiased β. But for this empirical stragegy to work, patents that had to be renewed before or after the fee increases must, in turn, be comparable.

My sample was restricted to and separated into two groups of patents — those whose latest possible dates of payment fall before March 19, 2013 (the low-fee group) and those whose earliest possible dates of payment falls after March 19, 2013 (the high-fee group) — and the belonging to either group is a function of the patent's date of issue only. The only way patentees could have sorted into either group is thus by strategically timing the issuance of their patents. However, since the youngest patents in my sample were issued in 2010, before the America Invents Act was even signed into law, there are no patents in my sample for which the owners could have reasonably timed the issue date on the basis of an expected fee increase.

The remaining threat to identification is the non-random assignment of patents to the low-fee and high-fee groups. This kind of sorting will bias my estimates if patents' issue dates are correlated with their potential to generate

[16]Patent valuation has proved to be a difficult problem and is itself the topic of a large literature in economics and management.

[17]Evidence supporting a negative correlation between patent value and the probability of lapse is discussed in Section 1.5.1.

more or less follow-on innovation. Patents are issued weekly, every Thursday, in the official journal of the PTO (the "Official Gazette") and the issuance of a patent is the culmination of a long process that begins with an application, continues with a substantive examination and, ultimately, results in an allowance by the examiner. Often, this process is disrupted by random events, such as the loss of documents, which affect the patent's issue date. Moreover, comparable patent applications can undergo very different examinations merely because they have been assigned to different examiners (Cockburn et al., 2002). Some examiners appear to be tougher than others, resulting in longer examinations and delayed issuance, and the assignment of applications to examiners follows a "first-in-first-out" principle based on filing dates and examiner availabilities (Sampat and Williams, 2019; Farre-Mensa et al., 2017). It should be clear from this discussion that there is no compelling reason to believe that there is a relationship between a patent's intrinsic potential to generate follow-on innovation and its issue date. As a result, the belonging to either the low- or high-fee group and a patent's ability to generate more or less follow-on innovation are, at least in principle, orthogonal.

To implement my empirical strategy, I use a two-step IV method with a Probit model in the first step and a two-stage least-square regression in the second step that uses the fitted probabilities from the Probit model to instrument for the endogenous variable. Wooldridge (2010) shows that this estimator, which is described in his Procedure 21.1 on page 939, has good efficiency and robustness properties. Moreover, both Galasso and Schankerman (2015) and Gaessler et al. (2017) use this procedure in their own empirical analyses of the effect of patents on follow-on innovation. In my context, the estimation procedure is the following:

Step 1: Estimate the following Probit model by maximum likelihood

$$Pr(Lapsed = 1|HighFees, X) = \Phi(\gamma HighFees + \Gamma X), \qquad (1.2)$$

where the dummy $HighFees = \{0, 1\}$ indicates whether the focal patent is in the high-fee group, Φ is the cumulative distribution of the standard normal distribution and X contains control variables. From this model, I obtain the fitted probabilities $\widehat{P}_i$.

Step 2: Estimate the following two-stage model

$$Lapsed_i = \alpha\widehat{P}_i + \theta X_i + u_i$$
$$\log(PostCites_i + 1) = \beta\widehat{Lapsed}_i + \eta X_i + \epsilon_i,$$

where X_i is the same set of controls in both stages.

1.5 Results

This section presents my empirical results. I first show how the 2013 increases in maintenance fees affected renewal decisions. I then report the estimated average effect of patents on follow-on innovation and, finally, I explore heterogeneity in this effect.

1.5.1 The Effect of the 2013 Maintenance Fee Increases on Renewal Decisions

In theory, higher maintenance fees should decrease the probability of renewal. The models of patent renewal behaviour discussed in Lanjouw et al. (1998) all incorporate this idea. In addition, existing empirical evidence consistently reveal a negative relationship between maintenance fees and renewal rates (de Rassenfosse and van Pottelsberghe de la Potterie, 2012). To investigate this relationship in my data, I estimate the Probit model (1.2) presented in Section 1.4. Given the theory and existing empirical evidence, the coefficient on $HighFees$, γ, should be positive.

Table 1.2 presents the results. In column (1), the Probit model is estimated without any other independent variables than $HighFees$. The estimated γ equals 0.064 and is highly significant. When the baseline set of independent variables is included, in column (2), the estimated coefficient for $HighFees$ equals 0.106 and is highly significant as well. This coefficient implies a marginal increase of about three percentage points in the rate at which patents are allowed to lapse due to the higher maintenance fees. The results from the specification in column (3), which includes additional patent and owner characteristics as controls, show that adding these additional controls has barely any impact on the estimated effect of $HighFees$, which now equals 0.118. A back-of-the-envelope calculation gives an estimated implied elasticity of patent renewals to maintenance fees of about -0.1, on par with estimates found in the literature and reported in de Rassenfosse and van Pottelsberghe de la Potterie (2012).[18] The other reported coefficients in Table 1.2 show that patents that are cited more and contain more claims are less likely to lapse. This suggests that more valuable patents are less likely to lapse, as we would expect.

Overall, these results show that the 2013 maintenance fee increases affected renewal decisions in a way that is consistent with both theory and evidence, and that they can be used to construct a meaningful instrument for $Lapsed$.

1.5.2 The Average Effect of Patents on Follow-On Innovation

This section presents the results of three separate regressions. The first and second are, respectively, an OLS regression and a two-step IV regression of the baseline model (1.1). The third is a two-step IV regression that includes additional controls as a robustness check.

Table 1.3 reports the results of these regressions. Column (1) shows that the OLS estimate of the effect of $Lapsed$ is significant and slightly negative. However, this estimate does not have a causal interpretation, as the renewal decision is endogenous. Column (2) reports the result of the estimation of the two-step IV method, in which $Lapsed$ is instrumented by the predicted probabilities obtained from the Probit model. The effect of $Lapsed$ is now positive and strongly significant. The difference between the IV and OLS estimates of β is consistent with the idea that more valuable patents are more likely to both be renewed and generate more follow-on innovation, which would explain the downward bias in the OLS estimate. Moreover, a formal endogeneity test strongly rejects the exogeneity of $Lapsed$.[19] Column (3) shows that including additional patent and owner characteristics in the regression has a very small impact on the IV estimates. Finally, as shown by the under-identification and weak identification tests reported in Table 1.3, the instrument does not seem to suffer from a weak instrument critique.

---

[18] The implied elasticity was calculated in the following way. First, I divided the three percentage points decrease in the renewal rate by the sample average renewal rate of 80%, giving an overall percentage decrease in the renewal rate of 3.75%. This figure was then divided by the average percentage increase in maintenance fees equal to 39%. The result of this calculation is the estimated implied elasticity, equal to -0.09.

[19] The endogeneity test is performed via Stata's `ivreg2` command with the `endog()` option and returns a p-value of ≈ 0.0000.

Table 1.2: The Effect of the 2013 Maintenance Fee Increases on Renewal Decisions

	(1)	(2)	(3)
Dependent Variable		Lapsed	
Estimation Method		Probit	
HighFees	0.064***	0.106***	0.118***
	(0.006)	(0.006)	(0.008)
log(PreCites+1)		-0.279***	-0.203***
		(0.009)	(0.009)
log(PreSelfCites+1)		-0.537***	-0.407***
		(0.025)	(0.025)
log(Claims)		-0.118***	-0.087***
		(0.004)	(0.005)
Technology and Maintenance Effects	No	Yes	Yes
Additional Controls	No	No	Yes
Observations	221077	221077	183699

Notes: Robust standard errors reported in parentheses. Lapsed is a binary variable indicating if the given patent was allowed to lapse by the patentee or not. HighFees equals 1 if the given patent was in the high-fee group and 0 if it was in the low-fee group. PreCites is the number of forward citations made by other assignees than the given patent's assignee before the maintenance event. PreSelfCites are the forward citations made by the same assignee as the given patent's assignee before the maintenance event. Claims is the number of claims in the given patent. Technology effects are the 37 technological subcategories defined in Hall et al. (2001). Maintenance effects are dummies indicating whether the given patent had to be renewed for the first, second or third time. For the list of variables included as additional controls, see Appendix 1.A. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

The estimated marginal effect of $Lapsed$, computed as $\exp(0.750)-1 = 1.12$, is 112%. This implies that the count of forward citations made by other assignees doubles within the next three years for patents that have lapsed, as compared to renewed patents. The magnitude of this finding is surprising. Forward citations might over-estimate the true amount of follow-on innova-

Table 1.3: The Average Effect of Patents on Follow-On Innovation

	(1)	(2)	(3)
Dependent Variable	log(PostCites+1)		
Estimation Method	OLS	IV	IV
Lapsed	-0.018***	0.750***	0.726***
	(0.001)	(0.038)	(0.048)
log(PreCites+1)	0.389***	0.439***	0.422***
	(0.002)	(0.004)	(0.004)
log(PreSelfCites+1)	0.044***	0.117***	0.097***
	(0.004)	(0.007)	(0.007)
log(Claims)	0.001**	0.026***	0.015***
	(0.001)	(0.002)	(0.002)
Technology and Maintenance Effects	Yes	Yes	Yes
Additional Controls	No	No	Yes
Observations	221077	221077	183699
Weak Identification Test		997.3	554.0
Underidentification Test		956.6	545.3

Notes: Robust standard errors reported in parentheses. PostCites is the number of forward citations made by different assignees than the given patent's assignee within three years following the the patent's maintenance event. Lapsed is a binary variable indicating if the given patent was allowed to lapse by the patentee or not. PreCites is the number of forward citations made by other assignees than the given patent's assignee before the maintenance event. PreSelfCites are the forward citations made by the same assignee as the given patent's assignee before the maintenance event. Claims is the number of claims in the given patent. Technology effects are the 37 technological subcategories defined in Hall et al. (2001). Maintenance effects are dummies indicating whether the given patent had to be renewed for the first, second or third time. For the list of variables included as additional controls, see Appendix 1.A. The weak identification test is the Kleibergen-Paap rk Wald F statistic and the underidentification test is the Kleibergen-Paap rk LM statistic reported by the `ivreg2` command in Stata. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

tion that is taking place, as would be the case when a patent is cited although it has not contributed to the creation of the new invention. To investigate this potential issue, I run IV regressions using two alternative outcome variables: (1) the log of one plus $PostTotCites$, the total number of forward citations (made by different assignees and the same assignee as the focal patent within the three years following the patent's maintenance event) and (2) the log of one plus $PostSelfCites$, the number of forward citations made by the same assignee as the focal patent within the three years following the patent's maintenance event.

Table 1.4 shows that the effect of $Lapsed$ is larger than the baseline result when total citations are considered and positive when citations by the same assignee only are considered. However, the point estimate for total citations in column (2) is relatively close to the baseline estimate reported in column (1), and the point estimate for citations by the same assignee, in column (3), is only about half the size of the baseline estimate. This suggests that while some of the baseline effect might be driven by over-estimation, my estimates still capture an effect on true follow-on innovation.

I further check the robustness of the main results by running four different tests. In the first test, I cluster my standard errors at the assignee level to account for possible correlation in the model errors within assignees, some of whom own several patents in my sample. In my second test, I remove from my sample any patent that has been transferred to a different assignee, which allows me to control for possible mismeasurement in the number of forward citations due to re-assignments.[20] In the third test, I add dummies to my baseline specification that indicate whether a given patent received no citation in the period that either preceded or followed its maintenance event. Finally, in the fourth test I use the sample that contains all patents whose latest dates of payment fall within a twenty-six weeks period prior to March 19, 2013 and to

[20]To see how re-assignments may introduce measurement error, consider the following simple example. Assignee A owns a patent that is later cited by assignee B. We will record this as one forward citation by a different assignee for patent A. But assume that, in fact, the patent owned by assignee A was bought by assignee B before the later patent cited the earlier one. Then, really, the forward citation is a self-citation and should not be counted as a citation made by a different assignee. Focusing on patents that have not been re-assigned prevents this kind of measurement error.

Table 1.4: Testing for Potential Overestimation of Follow-On Innovation

	(1)	(2)	(3)
Dep. Var. Based on	PostCites	PostTotCites	PostSelfCites
Estimation Method	Instrumental Variable		
Lapsed	0.750***	0.853***	0.336***
	(0.038)	(0.043)	(0.028)
log(PreCites+1)	0.439***	0.435***	0.019***
	(0.004)	(0.004)	(0.002)
log(PreSelfCites+1)	0.117***	0.340***	0.312***
	(0.007)	(0.008)	(0.007)
log(Claims)	0.026***	0.030***	0.010***
	(0.002)	(0.002)	(0.001)
Technology and Maintenance Effects	Yes	Yes	Yes
Additional Controls	No	No	No
Observations	221077	221077	221077
Weak Identification Test	997.3	997.3	997.3
Underidentification Test	956.6	956.6	956.6

Notes: Robust standard errors reported in parentheses. PostCites is the number of forward citations made by different assignees than the given patent's assignee within three years following the patent's maintenance event. PostTotCites is the total number of forward citations made by different assignees and by the same assignee as the given patent's assignee within three years following the patent's maintenance event. PostSelfCites is the number of forward citations made by the same assignee as the given patent's assignee within three years following the patent's maintenance event. In all three columns the dependent variable is the log of one plus the given measure it is based on. Lapsed is a binary variable indicating if the given patent was allowed to lapse by the patentee or not. PreCites is the number of forward citations made by other assignees than the given patent's assignee before the maintenance event. PreSelfCites are the forward citations made by the same assignee as the given patent's assignee before the maintenance event. Claims is the number of claims in the given patent. Technology effects are the 37 technological subcategories defined in Hall et al. (2001). Maintenance effects are dummies indicating whether the given patent had to be renewed for the first, second or third time. For the list of variables included as additional controls, see Appendix 1.A. The weak identification test is the Kleibergen-Paap rk Wald F statistic and the underidentification test is the Kleibergen-Paap rk LM statistic reported by the `ivreg2` command in Stata. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

all patents whose earliest dates of payment fall within a twenty-six weeks period on or after March 19, 2013, thus doubling my window size. Overall, the results are very robust and more detail on the tests and estimation results can be found in Appendix 1.C.

1.5.3 Variation Across Maintenance Stages

In the previous section, I assumed that the relationship in my baseline specification holds across all patents, as different as they might be. But the theory described in Section 1.2.3 indicates that the effect of patents might be different across the three maintenance stages due to the learning that occurs during the renewal process. I now assess this potential source of variation by splitting my sample into three groups, one for each of the three maintenance stages, and then running my two-step IV regressions over the three subsamples.

I start by investigating the relationship between the fee increases and renewal behaviour across the three maintenance stages by estimating the Probit model for the three relevant subsamples. The results are presented in Table 1.5 and show that the effect of $HighFees$ increases over the three maintenance stages. Thus, patentees seem to be more sensitive to fee increases when they face renewal decisions for older patents. This finding is consistent with the idea that patentees place an option value on their patents, especially at early stages (Pakes, 1986). Patentees are willing to keep young patents in force despite fee increases as they still know little about their potential value. But as their patents age, they are more informed on the potential value of maintaining the patent in force and hence more sensitive to increases in renewal costs. The associated marginal increases in the probability of lapse for the first, second and third maintenance stages are, respectively, approximately one, three and five percentage points. The implied elasticities are then -0.026 at the first maintenance stage, -0.17 at the second maintenance stage and -0.13 at the third maintenance stage.[21]

[21] The figure for the first maintenance stage is found by dividing the overall percentage change of $-(.01/.87) \approx -1\%$ by the increase in the first maintenance fee of 39%. The figure for the second maintenance stage is found by dividing the overall percentage change of $-(.03/.79) \approx -4\%$ by the increase in the first maintenance fee of 24%. The figure for the third maintenance stage is found by dividing the overall percentage change of $-(.05/.73) \approx -7\%$ by the increase in the first maintenance fee of 54%.

Table 1.5: Variation in the Effect of the 2013 Maintenance Fees on Renewal Decisions Across Maintenance Stages

	(1)	(2)	(3)
Dependent Variable		Lapsed	
Estimation Method		Probit	
Maintenance Stage	First	Second	Third
HighFees	0.066***	0.103***	0.157***
	(0.011)	(0.011)	(0.011)
log(PreCites+1)	-0.216***	-0.310***	-0.314***
	(0.016)	(0.015)	(0.014)
log(PreSelfCites+1)	-0.623***	-0.654***	-0.369***
	(0.042)	(0.050)	(0.040)
log(Claims)	-0.181***	-0.096***	-0.072***
	(0.007)	(0.007)	(0.007)
Technology and Maintenance Effects	Yes	Yes	Yes
Additional Controls	No	No	No
Observations	91677	69888	59512

Notes: Robust standard errors reported in parentheses. Lapsed is a binary variable indicating if the given patent was allowed to lapse by the patentee or not. HighFees equals 1 if the given patent was in the high-fee group and 0 if it was in the low-fee group. PreCites is the number of forward citations made by other assignees than the given patent's assignee before the maintenance event. PreSelfCites are the forward citations made by the same assignee as the given patent's assignee before the maintenance event. Claims is the number of claims in the given patent. Technology effects are the 37 technological subcategories defined in Hall et al. (2001). Maintenance effects are absorbed. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Table 1.6 reports the results of the two-step IV regression for each of the three maintenance stages. In column (1), the model is estimated over the subsample that consists of all first maintenance stage patents. The estimated coefficient for the effect of *Lapsed* is 0.360 and significant. In column (2), the sample consists of all second maintenance stage patents. For this subsample, the estimate for the effect of *Lapsed* is 0.171 and significant. Finally, column (3) shows the results for third maintenance stage patents. The estimate for $Lapsed$ now equals -0.157 and is significant again. These findings suggest that young patents have a blocking effect on follow-on innovation, that this blocking effect dissipates over time and that it gives place to a facilitating effect for older patents. Tables 1.8 and 1.9 in Appendix 1.D show that the results are, overall, robust to the inclusion of additional controls.

These findings can be understood in light of the theory discussed in Section 1.2.3. First, at the first and second maintenance stages, when upstream innovators are still uncertain of the value of potential downstream developments, bargaining failures between upstream and downstream innovators result in a blocking effect of patents on follow-on innovation. The learning that occurs during the renewal process decreases uncertainty and enables more efficient bargaining, resulting in less follow-on innovation being blocked over time. Second, since at the third maintenance stage bargaining frictions are no longer binding, the finding that patents promote follow-on innovation support the existence of coordination failures between downstream innovators.

1.5.4 Variation Across Technology Areas

Previous studies have shown that the effect of patents on follow-on innovation varies greatly across technology areas, but they disagree on the areas that are affected more than others. A major point of contention between these studies is the impact of technological complexity. Whereas Galasso and Schankerman (2015) find that patents in complex technology areas, where a given technology builds on many previous technologies, have a stronger blocking effect, Gaessler et al. (2017) find instead that patents block follow-on innovation more in discrete technologies. I revisit the heterogeneity across technology fields by splitting my sample according to the six NBER technology categories and then running IV regressions over the six corresponding subsamples. I then discuss the impact of technological complexity.

Figure 1.3 illustrates the results from the split sample analysis. The plot shows the two-step IV point estimate of $Lapsed$ for the relevant subsample, along with 95% confidence intervals. There is a lot of variation across the six NBER categories and, in contrast to Galasso and Schankerman (2015), but in line with Gaessler et al. (2017), I find that patents have a stronger blocking effect in technologies that are more commonly understood as being discrete. More specifically, the Mechanical, Drugs and Medical, and Chemical fields all

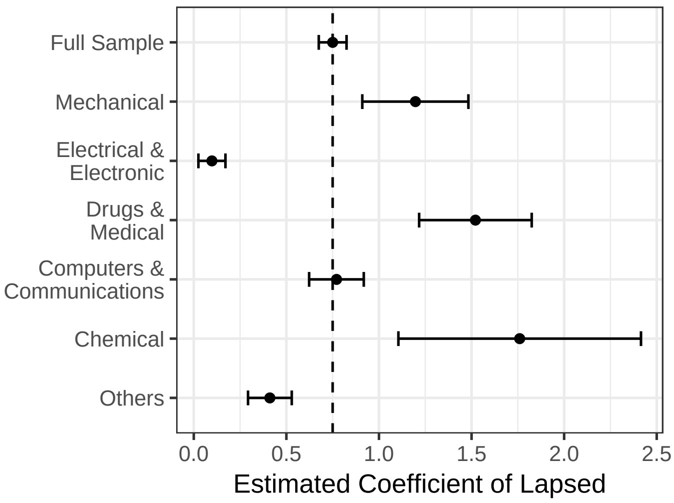
Table 1.6: Variation in the Effect of Patents on Follow-On Innovation Across Maintenance Stages

	(1)	(2)	(3)
Dependent Variable	log(PostCites+1)		
Estimation Method	Instrumental Variable		
Maintenance Stage	First	Second	Third
Lapsed	0.360***	0.171***	-0.157***
	(0.082)	(0.046)	(0.027)
log(PreCites+1)	0.523***	0.374***	0.272***
	(0.006)	(0.006)	(0.004)
log(PreSelfCites+1)	0.099***	0.050***	-0.014*
	(0.012)	(0.010)	(0.008)
log(Claims)	0.020***	0.007***	-0.004***
	(0.004)	(0.002)	(0.001)
Technology and Maintenance Effects	Yes	Yes	Yes
Additional Controls	No	No	No
Observations	91677	69888	59512
Weak Identification Test	248.9	346.8	257.0
Underidentification Test	237.0	323.9	254.0

Notes: Robust standard errors reported in parentheses. PostCites is the number of forward citations made by different assignees than the given patent's assignee within three years following the patent's maintenance event. Lapsed is a binary variable indicating if the given patent was allowed to lapse by the patentee or not. PreCites is the number of forward citations made by other assignees than the given patent's assignee before the maintenance event. PreSelfCites are the forward citations made by the same assignee as the given patent's assignee before the maintenance event. Claims is the number of claims in the given patent. Technology effects are the 37 technological subcategories defined in Hall et al. (2001). Maintenance effects are absorbed. The weak identification test is the Kleibergen-Paap rk Wald F statistic and the underidentification test is the Kleibergen-Paap rk LM statistic reported by the `ivreg2` command in Stata. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

seem to have stronger blocking effects than the other areas. In the Electrical and Electronics field, for instance, patents seem to have almost no impact on follow-on innovation. Figure 1.6 in Appendix 1.D shows that these findings are not affected by the inclusion of additional controls.

Figure 1.3: Variation in the Effect of Patents on Follow-On Innovation Across Technology Areas



Notes: This figure shows how the effect of *Lapsed* varies across the six NBER categories defined in Hall et al. (2001). Each point on the plot corresponds to the point estimate from the two-step IV regression taken over the relevant subsample or over the full sample. The plot also reports 95% confidence intervals for each point estimate. The vertical dotted line marks the point estimate from the IV regression taken over the full sample. All regressions include the variables in the baseline specification.

I then investigate variation across technology areas within maintenance stages along two dimensions, the complexity of technologies and the concentration of patent ownership. As in Galasso and Schankerman (2015), I consider that a patent belongs to a complex technology area if it is in either the Computers and Communications area (NBER category 2), the Electrical and Electronics area (NBER category 4) or the Medical Instruments and the Biotechnology areas (NBER subcategories 32 and 33). Patents in the Chemical area (NBER category 1), the Mechanical area (NBER category 5), Others area (NBER category 6) and Drugs area (NBER subcategory 31) are defined as being in a discrete technology field. I also use a similar measure as Galasso and Schankerman (2015) for patent ownership concentration, which I define as the share of patents owned by the four largest assignees in the given patent's International Patent Classification (IPC) class during the five years that preceded the patent's maintenance event.[22] Using this concentration measure, I divide the sample into two groups, the "Low" group that contains patents whose concentration levels lie below the first quartile of the concentration measure's sample distribution and the "Medium/High" group that contains the remaining patents.

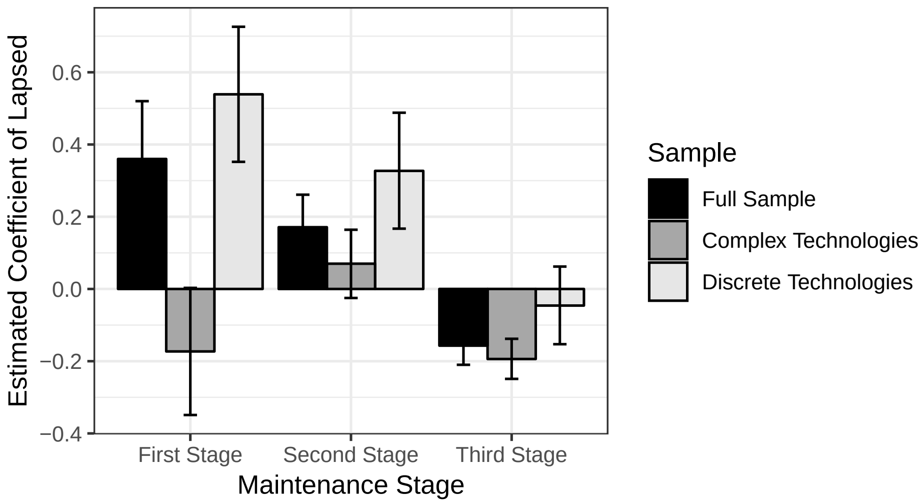
Figure 1.4 shows that patents in discrete technologies have a strong and significant blocking effect at the first two maintenance stages and a small but statistically insignificant positive impact on follow-on innovation at the third stage. Patents in complex technologies, on the other hand, have no statistically significant effects on follow-on innovation at the two first maintenance stages, but a significant facilitating effect at the third maintenance stage.

Figure 1.5 shows that patents in technology fields with low concentration levels exert a strong and significant blocking effect at the first two maintenance stages and does not have a statistically significant effect on follow-on innovation at the third maintenance stage. On the other hand, patents in fields with medium and high concentration levels do not appear to have significant blocking effects at the first and second maintenance stages, but have a statistically significant facilitating effect at the third maintenance stage.

To sum up, the blocking effect of patents, which is present at the first two maintenance stages, appears to be driven by patents in discrete technology fields and in which patent ownership is highly fragmented. We can interpret these findings in the following way. In complex technology areas, a new downstream innovation is more likely to be based on many existing upstream patents. When ownership is fragmented, licensing agreements with several patent holders are necessary to produce the downstream innovation. As a result, the lapse of one of the upstream patents is likely to have a low impact on follow-on innovation, given that many other patents can still exert a blocking effect on the production of the downstream innovation. In discrete technology fields, on the other hand, existing patents are more likely to be used alone or in combination with only a few other patents as inputs to a downstream innova-

[22]The IPC system places patents in relatively narrow technological groups and, therefore, allows me to measure ownership concentration at a finer level than the NBER categories and subcategories.

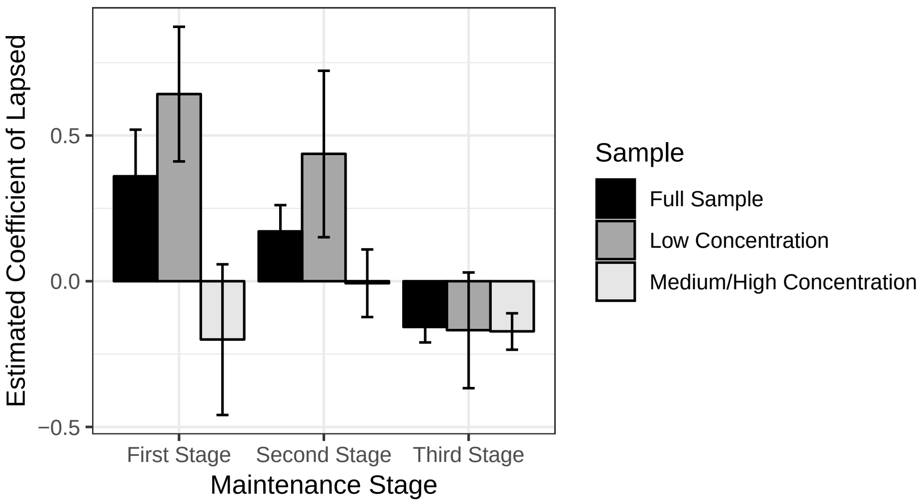
Figure 1.4: Variation in the Effect of Patents on Follow-On Innovation Across Levels of Technological Complexity Within Maintenance Stages



Notes: This figure shows how the effect of *Lapsed* varies across levels of technological complexity at the three maintenance stages. Each bar on the plot corresponds to the point estimate from the two-step IV regression taken over the relevant subsample or over the full sample. The plot also reports 95% confidence intervals for each point estimate. All regressions include the variables in the baseline specification. A patent is defined as being in a complex technology area if it belongs to either the Computers and Communications area (NBER category 2), the Electrical and Electronic area (NBER category 4), the Medical Instruments or the Biotechnology areas (NBER subcategories 32 and 33). Patents in the Chemical area (NBER category 1), the Mechanical area (NBER category 5), Others area (NBER category 6) and Drugs area (NBER subcategory 31) are defined as being in a discrete technology area.

tion. In these fields, the lapse of existing upstream patents is then more likely to result in large amounts of follow-on innovation. Thus, we should indeed expect the blocking effect of patents to be strong in discrete technology areas and when ownership is fragmented. At the third maintenance stage, patents have a facilitating effect on follow-on innovation, which suggests that bargaining frictions are not binding. As expected then, complexity and patent ownership, which both are characteristics of the bargaining environment, have no or very little impact on the (facilitating) effect of patents.

Figure 1.5: Variation in the Effect of Patents on Follow-On Innovation Across Levels of Patent Ownership Concentration Within Maintenance Stages



Notes: This figure shows how the effect of *Lapsed* varies across levels of patent ownership concentration at the three maintenance stages. Each bar on the plot corresponds to the point estimate from the two-step IV regression taken over the relevant subsample or over the full sample. The plot also reports 95% confidence intervals for each point estimate. All regressions include the variables in the baseline specification. Ownership concentration is defined as the share of patents owned by the four largest assignees in the given patent's IPC class during the five years that preceded the patent's maintenance event. Patents whose measure is below the first quartile of the measure's sample distribution are classified as "Low" and remaining patents are classified as "Medium/High."

1.6 Conclusion

This paper investigates the empirical effects of patents on follow-on innovation using U.S. data on patent renewals. I generate quasi-experimental variation in renewal decisions using the maintenance fee increases that came into effect in 2013 after the PTO gained fee-setting authority. I find that, on average, patents have a strong blocking effect on follow-on innovation. This effect, which is driven by patents in the first and second maintenance stages, dissipates over time and gives place to a facilitating effect at the third maintenance stage. I also find that the blocking effect is stronger in discrete technology areas and when patent ownership is fragmented.

My findings show that, in a world where technological advances are spurred mostly by existing technologies, the incentive effects of patent protection on follow-on innovation cannot be ignored. But they also indicate that adequate policy reforms should aim primarily at facilitating licensing agreements early on in patents' lives and for patents in specific technology fields, as it is in these cases that patent protection appears to have particularly detrimental effects on follow-on innovation.

Appendix

1.A Additional Control Variables

The set of additional patent and owner characteristics consists of the following variables:

- The log of the total number of IPC classes listed in the given patent, which is often used as an indicator of patent value;
- The log of the total number of U.S. Patent Classification classes listed in the patent, also used at times as an indicator of patent value;
- A dummy variable indicating whether the given patent has a foreign priority date, meaning it has been filed at a foreign patent office as well as at the PTO. Also a measure that is sometimes used as an indicator of patent value;
- A set of dummy variables for filing years of sample patents;
- A set of entity size dummies, indicating whether the assignee was recorded as a small entity at the PTO. Small entities are either individual inventors, universities, non-profit organisations or small companies (fewer than 500 employees) and they pay lower fees than large entities. Because the additional micro entity status was only added on March 19, 2013, I record it as small entity status in my data (this corresponds to an almost negligeable number of cases).
- A set of dummy variables indicating the assignee's type, which can be a U.S. company, a foreign company, a U.S. individual, a foreign individual, the U.S. government, a foreign government, a country government or a state government.

1.B Cleaning of the Data

The data had to be cleaned in the following way before being ready to be used in my empirical analysis. First, any patent for which a petition for reinstatement had been filed was removed from my sample. These are patents whose holders had unintentionally forgotten to pay the maintenance fee on time and wished to reinstate their patents. Such petitions are rare, representing about 0.3% of all occurences in the maintenance fee events dataset. Second, my sample contained a number of duplicate entries. More specifically, there were 84 duplicates in total, 83 of which were duplicated records of expiration due to failure to pay the maintenance fee. The other duplicate was due to a company who paid the wrong fee. These duplicates, but not their originals, are removed from the sample. Third, a negligible amount of patents, on the order of 0.1% of the sample, were removed from the sample because the date at which their maintenance event was recorded was inconsistent. More specifically, given the way the sample was selected, all sample patents should have event dates between December 18, 2011 and June 11, 2014, but these inconsistent patents had event dates outside of that time period. Another insignificant amount of patents were discarded because they either had an event date that preceeded their earliest possible dates of payment or had an event date that succeeded their latest possible dates of payment. Finally, two patents were discarded because they lacked information on their technology area and one additional patent was deleted because it had an incorrect technology area (it was recorded as in the NBER category 7, but only six NBER categories exist).

1.C Robustness Checks

Table 1.7 presents the results of a series of tests that I conduct to investigate the robustness of the main findings. In Column (1), the standard errors are clustered at the patent assignee level to account for the fact that some assignees own more than one patent in the sample. Whereas errors are plausibly independent across assignees, they could be correlated within assignees. Cameron and Miller (2015) explain that when there is correlation betwee model errors within clusters, the precision of estimators could be greatly overstimated if default standard errors are used. The results based on cluster-robust standard errors in column (1), however, show that the effect remains significant even when correcting for this potential issue. In Column (2), I verify that the estimates are not affected by the re-assignment of patents to new assignees. The assignee information recorded in the data dates back to when the patent was issued and patents might have been transferred to other assignees by the time the maintenance event takes place. This could introduce measurement error in forward citations made by the same versus different assignees. To control for this potential problem, I focus on the subsample of patents for which no post-issuance assignments were recorded at the PTO. More specifically, using the patent assignment research dataset provided by the PTO, I discard from my sample any patent for which an assignment was recorded. Note that I do not exclude employee-employer assignments as these are not real transfers between different entities, but I do count in assignments resulting from mergers. A limitation of this data is that assignees are not obligated to notify the PTO of an assignment. However, it is the only widely and publicly available data on re-assignments. Column (2) shows that there are only minor changes in the estimates when using this subsample instead of the full sample. In Column (3), I include dummies for patents that have received no forward citations before their maintenance event and after their maintenance event. The estimate for $Lapsed$ is larger and still highly significant. Finally, in Column (4), I focus on the sample that contains all patents whose latest dates of payment fall within a twenty-six weeks period prior to March 19, 2013 and to all patents whose earliest dates of payment fall within a twenty-six weeks period on or after March 19, 2013. The results are very similar to those based on the thirteen weeks window considered in the baseline analysis.

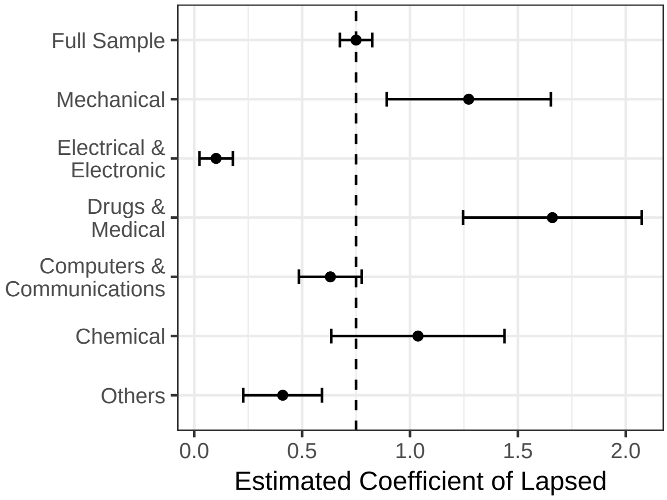
Table 1.7: Robustness Checks for the Average Effect of Patents on Follow-On Innovation

	(1)	(2)	(3)	(4)
Dependent Variable	log(PostCites+1)			
Estimation Method	Instrumental Variable			
Lapsed	0.680***	0.690***	1.152***	0.604***
	(0.108)	(0.044)	(0.049)	(0.025)
log(PreCites+1)	0.428***	0.423***	0.347***	0.425***
	(0.009)	(0.005)	(0.004)	(0.003)
log(PreSelfCites+1)	0.106***	0.131***	0.097***	0.104***
	(0.021)	(0.009)	(0.007)	(0.005)
log(Claims)	0.022***	0.025***	0.031***	0.020***
	(0.005)	(0.002)	(0.002)	(0.001)
Technology and Maintenance Effects	Yes	Yes	Yes	Yes
Additional Controls	No	No	No	No
Observations	203391	150242	221077	449180
Weak Identification Test	103.4	782.3	724.4	2050.3
Underidentification Test	87.5	709.2	716.7	1969.0

Notes: Robust standard errors reported in parentheses, except in column (1) where standard errors are clustered at the assignee level. Column (2) excludes patents that have been reassigned. Column (3) adds dummies for patents that have received no citations either prior to or after the maintenance event. In Column (4), the sample contains all patents whose latest dates of payment fall within a twenty-six weeks period prior to March 19, 2013 and to all patents whose earliest dates of payment fall within a twenty-six weeks period on or after March 19, 2013. PostCites is the number of forward citations made by different assignees than the given patent's assignee within three years following the patent's maintenance event. Lapsed is a binary variable indicating if the given patent was allowed to lapse by the patentee or not. PreCites is the number of forward citations made by other assignees than the given patent's assignee before the maintenance event. PreSelfCites are the forward citations made by the same assignee as the given patent's assignee before the maintenance event. Claims is the number of claims in the given patent. Technology effects are the 37 technological subcategories defined in Hall et al. (2001). Maintenance effects are dummies indicating whether the given patent had to be renewed for the first, second or third time. The weak identification test is the Kleibergen-Paap rk Wald F statistic and the underidentification test is the Kleibergen-Paap rk LM statistic reported by the `ivreg2` command in Stata. Significance levels: *$p<0.1$; **$p<0.05$; ***$p<0.01$.

1.D Additional Figures and Tables

Figure 1.6: Variation in the Effect of Patents on Follow-On Innovation Across Technology Areas — Robustness to the Inclusion of Additional Controls



Notes: This figure shows how the effect of *Lapsed* varies across the six NBER categories defined in Hall et al. (2001). Each point on the plot corresponds to the point estimate from the two-step IV regression taken over the relevant subsample or over the full sample. The plot also reports 95% confidence intervals for each point estimate. The vertical dotted line marks the point estimate from the IV regression taken over the full sample. All regressions include the variables in the baseline specification and the set of additional controls.

Table 1.8: Variation in the Effect of the 2013 Maintenance Fees on Renewal Decisions Across Maintenance Stages — Robustness to the Inclusion of Additional Controls

	(1)	(2)	(3)
Dependent Variable		Lapsed	
Estimation Method		Probit	
Maintenance Stage	First	Second	Third
HighFees	0.057***	0.133***	0.217***
	(0.013)	(0.014)	(0.016)
log(PreCites+1)	-0.148***	-0.221***	-0.248***
	(0.017)	(0.016)	(0.015)
log(PreSelfCites+1)	-0.508***	-0.484***	-0.272***
	(0.042)	(0.048)	(0.040)
log(Claims)	-0.152***	-0.060***	-0.047***
	(0.008)	(0.008)	(0.008)
Technology and Maintenance Effects	Yes	Yes	Yes
Additional Controls	Yes	Yes	Yes
Observations	74760	58067	50839

Notes: Robust standard errors reported in parentheses. Lapsed is a binary variable indicating if the given patent was allowed to lapse by the patentee or not. HighFees equals 1 if the given patent was in the high-fee group and 0 if it was in the low-fee group. PreCites is the number of forward citations made by other assignees than the given patent's assignee before the maintenance event. PreSelfCites are the forward citations made by the same assignee as the given patent's assignee before the maintenance event. Claims is the number of claims in the given patent. Technology effects are the 37 technological subcategories defined in Hall et al. (2001). Maintenance effects are absorbed. For the list of variables included as additional controls, see Appendix 1.A. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Table 1.9: Variation in the Effect of Patents on Follow-On Innovation Across Maintenance Stages — Robustness to the Inclusion of Additional Controls

	(1)	(2)	(3)
Dependent Variable	log(PostCites+1)		
Estimation Method	Instrumental Variable		
Maintenance Stage	First	Second	Third
Lapsed	0.642***	-0.012	-0.245***
	(0.157)	(0.066)	(0.039)
log(PreCites+1)	0.520***	0.360***	0.274***
	(0.007)	(0.006)	(0.005)
log(PreSelfCites+1)	0.108***	0.028***	-0.012
	(0.015)	(0.011)	(0.008)
log(Claims)	0.022***	-0.001	-0.004***
	(0.005)	(0.002)	(0.001)
Technology and Maintenance Effects	Yes	Yes	Yes
Additional Controls	Yes	Yes	Yes
Observations	74760	58067	50839
Weak Identification Test	74.8	106.5	136.6
Underidentification Test	73.3	105.2	135.8

Notes: Robust standard errors reported in parentheses. PostCites is the number of forward citations made by different assignees than the given patent's assignee within three years following the patent's maintenance event. Lapsed is a binary variable indicating if the given patent was allowed to lapse by the patentee or not. PreCites is the number of forward citations made by other assignees than the given patent's assignee before the maintenance event. PreSelfCites are the forward citations made by the same assignee as the given patent's assignee before the maintenance event. Claims is the number of claims in the given patent. Technology effects are the 37 technological subcategories defined in Hall et al. (2001). Maintenance effects are absorbed. For the list of variables included as additional controls, see Appendix 1.A. The weak identification test is the Kleibergen-Paap rk Wald F statistic and the underidentification test is the Kleibergen-Paap rk LM statistic reported by the `ivreg2` command in Stata. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Bibliography

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Bessen, J. (2004). Holdup and Licensing of Cumulative Innovations With Private Information. *Economics Letters*, 82(3):321–326.

Bessen, J. and Maskin, E. (2009). Sequential Innovation, Patents, and Imitation. *The RAND Journal of Economics*, 40(4):611–635.

Cameron, C. A. and Miller, D. L. (2015). A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources*, 50(2):317–372.

Cockburn, I., Kortum, S., and Stern, S. (2002). Are All Patent Examiners Equal? The Impact of Examiner Characteristics. NBER Working Papers 8980.

de Rassenfosse, G. and van Pottelsberghe de la Potterie, B. (2012). The Role of Fees in Patent Systems: Theory and Evidence. *Journal of Economic Surveys*, 27(4):696–716.

Farre-Mensa, J., Hegde, D., and Ljungqvist, A. (2017). What Is a Patent Worth? Evidence from the U.S. Patent "Lottery". USPTO Economic Working Paper 2015-5.

Gaessler, F., Harhoff, D., and Sorg, S. (2017). Patents and Cumulative Innovation: Evidence from Post-Grant Patent Oppositions. *Academy of Management Proceedings*, 2017(1).

Galasso, A. and Schankerman, M. (2010). Patent Thickets, Courts, and the Market for Innovation. *The RAND Journal of Economics*, 41(3):472–503.

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

Galasso, A. and Schankerman, M. (2018). Patent Rights, Innovation, and Firm Exit. *The RAND Journal of Economics*, 49(1):64–86.

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

Hall, B. H., Helmers, C., Rogers, M., and Sena, V. (2014). The Choice Between Formal and Informal Intellectual Property: A Review. *Journal of Economic Literature*, 52(2):375–423.

Hall, B. H., Jaffe, A. B., and Trajtenberg, M. (2001). The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools. NBER Working Papers 8498.

Huang, K. and Murray, F. (2009). Does Patent Strategy Shape the Long-Run Supply of Public Knowledge? Evidence from Human Genetics. 52:1193–1221.

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

Lanjouw, J. O. (1998). Patent Protection in the Shadow of Infringement: Simulation Estimations of Patent Value. *Review of Economic Studies*, 65(4):671 – 710.

Lanjouw, J. O., Pakes, A., and Putnam, J. (1998). How to Count Patents and Value Intellectual Property: The Uses of Patent Renewal and Application Data. *Journal of Industrial Economics*, 46(4):405–432.

Marco, A. C., Myers, A. F., Graham, S. J., D'Agostino, P. A., and Apple, K. (2015). The USPTO Patent Assignment Dataset: Descriptions and Analysis. *SSRN Electronic Journal*.

Moser, P. and Voena, A. (2012). Compulsory Licensing: Evidence from the Trading with the Enemy Act. *American Economic Review*, 102(1):396–427.

Murray, F. and Stern, S. (2007). Do Formal Intellectual Property Rights Hinder the Free Flow of Scientific Knowledge? *Journal of Economic Behavior & Organization*, 63(4):648–687.

Pakes, A. (1986). Patents as Options: Some Estimates of the Value of Holding European Patent Stocks. *Econometrica*, 54(4):755–784.

Pakes, A. and Schankerman, M. (1984). The Rate of Obsolescence of Patents, Research Gestation Lags, and the Private Rate of Return to Research Resources. In *R&D, Patents, and Productivity*, pages 73–88.

Sampat, B. and Williams, H. L. (2019). How Do Patents Affect Follow-On Innovation? Evidence from the Human Genome. *American Economic Review*, 109(1):203–236.

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

Spulber, D. F. (2015). How Patents Provide the Foundation of the Market for Inventions. *Journal of Competition Law and Economics*, 11(2):271–316.

United States Patent and Trademark Office (2013). Setting and Adjusting Patent Fees; Final Rule. *Federal Register*, 78(13).

Watzinger, M., Fackler, T. A., Nagler, M., and Schnitzer, M. (2017). How Antitrust Enforcement Can Spur Innovation: Bell Labs and the 1956 Consent Decree. CESifo Working Paper Series 6351.

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

Williams, H. L. (2017). How Do Patents Affect Research Investments? *Annual Review of Economics*, 9(1):441–469.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. The MIT Press.

Ziedonis, R. H. (2004). Don't Fence Me In: Fragmented Markets for Technology and the Patent Acquisition Strategies of Firms. *Management Science*, 50(6):804–820.

Chapter 2

The Examination of Continuation Applications and the Problem of Invalid Patents in the U.S.[1]

[1]I thank Tore Ellingsen, Brian Wright, Helen Sweeting and Stuart Graham for their comments and feedback. Financial support from the Jan Wallander and Tom Hedelius Foundation is gratefully acknowledged. All errors are my own.

2.1 Introduction

Many scholars believe that the U.S. Patent and Trademark Office (PTO) grants a large number of invalid patents. These patents cover either existing technologies or obvious technological advances and may stall innovation by creating a climate of uncertainty for innovators. Litigation is one way patent invalidity can be addressed. But it is often lengthy and prohibitively expensive and hence not necessarily the appropriate way of weeding out bad patents. Alternative measures have thus been taken to tackle the issue of invalid patents more efficiently. For instance, the America Invent Act, which was signed into law in 2011, instituted post-grant reviews allowing third parties to challenge the validity of patents within nine months after their issuance.

But scholars have argued that these measures do not tackle the root causes of invalid patents, which are ultimately linked to the patent system itself and, in particular, to the examination process (Frakes and Wasserman, 2017a). If examiners would not allow invalid patents in the first place, there would be no need for post-grant procedures or litigation. A growing body of research investigates causal relationships between features of the patent system and the granting of invalid patents using rich datasets and econometric methods. The main insight that is gained from these studies is that the problem of invalid patents can be addressed by changing some key administrative features of the PTO that bias examiners towards granting rather than rejecting patents.

However, because this body of research advocates large-scale changes to the patent system, it faces a limitation that can be traced back to an earlier idea promoted by Lemley (2001) on "rational ignorance" at the PTO. Lemley (2001) argued that, since few patents are of any value and are ever litigated, the costs that must be incurred to achieve across-the-board improvements of the patent system and examination process would far exceed the benefits from granting fewer invalid patents. Under this view, if anything, changes to patent examination should be implemented only for patents that present a significant risk of being invalid and later litigated.

In this paper, I study empirically the relationship between features of the examination process at PTO and examiners' practices in the context of continuation applications. A continuation application is a patent application that emanates from an existing application that is still pending at the PTO. It may contain new claims as long as it presents the same underlying invention as the earlier application. Thus, even if the earlier application ultimately results in a grant, a continuation application might not be granted if the new claims make it unpatentable. Lemley and Moore (2004) warn that continuation applications are used by patent applicants largely to abuse the patent system and obtain invalid patents. Several studies have shown that patent families that contain a large number of continuation applications are more likely to be involved in litigation (Lemley and Moore, 2004; Allison et al., 2009; United States Patent and Trademark Office, 2015).[2] In addition, continuation applications are common. In the fiscal year 2015, for instance, they accounted for about 20% of all filings at the PTO (Rantanen, 2015). The high risk of litigation and invalidity associated with continuation applications make them an ideal empirical setting to study the issue of patent invalidity at the PTO.[3]

A continuation application is automatically assigned to the examiner who examined the earlier application that it emanated from. I refer to this institutional feature of the U.S. continuation system as "relatedness," as in such circumstances the examiner and the continuation application are related via the earlier application. Lemley and Moore (2004) argue that relatedness leads to the wearing down of examiners and, ultimately, to the granting of invalid patents. Although this idea is well-known, to the best of my knowledge it has never been empirically tested. To study the implications of relatedness, I exploit the fact that, occasionally, examiners are assigned to continuation applications emanating from applications that they have not reviewed themselves. Because examiners and applicants exert no control over the assignment of applications, coupled with the fact that applications are randomly assigned within well-defined examiner working groups ("art units"), these occasional re-assignments are plausibly exogenous and can then be used to evaluate the causal effects of relatedness.

My identification strategy is thus based on a comparison of examination practices as examiners come across "related" versus "unrelated" continuations. I implement this strategy by estimating regressions with examiner fixed effects, which allow me to keep constant all unobserved examiner characteristics that

[2] A patent family is a group of patents that protect a single invention. It can consist of patents that contain new claims, or patents filed abroad, among other things.

[3] In particular, concerns such as those emerging from the rationale ignorance view are minimised in the present context.

might influence examination practices. As previous research has shown that there is a great deal of variation in leniency and examination efforts across examiners (Lemley and Sampat, 2012; Cockburn et al., 2002), controlling for examiner fixed effects is essential to prevent estimation bias due to unobservable examiner characteristics.

My paper has four different aims. The first aim is to obtain reduced-form estimates of the effects of relatedness on examination practices. I focus on the examiners' grant decisions, their efforts to limit the scope of protection sought by applicants in their applications and, finally, their efforts to search for relevant "prior art" during the examination process.[4] The second aim is to investigate whether the wearing down hypothesis formulated by Lemley and Moore (2004) is a plausible channel for the effect of relatedness. The third aim is to assess whether some technology areas, and especially those that rely more on continuation applications than others, are more exposed to the effects of relatedness. The fourth aim is to evaluate whether relatedness ultimately leads to the granting of invalid patents.

My findings show that relatedness increases grant rates, decreases examiners' efforts to narrow down the scope of protection claimed by patent applicants and decreases the examiners' search efforts for prior art. These findings suggest that examiners are more lenient towards related continuations or, at the very least, adopt "softer" examination practices when reviewing continuations of earlier applications that they were also assigned to. The estimates are economically meaningful, with magnitudes in the order of -30% to 6% depending on the outcome variable. However, I do not find evidence of a wearing down effect. If anything, continuations that are more likely to be "repeat" filings (replica of the earlier applications) are less likely to be granted when the examiner was also assigned to the earlier application, which is at odds with the wearing down hypothesis. Relatedness also does not appear to uniformly impact certain technology fields more than others. Finally, when I focus on continuations granted by the PTO and for which protection was also sought at the European Patent Office (EPO) and the Japan Patent Office (JPO), I find that relatedness makes a grant at the two foreign offices less likely. Although the three patent offices have similar patentability standards, the two foreign

[4]Prior art consists in any evidence that a given invention is already known by the public, such as previous patents or scientific publications.

offices invest more resources in patent examination (Picard and van Pottelsberghe de la Potterie, 2013; Lei and Wright, 2017). These findings suggest that relatedness results in a larger number of granted patents that are likely to be of more dubious validity.

This paper is mainly related to the growing body of studies that investigate the relationship between features of the PTO and examination practices and outcomes. Frakes and Wasserman (2017b) show that, as examiners get promoted, they face tighter time constraints to review patent applications and that makes them grant more applications of questionable validity. Frakes and Wasserman (2013) focus on the PTO's fee schedule and demonstrate that, since the PTO relies mostly on post-grant fees to fund itself, it has a natural tendency to over-grant. Finally, Frakes and Wasserman (2015a) propose to limit the number of repeat filings that applicants can make to decrease the PTO's backlog, which is another possible cause of over-granting at the patent office. This paper is also related to the study of Lemley and Sampat (2012), who showed that examiners become more lenient as they gain more experience within the PTO, a finding that was later backed by Frakes and Wasserman (2017b).

The paper is structured as follows. In Section 2.2, I discuss my empirical setting. Section 2.3 describes the data and reports summary statistics for the main variables used in the empirical analysis. In Sections 2.4 and 2.5, I introduce my empirical model and present its results. Section 2.6 concludes the paper.

2.2 Empirical Setting

In this section, I give all the necessary information on the institutional context. I start with a description of the patent examination process, I then move on to a discussion of continuation applications and, finally, I discuss the idea of relatedness and its potential effects on examiner behaviour.

2.2.1 The Patent Examination Process

When a new patent application is filed at the PTO, it is first checked by the Office of Patent Application Processing for completeness. An application is considered complete if it contains a specification, all necessary drawings and at least one claim.[5] Claims delineate the scope of protection conferred by a patent and, thus, constitute one of the most important parts of the application. If an application is complete, it is given a filing date and a technological class, which is then used to assign the application to an art unit.[6] Art units are work groups of examiners specialising in the examination of certain technologies.

Once the application reaches its art unit, it is assigned to an examiner. Several studies have documented the random nature of this assignment, which follows a "first-in-first-out" principle based on filing dates and examiner availabilities (Sampat and Williams, 2019; Farre-Mensa et al., 2017; Frakes and Wasserman, 2017b; Lemley and Sampat, 2012). The examination phase is then open and the assigned examiner checks for the patentability of the application. This involves a search for "prior art," which consists in all previous patents and non-patent publications relevant to the given application and against which the applications' claims are evaluated for novelty and non-obviousness. It also involves an evaluation of the specification. The written description contained in the specification and the claims should describe the claimed invention appropriately. If all these requirements are met, the examination results in an allowance and, eventually, in the grant of a patent if issue fees are duly paid by the applicant. If the examiner allows the application without at first rejecting it on the grounds of invalidating prior art or a misspecified application, the examination outcome is called a "first action allowance."

Typically, however, an examiner will first issue a non-final rejection, which generally consists in a written, formal, objection to one or more of the claims in the application. After receiving a non-final rejection, an applicant can either abandon the application, try to convince the examiner that the objected claims

[5]The specification must contain specific items, such as a title, an abstract and a detailed description of the invention. These are known as the written descriptions in a patent application. Also, a distinction is made at the PTO between provisional and non-provisional applications. The former do not have to contain claims and, unlike the latter, are never examined. This paper investigates examination practices and, therefore, focuses on non-provisional applications only.

[6]The PTO has its own classification system, the U.S. Patent Classification system (USPC), used to organise applications into relatively narrow technology groups (see United States Patent and Trademark Office (2012) for more information on the classification system used at the PTO).

are valid or amend the claims to satisfy the examiner's requests. Following the applicant's response, the examiner may decide to either allow the application if the response is deemed satisfactory or to issue a final rejection if some issues remain. As clarified by Lemley and Moore (2004), however, the term "final rejection" is a misnomer. In practice, examiners cannot finally reject applications because applicants have several options to re-open the examination process after final rejections. These include Requests for Continued Application (RCE), appeals filed at the Patent Trial and Appeal Board, and the filing of continuation applications. For a patent application to be finally rejected, it must thus be abandoned by the applicant, who can do so at any time during the examination process.

Most applicants seek broad claims and examiners often attempt to narrow them down during the examination process by issuing rejections. Thus, in a nutshell, the patent examination process is a back-and-forth negotiation over the breadth of the application's claims. To narrow down the breadth, examiners may simply ask the removal of some of the claims. More often, though, examiners will allow claims they initially objected to, on the condition that these claims are amended in a way that makes the application patentable.[7] As longer claims are generally more specific and, thus, less broad, claims narrowing often involves the addition of words to some key claims in the application. Scholars have thus considered the number of words added to claims during the examination process as a measure of the extent of claims narrowing. Marco et al. (2016b), for instance, consider the word-length difference in the shortest independent claim, while Kuhn and Thompson (2017) look at the first independent claim.[8] In practice, however, these two measures may be very similar, as for many patent applications the first independent claim is also the shortest one. Once the examiner and the applicant agree on the claims, the examiner allows the application, which then later issues as a patent.

[7] Typically, this will involve modifying the claims so as to not overlap with prior art.

[8] There are two types of claims: dependent and independent. Dependent claims always refer back to one or more independent claims, whereas independent claims do not refer to any other claims. Dependent claims are therefore, by definition, narrower than the independent claims they refer to, as they add specificity.

2.2.2 Continuation Applications

A continuation application is a patent application that emanates from a previously filed application that is still pending at the PTO. It always inherits the filing date of the application from which it emanates. If this latter application is itself a continuation of a previously filed application, then the date of the earlier application is the inherited filing date. More generally, the inherited filing date of any given continuation application is the filing date of the earliest filed application in the series of applications from which it emanates. The first application in that series is often called the "parent" application, and its filing date, which applies to all subsequent continuation applications, is referred to as the "priority" date.

Continuation applications may contain new claims, relative to their parent applications, as long as these new claims can be supported by their parents' specifications. Thus, information in a continuation application's specification that is used to make new claims ("new matter"), along with those additional claims, is not allowed and would be rejected during the examination of the continuation. In short, the underlying invention must be the same in a continuation application and its parent, but claims to that invention can differ across the applications. There is no requirement, however, that a continuation application should differ from its parent. In other words, applicants can file for exactly the same application over and over again in a string of continuations.[9]

Continuation applications allow applicants to keep the examination process on existing applications open. As long as the parent application is still pending, the applicant can try to obtain old or new claims by filing a continuation application that includes those claims. Because the continuation application inherits the filing date of its parent application, the newer claims are evalu-

[9] A popular alternative to continuation applications are Continuations-in-Parts (CIPs), which allow applicants to include new matter along with new claims in their applications. Because CIPs allow the inclusion of new matter, the content of a CIP might be significantly different than that of the parent application. Also, the rules for priority dates are a lot more intricate and depend on whether the new claims depend on old or new matter in the CIP. This, naturally, changes the way the examiner will assess the patentability of the claimed invention. Because continuation applications do not allow new matter and have simple rules for priority dates, they offer a much cleaner empirical setting to study relatedness and its effects on examination practices. CIPs are therefore not considered in the empirical analysis.

ated by the examiner in view of the existing prior art at the priority date. Thus, continuation applications allow inventors to protect new discoveries made after an initial patent application was filed for. But since continuation applications cannot include new matter, effectively limiting the scope of new claims, one of the more popular uses for continuation applications it to keep fighting for claims that were originally submitted but could not be obtained during the examination of the parent application.[10]

2.2.3 Relatedness and Its Effects

In the U.S., a continuation application is assigned to the same examiner who examined the parent application. In this paper, I refer to this institutional feature of the U.S. patent system as "relatedness." I also often talk about "related" continuation applications, which are defined as continuations that are examined by the same examiner who was assigned to their parent applications. Similarly, I say of an examiner that she is a "related" examiner of a given continuation application when she has also examined its parent application. In situations that are not characterised by relatedness, I refer to examiners and continuations as "unrelated."

Given that continuation applications are automatically assigned to related examiners, some scholars have postulated that applicants strategically file continuation applications to wear examiners down (Lemley and Moore, 2004). Under this "wearing down" hypothesis, applicants who are unsatisfied with the progress or expected outcome of their applications' examination file for continuations to obtain the claims that they were seeking in the parent applications. For instance, if an applicant feels that the examiner will not allow any of her claims, she might abandon that existing application after having filed an exact copy of it as a continuation. Similarly, if an applicant feels that the examiner is willing to allow some claims, but not others, she might settle for those claims in the initial application and file for a continuation to try and obtain the remaining claims. Either way, the wearing down hypothesis says that examiners will grant claims in continuations that they were unwilling to grant

[10] In practice, applicants can also continue the prosecution of their applications by filing requests for continued examination. RCEs are not "serialised" applications and, in turn, are not recorded as separate applications from the parent applications in the PTO records. Thus, they cannot be used to study the effects of relatedness on examiner practices.

in parent applications. This is why, according to Lemley and Moore (2004), continuation applications may play a role in the problem of invalid patents.

The wearing down hypothesis presupposes that unrelated examiners, for whom by definition no wearing down can operate, will not allow invalid claims. Another presumption of the wearing down hypothesis is that examiners will eventually allow claims that they previously deemed unallowable provided an applicant shows enough persistence. This is not an unreasonable assumption given the strong time pressures that examiners face and the fact that the only way an examiner can be sure to dispose an application is by allowing it. With applications accumulating, it might be in the examiner's interest to "give in" and allow claims in a continuation that she previously deemed unallowable.

Although the wearing down hypothesis is compelling, it is not clear that relatedness necessarily leads to too many patents or claims being granted. Relatedness can also plausibly lead to efficiencies if examiners become better at determining the patentability of continuations if they have also examined the parent applications. Indeed, an examiner who has already searched for relevant prior art and studied the technology underlying a given claimed invention may have a head start in assessing the patentability of a continuation application, as compared to a newly assigned examiner.

Whether relatedness has positive or negative consequences appears to be, ultimately, an empirical question. In this paper, I study its effects not only on the grant decision, but more generally on examination practices, such as claims narrowing and prior art search efforts. I describe in the following section the various outcomes studied in the empirical analysis.

2.3 Data

My objective is to construct a dataset of continuation applications that I can use to estimate an empirical model of the effects of relatedness with examiner fixed effects. To create such a dataset, I use a number of data files contained in the Patent Examination Research Dataset, PatentsView and the Patent Claims Research Dataset, all three released by the PTO. This section outlines the construction of my dataset. A more detailed discussion is provided in Appendix 2.A.

I first create families of continuation applications and their parents. A parent is any first-time application that has at least one continuation application. I focus on families that contain exactly one parent and one continuation, which account for about 80% of all continuation families. This choice is motivated by the desire to have, in the present paper, a manageable empirical setup to conduct the analysis, postponing a more general treatment to future work on the topic. The inclusion of families with three members or more would introduce complications in the estimation of the effect of relatedness, as in those cases this effect is likely to vary and spillover from one continuation to the next. Focusing on families with a single continuation allows for a cleaner interpretation of the effect of relatedness without great loss of generality in the empirical results.

For each family in my dataset, I can determine whether a given continuation and its parent have been examined by the same examiner by looking directly at the unique identification number assigned by the PTO to its examiners in its data. If the identification number matches for the continuation and its parent, then the examiner and the continuation application are defined as being related.

In terms of outcomes, I consider several variables that are related to three main aspects of examination practices: the grant decision, the extent of claims narrowing for granted applications and the intensity of search efforts to identify prior art for granted applications. The grant decision variable is available directly in the PTO data, which tells us whether an application was granted or not. However, using this outcome variable forces me to consider only applications filed after 2001, as prior to that year the PTO did not publish abandoned applications, meaning that only granted applications appear in the PTO's files for all years prior to 2001. Additionally, I narrow down my sample of applications to non-pending regular utility patent applications. Essentially, this removes all "special" types of applications, such as design or plant patents, reissues and applications destined to foreign filing (e.g. applications for which protection under the Patent Cooperation Treaty is sought). Applications that were still pending at the time the data was released are discarded, as for these applications, by definition, the grant decision variable does not exist yet.

Measures of claims narrowing should be based on a comparison of the claims desired by the applicant with the claims ultimately obtained, i.e. the granted claims. Only such measures may reflect the impact of examination on an application's claims. Note, however, that by definition such measures are only defined for applications that were ultimately granted, a subset of all applications in my sample. I measure efforts made by examiners to narrow down granted applications' claims during examination in two ways. First, I consider the incidence of first action allowances. The absence of rejections implies that the examiner did not require any narrowing of an application's claims before allowing it. Keeping the patentability of an application's claims constant, the absence of any rejection suggests that the examiner has put in lower efforts to narrow down the application's claims.[11] Second, I use the number of words added to the shortest independent claim. To construct this measure, I follow existing studies and compute the difference between the word-length of the shortest independent claim at the time of grant and at the time of publication. As argued by Marco et al. (2016b), although applications are published only 18 months after their filing date, due to backlogs and pendency times at the PTO, amendments required by examiners typically occur later than these initial 18 months. The claims in published applications can therefore be used to proxy the claims that were initially desired and submitted by applicants (which are unfortunately not available in the data).[12]

Finally, to measure examiners' search efforts to detect relevant prior art, I use both the number and share of citations made by examiners to prior art in the sample continuations. My main results are based on citations made to prior patents, mainly to preserve larger sample sizes as data on citations to the non-patent literature is not widely available. However, robustness checks confirm the results when using citations to the non-patent literature instead. As citations are not made available until a patent issues, the analysis of the search efforts by examiners is also based on the sample of continuation applications that were ultimately granted.

The final sample contains 98,483 continuation applications examined by a total of 7,858 individual examiners and filed over the 2001-2017 period. Ta-

[11] First action allowances can occur simply because the claims were fine as proposed by the applicant. But for two equally patentable applications, if one received rejections and the second did not, it reveals that the examiner of the second application exerted lower efforts towards narrowing down this application's claims.

[12] For a number of applications this measure is missing because the original data source used to construct the measure, the PTO's Patent Claims Research Dataset, ends in 2014, whereas my sample period ends in 2017. This explains the lower sample sizes used when estimating the effects on relatedness on these outcomes.

ble 2.1 gives summary statistics. Note that the sample sizes vary depending on the outcome considered. As about 80% of all continuation applications are eventually granted, the sample of granted continuations contains 78,141 observations. The incidence of relatedness is pretty high, with 84% of all continuations being examined by a related examiner, reflecting the established practice at the PTO of assigning continuations to the same examiner who examined the parent application. About 15% of granted continuations correspond to first action allowances, i.e. are granted with no rejections from their examiners. The average granted continuation application incurs an addition of 33 words to its shortest independent claim during its examination. Finally, examiners account for 37% of all citations made to prior patents and cite about three and a half prior patents.

Table 2.1: Summary Statistics

Variable	Mean	Standad Dev.	min	max
Related	0.84	0.37	0	1
Granted	0.79	0.40	0	1
First action allowance	0.15	0.35	0	1
Difference in word length of shortest independent claim	33.1	63.9	-2062	942
Share of examiner patent cites	0.37	0.37	0	1
Number of examiner patent cites	3.46	5.78	0	389

Notes: Statistics for Related and Granted are for the full sample of continuation applications (N=98,483). Statistics for first action allowances are for the sample of continuation applications that were granted (N=78,141). Statistics for the difference in word-length of the shortest independent claim is for the sample of continuations that were granted and for which claims data was available in the Patent Claims Research Dataset (N=45,293). Statistics for the share and count of examiner patent cites are for the sample of continuations that were granted and for which citation data was available in the PatentsView data (N=74,592).

2.4 Empirical Model

To investigate the effects of relatedness on examination practices, I estimate the following baseline specification:

$$y_{ait} = \alpha + \beta \times r_{ai} + \theta_i + \lambda_t + x'_{ait}\gamma + \epsilon_{ait}, \tag{2.1}$$

where a, i and t index continuation applications, examiners and years of disposal, respectively.[13] The dummy $r_{ai} = \{0, 1\}$ indicates whether examiner i and continuation application a are related. y_{ait} is either one of the five outcome variables that are considered in the analysis (the grant decision, the incidence of first action allowances, the difference in the word-length of the shortest independent claim between the granted and the published versions of the application, the number of examiner cites to patent prior art and the share of examiner cites to patent prior art). Examiner fixed effects θ_i and year-of-disposal fixed effects λ_t are included in all specifications. Additional controls are included in x_{ait} in a number of alternative specifications, to be discussed below, and model errors are denoted ϵ_{ait}.

The inclusion of examiner fixed effects is motivated by existing empirical work that shows that examination practices vary greatly across examiners. Lemley and Sampat (2012), for instance, show that examiner fixed effects explain a large share of the variation in examiner citations to patent prior art. Whereas a model with only art unit fixed effects explains about 20% of this variation, a model that, in addition to art unit effects, includes also examiner fixed effects, can explain 70% of this variation. They find similar results when looking at citations to non-patent prior art. Earlier work by Cockburn et al. (2002) had already revealed substantial heterogeneity across examiners in their leniency during the examination process. Including examiner fixed effects allows me to control for such unobserved examiner heterogeneity when studying the effect of relatedness on examination practices.

With examiner fixed effects, the coefficient of interest in the main specification, β, is identified using variation in relatedness within examiners. Although examiners typically examine related continuations, occasionally they are assigned to unrelated continuations, generating the within examiners variation

[13] The word "disposal" is commonly used to refer to the grant or abandon of an application. The year of disposal is thus the year of either the abandon or grant of a given application.

needed to estimate β. As examiners and applicants do not have any control over the assignment of applications, shocks to $Related$ are most likely driven by the notably high turn-over of examiners at the PTO or transfers of examiners from one art unit to another.[14]

The examiner fixed effects control for all time-invariant unobserved examiner-specific heterogeneity that might be correlated with relatedness. For instance, more lenient examiners might be assigned to continuations more often if they examine and dispose examinations faster. This could result in a higher number of assignments of unrelated continuations to lenient examiners, which would bias β in a simple OLS regression without examiner fixed effects.

Since identification is obtained by comparing the examination outcomes of related and unrelated continuation applications within examiners, it hinges on examiners who eventually are assigned to both related and unrelated continuation applications. These "switchers" account for approximately 75% of the 7,858 individual examiners in my sample.[15] Thus, although a minority of continuations are examined by unrelated examiners, only about one examiner in four in my sample never comes across an unrelated continuation application during the sample time period.

One potential threat to identification is that within-examiner shocks to relatedness could be correlated with time-varying examiner characteristics. Because the odds of being assigned to an unrelated application increases with an examiner's tenure, examiner experience might confound the results if it is left uncontrolled for.[16] Unfortunately, I do not observe experience directly for the examiners in my sample. One possible proxy variable for experience is examiner spell length, defined as the number of years separating the year of disposal of a given sample continuation and the first year of disposal observed in my sample for the examiner assigned to that continuation. For instance, if a continuation is granted in 2015 by an examiner who, within my sample, first

[14]The random assignment of applications within art units was first documented by Lemley and Sampat (2012) who conducted interviews with a number of examiners at the PTO.

[15]The term switchers should not be seen as meaning that examiners get to choose which applications they examine. Application assignment is determined by their supervisor and is out of their control, as clarified earlier. The term switchers simply refer to examiners who have examined both unrelated and related continuation applications in my sample.

[16]Lemley and Sampat (2012) show that examiners with more experience are more lenient.

disposed a continuation in 2012, experience equals three. But given that examiner fixed effects are included, this proxy variable would be perfectly collinear with year-of-disposal effects. In other words, the inclusion of year-of-disposal effects already accounts for some of the effects of experience, along with all unobserved temporal shocks that are common to all examiners.

I nonetheless consider an additional specification that includes, on top of year-of-disposal effects, a set of experience dummies specified into two-year blocks, in the spirit Frakes and Wasserman (2017b).[17] As these dummies are no longer perfectly collinear with year-of-disposal effects within examiners, they allow the effects of experience to be estimated separately. Finally, I also confirm in a robustness check in Section 2.5.1 that the main results are unchanged if I consider the subset of my sample for which examiner experience can be measured directly using the examiner rosters data used by Frakes and Wasserman (2017b).

A remaining concern is the potential sorting of continuations within examiners. If related and unrelated continuations differ in systematic ways, then β might capture the effect of these intrinsic differences on examination outcomes, rather than the true effect of relatedness itself on examination practices. Since applicants and examiners have no control over the assignment of continuations, it is unlikely that related continuations are systematically more or less patentable. Nonetheless, to mitigate potential non-random sorting, in a separate specification I add a set of detailed continuation application characteristics that are fixed at the time of publication to control for differences between applications.[18]

Finally, to account for possible serial correlation in model errors within examiners, I cluster my standard errors at the examiner level in all my regressions.

[17]The dummies are defined by group of experience levels between 0-1 year, 2-3 years, 4-5 years, 6-7 years and for 8 years and more.

[18]The set of application controls include: the size of the applicant (small or large), the total number and total word-length of dependent claims, the total number and total word-length of independent claims, the length of the shortest independent claim, the average word-length among dependent claims and the average word-length among independent claims.

2.5 Results

In this section, I first present estimates of the effects of relatedness on examination practices. I then evaluate the wearing down hypothesis. Next, I study variation in the effect of relatedness across various technology areas. Finally, I investigate whether relatedness results in over-granting of continuation applications.

2.5.1 The Effects of Relatedness on Examination Practices

For each of my outcome variables, I estimate the following four versions of specification 2.1. The first only includes examiner and year-of-disposal fixed effects. The second adds art unit fixed effects to control for any potential heterogeneity across art units that might be correlated with relatedness. The third adds a set of dummies for examiner experience to control for possible experience effects over time and within examiners that might correlate with relatedness. Finally, the fourth version adds a set of application controls to account for potential non-random sorting of continuation applications. While the fourth specification offers the advantage of having more controls, the third specification is based on a larger sample. I focus my discussions of the results on these two last specifications.

Table 2.2 presents the results for the relationship between relatedness and examiners' grant decisions. Related continuations are more likely to be granted. In other words, examiners are more likely to grant continuations of parent applications that they have themselves examined. The coefficient for $Related$ is positive, strongly significant and ranges between 4.8 and 5.8 percentage points across the four specifications. The estimates from my two preferred specifications, in columns (3) and (4), both imply that relatedness makes an allowance about 6% more likely relative to the sample mean of the dependent variable.

In Table 2.3, I investigate the relationship between relatedness and examiners' claims narrowing efforts. Panel A. presents the results when these efforts are measured by the incidence of first action allowances for granted applications. First action allowances are more likely for related continuations. That is, examiners tend to allow related continuations without first rejecting them more than unrelated applications. The estimates across the four specifications are all positive, strongly significant and vary between 3.9 and 4.8 percentage points. The estimates from my preferred specifications, in columns (3) and

Table 2.2: Relationship Between the Grant Decision and Relatedness

	(1)	(2)	(3)	(4)
Dependent Variable	Granted			
Related	0.058***	0.055***	0.051***	0.048***
	(0.004)	(0.004)	(0.004)	(0.004)
Mean of Dependent Variable	0.793	0.793	0.793	0.767
Observations	98483	98483	98483	73964
Examiner and Year Fixed Effects	Yes	Yes	Yes	Yes
Art Unit Fixed Effects	No	Yes	Yes	Yes
Experience Dummies	No	No	Yes	Yes
Application Controls	No	No	No	Yes

Notes: Standard errors reported in parentheses are clustered at the examiner level. The dependent variable, *Granted*, is a binary variable indicating if the given application was granted by the examiner or not. *Related* equals 1 if the examiner also examined the given application's parent application and 0 otherwise. Experience dummies are dummies for the number of years an examiner has been active in the sample period, specified into two-year blocks. Application controls include a number of application metrics at the time of publication (see footnote 18 for the list of variables included as application controls). Significance levels: *p<0.1; **p<0.05; ***p<0.01.

(4), both imply that relatedness makes a first action allowance about 30% more likely relative to the sample mean of the dependent variable. Panel B. presents the results when claims narrowing is measured by the difference in word-length of the shortest independent claim. The mean of the dependent variable is positive across the various specifications and reveal that about 33 words are added to the shortest independent claim for the average continuation. Fewer words are, however, being added to related applications, as the point estimate for *Related* is negative in all four specifications. The estimates are strongly significant and range between -8.871 and -7.553. For my preferred specifications, in columns (3) and (4), the point estimates imply that relatedness decreases the extent of claims narrowing by about 25% relative to the sample mean of the dependent variable when it is measured by the word-length difference of the shortest independent claim.

Lastly, I investigate the relationship between the intensity of examiners' prior art searches and relatedness. Table 2.4 presents the results for examiner citations to prior patents. In Panel A. of Table 2.4, the dependent variable is the number of examiner citations made to prior patents. In Panel B., the dependent variable is the share of examiner citations made to prior patents out of all citations (i.e. those made by both the applicant and the examiner). Examiners cite less prior art and account for a lower share of all citations for related continuation applications. The point estimates in Panel A. are all strongly significant and range between -1.079 and -1.066. For my two preferred specifications, in columns (3) and (4), the point estimates imply that relatedness reduces the number of examiner citations made to patent prior art by about 30% relative to the sample mean of the dependent variable. The point estimates in Panel B. are also all strongly significant and range between -0.034 and -0.029. The point estimates in columns (3) and (4) imply that relatedness decreases the share of examiner citations among all citations by about 10% relative to the sample mean of the dependent variable. Finally, Table 2.7 in Appendix 2.C shows that similar conclusions are reached when examiner citations to the non-patent literature, such as scientific publications, are considered instead of citations to prior patents.

To sum up, the results presented in this section show that related continuations are more likely to be granted, are more likely to receive first action allowances, have fewer words added to their shortest independent claims and contain fewer and lower shares of examiner citations to patent prior art.[19]

2.5.2 Testing the Wearing Down Hypothesis

The findings paint a consistent picture of the effects of relatedness, which appear to result in more favourable and more lenient examination practices. At first sight, these results suggest that a wearing down effect might be at play. To test the wearing down hypothesis, I compare the effects of relatedness for continuations that emanate from granted versus abandoned parent applications. If

[19]An additional robustness check, presented in Table 2.8, shows that the results are similar if a more precise measure of examiner experience, based on the examiner rosters data used by Frakes and Wasserman (2017b), is considered instead of within-sample examiner length spells.

Table 2.3: Relationship Between Examiners' Claims Narrowing Efforts and Relatedness

	(1)	(2)	(3)	(4)
Panel A.: **First Action Allowance**				
Related	0.046***	0.047***	0.048***	0.039***
	(0.004)	(0.004)	(0.004)	(0.004)
Mean of Dependent Variable	0.148	0.148	0.148	0.130
Observations	78141	78141	78141	56732
Panel B.: **Difference in Word-Length of Shortest Independent Claim**				
Related	-8.792***	-8.871***	-8.822***	-7.553***
	(0.948)	(0.950)	(0.952)	(0.848)
Mean of Dependent Variable	33.14	33.14	33.14	33.79
Observations	45293	45293	45293	42697
Examiner and Year Fixed Effects	Yes	Yes	Yes	Yes
Art Unit Fixed Effects	No	Yes	Yes	Yes
Experience Dummies	No	No	Yes	Yes
Application Controls	No	No	No	Yes

Notes: Standard errors reported in parentheses are clustered at the examiner level. In Panel A. the dependent variable is a binary variable indicating a first action allowance (the given application was granted by the examiner without first being rejected). In Panel B., the dependent variable is the difference in the word-length of the shortest independent claim in a given application between its published and issued versions. Related is 1 if the examiner also examined the given application's parent application and 0 otherwise. Experience dummies are dummies for the number of years an examiner has been active in the sample period, specified into two-year blocks. Application controls include a number of application metrics at the time of publication (see footnote 18 for the list of variables included as application controls). Significance levels: *p<0.1; **p<0.05; ***p<0.01.

Table 2.4: Relationship Between Examiners' Prior Art Search Efforts and Relatedness

	(1)	(2)	(3)	(4)
***Panel A.:* Number of Examiner Citations to Patents**				
Related	-1.073***	-1.079***	-1.070***	-1.066***
	(0.065)	(0.065)	(0.065)	(0.075)
Mean of Dependent Variable	3.465	3.465	3.465	3.603
Observations	74592	74592	74592	54708
***Panel B.:* Share of Examiner Citations to Patents**				
Related	-0.030***	-0.030***	-0.029***	-0.034***
	(0.004)	(0.004)	(0.004)	(0.005)
Mean of Dependent Variable	0.368	0.368	0.368	0.363
Observations	74592	74592	74592	54708
Examiner and Year Fixed Effects	Yes	Yes	Yes	Yes
Art Unit Fixed Effects	No	Yes	Yes	Yes
Experience Dummies	No	No	Yes	Yes
Application Controls	No	No	No	Yes

Notes: Standard errors reported in parentheses are clustered at the examiner level. Panel A. dependent variable is the number of citations to existing patents in a given continuation application. Panel B. dependent variable is the share of all patent citations made by the examiner for a given continuation. Related is 1 if the examiner also examined the given application's parent application and 0 otherwise. Experience dummies are dummies for the number of years an examiner is active in the sample period, specified into two-year blocks. Application controls include a number of application metrics at the time of publication (see footnote 18 for the list of variables included as application controls). Significance levels are defined as: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

a wearing down effect exists, one channel through which it should definitely operate are repeat filings. Suppose that an applicant wishes to obtain a given set of claims, but the examiner will not allow those claims. The applicant can then file for the same patent application as a continuation and abandon the existing application, which will restart the prosecution of the given patent. The wearing down hypothesis predicts that a related examiner is more likely to grant the continuation application than an unrelated examiner, for whom no wearing down can be at play.

I thus define a dummy that equals 1 if a continuation's parent application was granted and 0 otherwise and I then run fixed effects regressions that allow the impact of relatedness to vary depending on whether or not a continuation's parent application was granted. If the wearing down of examiners is driving the effect of relatedness, this effect should be particularly pronounced for the set of continuations that emanate from abandoned parent applications, which is the set of continuations for which the wearing down effect is more likely to be at play.

Table 2.5 presents the results of this test. In columns (1) and (2), split sample regressions of the specification that includes examiner, year-of-disposal and art unit fixed effects along with experience dummies show that relatedness has a strong negative effect on the likelihood of grant when the parent application has not been granted and a strong positive effect when the parent application was granted. Column (3) reports the results of a regression that uses the full sample and allows the effect of relatedness to vary depending on the grant status of the parent applications by including interaction terms. The results are similar to those obtained in the split sample analysis and show that, whereas relatedness increases the likelihood that a continuation is granted by about 6 percentage points when its parent application was granted, for continuations of parent applications that were not granted the likelihood of grant decreases by 4 percentage points. A Wald test for the equality of the two coefficients strongly rejects their equality. Table 2.9 in Appendix 2.C shows that the inclusion of application controls slightly attenuates the estimates but does not alter the conclusions.

This test presumes that the wearing down effect operates through all continuations emanating from abandoned parent applications and operates mostly through that channel. There are two limitations to this approach. First, some continuations that emanate from abandoned parent applications might have

Table 2.5: Test of the Wearing Down Hypothesis Based on the Grant Status of Parent Applications

	(1)	(2)	(3)
Dependent Variable		Granted	
Sample	Parent Not Granted	Parent Granted	Full
Related	-0.043***	0.065***	
	(0.013)	(0.004)	
Related × Parent Not Granted			-0.041***
			(0.010)
Related × Parent Granted			0.061***
			(0.004)
Equality test: p-value			$< .001$
Examiner and Year Fixed Effects	Yes	Yes	Yes
Art Unit Fixed Effects	Yes	Yes	Yes
Experience Dummies	Yes	Yes	Yes
Application Controls	No	No	No
Observations	13471	85012	98483

Notes: Standard errors reported in parentheses are clustered at the examiner level. The dependent variable is the grant decision, a binary variable indicating if the given application was granted by the examiner or not. Related is 1 if the examiner also examined the given application's parent application and 0 otherwise. Experience dummies are dummies for the number of years an examiner is active in the sample period, specified into two-year blocks. Column (3) also controls for the direct effect of the grant status of the parent application. The p-value of the equality test is obtained from testing equality between Related × Parent Not Granted and Related × Parent Granted with a Wald test. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

been filed for a different purpose than wearing examiners down. A more appropriate test would isolate the "true" repeat filings, i.e. those continuations that emanate from rejected parents and that are identical or nearly identical to their parent applications. Second, there might be other channels through which the wearing down of examiners takes place than repeat filings. In particular, applicants might settle over fewer, narrower, claims that the examiner is willing to allow in an initial application and subsequently file for a continuation to obtain the remaining, broader, claims. In Appendix 2.B, I provide additional tests that address, as much as possible, these limitations by isolating better the channels through which the wearing down of examiners might occur. The results from these additional tests are also hard to reconcile with the wearing down hypothesis.

In conclusion, I do not find empirical support for a wearing down effect. It appears that relatedness increases the odds of an allowance only for the continuations of granted parent applications. For abandoned parents' continuations, relatedness makes an allowance less likely. One possible explanation for these findings is an anchoring effect. Examiners might form opinions on the validity of applications the first time they assess them and then carry these opinions over to the continuations without considering or searching for new information, such as newly discovered prior art or claim amendments by applicants.

2.5.3 Variation Across Technology Areas

Previous studies have shown that inventors do not rely on continuation applications in the same way across different technology areas (Hegde et al., 2009; Frakes and Wasserman, 2015a). Strategic filings of continuation applications might then be more prevalent in some areas than in others and, as a result, the effects of relatedness could vary across technology fields. I assess this potential variation by splitting my sample according to the six NBER categories defined in Hall et al. (2001) and then running fixed effects regressions for my five outcome variables over the six subsamples. The NBER categories are the following: (1) Chemical, (2) Computers and Communications, (3) Drugs and Medical, (4) Electrical and Electronic, (5) Mechanical and (6) Others.[20]

Figure 2.1 illustrates the results from the split sample analysis. Each plot shows the point estimate of *Related* for each of the five outcome variables and the relevant subsample along with 95% confidence intervals. The main conclusion emerging from these plots is that the effects of relatedness do not vary much across the six technology areas. Not only do the differences between the point estimates tend to be minor, but most confidence intervals actually

[20] I run these regressions with examiner, year-of-disposal, art unit and examiner experience effects. I do not include application controls as this would reduce sample size substantially and thus the precision of the estimates over the subsamples.

overlap with the point estimates for the full sample. There are a few notable differences, though. First, relatedness affects the grant decision of examiners less in the Electrical and Electronic technology area, with a point estimate for the coefficient of $Related$ of around half the size of the full sample estimate. Second, the effect of relatedness on the incidence of first action allowances is smaller in the Chemical category, by about 75% relative to the estimate for the full sample. Finally, relatedness reduces the number of examiner cites to patent prior art more in the Mechanical technology field and less in the Computers and Communications field, relative to the full sample.

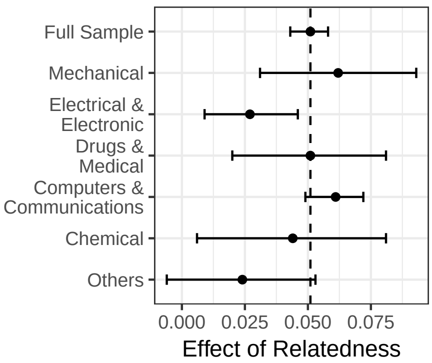
Overall, these findings show that, although there are a few noticeable differences in the effects of relatedness on examination practices across technology areas, no clear pattern emerges from the data.

2.5.4 Over-granting

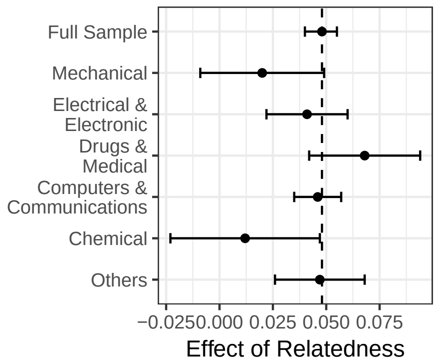
Relatedness is associated with more lenient and more favourable examination practices for continuation applications. However, this does not necessarily imply that it results in the granting of invalid patents. Examiners could be making the right decisions when facing related applications, despite putting less effort in searching for prior art and narrowing down these applications' claims. Indeed, an examiner might be better able to assess the validity of a related continuation application because she understands the underlying technology better the second time she comes across it, relative to a newly assigned examiner.

To understand whether relatedness results in over-granting or not, we need a benchmark for comparison. I follow the strategy used by Lemley and Sampat (2012), Lei and Wright (2017) and Frakes and Wasserman (2017b) in using outcomes at the EPO and the JPO as benchmarks. Applicants often seek protection for the same invention at different patent offices. So-called "triadic" patents are sets of patents filed at the PTO, the EPO and the JPO. While continuations filed at the PTO may be examined by related and unrelated examiners, their counterparts filed at the EPO and JPO will necessarily be reviewed be a new, unrelated, examiner. The EPO and JPO have similar patentability standards as the PTO, but spend more resources on patent examination (Picard and van Pottelsberghe de la Potterie, 2013; Lei and Wright, 2017), so we can use the outcomes at the two foreign patent offices to judge the validity of the continuations issued by examiners at the PTO. I thus consider the subset of continuations in my sample that were granted by the PTO and for which

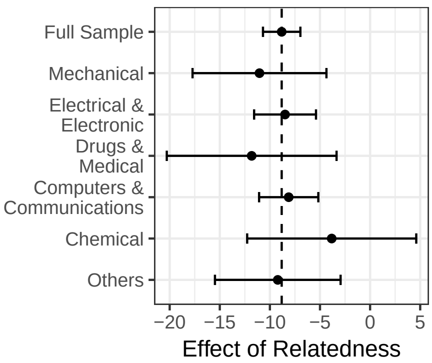
Figure 2.1: Variation in the Effects of Relatedness Across Technology Areas



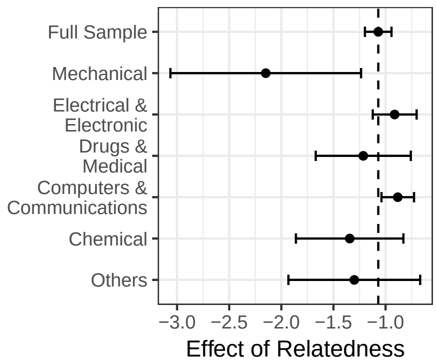
(a) Grant Decision



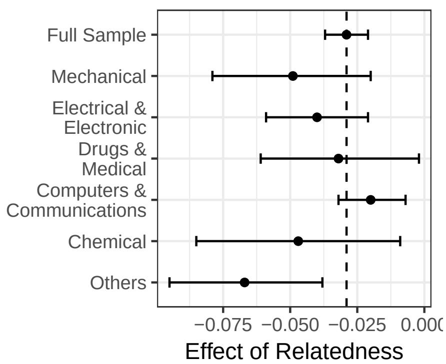
(b) First Action Allowance



(c) Difference in Word-Length of Shortest Independent Claim



(d) Number of Examiner Cites to Prior Patents



(e) Share of Examiner Cites to Prior Patents

Notes: This figure shows how the effect of relatedness on my five outcome variables varies across the six NBER categories defined in Hall et al. (2001). Each plot shows the point estimates from fixed effect regressions taken over the relevant subsample and for the relevant outcome variable, along with 95% confidence intervals. All regressions include examiner, year-of-disposal, art unit and examiner experience effects. The vertical dashed lines mark the point estimates of the effect of relatedness coming from the regressions over the full sample.

a corresponding application was filed at the EPO and the JPO. I then investigate the relationship between relatedness and the likelihood of grant at those foreign patent offices.

Table 2.6 presents the results of fixed effects regressions with three different outcome variables. The first variable is the grant of the corresponding application at both the EPO and the JPO. The second variable is the grant of the corresponding application at the EPO, but not at the JPO. The third variable is the grant of the corresponding application at the JPO, but not at the EPO. Continuations that were granted by related examiners are less likely to be granted by the two foreign patent offices. The negative point estimates in the three columns of Table 2.6 consistently show that relatedness makes an allowance in both or one of the two offices less likely by about 4-5 percentage points. Focusing on column (1), one can see that continuations granted by related examiners at the PTO were 5.1 percentage points less likely to be granted at both the EPO and JPO, equivalent to a drop of about 14% relative to the sample mean of the dependent variable.

The finding that relatedness results in patents that are less likely to be granted by foreign patent offices, taken together with the fact that relatedness results in a larger number of granted patents, less claims narrowing and shallower search efforts by examiners together suggest that relatedness leads to over-granting of continuation applications.

2.6 Conclusion

The PTO has been accused for a number of years of granting invalid patents, which tax consumers and put innovation at risk. Although there are ex-post means of checking the validity of patents, scholars believe that the root causes of the problem remain unaddressed. The key culprit is the examination process itself, which might bias examiners towards granting rather than rejecting patents if their incentives are inadequate.

In this paper, I have studied the relationship between patent office features and examination practices in the context of continuation applications. My analysis was focused on relatedness — the idea that continuation applications are typically assigned to the same examiner as the one who examined their parent applications — and its effects on examiner behaviour. My empirical analysis reveals that examiners are more likely to grant related continuations, they nar-

Table 2.6: Relationship Between Incidence of Grant at the EPO and JPO for Triadic Patents Granted at the PTO and Relatedness

Dependent Variable:	(1) Granted at both EPO and JPO	(2) Granted at EPO	(3) Granted at JPO
Related	-0.051**	-0.042*	-0.045**
	(0.021)	(0.022)	(0.021)
Mean of Dependent Variable	0.362	0.483	0.642
Observations	7540	7540	7540
Examiner and Year Fixed Effects	Yes	Yes	Yes
Art Unit Fixed Effects	Yes	Yes	Yes
Experience Dummies	Yes	Yes	Yes
Application Controls	Yes	Yes	Yes

Notes: Standard errors reported in parentheses are clustered at the examiner level. In column (1), the dependent variable is the incidence of grant at both the EPO and the JPO. In column (2), the dependent variable is the incidence of grant at the EPO. In column (3), the dependent variable is the incidence of grant at the JPO. Related is 1 if the examiner also examined the given application's parent application and 0 otherwise. Experience dummies are dummies for the number of years an examiner is active in the sample period, specified into two-year blocks. Application controls include a number of application metrics at the time of publication (see footnote 18 for the list of variables included as application controls). Significance levels are defined as: *p<0.1; **p<0.05; ***p<0.01.

row down the scope of related continuations less (by allowing them on first action more often and by adding fewer words to their shortest independent claims) and they search less for prior art when examining related continuation applications (as measured by their citations to prior art). Moreover, continuation applications granted by related examiners appear to be of more dubious validity. These findings together suggest that relatedness leads to softer examination practices and the granting of invalid patents.

The role played by continuation applications in the problem of patent invalidity has been emphasised for over a decade. In 2007, the PTO attempted to address this issue by limiting the amount of continuations that an applicant could file. Later, the PTO was ordered by the District Court for the Eastern District of Virginia to abandon this initiative. My findings suggest that, even if a single continuation could be filed, the continuation system might still fail at weeding out bad patents. An appropriate policy reform should tackle the issue at its root, which relates more to examiners' incentives than applicant behaviour. In the context of continuation applications, having an unrelated examiner review continuation applications, or a "second pair of eyes" checking the granting decisions of related examiners appear to be more adequate solutions.

Appendix

2.A Construction of the Dataset

This section provides more detail on the construction of the dataset used in the empirical analysis of the paper.

- Step 1: Creating the continuation families.

  To construct the continuation families, I use three data sources from the PTO's Patent Examination Research Dataset: the application data file, the child's continuity data file and the parent's continuity data file. From the child's continuity file, I keep only continuations, and therefore discard all divisionals and continuations-in-parts. From the application data file I keep only regular non-pending utility applications filed on and after 2001. I first identify all applications that have no parents, or "first time applications." Among these, I identify the subset that have continuations by looking through the child's continuity file. I then merge these continuations with the parents to form a dataset of continuation families, where a given family contains a parent and all its children (continuations).

- Step 2: Cleaning the continuation family dataset.

  The following cleaning operations were applied to the family dataset to prepare it for analysis:

  - A small number of duplicate entries of children were removed.
  - A small number of applications for which information on the examiner who was assigned to the application was missing were removed.

- Step 3: Focus on families with one parent and one continuation.

  I focus on the families that have only two members: one parent and one child. These families represent about 80% of all families. This results in a sample of about 100,000 continuation applications and the same number of parent applications.

- Step 4: Preparing the data for an examiner fixed effects specification.

  My baseline empirical specification is a regression with examiner fixed effects, where the observations are continuation applications. Thus, I first prepare a dataset of all continuation applications, making sure to keep certain pieces of information on their parent applications (e.g. whether they were granted or not). A small number of continuations (about 1%) are examined by examiners that do not examine any other application in my dataset. These continuations are thus not informative in my fixed effects regression and are discarded.

- Step 5: Creating experience variables.

  The analysis sometimes uses examiner experience as a control variable. I describe now how experience was defined.

  - Experience: I define examiner experience as examiner spell length within my sample. Thus, for each application, examiner experience is the number of years an examiner has been active within my sample period at the time the given application is disposed (that is, granted or abandoned).
  - Alternative measure of experience: For a subset of sample continuations, I am able to match the examiners in my sample to the examiner rosters data collected by Frakes and Wasserman (2017b). The rosters data contain examiners' years of entry at the PTO. I construct the alternative experience measure by taking the difference in years between the year of disposal of a given continuation application and the year of entry at the PTO of the examiner assigned to that given continuation.

- Step 6: Matching with rejection data to construct first action allowance outcome variable.

  To construct the first action allowance variable, I use the transactions file from the PTO Patent Examination Research Dataset. Using this file, I can compute, for all applications filed at the PTO, the number of rejections during the examination process. The absence of any rejection for a granted application is then recorded as a first action allowance in my dataset.

- Step 7: Matching with patent and publication claims dataset to construct the claims narrowing and the similarity measures.

To compute the total number of words added to the shortest independent claim between the published and granted versions of continuation applications, I use the patent and publication documents statistics datasets in the Patent Claims Research Dataset released by the PTO. To construct the measure of similarity between a continuation and its parent's claims (see Appendix 2.B for the use of this measure), I use a Jaccard index method, which compares the similarity of any two vectors. The vectors being compared in my case are 4-tuples, containing the total number of dependent claims, the total number of independent claims, the total number of words contained in all dependent claims and the total number of words contained in all independent claims.

- Step 8: Matching with patent citations data to construct the examiner citations measures.

  To compute the number of citations made by examiners to patent prior art during the examination and the share examiners account for in the total number of citations made to patent prior art, I use the patent citations data file available on PatentsView (May 28, 2018 version). In this file, every citation made within a given patent to prior patents can be traced back to either the examiner or the applicant. It is then straightforward to compute the desired measures for patent prior art using this file. For citations made to non-patent prior art, I rely on the data collected by Frakes and Wasserman (2017b). This data also contains information on the origin of the citations (applicants or examiners) and can therefore be used to construct the desired variables.

- Step 9: Matching with triadic patent family data

  When assessing whether there is over-granting of continuations at the PTO, I consider the set of continuation applications that were issued at the PTO, and also filed for patent protection at both the EPO and the JPO. I then look at the outcomes, whether the applications were granted or not, at those two foreign patent offices. To obtain this information, I follow Frakes and Wasserman (2017b) and use the OECD Triadic Patent Family database. The code to construct the Triadic Patent Families is borrowed from Frakes and Wasserman (2017b) (their replication code is available on Dataverse, see Frakes and Wasserman (2015b)) and the resulting data is then matched to my sample of continuation applications.

2.B Additional Tests of the Wearing Down Hypothesis

I conduct two additional tests of the wearing down hypothesis that try to capture more realistically the channels through which the wearing down of examienrs is likely to occur. I focus on two potential channels. First, wearing down is more likely to be at play for continuations that emanate from rejected parents and that are very similar to their parents, as these types of continuations are more likely to be repeat filings. Second, applicants may decide to settle over fewer, narrower, claims that the examiner is willing to allow in an initial application and subsequently file for a continuation to obtain the remaining, broader, claims. This second channel therefore operates via granted parent applications that have experienced high levels of claims narrowing.

The first test is based on a measure of similarity between a continuation application and its parent. The measure is constructed by comparing the total number of independent and dependent claims, as well as the total number of words contained in them, across parents and their continuations. The similarity measure is then computed as the Jaccard similarity index (for real-valued vectors), where any given vector contains the number of dependent claims, the number of independent claims, the total number of words contained in the dependent claims and the total number of words contained in the independent claims. I then define a dummy, *Similar*, that equals 1 if the continuation emanated from a rejected parent application and scores a similarity index larger than the median for all continuations that emanate from rejected parent applications. Finally, I run regressions that allow the effect of relatedness to vary depending on whether continuations emanate from a similar rejected parent or not.

The second test is based on a measure of claims narrowing incurred by parent applications. For each continuation that emanates from a granted parent application, I measure the extent of claims narrowing incurred in the parent application by computing the difference in the word-length of the parent's shortest independent claim. I then flag continuations that emanate from parent applications that incurred above-median narrowing and those that emanate, instead, from below-median parents. Finally, I run regressions allowing the effect of relatedness to vary depending on the extent of claims narrowing incurred by the parent applications.

Table 2.10 presents the results of these two additional tests. Columns (1) and (2) show the results of the first test without and with application controls. The estimated coefficients in these two columns on the interaction terms shows that continuation applications that emanate from rejected parents and are very similar to those parent applications are less likely to be granted by related examiners, whereas relatedness has a positive impact on the likelihood of grant for other applications (those either having dissimilar rejected parents or granted parents). In both columns, equality tests strongly reject the hypothesis that the two interaction terms have equal coefficients. Columns (3) and (4) present the results from the second test. The estimated effects seem unaffected by the extent of narrowing incurred by the parent applications. Equality tests show that, indeed, one cannot reject the null of equality of the two coefficients.

Overall, these two tests suggest that even when we pin down the channels through which wearing down is more likely to occur, the data does not lend support to the wearing down hypothesis.

2.C Additional Tables

Table 2.7: Relationship Between Examiners' Prior Art Search Efforts and Relatedness (Non-Patent Literature)

	(1)	(2)	(3)	(4)
***Panel A.:* Number of Examiner Cites to Non-Patent Literature**				
Related	-0.060***	-0.060***	-0.059***	-0.059***
	(0.013)	(0.013)	(0.013)	(0.013)
Mean of Dependent Variable	0.178	0.178	0.178	0.172
Observations	39913	39913	39913	34567
***Panel B.:* Share of Examiner Cites to Non-Patent Literature**				
Related	-0.027***	-0.027***	-0.026***	-0.027***
	(0.005)	(0.005)	(0.005)	(0.006)
Mean of Dependent Variable	0.076	0.076	0.076	0.074
Observations	21784	21784	21784	18965
Examiner and Year Fixed Effects	Yes	Yes	Yes	Yes
Art Unit Fixed Effects	No	Yes	Yes	Yes
Experience Dummies	No	No	Yes	Yes
Application Controls	No	No	No	Yes

Notes: Standard errors reported in parentheses are clustered at the examiner level. Panel A. dependent variable is the number of citations to non-patent publications in a given continuation application. Panel B. dependent variable is the share of all non-patent citations made by the examiner for a given continuation. Related is 1 if the examiner also examined the given application's parent application and 0 otherwise. Experience dummies are dummies for the number of years an examiner is active in the sample period, specified into two-year blocks. Application controls include a number of application metrics at the time of publication (see footnote 18 for the list of variables included as application controls). Significance levels are defined as: *p<0.1; **p<0.05; ***p<0.01.

Table 2.8: Robustness of the Main Results to an Alternative Measure of Examiner Experience

	(1)	(2)
Panel A.: **Granted**		
Related	0.042***	0.039***
	(0.005)	(0.005)
Observations	73564	56856
Panel B.: **First Action Allowance**		
Related	0.036***	0.031***
	(0.004)	(0.005)
Observations	58799	44097
Panel C.: **Difference in Word-Length of Shortest Independent Claim**		
Related	-7.464***	-6.667***
	(1.025)	(0.899)
Observations	37782	35522
Panel D.: **Number of Examiner Cites to Patents**		
Related	-1.080***	-1.091***
	(0.077)	(0.090)
Observations	56584	42730
Panel E.: **Share of Examiner Cites to Patents**		
Related	-0.031***	-0.035***
	(0.005)	(0.005)
Observations	56584	42730
Examiner and Year Fixed Effects	Yes	Yes
Art Unit Fixed Effects	Yes	Yes
Experience Dummies	Yes	Yes
Application Controls	No	Yes

Notes: Standard errors reported in parentheses are clustered at the examiner level. Experience dummies are dummies for the number of years of experience of an examiner when examining any given continuation application, based on their date of entry at the PTO and specified into two-year blocks (using the Frakes and Wasserman (2017b) rosters data). Significance levels are defined as: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table 2.9: Test of the Wearing Down Hypothesis Based on the Grant Status of Parent Applications — Robustness to the Inclusion of Application Controls

	(1)	(2)	(3)
Dependent Variable		Granted	
Sample	Parent Not Granted	Parent Granted	Full
Related	-0.039***	0.061***	
	(0.014)	(0.005)	
Related × Parent Not Granted			-0.030***
			(0.011)
Related × Parent Granted			0.057***
			(0.005)
Equality test: p-value			$< .001$
Examiner and Year Fixed Effects	Yes	Yes	Yes
Art Unit Fixed Effects	Yes	Yes	Yes
Experience Dummies	Yes	Yes	Yes
Application Controls	Yes	Yes	Yes
Observations	11334	62630	73964

Notes: Standard errors reported in parentheses are clustered at the examiner level. The dependent variable is the grant decision, a binary variable indicating if the given application was granted by the examiner or not. Related is 1 if the examiner also examined the given application's parent application and 0 otherwise. Experience dummies are dummies for the number of years an examiner is active in the sample period, specified into two-year blocks. Application controls include a number of application metrics at the time of publication (see footnote 18 for the list of variables included as application controls). Columns (3) also controls for the direct effect of the grant status of the parent application (granted or not). The p-value of the equality test is obtained from testing equality between Related × Parent Not Granted and Related × Parent Granted with a Wald test. Significance levels are defined as: *p<0.1; **p<0.05; ***p<0.01.

Table 2.10: The Impact of Similarity to Rejected Parents and Extent of Claims Narrowing of Granted Parents on the Effect of Relatedness

	(1)	(2)	(3)	(4)
Dependent Variable	Granted			
Related × Similar	-0.033**	-0.030*		
	(0.016)	(0.015)		
Related × Other	0.061***	0.054***		
	(0.005)	(0.005)		
Related × Highly-Narrowed			0.064***	0.057***
			(0.007)	(0.007)
Related × Slightly-Narrowed			0.069***	0.064***
			(0.007)	(0.007)
Equality test: p-value	$< .001$	$< .001$	0.59	0.52
Examiner and Year Fixed Effects	Yes	Yes	Yes	Yes
Art Unit Fixed Effects	Yes	Yes	Yes	Yes
Experience Dummies	Yes	Yes	Yes	Yes
Application Controls	No	Yes	No	Yes
Observations	72914	68705	62232	55911

Notes: Standard errors reported in parentheses are clustered at the examiner level. The dependent variable is the grant decision, a binary variable indicating if the given application was granted by the examiner or not. Related is 1 if the examiner also examined the given application's parent application and 0 otherwise. Experience dummies are dummies for the number of years an examiner is active in the sample period, specified into two-year blocks. Similar is a dummy variable that indicates whether the given continuation application was highly similar to its parent application, where the parent was not granted. Other is a dummy that equals 1 whenever Similar equals 0. Highly- and Slightly-Narrowed are dummies that indicate whether the continuation emanates from parent applications that has, respectively, above and below median levels of claims narrowing as measured by the word-length difference in their shortest independent claims. Direct effects of the dummies are included in all specifications. The p-values of the equality tests are obtained from testing equality between Related × Similar and Related × Other, on the one hand, and Related × Highly-Narrowed and Related × Slightly-Narrowed with a Wald tests. Significance levels are defined as: *p<0.1; **p<0.05; ***p<0.01.

Bibliography

Allison, J. R., Lemley, M. A., and Walker, J. (2009). Extreme Value or Trolls on Top: The Characteristics of the Most-Litigated Patents. *University of Pennsylvannia Law Review*, 158.

Carley, M., Hegde, D., and Marco, A. (2015). What is the Probability of Receiving a U.S. Patent? *Yale Journal of Law and Technology*, 17(1).

Cockburn, I. M., Kortum, S., and Stern, S. (2002). Are All Patent Examiners Equal? The Impact of Examiner Characteristics. Working Paper 8980, National Bureau of Economic Research.

Farre-Mensa, J., Hegde, D., and Ljungqvist, A. (2017). What Is a Patent Worth? Evidence from the U.S. Patent "Lottery". USPTO Economic Working Paper 2015-5.

Frakes, M. D. and Wasserman, M. F. (2013). Does Agency Funding Affect Decisionmaking? An Empirical Assessment of the PTO's Granting Patterns. *Vanderbilt Law Review*, 66:67–147.

Frakes, M. D. and Wasserman, M. F. (2015a). Does the U.S. Patent and Trademark Office Grant Too Many Bad Patents?: Evidence from a Quasi-Experiment. *Stanford Law Review*, 67:613–676.

Frakes, M. D. and Wasserman, M. F. (2015b). Replication data for: "Is the Time Allocated to Review Patent Applications Inducing Examiners to Grant Invalid Patents?: Evidence from Micro-Level Application Data". *Harvard Dataverse*, V1.

Frakes, M. D. and Wasserman, M. F. (2016). Empirical Scholarship on the Prosecution Process at the Patent Office. In Menell, P. S., Schwartz, D. L., and Depoorter, B., editors, *Research Handbook on the Law and Economics of Intellectual Property*. Edgar Elgar.

Frakes, M. D. and Wasserman, M. F. (2017a). Decreasing the Patent Office's Incentives to Grant Invalid Patents.

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

Ganzi, G. C. (2007). Patent Continuation Practice and Public Notice: Can They Coexist? *Journal of the Patent and Trademark Office Society*, 89.

Graham, S. J., Marco, A. C., and Miller, R. (2015). The USPTO Patent Examination Research Dataset: A Window on the Process of Patent Examination. Technical report.

Hall, B. H., Jaffe, A. B., and Trajtenberg, M. (2001). The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools. NBER Working Papers 8498.

Hegde, D., Mowery, D. C., and Graham, S. J. H. (2009). Pioneering Inventors or Thicket Builders: Which U.S. Firms Use Continuations in Patenting? *Management Science*, 55(7):1214–1226.

Kuhn, J. M. and Thompson, N. (2017). How to Measure and Draw Causal Inferences with Patent Scope. *International Journal of the Economics of Business, Forthcoming*.

Lei, Z. and Wright, B.D. (2017). Why Weak Patents? Testing the Examiner Ignorance Hypothesis. *Journal of Public Economics*, 148:43–56.

Lemley, M. A. (2001). Rational Ignorance at the Patent Office. *Northwestern University Law Review*, 95(4).

Lemley, M. A. and Moore, K. A. (2004). Ending Abuse of Patent Continuations. *Boston University Law Review*, 84.

Lemley, M. A. and Sampat, B. (2012). Examiner Characteristics and Patent Office Outcomes. *The Review of Economics and Statistics*, 94(3):817–827.

Lichtman, D. (2004). Rethinking Prosecution History Estoppel. *University of Chicago Law Review*, 71:151–182.

Marco, A. C., Carley, M., Jackson, S., and Myers, A. F. (2016a). The USPTO Historical Patent Data Files: Two Centuries of Innovation. Technical report.

Marco, A. C., Sarnoff, J. D., and deGrazia, C. (2016b). Patent Claims and Patent Scope. USPTO Economic Working Paper 2016-04.

Picard, P. M. and van Pottelsberghe de la Potterie, B. (2013). Patent Office Governance and Patent Examination Quality. *Journal of Public Economics*, 104:14–25.

Rantanen, J. (2015). To File a CON? Empirical Popularity and Prosecution Outcomes. *Patentlyo*.

Sampat, B. and Williams, H. L. (2019). How Do Patents Affect Follow-On Innovation? Evidence from the Human Genome. *American Economic Review*, 109(1):203–236.

United States Patent and Trademark Office (2012). Overview of the U.S. Patent Classification System.

United States Patent and Trademark Office (2015). Patent Litigation and USPTO Trials: Implications for Patent Examination Quality.

Chapter 3

Patent Length and Innovation Incentives for Industrial Designs[1]

[1]I thank Tore Ellingsen, Richard Friberg and Erik Lindqvist for their comments and feedback. I am also grateful to participants at the Stockholm School of Economics lunch seminar and Ph.D. workshop for helpful remarks. Financial support from the Jan Wallander and Tom Hedelius Foundation is gratefully acknowledged. All errors are my own.

3.1 Introduction

Patents reward inventors by granting them exclusivity rights over their inventions during a set period of time. But, as a result of these exclusivity rights, patents also increase prices above competitive levels during their lifespan. Since at least the work of Nordhaus (1969), scholars have understood how patent length can be used to achieve an optimal patent system.[2] Longer patents incentivise the development of more inventions, but they also increase the length of time during which supra-competitive pricing is applied over all inventions. Optimal patent length simply balances these benefits and costs at the margin.[3]

A key parameter needed to determine the optimal patent length is the sensitivity of innovation to the patent term (Budish et al., 2016). If inventors do not react to increases in the patent term, longer patents increase price distortions without generating new inventions, resulting in lower social welfare. Theoretical studies typically assume that inventors readily respond to increases in the patent term. Yet, empiricists agree that there is still a dearth of empirical evidence supporting that assumption.

In this paper, I explore how a longer patent term affects innovation in the context of design patents in the U.S. Design patents protect the way inventions look and, despite being relatively cheap and easy to obtain, they can in some cases provide strong protection, as illustrated by recent litigation cases, such as the Apple versus Samsung lawsuit.

In 2015, the Geneva Act of the Hague Agreement Concerning the International Registration of Industrial Designs came into force in the U.S. This treaty, whose purpose is to facilitate the process of obtaining patent protection for designs internationally, also increased the term of protection, from fourteen to fifteen years, for all design patents filed at the United States Patent and Trademark Office (PTO) after May 13, 2015. I study the response of inventors to the term increase, focusing on the number of PTO filings as a measure of

[2]The Nordhaus model is introduced in Chapter 5 of Nordhaus (1969) and further discussed in Scherer (1972).

[3]This argument abstracts from the effect of patent protection on follow-on innovation, which affects the optimal design of patents too. See Williams (2017) for a review of the literature that explores how patents affect follow-on innovation.

innovative output.[4]

I find that the term extension increased filings by 1.5-8% depending on the specification employed, though these estimated impacts are not statistically significant. There is no evidence that the impact is unambiguously stronger in patent classes, or industries, that rely more intensively on design patents. There is also no sign of variation across different types of applicants, such as individual inventors or companies, for whom design patents might be differentially valuable. Taken together, these findings suggest that inventors do not appear to be very sensitive to the term of design patents.

This paper is most closely related to the body of empirical studies that investigate the effects of stronger patent rights on innovation incentives using patent law changes as a source of variation.[5] Sakakibara and Branstetter (2001) explore how the Japanese patent reforms of 1988, which increased the scope of patent protection, have impacted innovators by looking at firms' R&D spendings and patenting activities. They find no evidence that stronger patents lead to more innovation. Qian (2007) studies the effects of introducing patent protection for pharmaceuticals in 26 countries over the 1978-2002 period and finds that national patent protection alone does not stimulate domestic innovation. Lerner (2009) explores the impact of major patent reforms in 60 countries over a period of 150 years and finds that reforms that increased the strength of patent protection did not have any positive effects on domestic innovation either. Another finding that emerges from Qian (2007) and Lerner (2009) is that, in countries that already have a strong intellectual property regime, additional strengthening of patent protection may lead to less, not more, innovation.[6] An exception to these negative results comes from Abrams (2009), who exploits variation in patent term adjustements across different patent classes

[4]Measures of innovative input, such as innovative investments into the creation of new designs, would have been desirable as well, but unfortunately are not readily available.

[5]Another line of research has used survey evidence to explore the relationship between patents and innovation (Mansfield, 1986; Levin et al., 1987; Cohen et al., 2000). More recently, scholars have also used variation in patent length induced by differences in clinical trial length to study the relationship between innovative investments and patent length (Budish et al., 2015).

[6]This finding also emerges from the economic history literature on patents and innovation (Moser, 2016).

following the implementation of the Agreement on Trade Related Aspects of Intellectual Property Rights (TRIPS) in the U.S.[7] He finds that patent classes that benefited from longer term adjustements experienced larger increases in patent filings, which suggests a positive relationship between patent term and innovation.

The paper is structured as follows. Section 3.2 describes my empirical setting and data. Section 3.3 and Section 3.4 provide, respectively, descriptive and regression analyses of the effect of the term extension for design patents. Section 3.5 discusses extensions and robustness checks. Section 3.6 provides a discussion of the results and Section 3.7 concludes.

3.2 Empirical Setting and Data

This section gives more detail on design patents and the accession of the U.S. to the Hague Agreement. It then discusses the data used in the empirical analysis.

3.2.1 Design Patents

Design patents protect the appearance of products, which differentiates them from the more popular form of patents, utility patents, that protect the way products work. They are simpler than utility patents, as they contain only one claim, which is visually depicted as the claimed design in a series of illustrations. They are also a lot cheaper and are generally faster to obtain than utility patents. Design and utility patents can be filed simultaneously for a given invention to protect both its appearance and its workings. Design patents are valid for a period of fifteen years counted from the date of grant, unlike utility patents, which have a term of twenty years counted from the date of issue.

Despite its simplicity, the design patent can be a strong form of intellectual property. A case in point is the recent legal battle between Apple and Samsung over some of Apple's design patents. Apple owns many design patents over the

[7] Before TRIPS, the patent term in the U.S. was of 17 years from the date of grant. After TRIPS, it became 20 years from the date of issue. To compensate applicants for delays caused by the PTO during the patent examination process, which would reduce the effective patent term, the PTO applies patent term adjustments to each patent. Abrams (2009) exploits the fact that, as patent application pendency differs greatly across patent classes, TRIPS generated variation in the effective patent term across different patent classes.

iPhone and iPad and, in 2011, sued Samsung for infringement of a number of these patents. Prior to an ultimate undisclosed settlement, the case was leaning in favour of Apple, as Samsung was ordered by U.S. courts to pay Apple $539 million in damages for infringing its patents. This example shows that, when design patents cover the ornemental features of an innovative and marketable product, such as the iPhone, they can be extremely valuable to their owners.

Another aspect of design patents that make them useful to inventors is that they can cover only sub-elements of a given design. Coming back to the Apple versus Samsung case, Apple left the back of the iPhone out of one of its claimed designs, which meant that Samsung would infringe Apple's design patents if it changed just the back and not the front of its own smartphones. By only claiming key parts of a single design, inventors can effectively appropriate a large part of the product space.

3.2.2 Accession to the Hague Agreement

On December 18, 2012, the U.S. signed the Geneva Act of the Hague Agreement into law, which came into force on May 13, 2015. This treaty allows applicants to obtain international protection for their designs under a single umbrella, by applying through either the PTO or the World Intellectual Property Organisation. Prior to the accession to the Hague Agreement, U.S. inventors seeking protection for their designs internationally had to file individual applications at each of the patent offices where protection was sought. With the Hague Agreement, they can obtain protection in all other member countries with a single application, resulting in cheaper and faster obtention of global protection of industrial designs.

The term of protection of international registrations under the Hague Agreement is five years with the option to be renewed up to expiry of the maximum term of protection in the relevant countries (i.e. where protection has been obtained). The maximum term of protection in member countries must therefore be a multiple of five. As a result, the U.S. had to extend its term for design patents from fourteen to fifteen years as a side-effect from its accession to the Hague Agreement. This term extension was implemented not only for international registrations, but also for domestic filings (for which protection abroad is not necessarily sought). More specifically, any design patent filed at the PTO after May 13, 2015, enjoyed a term of fifteen instead of fourteen years. The incidental nature of this term extension, which I often refer to as the "intervention" in the remaining of this paper, makes it an ideal event to study the response of inventors to an increase in patent length (I return to my empirical strategy in Section 3.4).

3.2.3 Data

I collected data on design patents using PatentsView, an online bulk download service provided by the PTO and the Patent Examination Research Dataset, also made available by the PTO.[8] My data contains information on filing and grant dates, patent classes and applicants.

My sample contains all design patents filed at most two years prior to and at most one year following the date of the intervention, May 13, 2015. In addition to switching from a first-to-invent to a first-to-file system in early 2013, the U.S. introduced new patent fees, which had noticeable impacts on patent filings.[9] I choose to restrict my pre-intervention window to a two-year period so that my data is not directly affected by those previous changes to the patent system.

The post-intervention period is chosen to limit the influence of truncation on patent counts. In the U.S., all design patents go through a substantive examination before being granted (or not), which results in delays between the dates of filing and of grant. Because design patents are never published before being granted, many applications filed in recent years are still pending at the PTO and are therefore missing from the data (which includes only design patents that have been granted). The severity of truncation in my sample can be evaluated by looking at recent pendency levels for design patents at the PTO. For all granted design patents filed within the five years that preceded the intervention, median pendency was 483 days, minimum pendency was 60 days and maximum pendency was 2,600 days. The 928 days that separate the last day covered by the PatentsView data (November 27, 2018) and the first anniversary of the intervention (May 13, 2016) coincides with the 94th percentile of the pendency distribution. This means that, at the upper limit of my sample

[8]This paper makes use of the November 27, 2018 release of the PatentsView data and the December, 2018 release of the Patent Examination dataset.

[9]In a first-to-file system, the applicant who files an application first is considered the rightful owner of the patent rights. This is not necessarily the case under a first-to-invent system, which is based on the date of conception of the invention, rather than the date of filing of the corresponding patent.

period, at most 6% of all applications are missing if pendency is assumed to have remained constant. At the time of the intervention, this figure equals a mere 1%. Thus, while truncation is an issue in the present study, it is limited, with few missing data mostly at the end of the sample period.

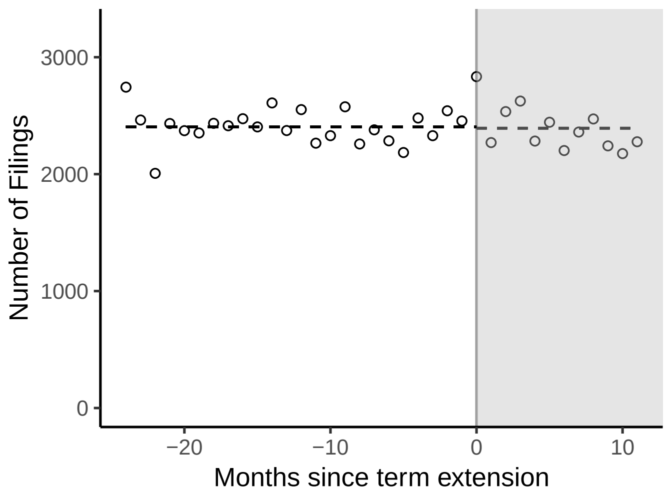
I compute counts of daily filings and aggregate them into monthly counts for my descriptive analysis in Section 3.3 and into weekly counts for my regression analysis in Section 3.4. The aggregation is performed with the date of the intervention as a starting point. In other words, all daily counts that take place in the month that immediately follows (and includes) May 13, 2015, are aggregated into one single count for month '0'. All daily counts that take place in the subsequent month are aggregated into a single count for month '1', all daily counts that took place the month that immediately precedes (and excludes) May 13, 2015, are aggregated into a single count for month '-1', and so forth until aggregate counts for the full window of 36 months are obtained. The same procedure is implemented to obtain weekly counts over the full window of 156 weeks. This procedure results in a neat separation of monthly and weekly counts during the pre- and post-intervention periods. Monthly counts are used for the descriptive analysis because daily counts are too volatile to depict insightful visual trends. Weekly counts are used for the regression analysis to avoid unnecessary volatility, such as weekend versus weekday effects, while maintaining a sample size that is large enough to achieve enough statistical precision and power. I confirm in Section 3.5 that the main results from the regression analysis hold when daily data is used instead.

3.3 Descriptive Analysis

I first explore visually the effect of the intervention on the creation of new designs by plotting design patent filings at the PTO around the date of the intervention. Figure 3.1 plots monthly filings during the three-year period considered. Filings show no sign of strong upward or downward trend and, instead, fluctuate around an average of 2,400 filings over the sample period. Filing behaviour does not seem to have been strongly affected by the patent term extension, except during the first month immediately following the intervention (month '0'), which shows a clear increase in filings. The plot also shows that the pre- and post-intervention means are almost identical.

The aggregate counts presented in Figure 3.1 could mask variation in the

Figure 3.1: Monthly Design Patent Filings Before and After the Intervention



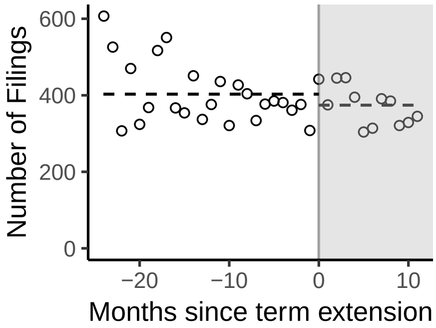
Notes: This plot shows monthly filings at the PTO before and after the term extension (which took place in month '0'). The shaded area represents the post-intervention period. The two dashed line segments show pre- and post-intervention means.

effect of the intervention. In particular, industries that rely more heavily on design patents to protect their intellectual property might be affected by a term extension more than others. Figure 3.2 explores this potential heterogeneity across the six most popular design classes in my sample. These classes are defined as the six United States Patent Classification (USPC) classes that had the largest number of filings in the five years that preceded the intervention. They accounted alone for about 45% of all filings among the 42 classes represented in the data within those five years. For these six classes, there is no visual evidence of strong effects from the intervention. The series seem to mostly follow their pre-intervention paths in the post-intervention period and a simple comparison of pre- and post-intervention means show that, in fact, the mean generally has decreased after the intervention.[10] Visual inspection of unreported plots confirms that there is no sign of strong responses from inventors to the intervention in the ten subsequent most popular classes either.

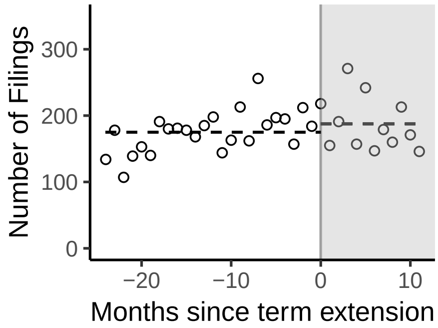
Different types of applicants might also have been affected differently by the intervention. First, a longer patent term might be more valuable to small

[10]Note that this decrease could be explained by the truncation issue discussed in Section 3.2.

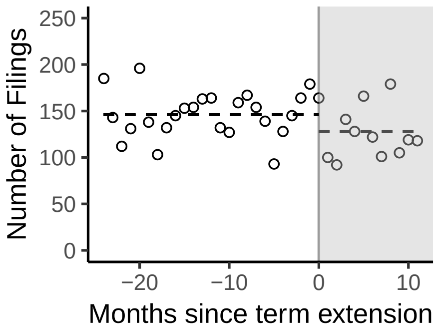
Figure 3.2: Monthly Filings Across the Six Most Popular Design Patent Classes



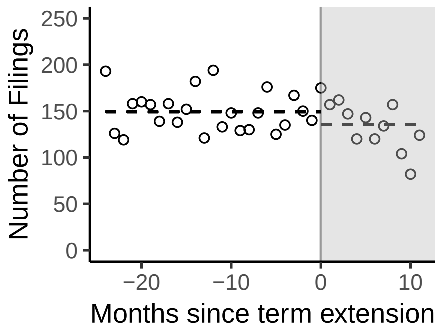
(a) Recording, communication, or information retrieval equipment (D14)



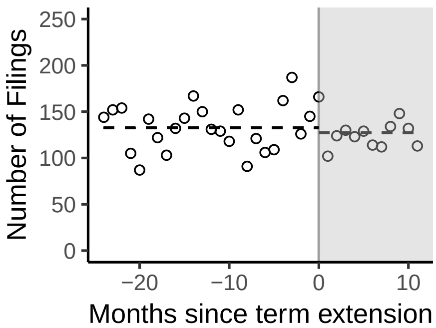
(b) Transportation (D12)



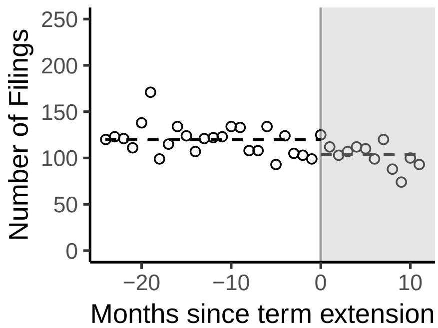
(c) Furnishings (D06)



(d) Medical and laboratory equipment (D24)



(e) Equipment for preparing or serving food or drink not elsewhere specified (D07)



(f) Packages and containers for goods (D09)

Notes: These plots show monthly filings in the six most popular USPC classes before and after the intervention (which took place in month '0'). The shaded area represents the post-intervention period. The two dashed line segments show pre- and post-intervention means.

companies and individual inventors than to large corporations who can more easily benefit from their intellectual property via other channels than design patents. I explore this potential source of variation in Figure 3.3. Plots (a) and (b) show that firms and individual inventors seem equally unaffected by the intervention. For both types of applicants, the pre- and post-intervention means are nearly identical and the series does not show any sign of strong changes around the time of the intervention. Second, foreign and domestic firms might have reacted differently to the term extension. Plots (c) and (d) show, however, that this does not seem to be the case. Although foreign firms' filings increase steadily during the first few months right after the intervention, they abruptly fall down and stay low thereafter. In fact, the post-intervention mean is even slightly smaller than the pre-intervention mean for foreign firms, but slightly larger for domestic firms. Finally, plots (e) and (f) show that applicant size, as reported to the PTO during the application process, does not matter either. Neither small nor large applicants seem to have been greatly affected by the intervention.

3.4 Regression Analysis

I now explore the effect of the intervention on the creation of new designs more formally by estimating the following baseline linear specification:

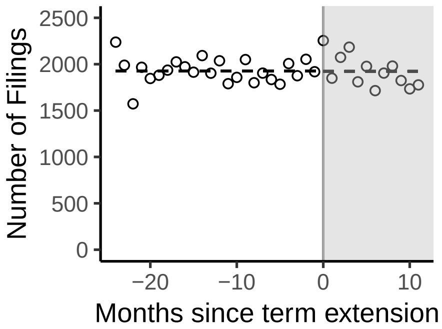
$$y_t = \alpha + \beta Post_t + \gamma t + \epsilon_t, \tag{3.1}$$

where $t = \{1, \cdots, 156\}$ indexes the given week since the start of the sample period and is included as a linear time trend in the regression, y_t is the count of design patents filed in week t, $Post_t$ is a dummy variable that indicates the pre-intervention period ($Post_t = 0$) and the post-intervention period ($Post_t = 1$) and ϵ_t denotes model errors.[11] The coefficient of interest, β, should be positive if there is an incentive effect resulting from the term extension.

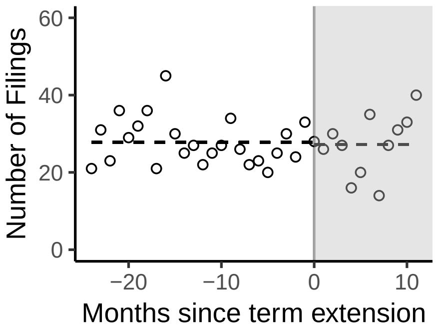
The use of a linear specification presents two potential issues. First, a linear model allows the conditional mean function to take negative values, while the outcome is a count variable, which is greater than 0 by definition. Second, it

[11]The time index starts at the beginning of the sample period, two years prior to the intervention, and daily counts are aggregated into weekly counts starting from the date of the intervention, as explained in Section 3.2.

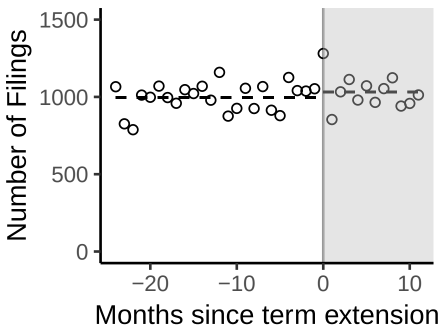
Figure 3.3: Monthly Filings Across Different Types of Applicants



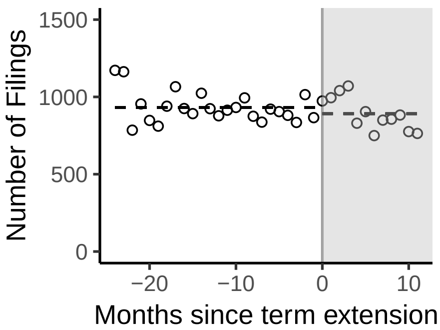
(a) U.S. and foreign firms



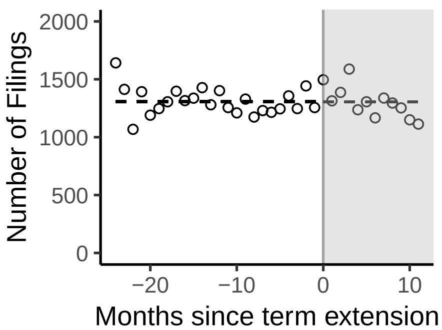
(b) U.S. and foreign individuals



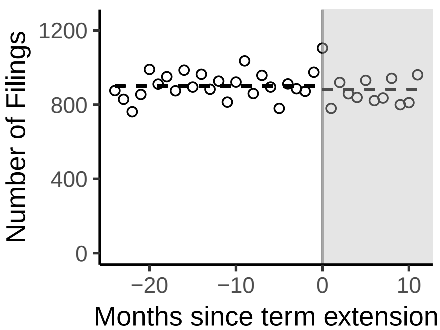
(c) U.S. firms



(d) Foreign firms



(e) Large entities



(f) Small entities

Notes: These plots show monthly filings for different types of patent applicants before and after the intervention (which took place in month '0'). Entity sizes are based on sizes reported to the PTO. Large entities consist of large firms (with more than 500 employees), while small entities consist of small firms, individual inventors and non-profit organisations, such as universities. The shaded area represents the post-intervention period. The two dashed line segments show pre- and post-intervention means.

assumes that the outcome is continuous, while counts are necessarily discrete. Though the linear model is generally a good starting point to get a sense of the relationship of interest in the data, special regression methods designed to accomodate the count nature of the data are available and should be used.[12]

The most basic regression model for count data is the Poisson regression model, which specifies the conditional distribution of the outcome variable as a Poisson distribution and an exponential conditional mean function.[13] In my context, this leads to the following non-linear specification of the conditional mean:

$$\mathbb{E}\left[y_t|Post_t, t\right] = \exp\left(\alpha + \beta Post_t + \gamma t\right), \tag{3.2}$$

whose parameters are typically estimated by maximum likelihood.

An important feature of the Poisson regression model is equidispersion, or the equality of the conditional variance and mean. When the data is not equidispersed, the standard errors of the Poisson maximum likelihood estimator are wrong, which leads to incorrect inference on the parameters. A typical solution is to consider the Poisson model with robust standard errors, known as the quasi-Poisson model, which yields corrected standard errors without having to model the dispersion in the data. Another typical solution for overdispersed data, that is when the variance is larger than the mean, is to use the Negative Binomial model, which specifies a Negative Binomial distribution for the outcome variable with conditional mean μ of the exponential form and the conditional variance equal to $\mu + \sigma\mu^2$.[14]

A simple comparison of the sample mean and variance of my outcome variable suggests that my data is overdispersed. The mean equals approximately 550 and the variance equals about $5,560$, and so is a little more than ten times larger than the mean. To accomodate for this overdispersion in my data, I present results based on the quasi-Poisson and Negative Binomial models instead of the standard Poisson model.

[12] Unless there are severe endogeneity concerns, which create a lot of complications for regressionX methods for count data.

[13] A popular alternative is to log-transform the dependent variable and then use a linear model. Given the wide availability of software for count regression models and the simplicity of the present empirical model, it is preferable to use count regression methods that take the count nature of the data into account.

[14] A formal treatment of these models can be found in Cameron and Trivedi (2013).

In addition, because the data consists of a time series, the possible presence of serial correlation in model errors may also lead to complications. If there is serial correlation, the parameters are consistently estimated but the standard errors are not. To address this issue, a quasi-Poisson model with heteroskedasticity and auto-correlation consistent standard errors may be used, but generally it is better to model the time dependency directly. I return to the issue of serial correlation in Section 3.5.

The identifying assumption that allows me to obtain unbiased estimates of β is that the pre-intervention trend can be used as a counterfactual for the post-intervention trend. In principle, there are a number of potential issues with this identification strategy. First, the existence of other policies on or around the same date as the intervention under study would confound the results. However, the last major patent reforms took place early 2013, before the sample period, and PTO sources list no other policies affecting design patents around the intervention, mitigating concerns related to confounding policies. Second, the existence of lobbies fighting for stronger intellectual property rights might reverse the direction of causality between the number of filings and policy reforms. However, since the term extension was a by-product of the U.S.'s entry into the Hague Agreement, whose purpose is to facilitate international filings rather than to strengthen patent rights in the U.S. per se, reverse causality is unlikely to be a major concern in my empirical setting. Finally, applicants might have sorted around the date of the intervention if those who had a new design ready before the intervention waited to file for their design patents to obtain the longer term. Since the U.S. patent system is a first-to-file system and designs are relatively easy and costless to imitate, waiting before filing would probably have seemed worthy only relatively close in time to the date of the intervention. To address the possibility of sorting around the date of the intervention, in Section 3.5 I present results based on a sample that excludes filings in time windows of four and eight weeks centered around the date of the intervention.

Table 3.1 presents the main results from the regression analysis. Column (1) shows the results from the linear model (3.1) with robust standard errors. Although the intervention appears to have increased the number of weekly filings by 8.5 units, this estimate is not statistically significant. Column (2) reports the results from the quasi-Poisson model. The exponentiated coefficients, shown in square brackets, suggest that the number of weekly filings post-intervention were about 1.6% larger than pre-intervention weekly filings.[15] The average marginal effect is estimated at 8.484, close to the Ordinary Least Square (OLS) estimate from the linear model in column (1). Again, however, the estimate is not statistically different from zero. The conclusions emerging from the Negative Binomial model, presented in column (3), are essentially the same as those emerging from the quasi-Poisson model. The estimated σ in the variance function is positive and strongly significant, which confirms the presence of overdispersion in the data.

Overall, these results reinforce the visual evidence presented in the previous section and show that the term extention for design patents has only had a limited impact on inventors' incentives to create new designs. In the next section, I extend the empirical model in various directions and check the robustness of these results.

3.5 Extensions and Robustness Checks

3.5.1 Cross-Industry Variation

The descriptive analysis in Section 3.3 showed that there is little variation in the effect of the intervention across the most popular patent classes. I revisit this question formally by conducting separate regression analyses for groups of patent classes with different levels of patenting intensity. It is reasonable to expect that inventors in patent classes that use design patents more intensively would react more to a term extension. To test this hypothesis, I measure a class's patenting intensity as the total number of filings in that given class during the five years that preceded the intervention and subsequently separate the counts of weekly filings into four groups based on the quartiles of the distribution of patenting intensity. I then conduct the regression analysis separately on each of the four resulting samples.

The results of this analysis are presented in Table 3.2. The intervention has had a bigger impact in the highly intensive patent classes, as we would ex-

[15] To see this, note the following relationship:

$$\frac{\mathbb{E}\left[y_t|Post_t=1,t\right]}{\mathbb{E}\left[y_t|Post_t=0,t\right]}=\exp(\beta)\frac{\mathbb{E}\left[y_t|Post_t=0,t\right]}{\mathbb{E}\left[y_t|Post_t=0,t\right]}=\exp(\beta).$$

Table 3.1: The Impact of the Patent Term Extension on the Number of Weekly Design Patent Filings

	(1)	(2)	(3)
Dependent Variable	Number of Weekly Design Patent Filings		
Model	OLS	Quasi-Poisson	Negative Binomial
t	-0.177	-0.000	-0.000
	(0.227)	(0.000)	(0.000)
		[1.000]	[1.000]
Post	8.503	0.015	0.015
	(22.832)	(0.042)	(0.039)
		[1.016]	[1.015]
Constant	562.289***	6.332***	6.332***
	(13.598)	(0.024)	(0.026)
		[562.352]	[562.450]
σ			0.016***
Observations	156	156	156

Notes: Robust standard errors reported in parentheses in columns (1) and (2) and usual standard errors in column (3). Exponentiated coefficients reported in square brackets in columns (2) and (3). In all columns, the dependent variable is the total number of design patents filed in a given week. t is a time variable that represents the given week since the start of the sample period. $Post$ is a dummy variable that equals one for all weeks including and following the intervention. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

pect. But the estimated impact declines at the next level of patenting intensity and then goes up again in less patent-intensive classes. This finding is at odds with the negative monotonic relationship we expect between patenting intensity and the magnitude of the intervention's impact. Finally, the overall lack of statistical significance once more suggest that across the various groups inventors have not been strongly affected by the intervention.

Table 3.2: The Impact of the Patent Term Extension on the Number of Weekly Design Patent Filings Across Patent Classes With Different Levels of Patenting Intensity

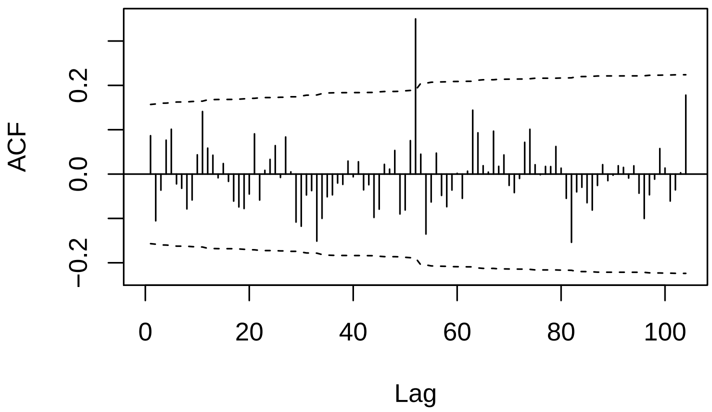
	(1)	(2)	(3)
Dependent Variable	Number of Weekly Design Patent Filings		
Model	OLS	Quasi-Poisson	Negative Binomial
High Patenting Intensity			
Post	10.835	0.082	0.081
	(8.684)	(0.067)	(0.069)
		[1.086]	[1.085]
Medium-High Patenting Intensity			
Post	-7.264	-0.063	-0.061
	(7.310)	(0.060)	(0.056)
		[0.939]	[0.941]
Medium-Low Patenting Intensity			
Post	1.201	0.007	0.007
	(8.751)	(0.059)	(0.056)
		[1.007]	[1.007]
Low Patenting Intensity			
Post	3.732	0.025	0.024
	(8.333)	(0.055)	(0.054)
		[1.025]	[1.024]
Observations	156	156	156

Notes: Robust standard errors reported in parentheses in columns (1) and (2) and usual standard errors in column (3). Exponentiated coefficients reported in square brackets in columns (2) and (3). In all columns, the dependent variable is the total number of design patents filed in a given week. All regressions include a linear time trend t and a constant term (coefficients not reported). $Post$ is a dummy variable that equals one for all weeks including and following the intervention. Significance levels: *$p<0.1$; **$p<0.05$; ***$p<0.01$.

3.5.2 Serial Correlation

I start by investigating whether serial correlation affects the main results. Serial correlation can be detected by inspecting the auto-correlation function of the regression residuals. I base my discussion on the residuals from the Negative Binomial model, as the two other models provide similar fits. Figure 3.4 plots the auto-correlation function of these residuals over 104 lags (each lag being a week). The auto-correlation is only statistically significant for one lag, the fifty-second one.

Figure 3.4: Auto-Correlation Function of the Residuals from the Negative Binomial Regression



Notes: Dashed lines represent 95% confidence intervals computed using Bartlett's formula.

To capture this yearly seasonal effect, I implement an Autoregressive Negative Binomial count model with a fifty-second order term, along with the linear time trend t and the post-intervention indicator variable $Post_t$.[16] The results, presented in Table 3.3, confirm a strong correlation between the outcome in time t and in time $t-52$, with an estimated coefficient on the fifty-second lag of 0.54 (with standard error equal to 0.09). The effect of the intervention is now estimated at about 3% in exponentiated terms, but is still not statistically significant. To check that the Autoregressive Negative Binomial model takes care

[16]This model is implemented in R using the `tscount` package described in Liboschik et al. (2017).

of the seasonality in the data, I plot the auto-correlation function of the residuals from this model in Figure 3.5, which shows no sign of serial correlation this time.

Table 3.3: The Impact of the Patent Term Extension on the Number of Weekly Design Patent Filings Using an Auto-Regressive Negative Binomial Model

Dependent Variable Model	Number of Weekly Design Patent Filings Auto-Regressive Negative Binomial
t	-0.000
	(0.000)
Post	0.028
	(0.036)
ρ_{52}	0.538
	(0.091)
Constant	2.925
	(0.227)
Observations	156

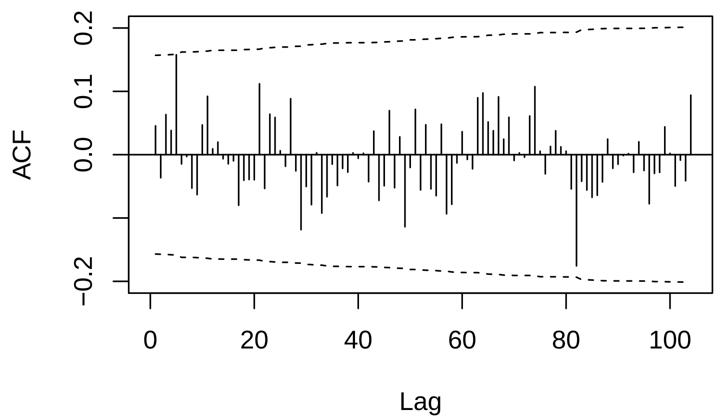
Notes: Standard errors obtained by normal approximation (via `tscount` statistical package in R). The dependent variable is the total number of design patents filed in a given week. t is a time variable that represents the given week since the start of the sample period. $Post$ is a dummy variable that equals one for all weeks including and following the intervention. ρ_{52} is a fifty-second order auto-regressive term. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

3.5.3 Polynomial Time Trend

The linear trend assumption made in Section 3.4 might be not capture accurately how filings evolve over time. As the particular choice of a linear trend can have an impact on the estimated effect of the intervention, it is a good idea to re-estimate models with polynomial time trends of various orders to check the robustness of the results based on a linear trend.

I thus run quasi-Poisson regressions containing gradually higher-order polynomials of t as regressors. Table 3.4 shows the exponentiated coefficients from polynomial regressions up to the eight order. In most cases, the estimated im-

Figure 3.5: Auto-Correlation Function of the Residuals from the Autoregressive Negative Binomial Regression



Notes: Dashed lines represent 95% confidence intervals computed using Bartlett's formula.

pact of the intervention increases but never enough to become statistically significant. In most cases, the exponentiated coefficient now lies in the 4-8% region.

3.5.4 Inner Window of Exclusion

If there was sorting of the type described in Section 3.4, with inventors waiting to file for patents for their new designs to obtain the longer patent term, the effect of the intervention would be overestimated. The spike in monthly filings that directly follows the intervention in Figure 3.1, along with the immediate return to prior levels of monthly filings thereafter, suggest that such sorting might have been at play.

A simple way to correct for this potential sorting is to exclude from the analysis all filings that took place within a window around the date of the intervention. Table 3.5 compares the results of quasi-Poisson regressions with no inner window, in column (1), and those with inner windows of four and eight weeks in, respectively, columns (2) and (3). As expected, the point estimate for the coefficient of *Post* decreases when an inner window of exclusion is considered. In particular, the point estimate is close to zero when the eight-week window is used.

Table 3.4: The Impact of the Patent Term Extension on the Number of Weekly Design Patent Filings With Various Polynomial Time Trends

Dependent Variable	Number of Weekly Design Patent Filings			
Model	Quasi-Poisson			
Polynomial Order	First	Second	Third	Fourth
Post	1.016	1.078	1.061	1.067
	(0.042)	(0.055)	(0.057)	(0.070)
Constant	562.352***	540.027***	557.289***	559.975***
	(13.729)	(16.895)	(23.169)	(31.004)
Polynomial Order	Fifth	Sixth	Seventh	Eigth
Post	1.063	0.994	1.045	1.037
	(0.073)	(0.072)	(0.079)	(0.077)
Constant	566.728***	625.945***	671.837***	642.732***
	(37.213)	(45.570)	(59.568)	(63.676)
Observations	156	156	156	156

Notes: Exponentiated coefficients reported in all columns with corresponding robust standard errors reported in parentheses. The dependent variable is the total number of design patents filed in a given week. *Post* is a dummy variable that equals one for all weeks including and following the intervention. Time variables are omitted from the table for readability. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

3.5.5 Daily Counts

Weekly counts have been used in the baseline analysis for convenience as they clean the data from unnecessary volatility. But a regression analysis based on daily data with appropriate controls for weekday effects would have worked as well. Table 3.6 presents the results of the OLS, quasi-Poisson and Negative Binomial models when daily data is used. Again, although the intervention seemed to have increased filings, the effect is not statistically significant. The exponentiated coefficients from the quasi-Poisson and Negative Binomial models are similar to those based on weekly data and suggest that post-intervention

Table 3.5: The Impact of the Patent Term Extension on the Number of Weekly Design Patent Filings With an Inner Window of Exclusion

	(1)	(2)	(3)
Dependent Variable	Number of Weekly Design Patent Filings		
Model	Quasi-Poisson		
Inner Window of Exclusion	None	4 Weeks	8 Weeks
t	1.000	1.000	1.000
	(0.000)	(0.000)	(0.000)
Post	1.016	1.012	1.001
	(0.042)	(0.044)	(0.044)
Constant	562.352***	561.872***	560.125***
	(13.729)	(14.007)	(14.238)
Observations	156	152	148

Notes: Exponentiated coefficients reported in all columns with corresponding robust standard errors reported in parentheses. The dependent variable is the total number of design patents filed in a given week. t is a time variable that represents the given week since the start of the sample period. $Post$ is a dummy variable that equals one for all weeks including and following the intervention. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

filings were 1.5-4% percent higher than pre-intervention filings.

3.6 Discussion

The results presented in Sections 3.4 and 3.5 show that the increase in the term of design patents had a positive, though not statistically significant, effect on the creation of new (patented) designs. The lack of statistical significance does not necessarily translate into effects that are not economically meaningful. Ignoring the possibility of sorting or bunching by applicants, the results imply that a one-year increase in the patent term leads to 1.5-8% more patented designs depending on the chosen specification. In the context of cancer drug development, Budish et al. (2015) estimate that a one-year increase in the patent term leads to a 7-23% increase in R&D investment. We would expect their estimates to be larger for at least two reasons. First, they focus on the phar-

Table 3.6: The Impact of the Patent Term Extension on the Number of Daily Design Patent Filings

	(1)	(2)	(3)
Dependent Variable	Number of Daily Design Patent Filings		
Model	OLS	Quasi-Poisson	Negative Binomial
t	-0.003	-0.000	-0.000
	(0.004)	(0.000)	(0.000)
		[1.000]	[1.000]
Post	1.148	0.014	0.040
	(2.784)	(0.035)	(0.045)
		[1.014]	[1.041]
Constant	124.185***	4.828***	4.840***
	(3.537)	(0.033)	(0.039)
		[124.938]	[126.509]
σ			0.13***
Observations	1082	1082	1082

Notes: Robust standard errors reported in parentheses in columns (1) and (2) and usual standard errors in column (3). Exponentiated coefficients reported in square brackets in columns (2) and (3). In all columns, the dependent variable is the total number of design patents filed on a given day. t is a time variable that represents the given day since the start of the sample period. $Post$ is a dummy variable that equals one for all days including and following the intervention. All regressions also include a full set of dummies for the different days of the week. Significance levels: *p<0.1; **p<0.05; ***p<0.01.

maceutical industry, which is highly dependent on patents. Inventors in that industry value patents greatly and are thus expected to be quite sensitive to patent length. Second, they study the sensitivity of R&D investment, a measure of innovative input, whereas I look at the sensitivity of patent filings, a measure of innovative output. As not all additional R&D translates into new patents, we should not be surprised if the latter is less sensitive to the patent term than the former.

The transition from a fourteen to a fifteen-year patent term represents a 7% increase in patent length. This implies an estimated elasticity of innovative output to the patent term equal to about 0.2 if we use the point estimates in columns (2) and (3) of Table 3.1.[17] Though we expect the term extension to increase innovative output, as longer patents confer stronger patent protection to inventors and results in the development of marginal inventions, one might wonder why it would do so less than proportionally. One possible explanation is that patented designs may in practice act as partial substitutes for each other, at least within given patent classes. Longer patents attracts new entrants and, as the number of competing designs increases, profit margins are compressed. In turn, the number of new entrants is dampened as the lower profit margins counter-act with the increase in revenues induced by the longer patent term. I illustrate this simple mechanism in the circular model of product differentiation of Salop (1979) in Appendix 3.A. With linear transportation costs and no discounting of future revenues, the model predicts an elasticity of innovation to the patent term equal to 0.5, not too far off from the elasticity implied by my estimates.[18]

Another possible explanation for the low elasticity lies in the timespan chosen for the empirical analysis. With a post-intervention period of one year, we can only hope to estimate the short-term impact of the intervention. The long-term effect could differ, as innovative investments might take time to translate into new patentable inventions. Though this issue might not be too serious in the context of design patents, for which probably less technical research is needed than for utility patents, the estimates obtained in this paper are nonetheless best understood as a lower bound for the overall effects of the intervention, which are likely to be larger.

3.7 Conclusion

In this paper, I study the relationship between patent length and innovation incentives in the context of design patents. I exploit an increase in the term of design patents that came into effect in 2015 as a result of the implementation of the Hague Agreement in the U.S. I find that the term extension increased the

[17] This figure was obtained by dividing the semi-elasticities 0.015 and 0.016 by the percent term increase 0.071.

[18] The same conceptual framework is used by Dubois et al. (2015) to explain the low elasticity of innovation to market size in the pharmaceutical sector.

number of design patent filings, though not in a statistically significant way. There is no apparent variation in this effect across different industries or types of applicants and patent classes that patent more intensively do not seem to be unambiguously more affected by the term increase than less patent-intensive classes. These results support existing empirical evidence that shows that inventors tend to not respond strongly to increases in the strength of patent protection.

Appendix

3.A Theoretical Model

I present a circular model of product differentiation based on the discussion in Tirole (1988), chapter 7, of the Salop (1979) model and show how it can generate an elasticity of innovation to the patent term of similar magnitude to the one obtained in my empirical analysis.

A unit mass of consumers is uniformly distributed around a circle of perimeter equal to 1. Innovators are located around the circle (and can locate to only one location) and consumers travel to them to purchase a single unit of their products. Each innovator sells a single patented design. The unit transportation cost incurred by consumers is equal to t and can be viewed as an index of differentiation. The higher t is, the more expensive it is for consumers to reach farther innovators, so they are more easily drawn to nearby innovators. Consumers wish to purchase at the lowest total cost, which equals the price of the purchased design plus the transportation cost. For simplicity, I assume that consumers will always obtain a sufficient surplus from the purchased designs and so will always proceed with their purchases. There is a fixed cost f incurred by the innovator for bringing a patented design to the market. The marginal cost of production of the design is assumed to be zero, for simplicity. Innovators have an outside option equal to zero if they decide not to enter the market. The length of the patent is l and innovators who have entered the market receive for the full patent term the per-period revenue p_iD_i where D_i is the demand for innovator i's design. After the expiration of the patent, imitation drives the price to marginal cost so there is no more per-period profit to be made. Assuming that innovators do not discount future revenues, the profit from bringing a new patented design to the market is given by $lp_iD_i - f$.[19]

In the first stage of the model, innovators choose to enter or not. The n entering innovators are exogenously placed at equi-distant locations, applying the principle of maximum differentiation. In the second stage, innovators choose their prices given their locations.

We start by solving the second stage for any given number of entering inno-

[19] If we assume that innovators discount future revenues, we can interpret l as discounted patent length instead of true patent length.

vators n and we assume that entrants play a symmetric equilibrium and set the same price p for their designs. Any innovator i faces two competitors, one on each of her sides. Consider a consumer located at point x between i and either one of her two surrounding competitors. This consumer will be indifferent between buying i's design and the design of the nearest competitor provided

$$p_i + tx = p + t(1/n - x),$$

where p_i is the price chosen by innovator i. Solving for x gives

$$x = \frac{p + t/n - p_i}{2t}.$$

Note that any consumer located closer to i than consumer x will chose to buy from i. Thus, the demand for i's design given her price p_i is given by

$$D_i = 2x = \frac{p + t/n - p_i}{t},$$

reflecting the fact that innovator i has consumers on either of her sides. Innovator i choses a price p_i that maximises profits, which are given by

$$lp_i \frac{p + t/n - p_i}{t} - f.$$

Solving for p_i and then setting $p_i = p$, as we are considering a symmetric equilibrium, gives the following expression for the unique equilibrium price p chosen by all innovators

$$p = t/n.$$

This expression shows that the price, which is also the profit margin in this simple model, decreases as the number of entrants gets larger. This is because a larger number of innovators share a constant mass of consumers, which puts a downward pressure on the price as competition between patented designs increases.

With the equilibrium price determined for every possible number of entrants n, we can solve the entry game that occurs in the first stage of the model. Free entry is assumed, which means that entry occurs until the profits are equal to zero. Thus, the equilibrium number of entering innovators solves

$$lpD - f = 0.$$

Substituting p and D with the expressions we derived above and solving for n gives

$$n = \sqrt{\frac{lt}{f}}.$$

As expected, a longer patent term will increase the equilibrium number of entering innovators and of patented designs. Taking the derivative of n with respect to l gives

$$\frac{dn}{dl} = \frac{1}{2}\sqrt{\frac{l}{fl}} = \frac{n}{2l},$$

which implies that the elasticity of the number of innovators, or equivalently patented designs, to the patent term equals 0.5. If prices were regulated, such that $p = \bar{p}$, the elasticity of innovation to the patent term would instead be equal to 1, as the dampening effect of entry occurs via changes to innovators' profit margins.

Bibliography

Abrams, D. S. (2009). Did TRIPS Spur Innovation? An Empirical Analysis of Patent Duration and Incentives to Innovate. *Faculty Scholarship at Penn Law*, 274.

Budish, E., Roin, B. N., and Williams, H. (2015). Do Firms Underinvest in Long-Term Research? Evidence from Cancer Clinical Trials. *American Economic Review*, 105(7).

Budish, E., Roin, B. N., and Williams, H. (2016). Patents and Research Investments: Assessing the Empirical Evidence. *American Economic Review P&P*, 106(5):183–87.

Cameron, A. C. and Trivedi, P. K. (2013). *Regression Analysis of Count Data*. Cambridge University Press.

Cohen, W. M., Nelson, R., and Walsh, J. P. (2000). Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not). NBER Working Paper 7552.

Dubois, P., de Mouzon, O., Scott-Morton, F., and Seabright, P. (2015). Market Size and Pharmaceutical Innovation. *The RAND Journal of Economics*, 46(4):844–871.

Gallini, N. T. (2002). The Economics of Patents: Lessons from Recent U.S. Patent Reform. *Journal of Economic Perspectives*, 16(2):131–154.

Graham, S. J., Marco, A. C., and Miller, R. (2015). The USPTO Patent Examination Research Dataset: A Window on the Process of Patent Examination. Technical report.

Lerner, J. (2002). Patent Protection and Innovation Over 150 Years. NBER Working Paper 8977.

Lerner, J. (2009). The Empirical Impact of Intellectual Property Rights on Innovation: Puzzles and Clues. *American Economic Review P&P*, 99(2):343–348.

Levin, R. C., Klevorick, A. K., Nelson, R., and Winter, S. (1987). Appropriating the Returns from Industrial Research and Development. *Brookings Papers on Economic Activity*, 18(3):783–832.

Liboschik, T., Fokianos, K., and Fried, R. (2017). tscount: An R Package for Analysis of Count Time Series Following Generalized Linear Models. *Journal of Statistical Software*, 82(5):1–51.

Mansfield, E. (1986). Patents and Innovation: An Empirical Study. *Management Science*, 32(2).

Moser, P. (2013). Patents and Innovation: Evidence from Economic History. *Journal of Economic Perspectives*, 27(1):23–44.

Moser, P. (2016). Patents and Innovation in Economic History. *Annual Review of Economics*, 8(1):241–258.

Nordhaus, W. D. (1969). *Invention, Growth, and Welfare: A Theoretical Treatment of Technological Change*, chapter 5. The MIT Press.

Qian, Y. (2007). Do National Patent Laws Stimulate Domestic Innovation in a Global Patenting Environment? A Cross-Country Analysis of Pharmaceutical Patent Protection, 19782002. *The Review of Economics and Statistics*, 89(3):436–453.

Rockett, K. (2010). Property Rights and Invention. In *Handbook of The Economics of Innovation, Vol. 1*, pages 315–380. Elsevier.

Sakakibara, M. and Branstetter, L. (2001). Do Stronger Patents Induce More Innovation? Evidence from the 1988 Japanese Patent Law Reforms. *RAND Journal of Economics*, 32(1):77–100.

Salop, S. C. (1979). Monopolistic Competition with Outside Goods. *The Bell Journal of Economics*, 10(1):141–156.

Scherer, F. M. (1972). Nordhaus' Theory of Optimal Patent Life: A Geometric Reinterpretation. *American Economic Review*, 62(3):422–427.

Scotchmer, S. (2004). *Innovation and Incentives*. MIT Press.

Tirole, J. (1988). *The Theory of Industrial Organization*, chapter 7. The MIT Press.

United States Patent and Trademark Office (2015). United States Deposits Instrument of Ratification to Geneva Act of the Hague Agreement Concerning the International Registration of Industrial Designs. //www.uspto.gov/about-us/news-updates/united-states-deposits-instrument-ratification-geneva-act-hague-agreement.

United States Patent and Trademark Office (2019). Hague Agreement Concerning the International Registration of Industrial Designs. https://www.uspto.gov/patent/initiatives/hague-agreement-concerning-international-registration-industrial-designs. Accessed: February 27, 2019.

Williams, H. L. (2017). How Do Patents Affect Research Investments? *Annual Review of Economics*, 9(1):441–469.