Thèse n° 10813

Essays on the Economics of Innovation: Incentives, Diffusion, and Disparity

Présentée le 30 août 2021

Collège du management de la technologie
Chaire en politiques d'innovation et de propriété intellectuelle
Programme doctoral en management de la technologie

pour l'obtention du grade de Docteur ès Sciences

par

Ling ZHOU

Acceptée sur proposition du jury

Prof. L. Lambertini, présidente du jury
Prof. G. J. A. de Rassenfosse, directeur de thèse
Prof. P. Gaulé, rapporteur
Prof. M. Mohnen, rapporteuse
Prof. D. Foray, rapporteur

École
polytechnique
fédérale
de Lausanne

2021

Acknowledgements

Completing the Ph.D. has been an enriching journey filled with the joys of learning and discovering and a heart of resilience. I would like to take this opportunity to thank many people without whose encouragement and support the writing of this thesis would have been impossible.

I would first like to express my deepest gratitude to my thesis advisor Gaétan de Rassenfosse for giving me the opportunity to learn from him and for believing in me. I consider myself lucky to have Gaétan as an advisor. He opened the door to the intriguing world of innovation and IP research for me and has always been very supportive when I needed help. Not only does Gaétan provide invaluable guidance on research ideas and methodologies, but he also allows me great freedom to explore my research interests.

I am thankful to the members of my dissertation committee Dominique Foray, Patrick Gaulé, Luisa Lambertini, and Myra Mohnen for their commitment and insightful comments. I wouldn't have had the opportunity to improve my thesis without their constructive feedback. My sincere thanks must go to the professors I met along the journey. In particular, I thank Dominique Foray for his advice and support especially towards the end of my thesis. His sense of humor during those CEMI-IIPP retreats in the Swiss mountains brought cheer to my Ph.D. life. I also thank Christopher Tucci for his encouragement during the job market season. His enthusiasm and erudition have always been an inspiration. Finally, I thank Deepak Hegde for hosting me at the seminars at NYU Stern and for giving helpful suggestions on the second chapter of my thesis.

I thank my coauthor Mary Kaltenberg for her willingness to test new ideas. Her inputs to the fourth chapter became an important part of writing up this thesis and thus deserve special credit.

I would also like to thank my colleagues and friends from the College of Management of

Technology and the Gerzensee cohort. I thank Omar Ballester and Charles Ayoubi for their support during my Ph.D. as well as the hand-me-down Gerzensee study materials. I thank Kyle Higham for all the interesting discussions on the patent system. I thank Gabriele Cristelli and Gabriele Pellegrino for our time spent co-organizing the EVIS seminars. I thank Matthias van den Heuvel, George Abi Younes, Orion Penner, David Portabella, and Emilio Raiteri for their help and valuable discussions in the past years. Special thanks to Ilona Ball for her heart-warming kindness and timely assistance. I also want to extend my thanks to Stefano Barufaldi, Annamaria Conti, Markus Simeth, and Fabiana Visentin. In addition, I am thankful to my brilliant office mates Farnaz Eslamishoar, Maximilian Hofer, Yara Kayyali, Raphael Klein, Rachel Lacroix, Philipp Schneider, Paul van Baal, Andrey Vasilyev, Andrea Xu, and everyone else with whom I shared many fun events and lively conversations during our lunches. I also thank Alice Antunes, Federica Braccioli, and Laura Nowzohour for their inspirational initiative on Women in Econ Léman. In particular, I feel lucky to have met Siyu Chen, Cho Ki Kwok, Qinli Lu, Ye Sun, Zilu Tian, Yu Wu, Kunpeng Yao, and Yihui Zuo during these wonderful years in Switzerland, and I cherish our memories dearly.

I especially thank Lipeng Chen, Xinyuan Liu, and Miaohan Yu. Their precious friendship plays an important role in my life.

Last but not least, my most special appreciation goes to my mom Yingjie Yang and dad Cunxu Zhou. I am forever grateful to my parents for selflessly encouraging me to pursue new directions and for always being there for me no matter what happens. Their unconditional love is what steers me through the difficult times in life.

Lausanne, 06 August, 2021 Ling Zhou

Abstract

As an important policy instrument for guiding innovation, the patent system involves a long-standing tension between creating economic rewards for the original innovators and stifling subsequent R&D activities. The availability of new data allows us to provide unique sights into understanding how effective patents are in providing incentives for innovation and how firms' disclosure of proprietary assets affects the dissemination of knowledge. On the other hand, innovation per se is the fruit of human intellect. The knowledge workforce is, therefore, a crucial component of a country's innovative capacity. The persistent gap in the participation in innovative activities between men and women reflects a misallocation of talented human capital and leaves opportunities for policy implementation. This dissertation ties the questions mentioned above and presents novel evidence on the incentives, diffusion, and disparity in innovative activities.

The first essay (chapter 2), in collaboration with Gaétan de Rassenfosse, provides empirical evidence on the real effect of patent expiration on monopoly prices and explores how this effect varies with product market competition. We compile a novel dataset that links patents to consumer products and retail prices for these products from Amazon.com. We find that patent expiration leads to a 7–8 percent drop in product prices. This effect is heterogeneous by patent type and importance and starts one year in advance. We also find a more substantial decline in prices in product markets where competition is intense, consistent with evidence from the pharmaceutical industry. We argue that incumbent innovators lower prices preemptively in the face of generic competition.

The second essay (chapter 3), in collaboration with Gaétan de Rassenfosse, studies how firms' disclosure of innovative assets affects the diffusion of inventions. We focus the empirical setting on firms' provision of constructive notice through virtual patent marking (VPM) and exploit patent-level variation in the marking dates. With data collected on marked patents from the VPM web documents of 16 firms, we find an overall small effect of VPM. However, conditional firms' strong reliance on patents as an appropriability regime, our findings suggest that VPM fends off diffusion to external firms and reduces similarity between inventions.

Last but not least, in the third essay (chapter 4), collaborated with Mary Kaltenberg, we explore the role of family policies in closing the innovation gender gap. By matching inventor data to web-scraped inventor age information, we exploit the staggered passage of maternity leave policies across the U.S. and evaluate its impact on the retention and productivity of female inventors. Our findings suggest that maternity leave policy supports women of reproductive ages stay longer in patenting but has little impact on their productivity.

Key words: The patent system, incentives, products, disclosure, invention diffusion, innovation gender gap, innovation policy

Résumé

En tant qu'instrument politique important pour guider l'innovation, le système des brevets implique une tension de longue date entre la création de récompenses économiques pour les innovateurs originaux et l'étouffement des activités de R&D ultérieures. La disponibilité de nouvelles données nous permet de fournir des vues uniques pour comprendre l'efficacité des brevets pour inciter à l'innovation et comment la divulgation par les entreprises des actifs exclusifs affecte la diffusion des connaissances. D'autre part, l'innovation est essentiellement le fruit de l'intellect humain. La main-d'œuvre du savoir est donc une composante cruciale de la capacité d'innovation d'un pays. L'écart persistant dans la participation aux activités innovantes entre les hommes et les femmes reflète une mauvaise affectation du capital humain talentueux et ouvre des opportunités pour la mise en œuvre de politiques publiques. Cette thèse relie les questions mentionnées ci-dessus et présente de nouvelles preuves sur les incitations, la diffusion et la disparité dans les activités innovantes.

Le premier essai (chapitre 2), en collaboration avec Gaétan de Rassenfosse, fournit des preuves empiriques de l'effet réel de l'expiration des brevets sur les prix de monopole et explore comment cet effet varie avec la concurrence sur les marchés de produits. Nous compilons un nouvel ensemble de données qui relie les brevets aux produits de consommation et aux prix de détail de ces produits sur Amazon.com. Nous constatons que l'expiration des brevets entraîne une baisse de 7 à 8% des prix des produits. Cet effet est hétérogène par type de brevet et par importance et se manifeste un an à l'avance. Nous constatons également une baisse plus importante des prix sur les marchés de produits où la concurrence est plus intense, conformément aux observations dans l'industrie pharmaceutique. Nous soutenons que les innovateurs en place baissent les prix de manière préventive face à la concurrence des génériques.

Le deuxième essai (chapitre 3), en collaboration avec Gaétan de Rassenfosse, étudie comment la divulgation d'actifs innovants par les entreprises affecte la diffusion des inventions. Nous concentrons le cadre empirique sur la fourniture de "public notice" par les entreprises par le biais du marquage virtuel des brevets (VPM) et exploitons la variation au niveau des brevets des dates de marquage. Avec les données collectées sur les brevets marqués à partir des

documents Web VPM de 16 entreprises, nous constatons que l'effet global du VPM est faible. Cependant, en raison de la forte dépendance des entreprises conditionnelles aux brevets en tant que régime d'appropriation, nos résultats suggèrent le VPM empêche la diffusion vers des entreprises externes et réduit la similitude entre les inventions.

Enfin, dans le troisième essai (chapitre 4), en collaboration avec Mary Kaltenberg, nous explorons le rôle des politiques familiales dans la réduction de l'écart entre les sexes en matière d'innovation. En associant les données des inventeurs aux informations sur l'âge des inventeurs extraites sur le Web, nous exploitons le passage échelonné des politiques de congé de maternité aux États-Unis et évaluons son impact sur la rétention et la productivité des inventrices. Nos résultats suggèrent que la politique de congé de maternité aide les femmes en âge de procréer à rester plus longtemps dans le brevet mais a peu d'impact sur leur productivité.

Mots clefs : Le système des brevets, les incitations, les produits, la divulgation, la diffusion des inventions, l'écart entre les sexes en matière d'innovation, la politique d'innovation

Contents

Acknowledgements **i**

Abstract (English/Français) **iii**

List of figures **xi**

List of tables **xiii**

1 Introduction **1**

Introduction **1**

1.1 Motivation . . . 1
1.2 The effectiveness of the patent system . . . 1
1.3 Talents and innovation . . . 3
1.4 Contribution to literature . . . 4

2 Patents and Supra-Competitive Prices: Evidence from Consumer Goods **7**

2.1 Introduction . . . 8
2.2 Background . . . 9
2.2.1 The monopoly pricing hypothesis of patents . . . 9
2.2.2 Patent protection and product price . . . 10
2.2.3 The case of the pharmaceutical industry . . . 12
2.3 Empirical approach . . . 13
2.3.1 Identification strategy . . . 13
2.3.2 Econometric model . . . 14
2.4 Data and variable construction . . . 15
2.4.1 Data on product-patent links . . . 16
2.4.2 Data on products . . . 18
2.4.3 Data on patents . . . 22
2.5 Econometric results . . . 26
2.5.1 The effect of patent expiry on product prices . . . 26
2.5.2 Accounting for product market competition . . . 34

2.5.3 Placebo tests 37
2.6 Concluding remarks 40

3 Public Notice and Invention Diffusion **43**
3.1 Introduction 44
3.2 Background 45
3.2.1 Virtual patent marking 45
3.2.2 Marking and follow-on inventions 46
3.3 Data and descriptive statistics 49
3.3.1 Patent marks 49
3.3.2 Measuring follow-on inventions 51
3.3.3 Indicators of patent importance 53
3.3.4 Descriptive statistics 54
3.4 Empirical approach 55
3.4.1 Identification strategy 55
3.4.2 Regression model 56
3.5 Baseline regression results 57
3.5.1 Marking and the strength of appropriability regime 57
3.5.2 Marking and invention importance 61
3.6 Conclusion 62

4 Does Motherhood Hold Back 'Marie Curies'? **65**
4.1 Introduction 66
4.2 Literature Review 68
4.3 Maternity leave policy and the labor market 69
4.3.1 Maternity leave policy in the U.S. 69
4.3.2 Fertility, motherhood, and labor market outcomes 71
4.4 Data and Empirical Strategy 72
4.4.1 Data 72
4.4.2 Empirical Strategy 73
4.4.3 Descriptive Statistics 77
4.5 Results 79
4.5.1 Inventors' Probability to Exit Patenting 79
4.5.2 Inventors' Productivity in Patenting 81
4.6 Conclusions 83

5 Conclusion **85**
Appendices 86

A Appendix **87**
A Appendix for chapter 1 88
B Appendix for chapter 2 91
C Appendix for chapter 3 92

Bibliography **106**

Curriculum Vitae **107**

List of Figures

2.1 Schematic representation of the measured effect . . . 14
2.2 Distributions of log of average product prices . . . 20
2.3 Products having been released vs. products tracked by *Keepa* . . . 21
2.4 Distribution of lapsed and expired patents . . . 24
2.5 Density of priority-to-release gap (left) and longest remaining patent life (right) 25
2.6 Evolution of product prices around patent expiry . . . 34
2.7 Non-monotonic effects on price of the number of substitutes (a) and competitors (b). . . . 37
2.8 Kernel density of $\hat{\beta}$'s estimated from placebo tests . . . 39

3.1 An example of a VPM document from Dyson . . . 46
3.2 Distribution of the timelag between the product launch date and the virtual marking date . . . 50
3.3 Proportion of virtually marked patents by assignee and appropriability regime 51
3.4 Average number of citations received by marked and not marked patents, by appropriability regime . . . 53
3.5 Average citation count before and after marking by appropriability . . . 53

4.1 Implementation of Maternity Leave Policy across the US . . . 70
4.2 Representation of Female Inventors and Patenting Activities in the US from 1970 to 2015 . . . 74
4.3 Distribution on the Exit Age for Women and Men Before vs. After Policy . . . 77
4.4 Productivity by Gender Before vs. After Policy . . . 78
4.5 Exit Decision around Maternity Leave Policy Enactment by Gender for Different Age Groups . . . 80
4.6 Kaplan-Meier Survival Estimates on Probability of Continuing Patenting . . . 81
4.7 Productivity around Maternity Leave Policy Enactment by Gender for Different Age Groups . . . 82

C.1 Timing of policy enactment and innovation capacity by state in 1992 . . . 92
C.2 share of women inventors by age group . . . 93

List of Tables

2.1 Summary of firms, products, and patents by product category 17
2.2 Patent and product intensity 18
2.3 Summary statistics for product-level variables 22
2.4 Summary statistics for patent-level variables 26
2.5 The effect of patent expiry on product prices, baseline specifications 27
2.6 The effect of patent expiry on product prices, alternative price variable 28
2.7 The effect of patent expiry on product prices by quartile of product prices . . . 29
2.8 The effect of patent expiry on product prices by patent type 30
2.9 The effect of patent expiry on product prices by patent importance 32
2.10 Product-level regression on the effect of patent expiry 33
2.11 Patent expiry and competition 36
2.12 Placebo test: the effect of patent expiry 40

3.1 Descriptive Statistics 55
3.2 The effect of patent marking on follow-on inventions 58
3.3 The moderating effect of appropriability regime on follow-on inventions 59
3.4 The effect of patent marking on follow-on inventions from new firms 60
3.5 The effect of patent marking on invention similarity 61
3.6 The moderating effect of patent quality on patent marking 62

4.1 Pairwise Correlation between Policy Enactment and Innovation Capacity 71
4.2 Descriptive Statistics 78

A.1 A list of representative product(s) by Amazon product catalog 88
A.2 Distribution of the sources of P^A and P^L at relevant periods, in percent 90
B.1 First stage result of PCA 91
C.1 Event study regressions on inventor retention 93
C.2 Event study regressions on inventor productivity 94

1 Introduction

1.1 Motivation

Over the past three decades, economists have established that technological progress, which at its core is the production and diffusion of knowledge, drives rapid economic growth in developed countries (Romer, 1990; Aghion & Howitt, 1992; Grossman & Helpman, 1994). Although it is desirable for private firms to invest in research and development (R&D), knowledge is essentially a public good that generates positive externalities. The know-how created by an innovative firm may flow across boundaries to other firms in the absence of protection, leaving the original innovator not sufficiently compensated of the costs of R&D. This market failure usually leads to a sub-optimal provision of new technologies to society (S. Martin & Scott, 2000; Veugelers et al., 2008).

Governments have long sought to provide incentives for stimulating R&D activities with a toolkit of policy interventions (Nelson & Langlois, 1983; B. Hall & Van Reenen, 2000; Gallini & Scotchmer, 2002; Bloom et al., 2019). These policy tools include procurement for R&D projects of defense and public interests, provision of grant fundings for basic research, tax credits targeting R&D spending, intellectual property rights (IPR), and policies to increase high-skilled human capital. This dissertation adds to the discussions on promoting innovation by concentrating chapter 2 and chapter 3 on the effectiveness of the patent system and chapter 4 on the inclusive participation in inventive activities.

1.2 The effectiveness of the patent system

The origin of modern patent systems dates back to Renaissance Venice, where the 1474 Venetian Patent Act explicitly offered exclusive rights to artisans and craftsmen on their ingenious inventions for a limited period (Nard & Morriss, 2006). Nowadays, patents still function similarly by conferring temporary monopoly rights over inventions for their holders to exclude others from commercializing the legally protected inventions. The theoretical justification for patents lies in the monetary rewards provided by the monopoly over inventions that motivate inventors to invest in research and develop new technologies (Arrow, 1962).

The economic rationale behind the patent system is to balance a trade-off between creating incentives for investment in new technologies and limiting the use of patented inventions (Nordhaus, 1969). Despite the social benefits of the patent system on R&D investments and technological change (Park & Ginarte, 1997; Moser, 2005), contentions regarding the benefits and costs of the patents abound. A somewhat controversial response from Boldrin and Levine (2013) argues in favor of abolishing the patent system, stating that patents play a minor role in fostering ground-breaking innovation but solicit rent-seeking. Hence, the first two essays evaluate two separate issues concerning the effectiveness of the patent system to advance our understanding of the core functionalities on how it affects innovation.

The first issue centers on the incentive mechanism whereby patents guarantee an imperfect level of appropriability owing to the intangible nature of knowledge. As Arrow (1962, p. 615) put it, "no amount of legal protection can make a thoroughly appropriable commodity of something so intangible as information." The existing evidence from the Yale Survey and the Carnegie Mellon Survey suggests that the effectiveness of patents is confined to a handful of industries such as pharmaceuticals and biotechnology (Levin et al., 1987; Cohen et al., 2000). A critical effort to understand how the incentive mechanism works is to estimate the private return to patenting; for example, Arora et al. (2008) find that an increase in the value of patent protection is positively associated with R&D spending. However, the appropriability of patents on general consumer goods remains unclear and has been empirically challenging to test due to the limitation of data. For this reason, chapter 2 aims to fill this gap by estimating the effect of losing patent protection on the prices of innovative consumer goods.

The second issue revolves around the disclosure of patents and knowledge spillovers. Inventors are required to disclose technical information upon applying for patents in return for obtaining exclusivity rights over the inventions. Thus, patents can be regarded as a medium that helps transforming tacit knowledge into codified information that facilitates the utilization of knowledge by subsequent inventors (Cowan & Foray, 1997; Cowan et al., 2000). Economic theories and empirical evidence suggest that disclosure has a positive impact on cumulative innovation and knowledge spillovers (Scotchmer, 1991; Furman et al., 2018); however, some scholars argue that the information revealed in a patent document may not disseminate sufficient technical details other than delineating the legal scope of an invention. Kitch (1977, p. 287) noted that "The purpose of the description in the patent is not to disclose the commercially relevant technology," which implies the limitations of patented information. Recent reforms in the patent system open the door to study an alternative form of disclosure that firms implement to signal the commercial values of their inventions, a notable example including the introduction of virtual patent marking. Chapter 3 examines heterogeneous effects of firms' disclosure of innovative portfolios through virtual patent marking on the diffusion of innovation.

1.3 Talents and innovation

New knowledge is created from the accumulative intellect of inventors and scientists who engage in R&D activities (Schmookler, 1957; Stephan, 2012; Toivanen & Väänänen, 2016). From a supply-side perspective, increasing the stock of science, technology, engineering, and mathematics (STEM) human capital is crucial for the improvement in the quantity and quality of innovative ideas that lead to productivity growth. In some scenarios, increasing the supply of inventors and scientists relies on policies favoring high-skilled immigration (Hunt & Gauthier-Loiselle, 2010; W. R. Kerr & Lincoln, 2010; Moser et al., 2014). In other scenarios, expanding the knowledge workforce requires improving the allocation of talents by reducing the barriers to entry into innovative occupations for underrepresented groups (Bell et al., 2019; Hsieh et al., 2019).

As of 2015, women only make up about 12% of the US inventor population. The differential participation in inventive activities between men and women may not only affect the direction and outcome of R&D but dampen the long-run productivity growth. Koning et al. (2020) show that biomedical research fields with a higher presence of female-led teams produce more female-focused inventions. One consequence of the lack of women in performing R&D activities involves biased innovation outcomes. For instance, the gender bias in cardiovascular clinical trials results in a higher wrong diagnosis rate for women in the population (Kim et al., 2010). On the other hand, the inclusion of talents is a notable contributor to economic prosperity. Hsieh et al. (2019) demonstrate that almost 40% of growth in the US GDP per capita from 1960 to 2010 is attributable to reduced barriers to entering high-skilled occupations for women and black people.

Understanding the contributors to the gender gap in patenting is vital for implementing policies that ensure inclusive participation in innovation and the sharing of inventions that benefit all groups. Yet gendered differences in STEM education and training only partly explain the under-representation of female inventors. In chapter 4, we take a new perspective on career interruptions related to motherhood and child-bearing obligations and explore the role of family policies on the participation and productivity of women in patenting.

1.4 Contribution to literature

Chapter 2 studies how product prices respond to the expiry of patents and the underlying mechanism for a group of more than 800 consumer products. We combine patent data with product data from Amazon.com and estimate the effect of patent expiry on various model specifications on product-patent pair and product level. Our findings show that patent expiry leads to a 7–8 percent drop in product prices and that the effect is heterogeneous regarding patent characteristics. A set of placebo tests where we build counterfactual expiry events suggest no such effect. In terms of the mechanism, the drop in product prices displays a U-shaped relationship with product market competition. Such evidence implies that firms may use preemptive pricing as an entry deterrence strategy for products with expired patents.

By linking patents to commercial products as well as product features, this paper makes the first attempt to estimating the monopoly prices conferred by patents for a sample of consumer goods. Our findings significantly contribute to understanding the incentive mechanism offered by patent protection and make way for future analysis on the welfare effects of patent protection. Our paper adds to the literature on the value of patent rights (Trajtenberg, 1990; Lanjouw, 1998; B. H. Hall et al., 2005; Kogan et al., 2017); we also complement the studies on the effects of IPR on prices and market structure (Caves et al., 1991; Grabowski & Vernon, 1992; Li et al., 2018; I. Reimers, 2019).

Chapter 3 investigates the effects of firms' disclosure of innovative portfolios on the diffusion of inventions. Firms communicate to the public their innovative portfolios and signal the commercial value of patents through virtual patent marking, which is a practice intended to prevent innocent infringement from competitors. We use a sample of 843 virtually marked patents and exploit the time-stamped variations of when patents were listed in web documents. We find heterogeneous effects of virtual marking conditional on appropriability regimes. While marking attracts more follow-on inventions for patents in weak regimes, competitors shun the marked inventions in strong regimes; moreover, the similarity between a citing patent and the focal patent also reduces for patents in strong regimes after marking. Patent importance also plays a role in exacerbating follow-on inventions for patents in strong regimes.

Past research on disclosure and knowledge spillovers predominately focuses on the divulgence of codified information, but few have looked into the impact of firms' signaling of valuable inventions, an alternative form of disclosure. This chapter is among the first papers studying the effect of virtual patent marking on follow-on inventions from competing firms. We add more empirical evidence to the research regarding information disclosure and cumulative innovations (Scotchmer & Green, 1990; Anton & Yao, 2004; Furman et al., 2018; Baruffaldi & Simeth, 2020; de Rassenfosse et al., 2020).

Chapter 4 examines the gender gap in inventive occupations by evaluating the effect of maternity leave policies on the retention and productivity of female inventors in the US from the early 80s to 90s. We use data on 1.4 million inventors and are able to identify the demographic information of inventors with web-scraped inventor ages. By exploiting variation in the timing of maternity leave passage across the US, the results of our event-study estimates suggest that job-protected maternity leave policies enable women of reproductive ages to stay longer in patenting-related careers. A further survival analysis confirms the dwindling gap in retention after policy implementation. However, the effect of maternity leave policies on productivity is limited, possibly due to the inaccessibility of childcare facilities for working mothers.

This chapter makes an effort to bridge the studies on participation in innovation to the studies on family policies and women's labor supply. Our findings align with the evidence from labor economics on childbirth, maternity leave, and women's labor market participation (Berger & Waldfogel, 2004; Bertrand et al., 2010; Goldin & Mitchell, 2017). In addition, this chapter contributes to the works on the gender gap in participation and productivity in science and innovation occupations (Ding et al., 2006; Hunt et al., 2013; Mairesse et al., 2019; Moser & Kim, 2020).

2 Patents and Supra-Competitive Prices: Evidence from Consumer Goods

This chapter is written in collaboration with Gaétan de Rassenfosse.[1]

Abstract:

The patent system is a central tool in innovation policy. The prospect of monopolistic pricing conferred by patent protection supposedly encourages firms to innovate. However, there is scant empirical evidence supporting the existence of higher markups for patent-protected products. Using an original data set that links a broad range of consumer products to the patents that protect them, we study the impact of patent protection on product prices. The empirical strategy exploits exogenous variations in patent status, namely the fall of the patent in the public domain after the statutory 20-year term limit is reached. We find that a loss of patent protection leads to a 7–8 percent drop in product prices. The price drop, which starts about one year before patent expiry, is larger for more important patents and is more pronounced in more competitive product markets. Finally, our findings also pass a set of robustness checks and placebo tests.

Key words: innovation, markup, patent system, product, R&D incentive

[1]We gratefully acknowledge Philippe Aghion, Stefano Baruffaldi, Fabian Gaessler, Alberto Galasso, Deepak Hegde, Martin Hetu, Adam Jaffe, Will Matcham, Ambre Nicolle, Christopher Tucci, and Beth Webster for their helpful comments. We also thank seminar and conference participants at HEC Paris, ETH Zurich, KU Leuven, the Max Planck Institute for Innovation and Competition, and the CCC Seminar at NYU Stern School of Business.

2.1 Introduction

Innovation, which is a key driver of productivity growth (Romer, 1990; Aghion & Howitt, 1992), is subject to several well-documented market failures that lead to under-investment in R&D activities (e.g., S. Martin & Scott, 2000; Bloom et al., 2019). Consequently, the social planner incentivizes R&D investments using a variety of policy instruments. One such instrument is the patent system, which offers inventors a temporary exclusion right over their inventions. This right allegedly allows them to charge monopolistic prices for their products in order to recoup their R&D investments (Plant, 1934; Arrow, 1962; Nordhaus, 1969)—we call this the 'monopoly pricing hypothesis.'

The theoretical literature assumes that monopoly over an invention translates into ability to charge supra-competitive prices in the product market. However, it is not clear that this is the case. For instance, competitors could invent around the original patented invention and offer a product that looks very similar to the end consumer, thereby breaking down market exclusivity. Moreover, recurring discussions about the poor 'quality' of issued patents (e.g., Lemley & Shapiro, 2005; Jaffe & Lerner, 2011), adds another reason to be skeptical. If patents do not allow innovators to sustain supra-competitive prices, the main argument about the effectiveness of the patent system in encouraging innovation collapses.

To the best of our knowledge, empirical research on the effect of patents on product prices has been limited to drugs. Yet, many observers would agree that drugs offer a very favorable setup for testing the effectiveness of patent protection. First, the active ingredient patent *is* the drug—the patent and the product are, therefore, virtually the same. Next, the costs of drug innovation are very high whereas the costs of imitation are comparatively low, making the industry prone to free-rider problems and patent protection all the more relevant. These arguments help explain why patent protection is particularly effective for the pharmaceutical industry compared to other industries (Mansfield, 1986; Levin et al., 1987; Harabi, 1995; Cohen et al., 2000). But the patent system has not been designed for drugs alone; innovators across all technology fields exploit it. Despite the centrality of the monopoly pricing hypothesis for justifying the existence of intellectual property rights, evidence on other industries is scant.

This chapter empirically examines the effect of patent protection on the price of an array of consumer products. We collect original data on patent-product associations and study the effect of an exogenous *loss* of patent protection on product prices. We have matched 2778 patents to 825 products available on the Amazon.com e-commerce website and have tracked the prices of these products for a period of up to eight years. We study the change in price around the time of patent expiry. Because patent protection is limited in time by law, patent expiry is exogenous to the quality of the underlying invention or to its commercial value. Furthermore, we are able to isolate the effect of patent expiry from the effect of product depreciation by controlling for product model displacement and product age. The empirical analysis then explores the heterogeneous effects of patent expiry across patent type and importance. It also portrays the price evolution around the time when patent terms expire. Finally, it considers how prices react to the intensity of the competition in the product market.

We find that patent expiry is associated with a 7–8 percent drop in product prices, and that the effect is larger for more important patents (i.e., patents protecting more products). We observe that the price starts dropping about one year *before* patent expiry, possibly suggesting strategic entry deterrence from the incumbent (Milgrom & Roberts, 1982; Goolsbee & Syverson, 2008). We also observe that the decline in price is more pronounced in more competitive markets, with some evidence of a U-shape relationship between the price drop and the level of the competitive pressure. Finally, placebo tests on samples of fake patent expiry events confirm the validity of our identification strategy.

The chapter adds to the long-standing debate on the effectiveness of intellectual property rights in stimulating innovation (summarized in B. H. Hall, 2007; Lerner, 2009; Budish et al., 2016). Overall, the results provide evidence supporting the monopoly pricing hypothesis—incumbents seem to be able to charge supra-competitive prices during patent protection. Furthermore, the estimates we obtain are important to quantify the extent of the subsidy conferred by the patent system (e.g., Schankerman, 1998). The 7–8 percent figure helps us understand the cost of the patent system that consumers bear in exchange of more innovative products. The chapter also adds to the literature on the economic valuation of patents. Scholars have proposed a variety of approaches to estimate patent value (e.g., B. H. Hall et al., 2007; Arora et al., 2008; J. Bessen, 2008; Kogan et al., 2017) but none have exploited the source of data we use.

The rest of this chapter is organized as follows. Section 2.2 provides background information on what we call the monopoly pricing hypothesis. Section 2.3 presents our empirical research design and Section 2.4 explains the construction of the dataset and introduces the main variables. Section 2.5 reports our findings. Section 2.6 offers concluding remarks.

2.2 Background

2.2.1 The monopoly pricing hypothesis of patents

Following Arrow (1962) and Nordhaus (1969), a vast theoretical literature has studied the design of patent systems. Contributions have looked into the optimal duration, strength, breadth and scope of patent protection under various industry structures and invention types (e.g., Kamien & Schwartz, 1974; Judd, 1985; Gilbert & Shapiro, 1990; Klemperer, 1990; Waterson, 1990; Denicolo, 1996; Matutes et al., 1996; O'Donoghue et al., 1998; Erkal, 2005; Acemoglu &

Akcigit, 2012).

Models of the patent system take different forms but the core principle works as follows. Knowledge is notoriously difficult to appropriate, which translates into a wide gap between the private returns to inventive activities and the social returns. As a result, competitive markets underincentivize private research investments compared to the social planner's preference. Governments intervene by granting a monopoly right over inventions in order to increase appropriability. The welfare loss created by the monopoly right is offset by the dynamic efficiency of increased investments in inventive activities.

The theoretical literature implicitly equates *monopoly over an invention* with *monopoly over a product*. That is, it assumes that patent protection (covering an invention) allows the firm to charge supra-competitive prices (for the product). This assumption is far from obvious. First, an invention does not come in the form of a finished product ready for sale. The inventor must undertake costly and risky development and testing to transform the invention into a commercially viable product (Sichelman, 2009).[2] Second, the U.S. Patent and Trademark Office (USPTO) has been criticized for issuing low-quality patents, in the sense that many patents would not stand up in court if litigated (Lemley & Shapiro, 2005; J. E. Bessen et al., 2008; Jaffe & Lerner, 2011). If patents are indeed "worthless" (Moore, 2005), the actual protection they offer might be substantially weaker than we assumed. Third, monopoly over an invention, even if a patent is 'solid', does not translate necessarily into monopoly over the final product. The next section explains this latter point in greater detail using the computer mouse as an example.

2.2.2 Patent protection and product price

Patent protection typically offers a monopoly over a specific feature of a final product, which may translate into an increase in product quality or a broadening of product variety (e.g., Horstmann et al., 1985; Waterson, 1990). These features may or may not allow the firm to charge supra-competitive prices.

To illustrate, let us consider the case of the computer mouse. Some inventions in this area are truly radical and pave the way for an entirely new product market. U.S. patent 3,541,541, entitled "X-Y Position Indicator for a Display System," falls in this category. The patent, filed by Douglas Engelbart in 1967, is known as the first computer mouse patent. The technology

[2] Note that invention owners may recoup their R&D investments not by commercialization in the product market but by licensing or selling their inventions to competitors (Arora et al., 2004). In markets for technologies, the actual invention *is* the 'product' being traded. Several studies have documented the prime role of patent protection in markets for technologies (Gans et al., 2008; de Rassenfosse et al., 2016). The present study focuses on product commercialization.

was licensed to Apple, Xerox, and a few other companies, creating *de facto* a market oligopoly.[3] Computer mice at the time sold between $200–$400, equivalent to $500–$1000 in 2020 price.[4] Since then, technological progress regarding the computer mouse has taken many forms.

Consider, first, the case of inventions that increase product quality. A radical technological shift occurred with the first optical mouse, which offered a superior solution compared to traditional mechanical mice—preventing dirt from getting stuck inside the mouse. The shift from mechanical to optical mouse was one of the main advances in this market, but optical mice still perform the same function as mechanical mice. This technology shift represents an improvement in product quality that can command a higher price. Another radical shift occurred with the first touchpad patent, U.S. Patent 5,305,017, which created a substitute technology—indeed, a new product, at least for the laptop market segment. However, new technologies do not necessarily improve product quality or create entirely new product families. For example, optical mice may rely either on lasers or on LEDs but function in the same way for the end user and offer otherwise similar features. The existence of two substitute technologies to address the same problem breaks down the exclusivity over optical mice, and exemplifies that exclusivity over an invention does not guarantee market exclusivity.

Next, consider the case of inventions that broaden product variety either by segmenting the market or by adding functionalities. Regarding market segmentation, adding more lasers on an optical mouse improves the tracking precision. This feature may appeal to a specific consumer segment such as gamers, who are willing to pay a higher price—but again, there are many ways to improve the tracking precision. Sometimes, inventions are developed to serve lower-end segments—indeed, 'frugal innovation' and 'innovation by subtraction' offer alternative ways of developing new products (e.g., Hart & Christensen, 2002). This is the case for Logitech's U.S. patent 7,030,857, which is typically associated with lower-end mice of the M series, such as the 'M100 Mouse.' Regarding functionalities, an invention may add a feature, which may turn out to be adopted widely, such as the scrolling wheel (U.S. patent 5,313,230), or abandoned, such as the side click.

In a nutshell, the relationship between patent protection and product price is complex: some patents can be invented around, others may cover lower-end versions of a product, and others may turn out to be a commercial flop. As far as we can ascertain, the effect of patents on product prices has not been tested empirically, to the notable exception of pharmaceuticals. The next section reviews the evidence in the pharmaceutical industry.

[3] Sadly for the inventor, the invention was not commercially viable until 1984 when Apple released the Macintosh, three years before the patent's expiration. See https://www.dougengelbart.org, last accessed on November 17, 2020.

[4] Source: https://www.macworld.com, last accessed on November 17, 2020.

2.2.3 The case of the pharmaceutical industry

The pharmaceutical industry offers an obvious set-up for studying the effect of patent protection on product prices. The drug discovery and development process is costly and risky. R&D expenditure for each new molecular entity is estimated at $1.8 billion; meanwhile, the average success rate from pre-clinical stage to launch is estimated at about 8 percent (Paul et al., 2010). Furthermore, patents are an efficient way to deter entry in this industry. Drugs are so-called 'discrete' products with a well identified 'invention' (i.e., an active ingredient) clearly described in the patent specification. However, production is relatively cheap, and patent-protected drugs are usually sold with a high markup (F. S. Morton & Kyle, 2011). This setup is particularly attractive for generic manufacturers, who enter the market as soon as drugs lose patent protection.

A host of studies has investigated the effect of patent protection on the price of drugs. This stream of research has been facilitated by data on the correspondence between drugs and patents compiled in the Orange Book Datafiles by the U.S. Food and Drugs Administration (FDA). Studies typically focus on the evolution of drug price around the time of patent expiry. Since patents are valid for a limited period of time, patent expiry is an exogenous event, allowing scholars to establish the causal impact of (a loss of) patent protection on price.

Using data on 30 drugs that lost patent protection in the 1976–87 period, Caves et al. (1991) estimate that the innovator's price declines by 4.5 percent on average. Furthermore, generic substitutes are sold about 17 percent below the innovator's pre-entry price. They attribute the relatively small price decline of the branded drug to the "loyalty-inducing goodwill" accumulated by the innovator during the period of patent protection. Grabowski and Vernon (1992) examine prices and market shares of 18 drugs turning off-patent after the implementation of the 1984 Drug Price Competition and Patent Term Restoration Act, which eased the testing requirements for entry by generic drugs in the United States. They find that prices for most branded drugs did not react strongly to entry; nominal prices continued to increase following roughly the same trend as during the pre-entry period. They attribute this result to the strength of brand loyalty for branded drugs. By contrast, generic drugs quote prices that are 39 percent lower than branded drugs at date of entry, and prices of generic drugs decrease sharply over time.

The low sensitivity of the price of off-patent branded drugs has been confirmed by most studies (cf. Wiggins & Maness, 2004), both in the U.S. market (Frank & Salkever, 1997) and the European market (Vandoros & Kanavos, 2013). However, the features of the drugs market make generalization to other product markets perilous. When drugs are for repeated use, consumers may have developed a strong preference for the branded version during patent protection. Besides, concerns about perceived quality for the generic versions and recommendations from doctors may exacerbate brand loyalty.

2.3 Empirical approach

The goal of the econometric analysis is to quantify the effect of a loss of patent protection on product prices.

2.3.1 Identification strategy

In an ideal experiment, one would observe time series of the price of products, each protected by exactly one patent. We would then let some patents lapse randomly. Products with a lapsed patent would form the treatment group, and products with an active patent throughout the study period would form the control group. Every control product would be assigned a fake treatment date of a hypothetical lapse. The average treatment effect would then be the difference in the change in price around the time of patent lapse (or hypothetical lapse) between the treatment and the control group. Needless to say, such an experiment cannot be implemented in practice—patent owners are reluctant to allow scholars to let lapse randomly commercially valuable patents.

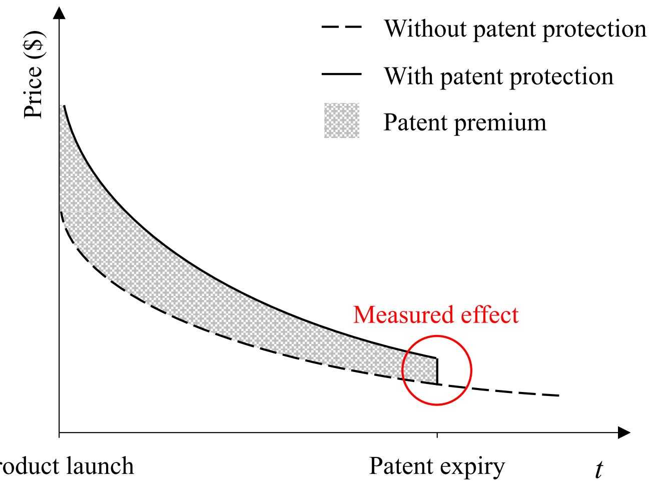
The present study exploits observational data on the price of patent-protected products that lose patent protection. There are three ways in which a product can lose patent protection. First, patents can be challenged in court and be invalidated. Galasso and Schankerman (2015) exploit data on invalidations to study the effect of patents on cumulative innovation. However, invalidations are rare events. Lemley and Shapiro (2005) estimate that a mere 0.1 percent of patents are litigated to trial. Second, the patent owner may decline to pay the renewal fees required to keep the patent in force. The patent consequently lapses and falls in the public domain—everyone is then free to use the invention. This source of variation is not appropriate for our purpose because the decision to let a patent lapse is presumably endogenous to the underlying product's commercial success and, therefore, to its price. Third, the patent is held active until the maximum allowed term (usually 20 years) and automatically expires after that period. This event is exogenous to product quality, and there is nothing that the firm can do to prevent expiry. Our identification strategy exploits variations in the time of patent expiry, as illustrated in Figure 2.1.[5]

Although the patent expiry event is exogenous to the firm, its exact date is known and the firm

[5] Arora et al. (2008) and P. H. Jensen et al. (2011) use the term 'patent premium' to indicate the proportional increase in value to an invention due to patent protection. In the context of the present analysis, the patent premium would correspond to the gray area in Figure 2.1. It is the overall surplus that the firm can extract throughout the life of the patent (the shapes of the price slopes and the gray area are arbitrary and only serve to illustrate the point).

can adapt accordingly. For instance, it could launch a new generation of the product in an attempt to capture the most profitable market segment (Chandy & Tellis, 1998; Van Heerde et al., 2010). Consequently, the econometric regression will control for potential confounding factors. Note that filing new patent applications to protect some features of the original product is not possible. Any unpatented invention embedded in the product would have long been part of prior art—and, therefore, no longer patentable—under U.S. patent law.

Figure 2.1 – Schematic representation of the measured effect



2.3.2 Econometric model

We exploit variations in patent status in a three-dimensional panel setting. The unit of analysis is the natural log price P in month t for product i protected by patent j.[6] The main panel specification is as follows:

$$\log P_{ijt} = \beta_0 + \beta_1 Expired_{jt} + \beta_2 ProdAge_{it} + \beta_3 NewGeneration_{it} + \boldsymbol{X\gamma} + \mu_{ij} + \epsilon_{ijt} \quad (2.1)$$

The variable of interest, $Expired_{jt}$, is a dummy variable that takes value 1 if patent j is expired in month t, and 0 if the patent is still active. All variables are formally introduced in the next section. The parameter β_1 captures the change in product price associated with patent expiry.

One empirical challenge lies in the fact that the price of a given product will tend to naturally decline over time. Therefore, the coefficient β_1 may simply capture the effect of the passing of time.[7] Our solution to this issue is twofold. First, we control for the effect of the passing

[6]It is common to model product prices in the log linear form (e.g., Brynjolfsson & Kemerer, 1996; Milyo & Waldfogel, 1999; Ashenfelter, 2008).

[7]A first-difference specification (ΔP_{ijt}) would not address this issue satisfactorily because the general price

of time using product age (variable $ProdAge_{it}$) as well as various non-linear specifications ($ProdAge^2_{it}$ and $\log ProdAge_{it}$). Second, we also perform a placebo test where we randomly assign a fake treatment date and compare placebo estimates with baseline estimates. The placebo estimates are subject to the natural price decline but not to the expiry events. Therefore, comparing estimates with placebo and actual dates informs us about the validity of the empirical setup.

Although patent expiry is exogenous to the firm and the product, the date of patent expiry is known. A firm can, therefore, release a new model of the product in anticipation of patent expiry. The regression model controls for the variable $NewGeneration_{it}$ to absorb the effect of product displacement. It takes value 1 if a newer version of product i is available in month t (and all the months afterwards), and 0 otherwise.

The vector $\boldsymbol{X}$ includes a set of control variables. Its exact composition varies depending on model specification. It includes the intensity of competition as well as patent-level variables. It also includes a set of dummy variables for each calendar month in order to control for seasonal sales and promotional offers. Finally, it includes a set of dummy variables that capture the source of the price information (variables $S^{A/L}_{1-4}$, defined below).

As the next section explains, our data are many-to-many matches between products and patents. Consequently, we are able to control for product-patent pair fixed effect (μ_{ij}) to capture time-invariant idiosyncratic characteristics such as the technological content of a patent, its importance for the product or other unobserved product characteristics. In alternative specifications, we will also include individual product and patent fixed effects (μ_i and μ_j, respectively).

Finally, ϵ_{ijt} is the error term. We estimate standard errors clustered at the product-patent level to account for potential serial correlations of prices within each unit. We have also estimated the regression models with standard errors clustered at the product level, with no change to the statistical significance of the main findings.

2.4 Data and variable construction

Studying the effect of patents on product prices calls for three elements: data on the products, data on the patents, and a way to link products to patents. Establishing the link between products and patents is the most challenging part, and we start by presenting our novel approach for doing so. We then turn to data on products and on patents. The final dataset is a monthly unbalanced panel of 489,878 observations associated with 14,621 product-patent pairs corresponding to 825 patented products (covered by 2778 patents) for the period 2011–

decline might not be constant over time, and the effect might not be contemporaneous to patent expiry.

2019.

2.4.1 Data on product-patent links

We collected data on the link between products and patents by manually searching for Virtual Patent Marking (VPM) web pages of consumer good companies. VPM was introduced in U.S. patent law under the 2011 Leahy-Smith America Invents Act (AIA). The AIA allows patentees to affix the word "patent" or "pat." on the product along with a URL of a web page that associates the patented product with the patent number(s). The marking statute enables patentees to give public notice that the article is patented, which can prove useful in infringement cases. de Rassenfosse (2018) explains that patentees have incentives to disclose information accurately because listing patents that do not cover a product exposes them to false marking suits.

Before delving further into the data, a note of caution is warranted. The marking statute provides firms with an incentive to list patents that they *own*. Manufacturing firms do not care as much if the patents they license from other firms are being infringed—indeed, it is usually the patent owner that files infringement suits, not the licensee. Thus, we may not have complete information on the patent coverage of products. Although licensors may require licensees to mark their products with the licensed patents, we cannot be sure that they do. To mitigate this concern, we purposefully excluded products that are well known to exploit licensed technologies.[8] Having noted this, a lack of data on licensed patents does not threaten our empirical analysis. Indeed, there is no reason to suspect that the timing of patent filing (and, therefore, expiry) for licensed patents exactly and systematically coincides with that of the innovator's own patents.

We obtained product-patent information for 825 products sold in the United States by 77 firms. Products are all consumer goods in a broad sense in that they are all available on the Amazon.com e-commerce website. We classify products using the 13 Amazon 'Departments' to which they belong (henceforth, product categories). For example, the 'Appliances' category includes the 'Dyson DC35 Cordless Stick Vacuum' and the 'Emerson CF830 Ceiling Fan.' Table 2.1 provides an overview of the number of firms, products and patents by product category. 'Electronics' is the most populated category, covering nearly 40 percent of products and 50 percent of patents. Appendix Table A.1 presents a list of representative products sold by each firm.

[8] For instance, we came across the VPM web page of mobile phone manufacturer BlackBerry. The company only lists patents that it owns and we have decided to exclude it from our sample. See: https://www.blackberry.com, last accessed on November 17, 2020.

Table 2.1 – Summary of firms, products, and patents by product category

Product category	**Number of firms**	**Number of products**	**Number of patents**
Appliances	4	52	335
Automotive parts	5	117	118
Baby Products	2	7	15
Clothing, Shoes & Jewelry	2	7	14
Electronics	23	310	1348
Health & Household	6	163	357
Industrial & Scientific	7	17	33
Musical Instruments	2	13	71
Office Products	5	21	81
Software	2	9	189
Sports & Outdoors	8	40	54
Tools & Home Improvement	10	36	135
Video Games	1	33	28
Total	77	825	2778

Table 2.2 shows the number of patents per product, which can be seen as a measure of the 'complexity' of products.[9] The median number of patents per product is 4, but the variable is highly skewed. In some categories, such as 'Electronics' and 'Software,' a quarter of products are covered by more than 77 and 66 patents, respectively. The table also presents the complementary figure, namely, the number of products protected by the same patent. It is a measure of patent importance. A patent protects a median number of two products in our sample.

[9]The literature offers several definitions of complex products. They are characterized by a "complex web of dependencies and interactions between the modules" (Sharman & Yassine, 2004), they are "high cost, engineering-intensive products" (Hobday, 1998), and their development involves a "large number of both physical components and design participants" (Sosa et al., 2004). In this chapter, we define complex products as products involving multiple patented components.

Table 2.2 – Patent and product intensity

Product category	Patents per product			Products per patent		
	Bottom 25%	Median	Top 25%	Bottom 25%	Median	Top 25%
Appliances	1.5	16	35	1	2	4
Automotive parts	1	1	2	1	1	2
Baby Products	1	2	2	1	1	1
Clothing, Shoes & Jewelry	4	5	6	1	3	3
Electronics	2	8	77	1	3	7
Health & Household	2	5	9	1	2	3
Industrial & Scientific	2	3	4	1	1	2
Musical Instruments	1	2	5	1	1	1
Office Product	1	4	6	1	1	1
Software	13	38	66	2	2	2
Sports & Outdoors	1	3	5	1	2	4
Tools & Home Improvement	1.5	2.5	5.5	1	1	2
Video Games	3	3	4	1	1	1
Total	1	4	11	1	2	5

2.4.2 Data on products

All products in our sample are (or were) available for purchase on the U.S. platform of Amazon.com. We manually searched for the products on Amazon.com, with a view of recovering the ASINs, the unique product identifiers.[10] We then collected various information about products in our sample.

Product price

We used the ASINs to obtain the Amazon price history for all products using *Keepa*, a commercial price-comparison web service that provides historic price data since 2011.[11] *Keepa* tracks Amazon's products several times per day and records their prices and the inventory status (in-stock and out-of-stock). *Keepa* records the prices for an item whenever a change occurs. Therefore, missing data on prices means either that the price has remained stable (such that no price was recorded) or that the item was temporarily out-of-stock (such that no price could have been recorded).

We use two prices: the Amazon price and the List price. The Amazon price is the actual sales price at which an article is sold. The List price is suggested by the manufacturer and does not always correspond to the Amazon price.[12] In order to have a balanced panel, we impute

[10] The Amazon Standard Identification Number (ASIN) is a 10-character alphanumeric unique identifier used for product identification within the Amazon organization.

[11] See http://www.keepa.com, last accessed on November 17, 2020. This service has already been used in academic research, see, e.g., I. C. Reimers and Waldfogel (2020).

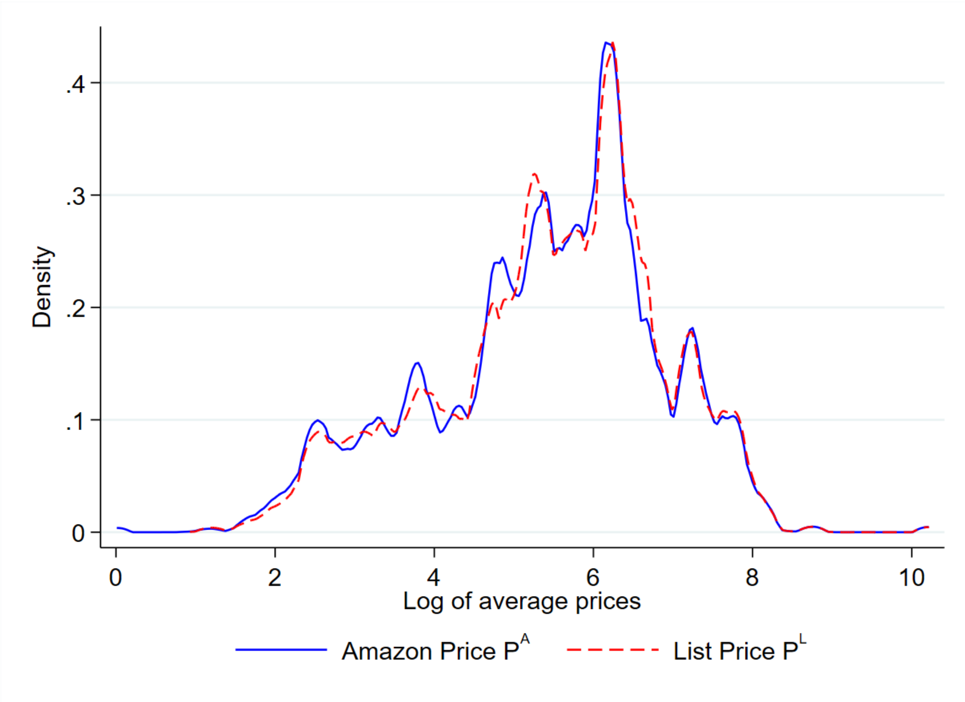
[12] On average, the Amazon price is 9.6 percent lower than the List price.

missing price data for both the Amazon price and the list price indices. For both indices, we first reconstruct the daily price series, which we then average by month. We follow some simple rules to impute missing price data for the daily series. Regarding the Amazon price index, if there is a gap in the Amazon price series while the product is in stock, we populate the missing data with the last known Amazon price. If there is a gap in the Amazon price series while the product is out of stock, we populate the missing data with the last known List price. If the List price is not available (out-of-stock), we again populate the missing data with the last known in-stock Amazon price. We perform the mirror operation for the list price index. Next, we average the daily prices by month and take the natural logarithm to obtain the dependent variables P^A and P^L.

Each price variable also comes with a set of five mutually exclusive and exhaustive dummy variables that indicate the main source of the price data in a given month (S^A_{0-4} and S^L_{0-4}). Regarding P^A, the variable S^A_0 takes value 1 if most of the daily prices in the given month are directly available from *Keepa*, the variable S^A_1 takes value 1 if most of the daily prices in the given month come from the in-stock Amazon prices with some out-of-stock prices imputed with Amazon prices, the variable S^A_2 takes value 1 if most of the daily prices in the given month come from the in-stock Amazon prices with some out-of-stock prices imputed with List prices, the variable S^A_3 takes value 1 if most of the daily prices in the given month come from the out-of-stock Amazon prices, and the variable S^A_4 takes value 1 if most of the daily prices in the given month come from the out-of-stock List prices. We perform the mirror operation for the S^L_{0-4} dummies. These variables will be used as controls in the regression analysis. In Appendix Table A.2, we report the prevalence of each price source at some relevant points in time. We find no particular pattern between the source dummies and the expiry event. Consequently, we are confident that the imputation method does not affect the validity of the estimates.

Figure 2.2 depicts the distributions of the P^A and P^L variables. To generate this figure, we pooled together the monthly prices across all time periods for each product-patent pair. The distributions of both price series largely overlap. On average, a product in our sample costs \$221 (minimum of \$2, median of \$270 and maximum of \$14,985). The P^A variable corresponds to the market price and, therefore, forms our baseline measure of price. However, we also report estimates performed using the P^L variable for the sake of robustness.

Figure 2.2 – Distributions of log of average product prices



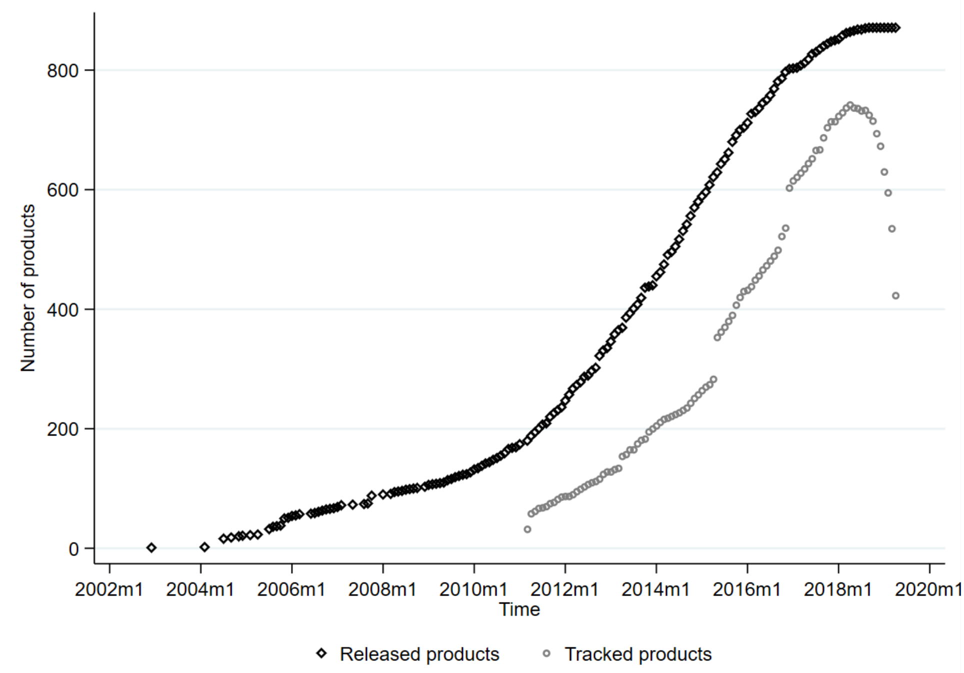
Product release date and new generation of product

We collected data on the product's release date as well as on the introduction of a new generation of product.

The product release date allows us to control for the product's age (variable $ProdAge$, in months), thereby accounting for the natural decline in price over time. The release date is set equal to the date at which the product was first available on the Amazon website or, if the information was missing, to the date of the first product review on the website. If no review is available, we set the release date equal to the date of the earliest sign of commercialization we could find online about that product.

The first product released in our data can be traced back to 2002, followed by the successive introduction of products to the market until late 2018, as shown in Figure 2.3. These products were first tracked by $Keepa$ in March 2011 and last observed in April 2019. The number of products tracked by $Keepa$ keeps growing until it peaks in early 2018. It then drops as products eventually exit the market.

Figure 2.3 – Products having been released vs. products tracked by *Keepa*



Notes: The 'released products' series indicates the number of products having been released up to a given month (cumulative variable). The 'tracked products' series indicates the number of products tracked by *Keepa* in a given month.

When a firm launches a new generation of a product, it may decide to adapt the price of the older generation. Since patent expiry may coincide with—or even trigger—new product introduction, the regression model controls for the availability of new products. We searched on Amazon.com and on other online resources for new product introduction. The dummy variable $NewGeneration$ takes value 1 when a new product generation becomes available, and value 0 as long as no new product generation exists.

Competitive pressure

We propose two measures of product market competition. The first measure ($Substitutes$) captures the number of alternative products of similar functionality sold by competitors. The second measure ($Competitors$) captures the number of competing firms selling substitute products (see, e.g., Bresnahan & Reiss, 1991).

We identify substitute products using Amazon's recommendation algorithm, which presents a menu of relevant items on the landing page of each product. This algorithm lists relevant products that a potential buyer might be interested in based on product similarity and the purchasing behavior of customers.[13] However, the algorithm itself does not distinguish complementary products from substitute ones when offering recommendations. For instance, a search for a Philips electric toothbrush returns not just electric toothbrushes from its rivals, but also toothbrush heads or toothbrush holders. We went through the list of all recommended items manually and only considered products that serve similar functional purposes as substitutes for the target products. When a product was clearly in a different price range, we did not consider it.

Overview of product-level variables

Table 2.3 provides descriptive statistics for all product-level variables. The unit of observation is a product-patent pair in a given month (N = 491,336). In our sample, the log of imputed monthly Amazon price (P^A) ranges from 0.1 to 10.22 with a mean of 5.39 (which corresponds to \$219). The variable P^L ranges from 0.43 to 10.22, with a mean of 5.44 (or \$230). Product age (variable $ProdAge$) counts the number of months between the product launch date and month t. It ranges from one month to 187 months (15.5 years) with a mean of 50 months. On average, 23 percent of the product-patent pairs are observed while an upgraded model is available on the market. In addition, a product faces an average of 15 substitutes with a maximum of 59 and a minimum of zero. On average, six firms compete in the same market segment, with a maximum of 30 competitors and a minimum of zero.

Table 2.3 – Summary statistics for product-level variables

	Mean	Standard error	Max	Min
log(imputed Amazon price)	5.39	1.47	10.22	0.01
log(imputed List price)	5.44	1.45	10.22	0.43
Product age (in months)	49.46	31.42	187	1
New generation	0.23	0.42	1	0
No. of substitutes	15.03	15.94	59	0
No. of competing firms	6.32	5.84	30	0
Month	-	-	2011. m3	2019. m4

2.4.3 Data on patents

We collected information on patents from three sources: the USPTO Patent Maintenance Fee Events dataset (last updated on August 26th, 2019), PatentsView.org, and the Patent Claims Research dataset.[14] We considered two types of patents, namely utility patents and design

[13] For an explanation of Amazon's product recommendation method, please refer to https://www.mageplaza.com, last accessed on November 17, 2020.

[14] The data are available on https://www.uspto.gov, last accessed on November 17, 2020.

patents. A ‘utility patent,’ sometimes called an invention patent, protects the way an article is used and works (its technical aspects), whereas a ‘design patent’ protects the way an article looks (its aesthetic aspects).

Patent expiry

Our variable of interest is a dummy variable that takes value 1 when the patent has expired, and 0 if the patent is still active ($Expired$). Expiration occurs when the patent has reached its maximal statutory life. According to the USPTO Manual of Patent Examining Procedure, design patents have a 15-year term limit from the grant date if filed as of May 13th, 2015, and a 14-year term limit if filed prior to that. No renewal fee is required for designs to be held active. Therefore, a design patent is expired when the statutory term limit is reached.

The case of utility patents is more complex: utility patents filed as of June 8th, 1995, have a term limit of 20 years from the patent priority date; for patents filed prior to that date, the patent term limit is either 20 years from the filing date or 17 years from the issue date, whichever is longer. Renewal fees are charged at three points in time: the fourth year, the eighth year, and the twelfth year after patent grant. A utility patent is active until the due date of the next payment, or at the termination of term if renewed at the twelfth year. Therefore, a utility patent is expired when all the renewal fees are paid, as indicated in the Patent Maintenance Fee Events dataset, and when the statutory term limit is reached.

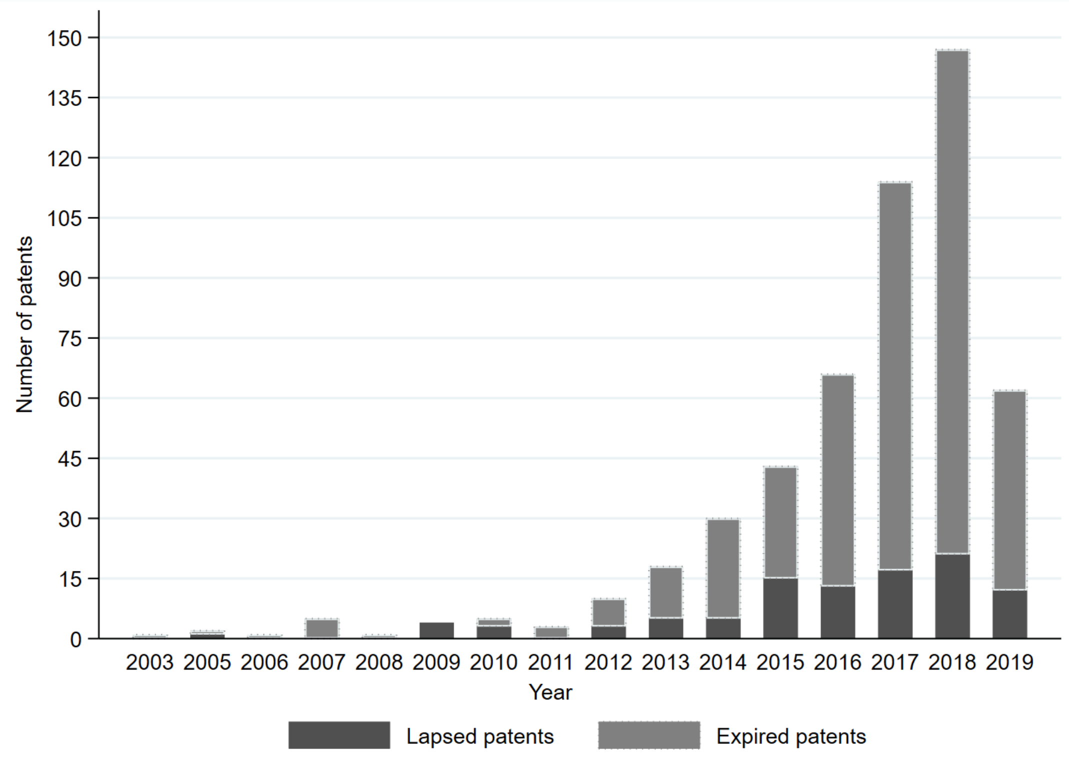
Recall that we do not exploit patents that lapse (which arise due to failure to pay the renewal fees). This is because the decision to let a patent lapse is driven, among others, by market consideration; it is likely to be endogenous to the price of the underlying product. By contrast, patent expiry after full term is clearly exogenous—there is nothing the firm can do to prolong patent life.[15]

Figure 2.4 provides a breakdown of the number of lapsed vs. expired patents in our sample. Overall, 99 patents lapsed and 394 patents expired in the period from 2003 to 2019 (the remaining 2285 patents remained active throughout the study period). Additional analysis (not reported) indicates that patent lapses occur predominantly in products that build on a large number of patents—unsurprisingly so, because the importance of any single patent presumably decreases as the number of patents protecting a product increases. In a robustness test, we find that excluding lapsed patents from the sample leads to similar results.[16]

[15] Although a patent’s term can be extended under certain circumstances, for example, in case of delays in the examination. We only find 22 expired patents with an extended term. Adjusting the term or excluding the extended patents doesn’t affect our results.

[16] Results are available upon request from the authors.

Figure 2.4 – Distribution of lapsed and expired patents



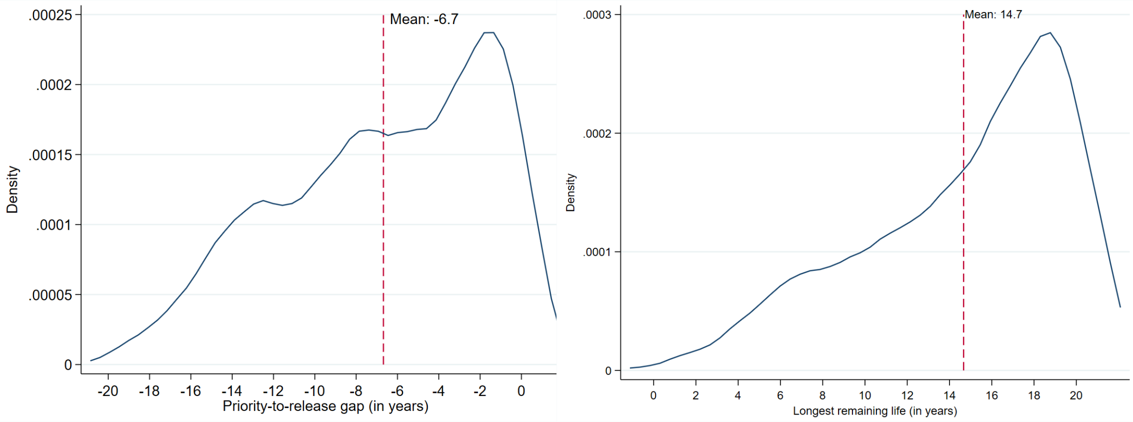
As mentioned earlier, the key date to determine patent expiry is the priority filing date, which, roughly speaking, corresponds to the first date at which the invention is disclosed through the patent system.[17] Combining the priority filing date and the product release date provides us with an estimate of the age of inventions at the time they reach the market. The left-hand side of Figure 2.5 depicts the number of days elapsed between the patent priority date and the release date of a focal product protected by that patent. When a patent covers more than one product, we select the earliest released product. On average, it takes 6.7 years for a patented *invention* to be commercialized into a *product* in our sample, with a mode at about two years.

The right-hand side of Figure 2.5 depicts the distribution of the longest remaining time period for which a product enjoys patent protection. It is counted as the number of days between the product release date and the last maximum expiry date among all patents protecting the product. On average, a product will be protected by at least one patent for a maximum of 14.7 years in our sample. As far as we can ascertain, it is the first time that such statistics have been

[17]According to AIA 35 U.S.C. 102(b)(1), an invention has a one-year grace period before the effective filing date during which disclosure in the form of public use or sale does not render the invention part of the prior art. In other words, inventions disclosed to the public must be submitted to the USPTO no more than 12 months after public disclosure to remain patentable. Refer to https://www.uspto.gov for more information, last accessed on November 17, 2020.

computable.

Figure 2.5 – Density of priority-to-release gap (left) and longest remaining patent life (right)



Notes: Left panel: we removed 252 observations whose patent priority date exceeds product release date by more than one year and 24 observations for which the product is released 20 years after the patent priority date. Right panel: we removed one observation for which the product is released 20 years after patent priority date and 125 observations for which the last maximum expiry date exceeds 21 years after the product release date, considering the one-year grace period of patent filing.

Other patent-level variables

We collected additional patent-level variables in order to capture the economic value or quality of patents (Lerner, 1994; Harhoff et al., 1999; B. H. Hall et al., 2005; Marco et al., 2019). We built four metrics of patent importance. The first is a dummy variable for whether a patent belongs to the top ten percent of forward citation distribution in our sample (*Top 10% citations*). The second variable is the number of distinct four digit IPC sub-classes in a patent (*IPC classes*). The third variable captures the number of independent claims in the patent (*Independent claims*). The fourth variable is the number of products protected by a patent (*Products per patent*).

Overview of patent-level variables

Table 2.4 provides descriptive statistics for all patent-level variables. In our sample (N = 491,336), about 14 percent of product-patent pairs are observed after a patent has expired. Design patents are relatively rare, comprising about 5 percent of the observations. Roughly 22 percent of the product-patent pairs contain a patent whose forward citation count belongs to the top ten percent of citation count. The number of IPC classes for patents in our sample varies from zero to 10 with a mean of 1.33. Zeroes are associated with design rights; utility patents always have at least one IPC class. The number of independent claims varies from 1 to 26 with a mean of 3.32. On average, a patent protects 33 products with a maximum of 97 products and a minimum of one. (The apparent difference with the median number reported in Table 2.2 arises from the skewed distribution of the variable.)

Table 2.4 – Summary statistics for patent-level variables

	Mean	Standard error	Max	Min
Expired	0.14	0.34	1	0
Design	0.05	0.22	1	0
Top 10% citations	0.22	0.42	1	0
IPC classes	1.33	1.22	10	0
Independent claims	3.32	2.53	26	1
Products per patent	32.78	37.21	97	1

2.5 Econometric results

2.5.1 The effect of patent expiry on product prices

Table 2.5 presents results for the baseline specification following equation (4.1). Columns (1)–(3) control for the product-patent pair fixed effect whereas columns (4)–(6) control for product and patent fixed effects separately. The dependent variable is the Amazon price (P^A).

In column (1), we only control for the product-patent pair fixed effect. The coefficient associated with the variable of interest reaches -0.149, meaning that the price is about 15 percent lower when the product loses patent protection. However, as explained previously, this figure may be inflated due to the natural decline in price over time. The regression results in column (2) controls for product age as well as for the availability of a new generation of the product. The coefficient of interest drops to -0.089. Finally, column (3) controls for month dummies as well as the sources of price imputation to absorb noise from seasonal sales and variable construction. On average, product prices decline by 8.2 percent after patent expiry, ceteris paribus. Results in columns (4)–(6) are quantitatively similar. The coefficient of interest settles to 7 percent in column (6).

A 7–8 percent drop in price due to patent expiry is a rather large effect. Indeed, innovative firms usually secure their product market position using multiple strategies besides patent protection (including, e.g., branding and advertising), which helps mitigate the price decline. Furthermore, patents expire fairly late in the product life cycle, presumably when markup has already eroded significantly. To put the 7–8 percent figure in perspective, it is substantially larger than comparable estimates obtained on branded drugs.

Table 2.5 – The effect of patent expiry on product prices, baseline specifications

	(1)	(2)	(3)	(4)	(5)	(6)
Expired	-0.149***	-0.089***	-0.082***	-0.128***	-0.076***	-0.070***
	(0.007)	(0.008)	(0.008)	(0.006)	(0.007)	(0.007)
Product age (in months)		-0.002***	-0.003***		-0.002***	-0.003***
		(0.000)	(0.000)		(0.000)	(0.000)
New generation		-0.165***	-0.161***		-0.165***	-0.161***
		(0.010)	(0.010)		(0.010)	(0.010)
Month dummies			YES			YES
Control for price sources			YES			YES
Pair FE	YES	YES	YES			
Patent FE				YES	YES	YES
Product FE				YES	YES	YES
Constant	5.411***	5.546***	5.560***	5.409***	5.546***	5.560***
	(0.001)	(0.008)	(0.010)	(0.001)	(0.008)	(0.010)
No. products	825	825	825	825	825	825
No. patents	2,778	2,778	2,778	2,778	2,778	2,778
No. pairs	14,621	14,621	14,621	14,621	14,621	14,621
Observations	491,336	491,336	491,336	491,336	491,336	491,336
R-squared	0.946	0.947	0.947	0.946	0.947	0.947

Notes: The dependent variable is P^A_{ijt}. Standard errors clustered at the product-patent-pair level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

In Table 2.6, we present estimates obtained using the List price as the dependent variable (P^L_{ijt}), following the same structure as in the previous table. Patent expiry results in a 8–9 percent drop in list prices, which is qualitatively similar to the baseline estimates. In the remainder of the analysis, we present estimates obtained with the Amazon price variable, but we note that all results are robust to the use of the List price variable.

Table 2.6 – The effect of patent expiry on product prices, alternative price variable

	(1)	(2)	(3)	(4)	(5)	(6)
Expired	-0.234***	-0.116***	-0.093***	-0.201***	-0.099***	-0.079***
	(0.009)	(0.010)	(0.009)	(0.008)	(0.008)	(0.008)
Product age (in months)		-0.005***	-0.008***		-0.005***	-0.008***
		(0.000)	(0.000)		(0.000)	(0.000)
New generation		-0.013	-0.097***		-0.012	-0.097***
		(0.009)	(0.008)		(0.009)	(0.008)
Month dummies			YES			YES
Control for price sources			YES			YES
Pair FE	YES	YES	YES			
Patent FE				YES	YES	YES
Product FE				YES	YES	YES
Constant	5.470***	5.700***	5.905***	5.466***	5.700***	5.905***
	(0.001)	(0.007)	(0.009)	(0.001)	(0.007)	(0.009)
No. products	825	825	825	825	825	825
No. patents	2,778	2,778	2,778	2,778	2,778	2,778
No. pairs	14,621	14,621	14,621	14,621	14,621	14,621
Observations	491,051	491,051	491,051	491,051	491,051	491,051
R-squared	0.960	0.962	0.965	0.960	0.962	0.965

Notes: Dependent variable is P^L_{ijt}. Standard errors clustered at the product-patent-pair level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

The 7–8 percent figure is an average obtained across a large variety of products of different prices. In the following exercise, we analyze whether the effect of expiry depends on the price of products. We break the sample into quartiles of product prices (averaged over the sample period) and report the estimates in Table 2.7. Patent expiry affects prices across the board, although the magnitude of the price drop varies by quartile. The drop in prices reaches 10.8 percent in the highest quartile (column 4), which corresponds to a loss of 153 dollars given the average price for products in this group. The effect is the smallest in the lowest price quartile, reaching about 3 percent.

Table 2.7 – The effect of patent expiry on product prices by quartile of product prices

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	1st quartile	2nd quartile	3rd quartile	4th quartile	1st quartile	2nd quartile	3rd quartile	4th quartile
Expired	-0.034***	-0.094***	-0.076***	-0.108***	-0.033***	-0.081***	-0.065***	-0.093***
	(0.011)	(0.019)	(0.010)	(0.016)	(0.011)	(0.017)	(0.009)	(0.014)
Product age (in months)	-0.002***	0.001	-0.005***	-0.007***	-0.002***	0.001	-0.005***	-0.007***
	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
New generation	-0.192***	-0.290***	-0.104***	0.140***	-0.192***	-0.290***	-0.104***	0.142***
	(0.018)	(0.027)	(0.008)	(0.034)	(0.018)	(0.027)	(0.008)	(0.034)
Month dummies	YES	YES	YES	YES	YES	YES	YES	YES
Control for price sources	YES	YES	YES	YES	YES	YES	YES	YES
Pair FE	YES	YES	YES	YES				
Patent FE					YES	YES	YES	YES
Product FE					YES	YES	YES	YES
Constant	3.576***	5.060***	6.206***	7.389***	3.576***	5.060***	6.206***	7.388***
	(0.011)	(0.019)	(0.011)	(0.021)	(0.011)	(0.020)	(0.011)	(0.021)
Average price ($)	39	174	405	1419	39	174	405	1419
No. products	494	105	92	134	494	105	92	134
No. patents	1,293	1,021	1,106	728	1,293	1,021	1,106	728
No. pairs	2,894	3,366	3,344	5,017	2,894	3,366	3,344	5,017
Observations	120,787	116,684	126,019	127,846	120,787	116,684	126,019	127,846
R-squared	0.934	0.351	0.427	0.859	0.934	0.351	0.427	0.859

Notes: The dependent variable is P^A_{ijt}. Standard errors clustered at the product-patent-pair level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Accounting for patent type and importance

So far, we have bundled together design patents and utility patents, even though they protect different features of a product and have a differentiated legal treatment. In Table 2.8, we split the sample by patent type. It clearly appears that prices react to the expiry of utility patents (column 2) but not to the expiry of design patents (column 4).

The lack of effect for design patents does not mean that design rights are worthless. The visual and ornamental features of a product contribute to its positioning and, hence, to its price (Eisenman, 2013). The lack of effect may suggest that these decorative features continue to uniquely identify the product even after the design rights have expired, which helps to sustain higher markups. This mechanism would be similar to that observed on drugs, where the brand name helps to sustain high drug prices after patent expiry.

Table 2.8 – The effect of patent expiry on product prices by patent type

	(1)	(2)	(3)	(4)
	Utility patents		Designs	
Expired	-0.086***	-0.079***	0.020	0.014
	(0.008)	(0.007)	(0.020)	(0.022)
Product age (in months)	-0.003***	-0.002***	-0.002***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
New generation	-0.163***	-0.167***	-0.121***	-0.109***
	(0.010)	(0.011)	(0.035)	(0.037)
Month dummies	YES	YES	YES	YES
Control for price sources	YES	YES	YES	YES
Pair FE	YES		YES	
Patent FE		YES		YES
Product FE		YES		YES
Constant	5.618***	5.611***	4.544***	4.427***
	(0.007)	(0.008)	(0.020)	(0.016)
No. products	745	745	285	285
No. patents	2,417	2,417	361	361
No. pairs	14,055	14,055	566	566
Observations	466,331	466,331	25,005	25,005
R-squared	0.943	0.943	0.987	0.986

*Notes:*The dependent variable is P^A_{ijt}. Standard errors clustered at the product-patent-pair level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

There is a large amount of heterogeneity in patent value. As Lemley and Shapiro (2005, p. 85) put it, “many patents are virtually worthless,” either because they cover technology that is not commercially viable, because they are impossible to enforce effectively, or because they are very unlikely to hold up in court if litigated. However, "a small number of patents are of enormous economic significance." Patents in our sample form a highly selected set of inventions that are commercially relevant and, in all logic, more valuable than the average U.S. patent. Nevertheless, patents in our sample also exhibit heterogeneity in their value, as suggested by the four indicators of patent importance in Table 2.4.

In Table 2.9, we test whether the effect of patent expiry on product price differs by the importance of the expired patent. As explained in Section 2.4.3, we measure importance in four ways: whether a patent belongs to the top ten percent of forward citation distribution in our sample; the number of distinct IPC sub-classes in a patent; the number of independent claims in a patent; and the number of products protected by a patent. Because these patent quality measures do not really make sense for design patents, we remove them from the regression sample. Columns (1)–(4) report estimates with the product-patent fixed effect and columns (5)–(8) report estimates with patent and product fixed effects. We interact the patent expiry dummy with each importance variable, and report the p-value associated with the F-test of joint statistical significance of the patent expiry dummy and the interaction term.[18]

The coefficients associated with the interaction terms are negative but not statistically significant when it comes to the standard measures of patent importance (columns 1–3 and 5–7). The estimates of the interaction terms in columns (4) and (8) are particularly interesting. The expiry of multi-product patents is associated with a significantly higher drop in price. These patents might cover technologies that are fundamental to the firms, leaving them particularly exposed when they expire.

[18]Note that we cannot include the level of the importance variables as in a traditional interaction model due to the use of fixed effects—they are wiped out by the fixed effects.

Table 2.9 – The effect of patent expiry on product prices by patent importance

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Expired	-0.084***	-0.048	-0.072***	-0.021	-0.072***	-0.037	-0.064***	-0.023
	(0.010)	(0.030)	(0.014)	(0.018)	(0.008)	(0.026)	(0.012)	(0.017)
Expired × Top 10%	-0.007				-0.004			
	(0.016)				(0.013)			
Expired × log(IPC classes)		-0.050				-0.046		
		(0.038)				(0.034)		
Expired × log(Independent claims)			-0.012				-0.008	
			(0.011)				(0.009)	
Expired × log(Products per patent)				-0.020***				-0.015***
				(0.005)				(0.004)
F-Stat	56.155	56.265	55.853	195.852	55.274	55.354	54.973	194.261
Prob > F	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Product age (in months)	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
New generation	-0.163***	-0.163***	-0.163***	-0.162***	-0.163***	-0.163***	-0.163***	-0.162***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Month dummies	YES	YES	YES	YES	YES	YES	YES	YES
Control for price sources	YES	YES	YES	YES	YES	YES	YES	YES
Pair FE	YES	YES	YES	YES				
Patent FE					YES	YES	YES	YES
Product FE					YES	YES	YES	YES
Constant	5.618***	5.618***	5.618***	5.619***	5.618***	5.618***	5.618***	5.619***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
No. products	745	745	745	745	745	745	745	745
No. patents	2,417	2,417	2,417	2,417	2,417	2,417	2,417	2,417
No. pairs	14,055	14,055	14,055	14,055	14,055	14,055	14,055	14,055
Observations	466,331	466,331	466,331	466,331	466,331	466,331	466,331	466,331
R-squared	0.943	0.943	0.943	0.944	0.943	0.943	0.943	0.943

Notes: The dependent variable is P^A_{ijt}. Standard errors clustered at the product-patent-pair level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Product-level estimates

Up to this point, we have exploited the many-to-many relationship between patents and products by conducting the analysis at the product-patent level. The next table reports product-level estimates. When two or more patents protect a product, the $Expiry_{it}$ variable is a continuous variable defined on the [0,1] interval capturing the proportion of patents expired at time t.[19] Columns (1)–(3) of Table 2.10 report the estimates for multi-patent products only whereas columns (4)–(6) consider all products. We find that product prices decline by 0.036–0.59 percent with a ten percentage point increase in the proportion of expired patents.

[19] The mean of $Expiry_{it}$ is 0.19 and the standard deviation is 0.39.

Table 2.10 – Product-level regression on the effect of patent expiry

	(1)	(2)	(3)	(4)	(5)	(6)
	Products with more than one patent			All products		
Expired	-0.109***	-0.064**	-0.059**	-0.091***	-0.042**	-0.036**
	(0.025)	(0.027)	(0.027)	(0.016)	(0.018)	(0.018)
Product age (in months)		-0.001***	-0.002***		-0.001***	-0.002***
		(0.000)	(0.000)		(0.000)	(0.000)
New generation		-0.147***	-0.154***		-0.142***	-0.155***
		(0.037)	(0.036)		(0.035)	(0.034)
Month dummies			YES			YES
Control for price sources			YES			YES
Product FE	YES	YES	YES	YES	YES	YES
Constant	4.490***	4.579***	4.622***	4.041***	4.126***	4.170***
	(0.004)	(0.018)	(0.030)	(0.003)	(0.014)	(0.024)
No. of Products	602	602	602	825	825	825
Observations	23,147	23,147	23,147	35,425	35,425	35,425
R-squared	0.982	0.983	0.983	0.984	0.984	0.984

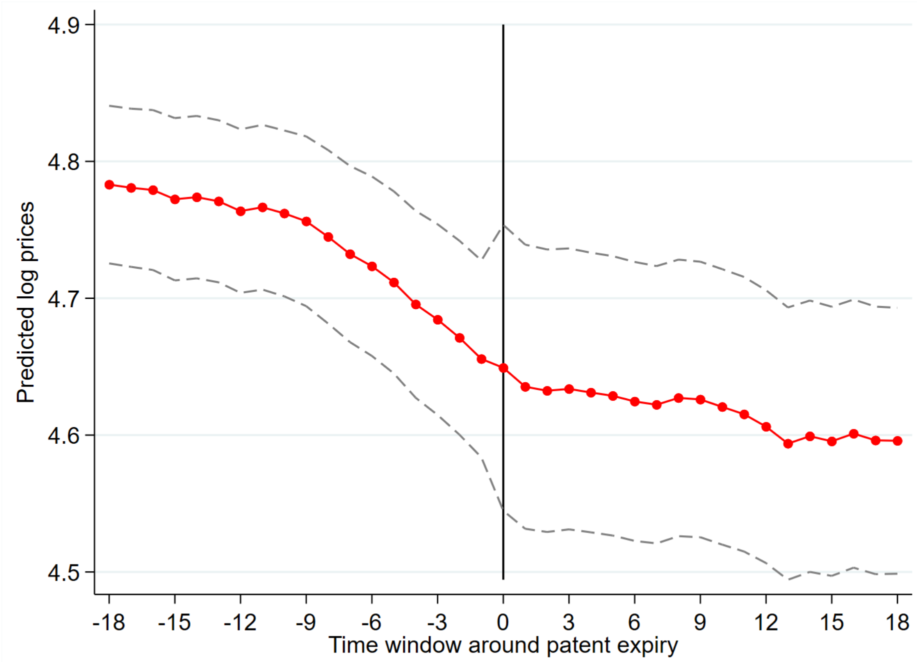
Notes: The dependent variable is P^A_{ijt}. Standard errors clustered at the product-patent-pair level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Temporal effect

In this section, we explore the temporal dimension of the decline in price with a view to shed light on its underlying mechanism. If prices react to imitation, the price drop should occur *after* patent expiry, when the invention is no longer protected and competing firms enter the market. By contrast, we expect a preemptive price decrease by the incumbent to manifest itself *before* patent expiry.

We use an event-study method and restrict our sample to observations that fall into an 18-month time window before and after patent expiry. We add a dummy variable for each leading/lagging month around patent expiry to our baseline model (corresponding to column 3 of Table 2.5). We then recover the predicted log price from the regression coefficients. The result, shown in Figure 2.6, indicates that the decline in prices starts about one year prior to patent expiry. It then seems to stabilize shortly after patent expiry. The drop in predicted prices over the ten months that precede expiry reaches about 10.40 percent ($(e^{4.76} - e^{4.65})/e^{4.76}$).

Figure 2.6 – Evolution of product prices around patent expiry



Notes: The sample is restricted to product-patent pairs that are either active or expired. The dashed lines depict the upper and lower bounds of the 95-percent confidence interval.

We see two possible reasons for the effect. It could result from strategic entry deterrence by the incumbent (Milgrom & Roberts, 1982; F. M. S. Morton, 2000; Goolsbee & Syverson, 2008). In this scenario, the incumbent proactively reduces the price to lower market attractiveness for would-be competitors. Alternatively, competitors might enter the market shortly before patent expiry, betting that the incumbent will not start a costly and lengthy infringement case. In the absence of a time-varying competition variable, we are left with conjectures.

2.5.2 Accounting for product market competition

So far, we have established that product prices react negatively to patent expiry. The price drop is more significant for more important patents and starts about one year before the actual expiry. In this section, we test the extent to which the price drop reacts to competitive pressure.

In theory, patent protection allows the innovator to fend off competition. Therefore, in markets with no-to-limited competitive pressure, the effect of patent protection must be limited. As competitive pressure increases, patent protection surely becomes more valuable to the firm. However, as competition further increases, the effect becomes ambiguous. On the one hand, in hyper-competitive markets competitors may have developed substitute technologies or may have invented around the patent to render the patent protection essentially useless. On the other hand, patent protection may be all the more important in such markets.

Although we cannot observe when a substitute product is launched or when a competitor enters the market, we have information on the current market structure. We proxy market competitiveness for each product in two ways: with the number of similar products sold by competitors ($Substitutes$) and with the number of competing firms selling the substitutes ($Competitors$).

To examine how market competitiveness moderates the effect of patent expiry, we interact patent expiry with each of the competition measures as well as their squared terms. Columns (1)–(4) of Table 2.11 report estimates with the product-patent fixed effect and columns (5)–(8) with patent and product fixed effects. We find that, by and large, the intensity of competition exacerbates the effect of patent expiry on price. However, the squared terms are positive and statistically significant, suggesting a U-shaped relationship.

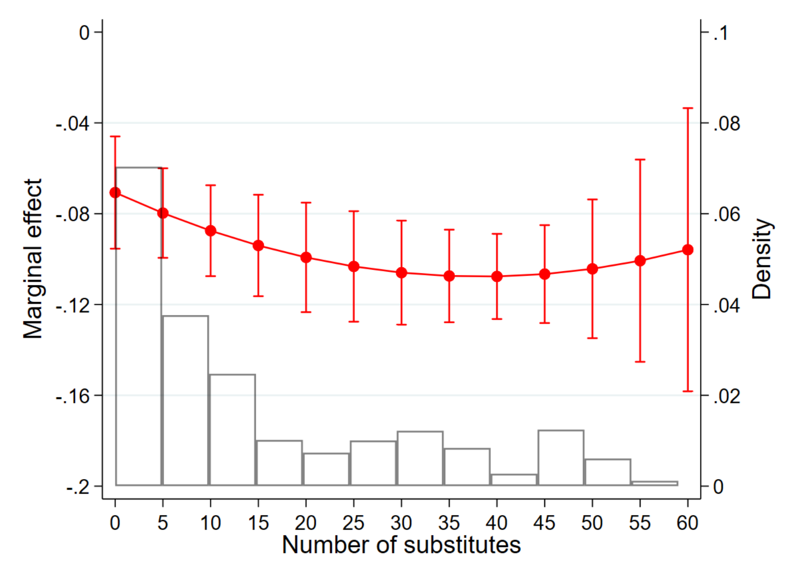
Figure 2.7 depicts the non-monotonic effect of competition on price using the models in columns (2) and (4). It also reports the distribution of substitute products and competitors in the sample. Most observations fall in the downward-sloping part of the effect, meaning that competition usually exacerbates the pressure on prices. The peak is reached at about 40 substitute products and ten competitors.

Table 2.11 – Patent expiry and competition

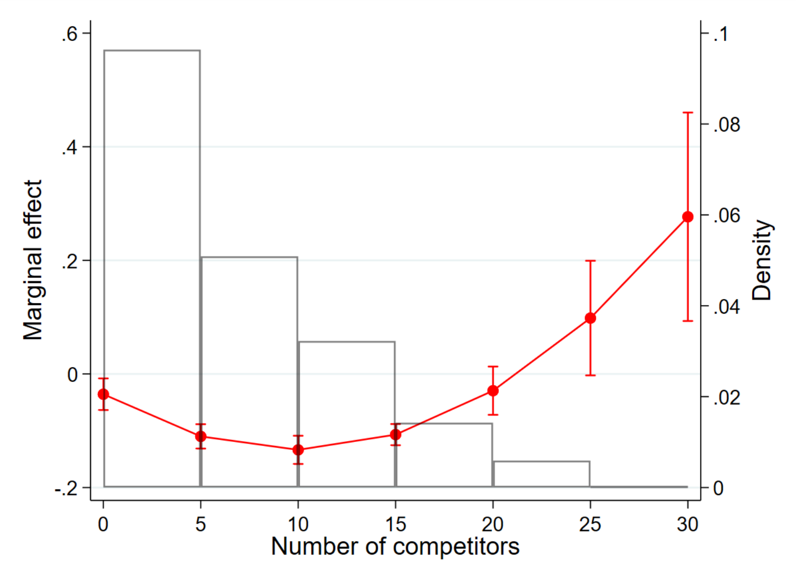
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Expired	-0.069***	-0.058***	-0.063***	-0.025*	-0.067***	-0.063***	-0.065***	-0.055***
	(0.011)	(0.012)	(0.012)	(0.014)	(0.007)	(0.007)	(0.007)	(0.007)
Expired × Substitutes	-0.001**	-0.003**			-0.000**	-0.001***		
	(0.000)	(0.001)			(0.000)	(0.000)		
Expired × Substitutes^2		0.000*				0.000**		
		(0.000)				(0.000)		
Expired × Competitors			-0.003***	-0.021***			-0.001***	-0.005***
			(0.001)	(0.004)			(0.000)	(0.001)
Expired × Competitors^2				0.001***				0.000***
				(0.000)				(0.000)
Product age (in months)	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***	-0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
New generation	-0.161***	-0.160***	-0.160***	-0.160***	-0.161***	-0.161***	-0.161***	-0.161***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
Month dummies	YES	YES	YES	YES	YES	YES	YES	YES
Control for price sources	YES	YES	YES	YES	YES	YES	YES	YES
Pair FE	YES	YES	YES	YES				
Patent FE					YES	YES	YES	YES
Product FE					YES	YES	YES	YES
Constant	5.560***	5.560***	5.560***	5.560***	5.560***	5.560***	5.560***	5.560***
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
No. products	825	825	825	825	825	825	825	825
No. patents	2,778	2,778	2,778	2,778	2,778	2,778	2,778	2,778
No. pairs	14,621	14,621	14,621	14,621	14,621	14,621	14,621	14,621
Observations	491,336	491,336	491,336	491,336	491,336	491,336	491,336	491,336
R-squared	0.947	0.947	0.947	0.947	0.947	0.947	0.947	0.947

*Notes:*The dependent variable is P^A_{ijt}. Standard errors clustered at the product-patent-pair level in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Figure 2.7 – Non-monotonic effects on price of the number of substitutes (a) and competitors (b).



(a)



(b)

Notes: 95-percent confidence intervals reported.

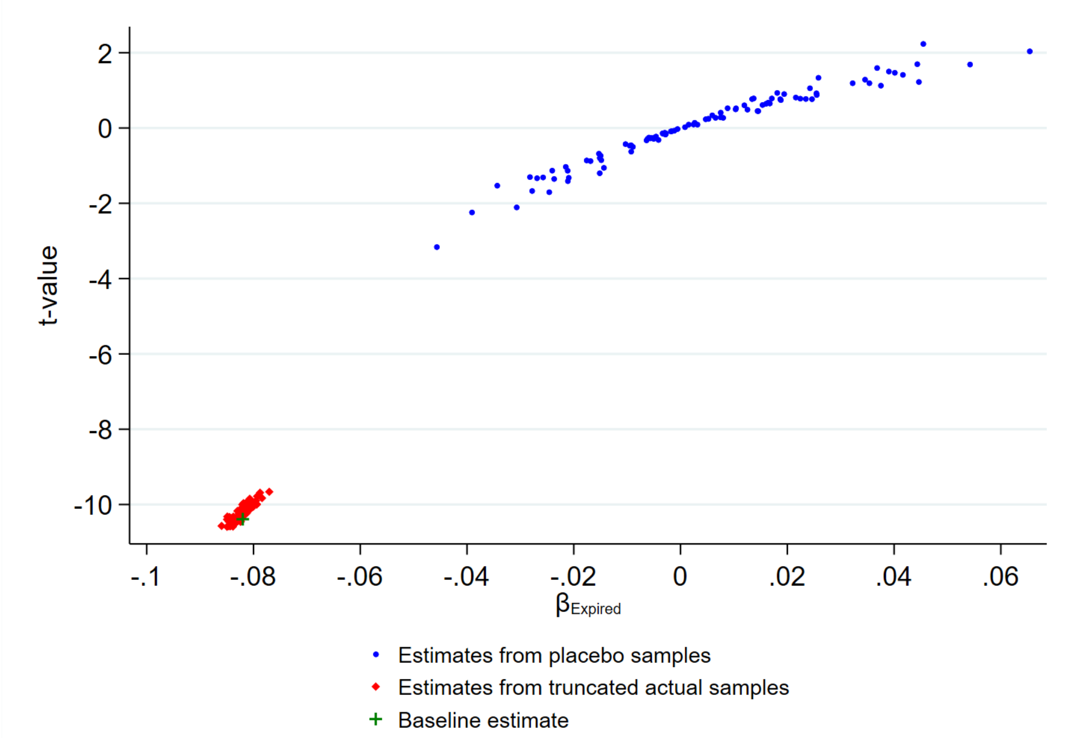
2.5.3 Placebo tests

We have accounted for confounding factors that are likely to threaten the validity of our estimates, namely product age and the introduction of a new product generation (as well as the fixed effects). In this section, we implement two placebo tests to further assess the validity of our results.

The first placebo test focuses on the sample of observations associated with patents that were active throughout the study period. We create fake expiry events on a random set of these patents and re-estimate equation (4.1). In all logic, these fake expiry events should not have any effect on product prices. We perform 100 estimates, every time randomly assigning fake expiry dates on a randomly selected set of active patents. We randomly select 17 percent (to mimic actual data) of the active patents and assign each of them an expiry date that is randomly and uniformly distributed between January, 2011 and April, 2019, leaving the other active patents unchanged. We plot the $\hat{\beta}$'s and the corresponding t-values associated with the placebo $Expired$ variable.

Next, we want to compare the placebo $\hat{\beta}$'s with the $\hat{\beta}$ estimated with the real data. Directly comparing the two quantities would be unfair, however, because sample sizes differ. Consequently, we also estimate equation (4.1) on the real data but randomly dropping 28 percent of the observations—so the placebo and the real-but-truncated samples are of similar size. Figure 2.8 reports the estimated $\hat{\beta}$'s as well as their t-values. The coefficients estimated with the placebo samples are typically insignificant with roughly half of them being positive. By contrast, the $\hat{\beta}$'s estimated from the randomly reduced samples are scattered closely around the baseline $\hat{\beta}$ (-0.082).

Figure 2.8 – Kernel density of $\hat{\beta}$'s estimated from placebo tests



Notes: Estimates marked by dots come from placebo samples whereas estimates marked by diamonds come from truncated actual samples. The cross reports the $\hat{\beta}$ and t-value obtained from the baseline model as in column (3) of Table 2.5.

In the second placebo test, we focus only on patents that have expired. We assign fake expiry events prior to the true expiry. To ensure a similar number of observations across the various samples, we restrict the samples to product-patent pairs observed within a one-and-a-half-year time window around the actual or the placebo expiry date. We then set the placebo expiry events to two and a half years, three years, and three and a half years prior to the actual expiry month, respectively. In column (1) of Table 2.12, the effect of true expiry event is statistically significant and its magnitude is close to the baseline estimate. By contrast, the coefficients estimated on the placebo expiry events are significantly positive as in columns (2)–(4). These figures confirm that the price decline that we observe once the patent expires does not merely reflect the effect of the passing of time.

To sum up, results from the placebo tests confirm that there is a genuine price drop that occurs around the time of patent expiry. We are not concerned by the possibility that our research design might drive the results.

Table 2.12 – Placebo test: the effect of patent expiry

	(1)	(2)	(3)	(4)
Expired	-0.087***			
	(0.024)			
Expired (2.5 years before)		0.210**		
		(0.088)		
Expired (3 years before)			0.244***	
			(0.084)	
Expired (3.5 years before)				0.120**
				(0.047)
Product age (in months)	-0.003***	-0.005***	-0.005***	-0.004***
	(0.001)	(0.001)	(0.001)	(0.001)
New generation	-0.096***	-0.287***	-0.320***	-0.228***
	(0.022)	(0.063)	(0.073)	(0.059)
Month dummies	YES	YES	YES	YES
Control for price sources	YES	YES	YES	YES
Pair FE	YES	YES	YES	YES
Constant	4.773***	4.524***	4.481***	4.419***
	(0.036)	(0.056)	(0.056)	(0.058)
No. products	111	96	89	82
No. patents	110	126	111	100
No. pairs	398	261	229	198
Observations	22,794	15,868	14,001	11,966
R-squared	0.956	0.893	0.885	0.915

Notes: The dependent variable is P^A_{ijt}. Standard errors clustered at the product-patent-pair level in parenthesis. *** p<0.01, ** p<0.05, * p<0.1.

2.6 Concluding remarks

This chapter provides evidence of the monopoly pricing power of patents. Specifically, it reveals a drop in the price of a sample of consumer products listed on Amazon.com around the time they lose patent protection. We find that patent expiry is associated with a 7–8 percent drop in product price. As far as we can ascertain, the present chapter is the first to report direct evidence of a markup for patent-protected consumer products. We have achieved this result thanks to a novel way of identifying the correspondence between patents and products. The results complement earlier (and mixed) findings related to patent-protected drugs (Caves et al., 1991; Grabowski & Vernon, 1992; Frank & Salkever, 1997; Wiggins & Maness, 2004; Vandoros & Kanavos, 2013) as well as copyrighted books (Li et al., 2018; I. Reimers, 2019)—an admittedly distinct type of intellectual property right.

The empirical analysis produces insights that inform us about the possible mechanisms at play.

We observe that prices start to decline about one year before actual patent expiry. This result is consistent with a preemptive price reduction by the incumbent with a view of deterring market entry as well as with other mechanisms. The price decline is also greater in more competitive markets, as measured by the number of offerings in a similar product category as well as by the number of firms proposing such offerings.

The econometric results also pass a series of sanity and robustness tests. We find that the price drop is larger for more important patents, as proxied by the number of products that the patent protects. We also find that product prices react only to the expiry of utility patents and not design patents. Design patents do not undergo a substantive examination and, therefore, offer a weaker form of protection. Finally, the results are robust to a range of alternative specifications and placebo tests of fake expiry events.

In passing, we were also able to compute, for the first time, statistics about the link between products and patents. Notably, we found that it takes on average six and half years before a patented invention is first commercialized. We found that patented products *de facto* enjoy an exclusivity of maximum 15 years (i.e., until the last patent in a product expires) from the time they are first released on the market.

The policy implications of the findings are clear: patents seem to provide *some* level of protection in the product market, thereby providing evidence that the patent system helps sustain supra-competitive prices for innovators. This finding represents an important step in our understanding of the functioning of patent systems. The 7–8 percent figure sheds light on the markup enjoyed by incumbent innovators. It is a measure of the welfare loss associated with the patent monopoly described in theoretical models—or, in other words, the subsidy rate paid by consumers. Nevertheless, the net welfare benefit (or cost) of monopoly pricing is beyond the scope of this chapter due to the constraint of data. Future research should find ways to observe the markup throughout the entire duration of patent life and combine it with sales data to estimate the patent premium. Such estimates should then be contrasted with the R&D cost associated with the underlying products in order to quantify the magnitude of the incentive effect that the patent premium represents. There is still a long way to go before fully understanding the welfare effects of the patent system, but we hope that the present paper will enable follow-on research on this topic.

3 Public Notice and Invention Diffusion

This chapter is written in collaboration with Gaétan de Rassenfosse.

Abstract:

This paper examines the effects of firms' disclosure of innovative portfolios on the diffusion of inventions. Firms signal the commercial value of inventions through the provision of constructive notice. We focus on a change in the marking statute, the virtual patent marking practice, and collect a sample of 843 virtually marked patents from 16 firms. By exploiting the time-stamped variations of when patents were listed in web documents, we compare the changes in external citations before and after marking. We find that while marking attracts more follow-on inventions for patents in weak regimes, competitors shun the marked inventions in strong regimes; moreover, the similarity between a citing patent and the focal patent also reduces for patents in strong regimes after marking. Patent importance also plays a role in exacerbating follow-on inventions for patents in strong regimes.

Key words: disclosure, virtual patent marking, diffusion, invention, similarity

3.1 Introduction

Conventional wisdom holds that the disclosure of information about R&D activities increases social welfare by encouraging cumulative innovation, improving efficiency in the market for ideas, and reducing duplicative R&D (Scotchmer, 1991; Gans et al., 2008; Lück et al., 2020). For firms, disclosing their R&D activities reduces information asymmetry with their investors, thereby facilitating equity financing (Botosan, 1997; Healy & Palepu, 2001; Lerner et al., 2011; He & Tian, 2013; Aggarwal & Hsu, 2014). Disclosure is also a pre-requisite to secure patent rights for their inventions (Ouellette, 2011). However, the revelation of valuable information comes at a cost, including a reduction in the private value of the information by facilitating imitation (Bhattacharya & Ritter, 1983; Anton & Yao, 2004). This core trade-off is well understood in the literature. But open questions remain. One of them is the extent to which information disclosure hinders or spurs follow-on inventions.

There is considerable heterogeneity in how much information firms disclose through patent documents. Key inventions can be buried in vast portfolios of hard-to-read documents, leading some scholars to claim that disclosure in patent law should be strengthened (Ouellette, 2011). Moreover, the information revealed in a patent document does not inform about which technologies are commercially relevant (Kitch, 1977). Some firms, however, go further in the disclosure of valuable information by explicitly signaling their important patents to competitors. They do so by means of 'patent marks', which provide constructive notice to the public that products are patented. Traditional patent marks take the form of patent number(s) physically printed on a product or its packaging; virtual patent marks take the form of web addresses that list the relevant product(s) and patent(s). Theoretically, the effect of marking on follow-on inventions is ambiguous. On the one hand, patent marks signal the commercial value of specific inventions, thus attracting the attention of competitors. On the other hand, patent marks clearly delineate innovative firms' intellectual property (IP) boundaries, thereby potentially pushing away competitors from the firm's turf. In addition, the effectiveness of patents as an appropriation tool may impose frictions on how competitors exploit the patented knowledge such that competitors' response to the marked inventions may differ according to the strength of appropriability regimes. It remains unclear under what conditions disclosure facilitates or impedes diffusion. As far as we ascertain, no study has ever investigated the effect of patent marking on invention diffusion.

The present paper fills this gap. It empirically studies how patent marking affects follow-on inventions in different appropriability regimes. The data comes from virtual patent marking (VPM) web pages and covers 843 patents virtually marked between 2011 and 2018. The key outcome variables for measuring invention diffusion are the count of forward patent citations coming from competing firms as well as state-of-the-art measures of patent similarity.

The empirical analysis exploits the within-patent change in marking status. Specifically, we implement a fixed-effect patent-level panel regression model in which the variable of interest is a dummy variable that signals *when* a patent has been publicly disclosed as protecting a product. We then track changes in the arrival rate of citations and changes in the similarity of patent documents following the change in marking status.

We find that signaling valuable patents is associated with more follow-on inventions in weak regimes. We also find that these follow-on patents become closer content-wise to the focal marked patents. Thus, it seems that competitors rush into the research path signaled by the firm. However, patent marking seems to deter competition in strong regimes: marked patents attract fewer follow-on inventions and become more dissimilar in nature. In addition, we also find that the negative effect is mitigated by patent importance: the effect of virtual marking in a strong regime is weaker for the more important patents. These findings suggest that virtually marking inventions may help innovative firms fend off spillovers from competitors when firms rely on patents as an effective appropriation tool.

We lay out the background on virtual patent marking and develop hypotheses in section 3.2. Section 3.3 describes the identification strategy and the data. Section 3.4 presents the regression models. Section 3.5 discusses the econometric results and section 3.6 concludes.

3.2 Background

3.2.1 Virtual patent marking

Title 35 of the U.S. Code Section 287, also known as the "marking statute", permits patentees to physically label products with the identification numbers of the patents that protect them, alongside the prefixes "patent" or "pat". The 2011 Leahy-Smith America Invents Act (AIA) amended the marking statute with the introduction of virtual patent marking, allowing patentees to mark products using a web link that provides information on patents protecting them.[1]

The purpose of patent marking is to prevent innocent infringement by providing constructive notice to the public that an article is patented. Firms that have marked their products can claim wilful infringement in case of litigation, which leads to treble damages—thereby providing firms with a powerful incentive to mark their products. Note, however, that firms cannot strategically manipulate patent marks. Should they list patents that, in fact, are not associated with the product, firms expose themselves to false marking suits. Thus, the patent marking statute encourages firms to signal valuable patents accurately. Overall, it is clear that patent

[1] See https://www.logitech.com/en-us/about/virtual-patent-marking.html for an example of a virtual patent marking web page.

marks lower competitors' search costs for identifying inventions of important commercial values.

As alluded to, there are two ways of marking: physically or virtually. Patent marks are difficult to track in scale as no database or repository exists. However, data on virtual marks are considerably easier to collect than data on physical marks since they are available online. Another advantage of virtual marks over physical marks is that we can observe the timing of the marks, which generates variations that we will exploit in the empirical analysis. Among the many VPM web documents that we found, few were time-stamped, owing to a lack of standardized formatting (de Rassenfosse & Higham, 2020). We turn our focus to patents listed in time-stamped VPM web documents. Figure 3.1 provides an example of Dyson's VPM document, where marking dates are clearly recorded as new patents are added.

Figure 3.1 – An example of a VPM document from Dyson

AM05 *This product is protected by the following patents and designs.* *Information correct as of* 22 February 2016 **Patents** US7931449; US7972111; US8052379; US8092166; US8197226; US8308432; US8348596; US8348597; US8348629; US8366403; US8403640; US8403650; US8430624; US8454322; US8469655; US8529203; US8708650; US8714937; US8734094; US8764412; US8784071; US8873940; US8932028; US9004878; US9249810 **Designs** USD598532; USD602144; USD643098; USD672023	**AM05** *This product is protected by the following patents and designs.* *Information correct as of* 20 December 2017 **Patents** US7931449; US7972111; US8052379; US8092166; US8197226; US8308432; US8348596; US8348597; US8348629; US8366403; US8403640; US8403650; US8430624; US8454322; US8469655; US8529203; US8708650; US8714937; US8734094; US8764412; US8784071; US8873940; US8932028; US9004878; US9249810; US9513028; US9745996; US9822778 **Designs** USD598532; USD602144; USD643098; USD672023

3.2.2 Marking and follow-on inventions

Existing knowledge is an essential input in the innovation process (Scotchmer, 1991). Disclosure of technical information facilitates the diffusion of knowledge and enables competitors to learn from the disclosing firm. A stream of recent works provides empirical evidence that disclosure through patent documents spurs follow-on invention. For example, Furman et al. (2018) find that the expansion of patent libraries across the United States between 1975 to 1997 increases local patenting by 17 percent and boosts local business and job creations. Taking American Inventors Protection Act (AIPA) as a natural experiment, Hegde et al. (2018) find that early disclosure of patent documents increases knowledge spillovers and follow-on patenting. Baruffaldi and Simeth (2020) further show that although AIPA increases knowledge flows measured by forward citations, the degree of diffusion is confined by the existing geographical and technological boundaries. On the contrary, another set of studies examine the effect of withholding technical information on follow-on inventions. Gross (2019) studies about 11,000 patents applied during World War II under USPTO's patent secrecy program. His findings suggest that a shorter secrecy term leads to a higher probability of getting cited, especially for patents of non-governmental interests; moreover, secrecy leads to less commercialization of inventions. Exploiting the Invention Secrecy Act on a sample of patents filed between 1982 and 2000, de Rassenfosse et al. (2020) find that the enforcement of a secrecy order has a significant negative effect on follow-inventions and the hindering effect is particularly salient for geographically distant inventions.

Patent marking provides an additional layer of disclosure. It signals which of the patents in a firm's portfolio have the most substantial commercial value, thereby attracting the attention of competitors to build on these valuable inventions. At the same time, however, the disclosing firm strengthens patent protection by laying the ground for treble damages in case of infringement, potentially threatening competitors not to imitate. Thus, a priori, the effect of patent marking on follow-on inventions remains ambiguous.

Marking in weak and strong appropriability regimes

We contend that the effect of marking on follow-on inventions depends on the strength of the 'appropriability regime' (Teece, 1986).

In strong regimes, IP protection is one of the preferred modes of appropriation. By contrast, in weak regimes, innovating firms resort to other means than formal IP rights to capture the returns from innovation. These other means include lead time and marketing capabilities (Cohen et al., 2000). The reason for which weak regimes are called 'weak' is precisely because patents do not offer strong protection for firms to appropriate returns in such regimes. The boundaries of patent rights are less clearly delineated in weak regimes than in strong regimes (think of software patents vs. drug patents), making it easier for competitors to invent around patented inventions.

Note that variations in the strength of the appropriability regime may come from cultural, legal, and technological factors. Regarding cultural and legal aspects, Cohen et al. (2002) find that intra-industry R&D spillovers are greater in Japan than in the United States, which they explain by differences in the appropriability of patented inventions among competitors. They argue that appropriability is weaker in Japan than in the United States, giving Japanese firms more leeway in exploiting competitors' inventions. The present analysis focuses on one country, namely the United States. The strength of the appropriability regime differs across technological groups (Mansfield, 1986; Cohen et al., 2000), making patent protection generate a higher premium in some industries than in others (Arora et al., 2008).

Because patents do not offer an iron-clad protection against imitation in weak regimes, signaling important inventions by way of patent marking will attract competitors and, therefore, facilitate follow-on inventions. From a resource-based perspective, weak regimes reduces firms' appropriability on cumulative inventions as competitors can easily build on the focal invention (Ahuja et al., 2013). Thus, we hypothesize:

Hypothesis 1a. *In weak appropriability regimes, patent marking is associated with an increase in follow-on inventions.*

Furthermore, competitors may rush into the research direction signaled by the innovative firm. We expect to see followers working in similar technologies to the focal invention. Thus, we pose the following hypothesis:

Hypothesis 1b. *In weak appropriability regimes, patent marking is associated with an increase in the similarity of follow-on inventions with the marked patent.*

Having established this hypothesis, it seems surprising that firms in weak appropriability regimes would practice patent marking. One must not forget that marking secures higher damages in case of infringement and may also benefit the firm as a marketing tool (by conveying to consumers the innovative nature of the products). Finally, it is also possible that firms may not be aware of the potential adverse effects of marking highlighted herein.

Turning now to the effect of patent marking in strong regimes, we expect a decrease in follow-on inventions for exactly the opposite reason. Although the innovative firm signals its valuable inventions, patent protection as an effective means of appropriation may prevent knowledge from flowing to competitors. Marking increases the cost of infringement (de Rassenfosse, 2018), giving further impetus for competitors to shy away from the protected area. We hypothesize:

Hypothesis 2a. *In strong appropriability regimes, patent marking is associated with a decrease in follow-on inventions.*

Hypothesis 2b. *In strong appropriability regimes, patent marking is associated with a decrease in the similarity of follow-on inventions with the marked patent.*

Marking of important patents

Patents significantly differ in their value, with a large number of patents being "worthless" (Moore, 2005; Duffy, 2018). This assertion is supported by surveys of patent owners, which consistently find that many patents are not being used (Blind et al., 2006; De Rassenfosse, 2012; Torrisi et al., 2016), or by hard data documenting low litigation rates and licensing rates

(Lemley, 2000).

A patent mark signals that the underlying invention is commercialized as associated with one or more products. Therefore, when considering the overall patent portfolio of a firm, we expect that patents that are marked are on average more valuable than patents that are not. We form the following hypothesis:

Hypothesis 3a. *Firms mark their most important patents.*

Patents that are more important naturally attract more follow-on inventions because of the inherent value. Therefore, once an innovative firm puts the spotlight on selected patents by marking them online, we expect more follow-in inventions for the most important of these patents. In weak regimes, patent importance may enhance the diffusing effect of virtual marking. In strong regimes, patent marks may not serve as a strong deterrence for competitors to build on the most important patents. Thus, patent importance may mitigate the negative effect of virtual marking on follow-on inventions. Therefore, we hypothesize:

Hypothesis 3b. *Marking increases follow-on inventions for the valuable important inventions.*

3.3 Data and descriptive statistics

The empirical analysis seeks to quantify the extent to which competitors rush into promising research directions signaled by the focal firm or whether they shy away from it. In order to quantify these effects, we need two critical pieces of information: an indication of the 'signal' as well as measures of follow-on inventions.

3.3.1 Patent marks

Patents listed in VPM web documents form our signal of important inventions. Since we need to observe when patent marks were communicated to the public, it is essential for us to observe time-stamped VPM, as illustrated in Figure 3.1.

Our sample comprises 16 firms that virtually marked patents with clearly recorded marking dates between March 2010 and April 2018.[2] Data on patent marks were recovered from 45 VPM documents belonging to these firms. Firms in our sample cover the sectors of consumer goods, drugs & medical devices, as well as electronics & mechanics. We manually collected information on the marked patent numbers and the marking dates. We removed design

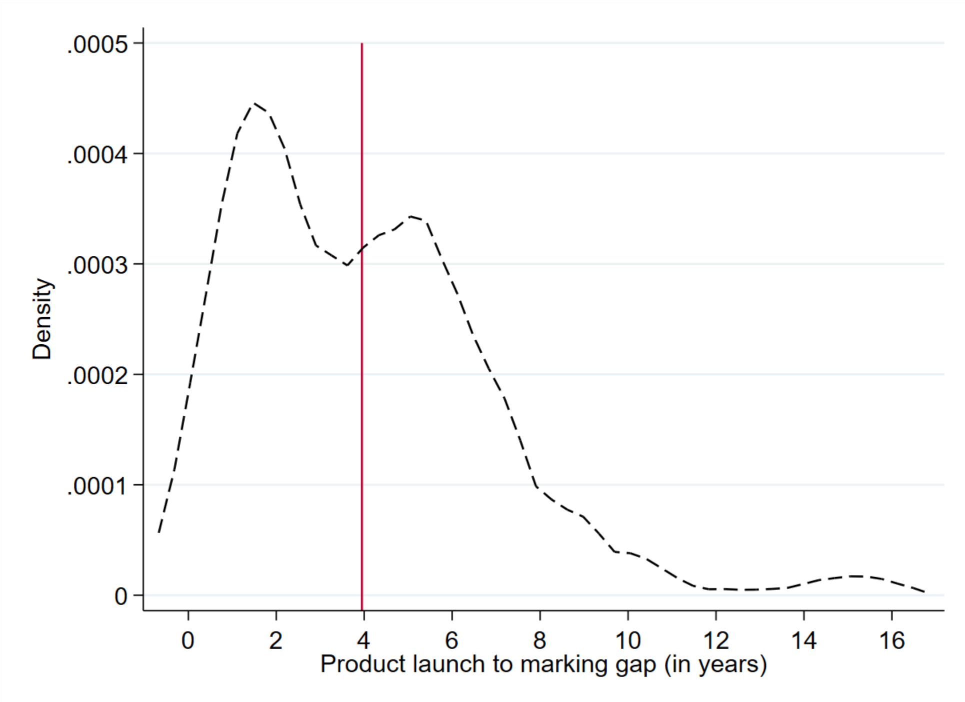
[2]Note that only two patents from Amgen, No. 5871740 and No. 7217689, were marked in March 2010. The rest of the patents were marked starting from 2011 November.

patents because they have different citation patterns than utility patents. The sample consists of 843 virtually marked patents.

Surveys on U.S. firms and research labs suggest that the pharmaceutical and medical equipment industries rely on patents as an effective appropriability mechanism (Mansfield, 1986; Cohen et al., 2000). We categorized firms into any of the two appropriability regimes. Firms in the pharmaceutical and medical device sector are in the strong appropriability regime, whereas all the other firms fall in the weak regime.

For every marked patent in our sample, we manually collected information on the earliest launch date of the corresponding product. If a patent is associated with more than one product, we selected the earliest launch date among all products. Figure 3.2 depicts the distribution of the time difference between the product launch date and the marking date. On average, patents are virtually marked about four years after product launch for the sample of products.

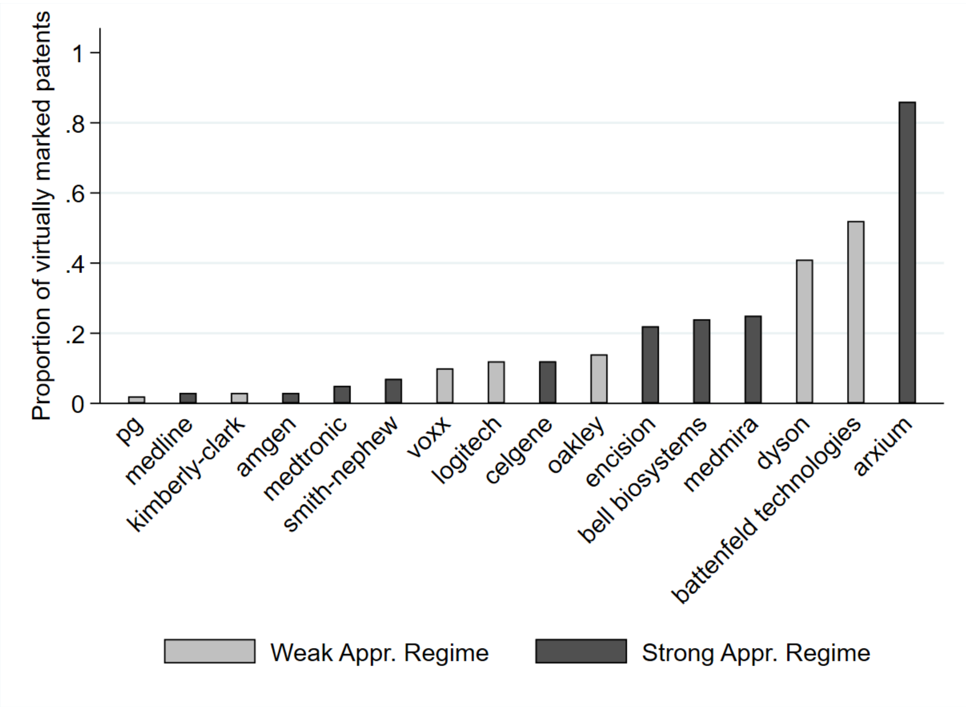
Figure 3.2 – Distribution of the timelag between the product launch date and the virtual marking date



We also extracted data on the patent portfolios for the 16 firms in our sample using PatentsView. Figure 3.3 shows the proportion of patents that are virtually marked by assignee and the strength of the appropriability regime. The proportions vary from less than 5 percent to more than 80 percent. Firms with large patent portfolios such as P&G, Kimberly-Clark, and

Medtronic have a relatively low share of patents marked.[3] There are two explanations for the varying proportion of marking practice across firms. First, the introduction of VPM was recent, and its adoption has been slow—about 12 percent of patent-holding firms have marked patents virtually as of 2018 (de Rassenfosse, 2018). In our sample, for example, Medtronic didn't start marking its patents until 2018. Second, a significant portion of patents remain commercially unused; these patents are either filed for strategic purpose or are so-called “sleeping" patents (Torrisi et al., 2016).

Figure 3.3 – Proportion of virtually marked patents by assignee and appropriability regime



3.3.2 Measuring follow-on inventions

Technological progress builds on existing innovation, but the intangible nature of innovation makes it challenging to gauge the cumulativeness of knowledge. The number of citations from subsequent inventions received by a patent depicts the trails of knowledge diffusion and, consequently, is used frequently as a measure of follow-on inventions (Jaffe et al., 1993; Belenzon, 2012; Galasso & Schankerman, 2015; Thompson & Kuhn, 2020). However, forward citations have limitations in capturing knowledge flows because they embed measurement errors from patent examination procedures or firms' citing strategies or because they tend not to reflect knowledge flows from basic scientific research (Alcácer et al., 2009; Sampat, 2010;

[3] P&G, Kimberly-Clark, and Medtronic are the top three firms with the largest patent portfolios in our sample, each having more than 5000 patents

Roach & Cohen, 2013; Bryan et al., 2020; Kuhn et al., 2020). Notwithstanding these limitations, we follow previous scholars in the field and rely on citations as our primary measure follow-on inventions. We source data from PatentsView and track citations coming from external firms; our measure excludes self-citations.

To alleviate some of the concerns raised in the literature, we also derive additional variables of citations. We track assignees citing the focal marking firm for the first time. That is, we exclude assignees that have previously cited any patents of the focal marking firm. The first variable is the number of citations from new firms, and the second is a binary indicator on whether a citation comes from a new firm, both counted in a six-month time window. These indicators are less granular than the total count of citations and are thus less impacted by strategic and legal considerations. We use these data to test H1a and H2a.

We characterize the nature of citing patents using a content-based measure of invention similarity. We extract the vectors of word embeddings for each citing and marked patents from the Google Patent Database on BigQuery. We then use these vectors to compute the cosine similarity score for each pair of a marked patent and a follow-on citing patent.[4] We use these data to test H1b and H2b.

Figure 3.4 plots the average number of citations received by patents from firms in the sample. Panel (a) reports self-citations, whereas panel (b) reports external citations, our primary dependent variable. The figure further differentiates between citations received by marked and not-marked patents for firms in strong and weak appropriability regimes.

Panel (a) offers a test of H3a. It suggests that firms mark their most important patents. Marked patents received significantly more self-citations than non-marked patents, implying that firms are more likely to build internal follow-on inventions on the marked ones. On average, self-citations are fewer than external citations for both marked and not marked patents.

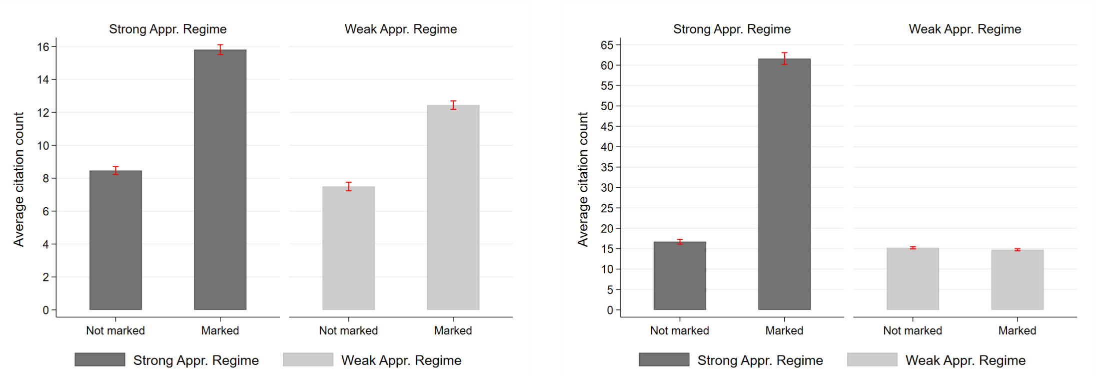
Turning to panel (b), the contrast between strong and weak appropriability regimes is particularly striking. In strong regimes, marked patents attract significantly more external citations than non-marked patents. In contrast, the differences in average citations between marked and non-marked patents are not statistically significant in weak regimes. However, this figure is silent competitors' response *once a patent becomes publicly marked.*

The next figure provides a first look at hypotheses H1a and H2a. It presents the average number of citations received in six-month intervals for the sample of marked patents before and after getting marked in both appropriability regimes. Panel (a) of Figure 3.5 reports the figure for self-citations for comparison purposes. Panel (b) shows that the average number of citations

[4] The Google word embeddings use WSABIE algorithms, see Weston et al. (2010) for more information. A recent example using cosine similarity score of word-embeddings can be found in de Rassenfosse et al. (2020).

Figure 3.4 – Average number of citations received by marked and not marked patents, by appropriability regime

(a) Self citations (b) External citations

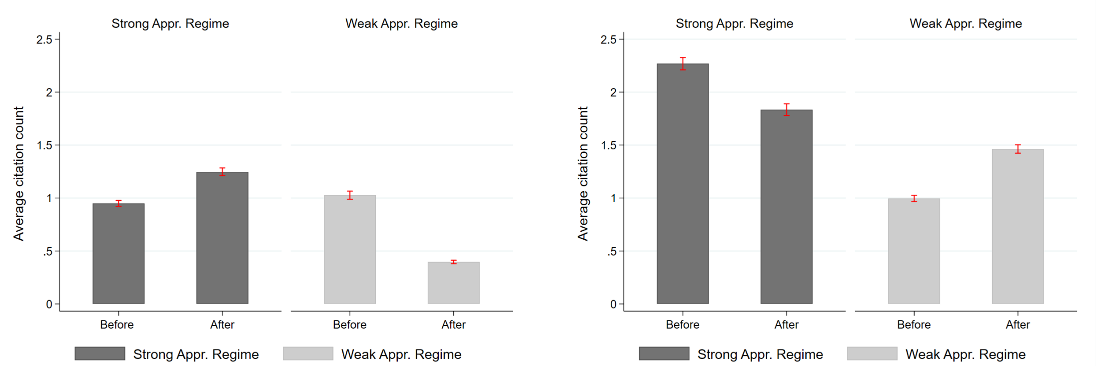


Notes: The bar plot depicts the average number of self (panel a) and external (panel b) citations and the errors bar represent the 95% confidence interval.

drops from 2.27 to 1.83 for patents in strong appropriability regimes, whereas that increases from 1.00 to 1.46 for patents of a weak appropriability regime. This finding represents *prima facie* evidence that patent marking may have different effects on the diffusion of inventions to external firms depending on the strength of the appropriability regime.

Figure 3.5 – Average citation count before and after marking by appropriability

(a) Self citations (b) External citations



Notes: The bar plot depicts the average count of self citations and external citations in a six-month window with 95% confidence interval.

3.3.3 Indicators of patent importance

Several indicators of patent importance have been proposed in the literature. A frequently used indicator is the count of citations received by the patent, which has been shown to correlate with the technical merit and the commercial value of the underlying invention

(Trajtenberg, 1990; Albert et al., 1991; Trajtenberg et al., 1997; B. H. Hall et al., 2005). In our context, given that we use the arrival rate of citations to measure follow-on inventions, we turn to alternative indicators of patent importance.

We consider many indicators, including the number of independent claims, the number of words in the first claim, the number of IPC classes, the patent originality, the number of scientific references listed in the patent document, and the patent family size. Various scholars have demonstrated the relevance of these indicators (Lerner, 1994; Harhoff et al., 2003; Marco et al., 2019; Higham et al., 2021). Note that all these indicators are available at the time of patent filing or shortly thereafter and are, therefore, not affected by patent citations.

We aggregate these variables into a single indicator of patent importance using principal component analysis (PCA), which we label $Importance$. As a robustness check, we also split the sample by the strength of the appropriability regime and estimate a regime-specific score, labeled $Importance_r$. Table B.1 reports the first-stage result on component eigenvalues and factor loadings for both variables. According to the Kaiser Rule, we drop components whose eigenvalues are below 1 and only retain the first component as it captures the highest proportion of variance (28%).

3.3.4 Descriptive statistics

Table 3.1 presents descriptive statistics for the sample of 843 marked patents. It reports the mean, standard deviation, as well as the 10th, 50th, and 90th percentile for all variables. The distributions for patent citations are skewed. On average, a patent receives 1.4 external citations (with the median being 1) and 0.18 citations from new firms every 6 months (with the median being 0). The mean of the similarity score (computed for each cited-citing pair) is 0.71 with a standard deviation of 0.16 (and a theoretical maximum at 1.00).

The primary variable of interest is the dummy variable $Marking$, for which the mean score suggests that about 23 percent of observations come from the post-marking periods. Similarly, about 49 percent of observations come from periods after product launch. The average age of patents in the sample is about 85 months. Furthermore, 29 percent of patents are filed by firms operating in strong appropriability regimes (and the remaining patents are in a weak regime). In terms of patent quality indicators, the sample of patents covers 4.6 IPC classes and has 4.7 independent claims on average. The average length of the first claim is 160 words. These patents also have an average geographical family size of 10.7 and an average of 16 citations to non-patent references. The originality score has a mean of 0.63 and a standard deviation of 0.24. The aggregated measure of patent importance $Importance$ has a mean of -0.01 and a standard deviation of 1.30; $Importance_r$ has a mean of -0.35 and a standard deviation of 1.55.

Table 3.1 – Descriptive Statistics

VARIABLES	Mean	SD	P10	P50	P90
Dependent variables					
External citations External citations	1.40	2.68	0	1	4
Citations from new firms	0.18	0.56	0	0	1
Similarity	0.71	0.16	0.47	0.74	0.89
Explanatory variables					
Marking	0.23	0.42	0	0	1
Product launch	0.49	0.50	0	1	1
Patent age	84.67	62.21	16	70	176
Strong appr. regime	0.29	0.45	0	0	1
Number of IPC classes	4.63	4.74	1	3	10
Number of independent claims	3.31	3.97	1	3	6
Number of words in the first claim	160.02	105.05	62	135	279
Patent family size	10.70	7.64	1	10	22
Number of non-patent references	15.93	26.74	0	4	55
Originality	0.63	0.24	0.26	0.69	0.89
$Importance$	-0.01	1.30	-1.75	-0	1.64
$Importance_r$	-0.35	1.55	-2.12	-0.25	1.30

3.4 Empirical approach

3.4.1 Identification strategy

We want to establish whether and to what extent patent marking leads to more or fewer follow-on inventions. Since patent marking is not a random event, we cannot rule out the presence of unobserved differences between these two groups of patents. Therefore, we cannot simply compare citation rates between marked and not marked patents.

Our identification strategy relies on two core elements. First, we will restrict the analysis to marked patents only and will focus on a *change* in the marking status of these patents (*i.e.*, from not-yet-marked to marked). Thus, we estimate the average difference in patent citations between the periods before and after getting marked. This setup allows us to control for patent fixed effects, thereby accounting for unobserved heterogeneity across patents.

Second, we will control for a series of confounding factors that may explain changes in citation rates over time. Most importantly, the regression will control for the launch date of the underlying product. Including this variable ensures that the change in citation rates is truly identified by the public signaling of inventions instead of the fact that the invention gets commercialized, which indicates that the technology is becoming market-ready. As figure 3.2 illustrates, there is considerable heterogeneity in the time lag between the commercialization and the marking dates. The regression model will also include a functional term for patent age to account for the effect of the passing of time on the arrival of citations.

3.4.2 Regression model

Assessing the effect of marking in weak and strong appropriability regimes

As explained, we implement a patent-level fixed effect regression model. Specifically, we estimate the effect of patent marking on citations received by patent i in a six-month time-window t using the following model:

$$\begin{aligned}\log(cit+\nu)_{it} =&\beta_0+\beta_1 Marking_{it}+\beta_2 Marking_{it}\times StrongPat_i \\ &+\phi ProdLaunch_{it}+f(PatAge_{it},\boldsymbol{\theta})+\lambda_i+\mu_y+\epsilon_{it},\end{aligned} \tag{3.1}$$

where the main dependent variable is the log number of external citations (cit) received in the time-window t. Since the citation variable contains a lot of zero values, we increment the number of citations with an arbitrarily small and strictly positive number ν. That is to say, the main dependent variable used in regressions is the log number of citations received by the focal patent from other firms (external citations) plus a non-zero term ν. In a first specification, ν is equal to 0.15 for all patents. In a second specification, ν is a random variable denoted by ν_{it}, which is drawn from a uniform distribution on the interval $[0.1, 0.2]$ (Bellego & Pape, 2019).

On the right hand side of equation (3.1), the variable of interest, $Marking_{it}$, is a binary indicator that takes the value of 1 if the focal patent has been virtually marked at time t and 0 if not. In order to test the moderating effect of the strength of the regime of appropriability, we add a dummy variable $StrongPat_i$ takes the value of 1 for patents by companies in a strong regime and 0 otherwise. We expect $\beta_1 > 0$ given H1a, whereas we expect $\beta_1+\beta_2 < 0$ given H2a.

Regarding control variables, $ProdLaunch_{it}$ is a dummy variable indicating whether the product associated with patent i has been launched at time t. We control for the age of the patent by including a quadratic specification of $PatAge_{it}$, which we compute as the number of months at time t since the filing month of patent i. In addition, the regression includes patent-level fixed effect λ_i and filing year fixed effect μ_y to account for unit- and time-varying heterogeneity. Note that the inclusion of the patent fixed effect prevents us from controlling for the level of the variable $StrongPat_i$ as this characteristic is constant for a patent over time. Furthermore, patent fixed effect also makes the inclusion of firm fixed effect irrelevant. Finally, ϵ_{it} is a heteroscedastic error term.

As a robustness test, we use an alternative approach to measure invention diffusion by looking at whether a citation comes from a new firm. Our definition of a new firm is one that has never cited any patents of the focal firm before. We then estimate equation (3.1) using the log number of citations from a new firm in a six-month window or a dummy variable on whether any citations come from a new firm in a six-month window. The rest of the right-hand side remains the same.

We test H1b and H2b (concerning invention similarity) using a modified version of equation (3.1). In particular, the unit of analysis is now a citing-cited pair. The estimating equation is specified as follows:

$$\begin{aligned} Similarity_{ijt} =& \beta_0 + \beta_1 Marking_{it} + \beta_2 Marking_{it} \times StrongPat_i \\ & + \phi ProdLaunch_{it} + f(PatAge_{it}, \boldsymbol{\theta}) + \lambda_i + \mu_y + \epsilon_{ijt}, \end{aligned} \tag{3.2}$$

where the dependent variable is the cosine similarity score between the focal patent i and its citing patent j. $Marking_{it}$ is a dummy variable that equals 1 if patent i has been marked at the time t. In line with our hypotheses, we expect $\beta_1 > 0$ (H1b) and $\beta_1 + \beta_1 < 0$ (H2b). All else remain the same as above.

Assessing the effect of marking of important patents

In order to test H3b, we adopt a split-sample approach and interact the $Marking$ variable with the indicator of patent importance. We split the sample by the strength of the appropriability regime and estimate:

$$\begin{aligned} \log(cit + \nu)_{it} =& \beta_0 + \beta_1 Marking_{it} + \beta_2 Marking_{it} \times Importance_i \\ & + \phi ProdLaunch_{it} + f(PatAge_{it}, \boldsymbol{\theta}) + \lambda_i + \mu_y + \epsilon_{it}, \;\; i \in \{S_w, S_s\} \end{aligned} \tag{3.3}$$

where S_w is the sample of patents in weak regimes and S_s is the sample of patents in strong regimes. All else remain the same as above. We expect $\beta_2 > 0$ per H3b.

3.5 Baseline regression results

3.5.1 Marking and the strength of appropriability regime

We start the analysis by estimating equation (3.1) without the interaction term to get a sense of the data. Panel A of Table 3.2 reports the OLS estimates with $\log(cit + 0.15)$ as the dependent variable, whereas Panel B reports the OLS estimates with $\log(cit + \nu_{it})$ as the dependent variable. We refrain from interpreting the coefficient associated with the variable $Marking$ since the overall effect depends on the composition of the sample (*i.e.*, firms in weak vs. strong appropriability regimes). The variable product launch is significantly negative in column (2) but drops in magnitude and significance when including patent and application year fixed effects (column 3 and column 4). Thus, simply commercializing a product does not seem to be associated with a change in the number of external citations received by the underlying patent(s). Note that using a fixed increment in the dependent variable or a random increment in the interval [0.1, 0.2] produces quantitatively similar results.

Table 3.2 – The effect of patent marking on follow-on inventions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
VARIABLES		$log(cit+0.15)$				$log(cit+v_{it})$		
		Panel A				Panel B		
Marking	0.267***	0.078**	0.056*	0.120***	0.272***	0.082**	0.063*	0.128***
	(0.027)	(0.032)	(0.033)	(0.040)	(0.027)	(0.032)	(0.034)	(0.041)
ProdLaunch		-0.253***	-0.018	0.037		-0.254***	-0.020	0.036
		(0.025)	(0.033)	(0.033)		(0.025)	(0.033)	(0.033)
PatAge		0.012***	0.011***	-0.025***		0.012***	0.011***	-0.025***
		(0.001)	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)
PatAge Squared		-0.000***	-0.000***	-0.000***		-0.000***	-0.000***	-0.000***
		(0.000)	(0.000)	(0.000)		(0.000)	(0.000)	(0.000)
Constant	-0.623***	-1.160***	-1.115***	1.929***	-0.632***	-1.175***	-1.129***	1.917***
	(0.013)	(0.025)	(0.027)	(0.112)	(0.013)	(0.025)	(0.028)	(0.113)
Patent FE	NO	NO	YES	YES	NO	NO	YES	YES
Year FE	NO	NO	NO	YES	NO	NO	NO	YES
Observations	15,578	15,578	15,578	15,578	15,578	15,578	15,578	15,578
R-squared	0.006	0.068	0.354	0.401	0.006	0.068	0.352	0.398

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

We now turn our attention to the moderating effect of the appropriability regime on follow-on inventions to test H1a and H2a. Table 3.3 reports estimates of equation (3.1) using several specifications and estimation methods. Column (1) reports OLS estimates with log($cit+0.15$) as a dependent variable and column (2) reports OLS estimates with log($cit+v_{it}$) as a dependent variable. In addition, considering that citation data are count data, we report the marginal effect from Poisson estimation (variable cit) in column (3). All regressions include the full set of control variables and fixed effects.

The coefficient associated with the variable $Marking$ is positive and significant in all specifications, providing empirical support for H1a. The number of citations increases by 22.3–22.8 percent in columns (1) and (2) when patents become marked. Given that the average number of citations in the six-month window before marking is 1.34, this figure corresponds to about

0.31 more citations. Marginal effects at mean of the Poisson estimates in column (3) suggest an increase of 0.137 citations. Conversely, the overall effect of patent marking in strong regimes is always negative and significant. In column (1), the magnitude of this effect reaches -0.174 (=0.228-0.402), and the p-value associated with the test of joint statistical significance leads to a rejection of the null hypothesis of no effect. In strong regimes, marking patents reduces follow-on inventions by 17.4 percent, corresponding to about 0.23 citations less in each time window.

Table 3.3 – The moderating effect of appropriability regime on follow-on inventions

	(1)	(2)	(3)
	OLS		Poisson
VARIABLES	$log(cit+0.15)$	$log(cit+v_{it})$	count of citations
Marking	0.228***	0.223***	0.137*
	(0.043)	(0.044)	(0.080)
Marking × StrongPat	-0.402***	-0.406***	-0.489***
	(0.052)	(0.053)	(0.124)
F-Stat on $\beta_1 + \beta_2$	33.228	32.322	538.62
Prob > F	0.000	0.000	0.000
ProdLaunch variable	YES	YES	YES
PatAge variables	YES	YES	YES
Patent FE	YES	YES	YES
Year FE	YES	YES	YES
Observations	15,578	15,578	15,578

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

As an additional robustness test, we test whether patent marking attracts more diffusion to new firms. The estimation follows equation (3.1) and results are reported in Table 3.4. In columns (1–2), our dependent variable is the log number of citations coming from new firms in a six-month window. In columns (3–4), we estimate a linear probability model where our dependent variable is a dummy variable $Newfirm$ that equals 1 if any citations in a six-month window come from a new citing firm. Again, the regression includes the full set of control variables and fixed effects.

Regardless of specifications, the results in columns (1) and (3) show that marking alone has almost no impact on attracting new citing firms. Once we add the interaction term between $Marking$ and $StrongPat$ in columns (2) and (4), marking induces more citations from new firms while having strong appropriability regimes prevents diffusion to new firms and these effects are significant on the 1% level. The findings in columns (2) and (4) further confirm H1a and H2a.

Table 3.4 – The effect of patent marking on follow-on inventions from new firms

	(1)	(2)	(3)	(4)
VARIABLES	$log(cit+v_{it})$		$Newfirm$	
Marking	-0.011	0.072***	-0.006	0.026**
	(0.025)	(0.025)	(0.011)	(0.011)
Marking × StrongPat		-0.315***		-0.123***
		(0.035)		(0.015)
ProdLaunch variable	YES	YES	YES	YES
PatAge variables	YES	YES	YES	YES
Patent FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Observations	15,655	15,655	15,655	15,655
R-squared	0.174	0.179	0.168	0.172

Notes: Results using $log(cit+0.15)$ as a dependent variable are similar. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3.5 reports estimates of equation (3.2), allowing us to test H1b and H2b regarding invention similarity. In columns (1–4) we gradually include the control variables and the fixed effects. We add the interaction term $Marking \times StrongPat$ in column (5) and report the joint test for $\beta_1 + \beta_2$.

Again, we find a stark difference between patents in weak and strong regimes. Follow-on patents become more similar to the focal marked patents after marking in weak regimes, the coefficient being 0.013 in column (5). This result provides empirical support for H1b and suggests that firms follow more closely the path signaled by the marked patent. However, the magnitude of the coefficient is small, corresponding to about one-tenth of the standard deviation. In strong regimes, the effect is negative and statistically significant (value of -0.005 = 0.013-0.018), in line with H2b. The coefficient is small in magnitude; nevertheless, this evidence suggests that information disclosure helps to reduce duplicative R&D efforts for patents in strong regimes.

Table 3.5 – The effect of patent marking on invention similarity

	(1)	(2)	(3)	(4)	(5)
VARIABLES			Cosine similarity score		
Marking	-0.029***	-0.019***	0.005*	0.006*	0.013***
	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)
Marking × StrongPat					-0.018***
					(0.005)
F-Stat on $\beta_1 + \beta_2$					10.119
Prob > F					0.000
ProdLaunch		-0.012***	-0.012***	-0.009***	-0.007**
		(0.002)	(0.003)	(0.003)	(0.003)
PatAge		-0.000***	-0.000***	-0.001***	-0.001***
		(0.000)	(0.000)	(0.000)	(0.000)
PatAge Squared		0.000***	0.000***	0.000***	0.000***
		(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.671***	0.698***	0.702***	0.737***	0.735***
	(0.001)	(0.003)	(0.003)	(0.026)	(0.026)
Patent FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Observations	22,002	22,002	22,002	22,002	22,002
R-squared	0.007	0.011	0.432	0.436	0.437

Notes: The results are the same when excluding patents from the same family. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

3.5.2 Marking and invention importance

In Table 3.6, we investigate the exacerbating effect of invention importance on follow-on inventions. We report results obtained using $\log(cit + v_{it})$ as the dependent variable and break down the sample according to the strength of the appropriability regime. We test two measures of patent importance: $Importance$, estimated from a PCA run on the full sample, and $Importance_r$, estimated from a PCA run on each subsample separately.

Overall, marking has a positive effect on citation rates in weak appropriability regimes in columns (1) and (2), as already illustrated. However, the interaction term between marking and patent importance is null, meaning that more important patents do not attract more follow-on inventions. There is a limit to how much follow-on inventions can occur, and the lack of a significant effect could be explained by the strength of the baseline effect.

Regarding strong appropriability regimes in columns (3) and (4), the baseline effect of marking is negative, as previously established. Interestingly, the interaction term is positive and highly statistically significant. In other words, the deterring effect of patent marking on follow-on inventions becomes weaker for more important patents—for instance, a 1.3 standard deviation increase in the patent importance is associated with 8.4 percent more external citations as in column (3). This evidence suggests patent marking may amplify the potential value of more important patents and attract competitors to build on these inventions. Overall, we find support for H3b in strong patent regimes only.

Table 3.6 – The moderating effect of patent quality on patent marking

	(1)	(2)	(3)	(4)
VARIABLES	$log(cit+v_{it})$			
	Weak appr. regime		Strong appr. regime	
Marking	0.170***	0.171***	-0.171**	-0.160**
	(0.050)	(0.050)	(0.076)	(0.076)
Marking × $Importance$	-0.003		0.084***	
	(0.030)		(0.024)	
Marking × $Importance_r$		0.001		0.093***
		(0.027)		(0.025)
ProdLaunch	0.061	0.061	0.257***	0.264***
	(0.039)	(0.039)	(0.068)	(0.068)
PatAge	-0.025***	-0.025***	-0.029***	-0.029***
	(0.002)	(0.002)	(0.003)	(0.003)
PatAge Squared	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Constant	1.628***	1.628***	2.893***	2.894***
	(0.118)	(0.118)	(0.251)	(0.251)
Patent FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Observations	11,106	11,106	4,484	4,484
R-squared	0.334	0.334	0.497	0.497

Notes: Results using $log(cit+0.15)$ as a dependent variable are similar. Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

3.6 Conclusion

The recent implementation of virtual patent marking allows firms to disclose product-associated patents in a cost-effective way to improve firms' ability to avoid potential innocent infringement. By frequently updating patent marks through URLs, firms deliberately signal to the public their valuable inventions. Whether and how virtual patent marking affects the diffusion of subsequent inventions has not been studied.

In this paper, we examine the effect of virtual patent marking on various outcomes of invention diffusion. The empirical setup exploits variations in patent marking over time. We show that signaling valuable patents is associated with more follow-on inventions in weak regimes such as consumer goods. We also find that these follow-on patents become closer content-wise to the focal patents. Thus, it seems that competitors rush into the research path signaled by the firm. In strong regimes (such as biotechnology), however, patent marking seems to deter competition: patents attract fewer follow-on inventions and of more dissimilar content, consistent with findings from Baruffaldi and Simeth (2020) and Lück et al. (2020). We also find patent importance plays a mitigating role in the negative effect of marking: the effect of marking in strong appropriability regimes is attenuated for the more important patents.

We acknowledge some limitations in this study. First, forward citations may present an imperfect measure of follow-on innovation and fail to capture cumulative R&D activities that are unpatentable or not patented (Williams, 2013). Second, our sample covers firms in consumer goods, electronics and mechanics, as well as pharmaceutical and medical equipment. It is also limited to firms that have timestamped their VPM web pages. We suggest caution when interpreting these findings for a more general sample.

4 Does Motherhood Hold Back 'Marie Curies'?

This chapter is written in collaboration with Mary Kaltenberg.[1]

Abstract:

Female invention participation has steadily grown in the U.S. over the past few decades, but the gender innovation gap remains substantial. This growth in participation corresponds with an overall increase of female labor force participation and changes in maternity leave policies. Using inventor data from patents from the U.S. Patent and Trademark Office, this chapter seeks to evaluate the impact of maternity leave policies in innovation-related jobs in two particular perspectives, exit decisions and productivity of female inventors of child-bearing age. Our findings suggest that maternity leave policies promote the retention of female inventors, but these policies have little impact on increasing productivity.

Key words: innovation, gender gap, maternity leave

[1] I am grateful to Gaétan de Rassenfosse for his valuable guidance and constructive feedbacks on this chapter. We also thank Stefano Breschi, Patrick Gaulé, Kyle Higham, as well as participants at the Research Policy Conference for Early Career Scholars for the helpful comments at an early stage of this project.

4.1 Introduction

Recent decades have witnessed a significant improvement in the labor force participation rate of women in various occupations (Goldin, 1989; Parkman, 1992; Goldin & Katz, 2002; Juhn & Potter, 2006; F. D. Blau et al., 2013). Yet what remains puzzling is the disproportionately low presence of women and their hampered productivity in science and engineering. As highlighted by Bell et al. (2019): "it will take another 118 years to reach gender parity" despite the dwindling gender gap in becoming inventors. Assuming that creativity and talents are uniformly distributed among men and women, the persistent under-representation of women in highly innovative occupations reflects an unlocked potential of high-skilled human capital that would have otherwise translated into discoveries of new ideas or inventions of new technologies.

While there are many explanations as to why women are under-represented in innovation, maternity and childcare obligations are an unignorable contributor (Preston, 1994; Hunt et al., 2013; Moser & Kim, 2020). The consequences of motherhood and childcare on career women vary from impaired productivity to halted careers. These interruptions hinder women who survived rigorous training from allocating their skills where best needed, thus putting the society at risk of losing 'Marie Curies'. However, it remains contentious how policies that intend to balance work and family will affect women in high-skilled occupations.

In this chapter, we study how maternity leave policy, particularly state maternity leave provisions and Family Maternity Leave Act, impacts labor market participation decisions of high-skilled inventors as well as its impact on productivity during peak reproductive ages. We obtain data on the universe of US-residing inventors and their patents from the United States Patent and Trademark Office (USPTO) for the period from 1970 to 2015 and match these inventors to the web-scraped data on inventor ages. By exploiting a staggered passage of maternity leave policy across the United States, we employ event-study models and estimate the effects of maternity leave policy on the probability of exit patenting and the productivity for male and female inventors in different age groups.

Our results suggest that maternity leave policies are most effective at improving the retention of R&D occupations for women of reproductive ages. After controlling for state and application year fixed effects and individual preferences, our event-study estimates suggest that the probability of leaving patenting for women aged 25-34 decreases by as much as 17.8 percent five years after the passage of maternity policy. In comparison, the maternity leave policy has little impact on older women and men of all age groups. A further Kaplan-Meier survival analysis confirms that the probability of continuing patenting has a sizable increase for women aged 25 to 45. after implementing the maternity leave policy. We also find that maternity leave policy is ineffective at increasing productivity in patenting for both women and men inventors. We speculate that this is because these types of policies do not help alleviate barriers in increasing productivity, such as the inaccessibility of childcare facilities.

Our findings build on and contribute to several strands of literature. First and foremost, our findings on the gendered exit patterns of inventors contribute to various studies on the disparity in innovation among different demographic groups such as age, gender, race, and ethnicity (Ding et al., 2006; Hunt et al., 2013; Cook, 2014; Jung & Ejermo, 2014; K. Jensen et al., 2018; Bell et al., 2019; Sarada et al., 2019; Koning et al., 2020; Moser & Kim, 2020; Kaltenberg, Jaffe, & Lachman, 2021). Some papers have documented that the under-representation of female inventors varies by field (Ding et al., 2006; Jung & Ejermo, 2014; Koning et al., 2020). What is worrying is that the gender gap in productivity over the life cycle never closes (Kaltenberg, Jaffe, & Lachman, 2021). One possible explanation of the patenting gender gap involves a lower enrollment of women in patent-intensive study fields and lower participation in development and design job tasks (Hunt et al., 2013). On the other hand, female inventors experience a disadvantage at obtaining patent rights—they are 7% less likely than men inventors to have patent applications accepted (K. Jensen et al., 2018). A recent paper by Moser and Kim (2020) proposes that motherhood might be a cause of the productivity gender gap among scientists: mothers experience a loss of productivity in patenting in their first 15 years of marriage.

Second, this chapter adds to the works on motherhood and women's labor market outcomes as well as the works on how family-friendly policies affect women's job attachment and productivity (Baum, 2003a, 2003b; Lalive & Zweimüller, 2009; Bertrand et al., 2010; Schönberg & Ludsteck, 2014). For example, Bertrand et al. (2010) show that female MBA graduates with children experience more career interruptions and shorter work hours, and they also earn less than their male counterparts. Likewise, the responsibility of family and children harms the productivity of female academics and scientists Ginther and Kahn (2004) and Mairesse et al. (2019). Nevertheless, policies aiming to protect mothers' labor market outcomes, such as parental leave policies, are found with ambiguous results; moreover, their impacts on high-skilled workers are little known. Our evaluation of the impact of maternal leave policies on female inventors bridges the gap between the literature on labor economics and the studies on the innovation gender gap, which draws policy implications, especially for high-educated and high-skilled knowledge workers.

The rest of this chapter proceeds as follows. Section 4.2 provides a literature review on gender gap in innovation. Section 4.3 illustrates the institutional background of maternity leave and its impacts on women's labor market outcomes. Section 4.4 describes data and empirical strategy. Section 4.5 presents the results and finally, Section 4.6 draws conclusions.

4.2 Literature Review

The under-representation of women in innovation can be traced back to gender differences at several milestones along the career path. A handful of papers point to the gender gap in entry into science, technology, engineering, and match (STEM) majors as well as into science and engineering careers (Xie et al., 2003; Ceci et al., 2014). Some show that males and females of equivalent capacities hold divergent preferences and beliefs regarding occupations, which prompts them to choose different majors upon starting college (Daymont & Andrisani, 1984; Zafar, 2013). On the other hand, women's performance and productivity in science and engineering fields as well as their persistence in doctoral training, a critical step for an R&D intensive career, is subject to various factors such as professor gender and peer gender composition (Bettinger & Long, 2005; Carrell et al., 2010; Bostwick & Weinberg, 2018; Gaule & Piacentini, 2018).

Barriers to entry only explain part of the "leaky pipeline" problem; there are many reasons as to why women don't stay in science and engineering occupations. Preston (1994) documents that women are twice as likely as men to leave science and engineering to other occupations or quit the labor force. Moreover, reaction to marriage and children explain their departure partly. In a careful study on the IT workforce, Stephan and Levin (2005) find that only 65.8% of women remained in IT-related occupations after six years compared to 73.2% men. In particular, women who parent young children are significantly more likely not to be working than men. Further studies suggest that family constraints are a major hurdle for women to continue inventing (Frehill, 2012; Hunt, 2016).

From giving birth to child-bearing, the obligation of motherhood could put working mothers at a disadvantage and contribute to a widened gender gap in terms of work hours, wages, and promotions. Even a sub-group of highly educated and highly skilled women who are presumably least susceptible to such interruptions cannot escape the motherhood penalty (Whittington, 2011). Previous research examining the gender gap in innovative occupations from the angle of parenthood mostly relies on survey data or bibliographical data on inventors and scientists. Drawing survey data on more than 9,000 inventors from 23 countries, Hoisl and Mariani (2017) show that women inventors earn 12–14% less than men and suffer from a loss in pay due to parenthood. Mairesse et al. (2019) collect longitudinal data on physicists at the Institute of Physics in France and find female physicists having a young child publish one journal article less. In an ongoing paper, Moser and Kim (2020) analyze bibliographical data on 83,000 American scientists during the Baby Boom, including 4,000 women, to study gender inequality in STEM. They find that women experience a decline in patenting productivity in their 20s and that mothers are 21% less likely to get tenure than fathers. A more recent study finds that female scientists in theU.S.and Europe are especially affected by the pandemic in the time of COVID, and they spend substantially less time on research because of the burden of taking care of small children (Myers et al., 2020). But little research has examined how motherhood, or policies aimed at promoting mothers' work-life balance, affect the retention of women in innovation.

4.3 Maternity leave policy and the labor market

This section elaborates on the institutional change in maternity policy in the U.S. and discusses its impact on women's labor market outcomes.

4.3.1 Maternity leave policy in the U.S.

Before 1993, the maternity leave policy in the U.S. was largely decided by state law and employer policies. Albeit varying lengths, maternity leave provisions were only adopted in thirteen states, including Washington D.C., among which Massachusetts pioneered its implementation in 1972. The Family and Medical Leave Act (FMLA) was enacted in January 1993 and went into effect in August in the same year. It mandated a nationwide 12-week unpaid maternity leave for eligible working women in companies with at least 50 employees for the first time. Employees entitled to the benefits should have worked for at least 1,250 hours in the previous 12 months. Figure 4.1 depicts the staggered enactment of maternity leave policy across the U.S. from the 70s to 1993 when FMLA became a federal mandate. It is worth mentioning that the FMLA mostly seeks to provide coverage to women who work full-time in corporations.[2] As of yet, there is no federal paternity leave policy in the U.S. with some exceptions that California passed the first paid parental leave in 2002 and New Jersey was the second state to do so in 2008.

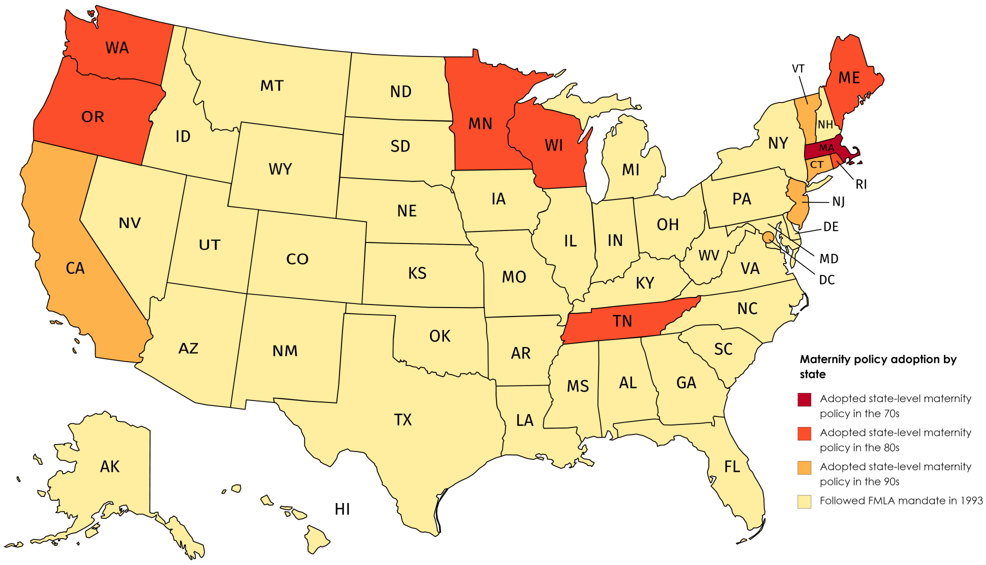
Other forms of mandated maternity benefits have been provided before FMLA, including comprehensive coverage of maternity by health insurance.

The ultimate purpose of maternity leave policies is to protect the maternity rights of employees at work and to ensure the well-being of mothers and children (Chatterji & Markowitz, 2004; Rossin, 2011).[3] We argue that the timing of maternity leave policy adoption is exogenous to state-level economic performance and thus unlikely to be confounded by factors that simultaneously influence regional innovative activities. To validate our argument, we present

[2] In large and medium-sized establishments, the share of full-time employees covered by unpaid leave rose from 37 percent in 1991 to 84 percent in 1995, while the share of paid leave coverage remained at two percent. See Waldfogel (1999).

[3] Other forms of mandated maternity benefits have been provided before FMLA, including comprehensive coverage of maternity by health insurance. However, there is little evidence on how those benefits relate to leave-taking behavior. The implication of mandated benefits is beyond the scope of our study. See Gruber (1994) for a detailed discussion on the welfare analysis of mandated benefits.

Figure 4.1 – Implementation of Maternity Leave Policy across the US



the pairwise correlations between the year of policy enactment and innovation capacity in the year before, the year at, and the year after federal enactment of FMLA across states. The measures of innovation capacity are state stock of inventors and patents. Table 4.1 shows that the correlation between the timing of policy enactment and state-level innovation capacity is insignificant.

Table 4.1 – Pairwise Correlation between Policy Enactment and Innovation Capacity

	Year of enactment	Total inventors	Total patents
1992			
Year of enactment	1.0000		
Total inventors	-0.1314 (0.3579)	1.0000	
Total patents	-0.1156 (0.4191)	0.9989 (0.0000)	1.0000
1993			
Year of enactment	1.0000		
Total inventors	-0.1378 (0.3349)	1.0000	
Total patents	-0.1171 (0.4130)	0.9989 (0.0000)	1.0000
1994			
Year of enactment	1.0000		
Total inventors	-0.1430 (0.3167)	1.0000	
Total patents	-0.1213 (0.3966)	0.9984 (0.0000)	1.0000

Notes: Innovation capacity is measured by the total number of inventors and the total number of patents in each state in the respective years from 1992 to 1994. P-values in parentheses.

As additional evidence, Figure C.1 in the appendix displays scatter plots on the timing of enactment and innovation capacity by the state in 1992, the year before federal enactment of FMLA. Out of the thirteen states that passed state-level maternity provisions before federal legislation, only three (California, New Jersey, Massachusetts) outperformed with more than two thousand patents and inventors in 1992. In contrast, other comparatively innovative states such as New York, Texas, and Michigan waited until the passage of FMLA.

4.3.2 Fertility, motherhood, and labor market outcomes

We propose two channels through which maternity leave policy affects women's labor market outcomes. The first channel operates on women's fertility decisions, which further determines how fast women return to work, if at all. Evidence from developed countries shows that the extension in parental leave policy raises birth rates and shortens the birth gap (Björklund, 2006; Lalive & Zweimüller, 2009). In the case of the US, the implementation of FMLA induces more births and encourages leave-taking among women (Averett & Whittington, 2001; Cannonier, 2014). In particular, mandated maternity leave is more effective at inducing leave-taking for women in a non-government sector and women with at least a college education (S. P. Kerr, 2016). Research in labor economics has established that fertility is associated with reduced labor supply and increased withdrawal from the labor market for women of reproductive ages. Therefore, it is reasonable to think that such family-friendly policies as maternity leave will likely cause more frequent interruptions in women's careers. Nevertheless, the duration and cash benefits offered by maternity leave also affect subsequent work-related decisions of women. For example, extensions of paid parental leave in Austria and Germany have been documented to have delayed return to work and reduced employment for mothers of newborns in the short-run (Lalive & Zweimüller, 2009; Schönberg & Ludsteck, 2014). Unlike other developed countries, the maternity leave policy in the US, whether by state provision or by the federal mandate, is relatively short and essentially unpaid. Despite its length, the job-protective nature of maternity leave policies arguably affects the labor market attachment of women. Most studies on the impact of FMLA as well as state-level provisions use data from the National Longitudinal Survey of Youth (NLSY) or Current Population Survey (CPS). For example, both Baum (2003b) and Berger and Waldfogel (2004) find that maternity leave is likely to increase women's return to work after childbirth. In contrast, other studies find an insignificant impact of FMLA on employment (Waldfogel, 1999; Baum, 2003a).

A second-order implication of maternity leave policy ascribes to the child-bearing obligation associated with motherhood. Women usually take the responsibility of providing care for children. Despite benefits provided on maternity, insufficient access to childcare facilities and support will particularly take a toll on working mothers.[4] A primary reason for the low entry into the employment for married women and high exit rate for women of child-bearing age is the high expenses of childcare (D. M. Blau & Robins, 1989; Ribar, 1992; Kimmel, 1998). Evidence suggests that high-skilled women experience the most severe wage penalty for becoming mothers. The high opportunity costs associated with job-specific human capital make even a small amount of time for child-bearing costly in terms of returns (Anderson et al., 2002; England et al., 2016). In addition, studies find that women in top-earning corporate occupations reduce their hours worked and exit employment after childbirth for family reasons (Bertrand et al., 2010; Ganguli et al., 2020).

4.4 Data and Empirical Strategy

4.4.1 Data

We use the '20200610' version of USPTO data from PatentsView as the main data source and obtain the universe of inventors that resided in the U.S. at the time of patent application for the years 1970 to 2015. We match the inventors to their patent applications, assignee, and geographic location (city and state). We next match this inventor-patent data to disambiguated inventor gender information from PatentsView, resulting in 1,465,934 inventors associated

[4] In 2014, the US ranks 20th out of 31 OECD countries in terms of the percentage of formal childcare enrollment for children aged 0-2 with 28 percent. See https://equitablegrowth.org/falling-behind-the-rest-of-the-world-childcare-in-the-united-states.

with 3,189,806 patents.

We connect this information to web-scraped inventor ages documented in (Kaltenberg, Jaffe, & Lachman, 2021). Using USPTO for inventors residing in the U.S., they scrape age information from three directory websites, Radaris, Spokeo, and Beenverified based on information included in patents, such as name and location. Relying on a scoring system of accuracy and verification through repeated age collection from multiple sites, the authors are able to identify the ages of 1,439,272 inventors. They verify their results of productivity and patent attribute patterns over the life course by comparing the full dataset to a subset of ages that are high confidence and low confidence. The high confidence subset includes inventors whose ages were verified across multiple web directories. The low confidence subset includes inventors who may have some disagreements between the ages they collected. They find little differences between the subsets, though the estimates at the extreme ends (below 20 and above 60) tend to be less precise. This should not be a concern in our results as the estimates from concerned groups do not rely on ages on either end of this spectrum. However, ages are collected through web directory websites and not through birth certificates or other verification systems, and thus, errors could remain.

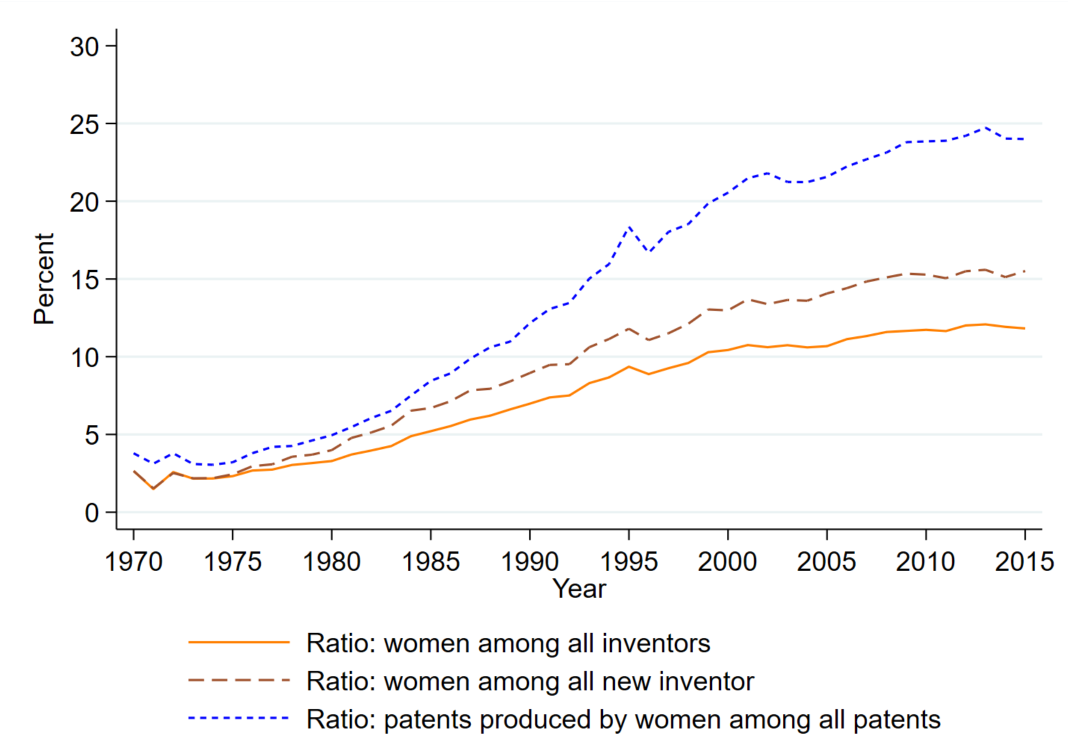
We link the age information with the inventor-patent data by matching on the persistent inventor I.D. crosswalk table from PatentsView and on first and last names, which resulted in 1,213,623 US-residing inventors associated with 2,640,156 patents. Among these inventors, 139,104 are women, which makes up about 11.46%. One particular drawback to the data set is that women are more likely to change names after marriage. The USPTO inventor disambiguation may not be able to identify that an inventor whose name changes is one consistent inventor rather than two inventors.

Female participation in invention has been steadily growing, as shown in Figure 4.2. Overall, the share of female inventors increased from 2.65% in 1970 to 11.82% in 2015. This increase is partly due to the increasing number of first-time U.S. residing female patent holders, which rose about five-fold from 1970-2015 to 15.51%. Correspondingly, the percentage of patents produced with at least one woman inventor increased from 3.79% to almost 24% in 2015 with a relatively faster change before 1995 than the period 1995–2015. These trends indicate that the growth of female participation in patenting slightly lagged behind the growth of their innovative output.

4.4.2 Empirical Strategy

We aim to evaluate the effects of state and federal legislation of maternity leave policy on the participation of women in innovation. Maternity leave policies target a demographically specific group of people who experience the incidences of motherhood, that is, women of

Figure 4.2 – Representation of Female Inventors and Patenting Activities in the US from 1970 to 2015



Notes: These statistics are computed on the sample of PatentsView inventor-patent data matched to inventor-age data, where 1,213,623 unique inventor ids are associated with 2,640,156 patents during 1970–2015. New inventors are identified as those who applied for a patent for the first time in each year.

child-bearing age. The family-friendly nature of such policy entitles mothers to the flexibility of spending time with their newborns. However, in dual-career households, mothers who work full-time are usually faced with decisions such as whether to compromise their careers for family responsibility or how to allocate their time between work and childcare if they stay at work. Therefore, it is reasonable to speculate that women's labor supply in the reproductive years is most affected by this policy.

We specifically focus on the inventors' decisions to exit patenting and their inventive productivity as the primary labor supply outcomes. We restrict the studied time window to five years around the enactment of maternity leave policies in respective states. Our identification strategy relies on variations in the timing of maternity leave policy provisions across the U.S., namely, state-level provisions and the FMLA. We employ an event study approach to estimate the effect of maternity leave policies on the retention and productivity of inventors with ordinary least squares (OLS) regressions. Based on the assumption that maternity and relevant leave policies may impact women and men of different age groups separately, we estimate the heterogeneous effects of maternity leave policies on subsamples of inventions belonging to different demographic groups. This approach allows us to see how a target group responds to the policy change in the years relative to policy adoption.

Although lacking information on whether inventors have children or the number of children they have, we approximate this information by their age at maternity leave policy adoption in the state of residence. Considering the average age of first birth and the average age at labor force exit, we restrict our sample to inventors between 25 and 64 years old and divide them into four groups based on their age at the time of policy adoption.[5] These groups comprise inventors aged between 25 and 34, inventors aged between 35 and 44, inventors aged between 45 and 54, inventors aged between 55 and 64.

We report the estimates for women and men inventors of different age groups for comparison. Supposedly, the work-related performances of men are subject to little influence of the maternity leave policy change. To be clear, we restrict our focus on career inventors that have patented more than once over their observed career.[6] We present the model specifications to be estimated under various subsections.

[5] The average age at first birth is around 25 for U.S. women (Mirowsky, 2005). According to S. P. Martin (2000), during mid 1970s to mid 1990s, women in the U.S. with a college degree have a tendency to postpone child-bearing past age 30. Besides, the mean age at retirement for U.S. men and women during 1965–1995 is between 62 and 66, see Gendell (1998) for more.

[6] In our sample, 80.50% of women are career inventors and 85.19% of men are career inventors. In Figure C.2, we plot the share of women inventors across the years relative to maternity leave policy enactment according to their ages. Removing one-shot inventors doesn't systematically bias the percentage of women among career inventors compared to their share among all inventors.

Event Study on Inventors' Probability to Exit

Given the high exit rate of women in science and engineering, examining how maternity leave policies affect the propensity of women of child-bearing age to stop patenting is critical for understanding the under-representation of women in inventive occupations. However, we caution that a women inventor's exit from patenting is a different notion than leaving the labor force. Women in favor of raising children usually leave R&D-intensive occupations for administrative roles or even part-time jobs (Preston, 1994). We consider an inventor to have exited patenting if we observe no more patent applications from that inventor after a certain time threshold. Therefore, we define "exit patenting" as when an inventor applied for a patent for the last time in their career.

Following Moser and Kim (2020), we use OLS regressions to estimate an event-study reduced-form in a 5-year window ($-5 \leq j \leq 5$) around the year of maternity leave policy adoption:

$$Pr(Exit_{it}^{G,A} = 1) = \alpha_0 + \sum_{\substack{-5\leq j\leq 5 \\ j\neq 0}} \alpha_j^{G,A} \cdot \mathbb{1}\{j = t - EventYear_s\} + \delta_s + \mu_t + \theta_i + \epsilon_{it}. \quad (4.1)$$

Our outcome variable $Exit_{ist}^{G,A}$ is a dummy variable that indicates an inventor i has exited patenting since the year t. The superscript G denotes the gender of inventor and we have $G = \{F, M\}$, where F stands for female and M for men. Likewise, the superscript A denotes the four subsets of previously defined age groups. On the left-hand side, α_0 is a constant term. t is the patent application year and $EventYear_s$ is the year when maternity leave was adopted in state s. j is a time variable for the event year that ranges from -5 to 5, and we omit the year in which maternity leave was adopted (in this case $j = 0$) as a benchmark. $\alpha_j^{G,A}$ are our estimates of interest, which can be interpreted as the probability to exit patenting for inventors of type $\{G, A\}$ in the event year j relative to maternity leave policy adoption.

In addition, δ_s captures the state-varying fixed effect such as economic and policy trends. μ_t is the patent application year fixed effect that captures trends in patenting activities over time. To account for individual heterogeneities in preferences for work and family, we include the inventor fixed effect θ_i. Finally, ϵ_{it} is an error term. We use robust standard errors to correct heteroscedasticity.

Event Study on Inventors' Productivity

We then look at whether inventors' productivity at work is affected by changes in maternity leave policy. We measure productivity as the log of annual patents produced. Granted, patenting is a discrete activity that may not precisely reflect the time and efforts put into work. In a similar fashion, we estimate the following event-study specification:

$$log(Pat)_{it}^{G,A} = \beta_0 + \sum_{\substack{-5 \leq j \leq 5 \\ j \neq 0}} \beta_j^{G,A} \cdot \mathbb{1}\{j = t - EventYear_s\} + \delta_s + \mu_t + \theta_i + \epsilon_{it}, \tag{4.2}$$

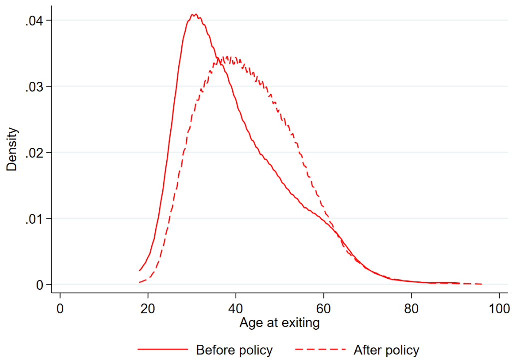
where $log(Pat)_{it}^{G,A}$ is our dependent variable on the log of annual patents for inventor i in year t of demographic group $\{G, A\}$. $\beta_j^{G,A}$ estimates the average change in patents produced by inventors belonging to $\{G, A\}$ in the event year j relative to the benchmark year. The rest of the left-hand side notions remain the same as in Equation 4.1.

4.4.3 Descriptive Statistics

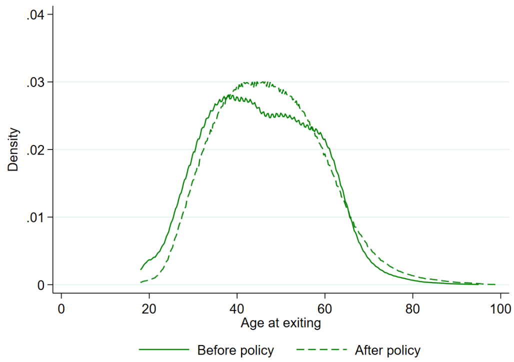
We first document the patterns of age at which women and men quit patenting before and after maternity leave policies. Figure 4.3 displays the densities of exit age for both women and men inventors in the years before and after maternity leave policy adoption; sample restricted to inventors aged between 18 and 100 at the year of patent application. The density of age at which women exit shifts rightward in the years following maternity leave policy passage. The average exit age for women increased from 39.6 to 43.1 years old, implying that women retained longer in a patenting career. In contrast, the densities of exit age in both periods for men largely overlap — the average exit age only slightly increased from 45.6 to 47.5 years old.

Figure 4.3 – Distribution on the Exit Age for Women and Men Before vs. After Policy

(a) Women inventors



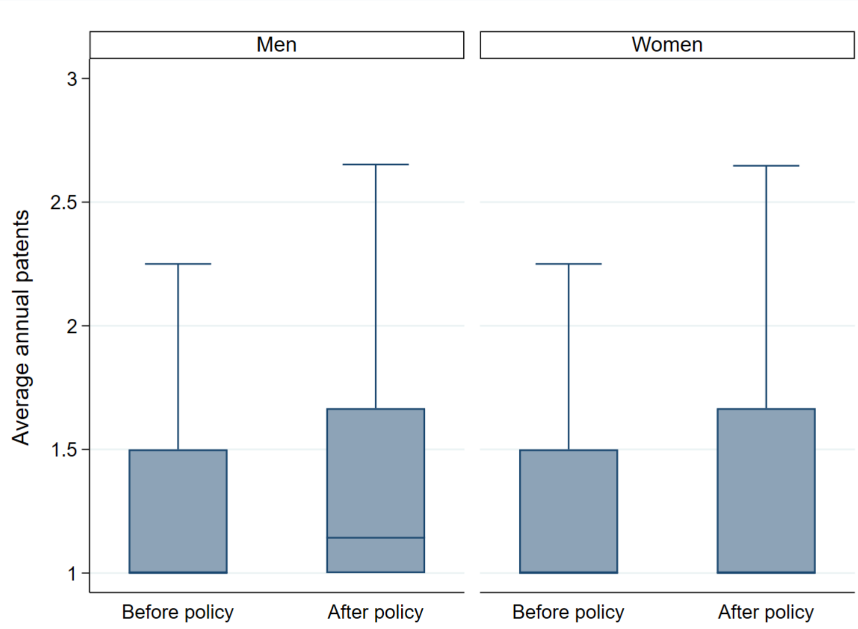
(b) Men inventors



Notes: Panel 4.3a displays the densities of exit age for women and panel 4.3b displays the densities of exit age for men. Years before maternity leave policy are represented by the solid line, whereas years after are represented by the dashed line.

We further look at the patterns of productivity by gender. We restrict career inventors to those aged 18 to 100 who patented before and after maternity leave policy. We then compute the average yearly number of patents produced for each inventor in both periods. The boxplots in Figure 4.4 show that the productivity increased slightly after maternity leave policy passage for men but not for women.

Figure 4.4 – Productivity by Gender Before vs. After Policy



Notes: Outlier values are excluded.

For the rest of the analysis, we restrict our sample to career inventors who were 25–64 at the time of maternity leave policy adoption in their state of residence. Table 4.2 presents the descriptive statistics for women and men inventors. Interestingly, the average age of women inventors is about four years lower than men inventors when maternity policy passed. Women also tend to be younger on average when they applied for a patent. On average, women inventors have a higher probability of exiting than men inventors, albeit with little differences in annual productivity by gender. Finally, a higher portion of women inventors occurred in the sample after maternity leave policies

Table 4.2 – Descriptive Statistics

	Women inventors				Men inventors			
	Mean	S.D.	Max.	Min.	Mean	S.D.	Max.	Min.
Age at policy adoption	37.88	8.52	64	25	41.62	9.71	64	25
Age at patent application	38.78	8.43	69	19	42.09	9.51	69	19
Exit	0.20	0.40	1	0	0.17	0.37	1	0
Log (annual patents)	0.29	8.52	4.41	0	0.30	0.50	5.18	0
Maternity	0.68	0.47	1	0	0.62	0.48	1	0

4.5 Results

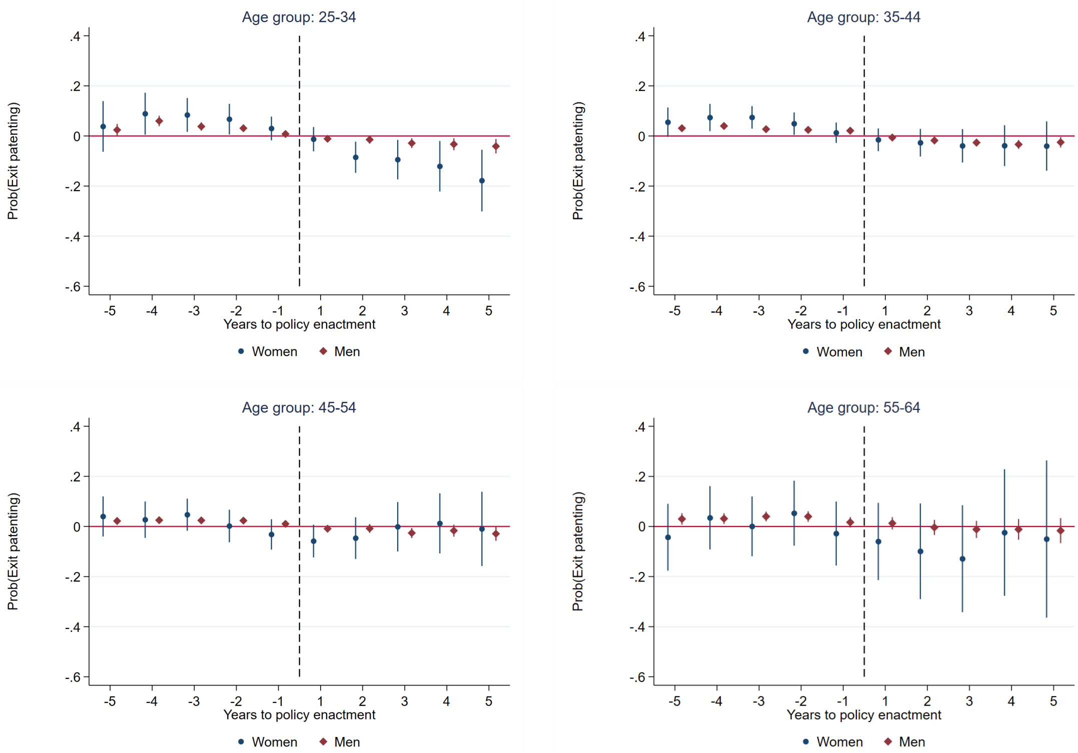
4.5.1 Inventors' Probability to Exit Patenting

Figure 4.5 depicts the patterns of exit in response to maternity leave policy change for women and men inventors at different stages of their lives. Among women inventors, those aged 25 to 34 have a significant decrease in the probability to exit patenting in the years after maternity leave adoption. Relative to their probability of exit at the year of leave policy adoption, these women are 8.3 percent ($p = 0.015$) more likely to quit patenting permanently three years before leave passage. On the opposite, they become 9.4 percent ($p = 0.019$) less likely to exit three years into policy change; and five years later, these inventors become 17.8 percent ($p = 0.005$) less likely to quit The decline in the probability to exit after policy enactment is much smaller for women aged 35 to 44 and becomes indifferent to zero for women aged 45 to 54. The wide confidence interval for women aged 55 to 64 can be explained by a relatively small sample of inventors in this category.

By contrast, the event-study estimates for men inventors are generally small in magnitude relative to the benchmark year. For example, men inventors aged 25 to 34 are only about 4.1 percent ($p = 0.005$) less likely to exit patenting after five years into policy enactment. These findings suggest that the implementation of maternity leave may help women inventors stay longer in a patenting career. The estimates obtained from the regressions can be found in Table C.1.

Event Study Estimates

Figure 4.5 – Exit Decision around Maternity Leave Policy Enactment by Gender for Different Age Groups



Notes: The event study estimates are obtained from the OLS regression on Equation 4.1 with 95% confidence intervals. We removed one-shot inventors that only occurred once in our sample. The age group 25–34 includes 7,523 women inventors and 61,518 men inventors. The age group 35–44 includes 6,031 women inventors and 67,905 men inventors. The age group 45–54 includes 2,439 women inventors and 46,236 men inventors. The age group 55–64 includes 703 women inventors and 23,208 men inventors.

Survival Analysis

To provide further evidence on how women and men inventors have survived in the years around the adoption of maternity leave policy, we then perform a non-parametric estimation on the survival function

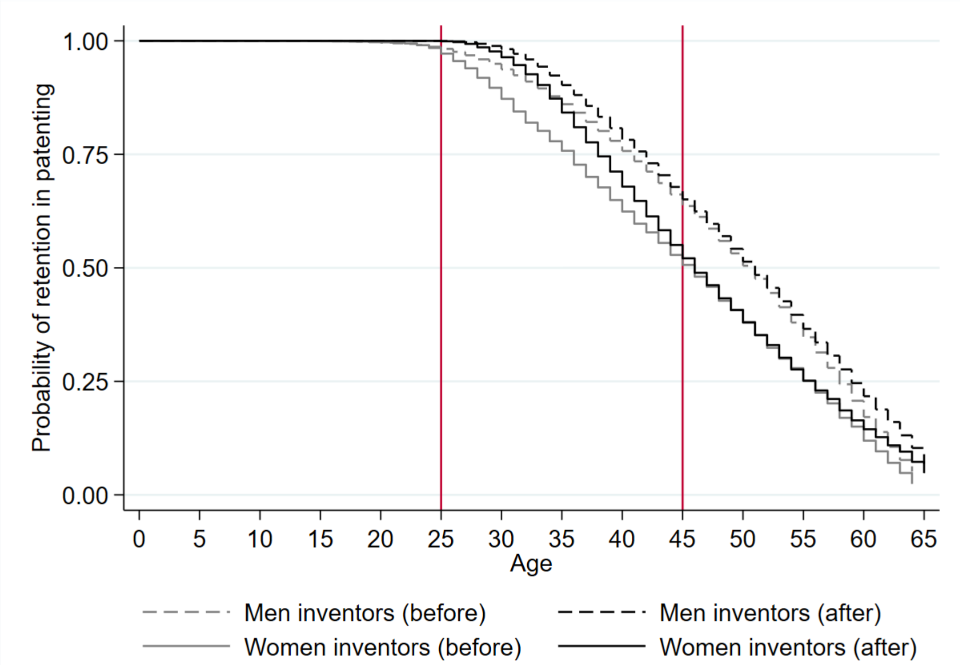
$$P(k) = \prod_{i:k_i \leq k} \left(1 - \frac{Exit_i}{N_i}\right)$$

where $P(k)$ is the probability of having retained in patenting at age k. $Exit_i$ is the number of incidences on inventor exit and N_i is the number of inventors having survived.

Figure 4.6 reports separate Kaplan-Meier survival curves for men (in dashed curves) and women (in solid curves) in the years before maternity leave policy adoption and afterward.

The y-axis displays the probability of survival in patenting, and the x-axis displays the age at which inventors stop patenting. We expand the time horizon to 10 years around the maternity leave policy adoption. At almost all ages, the curves for women are beneath those for men, revealing a persistent gender gap in retention in inventive careers. The plots for women show that maternity leave policy helps women of child-bearing age (between 25 and 44) stay in their job, consistent with previous evidence on the delayed exit in the years after policy adoption from descriptive statistics. The probability of staying in patenting substantially increased from 75.78 percent to 84.21 percent for women at 35. Moreover, in the period after the maternity leave policy, the gender gap in retention closes for inventors between 25 to 44. On the other hand, the plots for men large overlap between the two periods despite a slight increase in the likelihood to continue patenting for men aged 30 to 40.

Figure 4.6 – Kaplan-Meier Survival Estimates on Probability of Continuing Patenting

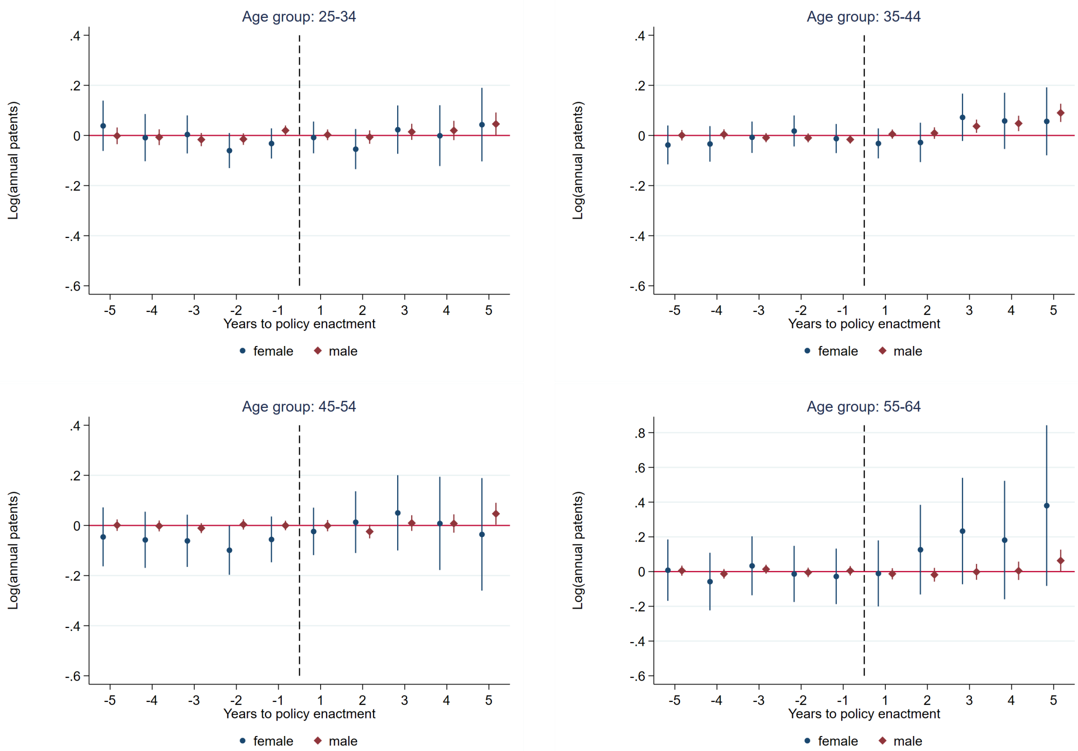


Notes: The stratified log-rank test shows $P < 0.01$, rejecting the null hypothesis that the survival functions among all groups are indifferent. We removed one-shot inventors that only occurred once in our sample. The sample includes 24,815 women inventors and 276,760 men inventors.

4.5.2 Inventors' Productivity in Patenting

Maternity leave may also affect women's productivity at work. Women who return to work after giving birth thanks to a job-protected maternity leave policy may still have to deal with interruptions. Figure 4.7 presents the event study estimates on the effect of maternity leave passage on productivity in patenting. Surprisingly, the change in patents produced per year is almost negligible for women inventors of all age groups. The insignificant estimates for women inventors aged 55 to 64 are likely to have resulted from a limited sample size. What seems interesting is that the productivity for men slightly increased a few years after the implementation of maternity leave. In particular, men of age 35 to 44 experienced an increase in productivity by about 9 percent ($p = 0.005$) five years after the passage of maternity leave. On the other hand, estimates on the sample of women of similar ages suggest increased productivity in patenting. This piece of evidence is indicative of Kaltenberg, Jaffe, and Lachman (2021)'s finding that the productivity of inventors peaked in their late 30s and early 40s for women and men, respectively. The estimates obtained from the regressions can be found in Table C.2.

Figure 4.7 – Productivity around Maternity Leave Policy Enactment by Gender for Different Age Groups



Notes: The event study estimates are obtained from the OLS regression on Equation 4.2 with 95% confidence intervals. We removed one-shot inventors that only occurred once in our sample. The age group 25–34 includes 7,523 women inventors and 61,518 men inventors. The age group 35–44 includes 6,031 women inventors and 67,905 men inventors. The age group 45–54 includes 2,439 women inventors and 46,236 men inventors. The age group 55–64 includes 703 women inventors and 23,208 men inventors.

4.6 Conclusions

Despite progress in the increase of female inventors over the past 40 years, women still only make up 15% of the inventor population in the US as of 2015. Women face several obstacles across the career pipeline in STEM and related scientific disciplines to patenting activities. One notable obstacle that women face is career interruptions that may occur with motherhood. Goldin and Mitchell (2017) show that women's labor force participation drops following the birth of their first child, but this decline is not as severe for women with access to maternity leave. Policies that aim to improve job attachment and the well-being of mothers are of particular interest to understand how these policies impact women in high-skilled occupations, particularly those that contribute to innovation activities that have positive impacts on the overall economy.

In this chapter, we examine the effect of maternity leave policy, notably state provisions and the FMLA, on the propensity to exit and productivity in patenting for men and women inventors. We find that the exit patterns for women aged 25 to 34 are significantly distinctive after the policy enactment—they are less likely to exit relative to the benchmark year. By contrast, the patterns of exit change little for men. Our survival estimates further confirm the increased probability for women aged 25 to 44 to stay in patenting. However, theses policies did not have an impact on the productivity of patenting for women. Though, patenting is a discrete outcome of years of work and may not reflect the consistent input of effort at work, thus compromising to be the best measure of productivity.

These results reflect that the maternity leave provisions and FMLA were successful in retaining new mothers in the innovation labor force. By securing employment for a few weeks, new mothers could choose to stay active members of the labor force and maintain consistent employment. However, women still exit more frequently than men during the reproductive years, and the policy was unable to close fully close the gap of quitting patenting between men and women.

Regarding the limitations, we are unable to identify which women had children as we only observe exit and productivity patterns based on ages. We also do not have information about if firm-specific maternity policies applied to women inventors. However, we suspect that our estimates are biased downward without these considerations.

While job-protective unpaid maternity leave improves the retention for women of reproductive ages, the absence of accessible childcare may still prevent mothers from maintaining or achieving high productivity. The most productive time period of inventors in terms of patenting activity is during their 30s–40s (Kaltenberg, Jaffe, & Lachman, 2021), and there is a gender gap of productivity that persists. We show no evidence that FMLA may bridge this productivity gap during the motherhood years. Childcare subsidies could potentially help increasing female labor participation (Haan & Wrohlich, 2011), but its impact on increasing productivity has been understudied. One should expect a higher level of innovative output had skilled women been able to continue inventing with the support of family-friendly policies that leave them free of interruptions.

5 Conclusion

Economic growth depends on the creation and diffusion of new technologies. Innovation policies that target technological progress are regarded as the cure for the increasing challenges in knowledge-based economies (David & Foray, 2003; Kremer & Williams, 2010; Mazzucato, 2016). The present dissertation seeks to provide empirical evidence on how institutions and policies support innovation, focusing on the current patent system and barriers to entry in R&D occupations.

Debates regarding the current patent system on contentious issues such as its efficiency to provide incentives or the strengthening of patent protection have been going on for years (Jaffe & Lerner, 2011). While scholars have known that patents may not be the most effective instrument for firms to appropriate returns in some industries (Mansfield, 1986; Levin et al., 1987; Cohen et al., 2000), it is largely unclear to what extent patent protection guarantees monopolistic pricing over the products. Moreover, despite the evidence that the disclosure of patented information may bring positive social externalities (Furman et al., 2018; Lück et al., 2020), it remains under-studies whether firms' signaling of patent portfolio affects knowledge diffusion to competitors. The optimal design of patent policies to induce innovation calls for new empirical evidence. Chapter 2 and chapter 3 of this dissertation thus concentrate on two sought-after questions on the patent system, namely, (i) does patent protection confer sufficient markups within its term for broadly-defined consumer goods and (ii) how does firms' signaling of innovative assets affect the diffusion of inventions across firms.

On the other hand, innovation depends on the knowledge workers who develop new products and processes. Diversity in the knowledge workforce not only boosts firm performance but also contributes to economic prosperity (Østergaard et al., 2011; Hsieh et al., 2019). However, women are underrepresented at several points of the innovation pipeline, from the entry into STEM majors to the participation in inventive occupations and commercial science (Xie et al., 2003; Murray & Graham, 2007; Hunt et al., 2013). Understanding the causes of how women leak out of the pipeline is thus crucial for designing policies to close the gender gap. Previous research shows women put substantial weight on work and life balance when it comes to career decisions (Barbulescu & Bidwell, 2013). Chapter 4 therefore focuses on the question of how job-protective family policies affect women's participation in inventive occupations.

Linking patents to commercial products and their prices, chapter 2 studies the effect of patent expiry on the product prices and explores the underlying mechanism. Our findings suggest that patents confer a certain level of appropriability for a group of consumer products by allowing firms to charge supra-competitive prices within the patent term. Firms may reduce prices preemptively to deter entry as the drop in product prices is more salient in markets where competition is more intense upon patent expiry. The novel data on product to patent concordance allows us to gauge the changes in monopoly pricing conferred by patents directly on the level of innovative output, which has been a challenging task in the literature. These pieces of evidence on the loss of price premium due to losing patent protection serves as a starting position to understand the social welfare associated with patent protection.

Taking advantage of a policy change in patent marking since 2011, chapter 3 exploits the time-stamped differences in virtual patent marking for 16 firms and analyzes how VPM affects follow-on inventions and their similarities. We find that conditional on patent appropriability regimes, signaling valuable inventions has heterogeneous effects on follow-on inventions from external firms. Virtual marking attracts more follow-on inventions in weak regimes such as consumer goods but fends off follow-on inventions in strong regimes. However, patent importance may play a mitigating role on follow-on invention in strong regimes as the underlying value may get amplified through marking. These findings help us understand how innovative firms' strategic disclosure of proprietary information brings externalities to competing firms and how appropriability affects the direction of knowledge flows.

Chapter 4 exploits the variation in the timing of maternity leave policies across the US and evaluates the impact on the retention and productivity of women inventors. Using data on 1.4 million inventors from USPTO, our results from event studies and survival analysis suggest that maternity leave policies are most effective at promoting the retention of women inventors during the reproductive ages—years close to their productivity peak. However, no evidence suggests that these policies help women to be more productive at patenting. These findings bring insights on how policies can support women of reproductive ages to continue an R&D career and call for more inclusive family policies such as childcare benefits and subsidies that support women at work.

A Appendix

This appendix includes supplementary materials in my thesis. The figures and tables that are referred to in each chapter can be found under relevant section titles.

A Appendix for chapter 1

Table A.1 – A list of representative product(s) by Amazon product catalog

Amazon product catalog	Subcatalog	Firm	Representative product(s)
Appliances	Vacuum	Dyson	AM08 / DC35
Appliances	Ceiling Fans & Accessories	Emerson	CF830 MONACO FAN
Appliances	Vacuum	Kaivac, Inc.	KaiVac
Appliances	Small Appliances	NuWave Now	NuWave® Precision Induction Cooktop (Flex)
Automotive parts	Replacement Parts	ANCO	A-14-M
Automotive parts	Replacement Parts	Bosch	Clear Advantage 28CA
Automotive parts	Replacement Parts	diono	Easy View Mirror
Automotive parts	Accessories	Lippert Components	FLIP™ jack foot
Automotive parts	Towing Products & Winches	Warn Industries	ProVantage Winches
Baby Products	Accessories	Munchkin	Bristle Brush
Baby Products	Strollers & Accessories	phil&teds	Verve Buggy
Clothing, Shoes & Jewelry	Shoes	KEEN	Yogui Arts
Clothing, Shoes & Jewelry	Shoes	Newton Running Company	Aha
Electronics	Camera & Photo	360fly	360FLYBLK
Electronics	Computers & Accessories	Advantech	EKI-2528PAI
Electronics	Cell Phones & Accessories	Belkin	F8Z442
Electronics	Cell Phones & Accessories	BlackBerry	BlackBerry® Classic
Electronics	Computers & Accessories	Brocade	Brocade NetIron CER 2000 Series
Electronics	Computers & Accessories	Cirque Corporation	Gen 3 and earlier
Electronics	Accessories & Supplies	CommScope	Cables Coaxial Braided
Electronics	Computers & Accessories	Control4	C4-TV120277
Electronics	Camera & Photo	Draper, Inc.	Micro Projector Lift
Electronics	Computers & Accessories	Elo Touch Solutions	Touch Screen
Electronics	Computers & Accessories	Honeywell	Voyager 1250g / Xenon 1900g General Duty Scanners
Electronics	Cell Phones & Accessories	HTC	HTC One ® (E8)
Electronics	Computers & Accessories	Kent Displays	Boogie Board™ Original 8.5 eWriter
Electronics	Television & Video	KING Connect	Tailgater® VQ2500
Electronics	Computers & Accessories	Logitech	Logitech G603 Mouse / Logitech K811 Keyboard
Electronics	Computers & Accessories	Mad Catz	Mad Catz V.7 Keyboard
Electronics	Computers & Accessories	Neonode	Neonode AirBar® sensor
Electronics	Computers & Accessories	Oki Data Americas, Inc.	ES3640e MFP
Electronics	Headphones	Skullcandy Inc.	Soundmine
Electronics	Portable Audio & Video	Sonos, Inc.	One
Electronics	Television & Video	Sound United	AV Receiver AVR-4520
Electronics	Computers & Accessories	tyconsystems	Tycon Systems 802.3at
Electronics	Camera & Photo	X-Rite	331C

Health & Household	Beauty & Personal Care	CND	Radical SolarNail™
Health & Household	Medical Supplies & Equipment	Game Ready	Straight Knee Wrap
Health & Household	Beauty & Personal Care	Kao Corporation	Jergens® Shea Butter
Health & Household	Household Supplies	Kimberly-Clark	COTTONELLE® CleanCare Toilet Paper
Health & Household	Household Supplies	Procter & Gamble	Power Razor
Health & Household	Household Supplies	RB	FINISH Powerball Quantum Max Capsules Ultra Degreaser
Industrial & Scientific	Industrial Electrical	American Radionic	Turbo® 200
Industrial & Scientific	Building Supplies	CleanAlert	FILTERSCAN WiFi (FS-245-C)
Industrial & Scientific	Additive Manufacturing Products	MakerBot®	MakerBot Replicator Z18 3D Printer
Industrial & Scientific	Lab & Scientific Products	Multisorb Technologies	TranSorb Humidity Absorber
Industrial & Scientific	Occupational Health & Safety Products	TCP Lighting	Exit Signs
Industrial & Scientific	Occupational Health & Safety Products	UltraTech	Ultra-Microbe Boom
Industrial & Scientific	Professional Medical Supplies	Welch Allyn	Diagnostic Otoscope
Musical Instruments	Electronic Music, DJ & Karaoke	Avid Technology	Pro Tools® | Sync HD
Musical Instruments	Electronic Music, DJ & Karaoke	Native Instruments	NI brand TRAKTOR
Office Product	Office & School Supplies	Avery Products	Addressing Labels
Office Product	Printer Ink & Toner	Epson America Inc.	T0971
Office Product	Accessories	ES Robbins	Mats/Matting
Office Product	Accessories	FireKing Security Group	Media Vault
Office Product	Office & School Supplies	Humanscale	Humanscale Keyboard Systems
Software	Video editing	Corel Corporation	Pinnacle Studio
Software	Antivirus & Security	Symantec	Norton Core
Sports & Outdoors	Electronics & Gadgets	Aqua Lung	i750TC
Sports & Outdoors	Golf Balls	Callaway Golf	Warbird 2.0
Sports & Outdoors	Accessories	CamelBak	Performance Bottle
Sports & Outdoors	Accessories	Everlast Climbing	Traverse Wall® Challenge Course
Sports & Outdoors	Accessories	Hobie	MirageDrive
Sports & Outdoors	Accessories	ISM Seat	Adamo Racing
Sports & Outdoors	Accessories	JumpSport	JumpSport PowerBounce Trampoline (with enclosure)
Sports & Outdoors	Accessories	Move Collective LLC	bobble
Tools & Home Improvement	Lighting	Colonial Tin Works Inc	Solar Lid Lights® 360318

Tools & Home Improvement	Power & Hand Tools	DeckWise	STANDARD Ipe Clip
Tools & Home Improvement	Lighting	Golight Inc.	GXL
Tools & Home Improvement	Accessories & Supplies	Gorilla Ladders	Slim-Fold Work Platform, GLWP-55A
Tools & Home Improvement	Accessories & Supplies	Legrand, North America	Wall Plates
Tools & Home Improvement	Power & Hand Tools	Max USA Corp	Rebar tying tool RB398
Tools & Home Improvement	Lighting	Nanoleaf	Nanoleaf One
Tools & Home Improvement	Power & Hand Tools	Rexair LLC	Rainbow Vacuum System
Tools & Home Improvement	Power & Hand Tools	Ridge Tool Company	V2 Press Ring Actuator
Tools & Home Improvement	Generators & Portable Power	SunPower Corporation	SunPower® Flexible Solar Panel
Video Games	Xbox One	Activision	Skylanders® Trap Team Triple Trap

Table A.2 – Distribution of the sources of P^A and P^L at relevant periods, in percent

Panel A: sources of P^A					
	S_0^A	S_1^A	S_2^A	S_3^A	S_4^A
The month one year before expiry	0.82	34.62	53.63	6.53	4.40
The month of expiry	0.53	26.61	62.68	7.17	3.01
The month one year after expiry	1.08	23.31	63.86	7.99	3.76
Panel B: sources of P^L					
	S_0^L	S_1^L	S_2^L	S_3^L	S_4^L
The month one year before expiry	0.94	7.79	75.57	0.08	15.62
The month of expiry	0.62	4.43	79.06	0.09	15.80
The month one year after expiry	1.03	4.38	77.21	0	17.38

B Appendix for chapter 2

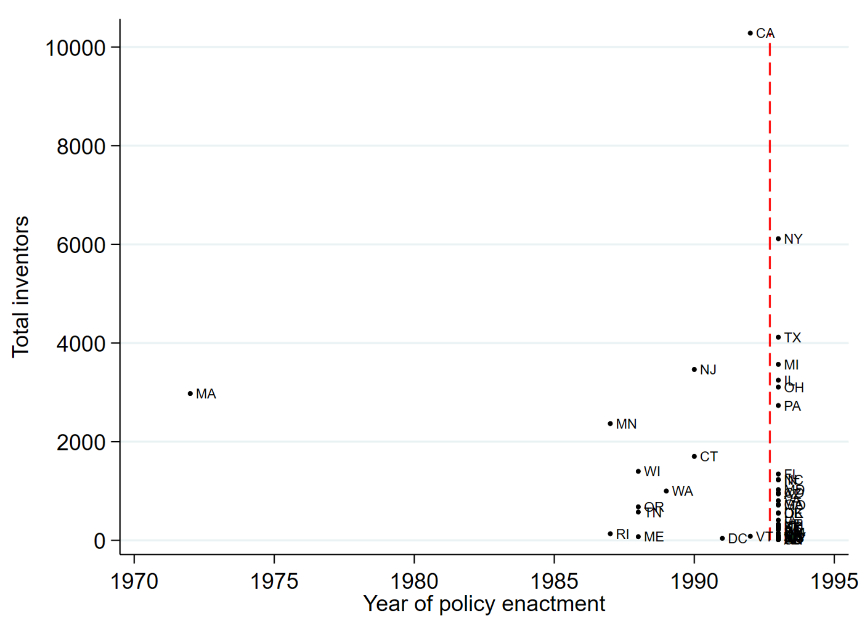
Table B.1 – First stage result of PCA

Principal components			Component loadings for the first component	
Component	Eigenvalue	Proportion	Variable	Comp1
		All patents		
Comp1	1.6845	0.2807	log independent claims	-0.0433
Comp2	1.23958	0.2066	log words in first claim	-0.4017
Comp3	1.08939	0.1816	log IPC class	0.5113
Comp4	0.834212	0.1390	log geographical family	0.5371
Comp5	0.603819	0.1006	log non-patent references	0.4159
Comp6	0.548497	0.0914	originality score	0.3374
		Patents of strong appropriability		
Comp1	2.79615	0.4660	log independent claims	-0.1547
Comp2	1.18016	0.1967	log words in first claim	-0.3565
Comp3	0.704526	0.1174	log IPC class	0.3759
Comp4	0.628635	0.1048	log geographical family	0.5088
Comp5	0.485838	0.0810	log non-patent references	0.5189
Comp6	0.204693	0.0341	originality score	0.4237
		Patents of weak appropriability		
Comp1	1.50852	0.2514	log independent claims	0.1335
Comp2	1.27717	0.2129	log words in first claim	-0.2044
Comp3	1.13316	0.1889	log IPC class	0.6319
Comp4	0.909327	0.1516	log geographical family	0.5631
Comp5	0.611369	0.1019	log non-patent references	-0.4379
Comp6	0.560459	0.0934	originality score	-0.1794

C Appendix for chapter 3

Figure C.1 – Timing of policy enactment and innovation capacity by state in 1992

(a) Total Inventors



(b) Total Patents

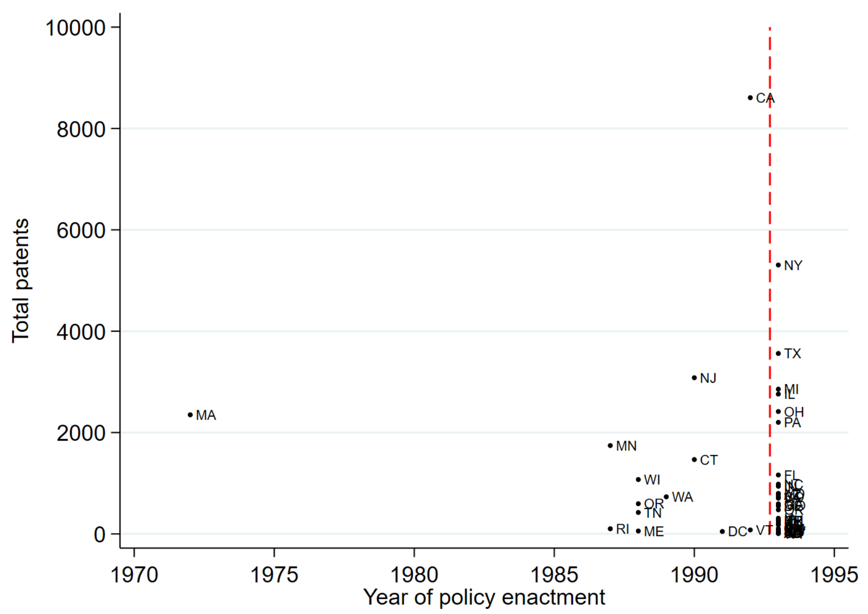


Figure C.2 – share of women inventors by age group

(a) All inventors (b) Career inventors

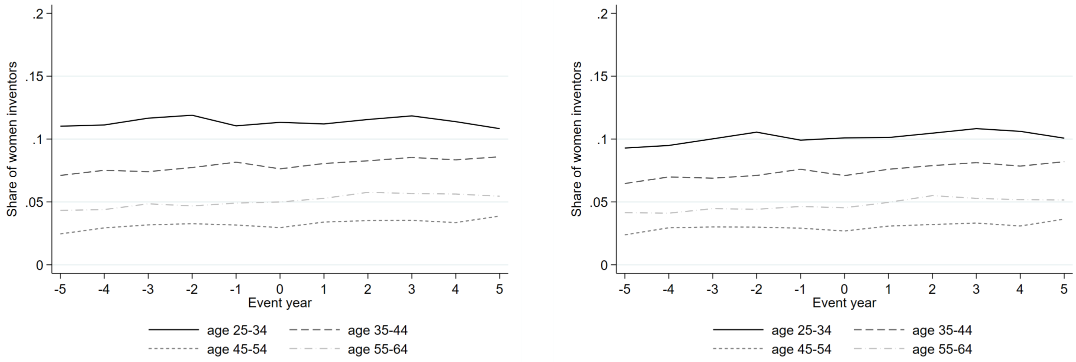


Table C.1 – Event study regressions on inventor retention

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	25–34	35–44	45–54	55-64	25–34	35–44	45–54	55-64
VARIABLES					$Exit = 1$			
	Female inventors				Male inventors			
5 years before	0.038	0.055*	0.040	-0.043	0.024**	0.031***	0.022***	0.030***
	(0.052)	(0.030)	(0.041)	(0.068)	(0.012)	(0.007)	(0.008)	(0.012)
4 years before	0.089**	0.074***	0.027	0.035	0.060***	0.040***	0.025***	0.032***
	(0.043)	(0.028)	(0.037)	(0.064)	(0.011)	(0.006)	(0.007)	(0.011)
3 years before	0.084**	0.074***	0.047	0.001	0.038***	0.027***	0.024***	0.040***
	(0.034)	(0.023)	(0.033)	(0.061)	(0.009)	(0.005)	(0.006)	(0.010)
2 years before	0.067**	0.049**	0.002	0.053	0.031***	0.024***	0.023***	0.040***
	(0.031)	(0.023)	(0.033)	(0.066)	(0.008)	(0.005)	(0.007)	(0.011)
1 year before	0.030	0.013	-0.032	-0.028	0.008	0.021***	0.010	0.017
	(0.024)	(0.021)	(0.031)	(0.065)	(0.006)	(0.005)	(0.006)	(0.011)
1 year after	-0.013	-0.015	-0.058*	-0.060	-0.011*	-0.006	-0.009	0.013
	(0.025)	(0.023)	(0.033)	(0.079)	(0.007)	(0.005)	(0.007)	(0.013)
2 years after	-0.085***	-0.027	-0.046	-0.099	-0.015*	-0.018***	-0.008	-0.004
	(0.032)	(0.028)	(0.042)	(0.097)	(0.008)	(0.007)	(0.009)	(0.015)
3 years after	-0.094**	-0.039	-0.001	-0.129	-0.029***	-0.026***	-0.026**	-0.012
	(0.040)	(0.034)	(0.050)	(0.109)	(0.010)	(0.008)	(0.010)	(0.018)
4 years after	-0.121**	-0.039	0.012	-0.024	-0.033***	-0.034***	-0.017	-0.011
	(0.051)	(0.042)	(0.061)	(0.129)	(0.012)	(0.009)	(0.012)	(0.021)
5 years after	-0.178***	-0.040	-0.010	-0.050	-0.041***	-0.025**	-0.029**	-0.017
	(0.063)	(0.050)	(0.076)	(0.160)	(0.015)	(0.011)	(0.014)	(0.025)
Constant	0.253***	0.168***	0.176***	0.237***	0.135***	0.126***	0.149***	0.193***
	(0.025)	(0.019)	(0.025)	(0.048)	(0.006)	(0.004)	(0.005)	(0.008)
Observations	12,038	11,636	5,175	1,572	105,720	143,404	101,654	49,660
R-squared	0.460	0.428	0.426	0.430	0.429	0.409	0.406	0.416
Inventor FE	YES	YES	YES	YES	YES	YES	YES	YES
State FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table C.2 – Event study regressions on inventor productivity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	25–34	35–44	45–54	55-64	25–34	35–44	45–54	55-64
VARIABLES	$log(patents+1)$							
	Female inventors				Male inventors			
5 years before	0.039	-0.038	-0.045	0.008	-0.001	0.001	0.001	0.005
	(0.051)	(0.040)	(0.060)	(0.090)	(0.017)	(0.011)	(0.012)	(0.015)
4 years before	-0.009	-0.034	-0.057	-0.057	-0.007	0.004	-0.002	-0.013
	(0.048)	(0.036)	(0.057)	(0.085)	(0.016)	(0.010)	(0.011)	(0.014)
3 years before	0.004	-0.007	-0.061	0.033	-0.016	-0.009	-0.011	0.014
	(0.039)	(0.032)	(0.053)	(0.087)	(0.013)	(0.009)	(0.010)	(0.013)
2 years before	-0.060*	0.018	-0.099**	-0.013	-0.014	-0.010	0.004	-0.005
	(0.036)	(0.032)	(0.050)	(0.082)	(0.012)	(0.009)	(0.010)	(0.014)
1 year before	-0.032	-0.012	-0.056	-0.027	0.020**	-0.016*	-0.000	0.005
	(0.031)	(0.030)	(0.047)	(0.081)	(0.010)	(0.008)	(0.010)	(0.014)
1 year after	-0.008	-0.032	-0.024	-0.010	0.003	0.005	-0.001	-0.013
	(0.032)	(0.031)	(0.048)	(0.097)	(0.011)	(0.009)	(0.012)	(0.016)
2 years after	-0.054	-0.028	0.013	0.126	-0.006	0.009	-0.024*	-0.018
	(0.041)	(0.040)	(0.063)	(0.131)	(0.014)	(0.012)	(0.014)	(0.020)
3 years after	0.023	0.072	0.050	0.234	0.014	0.037***	0.010	-0.002
	(0.049)	(0.048)	(0.077)	(0.156)	(0.016)	(0.013)	(0.016)	(0.023)
4 years after	-0.001	0.058	0.008	0.182	0.020	0.048***	0.008	0.005
	(0.062)	(0.057)	(0.095)	(0.174)	(0.020)	(0.016)	(0.019)	(0.027)
5 years after	0.043	0.056	-0.035	0.380	0.046*	0.090***	0.047**	0.063**
	(0.075)	(0.069)	(0.115)	(0.236)	(0.023)	(0.019)	(0.022)	(0.032)
Constant	0.292***	0.284***	0.324***	0.246***	0.309***	0.298***	0.308***	0.291***
	(0.032)	(0.026)	(0.038)	(0.067)	(0.010)	(0.007)	(0.008)	(0.010)
Observations	12,038	11,636	5,175	1,572	105,720	143,404	101,654	49,660
R-squared	0.504	0.466	0.460	0.492	0.476	0.446	0.446	0.456
Inventor FE	YES	YES	YES	YES	YES	YES	YES	YES
State FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Bibliography

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

Schmookler, J. (1957). Inventors past and present. *The Review of Economics and Statistics, 39*(3), 321–333.

Arrow, K. (1962). Economic Welfare and the Allocation of Resources for Invention. *The rate and direction of inventive activity: economic and social factors* (pp. 609–626). Princeton University Press.

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

Kamien, M. I., & Schwartz, N. L. (1974). Patent Life and R & D Rivalry. *American Economic Review, 64*(1), 183–187.

Kitch, E. W. (1977). The nature and function of the patent system. *The Journal of Law and Economics, 20*(2), 265–290.

Milgrom, P., & Roberts, J. (1982). Limit Pricing and Entry under Incomplete Information: An Equilibrium Analysis. *Econometrica*, 443–459.

Bhattacharya, S., & Ritter, J. R. (1983). Innovation and communication: signalling with partial disclosure. *The Review of Economic Studies, 50*(2), 331–346.

Nelson, R. R., & Langlois, R. N. (1983). Industrial innovation policy: lessons from american history. *Science, 219*(4586), 814–818.

Daymont, T. N., & Andrisani, P. J. (1984). Job preferences, college major, and the gender gap in earnings. *Journal of Human Resources*, 408–428.

Horstmann, I., MacDonald, G. M., & Slivinski, A. (1985). Patents as Information Transfer Mechanisms: to Patent or (Maybe) Not to Patent. *Journal of Political Economy, 93*(5), 837–858.

Judd, K. L. (1985). On the Performance of Patents. *Econometrica*, 567–585.

Mansfield, E. (1986). Patents and Innovation: An Empirical Study. *Management Science, 32*(2), 173–181.

Teece, D. J. (1986). Profiting from technological innovation: implications for integration, collaboration, licensing and public policy. *Research Policy, 15*(6), 285–305.

Levin, R. C., Klevorick, A. K., Nelson, R. R., Winter, S. G., Gilbert, R., & Griliches, Z. (1987). Appropriating the Returns from Industrial Research and Development. *Brookings Papers on Economic Activity*, *1987*(3), 783–831.

Blau, D. M., & Robins, P. K. (1989). Fertility, employment, and child-care costs. *Demography*, *26*(2), 287–299.

Goldin, C. (1989). Life-cycle labor-force participation of married women: historical evidence and implications. *Journal of Labor Economics*, *7*(1), 20–47.

Gilbert, R., & Shapiro, C. (1990). Optimal Patent Length and Breadth. *RAND Journal of Economics*, 106–112.

Klemperer, P. (1990). How Broad Should the Scope of Patent Protection Be? *RAND Journal of Economics*, 113–130.

Romer, P. M. (1990). Endogenous Technological Change. *Journal of political Economy*, *98*(5, Part 2), S71–S102.

Scotchmer, S., & Green, J. (1990). Novelty and disclosure in patent law. *The RAND Journal of Economics*, 131–146.

Trajtenberg, M. (1990). A penny for your quotes: patent citations and the value of innovations. *RAND Journal of Economics*, 172–187.

Waterson, M. (1990). The Economics of Product Patents. *American Economic Review*, 860–869.

Albert, M. B., Avery, D., Narin, F., & McAllister, P. (1991). Direct validation of citation counts as indicators of industrially important patents. *Research policy*, *20*(3), 251–259.

Bresnahan, T. F., & Reiss, P. C. (1991). Entry and Competition in Concentrated Markets. *Journal of Political Economy*, *99*(5), 977–1009.

Caves, R. E., Whinston, M. D., Hurwitz, M. A., Pakes, A., & Temin, P. (1991). Patent Expiration, Entry, and Competition in the US Pharmaceutical Industry. *Brookings Papers on Economic Activity. Microeconomics*, *1991*, 1–66.

Scotchmer, S. (1991). Standing on the shoulders of giants: cumulative research and the patent law. *Journal of economic perspectives*, *5*(1), 29–41.

Aghion, P., & Howitt, P. (1992). A Model of Growth Through Creative Destruction. *Econometrica*, *60*(2), 323–351.

Grabowski, H. G., & Vernon, J. M. (1992). Brand Loyalty, Entry, and Price Competition in Pharmaceuticals after the 1984 Drug Act. *Journal of Law and Economics*, *35*(2), 331–350.

Parkman, A. M. (1992). Unilateral divorce and the labor-force participation rate of married women, revisited. *The American Economic Review*, *82*(3), 671–678.

Ribar, D. C. (1992). Child care and the labor supply of married women: reduced form evidence. *Journal of human resources*, 134–165.

Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *the Quarterly journal of Economics*, *108*(3), 577–598.

Grossman, G. M., & Helpman, E. (1994). Endogenous innovation in the theory of growth. *Journal of Economic Perspectives*, *8*(1), 23–44.

Gruber, J. (1994). The incidence of mandated maternity benefits. *The American economic review*, 622–641.

Lerner, J. (1994). The Importance of Patent Scope: An Empirical Analysis. *RAND Journal of Economics*, 319–333.

Preston, A. E. (1994). Why have all the women gone? a study of exit of women from the science and engineering professions. *The American Economic Review*, *84*(5), 1446–1462.

Harabi, N. (1995). Appropriability of Technical Innovations: An Empirical Analysis. *Research Policy*, *24*(6), 981–992.

Brynjolfsson, E., & Kemerer, C. F. (1996). Network Externalities in Microcomputer Software: An Econometric Analysis of the Spreadsheet Market. *Management Science*, *42*(12), 1627–1647.

Denicolo, V. (1996). Patent Races and Optimal Patent Breadth and Length. *Journal of Industrial Economics*, 249–265.

Matutes, C., Regibeau, P., & Rockett, K. (1996). Optimal Patent Design and the Diffusion of Innovations. *RAND Journal of Economics*, 60–83.

Botosan, C. A. (1997). Disclosure level and the cost of equity capital. *Accounting review*, 323–349.

Cowan, R., & Foray, D. (1997). The economics of codification and the diffusion of knowledge. *Industrial and corporate change*, *6*(3), 595–622.

Frank, R. G., & Salkever, D. S. (1997). Generic Entry and the Pricing of Pharmaceuticals. *Journal of Economics & Management Strategy*, *6*(1), 75–90.

Park, W. G., & Ginarte, J. C. (1997). Intellectual property rights and economic growth. *Contemporary Economic Policy*, *15*(3), 51–61.

Trajtenberg, M., Henderson, R., & Jaffe, A. (1997). University versus corporate patents: a window on the basicness of invention. *Economics of Innovation and new technology*, *5*(1), 19–50.

Chandy, R. K., & Tellis, G. J. (1998). Organizing for Radical Product Innovation: the Overlooked Role of Willingness to Cannibalize. *Journal of Marketing Research*, *35*(4), 474–487.

Gendell, M. (1998). Trends in retirement age in four countries, 1965-95. *Monthly Lab. Rev.*, *121*, 20.

Hobday, M. (1998). Product Complexity, Innovation and Industrial Organisation. *Research Policy*, *26*(6), 689–710.

Kimmel, J. (1998). Child care costs as a barrier to employment for single and married mothers. *Review of Economics and Statistics*, *80*(2), 287–299.

Lanjouw, J. O. (1998). Patent protection in the shadow of infringement: simulation estimations of patent value. *Review of Economic Studies*, *65*(4), 671–710.

O'Donoghue, T., Scotchmer, S., & Thisse, J.-F. (1998). Patent Breadth, Patent Life, and the Pace of Technological Progress. *Journal of Economics & Management Strategy*, *7*(1), 1–32.

Schankerman, M. (1998). How valuable is patent protection? estimates by technology field. *RAND Journal of Economics*, 77–107.

Harhoff, D., Narin, F., Scherer, F. M., & Vopel, K. (1999). Citation Frequency and the Value of Patented Inventions. *Review of Economics and statistics*, *81*(3), 511–515.

Milyo, J., & Waldfogel, J. (1999). The Effect of Price Advertising on Prices: Evidence in the Wake of 44 Liquormart. *American Economic Review*, *89*(5), 1081–1096.

Waldfogel, J. (1999). Family leave coverage in the 1990s. *Monthly Lab. Rev.*, *122*, 13.

Cohen, W. M., Nelson, R. R., & Walsh, J. P. (2000). *Protecting Their Intellectual Assets: Appropriability Conditions and Why US Manufacturing Firms Patent (or Not)* (tech. rep.). National Bureau of Economic Research.

Cowan, R., David, P. A., & Foray, D. (2000). The explicit economics of knowledge codification and tacitness. *Industrial and corporate change*, *9*(2), 211–253.

Hall, B., & Van Reenen, J. (2000). How effective are fiscal incentives for r&d? a review of the evidence. *Research policy*, *29*(4-5), 449–469.

Lemley, M. A. (2000). Rational ignorance at the patent office. *Nw. UL Rev.*, *95*, 1495.

Martin, S., & Scott, J. T. (2000). The Nature of Innovation Market Failure and the Design of Public Support for Private Innovation. *Research Policy*, *29*(4-5), 437–447.

Martin, S. P. (2000). Diverging fertility among us women who delay childbearing past age 30. *Demography*, *37*(4), 523–533.

Morton, F. M. S. (2000). Barriers to Entry, Brand Advertising, and Generic Entry in the US Pharmaceutical Industry. *International Journal of Industrial Organization*, *18*(7), 1085–1104.

Averett, S. L., & Whittington, L. A. (2001). Does maternity leave induce births? *Southern Economic Journal*, 403–417.

Healy, P. M., & Palepu, K. G. (2001). Information asymmetry, corporate disclosure, and the capital markets: a review of the empirical disclosure literature. *Journal of accounting and economics*, *31*(1-3), 405–440.

Anderson, D. J., Binder, M., & Krause, K. (2002). The motherhood wage penalty: which mothers pay it and why? *American economic review*, *92*(2), 354–358.

Cohen, W. M., Goto, A., Nagata, A., Nelson, R. R., & Walsh, J. P. (2002). R&d spillovers, patents and the incentives to innovate in japan and the united states. *Research policy*, *31*(8-9), 1349–1367.

Gallini, N., & Scotchmer, S. (2002). Intellectual property: when is it the best incentive system? *Innovation policy and the economy*, *2*, 51–77.

Goldin, C., & Katz, L. F. (2002). The power of the pill: oral contraceptives and women's career and marriage decisions. *Journal of political Economy*, *110*(4), 730–770.

Hart, S. L., & Christensen, C. M. (2002). The Great Leap: driving Innovation from the Base of the Pyramid. *MIT Sloan Management Review*, *44*(1), 51.

Baum, C. L. (2003a). The effect of state maternity leave legislation and the 1993 family and medical leave act on employment and wages. *Labour Economics*, *10*(5), 573–596.

Baum, C. L. (2003b). The effects of maternity leave legislation on mothers' labor supply after childbirth. *Southern Economic Journal*, 772–799.

David, P. A., & Foray, D. (2003). Economic fundamentals of the knowledge society. *Policy futures in education*, *1*(1), 20–49.

Harhoff, D., Scherer, F. M., & Vopel, K. (2003). Citations, family size, opposition and the value of patent rights. *Research policy*, *32*(8), 1343–1363.

Xie, Y., Shauman, K. A., & Shauman, K. A. (2003). *Women in science: career processes and outcomes* (Vol. 26). Harvard university press Cambridge, MA.

Anton, J. J., & Yao, D. A. (2004). Little patents and big secrets: managing intellectual property. *RAND Journal of Economics*, 1–22.

Arora, A., Fosfuri, A., & Gambardella, A. (2004). *Markets for Technology: the Economics of Innovation and Corporate Strategy*. MIT press.

Berger, L. M., & Waldfogel, J. (2004). Maternity leave and the employment of new mothers in the united states. *Journal of Population Economics*, *17*(2), 331–349.

Chatterji, P., & Markowitz, S. (2004). *Does the length of maternity leave affect maternal health?* (Tech. rep.). National Bureau of Economic Research.

Ginther, D. K., & Kahn, S. (2004). Women in economics: moving up or falling off the academic career ladder? *Journal of Economic perspectives*, *18*(3), 193–214.

Sharman, D. M., & Yassine, A. A. (2004). Characterizing Complex Product Architectures. *Systems Engineering*, *7*(1), 35–60.

Sosa, M. E., Eppinger, S. D., & Rowles, C. M. (2004). The Misalignment of Product Architecture and Organizational Structure in Complex Product Development. *Management Science*, *50*(12), 1674–1689.

Wiggins, S. N., & Maness, R. (2004). Price Competition in Pharmaceuticals: The Case of Anti-Infectives. *Economic Inquiry*, *42*(2), 247–263.

Bettinger, E. P., & Long, B. T. (2005). Do faculty serve as role models? the impact of instructor gender on female students. *American Economic Review*, *95*(2), 152–157.

Erkal, N. (2005). The Decision to Patent, Cumulative Innovation, and Optimal Policy. *International Journal of Industrial Organization*, *23*(7-8), 535–562.

Hall, B. H., Jaffe, A., & Trajtenberg, M. (2005). Market Value and Patent Citations. *RAND Journal of Economics*, 16–38.

Lemley, M. A., & Shapiro, C. (2005). Probabilistic Patents. *Journal of Economic Perspectives*, *19*(2), 75–98.

Mirowsky, J. (2005). Age at first birth, health, and mortality. *Journal of health and social behavior*, *46*(1), 32–50.

Moore, K. A. (2005). Worthless patents. *Berkeley Technology Law Journal*, *20*, 1521–1552.

Moser, P. (2005). How do patent laws influence innovation? evidence from nineteenth-century world's fairs. *American Economic Review*, *95*(4), 1214–1236.

Stephan, P. E., & Levin, S. G. (2005). Leaving careers in it: gender differences in retention. *The Journal of Technology Transfer*, *30*(4), 383–396.

Björklund, A. (2006). Does family policy affect fertility? *Journal of population economics*, *19*(1), 3–24.

Blind, K., Edler, J., Frietsch, R., & Schmoch, U. (2006). Motives to patent: empirical evidence from germany. *Research Policy*, *35*(5), 655–672.

Ding, W. W., Murray, F., & Stuart, T. E. (2006). Gender differences in patenting in the academic life sciences. *science*, *313*(5787), 665–667.

Juhn, C., & Potter, S. (2006). Changes in labor force participation in the united states. *Journal of Economic Perspectives*, *20*(3), 27–46.

Nard, C. A., & Morriss, A. P. (2006). Constitutionalizing patents: from venice to philadelphia. *Review of Law & Economics*, *2*(2), 223–321.

Hall, B. H. (2007). Patents and Patent Policy. *Oxford Review of Economic Policy*, *23*(4), 568–587.

Hall, B. H., Thoma, G., & Torrisi, S. (2007). The Market Value of Patents and R&D: evidence from European Firms. *Academy of Management Proceedings*, *2007*(1), 1–6.

Murray, F., & Graham, L. (2007). Buying science and selling science: gender differences in the market for commercial science. *Industrial and Corporate Change*, *16*(4), 657–689.

Arora, A., Ceccagnoli, M., & Cohen, W. M. (2008). R&D and the Patent Premium. *International Journal of Industrial Organization*, *26*(5), 1153–1179.

Ashenfelter, O. (2008). Predicting the Quality and Prices of Bordeaux Wine. *Economic Journal*, *118*(529), F174–F184.

Bessen, J. (2008). The Value of US Patents by Owner and Patent Characteristics. *Research Policy*, *37*(5), 932–945.

Bessen, J. E., Bessen, J., & Meurer, M. J. (2008). *Patent failure: how judges, bureaucrats, and lawyers put innovators at risk*. Princeton University Press.

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

Goolsbee, A., & Syverson, C. (2008). How Do Incumbents Respond to the Threat of Entry? evidence from the Major Airlines. *Quarterly Journal of Economics*, *123*(4), 1611–1633.

Veugelers, R. et al. (2008). The role of smes in innovation in the eu: a case for policy intervention? *Review of business and economics*, *53*(3), 239–262.

Alcácer, J., Gittelman, M., & Sampat, B. (2009). Applicant and examiner citations in us patents: an overview and analysis. *Research Policy*, *38*(2), 415–427.

Lalive, R., & Zweimüller, J. (2009). How does parental leave affect fertility and return to work? evidence from two natural experiments. *The Quarterly Journal of Economics*, *124*(3), 1363–1402.

Lerner, J. (2009). The Empirical Impact of Intellectual Property Rights on Innovation: Puzzles and Clues. *American Economic Review*, *99*(2), 343–48.

Sichelman, T. (2009). Commercializing patents. *Stanford Law Review*, *62*, 341–414.

Bertrand, M., Goldin, C., & Katz, L. F. (2010). Dynamics of the gender gap for young professionals in the financial and corporate sectors. *American economic journal: applied economics*, *2*(3), 228–55.

Carrell, S. E., Page, M. E., & West, J. E. (2010). Sex and science: how professor gender perpetuates the gender gap. *The Quarterly Journal of Economics*, *125*(3), 1101–1144.

Hunt, J., & Gauthier-Loiselle, M. (2010). How much does immigration boost innovation? *American Economic Journal: Macroeconomics*, *2*(2), 31–56.

Kerr, W. R., & Lincoln, W. F. (2010). The supply side of innovation: h-1b visa reforms and us ethnic invention. *Journal of Labor Economics*, *28*(3), 473–508.

Kim, A. M., Tingen, C. M., & Woodruff, T. K. (2010). Sex bias in trials and treatment must end. *Nature*, *465*(7299), 688–689.

Kremer, M., & Williams, H. (2010). Incentivizing innovation: adding to the tool kit. *Innovation policy and the economy*, *10*(1), 1–17.

Paul, S. M., Mytelka, D. S., Dunwiddie, C. T., Persinger, C. C., Munos, B. H., Lindborg, S. R., & Schacht, A. L. (2010). How to Improve R&D Productivity: the Pharmaceutical Industry's Grand Challenge. *Nature Reviews Drug Discovery*, *9*(3), 203–214.

Sampat, B. N. (2010). When do applicants search for prior art? *The Journal of Law and Economics*, *53*(2), 399–416.

Van Heerde, H. J., Srinivasan, S., & Dekimpe, M. G. (2010). Estimating Cannibalization Rates for Pioneering Innovations. *Marketing Science*, *29*(6), 1024–1039.

Weston, J., Bengio, S., & Usunier, N. (2010). Large scale image annotation: learning to rank with joint word-image embeddings. *Machine learning*, *81*(1), 21–35.

Haan, P., & Wrohlich, K. (2011). Can child care policy encourage employment and fertility?: evidence from a structural model. *Labour Economics*, *18*(4), 498–512.

Jaffe, A. B., & Lerner, J. (2011). *Innovation and its discontents: how our broken patent system is endangering innovation and progress, and what to do about it*. Princeton University Press.

Jensen, P. H., Thomson, R., & Yong, J. (2011). Estimating the Patent Premium: evidence from the Australian Inventor Survey. *Strategic Management Journal*, *32*(10), 1128–1138.

Lerner, J., Sorensen, M., & Strömberg, P. (2011). Private equity and long-run investment: the case of innovation. *The Journal of Finance*, *66*(2), 445–477.

Morton, F. S., & Kyle, M. (2011). Markets for Pharmaceutical Products. *Handbook of health economics* (pp. 763–823). Elsevier.

Østergaard, C. R., Timmermans, B., & Kristinsson, K. (2011). Does a different view create something new? the effect of employee diversity on innovation. *Research policy*, *40*(3), 500–509.

Ouellette, L. L. (2011). Do patents disclose useful information. *Harv. JL & Tech.*, *25*, 545.

Rossin, M. (2011). The effects of maternity leave on children's birth and infant health outcomes in the united states. *Journal of health Economics*, *30*(2), 221–239.

Whittington, K. B. (2011). Mothers of invention? gender, motherhood, and new dimensions of productivity in the science profession. *Work and Occupations*, *38*(3), 417–456.

Acemoglu, D., & Akcigit, U. (2012). Intellectual Property Rights Policy, Competition and Innovation. *Journal of the European Economic Association*, *10*(1), 1–42.

Belenzon, S. (2012). Cumulative innovation and market value: evidence from patent citations. *The Economic Journal*, *122*(559), 265–285.

De Rassenfosse, G. (2012). How smes exploit their intellectual property assets: evidence from survey data. *Small Business Economics*, *39*(2), 437–452.

Frehill, L. M. (2012). Gender and career outcomes of us engineers. *International Journal of Gender, Science and Technology*, *4*(2), 148–166.

Stephan, P. E. (2012). *How economics shapes science* (Vol. 1). Harvard University Press Cambridge, MA.

Ahuja, G., Lampert, C. M., & Novelli, E. (2013). The second face of appropriability: generative appropriability and its determinants. *Academy of Management Review*, *38*(2), 248–269.

Barbulescu, R., & Bidwell, M. (2013). Do women choose different jobs from men? mechanisms of application segregation in the market for managerial workers. *Organization Science*, *24*(3), 737–756.

Blau, F. D., Brummund, P., & Liu, A. Y.-H. (2013). Trends in occupational segregation by gender 1970–2009: adjusting for the impact of changes in the occupational coding system. *Demography*, *50*(2), 471–492.

Boldrin, M., & Levine, D. K. (2013). The case against patents. *Journal of Economic Perspectives*, *27*(1), 3–22.

Eisenman, M. (2013). Understanding Aesthetic Innovation in the Context of Technological Evolution. *Academy of Management Review*, *38*(3), 332–351.

He, J. J., & Tian, X. (2013). The dark side of analyst coverage: the case of innovation. *Journal of Financial Economics*, *109*(3), 856–878.

Hunt, J., Garant, J.-P., Herman, H., & Munroe, D. J. (2013). Why are women underrepresented amongst patentees? *Research Policy*, *42*(4), 831–843.

Roach, M., & Cohen, W. M. (2013). Lens or prism? patent citations as a measure of knowledge flows from public research. *Management Science*, *59*(2), 504–525.

Vandoros, S., & Kanavos, P. (2013). The Generics Paradox Revisited: Empirical Evidence from Regulated Markets. *Applied Economics*, *45*(22), 3230–3239.

Williams, H. L. (2013). Intellectual property rights and innovation: evidence from the human genome. *Journal of Political Economy*, *121*(1), 1–27.

Zafar, B. (2013). College major choice and the gender gap. *Journal of Human Resources*, *48*(3), 545–595.

Aggarwal, V. A., & Hsu, D. H. (2014). Entrepreneurial exits and innovation. *Management Science*, *60*(4), 867–887.

Cannonier, C. (2014). Does the family and medical leave act (fmla) increase fertility behavior? *Journal of Labor Research*, *35*(2), 105–132.

Ceci, S. J., Ginther, D. K., Kahn, S., & Williams, W. M. (2014). Women in academic science: a changing landscape. *Psychological Science in the Public Interest*, *15*(3), 75–141.

Cook, L. D. (2014). Violence and economic activity: evidence from african american patents, 1870–1940. *Journal of Economic Growth*, *19*(2), 221–257.

Jung, T., & Ejermo, O. (2014). Demographic patterns and trends in patenting: gender, age, and education of inventors. *Technological Forecasting and Social Change*, *86*, 110–124.

Moser, P., Voena, A., & Waldinger, F. (2014). German jewish émigrés and us invention. *American Economic Review*, *104*(10), 3222–55.

Schönberg, U., & Ludsteck, J. (2014). Expansions in maternity leave coverage and mothers' labor market outcomes after childbirth. *Journal of Labor Economics*, *32*(3), 469–505.

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

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

de Rassenfosse, G., Palangkaraya, A., & Webster, E. (2016). Why Do Patents Facilitate Trade in Technology? testing the Disclosure and Appropriation Effects. *Research Policy*, *45*(7), 1326–1336.

England, P., Bearak, J., Budig, M. J., & Hodges, M. J. (2016). Do highly paid, highly skilled women experience the largest motherhood penalty? *American Sociological Review*, *81*(6), 1161–1189.

Hunt, J. (2016). Why do women leave science and engineering? *ILR Review*, *69*(1), 199–226.

Kerr, S. P. (2016). Parental leave legislation and women's work: a story of unequal opportunities. *Journal of Policy Analysis and Management*, *35*(1), 117–144.

Mazzucato, M. (2016). From market fixing to market-creating: a new framework for innovation policy. *Industry and Innovation*, *23*(2), 140–156.

Toivanen, O., & Väänänen, L. (2016). Education and invention. *Review of Economics and Statistics*, *98*(2), 382–396.

Torrisi, S., Gambardella, A., Giuri, P., Harhoff, D., Hoisl, K., & Mariani, M. (2016). Used, blocking and sleeping patents: empirical evidence from a large-scale inventor survey. *Research policy*, *45*(7), 1374–1385.

Goldin, C., & Mitchell, J. (2017). The new life cycle of women's employment: disappearing humps, sagging middles, expanding tops. *Journal of Economic Perspectives*, *31*(1), 161–82.

Hoisl, K., & Mariani, M. (2017). It's a man's job: income and the gender gap in industrial research. *Management Science*, *63*(3), 766–790.

Kogan, L., Papanikolaou, D., Seru, A., & Stoffman, N. (2017). Technological innovation, resource allocation, and growth. *Quarterly Journal of Economics*, *132*(2), 665–712.

Bostwick, V. K., & Weinberg, B. A. (2018). *Nevertheless she persisted? gender peer effects in doctoral stem programs* (tech. rep.). National Bureau of Economic Research.

de Rassenfosse, G. (2018). *Notice Failure Revisited: Evidence on the Use of Virtual Patent Marking* (tech. rep.). National Bureau of Economic Research.

Duffy, J. F. (2018). Reasoned decisionmaking vs. rational ignorance at the patent office. *Iowa L. Rev.*, *104*, 2351.

Furman, J. L., Nagler, M., & Watzinger, M. (2018). *Disclosure and subsequent innovation: evidence from the patent depository library program* (tech. rep.). National Bureau of Economic Research.

Gaule, P., & Piacentini, M. (2018). An advisor like me? advisor gender and post-graduate careers in science. *Research Policy*, *47*(4), 805–813.

Hegde, D., Herkenhoff, K., & Zhu, C. (2018). Patent disclosure and innovation. *Available at SSRN 3158031*.

Jensen, K., Kovács, B., & Sorenson, O. (2018). Gender differences in obtaining and maintaining patent rights. *Nature biotechnology*, *36*(4), 307–309.

Li, X., MacGarvie, M., & Moser, P. (2018). Dead Poets' Property—How Does Copyright Influence Price? *RAND Journal of Economics*, *49*(1), 181–205.

Bell, A., Chetty, R., Jaravel, X., Petkova, N., & Van Reenen, J. (2019). Who becomes an inventor in america? the importance of exposure to innovation. *The Quarterly Journal of Economics*, *134*(2), 647–713.

Bellego, C., & Pape, L.-D. (2019). Dealing with logs and zeros in regression models. *Série des Documents de Travail*, (2019-13).

Bloom, N., Van Reenen, J., & Williams, H. (2019). A toolkit of policies to promote innovation. *Journal of Economic Perspectives*, *33*(3), 163–84.

Gross, D. P. (2019). *The consequences of invention secrecy: evidence from the uspto patent secrecy program in world war ii* (tech. rep.). National Bureau of Economic Research.

Hsieh, C.-T., Hurst, E., Jones, C. I., & Klenow, P. J. (2019). The allocation of talent and us economic growth. *Econometrica*, *87*(5), 1439–1474.

Mairesse, J., Pezzoni, M., & Visentin, F. (2019). Impact of family characteristics on the gender publication gap: evidence for physicists in france. *Interdisciplinary Science Reviews*, *44*(2), 204–220.

Marco, A. C., Sarnoff, J. D., & Charles, A. (2019). Patent Claims and Patent Scope. *Research Policy*, *48*(9), 103790.

Reimers, I. (2019). Copyright and Generic Entry in Book Publishing. *American Economic Journal: Microeconomics*, *11*(3), 257–84.

Sarada, S., Andrews, M. J., & Ziebarth, N. L. (2019). Changes in the demographics of american inventors, 1870–1940. *Explorations in Economic History*, *74*, 101275.

Baruffaldi, S. H., & Simeth, M. (2020). Patents and knowledge diffusion: the effect of early disclosure. *Research Policy*, *49*(4), 103927.

Bryan, K. A., Ozcan, Y., & Sampat, B. (2020). In-text patent citations: a user's guide. *Research Policy*, *49*(4), 103946.

de Rassenfosse, G., & Higham, K. (2020). Wanted: a standard for virtual patent marking. *Available at SSRN 3562133*.

de Rassenfosse, G., Pellegrino, G., & Raiteri, E. (2020). Do patents enable disclosure? evidence from the invention secrecy act. *Evidence from the Invention Secrecy Act (March 26, 2020)*.

Ganguli, I., Hausmann, R., & Viarengo, M. (2020). Gender differences in professional career dynamics: new evidence from a global law firm. *Economica*.

Koning, R., Samila, S., & Ferguson, J.-P. (2020). Inventor gender and the direction of invention. *AEA Papers and Proceedings*, *110*, 250–54.

Kuhn, J., Younge, K., & Marco, A. (2020). Patent citations reexamined. *The RAND Journal of Economics*, *51*(1), 109–132.

Lück, S., Balsmeier, B., Seliger, F., & Fleming, L. (2020). Early disclosure of invention and reduced duplication: an empirical test. *Management Science*, *66*(6), 2677–2685.

Moser, P., & Kim, S. (2020). Women in science: lessons from baby boom. *Working paper*.

Myers, K. R., Tham, W. Y., Yin, Y., Cohodes, N., Thursby, J. G., Thursby, M. C., Schiffer, P., Walsh, J. T., Lakhani, K. R., & Wang, D. (2020). Unequal effects of the covid-19 pandemic on scientists. *Nature Human Behaviour*, 1–4.

Reimers, I. C., & Waldfogel, J. (2020). *Digitization and Pre-Purchase Information: the Causal and Welfare Impacts of Reviews and Crowd Ratings* (tech. rep.). National Bureau of Economic Research.

Thompson, N. C., & Kuhn, J. M. (2020). Does winning a patent race lead to more follow-on innovation? *Journal of Legal Analysis*, *12*, 183–220.

Higham, K., De Rassenfosse, G., & Jaffe, A. B. (2021). Patent quality: towards a systematic framework for analysis and measurement. *Research Policy*, *50*(4), 104215.

Kaltenberg, M., Jaffe, A. B., & Lachman, M. (2021). *The age of invention: matching inventor ages to patents based on web-scraped sources* (tech. rep.). National Bureau of Economic Research.

Kaltenberg, M., Jaffe, A. B., & Lachman, M. E. (2021). *Invention and the life course: age differences in patenting* (tech. rep.). National Bureau of Economic Research.

LING ZHOU

Nationality: China ⋄ Date of Birth: August 19th 1993
ODY 4 15 (Odyssea), Station 5, EPFL, CH-1015 Lausanne, Switzerland
(+41) 078-732-0118 ⋄ ling.zhou@epfl.ch

EDUCATION

Ph.D., Management of Technology, Expected completion: August 2021.
Swiss Federal Institute of Technology (EPFL), Switzerland

Consortium on Competitiveness and Cooperation Virtual visiting Ph.D. student, 2020 Fall
NYU Stern School of Business (Host: Prof. Deepak Hegde), USA

Swiss Program for Beginning Doctoral Students in Economics, 2017 - 2018
Swiss National Bank, Switzerland

M.A., Economics, 2016 - 2017.
Simon Fraser University, Canada

B.A., Economics, 2012 - 2016.
Nanjing University, China

Visiting student, 2014 - 2015
Utrecht University, the Netherlands.

RESEARCH FIELDS

Economics of Innovation, Intellectual Property Rights, Applied Microeconomics

REFERENCES

Prof. Gaétan de Rassenfosse (supervisor)
Chair of Innovation and IP Policy
EPFL
gaetan.derassenfosse@epfl.ch

Prof. Dominique Foray
Chair of Economics and Management of Innovation
EPFL
dominique.foray@epfl.ch

Prof. Stefano Baruffaldi
Centre for Research in Entrepreneurship and Innovation
University of Bath
shb40@bath.ac.uk

Prof. Christopher Tucci
Imperial College Business School
c.tucci@imperial.ac.uk

l:

WORKING PAPERS

Patents and Supra-Competitive Prices: Evidence from Consumer Products, *with Gaétan de Rassenfosse* (submitted).

(Original title: A Test of the Monopoly Pricing Hypothesis of Patents)

Public Notice and Invention Diffusion, *with Gaétan de Rassenfosse*

Does Motherhood Hold Back 'Marie Curies'? , *with Mary Kaltenberg*

PUBLICATIONS

COVID-19: Insights from Innovation Economists, *with George Abi Younes et al.*
Science and Public Policy (2020)

CONFERENCES AND WORKSHOPS

CCC Visiting Student Seminar at NYU Stern, Online	April 2021
Research Policy Online Conference for Early Career Researchers, Online	April 2021
Gerzensee Alumni Conference, Swiss National Bank	October 2020
KUL Summer School: Data & Algorithms for STI Studies, Leuven	September 2020
2nd Research on Innovation, Science and Entrepreneurship Workshop, Munich	December 2019
KUL&EPO Summer School: Data & Algorithms for STI Studies, Vienna	September 2019
European Policy for Intellectual Property, Zurich	September 2019
R&D Management, Paris	June 2019
Competition and Innovation Summer School, Montenegro	May 2019

ACADEMIC SERVICES

Activities

Co-organizer of "EPFL Virtual Innovation Seminar" May. 2020 - Present
College of Management of Technology, EPFL

Co-organizer of "Innovation Reading Group" Sep. 2018 - May. 2020
College of Management of Technology, EPFL

Ad hoc Review

Organization Science, Journal of Economics & Management Strategy, Academy of Management Conference 2021

TEACHING EXPERIENCES

Teaching Assistant Mar. 2018 - Present
College of Management of Technology, EPFL

- Economics of Innovation and Intellectual Property (master course)
- Principles of Microeconomics (master course)
- Introduction to Econometrics (master course)

Teaching Assistant Sep. 2016 - Aug. 2017
Department of Economics, Simon Fraser University

- Principles of Microeconomics(undergraduate course)
- Labor Economics (undergraduate course)
- International Trade (undergraduate course)

AWARDS & SCHOLARSHIP

4iP Council Research Award 1st, 2021; Doctoral Assistantship, 2017-2021; Herbert G. Grubel Award, 2017; The Second Grade Award in the 19th Forum of Science Arts of Nanjing University, 2016; Renmin Scholarship, 2014.

SKILLS & INTERESTS

Languages	English (fluent), Chinese (native), French (beginner)
Programming & Software	Stata, Python, R, MySQL, MS Office, LaTeX
Hobbies	Hiking, Running, Badminton

Ce document a été imprimé au Centre d’impression EPFL,
imprimerie climatiquement neutre, certifiée myClimate.