Title
Essay on Innovation and Finance

Author
Jurek, Dominik

Publication Date
2023

Essay on Innovation and Finance

by

Dominik Jurek

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Business Administration

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Gustavo Manso, Co-chair
Associate Professor Timothy McQuade, Co-chair
Professor Ross Levine
Professor Lee Fleming

Spring 2023

Essay on Innovation and Finance

Abstract

Essay on Innovation and Finance

by

Dominik Jurek

Doctor of Philosophy in Business Administration

University of California, Berkeley

Professor Gustavo Manso, Co-chair

Associate Professor Timothy McQuade, Co-chair

Do patents facilitate market entry and job creation? Using a 2014 Supreme Court decision that limited patent eligibility and natural language processing methods to identify invalid patents, I find that large treated firms reduce job creation and create fewer new establishments in response, with no effect on new firm entry. Moreover, companies shift toward innovation aimed at improving existing products consistent with the view that patents incentivize creative destruction.

To Ruochen

Contents

List of Figures

List of Tables

Acknowledgments

Disclaimer: Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau's Disclosure Review Board and Disclosure Avoidance Officers have reviewed this information product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2669. (CBDRB-FY22-P2669-R10037)

Chapter 1

Introduction

Do patents facilitate market entry of new ideas by protecting rents on novel technologies? Or do they stifle innovation by allowing incumbents to maintain excessive legal monopolies? Influential research suggests that patents facilitate the market entry of new ideas by granting innovators property rights over their inventions and incentivizing costly, risky investments in innovation (e.g., Schumpeter 1942; Arrow 1962; Nordhaus 1969). In this way, strong patent protection can spur start-ups and incumbents to introduce new technologies to the market, accelerating economic growth. However, research also suggests that poorly designed patent systems that grant excessively broad property rights can stifle innovation and new entrants (e.g., Boldrin and Levine 2013). For example, incumbent firms can exploit such systems by acquiring patents that block new firms from entering the market, allowing incumbents to extract rents from their existing products rather than developing new technologies to compete with other firms. From this perspective, patents can decrease entry, increase markups for incumbents, and reduce business dynamism and growth (Akcigit and Ates 2019).

In this paper, I contribute to research on how patents affect innovation by evaluating the effect of a recent court decision that limited the patent eligibility of software patents. The 2014 *Alice Corp. v. CLS Bank International* Supreme Court decision[1] (*Alice* decision) weakened the enforceability of existing software patents and limited the patentability of new software-related innovations. I assess the impact of the *Alice* decision on job creation, firm entry and exit, and innovation efforts.

Identifying patents that are affected by *Alice* is challenging because few patents are litigated and software patents fit into multiple broad patent classes. Instead, I introduce a novel approach that directly uses the patent claim language. I use texts of patent applications that were rejected due to the *Alice* decision to train a binary natural language processing (NLP) classification algorithm and identify claim language that is not eligible after *Alice* anymore. The advantage of this method

[1] 573 U.S. 208 (2014)

is that I can predict for a large set of patents whether they are invalid and identify treated firms by measuring the share of invalid patents in patent portfolios before the *Alice* shock.

To measure innovation and growth-relevant outcomes, I leverage restricted-use U.S. Census microdata on establishment entry and job creation, surveys on research and development (R&D) expenditures, and patent data on innovation quality and direction. Using industry- and firm-level data and my NLP method for identifying treated firms, I estimate a difference-in-differences (DiD) model and address three sets of questions that measure the effect of patentability on innovation and growth: First, how do entry and exit change for new ventures in treated industries? Second, how do establishment entry, job creation and destruction, and innovation quality and direction change for incumbents if they receive fewer patents for their inventions? Third, which types of firms are most affected by the *Alice* shock to patentability?

I estimate the impact of *Alice* on firm entry and exit by aggregating Census and patent portfolio data to the industry level. I find a significant decrease of 3%–4% for establishment entry by large incumbents, but no significant change in entry or exit for small firms and new ventures. We can interpret this result as new entrants being more interested in getting their products on the market even without the ability to receive patents (see Boldrin and Levine 2013), while large incumbents reduce the exploration and market entry of new technologies.

R&D surveys conducted by the Census Bureau and the National Science Foundation give further insight into how innovation strategies have changed after *Alice*. I find no significant effect for overall R&D spending, but the share of R&D spending on development increases by 0.27% per one percent of treatment intensity, and relative spending on research decreases by 0.32%. Thus, innovation does not stop due to *Alice* but rather becomes more focused on the improvement of existing products. This is consistent with the aggregate industry results showing less establishment creation; improving existing technologies rather than entering into new ones reduces establishment entry (see Garcia-Macia et al. 2019; Klenow and Li 2021).

Overall, industry results show that incumbents engage more in own innovation by improving existing products rather than exploring new technologies after *Alice*. Based on this shift in innovation direction, growth theory as in Peters (2020) predicts a decrease in establishment entry and job creation and an increase in markups at the firm level. Consistent with this, I find negative treatment effects of 5.5% for job creation and 3.4% for new establishment entry, but a positive treatment effect of 2.4% for markups. The shift toward own innovation is also reflected in patenting data: while *Alice* has a negative effect on overall issuances, patents have become more narrow, less experimental, and focused more on internal rather than external innovation. *Alice* has changed which innovations can be patented, and firms have shifted their innovation strategy toward improving technologies they already have in place.

Depending on which incumbents shift their innovation strategy toward own innovation, the implications for innovation and growth might be different. If large market leaders reduce creative destruction, the long-term effects on competition and growth are smaller than if followers limit innovation directed at acquiring new technologies to replace leading firms. I find that the largest firms in the 90th percentile of the patent portfolio and firm employment size distribution drive the results: new establishment creation decreases by 9.8% while overall patent issuances decrease by 10.9%. Markups increase by 2.1%, and patent references to internal knowledge increase relative to overall patent citations. The *Alice* shock reduces the patentability of innovations, and large incumbents substitute new technology exploration with product improvements. Thus, the results are consistent with patents incentivizing incumbents to innovate by entering into new technologies, which in turn promotes creative destruction and growth.

My results are robust to variations in the regression specification and sample selection. I implement several alternative definitions for treated and control firms, estimate economic outcomes with different variables, and implement various model specifications. Furthermore, my approach fulfills all assumptions for the causal identification of treatment effects. For the difference-in-differences setting to be valid, the key assumption is that changes in outcome variables are different for the treated group only due to *Alice*. The *Alice* decision is the most significant change in software patentability since 1998, and no patent regime change affected the control group in the sample period. Thus, only the treated group experiences a consistent shock to patentability. Another concern might be unobserved industry shock affecting the treated firms. I find that *Alice*-treated patents are concentrated among large firms in computer-related service industries, but my main results are robust to the inclusion of sector-year fixed effects and controlling for size-related covariates, limiting concerns that the treatment effects are driven by other channels than *Alice* (Goldsmith-Pinkham et al. 2020). Econometrically, the treatment variable is uncorrelated with covariates that are related to changes in outcomes, and graphical analysis of event study outcomes does not reject the parallel trend assumption. Thus, there is no economic or statistical reason why the exclusion restriction should be violated.

This paper contributes to the literature in three ways. First, I show consistent microdata-based evidence for Schumpeterian growth theory with heterogeneous innovations (Akcigit and Kerr 2018; Garcia-Macia et al. 2019; Klenow and Li 2021). Large firms reduce the exploration of new technologies in favor of improving existing products after the *Alice* shock. This fits into the framework introduced by Peters (2020). The core idea is that firms improve their current products to gain market power and increase markups, but also engage in creative destruction to acquire new technologies. My results are consistent with patents reducing expansion costs for incumbents. That is, without the ability to receive patents for their innovations, incumbent firms face higher costs of breaking into new product markets and thus increase own innovation of their current products relative to creative destruction. I show that this is most important for large firms. As a consequence, a simple model calibrated with my estimated treatment effects shows that the *Alice* shock decreases welfare by ca. 1.3% for the most treated sectors. As in Peters (2020), the effect is small because firms substitute innovation types rather than limit innovation efforts, leading to a relatively small estimate for the elasticity of growth to patentability of 0.2.

Second, I introduce a novel identification approach combining machine learning, natural language processing, and patent data. Recent literature has used firm and patent texts to measure the similarity of companies and technologies (Hoberg and Phillips 2016; Frésard et al. 2020; Kelly et al. 2021) and group firms and patents based on text analysis concepts such as tone, sentiment, or readability (Loughran and Mcdonald 2011; Ke et al. 2019; Kong et al. 2020). My approach is the first application of NLP-based patent classification to identify causal treatment effects in response to a patent regime change. Moreover, my method provides new insights into the importance of software and business method patents for economic growth (see, e.g., Bessen and Hunt 2007; Webb et al. 2018; Lerner et al. 2021). The role of software patents is best understood in the context of why some low-citation patents have a high value for patent holders. The number of citations a patent receives is generally considered to be a proxy for the value of the innovation (Harhoff et al. 1999; Hall et al. 2005). However, Abrams et al. (2013) find that some high-value patents receive few citations and model low-citation patents as additional protection against infringement for other valuable inventions. Consistent with this idea, I give direct empirical evidence that the additional protection from low-citation patents is important for new establishment creation. Thus, software patents, which are generally considered to be low-value patents (Hall and MacGarvie 2010), are economically important for their value in protecting innovations.

Finally, I provide empirical support for a positive causal link between patenting and entry with my core insight that large firms protect the market entry of new technologies with patents. This is an important new perspective on the longstanding question of how patentability incentivizes innovation: early theoretical literature provides no clear answer to whether a broad or narrow patenting regime is optimal (see, e.g., Gilbert and Shapiro 1990; Klemperer 1990; Gallini 1992). Empirical evidence on the question is scarce with some studies suggesting that broad patent rights increase innovation costs and limit entry (Cockburn and MacGarvie 2011; Hall et al. 2021), while other studies find little effect of patentability on innovation and follow-on inventions (Moser 2011; Sampat and Williams 2019). On the other hand, studies such as Gans and Stern (2003), Aghion et al. (2014), and Aghion et al. (2015) stress the importance of strong patent protection to incentivize innovation investment.[2] My empirical results support a more differentiated view of patents: patentability incentivizes innovation that is linked to new establishment entry and job creation for large firms but has no effect on innovation by small firms. A more narrow patent regime leads large firms to fo-

[2]My research is closely related to Cockburn and MacGarvie (2011) who ask whether patents on relevant technologies keep firms out of software product markets. In a sample from 1990 to 2004, they find that a 10% increase in the number of patents relevant to a market reduces the rate of entry by 3%–8%, which intensifies following expansions to the patentability of software in the mid-1990s. I directly build and improve upon this result in two ways: First, I capture better how large firms can change their innovation strategy to own innovation. Second, unlike previous studies that look at expansions of what can be patented (Cockburn and MacGarvie 2009; Noel and Schankerman 2013), my study focuses on a regime change in the opposite direction. This distinction is economically important because the response to patenting regime changes might not be symmetric.

cus on a different type of innovation (own innovation) rather than reduce innovation efforts. Thus, patents are positively related to entry and job creation through the exploration of new technologies.

Chapter 2

Methodoloy and Data

I exploit a recent Supreme Court decision that limited the patent eligibility of software patents as a natural experiment to identify the causal link between patentability and innovation-relevant outcomes. I develop a novel NLP-based classification method on patent texts to identify technologies that became patent ineligible following the Supreme Court decision. I define treated industries and firms based on their share of affected patents and estimate treatment effects with a difference-in-differences (DiD) approach. To measure growth and innovation-relevant outcomes, I use restricted-use U.S. Census data on establishment counts and employment at the firm level, R&D surveys, and patent quality measures from the innovation literature.

2.1. Identification

Patent-related measures can be used as endogenous proxies for innovation outcomes. To identify the causal impact of patentability on innovation, I need to find an exogenous shock that changed the ability to receive patents for innovations.

Software-related patents are a good starting point to identify the causal effect of patentability: the particular difficulty of describing software concepts in patent claims leads to tradeoffs when defining the right scope of patent-eligible subject matters, that is which inventions can receive patent protection. On the one hand, the incentive to innovate needs to be preserved while avoiding on the other hand ‘patent thickets’ that stifle competition and limit innovation by other inventors (Stroud and Kim 2017). As a consequence, Supreme Court decisions led to several regime shifts over the decades on what is and is not patent-eligible (figure 2.1 summarized the timeline of some of the most important decisions related to software patents). The *State Street* decision in 1998 allowed for a relatively broad interpretation of what is patent eligible. In the following years, overly broad software patents were linked to the rise of patent trolls and nonpracticing entities; a type of patent holder that generates income by claiming license fees and litigating against genuine innovators (see, e.g. Appel et al. 2019, Cohen et al. 2019, Lee et al. 2019, Lemley and Feldman 2016).

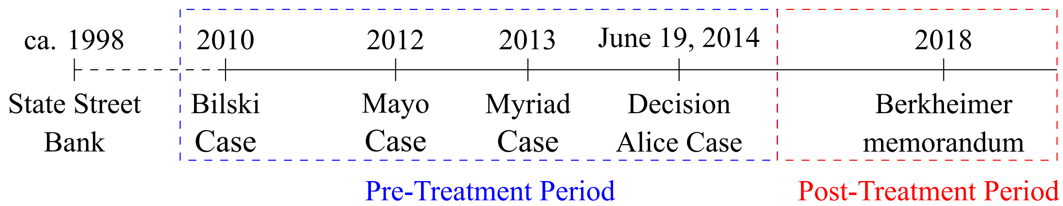


Fig. 2.1. Supreme Court decisions timeline on software and business method patents

On June 19, 2014, the U.S. Supreme Court affirmed the judgment of a lower court that the method claims directed at a scheme/method for mitigating settlement risk were ineligible because the claims just implemented an abstract idea (Ren and Duprez 2019). This decision, as many commentators were quick to point out, increased uncertainty around the patent eligibility of software and business method claims and limit their scope, reversing the broad patentability regime for software innovations (see, e.g. DiNizo 2018, Tran 2015, Tran 2016, Daily 2017, Stroud and Kim 2017, Chien 2016, Craig 2017). Recent empirical research has confirmed the impact of *Alice* on patenting and litigation: Chien and Wu (2018), Kesan and Wang (2020), and Toole and Pairolero (2020) all find that rejection rates and application abandonments among software and business method application have spiked after *Alice*, with some rejection rate for business methods growing from 25% to 81% in the month after *Alice*.

Overall, the *Alice* decision is one of the most important recent regime shifts in the patentability of certain types of innovations. For individual firms, this means an exogenous shock to the ability to file for patent protection of new inventions and enforce existing software patents.

To use this regime shift to identify the causal effect of patentability, I need to classify patents based on their exposure to the *Alice* decision. This is challenging since not a particular type of patent or patent class was invalidated, rather a certain type of language describing innovations was ruled to be patent-ineligible. Prior research uses United States Patent Classification (USPC) patent class 705 to identify the most treated patents (see, e.g., Wagner and Cockburn 2010, Contigiani 2020). I can confirm that this class is indeed the most affected USPC class by *Alice*, but USPC classes are not useful going forward since they were discontinued by the United States Patent and Trademark Office (USPTO) in 2013. Using USPC classes thus does not allow us to validate the empirical approach and identify treated patents after the court decision. The Cooperative Patent Classification (CPC) groups, which replaced the USPC system, have several drawbacks; there is no direct concordance between USPC and CPC, and CPC groups are much broader than USPC classes, thus covering more technologies than which became patent-ineligible after *Alice*. Another approach to identify treated patents would be to directly observe litigated patents after *Alice* (Galasso and

Schankerman 2015). While this certainly allows us to define which patents are not eligible under *Alice*, few patents are litigated.

To allow for a consistent measure of treated patents across the entire sample, I develop and train an NLP model to identify claim language that is invalid under *Alice* and classify existing patents. I use as a sample of text formulations that are not eligible after the decision claims in applications that were rejected due to *Alice* in the Office Action Research Data from Lu et al. (2017). I select issued patents from the same filing years and patent classes as the rejected applications to obtain control claims with eligible language. I extract the rejected independent claim texts from the application publications and collect from PatentsView as control texts the published independent claims of issued patents. I then train a support-vector machine on a TF-IDF matrix of the rejected and control claim texts, which allows me to predict for a given claim text whether it would be rejected under *Alice* or not. I use this NLP model to classify all independent claims of issued patents since 1990 in PatentsView in CPC groups related to the training data. In total, this results in 3,359,812 classified independent claims in 1,062,897 patents.[1] Appendix C provides a detailed description of the NLP methodology, analysis of the language features of *Alice*-affected claim texts, and validation tests for the quality of the patent classification. Overall, all performance metrics point toward a very good classification model with high precision (91.7%), high recall (94.1%), and more than 85% Matthews correlation coefficient (MCC), meaning a very high correlation between the true and predicted labels. Eventually, I can thus predict for each patent in related patent classes if it is affected by *Alice* or not. I define a patent as treated if my NLP method predicts that its first claim is affected by *Alice*. The first claim is generally the broadest claim of a patent and thus the most relevant text when defining which technologies are covered by the patent (see, e.g., Kuhn and Thompson 2019). The following analysis is robust to alternative definitions of treatment, such as requiring at least one independent claim to be predicted as affected by *Alice*.

To validate my approach, figure 2.2 shows the indices of treated patent issuances defined by my NLP method (*Alice* NLP treated), patents in the CPC groups closely related to *Alice* (*Alice*-related CPC groups), and all utility patent issuances in a given year (Actual total issues). Following 2014, there was a clear downward shock for both, the *Alice*-related CPC groups as well as my *Alice* NLP-treated patent, while actual total issues did not move a lot. My NLP classification method identifies treated patents better than the affected CPC groups alone, which are broad and cover more technologies than are treated by *Alice*: the *Alice*-related CPC groups saw a decline in patent

[1] Note that the USPTO moved away from USPC classes in 2013 and adopted CPC classification instead. I use the older classification to find control claims since applications around *Alice* were still mainly classified under USPC. CPC classifications are also assigned to older patents, thus I can follow the CPC groups closest related to the USPC classes from my training data set over the entire sample period. I restrict the classification to the five CPC groups most consistent with the USPC classes from the training set, i.e., the classes with the most rejections due to *Alice*, which are CPC groups A63F (video games), G07F (coin-freed or like apparatus), G06F (digital data processing), H04L (transmission of digital information), and G06Q (data processing systems), representing around 50% of the affected patents in the most treated USPC classes 705, 463, 434, and 702.

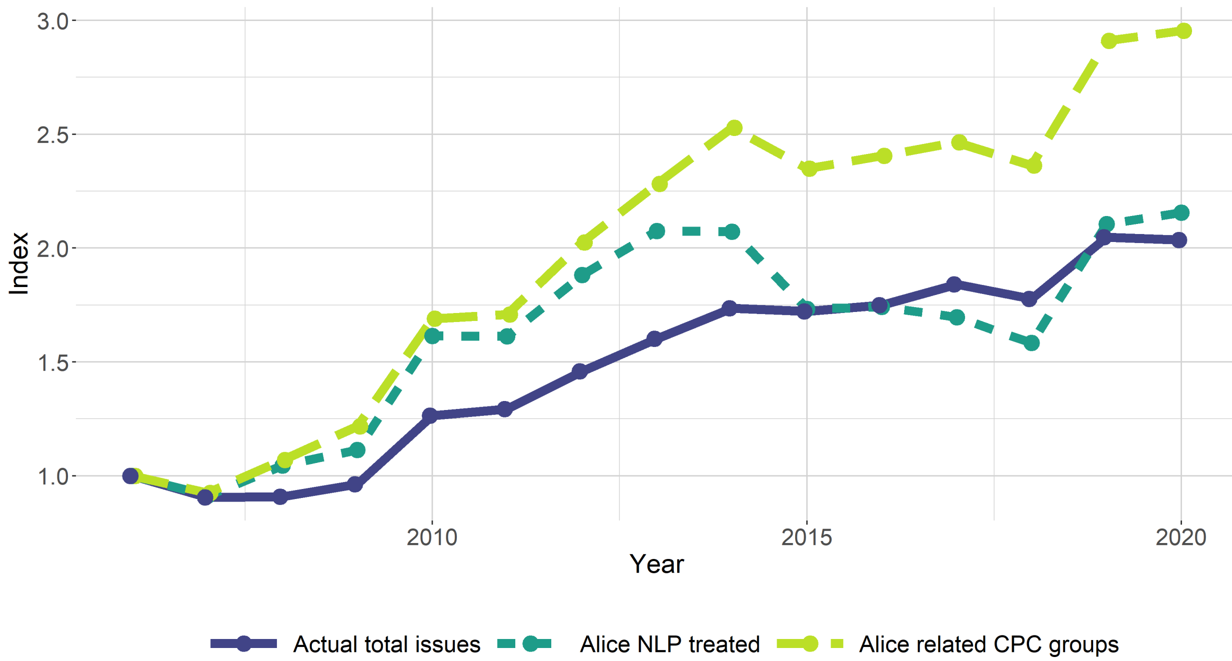


Fig. 2.2. Normalized patenting index 2006 to 2020

Notes: Count of issued utility patents by year, normalized to 2006-levels. The three patent indices are the actual count of all issued utility patents, the count of patents that are classified by the NLP algorithm as treated under *Alice*, and the count of issued patents in the five CPC groups closely related to *Alice*–A63F (video games), G07F (coin-freed or like apparatus), G06F (digital data processing), H04L (transmission of digital information), and G06Q (data processing systems). Patent issuance data are sourced from PatentsView.

issuance from 2014 to 2015 of 7.1%, while *Alice* NLP-treated patent issuances declined by more than 16.3%.

Several other observations are interesting in this graphic: first, software-related patent issuances are rising faster than other types of patents before *Alice*, consistent with the evidence in Lerner et al. (2021) and Webb et al. (2018). A jump in annual issuances in 2010 is followed by a flattening of the trend relative to total patent issuances. This is consistent with several rulings by the Supreme Court building up to the *Alice* decision, starting with the *Bilski* case in 2010 (see also figure 2.1). The flatter trend line after 2010 supports the parallel trend assumption before 2014 for my DiD approach, confirming that my pre-period should be defined as the years between 2010 to 2014. We also see an uptick in software patents after 2019, consistent with Toole and Pairolero (2020) describing that the recent subject matter eligibility decision in *Berkheim v. HP, Inc* in 2018 and the

2019 revised subject matter eligibility guidance by the USPTO reversed the trend in subject matter eligibility rejections and decreased uncertainty. Thus, the best treatment period is the years 2015 to 2019.

2.2. Data

To answer my research question, I combine data on patents, R&D surveys, and restricted-use Census data to identify treated firms and industries, and measure relevant growth and innovation outcomes. All variable definitions are summarized in the appendix table A.

My main economic data set is the redesigned Longitudinal Business Database (LBD) (Chow et al. 2021 and Jarmin and Miranda 2002). LBD is a restricted-use data set provided by the U.S. Census Bureau covering private, non-farm business establishments with employees that operated in the United States since 1976. One of the main goals of the LBD is to measure the entrance and exit of businesses across industries and yearly employment changes. The data are based on the Business Register and use among other source tax data. LBD thus gives a comprehensive picture of even small establishments in the US. I focus on the years 2010 to 2019 in LBD; this is consistent with the pre-period of 2010 to 2014 where the parallel trend assumptions for treated patent issuances in figure 2.2 seems to hold well, and the post-period ending in 2019.[2]

To estimate the effect of *Alice* on new firm creation and destruction, I aggregate annual establishment entry and exit from the LBD to the industry level. I use the vintage-consistent four-digit NAICS codes from Fort and Klimek (2018) to define industries. LBD only tracks non-farm businesses, following Kerr and Nanda (2009), I additionally exclude all establishments in agriculture, forestry, fishing (NAICS 11), public administration (NAICS 92), private households (NAICS 814), U.S. Postal Service (NAICS 491), restaurants and food stores (NAICS 722 and 445), hospitals (NAICS 622), education services (NAICS 61), and social services (NAICS 624 and 813). Many finance papers also exclude observations in utilities (NAICS 22) and finance and real estate (NAICS 52 and 53), due to these industries being highly regulated (see, e.g., Gutiérrez and Philippon 2016); for this analysis, there is no immediate reason why these sectors should be excluded when it comes to intellectual property rights and innovation (the results remain unchanged even if utilities and finance and real estate industries are excluded).

For the firm-level analysis, I aggregate establishment data in LBD to the firm-year level and add revenue data from Haltiwanger et al. (2019). The revenue data are based on tax receipts in

[2] My analysis uses the 2019 LBD vintage, which includes two years of observations after the last Economic Census in 2017. Substantial efforts are spent to correctly time the birth and death of establishments between Economic Census years, this is not possible though for the last two years of my sample. However, re-timing is most relevant for smaller incumbents. Large firms are likely to be consecutively surveyed by, e.g., the Corporate Organization Survey (COS), and newly created firms are immediately recognized in the underlying tax data. In any case, controlling for re-timing and re-activation of establishments after some years without employment does not change my results.

the Standard Statistical Establishment List and the Business Register and are available for ca. 80% of all firms up to 2018. I assign firms to the four-digit NAICS industry accounting for the largest employment share.[3]

I use the novel patent assignee-firm link from Dreisigmeyer et al. (2018) to match patent data to firms. This bridge builds on prior work by Graham et al. (2018) and uses a triangulation method that combines string matching of company names to assignee names and inventor names to employee data in restricted use Census data creating a high-quality match of patents to firms. I can use this match to define the treatment variable as the share of *Alice*-treated patents relative to all issued patents to the firm between 2000 and 2009. This ensures that the pre-treatment period with several court cases and decisions leading up to *Alice* does not lead to sample selection bias through firms selecting into treatment by filing *Alice*-affected patents they expect to be invalidated in the future. Furthermore, the parallel trend assumption can be tested for the years between 2010 and 2014. The only caveat of this definition of treatment is the limitation to incumbents patenting before 2010.

I can use the patenting statistics on the firm level to identify treatment at the industry level. First, I select industries that had at least 20 unique firms with patent assignments between 2000 and 2009. While the precise number of patenting firms in an industry does not change the results, having too few patents and firms with patents in an industry poses issues when identifying treatment. If there

[3]It should be noted that Compustat as a data source for publicly listed firms is a poor substitute for the following analysis: using data from Kogan et al. (2017) to identify patenting firms in Compustat, there are only 1,586 firms with positive patent assignments between 2000 and 2009 in the cross-section in 2010. Of those firms, ca. 1,100 are in manufacturing sectors, thus outside the industries that are mostly treated by *Alice*. In the cross-section of 2010, 48% of firms with patents assigned to them over the prior 10 years in Compustat are in (2-digit NAICS) manufacturing sector 33 alone, which also accounts for 71% of total patent assignments over the prior 10 years to Compustat firms. With another 22% of patenting firms and 12% of total patent assignments from sector 32, manufacturing sectors overall account for more than 4 out of 5 patents assigned to Compustat firms between 2000 and 2009. Furthermore, the largest sample of firm-year observations between 2010 and 2019 for patenting Compustat firms is 12,100 observations, far smaller than the 197,000 observations using LBD. For the largest size bracket with more than 1000 employees and more than 20 patents assigned between 2000 and 2009, only 4,280 observations can be found in Compustat, less than half the sample size in LBD. Also here, the sample is heavily tilted toward manufacturing firms with 3,550 observations in manufacturing industries. Thus, with the far smaller sample size and the heavy focus on manufacturing, Compustat data are not equivalent to LBD, especially since innovation- and growth-relevant measures such as establishment entry are not directly observable. Note also that the bias toward manufacturing firms is even more pronounced when just looking at treated patents. In the 2010 cross-section in Compustat of companies that had at least one treated patent assigned to them over the prior 10 years, we find more than twice as many firms from the manufacturing sector 33 than from the information sector 51 (207 vs. 93) and few from sector 54 (23). We could try to weight observations based on how important patents are for the company by, e.g., using the ratio of patent portfolio size relative to sales. However, this would increase the bias toward manufacturing firms since the average manufacturing firm with treated patents in sector 33 has a patent portfolio-to-sales ratio of 0.35, while firms in sectors 51 and 54 have ratios of 0.09 and 0.19, respectively, and the median of the patent portfolio-to-sales ratio for treated firms in sector 33 is 0.174 and for firms in sectors 51 and 54 0.035 and 0.055, respectively. Overall, Compustat is a poor alternative to LBD since the selection of treated and patenting firms is heavily tilted toward manufacturing firms, even with reweighting of treatment this cannot be corrected for.

are only very few firms with patents in an industry, the definition of treatment might be driven by outliers and measurement errors.[4] The treatment variable is the share of treated patents relative to all patents issued to firms in the four-digit industry between 2000 and 2009. While this treatment definition is based on the patent portfolios of incumbents, it identifies industries that patented in *Alice*-affected technologies and for which patents are important.

I use survey responses in the Business R&D and Innovation Survey (BRDIS) and firm-level patent data from PatentsView to analyze how innovation efforts and direction change after *Alice*. BRDIS is designed to be the 'primary source of information on research and development performed or funded by businesses within the United States'[5]. The results thus are supposed to reflect broader trends and are most useful for industry analysis.

PatentsView data are already used to identify patent issuances and define treatment. I follow prior literature to define measures for novelty, scope, and quality of innovations on the firm level. For these measures, I focus on patent grant years rather than application years since *Alice* is primarily an issuance shock.[6] Following Balsmeier et al. (2017), I measure the number of backward citations, i.e., the number of citations to prior patents, and self-citations, citations to prior patents of the same firm. Backward citations proxy for the prior art and overall knowledge that inventions refer to, and self-citations indicate references to internal knowledge. Patents filed in new CPC groups to the firm and the technology proximity to prior patents filed by the firm measure the exploration of new technology areas. The total number of claims is related to the scope of patented inventions and the number of citations patents receive (forward citations) is a measure of innovation quality.[7]

I differentiate the issued patents further into four groups according to the distribution of forward citations of patents from the same filing year and CPC group: I count the number of patents that are in the top 1% of the forward citations distribution, that are in the 2-10% percentile, that are

[4]For example, if we can find only one firm with one single treated patent assigned to them, the industry-level treatment would be 100%. Despite the high quality of the patent-firm match in Dreisigmeyer et al. (2018), the precision of more than 90% is high but not perfect, thus measurement error might also distort results if there are too few distinct firms used to identify treated industries.

[5]See `https://www.nsf.gov/statistics/srvyberd/#sd`

[6]The results are consistent if I use application years instead. There is a noticeable pre-trend, though, since applications filed in 2012 and later, which would have been granted in 2014 and after, gradually start to be affected by *Alice*.

[7]PatentsView creates time-consistent unique identifiers for assignees of granted patents based on a string-matching protocol. I use these assignee data to identify self-citations, technology proximity, and new and known CPC groups for firms rather than the patent-firm link by Dreisigmeyer, Goldschlag, Krylova, Ouyang, and Perlman (2018) since the coverage of PatentsView is much longer (1976 to 2021 compared to 2000 to 2019).

The overall coverage is very high between the firms identified by PatentsView and the patent-firm link in Census data: 84% of assignees in PatentsView have a unique firm in the patent-firm link that can be assigned to them, and 81% of firms in patent-firm link have a unique assignee-ID in PatentsView. 94% of the patents covered by the patent-firm link have a unique assignee in PatentsView data. In total, at least 3 million patents are covered by both the patent-firm link and PatentsView.

below the top decile but have received at least one forward citation, and patents that have received no citations. According to Balsmeier, Fleming, and Manso (2017), this allows us to measure breakthrough, important, incremental, and failed inventions. I further follow Lerner and Seru (2021) and normalize forward citations by the average number of citations received by patents filed in the same year and CPC group, giving a more continuous measure of quality.[8]

I build on Galasso and Schankerman (2015) and measure follow-on citation as the total forward citations received by patents within three years after issuance. Since PatentsView data are available until 2021, citations within three years after issuance can be measured up to 2018 and partially for 2019. I follow the empirical approach in the working paper to Akcigit and Kerr (2018) and distinguish follow-on citations made by the firm itself as self-follow-on citations. The overall purpose of follow-on measures is to understand how important patented inventions are for innovations by the firm itself (own innovation) and other innovators (creative destruction).

I measure the novelty of inventions by counting the pairwise combinations of CPC subgroups that appear for the first time in the patent database, following Arts and Fleming (2018). I measure the scope of patents by counting the number of words in the first claim and normalize by the average number of words in the first claim for patents filed in the same CPC group and year, following Kuhn and Thompson (2019). I include data from Marx and Fuegi (2019) counting the number of citations to scientific papers (non-patent literature, NPL citations) as a measure of how focused on basic research inventions are. I also include data from Bena et al. (2021) to measure how many process and product innovations firms patent.

2.3. Methodology

I use a difference-in-differences setting to estimate the causal treatment effect of *Alice* on innovation and growth outcomes. I estimate a two-way fixed effects (TWFE) regression model on firm- and industry-level LBD data for the years between 2010 and 2019, with the post period starting in 2015.[9] On the firm level, the treatment variable is the share of treated patents among all patents issued to the firm between 2000 and 2009. On the industry level, I similarly use the share of treated patents to all patents assigned to firms in the four-digit NAICS industry between 2000 and 2009.[10]

The main regression model thus is:

[8] I use filing years to normalize the quality distribution of patents but grant years for the outcomes since I want to know the quality of eligible innovations relative to the time of the invention rather than the time of grant.

[9] For the following analysis, it should be noted that LBD variables such as employment and establishment count are measured in March for a given year, while the innovation and patenting outcomes are measured at the end of the calendar year. Thus, the first full year for the post-period for LBD variables is 2016.

[10] The results remain unchanged if I define treatment instead as the share of patents in *Alice*-related CPC-groups from figure 2.2 alone, without further classification using the NLP method. Since software-related CPC groups include more technologies than which are affected by Alice, the treatment effects are smaller and less precisely estimated.

$$Y_{i,t} = \alpha_i + \lambda_t + \delta_{Alice} * d_i * Post + X'_{i,t} * \beta + \epsilon_{i,t} \tag{2.1}$$

where $Y_{i,t}$ is the outcome variable, d_i is the treated patent share, $Post$ the post-period dummy, and the coefficient δ_{Alice} measures the post-treatment effect. α_i and λ_t are fixed effects for the firm or four-digit NAICS industry and year. Depending on the specification, firm-level controls, $X'_{i,t}$, are the lagged overall employment of the company, the patent portfolio size, and the firm age, all of which are determined at the same time as the continuous treatment (at the beginning of 2010) and interacted with the post-dummy or year dummies (Goldsmith-Pinkham et al. 2020).

To test the necessary parallel trend assumption for identification, I run an event study model for the years between 2010 and 2019:

$$Y_{i,t} = \alpha_i + \lambda_t + \sum_{q=-4}^{5} \delta_{Alice,q} * d_{i,q} + X'_{i,t} * \beta + \epsilon_{i,t} \tag{2.2}$$

where $d_{i,q}$ is the treatment variable interacted with a dummy for the respective year q relative to 2014. For the parallel trend assumption not to be rejected, we need to have no significant effects for the coefficients $\delta_{Alice,q}$ in the years before 2014 and only significant effects in the post-period. I normalize the coefficient for year 2013 to zero, the year immediately before *Alice*, i.e., $\delta_{Alice,-1} = 0$.

One challenge for DiD models like these is estimating standard errors (Bertrand, Duflo, and Mullainathan 2004). In general, my setting fits well to what Abadie, Athey, Imbens, and Wooldridge 2017 describe as a design problem for which clustering on treatment level, i.e., along industry or firm, is appropriate. However, we might also be worried about correlated standard errors within industries and sectors. Thus, in specification variations for the firm-level model, I will estimate clustered standard errors on the industry level instead of the firm level as for my main specification.[11]

[11] We might be concerned that larger clusters might increase the bias of standard error estimates. Angrist and Pischke (2009) suggest that even for fewer than 30 clusters, clustering on group level leads to reasonable results (see, e.g., Bertrand, Duflo, and Mullainathan 2004, Hansen 2007).

Chapter 3

Results

My study provides three main results: first, *Alice* decreases new establishment entry by incumbent multi-unit firms but leaves new venture creation and single-unit entry unchanged. Thus, patentability has little effect along the extensive margin of firm creation but is relevant for incumbent firm growth. Second, companies receive fewer, more narrow patents, and citation measures show more usage of internal knowledge. The results are consistent with innovation becoming more focused on own innovation after *Alice*. Third, the largest, most innovative firms drive the results. Firms in the top decile in terms of employment and patent portfolio size decrease new establishment and job creation following the limited patentability of inventions and cite more internal knowledge. Overall, the three results are consistent with patentability incentivizing creative destruction for large incumbents.

3.1. Industry Results

We can use the *Alice* shock to ask how patenting affects new firm creation and firm exit on the industry level. I use firm-level statistics to identify the industries with the largest share of treated patents. First, I select industries that had at least 20 distinct firms with patent assignments between 2000 and 2009, allowing for a pre-period to test the parallel trend assumption and avoiding issues with selection due to the precursor decisions to *Alice*. While the precise number of patenting firms in an industry is not relevant to the result, having too few patents and firms with patents in an industry poses issues for identification. As with the firm level, the treatment variable can take extreme values if the numbers of patents involved become too small, and we might include industries in the treatment or control groups for which patenting is not important if too few firms receive patents. The treatment variable then is the share of treated patents to all patents assigned to firms in the four-digit NAICS industry between 2000 and 2009. While this treatment definition is based on the patent portfolios of incumbents, it identifies industries for which *Alice* is relevant and the patentability of inventions is important.[1]

We can use the difference-in-differences setting from equation 2.1 to measure the effect of *Alice* on industry entry and exit. The observations are on the industry-year level for four-digit NAICS industries between 2010 and 2019. I follow Kerr and Nanda (2009) and use as outcomes the log of the sum of entering and exiting establishments and total employment of entering and exiting establishments. As in Kerr and Nanda (2009), I fill zero values on the left-hand side to one and add a dummy variable to control for those cases.[2] Furthermore, I can distinguish between single-unit and multi-unit firms. Single-unit firms have only one establishment, thus single-unit entries and exits are in most cases the same as new venture creation and small firm exit.[3] Multi-unit entry and exit, on the other hand, refers to the entry and exit of establishments of firms that are predominantly large incumbents and have multiple other locations.

Tables 3.1 and 3.2 show the DiD results for industry-level entry and exit. Overall, establishment entry decreases, especially weighted by employment. The treatment effect of -0.63 log points is almost completely driven by a decrease in multi-unit entry with a treatment effect of -1.14 for multi-unit entry and a similarly-sized effect of -0.92 for employment-weighted multi-unit entry. With an average continuous treatment of 0.034 for treated industries, this translates to a decrease in multi-unit entry of 3.1 to 3.8%. In no specification does single-unit entry seem to change following *Alice*. This result is robust to the inclusion of sector-year fixed effects and sector-trends as controls, using non-retimed entry, and using a binary treatment variable for industries above the median of the treated patent share. Thus, our first important result is that *Alice* had no effect on small firm entry or exit, but reduced new establishment creation by large multi-unit firms.

---

[1] For robustness, I also implement an approach that is independent of firm-level links and additional assumptions for how to aggregate firm patent portfolios to industry level to identify treated industries: I directly use the ‘Algorithmic Links with Probabilities’ (ALP) patent crosswalk from Goldschlag, Lybbert, and Zolas (2016) and Lybbert and Zolas (2012). This approach is based on NLP methods and provides a direct link between CPC/USPC classes and industry NAICS codes by matching keywords from patent abstracts to industry and product classification descriptions. This means patent classes can be matched according to how important the described technologies are for industries. Since I classify all patents for specific CPC groups, I can define aggregate statistics on treatment and overall patent issuances on CPC group level. I focus on the ten years between 2000 and 2009 and the sum of all issued patents for each CPC group and the sum of all treated patents within each of the groups that I classified. I then use the NAICS-CPC bridge and calculate the probability-weighted total count of issued patents and the count of treated patents for each four-digit NAICS industry. I restrict to industries above the median number of total issued patents. I define as treated industries those with an above median share of treated patents relative to all issued patents and use the rest as control industries. Since I first condition on the total number of assigned patents, I ensure that patents are relevant for both the treated and control industries and that the *Alice*-treated share is not driven by outliers with few relevant patents overall. The main results remain unchanged with no increase in single-unit entry and a strong decrease in multi-unit establishment entry.

[2] the results remain unchanged if I use instead ‘log(y + 1)’ on the LHS, as for most of the firm-level results in section 3.2.

[3] Only a small share of entering and exiting firms have multiple establishments at birth or death, and most firms in

Table 3.1. Industry entry

	Entry	Entry Emp.	MU Entry	SU Entry	MU Entry Emp.	SU Entry Emp.
Cont. Treatment * Post	-0.300	-0.627*	-1.135***	0.245	-0.924**	0.054
	(0.409)	(0.371)	(0.391)	(0.501)	(0.443)	(0.444)
Fixed Effects	Ind. + Year	Ind. + Year	Ind. + Year	Ind. + Year	Ind. + Year	Ind. + Year
S.E. Cluster	Ind.	Ind.	Ind.	Ind.	Ind.	Ind.
N	1,800	1,800	1,800	1,800	1,800	1,800
adj. R sq.	0.99	0.95	0.96	0.99	0.89	0.96

Notes: Difference-in-differences for industry entry. Observations are on the industry-year level. The dependent variables are the total of all entering establishments, the sum of overall employment of entering establishments, the total of all multi-unit establishment entries, the total of single-unit entries, the sum of employment of multi-unit establishment entries, and the sum of employment of single-unit entrants. Single-unit entries are new firm formations, multi-unit entries are new facilities created by existing firms. All dependent variables are in log terms. All regressions include a dummy for zero-values of the dependent variable; the respective observations are replaced with one. Treatment is the share of *Alice*-treated patents relative to all assigned patents between 2000 and 2009 to firms in the industry, the post-dummy is positive for years after 2014. Clustered standard errors are reported in parenthesis, $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$.

Table 3.2. Industry exit

	Exit	Exit Emp.	MU Exit	SU Exit	MU Exit Emp.	SU Exit Emp.
Cont. Treatment * Post	-0.025	-0.136	-0.048	0.128	-0.049	0.044
	(0.399)	(0.480)	(0.405)	(0.415)	(0.610)	(0.373)
Fixed Effects	Ind. + Year	Ind. + Year	Ind. + Year	Ind. + Year	Ind. + Year	Ind. + Year
S.E. Cluster	Ind.	Ind.	Ind.	Ind.	Ind.	Ind.
N	1,800	1,800	1,800	1,800	1,800	1,800
adj. R sq.	0.99	0.95	0.97	0.99	0.88	0.97

Notes: Difference-in-differences for industry exit. Observations are on the industry-year level. The dependent variables are the total of all exiting establishments, the sum of overall previous year employment of exiting establishments, the total of all multi-unit establishment exits, the total of single-unit exits, the sum of previous year employment of multi-unit establishment exits, and the sum of previous year employment of single-unit exits. Single-unit exits are firm with one location ceasing operations, multi-unit exits are the closing of facilities by firms with other continuing locations. All regressions include a dummy for zero-values of the dependent variable; the respective observations are replaced with one. Treatment is the share of *Alice*-treated patents relative to all assigned patents between 2000 and 2009 to firms in the industry, the post-dummy is positive for years after 2014. Clustered standard errors are reported in parenthesis, $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$.

Table 3.3. Patenting by entrants

	Count Issue	Sum Issue	Count Issue Treated	Sum Issue Treated
Cont. Treatment * post-dummy	-1.763***	-0.971	-3.792***	-5.546***
	(0.650)	(1.078)	(0.924)	(1.421)
Fixed Effects	Ind. + Year	Ind. + Year	Ind. + Year	Ind. + Year
S.E. Cluster	Ind.	Ind.	Ind.	Ind.
N	1,800	1,800	1,800	1,800
adj. R sq.	0.87	0.78	0.81	0.84

Notes: Difference-in-differences for the patenting of entrants. Observations are on the industry-year level. The dependent variables are the count of entering establishments with at least one patent assigned in the year of entry, the sum of patents issued to entrants, the count of entrants with at least one *Alice*-treated patents assigned in the year of entry, and the sum of *Alice*-treated patents assigned to entrants. All dependent variables are in log terms. All regressions include a dummy for zero-values of the dependent variable; the respective observations are replaced with one. Treatment is the share of *Alice*-treated patents relative to all assigned patents between 2000 and 2009 to firms in the industry, the post-dummy is positive for years after 2014. Clustered standard errors are reported in parenthesis, $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$.

Using the patent assignee-firm link from Dreisigmeyer et al. (2018) and the firm identifiers in LBD, we can observe entering establishments that received patents in the year of entry and how many patents are assigned. I follow the same econometric approach as for the entry counts above and use the overall count of entering establishments that received at least one patent in the year of entry, the total number of patents received, and also the count of entrants with *Alice* treated patents assigned and the sum of treated patents assigned to entering establishments. The results change little if we instead use the patents received in the year before or after entry.[4] Table 3.3 shows that for all measures except the sum of patents issued we find significant negative treatment effects. The average treatment effect for entrants with at least one patent assigned to them of 5.98% (DiD coefficient -1.76 * 3.4% treatment variable average for treated industries) is even larger than the effect on multi-unit entry in table 3.1. Furthermore, since the effect on the total number of patents received by entering establishments is insignificant, firms that want to create new establishments seem to require more patents than pre-*Alice* for entry. We also see that the magnitude of the treatment effects for *Alice*-treated patents are much larger than for total patent issuances. Thus, the decrease in the patentability of *Alice* related inventions drives the decrease in patents received by entrants. Overall, our second important result in this section shows that the

the economy are single-unit establishments.

[4]The only estimate that changes is the treatment effect for the total number of patents assigned to entrants, which becomes significantly negative.

decrease in establishment entry is accompanied by a decrease in entrants that receive patents in the year of entry.

Taking stock, we find that *Alice* led to a reduction in new establishment creation by large firms. New ventures seem to enter with new products no matter whether they own a patent or not. Thus, in section 3.2, I focus on how incumbent firms are changing their growth and innovation strategy following *Alice*. Since I also find no significant effects on firm exit, *Alice* was a shock to innovation rather than the continuation of current products. In the following section 3.1, I look at industry-level R&D surveys to measure how innovation strategies have changed after *Alice*.

R&D Results

I use R&D survey responses from BRDIS years 2010 to 2019 to gain insight into innovation efforts on the industry level. I merge survey responses with the industry-level patenting statistics from section 3.1 to identify treated industries. I adapt the model 2.1 and weight observations by their respective survey weights.[5] Thus, the weighted observations are representative of industry-wide innovation efforts and estimated treatment effects measure industry-level responses to *Alice*.[6]

It should be noted that survey responses go through multiple stages of editing and adjustments before weights are assigned.[7] Also, the survey design has changed multiple times over the years, leading to potential inconsistencies across the years. The results, thus, should be taken as evidence for the directional effect of *Alice* on innovation outcomes rather than as precise estimates.

First, in tables 3.4 and 3.5, I show results for different measures of R&D expenses relative to sales and in absolute log terms for both, worldwide and domestics R&D expenses. In no specification, I find significant treatment effects.

[5]Aggregating to industry-level first would overstate the precision of observations measured with few survey responses, nevertheless we find also here no significant change in R&D spending.

[6]Since large innovative firms are more likely to be sampled by BRDIS, we can link certain firms directly to their R&D responses. This is somewhat problematic since BRDIS is not longitudinal and we thus add sampling error to our analysis. Furthermore, BRDIS samples heavily from manufacturing firms and the DiD setting only works reasonably well with multiple pre- and post-period observations, restricting the number of observations even more. Nevertheless, the results from tables 3.4, 3.5, and 3.6 hold up with no clear change in R&D efforts and a directional change away from research. It should be noted that the absolute dollar amounts of R&D expenses do show a positive treatment effect of ca. 0.5, but not the relative measures, leaving no conclusive evidence that R&D efforts change in the end.

[7]The results remain unchanged when using unweighted observations.

Table 3.4. Industry worldwide R&D spending

	Log R&D Exp.	Log R&D Perf.	Log R&D Emp.	R&D Exp. to Sales	R&D Perf. to Sales	R&D Emp. to Emp.
Cont. Treatment * Post	-0.237	-0.202	0.170	-0.001	0.003	-0.006
	(0.391)	(0.402)	(0.137)	(0.010)	(0.011)	(0.024)
Fixed Effects	Ind. + Year	Ind. + Year	Ind. + Year	Ind. + Year	Ind. + Year	Ind. + Year
S.E. Cluster	Ind.	Ind.	Ind.	Ind.	Ind.	Ind.
N	167,000	167,000	167,000	167,000	167,000	167,000
adj. R sq.	0.39	0.39	0.31	0.17	0.16	0.27

Notes: Difference-in-differences for the weighted annual R&D survey responses for worldwide variables. The dependent variables are total R&D expenses, total R&D performed, total R&D employment, R&D expenses to sales, R&D performed to sales, and R&D employment to total employment, all variables are worldwide measures. Dependent variables in log terms include '+1'. Treatment is the share of *Alice*-treated patents relative to all assigned patents between 2000 and 2009 to firms in the industry, the post-dummy is positive for years after 2014. All observations are weighted by the sample weight assigned by BRDIS, calculated as the ratio of weighted worldwide sales to adjusted worldwide sales. Clustered standard errors are reported in parenthesis, $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

Table 3.5. Industry domestic R&D spending

	Log U.S. R&D Perf.	Log U.S. R&D Emp.	U.S. R&D Perf. to U.S. Sales	U.S. R&D Emp. to U.S. Emp.
Cont. Treatment * Post	-0.204	0.159	0.004	-0.004
	(0.404)	(0.138)	(0.014)	(0.025)
Fixed Effects	Ind. + Year	Ind. + Year	Ind. + Year	Ind. + Year
S.E. Cluster	Ind.	Ind.	Ind.	Ind.
N	167,000	167,000	167,000	167,000
adj. R sq.	0.39	0.31	0.00	0.27

Notes: Difference-in-differences for the weighted annual R&D survey responses for domestic/US variables. The dependent variables are total R&D performed, R&D employment, R&D performed to sales, and R&D employment to total employment, all variables are US/domestic measures. The regression sample and specifications are the same as in table 3.4. Clustered standard errors are reported in parenthesis, $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

BRDIS also asks for the type of R&D spending. In table 3.6, I show results for the four R&D spending types measured relative to worldwide R&D expenses. I normalize with the worldwide R&D spending to measure the direction of R&D efforts rather than their absolute size in Dollar terms. Here, research is related to the exploration of new ideas, and BRDIS distinguished research further into basic research as “activities aimed at acquiring new knowledge or understanding without specific immediate commercial applications” and applied research “aimed at solving a specific problem or meeting a specific commercial objective”. Development is more focused on the use of internal knowledge and improving current products through the “systematic use of research and practical experience to produce new or significantly improved goods, services, or processes”.[8].

Development spending has a positive DiD coefficient of 0.26, while research, applied research, and basic research spending all decrease with DiD coefficients of -0.32 to -0.15.[9] The overall magnitude of the effects is small, though: the mean for the treatment variable is 3.1%, thus, the average treatment effect for the increase in development spending is 0.8% and for the decrease in research spending ca. 1%. While the limited patentability of innovations after *Alice* has not changed innovation efforts in affected industries, it did change the direction of innovation away from exploration-oriented research and more directed toward internal product improvement-oriented development.[10]

To summarize the industry results, we found a clear decrease in new establishment creation by large incumbent firms after *Alice* in the most affected industries. Firms are also less likely to receive a patent in the year of entry, providing direct evidence that the limited patentability of inventions is linked to the decrease in new establishment creation. R&D survey responses indicate no decrease in R&D effort, but rather a redirection of innovation efforts toward development rather than research. Overall, this is in line with the limited patentability of innovations incentivizing incumbents to focus more on improving current products rather than exploring new ideas, leading to less creative destruction in favor of more own innovation.

3.2. Firm Results

Following the insight from section 3.1 that new firm creation is not affected by *Alice* and multi-unit establishment entry decreases, we can focus in this section on the effect of *Alice* on growth and innovation outcomes for incumbent firms in the sample period of 2010 to 2019. I restrict to

[8] See: `https://www.census.gov/programs-surveys/brds/about/faq.html`

[9] Note that the changes in innovation spending types do not add up to 1 since the survey design does neither impose a clear cut differentiation between the spending types nor that the responses have to add up to the total R&D spending.

[10] Less research and more development do not mean that *Alice*-treated patents are primarily research related. Patents are relevant for both, research and development, and we can use patent measures in section 3.2 to analyze how the innovation direction has changed after *Alice*.

Table 3.6. Industry R&D spending type (relative to R&D spending)

	Research	Development	Applied Research	Basic Research
Cont. Treatment * Post	-0.321***	0.265***	-0.145*	-0.174**
	(0.082)	(0.087)	(0.080)	(0.077)
Fixed Effects	Ind. + Year	Ind. + Year	Ind. + Year	Ind. + Year
S.E. Cluster	Ind.	Ind.	Ind.	Ind.
N	88,500	88,500	88,500	88,500
adj. R sq.	0.15	0.15	0.10	0.11

Notes: Difference-in-differences for the weighted annual R&D survey responses for R&D spending types relative to worldwide R&D expenses. The dependent variables are research spending relative to worldwide R&D expenses, development spending relative to worldwide R&D expenses, applied research spending relative to worldwide R&D expenses, and basic research spending relative to worldwide R&D expenses. Treatment is the share of *Alice*-treated patents relative to all assigned patents between 2000 and 2009 to firms in the industry, the post-dummy is positive for years after 2014. All observations are weighted by the sample weight assigned by BRDIS, calculated as the ratio of weighted worldwide sales to adjusted worldwide sales. Clustered standard errors are reported in parenthesis, ${}^{*}p < 0.1$, ${}^{**}p < 0.05$, ${}^{***}p < 0.01$.

active firms with positive employment that received at least one patent between 2000 and 2009.[11] As in section 3.1, I restrict to firms in industries with at least 20 patenting firms, that is industries with at least 20 unique firms that received patent assignments between 2000 and 2009. The logic is the same as above; I limit the influence of outliers and focus on industries for which patenting is relevant. Furthermore, in section 3.2, I split the sample along the within-industry distribution of patent portfolios in 2010. Thus, I need to have a sufficient number of firms within each industry to define a patent portfolio size distribution.

In the following, I generally refer to the estimated difference-in-difference coefficients as treatment effects. The average for the treatment variable in the full sample is 0.44 for treated firms and for the largest patenting/firm size group in subsections 3.2 and 3.2 the average is 0.087 for treated firms. The average treatment effect for the treated, thus, is the DiD coefficient multiplied by these averages of the continuous treatment variable (see Callaway and Sant'Anna 2021 for more details on the assumptions and interpretation of average treatment effects in DiD settings with continuous treatment).

[11] Note that restricting to firms with at least one establishment with positive employment does not mean that firms cannot drop out of the sample over the years due to, e.g., mergers or firm death. For robustness, I run the same analysis as below for continuing firms that have positive employment throughout the sample period. The results remain unchanged.

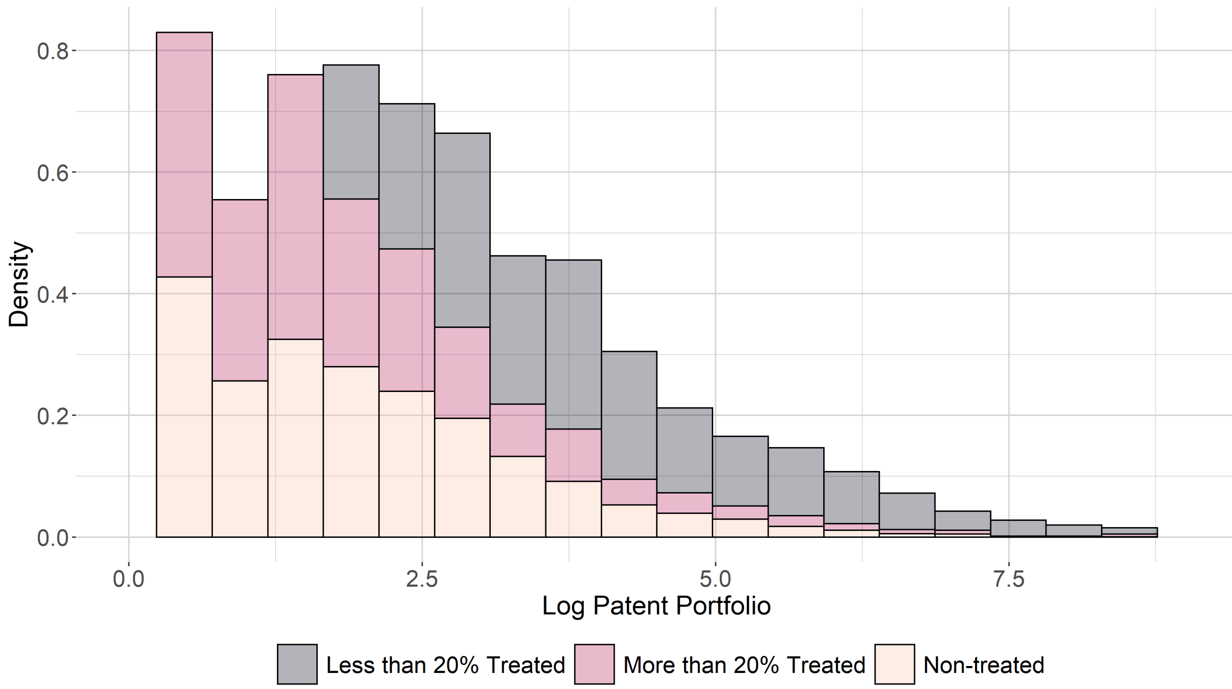


Fig. 3.1. Patent portfolio distribution - 2010

Notes: Density-scale histogram for cross-section of log patent portfolios size in 2010, grouped by *Alice*-treated patent shares. Patent portfolios are the unweighted total of all patents issued to the same U.S. company assignee between 2000 and 2009. The three groups of portfolios are portfolios with less than 20% of issued patents classified as *Alice*-treated by the NLP algorithm, portfolios with more than 20% of the patent portfolio treated, and portfolios without treated patents. The data are sourced from PatentsView.

Balancing analysis

Before running my main analysis, it is important to understand how the *Alice* treatment is distributed across firms and industries, and test whether covariates might have a confounding effect for identification.

First, I use PatentsView data to plot the histogram of the cross-sectional patent portfolio size distribution in 2010. Patent portfolios are the total count of patents assigned to the same company assignees between 2000 and 2009. In figure 3.1, I distinguish between patent portfolios without treated patents, portfolios with less than 20% treated patents, and portfolios with more than 20% treated patents for the year 2010; at the beginning of the pre-period.

Larger treatment shares are related to smaller patent portfolios. This makes sense since with fewer patents the treatment shares are more like discrete steps. For example, a patent portfolio of two patents can have a treatment share of 0%, 50%, or 100%. For larger patent portfolios, the distribution is more continuous. Patent portfolios with small treated patent shares are overall much large than the average non-treated portfolio.[12] This seems reasonable since firms with many patents are more likely to hold also a treated patent. Furthermore, as hinted on in figure 2.2 and described in, e.g., Webb et al. (2018) and Forman and Goldfarb (2020), software patents have been rising dramatically up to *Alice* and the concentration of these patents has increased. The identifying parallel trend assumption for my DiD method is not violated, though, if the distribution of the treatment variables is correlated with the level of covariates. Going forward, I can include patent portfolio size as a covariate in my TWFE regression model in equation 2.1 and match treated and control firms based on patent portfolio size.

I follow Goldsmith-Pinkham et al. (2020) and run in table 3.7 cross-sectional regressions of the treatment variable in 2010 on different covariates related to size and change in outcomes.

Patent portfolio and firm size, measured in log terms, are correlated with treatment, as is firm age. With the unconditional average of the treatment variable around 0.05 in 2010, the magnitudes of the coefficients are overall small: for example, doubling the firm size increases the expected treatment value by just 0.002 to 0.004, between 5 and 10% of the unconditional average. The low R^2s are also in line with the wide distribution of covariates for treated patent portfolios in figure 3.1. Importantly, none of the covariates related to changes in outcomes–employment growth, job creation and destruction rates, establishment entry and exit rates–are significantly correlated with treatment.

Overall, while treated firms are larger than non-treated firms, the share of treated patents is not very closely correlated with any covariate and seems to be uncorrelated with covariates related to changes in outcomes. Especially the latter is important for not rejecting the parallel trend assumption. The correlation table for the covariates in table 3.8 confirms this: no correlation coefficient of any covariate with the treatment variable is larger than 0.081, much smaller than the correlation between, e.g., the patent portfolio and firm size of 0.46.

Finally, it is important to understand the distribution of treated firms across industries. In table 3.9, I show the five industries with the highest share of the treatment variable in 2010. That is the sum of the treatment variables within each industry relative to the cross-sectional sum of the treated variable (this is related to the Rotemberg weights from Goldsmith-Pinkham et al. 2020). I also include the average of the treatment variables as well as the share of treated firms (firms with at least one treated patent) and treated patents assigned between 2000 and 2009 each industry accounts for.

[12] Consistent with larger firms having overall smaller treatment variables, the coefficient for the patent portfolio size in table 3.7 for the pooled regression in the right-most column is negative.

Table 3.7. Cross-sectional regression - continuous treatment on covariates

Log PatentPF	-0.002***								-0.006***
	(0.001)								(0.001)
Log Emp - Lag		0.003***							0.006***
		(0.001)							(0.001)
Firm Age			0.000**						-0.001***
			(0.000)						(0.000)
Emp. Growth				0.004					0.009
				(0.003)					(0.016)
Entry Rate					0.000				-0.009
					(0.014)				(0.018)
Exit Rate						0.003			-0.005
						(0.010)			(0.013)
Job Creation Rate							0.003		-0.004
							(0.005)		(0.017)
Job Destruction Rate								-0.004	0.003
								(0.004)	(0.016)
Fixed Effects	Ind.	Ind.	Ind.	Ind.	Ind.	Ind.	Ind.	Ind.	Ind.
S.E. Cluster	Ind.	Ind.	Ind.	Ind.	Ind.	Ind.	Ind.	Ind.	Ind.
N	25,500	25,500	25,500	25,500	25,500	25,500	25,500	25,500	25,500
adj. R sq.	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13

Notes: Cross-sectional regression of the treatment variables on firm covariates in 2010. The dependent variable is the share of *Alice*-treated patents relative to all patents issued to the firm between 2000 and 2009. ‘Log PatentPF’ is the log count of patents assigned to the firm between 2000 and 2009, ‘Log Emp - Lag’ is the log of total firm employment in 2009, all other covariates are 2010 values; definitions are in Appendix A. Clustered standard errors are reported in parenthesis, $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$.

Table 3.8. Correlation table covariates - 2010 cross-section

Treatment	1								
Log PatentPF	-0.043	1							
Log Emp - Lag	0.017	0.456	1						
Firm Age	-0.081	0.197	0.520	1					
Emp. Growth	0.007	-0.019	-0.088	-0.047	1				
Entry Rate	0.019	0.110	0.193	0.081	0.026	1			
Exit Rate	0.015	0.168	0.273	0.143	-0.115	0.199	1		
JC Rate	0.018	-0.013	-0.084	-0.118	0.631	0.265	0.085	1	
JD Rate	0.004	0.019	0.052	-0.027	-0.809	0.074	0.200	-0.151	1

Notes: Correlation table for the cross-sectional variables used in table 3.7.

Table 3.9. Cross-industry distribution of treatment

NAICS	Firms	Avg. Treatment	Share Treatment	Share Treated Firms	Share Patents	Share Treated Pat.
Computer Systems Design and Related Services (5415)	1,200	0.162	0.220	0.159	0.070	0.262
Software Publishers (5112)	450	0.210	0.104	0.085	0.022	0.117
Data Processing, Hosting, and Related Services (5182)	250	0.243	0.062	0.043	0.007	0.011
Management, Scientific, and Technical Consulting Services (5416)	450	0.091	0.047	0.028	0.002	0.006
Other Financial Investment Activities (5239)	100	0.288	0.034	0.019	0.001	0.003

Notes: Cross-sectional distribution of treatment for the most treated industries. All variables are measured in 2010. Column two shows the number of patenting firms in the industry and column three the average of the treatment variable in the industry. Column four shows the sum of the treatment variable in the industry relative to the overall sum of the treatment across industries. Column five shows the number of treated firms in the industry relative to the total number of treated firms across industries. Treated firms have at least one treated patent assigned to them between 2000 and 2009. Column six shows the sum of patents received by firms in the industry relative to all patents assigned to firms between 2000 and 2009, and column seven the sum of treated patents assigned to firms in the industry relative to all treated patents assigned between 2000 and 2009.

The most treated industries all belong to service sectors related to software innovations. This is consistent with *Alice* invalidating software and business method patents. Computer Systems Design and Related Services (5415) stands out and accounts for 22% of the total treatment and 26% of all treated patents assigned to firms. Treated firms, i.e. firms with at least treated patents assigned between 2000 and 2009, are less concentrated than the treatment variable itself; industry 5415 only account for 15.9% of all treated firms. The second most affected industry, Software Publishers (5112), accounts for ca. 10.4% of the total treatment, 11.7% of all treated patents, and 8.5% of treated firms, with all other sectors being in the single-digits. Thus, the firms in the two most treated industries 5415 and 5112 alone account for ca. 1/3 of the overall treatment.

Having treatment concentrated in software-related service sectors does not invalidate the parallel trend assumption, but we need to be careful and think about robustness analysis. First, excluding the most treated industry 5415 does not change the results. Including four-digit-industry-year fixed effects as controls generally consumes most of the treatment effects due to the relatively high concentration of the treatment variable in just two four-digit industries. However, including higher-level sector/two-digit-industry-year fixed does not change the results. Similar to including lower-level industry-year fixed effects, restricting the sample to service sectors does not qualitatively change the results but reduces the statistical significance of the DiD coefficients. Overall, treatment is correlated with industry but the main results, especially in section 3.2, are generally robust controlling for industry-wide trends.[13]

Patenting results

I start the firm-level analysis by estimating the treatment effects of *Alice* on patenting measures related to innovation effort and direction for the full firm sample. In tables 3.10, 3.11, 3.13, and 3.12, I report the DiD results for measures of patenting quality, novelty, scope, and follow-on innovations. As a first-stage result, table 3.10 confirms that *Alice* decreased the ability of firms to patent inventions in software-related technologies: the annual number of issued patents has a negative treatment effect of 9%, with the same estimate for treated patents being 13.3%. The difference in treatment effects for forward-citation weighted patent issuances is even stronger (15.3% treatment effect for all patent issuance and 35.5% for treated patents), confirming that *Alice* limited the patentability of software-related innovations.

Table 3.11 adds to this result by showing that low-citation patents are most affected by *Alice*: patents issuances in the top 10% of the forward citations distribution do not change for treated firms, patent issuances receiving few or no forward citations, however, have a negative treatment effect

[13] It should be noted that my treatment is not directly fit to distinguish treatment intensity within industries. Besides the measurement errors from the NLP model and the patent-firm link, treatment is more discrete with more extreme values for smaller firms. Furthermore, a control firm in an affected industry that has not filed *Alice*-treated patents before 2009 might intend to innovate in *Alice*-related technologies later.

Table 3.10. Firm-level patenting - absolute measures

	Total Patent Issue	Treated Pat. Issue	Fwd. Cites	Fwd. Cites Treated
Cont. Treatment * Post	-0.090***	-0.133***	-0.153***	-0.355***
	(0.021)	(0.016)	(0.041)	(0.034)
Fixed Effects	Firm+Year	Firm+Year	Firm+Year	Firm+Year
S.E. Cluster	Firm	Firm	Firm	Firm
N	197,000	197,000	197,000	197,000
adj. R sq.	0.85	0.83	0.73	0.67

Notes: Difference-in-differences for absolute patenting outcomes. Observations are on the firm-year level for the year of patent issuance. The dependent variables are the total number of issued patents to the firm, total forward citations to issued patents, total number of *Alice*-treated patents issued to the firm, and the total forward citations to issued *Alice*-treated patents. Dependent variables are in log terms and include '+1'. Treatment is the share of *Alice*-treated patents relative to all assigned patents to the firm between 2000 and 2009, the post-dummy is positive for years after 2014. Clustered standard errors are reported in parenthesis, $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

of ca. 5%. Normalized forward citations show a negative treatment effect of 4.8%, around half the size of the treatment effect for overall issuances in table 3.10, meaning that the patenting quality has decreased less than the absolute number of issued patents. Software and business method patents are generally considered to be of low value (Hall and MacGarvie 2010), thus we would expect the *Alice* shock to affect lower-value patents by raising the patentability threshold of innovations rather than breakthrough and relevant inventions. This is also consistent with the industry-level R&D results from section 3.1, overall innovation efforts do not change, but rather the direction of innovation.

Furthermore, my results show a direct link between low-citation patents and establishment entry and job creation. Recent research by Abrams et al. (2019) and Abrams et al. (2013) show an inverted U-shape for the relation between forward citations and patent value with some high-value patents receiving few citations, arguing that low-citation patents add additional protection against infringement for other valuable inventions. I add to this idea by showing empirically how low-citation *Alice* patents strengthen the patent protection of new ideas and incentivize large companies to explore new technologies and create new establishments.

Tables 3.13 and 3.12 show that patented innovations are becoming narrower in scope and more focused on incremental own innovation. The number of total claims decreases sharply with a treatment effect of -29.1%, while the normalized number of words in the first claim has a treatment effect of -6.2%, which is smaller than the overall decrease in issued patents. With relatively fewer claims and more words in the first claim, the scope of patents is decreased (Kuhn and Thompson

Table 3.11. Firm-level patenting - quality measures

	Top 1% Cited	Top 2-10% Cited	Cited Patents	Uncited Patents	Norm. Cites
Cont. Treatment * Post	-0.002	-0.012	-0.050***	-0.047***	-0.048**
	(0.004)	(0.010)	(0.017)	(0.016)	(0.021)
Fixed Effects	Firm+Year	Firm+Year	Firm+Year	Firm+Year	Firm+Year
S.E. Cluster	Firm	Firm	Firm	Firm	Firm
N	197,000	197,000	197,000	197,000	197,000
adj. R sq.	0.67	0.80	0.81	0.79	0.79

Notes: Difference-in-differences for quality adjusted patenting outcomes. Observations are on the firm-year level for the year of patent issuance. The dependent variables are the total number of issued patents in the top 1% of the forward citations distribution for patents in the same CPC-group and filing year, the patents which are in the top 10% of the CPC-group-filing year forward citation distribution but not in the top 1%, the total number of patents that are outside of the top 10% of the CPC-group-filing year forward citation distribution but receive at least one citation, the number of patents that do not receive forward citations, and the total sum of forward citations normalized with the average number of forward citations in the CPC-group-filing year. Dependent variables are in log terms and include '+1'. Treatment is the share of *Alice*-treated patents relative to all assigned patents to the firm between 2000 and 2009, the post-dummy is positive for years after 2014. Clustered standard errors are reported in parenthesis, $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$.

2019).

Alice did not lead to the exploration of novel technologies since the number of patents filed in CPC groups the firm has not filed in before slightly decreases, as does the number of new CPC-subgroup combinations appearing for the first time in the data and the number of product patents in table 3.13. A decrease in novelty and exploration is also consistent with the strong negative treatment effect for citations to scientific papers and basic research of 12.5%(Marx and Fuegi 2019).[14]

Table 3.13 shows a strong negative treatment effect for follow-on citations of 10.6%, but no significant effect for self-follow-on citations. That is, patents receive within the first three years after issuance fewer citations from outside inventors, but the number of citations from other inventions by the firm does not significantly decrease. We can interpret this as a decrease in creative destruction and a relative increase in own innovation since the inventions patented by the firm are less relevant for potential outside inventors compared to follow-on innovation by the firm itself. This

[14]I find also a strong negative treatment effect for technology proximity and the number of patents issued in CPC groups known to the firm (Known CPC-Group) of -0.062 and -0.085, respectively. While the magnitude is close to the overall treatment effect for patent issuance in table 3.10, these two results are partially mechanical since *Alice* is concentrated among specific CPC groups. Treated firms by definition hold patents in *Alice*-affected CPC groups and thus have to receive fewer patents in these groups following *Alice*. Neither effect is larger than the overall treatment effect of issued patents, further supporting that *Alice* did not induce exploration of new technologies.

Table 3.12. Firm-level patenting - scope and innovation direction measures

	Claims	Bwd. Cites	Self-Cites	New CPC Grp.	Wd. Claim	Norm. Wd. Claim
Cont. Treatment * Post	-0.291***	-0.246***	0.001	-0.014*	-14.840***	-0.062***
	(0.054)	(0.052)	(0.025)	(0.008)	(3.682)	(0.016)
Fixed Effects	Firm+Year	Firm+Year	Firm+Year	Firm+Year	Firm+Year	Firm+Year
S.E. Cluster	Firm	Firm	Firm	Firm	Firm	Firm
N	197,000	197,000	197,000	197,000	197,000	197,000
adj. R sq.	0.73	0.72	0.80	0.63	0.44	0.44

Notes: Difference-in-differences for patenting scope and innovation direction outcomes. Observations are on the firm-year level for the year of patent issuance. The dependent variables are the number of claims, the total number of backward citations, the number of self-citations, the number of patents in new CPC-groups the firm has not filed in before, the average number of words in the first claim, and the average number of words in the first claim normalized with the average for patents in the same CPC-group-filing year. Dependent variables are in log terms and include '+1'. Treatment is the share of *Alice*-treated patents relative to all assigned patents to the firm between 2000 and 2009, the post-dummy is positive for years after 2014. Clustered standard errors are reported in parenthesis, $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$.

is also confirmed when looking at the prior art patents cite in table 3.12: while overall backward citations have a strong negative treatment effect of -24.6%, self-citations show no significant effect, meaning that less external knowledge is referenced while prior innovations by the firm itself remain relevant. Following Akcigit and Kerr (2018), this is consistent with treated firms substituting new product development in favor of the improvement of existing products.

Table 3.13. Firm-level patenting - novelty, process patents, and follow-on innovation measures

	NPL-Science Bwd. Cites	Follow-on Cites	Self-Follow-on Cites	Non-Process Patents	Process Patents	New Combinations
Cont. Treatment * Post	-0.125***	-0.106***	0.000	-0.022**	-0.036***	-0.043*
	(0.035)	(0.025)	(0.016)	(0.011)	(0.010)	(0.024)
Fixed Effects	Firm+Year	Firm+Year	Firm+Year	Firm+Year	Firm+Year	Firm+Year
S.E. Cluster	Firm	Firm	Firm	Firm	Firm	Firm
N	197,000	197,000	197,000	197,000	197,000	197,000
adj. R sq.	0.79	0.79	0.79	0.82	0.80	0.70

Notes: Difference-in-differences for patenting novelty, process patents, and follow-on innovation outcomes. Observations are on the firm-year level for the year of patent issuance. The dependent variables are the number of backward citations to scientific articles, the sum of follow-on (forward) citations received within three years after issuance, the sum of follow-on (forward) citations by other patents issued to the same firm within three years, the number of non-process, the number process patents, and the number of pairwise CPC-subgroup combinations that appear for the first time in the data. Dependent variables are in log terms and include '+1'. Treatment is the share of *Alice*-treated patents relative to all assigned patents to the firm between 2000 and 2009, the post-dummy is positive for years after 2014. Clustered standard errors are reported in parenthesis, $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Overall, the patenting results confirm that treated firms receive fewer patents after *Alice*. This is concentrated among low-citation patents. While high-value innovations still receive patents, references to external knowledge decrease while references to inventions by the firm itself become relatively more relevant, consistent with the marginal innovation direction shifting from creative destruction toward own innovation.

Firm-level growth

In tables 3.14 and 3.15, I show the DiD regression results for a broad range of variables from the LBD related to growth, employment, and establishment creation and destruction on the full firm sample. First, establishment entry and job creation decrease for treated firms after *Alice*. Job creation has a negative treatment effect of 12.5%, driven in part by the decrease in employment from entering establishments with a negative treatment effect of 7.8%. Interesting is also that wage share in table 3.14 has a negative treatment effect of 5.5% and job destruction a positive effect of 7.5%, though the last result is not robust in all specifications and only significant at the 10% level. As in Peters (2020) and Peters and Zilibotti (2021), we can interpret the decrease in the wage share as an increase in markups.[15] Taken together, the decrease in establishment entry and job creation and increase in markups are overall consistent with incumbent firms shifting away from the exploration and introduction of new technologies and toward improving existing products to increase the pricing power.[16]

[15]For around 80% of the firm in the LBD data, we have data on revenue and the total wage bill. We can thus define the wage share as the ratio of the total wage bill to revenue. Following De Loecker and Warzynski (2012), markups are the product of the output elasticity of labor and the inverse of the wage share, assuming labor is the only input. If the output elasticity is constant, we can measure changes in the log markups through our DiD setting using the log of the inverse of the wage share as our outcome. In our TWFE model, the time and firm fixed effects control for the constant output elasticity. This is also consistent with De Loecker et al. (2020) who find firm-level evidence for the direct inverse relation between markups and wage share. Thus, our TWFE model estimates treatment effects for markups if the LHS is the inverse of the wage share.

[16]The results are robust for the same specification variations as described in section 3.2.

Table 3.14. Firm revenue, wage share, and market share

	Real Rev.	Wage Share	Market Share (Emp.)
Cont. Treatment * Post	0.044	-0.055***	-0.039
	(0.028)	(0.020)	(0.032)
Fixed Effects	Firm+Year	Firm+Year	Firm+Year
S.E. Cluster	Firm	Firm	Firm
N	146,000	146,000	197,000
adj. R sq.	0.98	0.81	0.96

Notes: Difference-in-differences for revenue, wage share, and market share. Observations are on the firm-year level. The dependent variables are the log of real revenue, the log of the wage share, and the log of the employment market share within the 4-digit-industry. Treatment is the share of *Alice*-treated patents relative to all assigned patents to the firm between 2000 and 2009, the post-dummy is positive for years after 2014. Clustered standard errors are reported in parenthesis, $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$.

Table 3.15. Firm establishments

	Estabs (pos. emp.)	Emp.	Entry	Entry Emp.	Exit	Exit Emp.	Job Creation	Job Destruction
Cont. Treatment * Post	0.007	0.000	-0.032	-0.078**	-0.012	-0.023	-0.125***	0.075*
	(0.014)	(0.028)	(0.020)	(0.038)	(0.019)	(0.038)	(0.047)	(0.043)
Fixed Effects	Firm+Year	Firm+Year	Firm+Year	Firm+Year	Firm+Year	Firm+Year	Firm+Year	Firm+Year
S.E. Cluster	Firm	Firm	Firm	Firm	Firm	Firm	Firm	Firm
N	197,000	197,000	197,000	197,000	197,000	197,000	197,000	197,000
adj. R sq.	0.98	0.97	0.64	0.53	0.69	0.59	0.66	0.65

Notes: Difference-in-differences for establishment-related outcomes. Observations are on the firm-year level. The dependent variables are the total number of establishments with positive employment, total employment, the number of entering establishments, the employment of entering establishments, the number of exiting establishments, the employment in t-1 of exiting establishments, and total job creation and destruction. Dependent variables are in log terms and include '+1', except for the log of the count of establishments and total employment which are required to be positive. Treatment is the share of *Alice*-treated patents relative to all assigned patents to the firm between 2000 and 2009, the post-dummy is positive for years after 2014. Clustered standard errors are reported in parenthesis, $^{*}p < 0.1,^{**}p < 0.05,^{***}p < 0.01$.

Firm-size variation

To better understand which types of firms are driving the effect in section 3.2, I split my sample along two dimensions: the size of the patent portfolio and total firm employment at the beginning of 2010, i.e., at the beginning of the sample period. Within each four-digit NAICS industry, I split the sample into firms below the 50th percentile, firms in the 50th to 90th percentile, and firms in the top 10 percent of the patent and employment size distribution, respectively. Table 3.16 show the means of relevant variables within the different brackets. As described in section 3.2, firms are very heterogeneous in terms of patent portfolio and employment size; while the smallest firms have on average less than 40 employees and between one and four patents in their portfolio, companies in the top brackets have several thousands of employees and hundreds of patents in their portfolio on average. Particularly firms that are above the 90th percentile of the patent portfolio and also above the 90th percentile of employment size distribution stand out, accounting for almost half of all new establishment entries and close to 80% of all patent issuances. This might also explain why small firm entry and exit on the industry level does not change in section 3.1: small firms with one establishment make up a large share of entering and exiting establishments, but hardly file patents in the first place. Patenting seems less relevant for small firms in *Alice* affected industries.

Table 3.16. Summary statistics by employment and patent portfolio size

Patent Tier	Emp. Tier	N	Treatment	Estabs.	Emp.	Patent PF	Mkt. Share	Entry	Entry Emp	JC	Wage Share	Pat. Issue	Follow-on	Self-F/O	Bwd. Cit.	Self-Cit.
<0.5 Pctl.	<0.5 Pctl.	37,000	0.018	1.63	45.96	1.55	0.000	0.04	0.84	4.77	0.29	0.22	0.32	0.11	7.57	0.27
<0.5 Pctl.	0.5-0.9 Pctl.	19,000	0.020	8.74	369.80	2.85	0.001	0.59	10.02	37.79	0.28	0.60	0.92	0.37	15.62	0.88
<0.5 Pctl.	>0.9 Pctl.	2,100	0.025	85.66	2,194.00	3.40	0.009	6.50	65.74	246.90	0.26	1.02	1.73	0.92	30.00	1.45
0.5-0.9 Pctl.	<0.5 Pctl.	53,500	0.036	1.46	35.45	4.36	0.000	0.04	0.70	3.76	0.30	0.59	0.69	0.25	13.86	0.88
0.5-0.9 Pctl.	0.5-0.9 Pctl.	53,000	0.041	10.42	518.80	10.86	0.001	0.62	12.58	55.08	0.30	1.69	2.34	0.86	46.88	3.07
0.5-0.9 Pctl.	>0.9 Pctl.	9,900	0.052	136.70	5,735.00	28.95	0.012	10.34	202.50	690.80	0.27	6.61	7.99	3.37	140.30	9.89
>0.9 Pctl.	<0.5 Pctl.	3,200	0.016	1.62	39.96	48.53	0.000	0.06	0.50	3.70	0.30	5.33	4.73	1.50	79.96	6.43
>0.9 Pctl.	0.5-0.9 Pctl.	8,400	0.025	28.40	909.10	124.20	0.003	2.58	27.24	99.33	0.26	15.27	20.44	10.10	292.00	48.00
>0.9 Pctl.	>0.9 Pctl.	10,500	0.032	229.60	10,950.00	857.00	0.028	19.31	337.90	1,235.00	0.23	115.40	158.60	63.51	2,295.00	259.40

Notes: Summary statistics by patent portfolio and firm employment size groups. Firms are grouped based on their rank in the patent portfolio and firm employment size distribution within their 4-digit NAICS industry. Within each industry, firms either are below the 50th percentile, between the 50th and 90th percentile, or above the 90th percentile in the distribution of the total number of patents assigned to firms between 2000 and 2009, and independent from this below the 50th percentile, between the 50th and 90th percentile, or above the 90th percentile in the total firm employment size distribution in 2009. N is the total number of observations, treatment is the share of *Alice*-treated patents relative to all assigned patents to the firm between 2000 and 2009, all other variables are defined in Appendix A.

We can now perform the analysis from above within each patent portfolio /employment size bracket. This also improves sample balancing since we now explicitly control for firm employment and patent portfolio size.[17] In tables 3.17a and 3.17b, I report the treatment effects for patent issuances and establishment entry for different patent portfolio and employment size groups. The treatment effects for patenting are negative and significant for multiple patent portfolio/employment size groups since *Alice* limited the overall patentability of software-related innovations. However, the largest patent portfolio/employment size group, firms above the 90th percentile of the patent portfolio as well as employment size distribution, are most affected with a negative treatment effect of 1.256. Only this group has a significant treatment effect of -1.069 for establishment entry. This shows how the limitation of patentability of innovations is most relevant for the largest patenting firms and leads to a decrease in the exploration of new technologies and establishment entry.[18]

[17]Performing the same cross-sectional regression of treatment on covariates as in table 3.7, no covariates are significant for the largest patent portfolio/employment size group.

[18]For robustness, I implement an alternative sample split with absolute patent portfolio and firm employment size cutoffs. I group firms in patent portfolio brackets counting the number of patents assigned to them between 2000 and 2009: 1 to 5 patents, 6 to 19, or more than 20 patents. Independent from this, I define employment size brackets based on the total firm employment in 2009: 1 to 99 employees, 100 to 999 employees, or more than 1000 employees. The cutoffs are selected to allow for sufficient sample size within each bracket; the cutoffs for the top brackets are close to the 90-95th percentile in the respective patent portfolio and employment size distribution. The results confirm that the patenting shock is more widely distributed across the different size brackets. Also in this sample split, the treatment effect for establishment entry is only significant for the largest patent portfolio and employment size bracket, with the magnitude of -0.84 close to the magnitude from table 3.17b for the largest patent portfolio /employment size group.

Table 3.17. Firm-level entry and patenting by employment and patent portfolio size

(a) Firm-level patent issuances by employment and patent portfolio size

Total Patent Issue	<0.5 Pctl. Emp.	0.5-0.9 Pctl. Emp.	>0.9 Pctl. Emp.
<0.5 Pctl. Pat.	-0.030	-0.015	-0.131
	(0.054)	(0.093)	(0.106)
0.5-0.9 Pctl. Pat.	-0.049***	-0.108***	0.001
	(0.018)	(0.033)	(0.100)
>0.9 Pctl. Pat.	-0.522	-0.869**	-1.256***
	(0.430)	(0.343)	(0.296)
Fixed Effects	Firm+Year	Firm+Year	Firm+Year
S.E. Cluster	Firm	Firm	Firm

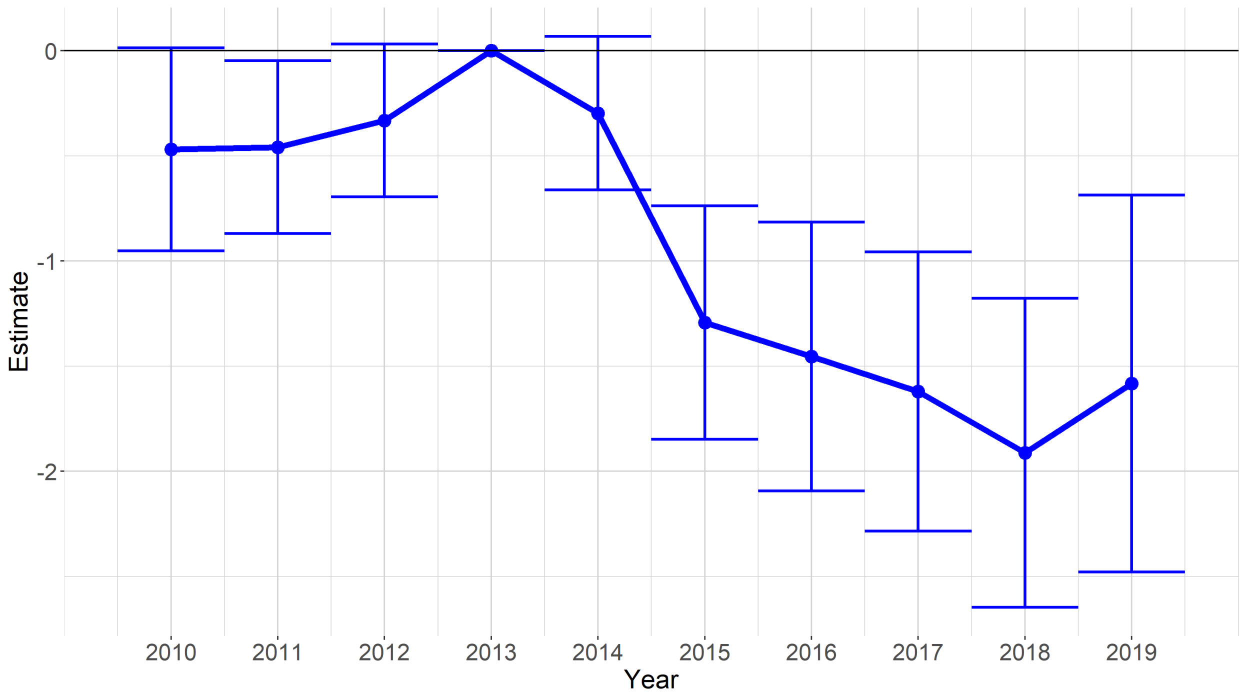
Notes: Difference-in-differences for total annual patent issuances for different patent portfolio and firm employment size groups. Within each 4-digit NAICS industry, firms are ranked below the 50th percentile, between the 50th and 90th percentile, or above the 90th percentile based on the total number of patents assigned between 2000 and 2009 (row), and independent from this based on total firm employment in 2009 (column). For each of the possible nine sub-sample groups, the difference-in-differences coefficient for the setting in table 3.10 is reported. Dependent variables are in log terms and include '+1'. Clustered standard errors are reported in parenthesis, $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

(b) Firm-level establishment entry by employment and patent portfolio size

Entry	<0.5 Pctl. Emp.	0.5-0.9 Pctl. Emp.	>0.9 Pctl. Emp.
<0.5 Pctl. Pat.	0.095	0.079	-0.145
	(0.065)	(0.066)	(0.173)
0.5-0.9 Pctl. Pat.	0.013	-0.029	-0.222
	(0.012)	(0.026)	(0.135)
>0.9 Pctl. Pat.	0.004	0.053	-1.069***
	(0.043)	(0.231)	(0.335)
Fixed Effects	Firm+Year	Firm+Year	Firm+Year
S.E. Cluster	Firm	Firm	Firm

Notes: Difference-in-differences for annual establishment entry for different patent portfolio and firm employment size groups. Within each 4-digit NAICS industry, firms are ranked below the 50th percentile, between the 50th and 90th percentile, or above the 90th percentile based on the total number of patents assigned between 2000 and 2009 (row), and independent from this based on total firm employment in 2009 (column). For each of the possible nine sub-sample groups, the difference-in-differences coefficient for the setting in table 3.15 is reported. Dependent variables are in log terms and include '+1'. Clustered standard errors are reported in parenthesis, $^{*}p < 0.1, ^{**}p < 0.05, ^{***}p < 0.01$.

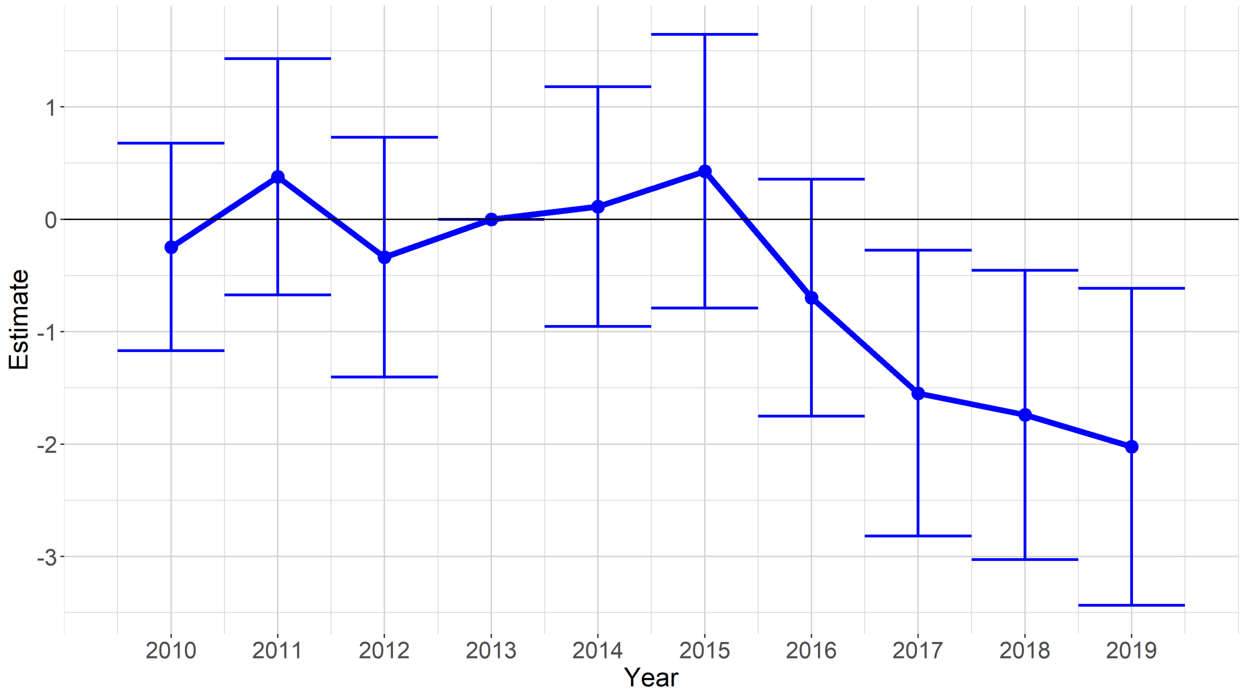
Fig. 3.2. Event-study - Total Patent Issue - >90th Pctl. Pat. - >90th Pctl. Emp.



Notes: Event study for total annual patent issuances for the group of firms above the within-industry 90th percentile in the patent portfolio size distribution as well as firm employment size distribution. The variables and sample are defined in table 3.17a, treatment effects are estimated with the event-study model in 2.2. Dependent variables are in log terms and include '+1'. Coefficients for the interaction of the treatment variable with the year dummies are shown with error collars for the 95% confidence interval. Standard errors are clustered on firm level.

To test the parallel trend assumption for the most treated group of firms above the 90th percentile of the patent portfolio as well as employment size distribution, I use model 2.1 to plot in figures 3.2 and B.1 the event study results for patent issuances and establishment entry. For neither outcome variable, we find a pre-trend or an inverted "Ashenfelter's Dip" which could indicate a violation of our identifying assumption. We also see that the treatment effect for patent issuances sets in immediately in 2015. The effect on establishment entry gradually increases starting in 2016. This is consistent with the reduced patentability of innovations causing establishment entry to decrease over time. Note also that LBD data measure employment in March and the *Alice* decision was in June 2014, thus, the first full year after *Alice* is 2016.

Fig. 3.3. Event-study - Entry - >90th Pctl. Pat. - >90th Pctl. Emp.



Notes: Event study for establishment entry for the group of firms above the within-industry 90th percentile in the patent portfolio size distribution as well as firm employment size distribution. The variables and sample are defined in table 3.17b, treatment effects are estimated with the event-study model in 2.2. Dependent variables are in log terms and include '+1'. Coefficients for the interaction of the treatment variable with the year dummies are shown with error collars for the 95% confidence interval. Standard errors are clustered on firm level.

Large patenting firms

The results in section 3.2 show that firms in above the 90th percentile of the patent portfolio and employment size distribution within industries are most affected by *Alice*. In this section, I implement different variations of the DiD model for this group of firms to estimate robust treatment effects for the most relevant outcome variables. Tables 3.18, 3.19, and 3.20 report treatment effects for four model variants: the base specification without covariates, the base specification with standard errors clustered on industry level, and two specifications that include log firm employment size, log patenting portfolio size, and firm age, all measured in 2009, as covariates. The first specification with covariates interacts the controls with the post-dummy, the second with year dummies. These are the only significant covariates in the cross-sectional regression in table 3.9, including them as controls might help to improve precision of the DiD coefficient estimate (Angrist and Pischke

2009).[19] In appendix B, I show the event-study results for the base specification and confirm that there are no pre-trends, except for the follow-on citations with a marginally significant coefficient for the year 2010. This might be related to fewer patents being issued after *Alice* that could cite a given patent. Nevertheless, the treatment effects for follow-on citations are highly significant after 2015 and the magnitude remains much larger than for the self-follow-on citations.[20] The average of the treatment variable for the treated firms is 0.087, thus the average treatment effect for the treated can be calculated by multiplying the difference-in-difference coefficients with this average (I generally refer to the estimated difference-in-difference coefficients as treatment effects).

The results in table 3.18 show a strong shock to patent issuances following *Alice* with treatment effects between -1.40 and -1.26. We also see a clear gap in the magnitudes of the treatment effects for total follow-on citations (-1.53 to -1.42) and self-follow-on citations (-0.83 to -0.64). The difference between the treatment effects for backward citations (-2.56 to -2.47) and self-citation (-0.97 to -0.75) is even larger than between follow-on citations and self-follow-on citations. Overall, this is consistent with treated firms focusing more on internal knowledge relative to external knowledge. Following Akcigit and Kerr (2018), we can interpret this as treated firms shifting away from creative destruction innovation toward own innovation.

Table 3.19 shows negative treatment effects for establishment entry (-1.07 to -0.64), employment from establishment entry (-1.84 to -1.29), and job creation (-1.57 to -1.14).[21] Thus, all results confirm that *Alice* has a negative effect on the exploration of new technologies; the limited patentability of inventions leads treated firms to introduce fewer new technologies and thus reduces the creation of new establishments and job creation.

Note also that we have a negative treatment effect for job destruction, which is significant at the 10%-level for the specification with standard errors clustered on the industry level. In table 3.15, however, we have a positive treatment effect for job destruction in the full sample. Unreported results confirm that we find for firms in the mid-tiers of the patent portfolio and employment size distribution positive treatment effects for job destruction (which are not robust, though, to the specification variations in this section). Thus, *Alice* had a different effect for smaller treated firms: *Alice* limited the enforceability of some of the patents in their portfolio, leading to a higher likelihood that these firms face more competition in some of their products, which increases job destruction.

[19] Since the treatment is continuous and the shock nation-wide, there are few group-level controls that are useful to include. The results are robust to excluding observations in industry 5415, winsorizing the top 1% of observations, and restricting to firms with non-missing observations for the entire sample period of 2010 to 2019. The results remain largely unchanged when including three-digit-industry-year fixed effects. Furthermore, I can use Coarsened Exact Matching (CEM) on the log firm employment size, log patenting portfolio size, firm age, and sector dummies to balance the covariate distribution of treated and control firms at the beginning of the pre-period in 2010. The results remain largely unchanged when using the CEM-balanced sample for the analysis.

[20] Also event study results for the full sample show no pre-trend except for follow-on citations.

[21] Establishment entry-related results are robust to controlling for re-timed entry and excluding re-activated establishments that temporarily had no employees for one to seven years before entry.

Table 3.18. Firm-level patenting variables - >90th Pctl. Pat. & >90th Pctl. Emp. - Specification variations

Total Patent Issue				
Cont. Treatment * Post	-1.256***	-1.256***	-1.399***	-1.401***
	(0.296)	(0.266)	(0.295)	(0.295)
N	10,500	10,500	10,500	10,500
adj. R sq.	0.90	0.90	0.90	0.90
Backward Citations				
Cont. Treatment * Post	-2.470***	-2.470***	-2.557***	-2.564***
	(0.466)	(0.369)	(0.471)	(0.471)
N	10,500	10,500	10,500	10,500
adj. R sq.	0.81	0.81	0.81	0.81
Self-Citations				
Cont. Treatment * Post	-0.747*	-0.747**	-0.966**	-0.970**
	(0.423)	(0.381)	(0.422)	(0.422)
N	10,500	10,500	10,500	10,500
adj. R sq.	0.86	0.86	0.86	0.86
Total Follow-on				
Cont. Treatment * Post	-1.417***	-1.417***	-1.532***	-1.534***
	(0.323)	(0.270)	(0.317)	(0.317)
N	10,500	10,500	10,500	10,500
adj. R sq.	0.86	0.86	0.86	0.86
Self-Follow-on				
Cont. Treatment * Post	-0.638**	-0.638***	-0.823***	-0.826***
	(0.308)	(0.228)	(0.303)	(0.303)
N	10,500	10,500	10,500	10,500
adj. R sq.	0.84	0.84	0.85	0.85
Fixed Effects	Firm+Year	Firm+Year	Firm+Year	Firm+Year
S.E. Cluster	Firm	Industry	Firm	Firm
Control	N	N	Post-interaction	Year-interaction

Notes: Difference-in-differences results for different model specifications - patenting outcomes. The dependent variables are defined as in tables 3.10 and 3.13. Results are shown for the group of firms above the within-industry 90th percentile in the patent portfolio size distribution as well as employment size distribution, as defined in table 3.17b. The first column shows results for the base DiD setting as used above. Column two clusters standard errors on the industry level without additional covariates. Columns three and four include as covariates the log firm employment size, log patent portfolio size, and firm age, all measured in 2009. Column three interacts the covariates with the post-dummy, column four interacts the covariates with year dummies. Clustered standard errors are reported in parenthesis, $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$.

The firms in the largest patent portfolio/employment size group, on the other hand, have on average 857 patents assigned to them between 2000 and 2009; even with a shock to the enforceability of software-related patents, they still hold large patent portfolios of valid patents that are unlikely to be more successfully challenged after *Alice* (see also Choi and Gerlach 2015). Instead, the negative treatment effect for job destruction for this group of firms is more consistent with the transition into a new equilibrium with less creative destruction. As shown in figures 3.2 and B.1 in section 3.2, the decrease in patent issuances after *Alice* leads to a reduction in establishment entry over time. If firms move over time into a new equilibrium with less creative destruction, we would expect also job destruction to eventually follow. Since the treatment effect for job destruction in the largest patent portfolio/employment size group is only marginally significant, the new equilibrium might not have been reached yet by the end of the sample period in 2019.[22]

Finally, table 3.20 shows that the negative treatment effects for establishment count and total employment only survive for the specifications without controls. The serial correlation of establishment count and firm employment with the size-related covariates might explain the insignificant estimates. The magnitude of the treatment effects for the significant estimates without covariates is for both outcomes in line with the treatment effect for the market share (-0.64 to -0.57). The decrease in creative destruction following *Alice* eventually leads to smaller market shares for treated firms.[23] Since I find no clear positive treatment effect for the market share in any other patent portfolio/employment size group, it seems likely that the treated firms in the largest patent portfolio/employment size group lose market share to a wider range of competitors.

Note also that the treatment effect for the market share is larger than what we would expect from a decrease in new job creation alone (considering that the job creation rate is closer to 10% of total firm employment and the magnitude of the treatment effect for job creation is only around twice as large as the treatment effect for the market share). One explanation could be that treated firms sell operations related to new technology exploration through M&A activities, which is consistent with unreported results showing that the number of establishments that move to a new owner after *Alice* increases (see, e.g., Garcia-Macia et al. 2019). Overall, the relatively large decrease in market share for treated firms is not inconsistent with the interpretation that *Alice* reduces the incentive to explore new technologies.

[22] I do not report treatment effects for establishment exit since I cannot control for re-timed and re-activated establishment exits as for establishment entries. The coverage of the relevant re-timing flag for exiting establishments is much lower than for entering establishments. Re-timing is overall more of a challenge for exits, as can be seen in spikes for establishment exits around Economic Census years. Since a large part of the treatment effect is driven by the last two sample years (after the last Economic Census in 2017), I focus on job destruction which is more directly observable for those years. Nevertheless, I find a significant negative treatment trend for establishment exits (-0.206 per year after 2014) and employment of exiting establishments (-0.381 per year after 2014) for the largest patent portfolio/employment size group of firms, which is consistent with a long-term decrease in creative destruction.

[23] Note that overall the effects are too small to change the industry-wide HHI and other concentration measures.

Table 3.19. Firm-level growth variables - >90th Pctl. Pat. & >90th Pctl. Emp. - Specification Variations

Entry				
Cont. Treatment * Post	-1.069***	-1.069***	-0.652**	-0.640**
	(0.335)	(0.323)	(0.315)	(0.314)
N	10,500	10,500	10,500	10,500
adj. R sq.	0.59	0.59	0.60	0.60
Entry Emp				
Cont. Treatment * Post	-1.840***	-1.840***	-1.289**	-1.267**
	(0.544)	(0.496)	(0.535)	(0.535)
N	10,500	10,500	10,500	10,500
adj. R sq.	0.46	0.46	0.46	0.46
Job Creation				
Cont. Treatment * Post	-1.571***	-1.571***	-1.161**	-1.146*
	(0.594)	(0.510)	(0.591)	(0.591)
N	10,500	10,500	10,500	10,500
adj. R sq.	0.53	0.53	0.53	0.54
Job Destruction				
Cont. Treatment * Post	-0.710	-0.710*	-0.342	-0.333
	(0.505)	(0.383)	(0.508)	(0.509)
N	10,500	10,500	10,500	10,500
adj. R sq.	0.56	0.56	0.56	0.56
Fixed Effects	Firm+Year	Firm+Year	Firm+Year	Firm+Year
S.E. Cluster	Firm	Industry	Firm	Firm
Control	N	N	Post-interaction	Year-interaction

Notes: Difference-in-differences results for different model specifications - growth outcomes. The dependent variables are defined as in table 3.15. Results are shown for the group of firms above the within-industry 90th percentile in the patent portfolio size distribution as well as employment size distribution, as defined in table 3.17b. The first column shows results for the base DiD setting as used above. Column two clusters standard errors on the industry level without additional covariates. Columns three and four include as covariates the log firm employment size, log patent portfolio size, and firm age, all measured in 2009. Column three interacts the covariates with the post-dummy, column four interacts the covariates with year dummies. Clustered standard errors are reported in parenthesis, $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$.

Table 3.20 also shows a positive treatment effect of 0.22 to 0.25 for markups. The estimate is insignificant for one of the specifications with covariates and only significant at the 5%-level for standard errors clustered at the industry level. It should be noted, though, that the sample size is more than 20% smaller for markups than for the other outcomes since the last available year is 2018, and the magnitude of the estimates does not change much across variations. The increase in markups for treated firms is consistent with the shift of innovation efforts toward own innovation. Improving the current products in place increased the gap to the next-best competitor and increase the pricing power of incumbent firms, thus allowing for higher markups.

3.3. Model

I use a simplified Schumpeterian growth model based on Peters (2020) to quantify the welfare effect of the *Alice* decision. I calibrate changes to the innovation rates of technology exploration and own innovation with the DiD results from section 3.2.

The economy has a representative household with log-utilities in continuous time, inelastic labor supply of L, wage level w, and risk-neutral interest rate r. Household consumption C_t is a fixed share of the final good Y_t. Y_t is produced by a unit mass of intermediate products. I assume for simplicity that the elasticity of substitution across varieties of intermediate goods is one. $y(i)$ is the quantity and $q(i)$ the quality of intermediate product i.[24]

$$U = \int_0^\infty e^{-rt} \ln C_t dt$$

$$\ln Y_t = \int_0^1 \ln(q_t(i) y_t(i)) di$$

J incumbent firms, which are all price takers on the labor market, only use labor as input to produce product i such that $y(i) = l(i)$. N_j is the set of products owned by firm j and $n_j = |N_j|$. There is no entry since the results in section 3.1 show that new firm creation is not affected by *Alice*.

Product quality $q(i)$ is a quality ladder with proportional improvements of size $\lambda > 1$. Firm j is the quality leader for product i and the quality gap to the second-highest quality producer is $\lambda^{\Delta(i)}$. That is, firm j is $\Delta(i)$ rungs above the second-highest quality producer in the quality ladder. Firms engage in Bertrand competition, which implies that only the quality leader j is active. The markup $\mu(i)$ that j can charge for product i depends on its quality lead:

$$\mu(i) = \lambda^{\Delta(i)}$$

[24] I drop in the following the time subscript if there is no ambiguity.

Table 3.20. Firm-level size variables and markup - >90th Pctl. Pat. & >90th Pctl. Emp. - Specification Variations

Estabs. (pos. emp.)				
Cont. Treatment * Post	-0.425*	-0.425***	-0.385	-0.383
	(0.241)	(0.147)	(0.244)	(0.244)
N	10,500	10,500	10,500	10,500
adj. R sq.	0.95	0.95	0.95	0.95
Emp.				
Cont. Treatment * Post	-0.461**	-0.461**	-0.355	-0.353
	(0.225)	(0.211)	(0.220)	(0.220)
N	10,500	10,500	10,500	10,500
adj. R sq.	0.92	0.92	0.92	0.92
Markup				
Cont. Treatment * Post	0.246*	0.246**	0.219	0.221*
	(0.128)	(0.111)	(0.134)	(0.134)
N	7,900	7,900	7,900	7,900
adj. R sq.	0.85	0.85	0.85	0.85
Market Share (Emp.)				
Cont. Treatment * Post	-0.637**	-0.637**	-0.570**	-0.567**
	(0.270)	(0.290)	(0.271)	(0.271)
N	10,500	10,500	10,500	10,500
adj. R sq.	0.88	0.88	0.88	0.88
Fixed Effects	Firm+Year	Firm+Year	Firm+Year	Firm+Year
S.E. Cluster	Firm	Industry	Firm	Firm
Control	N	N	Post-interaction	Year-interaction

Notes: Difference-in-differences results for different model specifications - size outcomes. The dependent variables are defined as in tables 3.14 and 3.15; Markup is the inverse of Wage Share (in log-terms). Results are shown for the group of firms above the within-industry 90th percentile in the patent portfolio size distribution as well as employment size distribution, as defined in table 3.17b. The first column shows results for the base DiD setting as used above. Column two clusters standard errors on the industry level without additional covariates. Columns three and four include as covariates the log firm employment size, log patent portfolio size, and firm age, all measured in 2009. Column three interacts the covariates with the post-dummy, column four interacts the covariates with year dummies. Clustered standard errors are reported in parenthesis, $^{*}p < 0.1$,$^{**}p < 0.05$,$^{***}p < 0.01$.

The flow profits for firm j in product i are $\pi(i) = \left(1 - \frac{1}{\mu(i)}\right) = \left(1 - \frac{1}{\lambda^{\Delta(i)}}\right)$, where I use that with elasticity of substitution of unity the household spends the same amount on each product i. I normalize the amount spent on each product to one.[25]

Firm j can lose product i through creative destruction if a competitor successfully innovates on product i and becomes the new quality leader. I model creative destruction as the flow rate τ. Since I observe no significant effect of *Alice* on establishment exit or job destruction, I take τ as exogenous.[26]

We can now calculate the present value (PV) $V(i)$ of product i. The discount rate combines the risk-neutral interest rate r and the rate of creative destruction τ. In other words, firm j takes into account the time value of expected future profits and the probability of losing the product to a competitor.

$$V(i) = \int_0^\infty \pi(i) e^{-(r+\tau)s} ds = \frac{\pi(i)}{r+\tau} = \frac{\left(1 - \frac{1}{\mu(i)}\right)}{r+\tau} \tag{3.1}$$

Firm j can engage in two types of innovation: first, improve product quality $q(i)$ by own innovation and increase the gap to the second-highest quality producer from $\Delta(i)$ to $\Delta(i)+1$. Second, engage in new technology exploration and expand into a new product k. If expansion innovation is successful and the firm takes over product k, it will take the quality lead only by a single rung above the second-highest quality producer such that $\mu(k) = \lambda$ and the quality gap is $\Delta(k) = 1$. In equilibrium, the sum of the expansion innovation efforts by firms would be equal to the overall rate of creative destruction: $\tau = \int_0^1 x(j)dj$.

Both types of innovation are random and firm j decides on the company-wide Poisson arrival rate for own innovation I_j and expansion innovation x_j. For simplicity, I formulate the own innovation and expansion rates as per product rates. The cost functions for x_j and I_j are quadratic and also scale with the number of products n_j: $C(I_j, x_j, n_j) = n_j \frac{\xi_I}{2} I_j^2 + n_j \frac{\xi_x}{2} x_j^2$, where ξ_I and ξ_x are scale parameters.

We can use equation 3.1 to find expressions for the optimal own innovation and expansion rates per product, I and x, by setting the expected marginal increase in $V(i)$ equal to the marginal cost of either innovation effort.

[25] Labor demand for product i then is equal to $l(i) = \frac{1}{w}\frac{1}{\mu(i)}$. We can use this to define firm-level labor demand for firm j as $l_j = \frac{1}{w} n_j \left(\frac{1}{n_j}\sum_{i\in N_j} \frac{1}{\mu(i)}\right) = \frac{1}{w} n_j \frac{1}{\mu_j}$. Firm-level markup μ_j is the harmonic mean over the products of firm j: $\mu_j^{-1} = \frac{1}{n_j}\sum_{i\in N_j}\frac{1}{\mu(i)}$.

[26] As mentioned in section 3.2, I observe toward the end of my sample period a negative treatment effect for job destruction and establishment exit, which would be consistent with a decrease in τ and a new equilibrium over the long run.

$$\frac{\left(1-\frac{1}{\lambda}\right)}{r+\tau}=\xi_x x$$
$$\frac{\left(\frac{1}{\lambda^{\bar{\Delta}}}-\frac{1}{\lambda^{\bar{\Delta}+1}}\right)}{r+\tau}=\xi_I I$$

where $\bar{\Delta}$ is the average quality lead of firm j across its products.

Acquiring a new product is in general more profitable than improving products since $1-\frac{1}{\lambda}>\frac{1}{\lambda^{\bar{\Delta}}}-\frac{1}{\lambda^{\bar{\Delta}+1}}$, i.e., the flow profits $\pi(k)$ from expanding into new product k are larger than the increase in flow profits for existing product i for successful own innovation. We can assume that expansion innovation is also more costly since firm j needs to acquire more external knowledge and faces higher imitation risk by competitors in the new market.

My empirical results in section 3.1 show no change in overall R&D efforts for treated industries, but rather a substitution effect toward product development. My patenting results in section 3.2 also confirm that *Alice* did not prevent breakthrough innovations from being patent eligible, but rather led to less exploration and more internal product improvement. To capture the substitution of expansion innovation with own innovation, I assume that the firms are budget constrained and firm j substitutes the least valuable expansion projects with the next most valuable own innovation projects. Formally, firm j optimizes the following Lagrangian:[27]

$$\max_{x,I}\left\{x\frac{\left(1-\frac{1}{\lambda}\right)}{r+\tau}+I\frac{\left(\frac{1}{\lambda^{\bar{\Delta}}}-\frac{1}{\lambda^{\bar{\Delta}+1}}\right)}{r+\tau}+\Lambda\left(B-\frac{\xi_I}{2}I^2-\frac{\xi_x}{2}x^2\right)\right\}$$

where B is the R&D budget and the constraint is binding. The solution shows that the ratio of expected PVs from innovation is equal to the ratio of the marginal innovation costs:

$$\frac{\left(1-\frac{1}{\lambda}\right)}{\left(\frac{1}{\lambda^{\bar{\Delta}}}-\frac{1}{\lambda^{\bar{\Delta}+1}}\right)}=\frac{\xi_x x}{\xi_I I} \tag{3.2}$$

What does *Alice* mean for the decision to invest into own innovation versus expansion innovation? Decreasing patent protection means that expansion costs increase since entering a new market requires additional efforts by the firm to protect the invention against imitation risk and license external knowledge. In equation 3.2, increasing ξ_x relative to ξ_I on the right-hand side of the equation predicts the same shift away from expansion innovation and new technology exploration toward own innovation as I find in my empirical results after the *Alice* shock.

[27] We can ignore the scaling by the number of products n_j since investment cost, innovation rates, and the project value scale with the product count.

Welfare impact calibration

I can use the estimated treatment effects from section 3.2 and measure how welfare changes when the largest treated firms shift their innovation direction away from technology exploration toward own innovation.

Following Peters (2020), aggregate output is proportional to the quality index of products, $Q = \exp\left(\int_0^1 \ln q(i)\, di\right)$.[28] Growth, thus, depends only on the overall quality improvements to the intermediate products: $g = \sum_J s_j(x+I)\ln(\lambda)$, where s_j is the market share of firm j.[29]

Since Y_t only grows through quality improvements at rate g and $C_t \propto Y_t$, i.e., consumption is a fixed share of total output, the change in welfare following *Alice* is the present value of the (permanent) change in the growth rate $\frac{\Delta g}{r}$.

Table 3.21 summarizes the parameters I use to calibrate my model. First, I set the risk-neutral interest rate r to 3%, similar to the value used in Akcigit and Kerr (2018). Next, I need an estimate for the innovation-step size λ. I follow Aghion et al. (2019) and use the inverse of the markups estimated in Hall (2018) to calibrate λ.[30] Since table 3.9 shows that the two most treated sectors are 5415 and 5112, I focus on the markup estimates for Professional, Scientific, and Technical Services (NAICS 54) and Information (NAICS 51) in Hall (2018), which are 1.31 and 1.39, respectively. I round to 1.3 to account for estimation errors.

Next, we can use the establishment entry rate as proxy for the rate of expansion innovation x, similar to Klenow and Li (2021) and Garcia-Macia et al. (2019). I use industry-wide entry

[28] Aggregate labor demand L is the sum of the individual firm-level labor demand: $L = \sum_J l_j = \frac{Y}{w}\sum_J n_j \frac{1}{\mu_j}$. We can use this to find an expression for the real wage w: $\ln Y = \int_0^1 \ln\left(q(i)y(i)\right) di = \int_0^1 \ln q(i) di + \int_0^1 \ln y(i) di = \ln Q + \int_0^1 \ln\left(\frac{Y}{w}\frac{1}{\mu(i)}\right) di = \ln Q + \ln Y - \ln w - \int_0^1 \ln \mu(i) di$. Thus, the equilibrium wage w is given as $w = Q \exp\left(-\int_0^1 \ln \mu(i) di\right)$. We can substitute this into the expression for the aggregate labor demand and define the misallocation measure $M = \frac{\exp\left(-\int_0^1 \ln \mu(i) di\right)}{\sum_J n_j \mu_j^{-1}}$, such that $Y = QML$. Since I am interested in changes to growth rather than the full equilibrium, I can focus just on the change in overall quality Q and take L and M as constant.

[29] I define as the market share of firm j the total employment of firm j relative to total employment in the four-digit NAICS industry. Note that $\frac{l_j}{L} = \frac{n_j \frac{1}{\mu_j}}{\sum_J n_j \frac{1}{\mu_j}}$; if I assume that before *Alice* markups are relatively evenly distributed across companies, this simplified to $\frac{l_j}{L} = \frac{n_j}{\sum_J n_j}$.

[30] Directly using the inverse of the wage share from my data as an estimate for the markup would overestimate the actual innovation step size since large incumbents with successful own innovations in the past have higher markups than the simple inverse of the step size. The wage share for the largest patent portfolio/employment size group in the summary statistics in table 3.16 is 0.23 and for all other size groups ca. 0.3. If I calibrate λ to 1.3, the ratio of the inverse of the wage share for the firms in the largest patent portfolio/employment size group and the rest indicates an average innovation lead of $\Delta' = 2$ for the largest firms; $\frac{\mu'}{\mu} = \frac{\lambda^{\Delta'}}{\lambda^{\Delta}} = \frac{\frac{1}{0.23}}{\frac{1}{0.3}} = 1.3 = \lambda^{2-1}$.

Table 3.21. Parameter calibration - welfare analysis

Parameter	Name	Estimate	Source
r	Risk-neutral interest rate	0.03	Akcigit and Kerr (2018)
λ	Innovation step-size	1.3	Hall (2018)
x	Expansion innovation rate	0.08	BDS Data - Establishment Entry Rate
I	own innovation rate	0.05	Forward-citation calibration
g	Growth rate	0.034	BLS Data - TFP Growth
s_j	Market share	0.26	Pre-period average for treated firms above the 90th percentile of the patent portfolio and employment size distribution as defined in table 3.17a
d_j	Continuous treatment	0.16	Pre-period average for treated firms above the 90th percentile of the patent portfolio and employment size distribution as defined in table 3.17a
	Treatment effect expansion innovation	-0.6	DiD-coefficient for establishment entry in table 3.19
	Treatment effect own innovation	+0.2	DiD-coefficient for markup in table 3.20

Notes: Parameter calibration for welfare analysis.

rates for sectors 511 and 541 from the Business Dynamics Statistics (BDS). The average annual establishment entry rates for the pre-period (2010 to 2014) are 7.85% and 12.2%, respectively. I approximate the expansion innovation rate with 8%, closer to the more conservative entry rate.[31] Note that in equilibrium, the expansion innovation rate is equal to the rate of creative destruction.

I follow an approach similar to Akcigit and Kerr (2018) and use self-citations to proxy for the own innovation rate I. Since backward citations refer to all prior art and might cite more than the relevant technologies (Kuhn et al. 2020), I use follow-on citations for the calibration. I estimate a share of ca. 40% of self-follow-on citations to total follow-on citations and thus set I equal to 0.05.[32]

With these calibrated parameters, I find an estimate for the total growth rate of $(x + I)\ln\lambda = (0.08 + 0.05) * \ln 1.3 = 0.034$. I compare this to the average total factor productivity (TFP) growth rate estimates from the Bureau of Labor Statistics (BLS) for the pre-period for the two most treated

[31]The estimate based on the establishment entry rate lies between the employment-weighted entry rates (3.82% for 511 and 5.66% for 541) and the job creation rates (11.2% for 511 and 16.3% for 541).

[32]The share of self-follow-on citations is equal to $\frac{I}{I+x} = 0.4$, thus $I = \frac{2}{3} * x = 0.05$.

industries: 2.68% for industry 511 and 3.42% for 5415. Thus, the calibrated parameters lead to a realistic estimate close to the observable TFP growth rate for the most treated industries.

I focus on the treatment effects from section 3.2 for large incumbent firms above the 90th percentile of the patent portfolio and employment size distribution to estimate the welfare impact of *Alice*. The market shares of these large treated firms in sectors 51 and 54 are 0.37 and 0.15, respectively, and the average of the continuous treatment d_j is 0.22 and 0.10. Thus, there is a relatively wide distribution of market shares and treatment across industries. For this analysis, I take the mid-points of both measures, setting the market share of treated firms to 0.26 and the average treatment to 0.16.

I use the treatment effects estimated in tables 3.19 and 3.20 of -0.6 for establishment entry and +0.2 for markups to measure the decrease in expansion innovation and the increase in own innovation for treated firms after *Alice*.[33]

To calculate the effect of *Alice* on expansion innovation, I multiply the treatment effect for establishment entry with the calibrated pre-period expansion innovation rate and weigh the effect with the average continuous treatment variable and the market share of the treated firms: $\Delta x = (-0.6 * 0.08) * (0.26 * 0.16) = -0.0020$. Similarly, I can estimate the increase in own innovation using the treatment effect for markups and the pre-period own innovation rate: $\Delta I = (+0.2{*}0.05){*}(0.26 * 0.16) = 0.0004$. The net-change to overall growth then combines the two effect, multiplied with the innovation step-size: $\Delta g = (\Delta x + \Delta I) \ln \lambda = (-0.0020 + 0.0004) * \ln 1.3 = -0.0004$. The total welfare loss is the present value of the decrease in consumption growth: $\frac{\Delta g}{r} = \frac{-0.0004}{0.03} = -0.013$. Thus, *Alice* leads to a decrease in welfare of ca. 1.3%.

Furthermore, I can use the *Alice* shock to estimate the elasticity of growth to patentability. For my calibrated model, the decrease in the growth rates is equal to $\frac{\Delta g}{g} = \frac{-0.0004}{0.034} = -1.18\%$. Table 3.18 shows a treatment effect for total patent issuances of ca. -1.4. Since the dependent variable is in log terms, patent issuances decrease by $-1.4{*}(0.26{*}0.16) = -5.82\%$, using also the market share of treated firms of 0.26 and the average treatment of 0.16 from above. This implies an elasticity of $-1.18/-5.82 = 0.2$ for the relationship of growth to patentability. One percent decrease in the patentability of innovations leads to a decrease in growth of 0.2% due to large firms shifting from new technology exploration toward own innovation.

Overall, the decrease in welfare of ca. 1.3% after *Alice* is relatively moderate since large firms change the innovation direction toward more own innovation rather than decrease overall innovation efforts.

[33]The treatment effects for employment from establishment entry and job creation are much larger, between -1.15 and -1.27 for the specification with covariates interacted with year dummies. Thus, using the effect on the establishment entry provides a more conservative estimate for the welfare impact of *Alice*.

Chapter 4

Conclusion

Patents, market entry, and innovation are closely related: patents provide a temporary monopoly on inventions to incentivize the exploration of new ideas. New technologies such as computer software challenge the balance between what defines a patent-eligible innovation and what is an abstract idea giving unfair competitive advantages if patented. The Supreme Court decision in *Alice Corp. v. CLS Bank International* limited the patent eligibility of software and business method claims and significantly changed the patenting landscape for inventors. I use this shock to identify the effects of patentability on firm innovation and growth.

I develop a novel NLP-based approach to classify patents if they are at risk of being invalid under *Alice*. I identify treated firms and industries in restricted-use U.S. Census data and use a difference-in-differences approach to estimate the causal effect of reduced patentability on growth and innovation outcomes. I find no significant change in the entry or exit of new ventures, but a significant decrease in new establishment creation by incumbents. On the firm level, I find negative treatment effects for overall patent issuances and job creation but positive effects for markups. Patent citations to internal knowledge become more important and R&D spending shifts toward development. Large patenting firms drive the effects and focus on improving the products they already have in place rather than introducing new technologies after the court decision limited the patentability of new inventions. Overall, my results show that patents incentivize incumbents to explore new technologies, thus increasing overall growth through creative destruction.

My study opens a new perspective on how innovation, patentability, and growth are empirically connected. I contribute to multiple strains of literature with my novel identification and micro-based evidence for Schumpeterian growth with heterogeneous innovations. There are several limitations to this analysis, though: patent applications typically take two to three years to be granted and my sample only has five years of post-treatment observations. Thus, my analysis does not fully capture the long-run effect of *Alice*. *Alice* is also industry specific to software-oriented service sectors, and less than 3% of all patents are treated. Nevertheless, *Alice* is relevant for affected firms and significantly changes their innovation and growth trajectory. This raises new questions about how patentability might affect investment, financing, and labor demand. Which type of workers benefits from the shift toward own innovation? Are fixed assets substitutes or complements for patent-protected inventions? Does creative destruction raise funding costs? The purpose of this study is to set the framework for future studies to answer these questions and deepen our understanding of patentability, innovation, and growth.

References

Abadie, A., S. Athey, G. W. Imbens, and J. Wooldridge (2017, November). When Should You Adjust Standard Errors for Clustering? Working Paper 24003, National Bureau of Economic Research. Series: Working Paper Series.

Abrams, D. S., U. Akcigit, and J. Grennan (2013, November). Patent Value and Citations: Creative Destruction or Strategic Disruption? Working Paper 19647, National Bureau of Economic Research. Series: Working Paper Series.

Abrams, D. S., U. Akcigit, G. Oz, and J. G. Pearce (2019, March). The Patent Troll: Benign Middleman or Stick-Up Artist? Working Paper 25713, National Bureau of Economic Research. Series: Working Paper Series.

Aghion, P., U. Akcigit, and P. Howitt (2014, January). Chapter 1 - What Do We Learn From Schumpeterian Growth Theory? In P. Aghion and S. N. Durlauf (Eds.), *Handbook of Economic Growth*, Volume 2 of *Handbook of Economic Growth*, pp. 515–563. Elsevier.

Aghion, P., A. Bergeaud, T. Boppart, P. J. Klenow, and H. Li (2019, November). A Theory of Falling Growth and Rising Rents. Working Paper 26448, National Bureau of Economic Research.

Aghion, P., P. Howitt, and S. Prantl (2015, September). Patent rights, product market reforms, and innovation. *Journal of Economic Growth 20*(3), 223–262.

Akcigit, U. and S. T. Ates (2019, April). What Happened to U.S. Business Dynamism? Working Paper 25756, National Bureau of Economic Research. Series: Working Paper Series.

Akcigit, U. and W. R. Kerr (2018, March). Growth through Heterogeneous Innovations. *Journal of Political Economy 126*(4), 1374–1443. Publisher: The University of Chicago Press.

Angrist, J. and J.-S. Pischke (2009, January). *Mostly Harmless Econometrics* (1 edition). Princeton University Press.

Appel, I., J. Farre-Mensa, and E. Simintzi (2019, September). Patent trolls and startup employment. *Journal of Financial Economics 133*(3), 708–725.

Arrow, K. (1962, January). Economic Welfare and the Allocation of Resources for Invention. Working Paper c2144, National Bureau of Economic Research.

Arts, S. and L. Fleming (2018, December). Paradise of Novelty—Or Loss of Human Capital? Exploring New Fields and Inventive Output. *Organization Science 29*(6), 1074–1092.

Balsmeier, B., L. Fleming, and G. Manso (2017, March). Independent boards and innovation. *Journal of Financial Economics 123*(3), 536–557.

Bena, J., H. Ortiz-Molina, and E. Simintzi (2021, October). Shielding Firm Value: Employment Protection and Process Innovation. *Journal of Financial Economics*.

Bertrand, M., E. Duflo, and S. Mullainathan (2004, February). How Much Should We Trust Differences-In-Differences Estimates?*. *The Quarterly Journal of Economics 119*(1), 249–275.

Bessen, J. and R. M. Hunt (2007). An Empirical Look at Software Patents. *Journal of Economics & Management Strategy 16*(1), 157–189. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1530-9134.2007.00136.x.

Boldrin, M. and D. K. Levine (2013, February). The Case Against Patents. *Journal of Economic Perspectives 27*(1), 3–22.

Callaway, B. and P. H. C. Sant'Anna (2021, December). Difference-in-Differences with multiple time periods. *Journal of Econometrics 225*(2), 200–230.

Chien, C. V. (2016). Software Patents As A Currency, Not Tax, On Innovation. *Berkeley Technology Law Journal 31*(3), 1669–1723. Publisher: [University of California, Berkeley, University of California, Berkeley, School of Law].

Chien, C. V. and J. Y. Wu (2018, October). Decoding Patentable Subject Matter. SSRN Scholarly Paper ID 3267742, Social Science Research Network, Rochester, NY.

Choi, J. P. and H. Gerlach (2015). Patent pools, litigation, and innovation. *The RAND Journal of Economics 46*(3), 499–523.

Chow, M. C., T. C. Fort, C. Goetz, N. Goldschlag, J. Lawrence, E. R. Perlman, M. Stinson, and T. K. White (2021, May). Redesigning the Longitudinal Business Database. Working Paper w28839, National Bureau of Economic Research.

Cockburn, I. M. and M. J. MacGarvie (2009). Patents, Thickets and the Financing of Early-Stage Firms: Evidence from the Software Industry. *Journal of Economics & Management Strategy 18*(3), 729–773. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1530-9134.2009.00228.x.

Cockburn, I. M. and M. J. MacGarvie (2011, April). Entry and Patenting in the Software Industry. *Management Science 57*(5), 915–933. Publisher: INFORMS.

Cohen, L., U. G. Gurun, and S. D. Kominers (2019, September). Patent Trolls: Evidence from Targeted Firms. *Management Science*, mnsc.2018.3147.

Contigiani, A. (2020, April). Experimentation and Appropriability in Early-Stage Ventures: Evidence from the Software Industry. SSRN Scholarly Paper ID 3282261, Social Science Research Network, Rochester, NY.

Craig, J. A. (2017). Deconstructing Wonderland: Making Sense of Software Patents in a Post-Alice World Patent Law. *Berkeley Technology Law Journal 32*(Annual Review Issue), 359–378.

Daily, J. E. (2017). Alice's Aftermath: Changes in Patentee Behavior since Alice v. CLS Bank Bridging the Gap between the Federal Courts and the United States Patent & Trademark Office: The Journal of Science and Technology Law Symposium. *Boston University Journal of Science and Technology Law 23*(2), 284–303.

De Loecker, J., J. Eeckhout, and G. Unger (2020, May). The Rise of Market Power and the Macroeconomic Implications. *The Quarterly Journal of Economics 135*(2), 561–644.

De Loecker, J. and F. Warzynski (2012, October). Markups and Firm-Level Export Status. *American Economic Review 102*(6), 2437–2471.

DiNizo, A. M. J. (2018). From Alice to Bob: The Patent Eligibility of Blockchain in a post CLS Bank World. *Case Western Reserve Journal of Law, Technology and the Internet 9*, 1–28.

Dreisigmeyer, D., N. Goldschlag, M. Krylova, W. Ouyang, and E. Perlman (2018, July). Building a Better Bridge: Improving Patent Assignee-Firm Links. CES Technical Notes Series, Center for Economic Studies, U.S. Census Bureau.

Dugan, B. (2018). Mechanizing Alice: Automating the Subject Matter Eligibility Test of Alice v. CLS Bank. *University of Illinois Journal of Law, Technology & Policy 2018*(1), 33–80.

Forman, C. and A. Goldfarb (2020, June). Concentration and Agglomeration of IT Innovation and Entrepreneurship: Evidence from Patenting. Working Paper w27338, National Bureau of Economic Research.

Fort, T. and S. Klimek (2018, June). The Effects of Industry Classification Changes on US Employment Composition. Working Paper, U.S. Census Bureau, Center for Economic Studies.

Frésard, L., G. Hoberg, and G. M. Phillips (2020, July). Innovation Activities and Integration through Vertical Acquisitions. *The Review of Financial Studies 33*(7), 2937–2976.

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

Gallini, N. T. (1992). Patent Policy and Costly Imitation. *The RAND Journal of Economics 23*(1), 52–63. Publisher: [RAND Corporation, Wiley].

Gans, J. S. and S. Stern (2003, February). The product market and the market for "ideas": commercialization strategies for technology entrepreneurs. *Research Policy 32*(2), 333–350.

Garcia-Macia, D., C.-T. Hsieh, and P. J. Klenow (2019). How Destructive Is Innovation? *Econometrica 87*(5), 1507–1541. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA14930.

Gilbert, R. and C. Shapiro (1990). Optimal Patent Length and Breadth. *The RAND Journal of Economics 21*(1), 106–112. Publisher: [RAND Corporation, Wiley].

Goldschlag, N., T. J. Lybbert, and N. J. Zolas (2016, March). An 'Algorithmic Links with Probabilities' Crosswalk for USPC and CPC Patent Classifications with an Application Towards Industrial Technology Composition. SSRN Scholarly Paper ID 2749287, Social Science Research Network, Rochester, NY.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020, August). Bartik Instruments: What, When, Why, and How. *American Economic Review 110*(8), 2586–2624.

Graham, S. J. H., C. Grim, T. Islam, A. C. Marco, and J. Miranda (2018). Business Dynamics of Innovating Firms: Linking U.S. Patents with Administrative Data on Workers and Firms. SSRN Scholarly Paper ID 3224865, Social Science Research Network, Rochester, NY.

Graham, S. J. H., A. C. Marco, and R. Miller (2015, November). The USPTO Patent Examination Research Dataset: A Window on the Process of Patent Examination. SSRN Scholarly Paper ID 2702637, Social Science Research Network, Rochester, NY.

Gutiérrez, G. and T. Philippon (2016, December). Investment-less Growth: An Empirical Investigation. Working Paper w22897, National Bureau of Economic Research.

Hall, B. H., G. v. Graevenitz, and C. Helmers (2021, April). Technology entry in the presence of patent thickets. *Oxford Economic Papers 73*(2), 903–926.

Hall, B. H., A. Jaffe, and M. Trajtenberg (2005). Market Value and Patent Citations. *The RAND Journal of Economics 36*(1), 16–38. Publisher: [RAND Corporation, Wiley].

Hall, B. H. and M. MacGarvie (2010, September). The private value of software patents. *Research Policy 39*(7), 994–1009.

Hall, R. E. (2018, May). New Evidence on the Markup of Prices over Marginal Costs and the Role of Mega-Firms in the US Economy. Working Paper 24574, National Bureau of Economic Research.

Haltiwanger, J., R. Jarmin, R. Kulick, J. Miranda, V. Penciakova, and C. Tello-Trillo (2019, August). Firm-level Revenue Dataset. Working Paper 19-02, Center for Economic Studies, U.S. Census Bureau. Publication Title: CES Technical Notes Series.

Hansen, C. B. (2007, October). Generalized least squares inference in panel and multilevel models with serial correlation and fixed effects. *Journal of Econometrics 140*(2), 670–694.

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

Hoberg, G. and G. Phillips (2016, October). Text-Based Network Industries and Endogenous Product Differentiation. *Journal of Political Economy 124*(4), 1423–1465.

Jarmin, R. S. and J. Miranda (2002). The Longitudinal Business Database. *SSRN Electronic Journal*.

Ke, Z. T., B. T. Kelly, and D. Xiu (2019, July). Predicting Returns with Text Data. SSRN Scholarly Paper ID 3389884, Social Science Research Network, Rochester, NY.

Kelly, B., D. Papanikolaou, A. Seru, and M. Taddy (2021, September). Measuring Technological Innovation over the Long Run. *American Economic Review: Insights 3*(3), 303–320.

Kerr, W. R. and R. Nanda (2009). Democratizing entry: Banking deregulations, financing constraints, and entrepreneurship$. *Journal of Financial Economics 94*(1), 124–149.

Kesan, J. P. and R. Wang (2020, March). Eligible Subject Matter at the Patent Office: An Empirical Study of the Influence of Alice on Patent Examiners and Patent Applicants. SSRN Scholarly Paper ID 3556216, Social Science Research Network, Rochester, NY.

Klemperer, P. (1990). How Broad Should the Scope of Patent Protection Be? *The RAND Journal of Economics 21*(1), 113–130. Publisher: [RAND Corporation, Wiley].

Klenow, P. J. and H. Li (2021, May). Innovative Growth Accounting. *NBER Macroeconomics Annual 35*, 245–295. Publisher: The University of Chicago Press.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman (2017, May). Technological Innovation, Resource Allocation, and Growth*. *The Quarterly Journal of Economics 132*(2), 665–712.

Kong, N., U. Dulleck, A. B. Jaffe, S. Sun, and S. Vajjala (2020, September). Linguistic Metrics for Patent Disclosure: Evidence from University Versus Corporate Patents. Working Paper 27803, National Bureau of Economic Research. Series: Working Paper Series.

Kuhn, J., K. Younge, and A. Marco (2020). Patent citations reexamined. *The RAND Journal of Economics 51*(1), 109–132. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1756-2171.12307.

Kuhn, J. M. and N. C. Thompson (2019, January). How to Measure and Draw Causal Inferences with Patent Scope. *International Journal of the Economics of Business 26*(1), 5–38.

Lee, J.-S., N. Kim, and Z.-T. Bae (2019, December). The effects of patent litigation involving NPEs on firms' patent strategies. *Technological Forecasting and Social Change 149*, 119758.

Lemley, M. A. and R. Feldman (2016, May). Patent Licensing, Technology Transfer, and Innovation. *American Economic Review 106*(5), 188–192.

Lerner, J. and A. Seru (2021, July). The Use and Misuse of Patent Data: Issues for Finance and Beyond. *The Review of Financial Studies* (hhab084).

Lerner, J., A. Seru, N. Short, and Y. Sun (2021, July). Financial Innovation in the 21st Century: Evidence from U.S. Patents. Working Paper 28980, National Bureau of Economic Research. Series: Working Paper Series.

Loughran, T. and B. Mcdonald (2011). When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *The Journal of Finance 66*(1), 35–65.

Lu, Q., A. Myers, and S. Beliveau (2017, November). USPTO Patent Prosecution Research Data: Unlocking Office Action Traits. SSRN Scholarly Paper ID 3024621, Social Science Research Network, Rochester, NY.

Lybbert, T. J. and N. J. Zolas (2012, September). Getting Patents and Economic Data to Speak to Each Other: An 'Algorithmic Links with Probabilities' Approach for Joint Analyses of Patenting and Economic Activity. SSRN Scholarly Paper ID 2161982, Social Science Research Network, Rochester, NY.

Marx, M. and A. Fuegi (2019, November). Reliance on Science: Worldwide Front-Page Patent Citations to Scientific Articles. SSRN Scholarly Paper ID 3331686, Social Science Research Network, Rochester, NY.

Miller, S. P., A. Aravind, B. Bengfort, C. De La Cerda, M. Dragoni, K. Gibson, A. Itai, C. Johnson, D. Kannappan, E. Kehoe, H. Kim, K. Mladinich, R. Pinho, J. Polansky, and B. Weissenberg (2018). Who's Suing Us? Decoding Patent Plaintiffs since 2000 with the Stanford NPE Litigation Dataset. *Stanford Technology Law Review 21*(2), 235–275.

Moser, P. (2011, April). Innovation Without Patents - Evidence from the World Fairs. SSRN Scholarly Paper ID 930241, Social Science Research Network, Rochester, NY.

Noel, M. and M. Schankerman (2013). Strategic Patenting and Software Innovation. *The Journal of Industrial Economics 61*(3), 481–520.

Nordhaus, W. D. (1969). An Economic Theory of Technological Change. *The American Economic Review 59*(2), 18–28. Publisher: American Economic Association.

Peters, M. (2020). Heterogeneous Markups, Growth, and Endogenous Misallocation. *Econometrica 88*(5), 2037–2073. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA15565.

Peters, M. and F. Zilibotti (2021, October). Creative Destruction, Distance to Frontier, and Economic Development. Working Paper 29333, National Bureau of Economic Research. Series: Working Paper Series.

Ren, X. and R. Duprez (2019, May). Financial impact of legal practice on software patent subject matter eligibility: Effectiveness of a workable standard. *The Journal of High Technology Management Research 30*(1), 50–58.

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

Schumpeter, J. A. (1942). Capitalism, Socialism, and Democracy. SSRN Scholarly Paper ID 1496200, Social Science Research Network, Rochester, NY.

Stroud, J. and D. M. Kim (2017). Debugging Software Patents after Alice. *South Carolina Law Review 69*(1), 177–220.

Toole, A. A. and N. A. Pairolero (2020, April). Adjusting to Alice. Working Paper 3, United States Patent and Trademark Office.

Tran, J. L. (2015). Software Patents: A One-Year Review of Alice v. CLS Bank. *Journal of the Patent and Trademark Office Society 97*(3), 532–550.

Tran, J. L. (2016). Two Years after Alice v. CLS Bank. *Journal of the Patent and Trademark Office Society 98*(3), 354–373.

Tran, J. L. and J. S. Benevento (2019). Alice at Five. SSRN Scholarly Paper ID 3454922, Social Science Research Network, Rochester, NY.

Wagner, S. and I. Cockburn (2010, March). Patents and the survival of Internet-related IPOs. *Research Policy 39*(2), 214–228.

Webb, M., N. Short, N. Bloom, and J. Lerner (2018, July). Some Facts of High-Tech Patenting. Working Paper w24793, National Bureau of Economic Research.

Appendix A

Appendix: Data and Variable Definition

The following table summarized the definitions of variables on firm-level and shows a brief reference for the definition.[1]

Table A.1. Variable Definitions

Variable	Definitions	Source/Reference
Employment (Emp.)	Sum of total employment of all establishments of the firm in t.	LBD
Market Share (Emp.) (Mkt. Share)	Employment of the firm in t relative to the total employment of all firms in t in the same 4-digit NAICS industry.	LBD
Employment Growth	Change in Employment from t-1 to t over the average employment of the firm in t and t-1.	LBD
Establishments	Sum of all establishments of the firm in t. 'An establishment is a single physical location at which business is conducted or services or industrial operations are performed.'[2]	LBD
Establishments (Pos. Emp.) (Estabs.)	Sum of all establishments with at least one active employee of the firm in t.	LBD

[1] I refer to the current year as 't'.
[2] see https://www.census.gov/programs-surveys/susb/about/glossary.html

Firm Age	Difference between the observation year and the first year the firm is observed with positive employment	LBD, Chow et al. (2021)
Firm Death	Last year that the firm is observed with positive employment	LBD, Chow et al. (2021)
Firm Size	Firm total employment in current year for in scope establishments for BDS	LBD, Chow et al. (2021)
Entry	Sum of all establishments of firm where t-1 employment is zero and current year employment is >0	LBD, Chow et al. (2021)
Entry Rate	Sum of Entry relative to the average number of establishments of firm in t and t-1. The average number of establishments in t-1 is calculated by subtracting from the establishment count in t the number of entering establishments and adding the number of exiting establishments.	LBD, Chow et al. (2021)
Exit	Sum of all establishments of firm where employment in t-1 is >0 and employment in t is 0	LBD. Chow et al. (2021)
Exit Rate	Sum of Exit relative to the average number of establishments of firm in t and t-1. The average number of establishments in t-1 is calculated by subtracting from the establishment count in t the number of entering establishments and adding the number of exiting establishments.	LBD, Chow et al. (2021)
Entry Employment	Sum of t employment of entering establishments of firm.	LBD, Chow et al. (2021)
Exit Employment	Sum of t-1 employment of exiting establishments of firm.	LBD, Chow et al. (2021)
Job Creation (JC)	Sum of positive employment change of firm establishments from t-1 to 1.	LBD, Chow et al. (2021)
Job Creation Rate (JC Rate)	Job Creation relative to the average employment of the firm for t and t-1.	LBD, Chow et al. (2021)

Job Destruction (JD)	Sum of negative employment change of firm establishments from t-1 to t.	LBD, Chow et al. (2021)
Job Destruction Rate (JD Rate)	Job Destruction relative to the average employment of the firm for t and t-1.	LBD, Chow et al. (2021)
Wage Share (W/S)	Ratio of total payroll to nominal revenue of the firm in t.	LBD, Haltiwanger et al. (2019)
Markup	Inverse of Wage Share.	LBD, Haltiwanger et al. (2019)
Real Revenue (Real Rev.)	Total revenue of firm in t adjusted for inflation to the base year 2009.	Haltiwanger et al. (2019)
Real Revenue Growth	Change of Real Revenue from t-1 to t relative to the average Real Revenue between t-1 and t of the firm.	Haltiwanger et al. (2019)
Patent Portfolio (Patent PF)	Sum of all issued patents to firm between 2000 and 2009.	PatentsView
Treatment Share	Sum of *Alice* treated patents issued to firm between 2000 and 2009 relative to Patent Portfolio.	PatentsView
Patent Issue (Pat. Issue)	Sum of all patents issued to firm in t.	PatentsView
Forward Citations	Sum of total citations to patents issued to firm in t.	PatentsView, Balsmeier et al. (2017)
Top 1% Forward Citations (Top 1% Cited)	Sum of all patents issued to firm in t that are in the 99th percentile of the forward citation distribution of patents in the same filing year and CPC group.	PatentsView, Balsmeier et al. (2017)
Top 2% to 10% Forward Citations (Top 2-10% Cited)	Sum of all patents issued to firm in t that are between the 90th and 99th percentile of the forward citation distribution of patents in the same filing year and CPC group.	PatentsView, Balsmeier et al. (2017)
Cited Patents	Sum of all patents issued to firm in t that are below the 90th percentile of the forward citation distribution of patents in the same filing year and CPC group but received at least one forward citation.	PatentsView, Balsmeier et al. (2017)

Uncited Patents	sum of all patents issued to firm in t that have not received any forward citations.	PatentsView, Balsmeier et al. (2017)
Backward Citations (Bwd. Cit.)	Sum of citations by patents issued to firm in t.	PatentsView
Self-Citations (Self-Cit.)	Sum of citations by patents issued to firm in t to patent assigned to the same firm.	PatentsView, Balsmeier et al. (2017)
Claims	Sum of number of independent claims in patents issued to firm in t.	PatentsView, Balsmeier et al. (2017)
New CPC-Group (New CPC Grp.)	Sum of patents issue to firm in t in CPC-groups the firms has not patented in before at time of filing.	PatentsView, Balsmeier et al. (2017)
Known CPC-Group	Sum of patents issue to firm in t in CPC-groups the firms has patented in before at time of filing.	PatentsView, Balsmeier et al. (2017)
Technology Proximity	Average cosine-similarity of issued patents to prior patents of the firm in t. Cosine-similarity is calculated as the count vector of CPC-groups for a given filing year relative to the count vector of CPC-groups of all patents filed by the firm in the 10 years before.	PatentsView, Balsmeier et al. (2017)
New CPC-subgroup Combinations	Sum of total pairwise CPC-subgroup combinations by patents issued to firm in t that have not appeared before in the data.	PatentsView, Arts and Fleming (2018)
Normalized Forward Citations (Norm. Cites)	Sum of forward citations received by patents issued to firm in t relative to average number of forward citations received by patents in same filing year and CPC group.	Lerner and Seru (2021)
Process Patents	Sum of patents issued to firm in t that have all independent claims classified as process claim.	Bena et al. (2021)

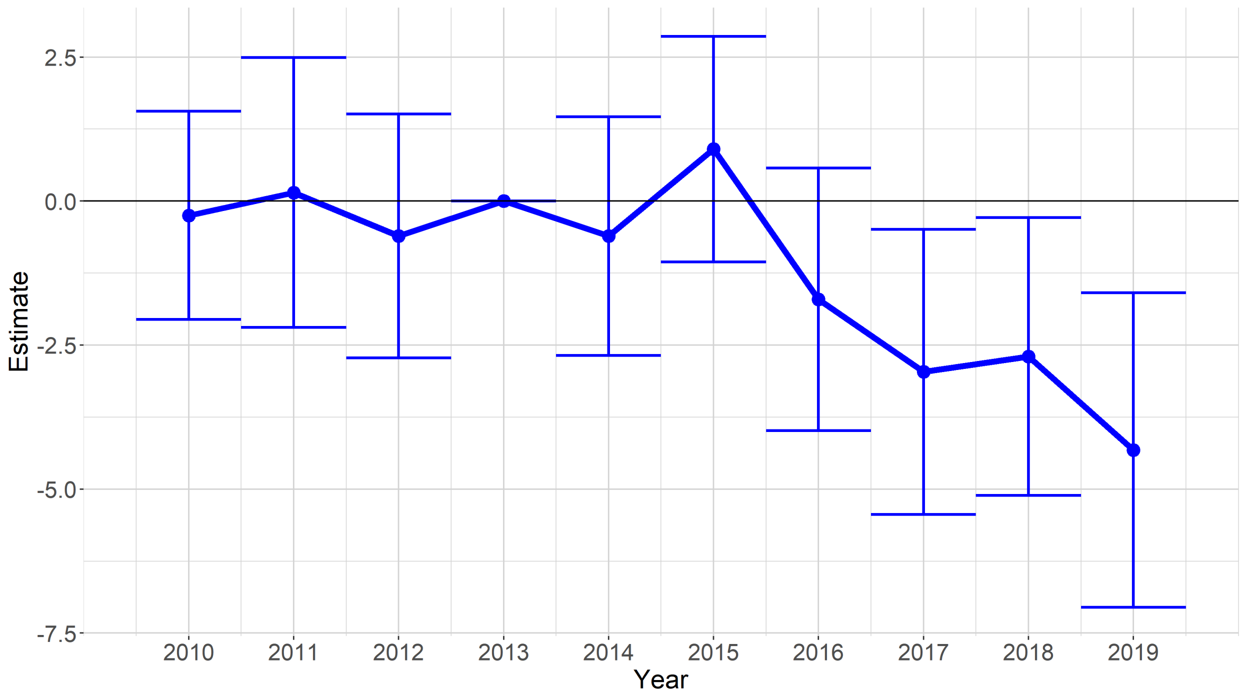
Non-process Patents	Sum of patents issued to firm in t that have no process claims.	Bena et al. (2021)
Average Words First Claim (Wd. Claim)	Average number of words in the first claim in patents issued to firm in t.	PatentsView, Kuhn and Thompson (2019)
Normalized Average Words First Claim (Norm. Wd. Claim)	Average Words First Claim in patents issued to firm in t normalized by the average for patents in the same filing year and CPC group.	PatentsView, Kuhn and Thompson (2019)
Follow-on Citations (Follow-on)	Sum of all forward citations received by patents issued to firm in t within three years after issuance.	PatentsView
Self-Follow-on Citations (Self-F/O)	Sum of all forward citations received by patents issued to firm in t within three years after issuance by other patents issued to the firm.	PatentsView

Appendix B

Appendix: Event Study Graphics

I present in this appendix the results for event study analysis using model 2.2 for different outcomes variables used in the main text. I summarize the results by plotting graphics for the year-treatment coefficients with collars for the 95%-interval. The models do not contain control variables, but their inclusion changes little for the overall results.

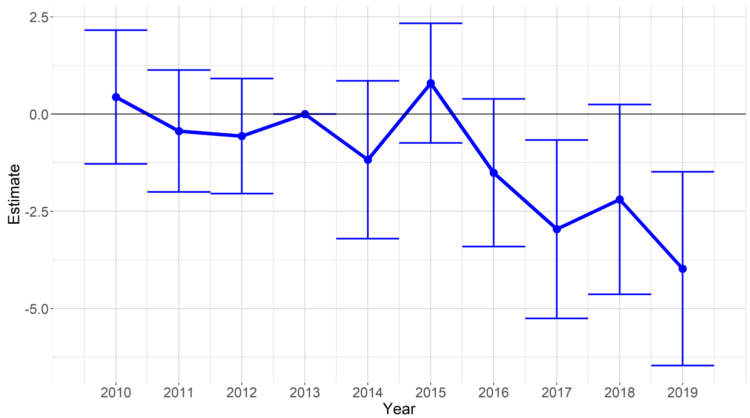
Fig. B.1. Event-study - Entry Emp - >90th Pctl. Pat. - >90th Pctl. Emp.



Event study for entry employment for the group of firms above the top 90th percentile of the patenting distribution and employment size distribution in their industry in 2009, respectively. The setting is defined as in table 3.17a, estimated with the event-study model in 2.2. Dependent variable is in log terms and includes '+1', defined in Appendix A. Coefficients for the interaction of the treatment variable with year dummies are shown with error collars for the 95% confidence interval.

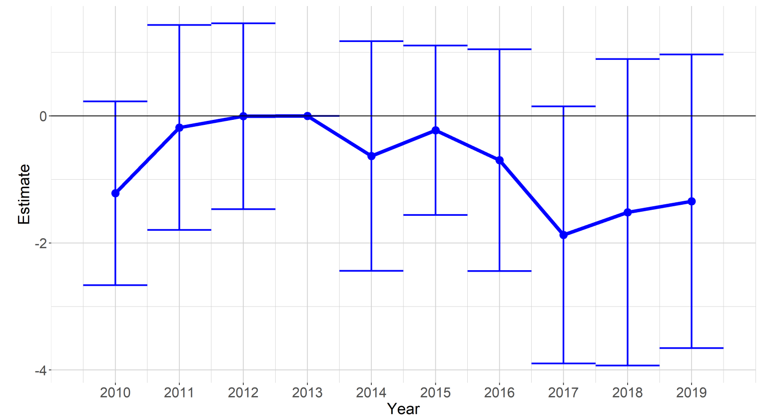
Fig. B.2. Event-study - Job Creation and Destruction - >90th Pctl. Pat. - >90th Pctl. Emp.

(a) Job Creation - >90th Pctl. Pat. - >90th Pctl. Emp.



Notes: Event study for job creation for the group of firms above the top 90th percentile of the patenting distribution and employment size distribution in their industry in 2009, respectively. The setting is defined as in table 3.17a, estimated with the event-study model in 2.2. Dependent variable is in log terms and includes '+1', defined in Appendix A. Coefficients for the interaction of the treatment variable with year dummies are shown with error collars for the 95% confidence interval.

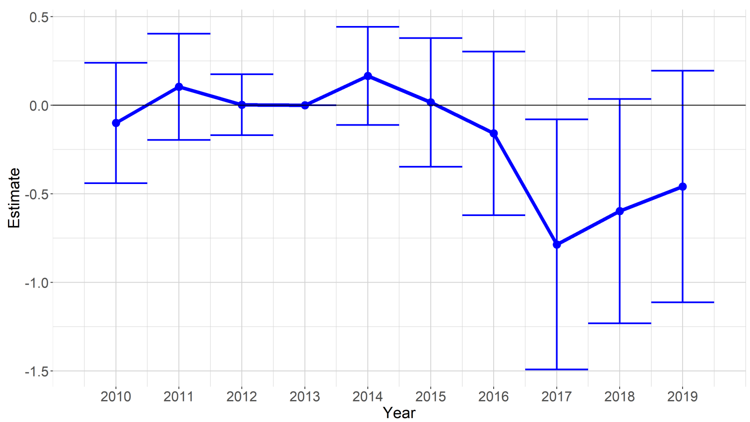
(b) Job Destruction - >90th Pctl. Pat. - >90th Pctl. Emp.



Notes: Event study for job destruction for the group of firms above the top 90th percentile of the patenting distribution and employment size distribution in their industry in 2009, respectively. The setting is defined as in table 3.17a, estimated with the event-study model in 2.2. Dependent variable is in log terms and includes '+1', defined in Appendix A. Coefficients for the interaction of the treatment variable with year dummies are shown with error collars for the 95% confidence interval.

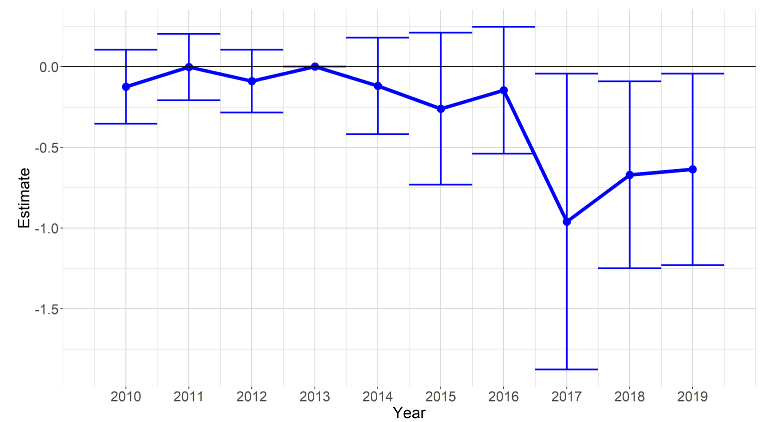
Fig. B.3. Event-study - Establishment and Employment - >90th Pctl. Pat. - >90th Pctl. Emp.

(a) Establishments (Pos. Emp.) - >90th Pctl. Pat. - >90th Pctl. Emp.



Notes: Event study for the total establishments with positive employment for the group of firms above the top 90th percentile of the patenting distribution and employment size distribution in their industry in 2009, respectively. The setting is defined as in table 3.17a, estimated with the event-study model in 2.2. Dependent variable is in log terms, defined in Appendix A. Coefficients for the interaction of the treatment variable with year dummies are shown with error collars for the 95% confidence interval.

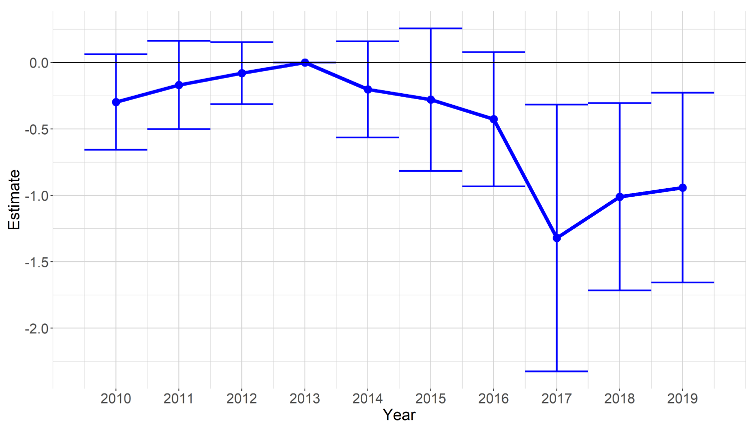
(b) Employment - >90th Pctl. Pat. - >90th Pctl. Emp.



Notes: Event study for total employment for the group of firms above the top 90th percentile of the patenting distribution and employment size distribution in their industry in 2009, respectively. The setting is defined as in table 3.17a, estimated with the event-study model in 2.2. Dependent variable is in log terms, defined in Appendix A. Coefficients for the interaction of the treatment variable with year dummies are shown with error collars for the 95% confidence interval.

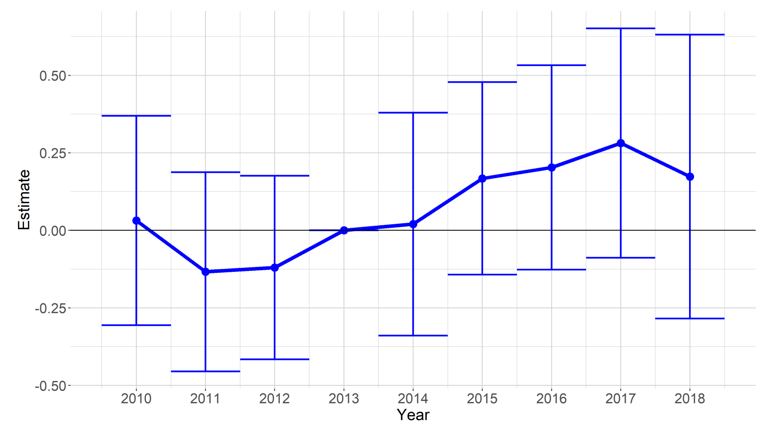
Fig. B.4. Event-study - Market Share and Markup - >90th Pctl. Pat. - >90th Pctl. Emp.

(a) Market Share (Emp.) - >90th Pctl. Pat. - >90th Pctl. Emp.



Notes: Event study for market share for the group of firms above the top 90th percentile of the patenting distribution and employment size distribution in their industry in 2009, respectively. The setting is defined as in table 3.17a, estimated with the event-study model in 2.2. Dependent variable is in log terms, defined in Appendix A. Coefficients for the interaction of the treatment variable with year dummies are shown with error collars for the 95% confidence interval.

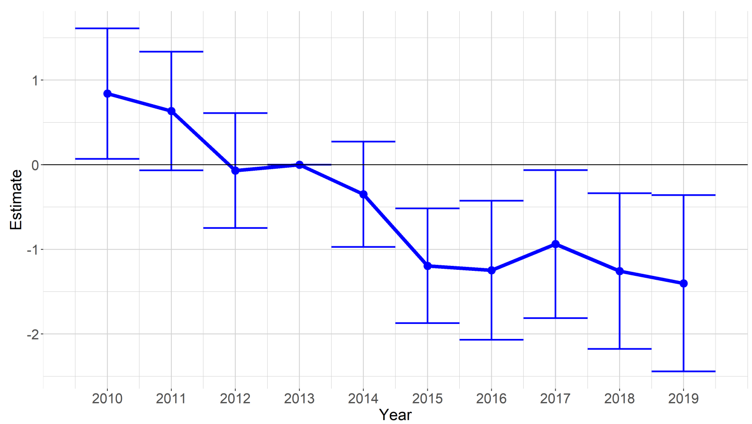
(b) Markup - >90th Pctl. Pat. - >90th Pctl. Emp.



Event study for markups for the group of firms above the top 90th percentile of the patenting distribution and employment size distribution in their industry in 2009, respectively. The setting is defined as in table 3.17a, estimated with the event-study model in 2.2. Dependent variable is in log terms, defined in Appendix A. Coefficients for the interaction of the treatment variable with year dummies are shown with error collars for the 95% confidence interval.

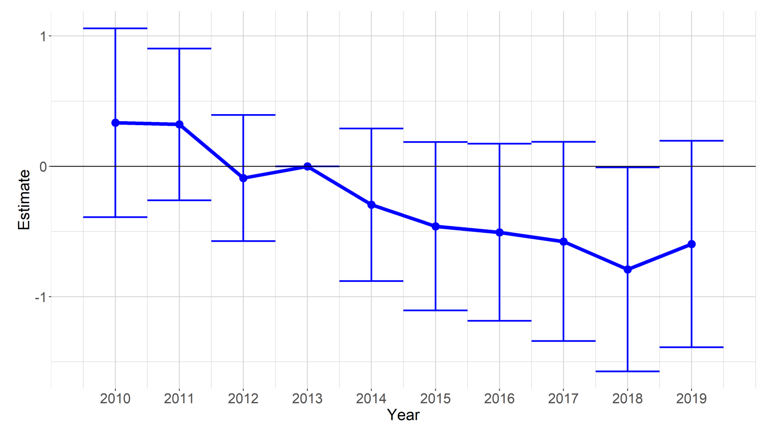
Fig. B.5. Event-study - Follow-on and Self-follow-on - >90th Pctl. Pat. - >90th Pctl. Emp.

(a) Total Follow-on - >90th Pctl. Pat. - >90th Pctl. Emp.



Event study for total follow-on citations for the group of firms above the top 90th percentile of the patenting distribution and employment size distribution in their industry in 2009, respectively. The setting is defined as in table 3.17a, estimated with the event-study model in 2.2. Dependent variable is in log terms and includes '+1', defined in Appendix A. Coefficients for the interaction of the treatment variable with year dummies are shown with error collars for the 95% confidence interval.

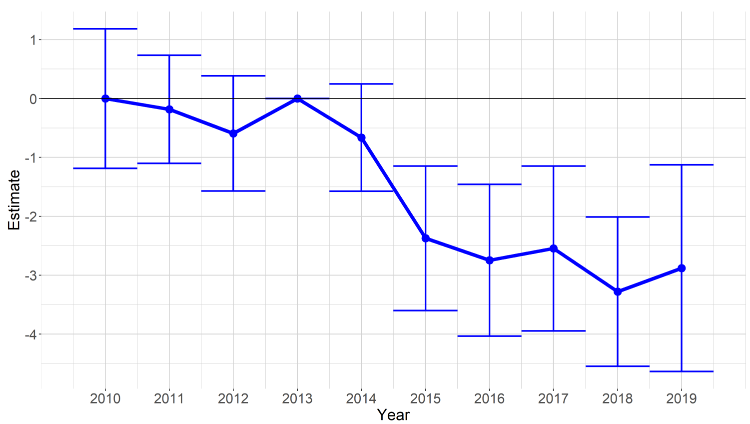
(b) Self-Follow-on - >90th Pctl. Pat. - >90th Pctl. Emp.



Event study for self-follow-on citations for the group of firms above the top 90th percentile of the patenting distribution and employment size distribution in their industry in 2009, respectively. The setting is defined as in table 3.17a, estimated with the event-study model in 2.2. Dependent variable is in log terms and includes '+1', defined in Appendix A. Coefficients for the interaction of the treatment variable with year dummies are shown with error collars for the 95% confidence interval.

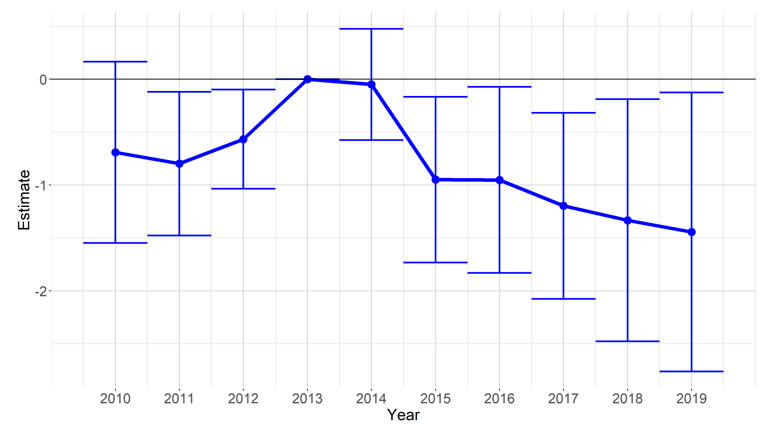
Fig. B.6. Event-study - Backward and Self-citations - >90th Pctl. Pat. - >90th Pctl. Emp.

(a) Backward Citations - >90th Pctl. Pat. - >90th Pctl. Emp.



Event study for backward citations for the group of firms above the top 90th percentile of the patenting distribution and employment size distribution in their industry in 2009, respectively. The setting is defined as in table 3.17a, estimated with the event-study model in 2.2. Dependent variable is in log terms and includes '+1', defined in Appendix A. Coefficients for the interaction of the treatment variable with year dummies are shown with error collars for the 95% confidence interval.

(b) Self-Citations - >90th Pctl. Pat. - >90th Pctl. Emp.



Event study for self-citations for the group of firms above the top 90th percentile of the patenting distribution and employment size distribution in their industry in 2009, respectively. The setting is defined as in table 3.17a, estimated with the event-study model in 2.2. Dependent variable is in log terms and includes '+1', defined in Appendix A. Coefficients for the interaction of the treatment variable with year dummies are shown with error collars for the 95% confidence interval.

Appendix C

Appendix: NLP Methodology

Since my research question deals with the impact of a Supreme Court decision on the validity of patent claims, I need to define the intellectual property portfolio held by firms and their exposure to the *Alice* decision. I first develop and train a NLP model to identify claim language that is invalid under *Alice*. Second, I classify existing patents and create annual patent portfolios using patent assignment data to firm. Finally, I use this portfolio in a difference-in-differences approach to quantify the impact of *Alice* on entry-exit, market shares, and innovation decisions.

C.1. Natural language processing approach for identification

Overview

I build a natural language processing (NLP) model to identify claim language that is affected by the *Alice* decision. For this, I use USPTO rejections in office actions mentioning the *Alice* decision as reason to find examples of claim text formulations that are not eligible. I select approved applications from the same filing years and patent classes as the rejected claims to train my binary classification model on the respective wording. Finally, I use this trained classification model on issued patents to predict their exposure to potential invalidation due to *Alice*. Figure C.1 illustrates the building and the classification process of patent claims to define a firm-wide *Alice* exposure score.

Data and training sets

As a first step, I need to identify text examples of claims that are affected by the *Alice* decision. For this, I use the office action rejection data provided by the USTPO, based on the work by Lu et al.

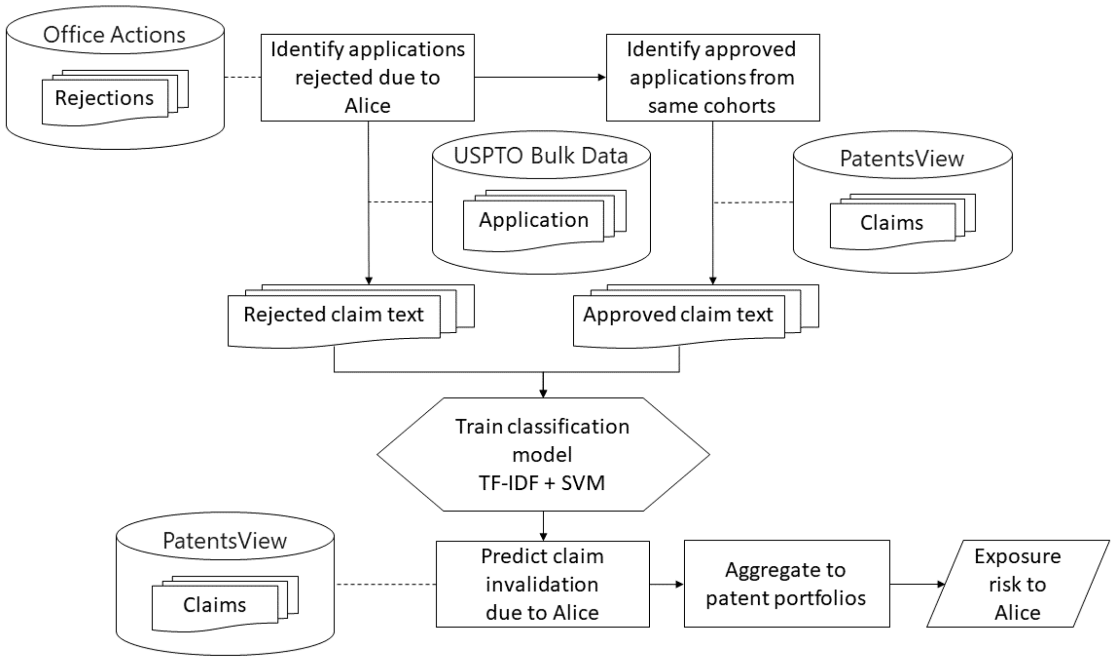


Fig. C.1. Process diagram patent claim classification

(2017).[1]

In this paper, the authors use natural language processing tools on "office actions" to systematically extract and classify the reasons for any rejections, objections, or requirements. Office actions are the written notifications to the applicants of the examiners' decision on patentability. Since these office actions include references that the applicant may find useful for responding to the examiner and deciding whether to continue prosecuting the application, Lu et al. (2017) can identify rejections due to the Supreme Court decision in *Alice Corp. v. CLS Bank International*. The authors restrict to published applications, thus I can use the Patent Examination Research Dataset (PatEx)[2] and published applications[3] to extract the respective independent claim texts and additional application details such as filing date, patent class, and application status.[4] I identify the four

[1]The data are sourced from the Office of the Chief Economist (OCE), URL: `https://www.uspto.gov/learning-and-resources/electronic-data-products/office-action-research-dataset-patents`.

[2]Available from USPTO, URL: `https://www.uspto.gov/learning-and-resources/electronic-data-products/patent-examination-research-dataset-public-pair`.

[3]Applications received since November 2000 are generally published within 18 month since the American Inventors Protection Act (AIPA) (Graham et al. (2015)). XML files with weekly pre-grant publications of applications are available under `https://bulkdata.uspto.gov/`.

[4]I can find for more than 90% of all applications with *Alice* rejections the respective claim publications. If ap-

USPC patent classes with the most rejections due to *Alice*, accounting for more than 91% of all *Alice*-based claim rejections in my sample.[5] Treated claims are those rejected claims that had no other issue identified in the office action and were either final rejections, were abandoned due to failure to respond to the office action, or were abandoned after the examiner's answer or board of appeals decision. In total, this renders 5,125 unique treated claim texts.

For the classification algorithm to be trained, I need to have a control group of claim texts that are valid under *Alice*. I select applications from the same USPC classes and the same filing years as the treated applications. I limit to applications that were granted without office action due to 35 U.S.C. §101, that is without issues referring to subject matter eligibility that could relate to the *Alice* decision.[6] I restrict to utility patents issued after the *Alice* decision in June 2014 and use the final claim text from PatentsView; I can be reasonably certain that this issued claim text was deemed eligible under *Alice* by the patent examiner. To balance my training sets, I randomly draw independent claims as controls, proportional to the distribution of USPC classes and filing years as in my treated claims set.

Figures C.2a and C.2b show word clouds of the most frequent words in the treated and control claim corpa, with more frequent words being larger. The clouds look quite similar, which is to be expected for claims from the same patent classes and the same cohorts. The terms are weighted by raw term frequencies without any pre-selection of terms, thus some of the most prominent terms in

plicants only file in the U.S. and don't seek protection in other national jurisdictions, however, they can request non-publication of an application, which might introduce selection bias into my observable treated claims. Overall, only 3.4% of all *Alice* treated applications do not have a pre-grant publication (making up around 36.9% of all missing observations). Furthermore, looking at the sets of applications for which I am not able to find the published pre-grant publication, I do not find evidence for any selection: the average time between filing and recorded pre-grant publication are for both groups very similar with on average 317 days for the publications I found versus 284 days for the missing application, likewise the likelihood of being ultimately granted are 57% and 50%, respectively. For the ultimately granted applications, the average time between filing and patent issue is statistically not different between the groups, with on average 1,519 days for the found applications and 1,474 days for the missing observations. This rejects the notion that applicants systematically defer publication of *Alice*-affected claims if the issues are more likely to be resolved. Overall, the applications to which I cannot link claim texts are very similar to the applications with treated claims in terms of USPC classes and filing years, suggesting that the missing claim publications are due to issues like typos and errors in the underlying XML files.

[5]Those classes are: 705 (Data processing: financial, business practice, management, or cost/price determination), 463 (Amusement devices: games), 702 (Data processing: measuring, calibrating, or testing), 434 (Education and demonstration); with the 705 class making up 82.4% of all rejected claims, followed by 9.9% for class 463, 4.6% for class 702, and 3.1% for class 434, each containing at least 100 claims.

[6]The paragraph reads: 'Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.' The *Alice* decision significantly broadened the scope of ineligible subject matters by holding that the financial settlement system in the disputed patent is ineligible as mere computer-based implementation of an abstract ideas.

(a) Word frequency Alice-affected claims

(b) Word frequency control claims

Fig. C.2. Word clouds for raw word frequencies

both groups are terms with little individual meaning such as ‘plurality’, ‘least’, ‘wherein’, reflecting the legal nature of the claim formulations.

Figures C.3a and C.3b show word clouds by weighting based on the relative frequency differences of terms, i.e. how more frequent is a word in one class compared to the other class. Words like ‘amount’, ‘financial’, ‘risk’, and ‘account’ are more prominent for *Alice*-affected claims, since the invalidated patents in the case were about an electronic escrow service for facilitating financial transactions. While the eligible claims contain more diverse words including specifications of physical implementations like ‘device’, ‘server’, ‘display’, *Alice*-affected claims prominently feature abstract descriptions about business processes such as ‘computer’, ‘determining’, ‘method’, and ‘program’. This is in line with the Supreme Court ruling stating that abstract claims are not

(a) Relative word frequency Alice-affected claims

(b) Relative word frequency control claims

Fig. C.3. Word clouds with differential frequency weighting

eligible just by introducing a computer into the claim.[7] The wording of the treated and control claims thus have distinct characteristics which are not merely explained by patent class or filing cohort.

[7]See the Syllabus to *Alice Corporation v. CLS Bank International*, Supreme Court, No. 13–298.: '"Simply appending conventional steps, specified at a high level of generality," to a method already "well known in the art" is not "enough" to supply the " 'inventive concept' " needed to make this transformation [from an abstract idea into a patent-eligible invention]. [...] . The introduction of a computer into the claims does not alter the analysis. Neither stating an abstract idea "while adding the words 'apply it,' "'

Model building and evaluation

I need to construct a binary classification mechanism and train it on the rejected claims with the eligible claims as control set to classify patent claims based on their exposure to the *Alice* decision. This is a typical natural language processing application in machine learning.[8]

Since most classification algorithms require numerical features as input, I need to vectorize claim texts into arrays. I split the text of the claims into separate words (tokens), remove all numbers, special characters, HTML tags, punctuation, and words shorter than three characters. I also remove stop words, i.e., common words that occur with high frequency but carry little substantive information like 'the', 'and', 'like'. Finally, I convert the resulting strings to lowercase and apply the Porter stemming algorithm, e.g., the words 'playing', 'played', and 'play' all have the same stem 'plai'.[9]

To turn the resulting arrays of tokens into a numeric vector representation, I count the occurrences of each token found in the corpus for each document. The collection of claims thus can be represented by a matrix with one row per claim and one column per token (e.g. 'word') or n-gram (i.e. a sequence of n tokens) occurring in the corpus. I use here uni- and bigrams, i.e., sequences of up to two tokens as features. A single claim thus is represented as a matrix row with the number of occurrences of a token / n-gram, i.e., the term-frequency, as the cell entry in the respective column. Since tokens that are very common among claims have little value in classifying them, I weight the term-frequency with the inverse document-frequency. This means the frequency of occurrences of a token / n-gram is multiplied by a factor decreasing with the frequency of the term in the overall corpus.[10] The text of treated and control claims thus is represented as numeric vectors, underweighting common terms, in a TF-IDF matrix, a Term Frequency–Inverse Document Frequency matrix.

Having a vector space representation of the treated and control claims in the TF-IDF matrix, I use a support vector machine to construct a hyper-plane to separate eligible from ineligible claims. I use a support vector machine with quadratic kernel to solve for non-linear classifications.[11] The full model thus takes as input the full-text of a claim and gives a predicted probability between 0 and 1 of whether this claim was affected by the *Alice* decision. I use 85% of the training claim data

[8] All calculations and executions were performed in Python 3.7.8 on a GNU/Linux operating system, release 3.10.0-1062.18.1.el7.x86_64

[9] All these pre-processing steps are performed using the 'gensim' library in Python 3 with the function 'preprocess_string'.

[10] The specific weighting term is $idf(t) = \ln \frac{1+n}{1+df(t)} + 1$ with n being the total number of claims in the training set and $df(t)$ the number of claims in the corpus containing the term t. The resulting vectors are also normalized to unit length.

[11] Using a linear or rbf kernel doesn't change the results. I use here the implementation for the TF-IDF matrix and the support vector classification (SVC) in sklearn.

for model fitting and the rest for evaluating the fit quality. Table C.1 reports the performance of my classification model for this test dataset, figures C.4c, C.4a, and C.4b give a graphical interpretation with the confusion matrix, precision-recall curve, and receiver operating characteristic (ROC) curve.

The confusion matrix plots for the test sample the relative share of correctly predicted labels in the diagonal from the upper left to the bottom right corner. Thus, the relative number of true negatives and true positives (since I define *Alice* rejections as treated) are in the diagonals. Related measures of these are precision, defined as the number of true positives over the number of true positives plus the number of false positives, and recall, defined as the number of true positives over the number of true positives plus the number of false negatives. I want to achieve a high precision, i.e., my model correctly classifies claims as affected, and a high recall, i.e., my model finds most instances of *Alice*-affected claims.

	precision	recall	f1-score	support
Valid Claims	0.940	0.915	0.927	696
Invalid Claims	0.917	0.941	0.929	696
macro avg	0.928	0.928	0.928	1392
weighted avg	0.928	0.928	0.928	1392
Precision Score	0.917			
Recall Score	0.941			
Accuracy Score	0.928			
F1 Score	0.929			
MCC	0.857			

Table C.1. Classification report of prediction model

There is a trade-off between these two measures, since higher thresholds for being classified as affected return fewer results, but most of the predicted labels are correct when compared to the training labels, i.e., a high precision. On the other hand, if the threshold is very low, the model returns many results, i.e., has a high recall, but most of its predicted labels are incorrect. The precision-recall curve plots this trade-off for different threshold levels, with a larger area under the curve indicating a better classification. The receiver operating characteristic (ROC) curve similarly plots the fraction of true positives out of the invalid claims against the fraction of false positives out of the valid claims at various threshold settings. The 45 degree line in this ROC curve plot would represent a random guessing model, while a good model would be in the upper left corner with a high true positive rate and a low false positive rate.

All the relevant performance measures for each class are shown in the classification report in table C.1. Macro and weighted averages of precision and recall are identical since I use a balanced training and test set. In the lower panel of the classification report, I also report the overall precision and recall scores, as well as the accuracy score, which is the average of correctly classified labels. A common measure combining precision and recall is the F1-score, defined as $F1 = 2\frac{precision*recall}{precision+recall}$. Furthermore, the Matthews correlation coefficient (MCC) is frequently used to measure the quality of binary classification models.

Overall, all measures show a very good classification model with high precision, high recall, and more than 85% MCC correlation, meaning a very high correlation between the true and predicted labels.

To further validate my model choice, I use different performance metrics in table C.2 to compare popular classification models on my training and test data. All models were used in their default features, nevertheless my approach using a support vector machines proves to be best among the listed classifiers.

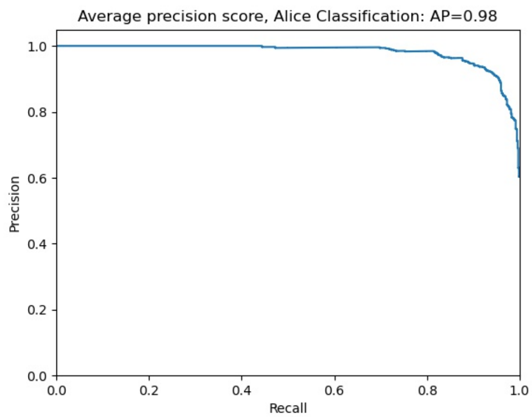
Model Name	Precision Score	Recall Score	F1 Score	Accuracy Score	MCC
Logistic Regression	0.813	0.829	0.821	0.819	0.638
Naive Bayes	0.802	0.819	0.810	0.808	0.617
Support Vector Machine	0.893	0.935	0.914	0.912	0.824
Decision Tree	0.746	0.777	0.761	0.756	0.513
Random Forest	0.876	0.905	0.890	0.889	0.778
K-nearest Neighbors	0.769	0.773	0.771	0.770	0.540
Stochastic Gradient Descent	0.852	0.869	0.861	0.859	0.719
AdaBoost	0.724	0.739	0.731	0.728	0.457
Gradient Boosting	0.767	0.792	0.779	0.776	0.552

Table C.2. Comparison of different classification models

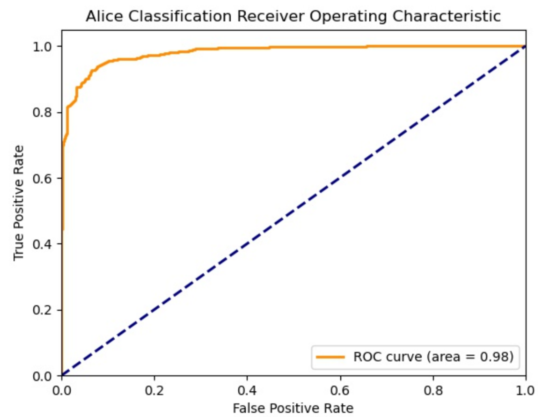
C.2. Patent claim classification

Using my trained classification model, I calculate the predicted probability of being affected by *Alice* for independent patent claims in the treated USPC classes. I do this for all patents issued since 1990 using the claims data from PatentsView.[12] This renders a total 393,100 classified claims in 115,630 unique patents. Since the USPTO moved to CPC classification in 2013, I repeat the

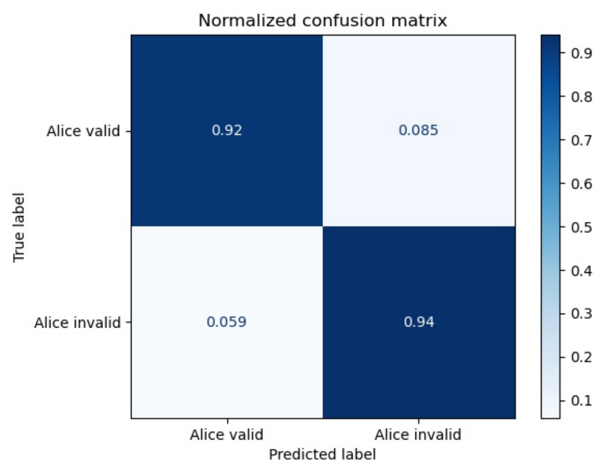
[12]URL: `https://www.patentsview.org/download/claims.html`.



(a) Precision-recall curve



(b) ROC curve



(c) Confusion matrix

Fig. C.4. Model performance graphics

classification process for the five CPC groups most consistent with my USPC classes.[13] Classifying patent claims based on CPC groups gives 3,236,420 classified independent claims in 983,229 patents. This is notably more than using the USPC classification, which can be mainly led back to the USPC classification at issue ending in 2015 in the PatentsView data and the affected CPC

[13]There is no direct mapping from the USPC classes into CPC, however the groups A63F (video games), G07F (coin-freed or like apparatus), G06F (digital data processing), H04L (transmission of digital information), and G06Q (data processing systems) represent around 50% of the affected patents in the USPC classes 705, 463, 434, and 702.

groups being much broader than the USPC classes.[14] This broader definition of treated patents classes under CPC can also be seen in figures C.5a, C.5b, C.5c, and C.5d where I generate word clouds from the claim texts that are predicted to be treated and untreated under *Alice* for the CPC and USPC classes: the CPC group clouds are in general more diverse, though the overall image is similar with words like ‘computer’, ‘means’, ‘value’, ‘device’, ‘signal’, and ‘display’ showing prominently in both class groups.[15]

(a) Relative word frequency of predicted Alice treated patents in main CPC groups

(b) Relative word frequency of predicted Alice treated patents in main USPC classes

(c) Relative word frequency of predicted untreated patents in main CPC groups

(d) Relative word frequency of predicted untreated patents in main USPC classes

Fig. C.5. Word clouds for classified patent claims

To understand better which words are most important for the classification decision, I use a Local Interpretable Model-agnostic Explanations (LIME) algorithm on 1000 randomly selected patent claims. The basic idea behind the LIME algorithm is to perturb the underlying text by leaving out words and measuring how much and in which direction this changes the classification. Doing this gives for each test text a list of words with the highest positive or negative impact on the classification. In table C.3, I show the 20 most frequent words associated with a classification as eligible and ineligible. The general picture from the word clouds of figures C.3a and C.3b is

[14]The four treated USPC classes account for a total of around 1.1% of class assignments, while the five corresponding CPC groups make up around 9.2% of all CPC assignments.

[15]I restrict to a sample of 1 million claims from the CPC groups to limit memory issues in the execution.

confirmed: more precise wording with references to the actual implementation of claims improves the eligibility. More abstract and general formulations of methods and processes render claims ineligible. This is consistent with the findings of Dugan (2018) and the literature stating that more specific formulations in claims are eligible under *Alice* compared to abstract ideas and processes (DiNizo (2018), Tran and Benevento (2019), Craig (2017)). Notably, the differences between the respective CPC and USPC classes are small, giving confidence in ability of the prediction algorithm to correctly predict *Alice*-affected patents in the larger CPC groups.

The most frequent word with impact on the classification are 'said' for ineligible texts and 'second' for eligible language. These words also show up in the respective word clouds in figures C.3a, C.3b, C.5b, C.5a, C.5d, and C.5c. 'Said' and 'second' are examples of sentence structures in eligible and ineligible claims that cannot be captures by, e.g., key word search: while 'said' frequently refers back to a broader subject at the beginning of a claim, 'second' is common in sentences about how several components of an invention are interacting, thus giving a more concrete description of the operations.

Since more patent granted since 2013 have a CPC classification rather than a USPC classification, I focus on the CPC based patent classification in the following. Finally, I define a patent as being treated if its first claim is predicted to be invalid under *Alice*. The first claim is the broadest claim of a patent and thus sets the scope for the technology protected by the patent (see, e.g., Kuhn and Thompson 2019). Alternative specifications of treatment are as follows: a patent is treated if the majority of its independent claims are predicted to be *Alice*-treated, the average (continuous) predicted treatment probability for all independent claims is above 50%, or at least one independent claim is predicted to be *Alice*-treated. The different treatment definitions on patent-level are highly correlated with correlation coefficient between 71.1% and 88.0%.[16] The only effective difference is in how many patents are predicted to be treated: of the 1,062,883 classified patents, 18.5% are predicted to be treated using the first claim. Using the average predicted treatment probability leads to the fewest patents being *Alice*-treated (14.4% of classified patents) and minimum one claim treated to the most treated patents (24.9% of classified patents), in any case the number of treated patents is large with more than 150,000. In practice, these variation in the definition of treated patents have little effect on the economic analysis.

Patent Litigation - Validation Analysis

I use the Stanford NPE Litigation Dataset, based on Miller et al. (2018), to assess if my classification mechanism can predict real world patent challenges. The Stanford NPE Litigation Dataset is the 'first ever publicly available database to track comprehensively how practicing entities, non-practicing entities, and patent assertion entities (PAEs) claim patent ownership rights in litigation.

[16]The main definition of the first claim treated has a correlation coefficient of 82.6% with the minimum one claim treated definition and 88.0% with most claims treated, thus the first claim is very predictive for the rest of the patent.

USPC classification		CPC classification	
Ineligible words	Eligible words	Ineligible words	Eligible words
said	second	said	second
means	device	method	device
method	data	means	data
value	signal	value	signal
steps	plurality	program	control
program	display	memory	plurality
player	game	comprising	network
function	control	steps	user
processor	signals	file	information
step	user	group	display
processing	sensor	module	request
condition	information	processing	signals
module	position	processor	application
transaction	server	code	content
comprising	image	level	server
reference	time	sequence	image
level	network	bit	circuit
number	electronic	bits	node
selected	application	input	address
test	circuit	logic	virtual

Table C.3. LIME important words for classification in 1000 claims sample

[...] [R]esearchers are tracking every lawsuit filed in U.S. district courts from 2000 to the present and identifying each patent plaintiff as either a practicing entity or as one of eleven types of NPEs.'[17] Since confounding legal changes around 2014/2015 have limited the ability for NPEs to file lawsuits (Appel, Farre-Mensa, and Simintzi 2019), I focus on lawsuits with product companies as plaintiff. Product companies manufacture products, sell products, or deliver services generally, or are IP enforcement subsidiaries of practicing entities. This category account for around 52.4% of all patent litigation cases in the database. Lawsuits by product companies are also less likely to be 'opportunistic' (Cohen, Gurun, and Kominers 2019) and the patents defended are related to actual products and used technologies. Thus, if I find that patents predicted to be treated by the NLP method are less likely to be litigated by product companies after 2014, I can confirm that my algorithm identifies patents that are less enforceable after *Alice*.

[17]URL: `https://law.stanford.edu/projects/stanford-npe-litigation-database/`

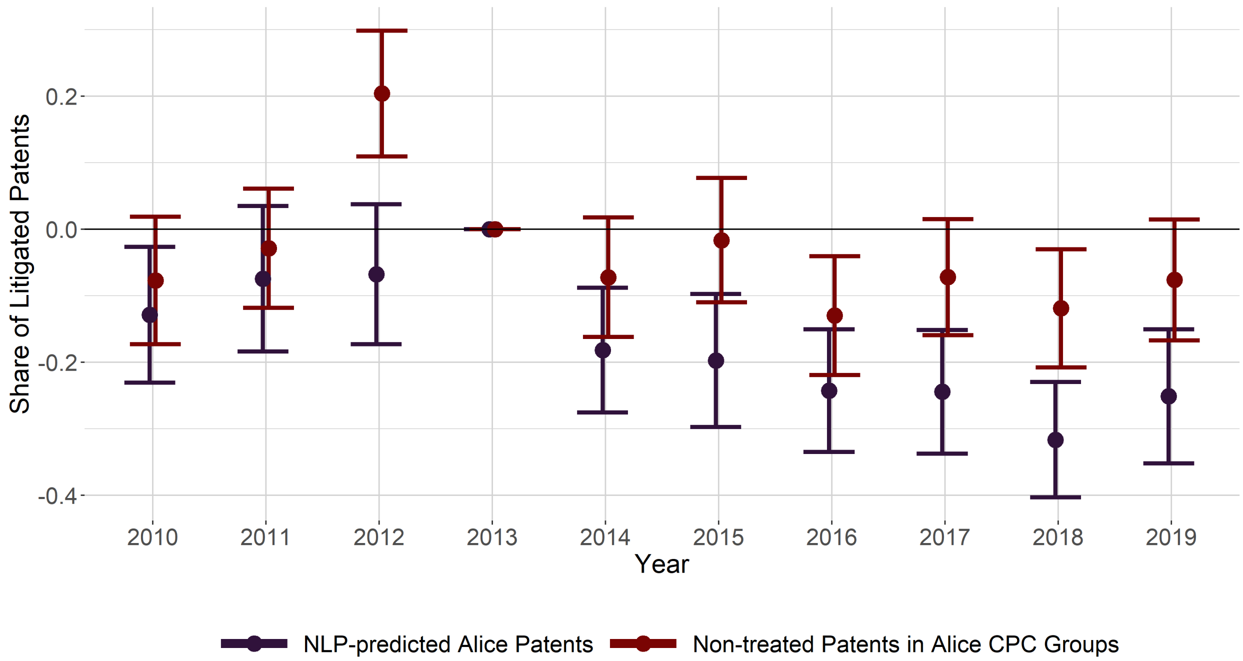


Fig. C.6. Share of treated and non-treated patents in product company litigation

In figure C.6, I plot year dummy regression coefficients with 95% confidence intervals for the share of treated patents among all litigated patents in the filing years and the share of patents that are in the CPC-groups that are classified but predicted to be not treated. The regression is restricted to lawsuits filed by product companies between 2000 and 2019 and the dependent variable is scaled to zero mean and unit variance. Coefficients are re-leveled with respect to 2013, immediately before *Alice*. While the share of non-treated patents in product company lawsuits remains largely unchanged, we see a clear downward shift after 2014 for the share of treated patents (there is one outlier for non-treated patent shares in 2012, which is not part of trend though). The magnitude of the shock is large; the share of treated patents before *Alice* was 4.82%, this falls by 2.33%, a drop of more than 48%. The share of non-treated patents was 16.9% and falls by 2.91%, a much smaller decrease of just 17.2%. Thus, my method can predict the enforceability of patents after *Alice*, even for patents filed within the same CPC-groups. The prediction method is not perfect, though, thus I still expect a negative treatment effect for patents that are predicted to be non-treated. However, the differences in magnitudes is large and my method more reliably predicts treatment than CPC-groups alone.[18]

[18]A more detailed investigation of whether my model can predict invalidation of patents in infringement cases is beyond the scope of this paper. To do this, for each patent, PACER data would need to be searched and analyzed for

Overall, my NLP algorithm is a reliable method for predicting the exposure of patents to *Alice*.

references to the *Alice* decision. It also could be that a higher exposure to *Alice* decreases the likelihood of suing for patent infringement, since the court could strike down the patent as ineligible.