Intellectual Property Protection Lost and Competition: An Examination Using Machine Learning

Utku U. Acikalin, Tolga Caskurlu, Gerard Hoberg and Gordon M. Phillips*

November 17, 2022

Abstract

We examine the impact of lost intellectual property protection on innovation, competition, acquisitions, lawsuits and employment agreements. We consider firms whose ability to protect intellectual property (IP) using patents is weakened following the Alice Corp. vs. CLS Bank International Supreme Court decision. This decision has impacted patents in multiple areas including business methods, software, and bioinformatics. We use state-of-the-art machine learning techniques to identify firms' existing patent portfolios' potential exposure to the Alice decision. While all affected firms decrease patenting post-Alice, we find an unequal impact of decreased patent protection. Large affected firms benefit as their sales and market valuations increase, and their exposure to lawsuits decreases. They also acquire fewer firms post-Alice. Small affected firms lose as they face increased competition, product-market encroachment, and lower profits and valuations. They increase R&D and have their employees sign more nondisclosure agreements.

Keywords: Patents, intellectual property protection, innovation, competition, litigation, Alice. **[JEL Codes: O31, O34, D43, F13]**

---

*TOBB University of Economics and Technology, University of Amsterdam, University of Southern California Marshall School of Business, and Tuck School of Business at Dartmouth College and National Bureau of Economic Research, respectively. Acikalin can be reached at u.acikalin@etu.edu.tr, Caskurlu can be reached at t.caskurlu@uva.nl, Hoberg can be reached at hoberg@marshall.usc.edu, and Phillips can be reached at gordon.m.phillips@tuck.dartmouth.edu. We thank Raj Aggarwal, Janet Gao, Josh Lerner, Alexander Ljungqvist, Tim Loughran, N.R. Prabhala, Carl Shapiro, Andrei Shleifer, Richard Thakor and seminar participants at Alberta, Amsterdam, Columbia, Harvard Business School, European Finance Association conference, Johns Hopkins, Minnesota, NBER Innovation and Big Data conferences, Notre Dame, Peking University, UCLA and Yeshiva University for helpful comments and discussions. All errors are the authors' alone.

What though the field be lost? All is not lost

Paradise Lost, John Milton, 1674.

1 Introduction

Intellectual property protection is at the core of innovation and competition policy. Economic and legal scholars have debated extensively whether intellectual property (IP) protection increases the incentives of firms to innovate and conduct R&D. The general consensus by many economists has been that patents stifle innovation as Boldrin and Levine (2013) describe in their survey article. Galasso and Schankerman (2015) reinforce this view by documenting a positive impact on small firm innovation following patent invalidation of patents by large patentees. Examining 60 countries over 150 years, Lerner (2002) also finds limited benefits of increasing patent protection. He finds decreased domestic patenting following increased IP protection but increases in foreign patenting, suggesting foreign competitors enter with the increased protection. Yet not all studies agree that IP protection is harmful to innovation. Budish et al. (2015) models how the length of patent protection should optimally increase for long-term costly innovation when commercialization occurs later, otherwise companies may not have enough incentives to innovate.

Thus, a natural question is how strong to make IP protection? The theories behind optimal IP protection begin with Nordhaus (1969). In that study, the debate is about the trade-off between giving patents to encourage innovation and the cost of reducing subsequent competition resulting from giving the patentee a local monopoly over the life of the patent. There are also issues with the scope of the patent. If patent protection is too broad, new entrants and new innovation may be discouraged as the protected scope of existing innovation might imply high entry barriers. Monopoly profits that arise from IP protection would also be high, harming consumers. If too weak, then firms would be discouraged from engaging in costly innovation as the fruits of that innovation would be potentially available to all to copy without incurring the costs of discovery.

Our study examines the consequences of weakened IP protection across multiple categories in a setting that shocked both existing patents and also incentives for *future* innovation and patenting in the U.S. in multiple patent categories. We examine firms whose patents are exogenously weakened by the Alice Corp v. CLS Bank International, 573 U.S. 208 (2014) Supreme Court decision (Alice, henceforth). This decision revoked patent eligibility in multiple patent areas. We examine the impact of lost intellectual property protection on a wide array of future firm decisions including firm innovation, competitive entry, acquisitions, lawsuits, patent trolls, and secrecy via non-disclosure agreements.

The Alice decision revoked patent protection on business methods patents whose fundamental idea is considered abstract with a transformation that is not novel. As part of this decision, the Supreme Court also ruled that the media and systems claims are similar to the business methods claims, and they are also patent ineligible. Thus, the Alice decision impacted multiple industries with patenting including data processing, software, finance, games, and measuring or testing in microbiology and enzymology. The outcome of this decision was in doubt given prior court decisions, and we show that it had a large impact after the ruling. In the next section, we summarize the extensive lower court disagreements on this case preceding the final Supreme Court ruling.[1]

We document that Alice has had a large impact on patent rejections, and it led to further decreases in patenting in exposed areas in the years after 2014 in the above-mentioned industries. Even post-Alice, there is considerable uncertainty about whether a particular patent sufficiently transforms an abstract idea enough to make it patent-eligible. Rejections based on Alice represented approximately 20% of the patent subject matter rejections overall in 2015 and 2016. For example, in the commerce and data processing methods industry,

[1] Indeed when the case was being considered at the Supreme Court, there were extensive Amicus briefs filed on both sides. Amicus briefs filed with the Supreme Court in support of CLS Bank included briefs filed by Google, Amazon, Dell, LinkedIn, Verizon, Microsoft, Checkpoint Software, the Software Freedom Law Center and Opensource Initiative, and prominent lawyers and economists. There were over 20 Amicus briefs in support of Alice Corp., including Advanced Biological Laboratories, IBM, Trading Technologies International, Inc., and prominent lawyers and economists. See http://www.alicecorp.com/fs_patents.html

36.2% of patents filed in 2013 were rejected citing Alice. We also conservatively estimate Alice's impact on future patents and find that Alice resulted in at least 3,237 fewer patents being granted per year after Alice.

While the decision had a large ex post impact on patenting, there was and still exists uncertainty about whether an existing or proposed patent transforms an idea sufficiently to be granted patent protection. Given the uncertain impact on each patent, we apply novel machine learning techniques on regulatory and patent textual corpora to assess how much a given firm's patent portfolio is exposed to Alice. Many legal scholars have written about the Alice decision and the difficulties of measuring and deciding whether there is sufficient transformation of an abstract idea to warrant a patent.[2]

We examine all patents in Alice impacted areas that were granted by 2014 (the date of the Alice decision). Some of these patents are likely to fail to provide protection if challenged in a court in the post-Alice period. This is a challenging computational task as there are more than 3.8 million patents granted between 1994 and 2014. We thus focus on the patents with the same primary Cooperative Patent Classification (CPC) as those rejected by the United States Patent and Trademark Office (USPTO) per Supreme Court's Alice criteria. Given the uncertainty about whether a given patent will be rejected, we use machine learning to gauge each patent's textual semantic similarity to patents previously rejected under Alice.

We use a deep learning-based language model called Longformer to predict the likelihood that each of the pre-Alice granted patents in the sample may become weakened and patent-ineligible due to the Alice decision. Longformer is a transformer model that is an improvement for long texts on the BERT model, which was released by Google in 2019 and achieves state-of-the-art performance on various Natural Language Processing (NLP) tasks (Devlin et al. (2019)). The model is also used in Google search queries, and Google argues that transformer models help Google Search better understand one in ten searches in the

[2] See Kesan and Wang (2020) and Lim (2020) for an extensive discussion of these debates and issues. These difficulties and the impact of Alice gave rise to U.S. Senate subcommittee hearings promoted in 2019 on potential revisions to strengthen intellectual property law in the "Stronger Patents Act."

U.S. in English.[3] The breakthrough innovation of models like Longformer and BERT is that they process words in relation to all the other words in a sentence, rather than one-by-one in order or in a fixed-sized sliding window approach.[4] These models can examine the full context of a word by looking at the words that come before and after it.

We find a large impact of Alice on future patenting and innovation. We verify that ex post patenting by firms whose patent stock is exposed to Alice significantly decreases for all impacted firms. We then split firms by size (and also by market share) as we hypothesize that smaller firms may be hurt more by the lost patent protection as they have fewer resources (managerial, financial and organizational) to defend their product spaces, while large firms are leading firms with more resources who can defend their product areas.[5] When we examine R&D, we find no change for large firms but find a significant increase in R&D for small firms. These results are consistent with small firms' attempting to replenish their innovative portfolio to escape competition from lost IP protection and to rebuild product differentiation. Examining ex-post changes in sales growth and profitability along with firm value, we find unequal impact. Large firms gain and small firms lose. Exposed large firms increase sales and experience insignificant gains in valuation. Small firms whose patent portfolio is exposed to Alice experience a decrease in operating margins and their market valuations.

It is perhaps not surprising that small firms lose when they experience losses in IP protection. Farre-Mensa et al. (2020) show that small firms gain from patent protection beyond the value of the idea itself using an instrument of random assignment of patent examiners from Sampat and Williams (2019). They show that small firms gain access to increased funding post-patent. Previous research by Galasso and Schankerman (2015) also showed that when larger firms' patents were invalidated, small firms increased innovation. Our setting is different, and we examine who gains and loses after a change in intellectual

[3] https://blog.google/products/search/search-language-understanding-bert/

[4] We use the Longformer model as it outperforms the BERT model in our context. Results using SciBERT, a version of BERT pretrained on a scientific corpus, were in an earlier version and are available from the authors. Our conclusions using Longformer are fully robust using the SciBERT model.

[5] In the Internet Appendix, we provide results from classifying firms based on their market shares, and thus by sales shares instead of assets, and the results are qualitatively similar.

property protection that impacts whole areas of technology indefinitely for both large and small firms. This differs from the prior focus on individual firms losing patent protection.

We show that differential losses for small firms are related to changes in competition relating to decreased IP protection. These small firms face increased competition on a number of different measures. Both large and small firms face increased venture capital financed entry into their product space, but this entry is more severe for small firms. Small firms face increased product similarity with existing competitors, and they complain more about increased competition. We also find using textual analysis of 10-Ks that small firms resort to non-disclosure agreements as an alternative way to protect IP post-Alice. Thus, small firms turn to increased secrecy to defend new IP in the face of lost patent IP protection. This shows that disclosure is important, which was also noted by Sampat and Williams (2019) in the case of technologies that shift from patentable to unpatentable.

We next examine patent infringement and intellectual property risk. We find that large firms experience fewer claims that they infringe on other firms. This decrease is significant for both patent-troll lawsuits and lawsuits by operating companies. This is intuitive as firms would be less likely to sue a deep-pocket firm when the validity of the patents is questionable. However, we find only weak results with the opposite sign for small firms. Our results are consistent with losses in IP protection enabling large firms to increase product market power at the cost of established smaller firms in their markets.

Regarding acquisitions post-Alice, we find that large firms decrease their acquisition activity. This is consistent with the theoretical arguments and empirical evidence in Phillips and Zhdanov (2013). They model how high market share firms may buy small firms after they have successfully patented an innovation. Large firms buy smaller firms to access their technology to then apply it to their larger customer base. Without a patent, there is less reason for a large firm to buy a smaller firm. If a large firm can forecast that a small firm's patent might lose protection post-Alice, there is less incentive to buy small firms for their technology as they can implement it for free without infringing the smaller firm's patents.

We document the impact of lost IP protection for firms in an entire area and examine future firm performance, litigation, competition, secrecy, and acquisitions. Empirically we show how and why small firms lose more from lost intellectual property protection. Small firms lose as they face increased competition. They increase R&D and increase secrecy as they use more nondisclosure agreements with their employees. In contrast, large firms benefit from area-wide patent ineligibility as their sales increase while lawsuits against them and costly acquisitions decrease. These results are consistent with large firms having more resources - technological, financial and managerial - which allows them to protect their product market position. The results are also consistent with the Schumpeterian effect dominating, with increased innovation after the shock being preformed by laggard small firms with low profits as Acemoglu et al. (2010) note. We thus conclude that patent protection is particularly important for small firms competing with larger firms.

Our paper also contributes methodologically by applying machine learning techniques to a difficult legal environment where the impact of Supreme Court decisions on individual firms is not known until after a patent is litigated. Our paper points to the benefits of increased competition and fewer lawsuits from reduced patent protection but also the costs for existing small firms who most directly face the impact of increased competition from large firms and new entrants. Our results thus show both costs and benefits of decreased IP protection.

2 Innovation and Alice *v.* CLS Bank International

There is a substantial debate on how strong to make IP protection. The general academic consensus is summarized by Boldrin and Levine (2013), who state that there is no empirical evidence that patents increase innovation and productivity. They advocate for abolishing patents entirely and using other legislative instruments to increase innovation. Galasso and Schankerman (2015) document that patent invalidation of large patentees triggers follow on innovation by smaller firms. However, these were exogenous invalidations of particular existing patents and these tests are not about indefinite forward-looking changes to entire patent areas as is the case for Alice. Lerner (2002)'s comprehensive study of over 60 countries used patent law changes and showed some benefits of strengthening patent protection for countries with initially weaker patent protection. Over time, however, domestic innovation declines with increases in IP protection while foreign patenting goes up. Frequently, however, such expansions of IP protection have been enacted simultaneously with relaxations of trade protections.[6] There is also evidence (see Budish et al. (2015) for example using cancer clinical trials) that maintaining incentives to engage in innovation is important if the ideas take a long time and are costly to develop.

We examine firm outcomes and competition after the landmark Supreme Court case, Alice Corp *v.* CLS Bank International, 573 U.S. 208 (2014). This decision impacted large industry areas - and key for us, not just a subset of an area. These areas previously had substantial patenting activity. Kesan and Wang (2020) review the impact of Alice and document large decreases in 11 patent categories including bioinformatics, business methods, business methods of finance, business methods of e-commerce, software (in general), databases and file management, cryptography and security, telemetry and code generation, digital cameras, computer networks, and digital and optical communications. They showed significant rejections of patents under Alice based on whether the proposed invention sufficiently transforms an abstract idea or natural law. Section 101 of the Patent Statute specifies four categories of the invention that are patent eligible: process, machine, manufacture, and composition of matter. However, there are, three court made exclusions to these categories that limit patent-eligibility: laws of nature, natural phenomena, and abstract ideas.

2.1 Legal Background of the Alice Case

In 2014, the Supreme Court of the United States decided a landmark case, Alice Corp *v.* CLS Bank International, 573 U.S. 208 (2014). It had a major effect on patent eligi-

[6] Lerner uses an indicator for whether the change took place in the aftermath of the Paris Convention of 1883 or the TRIPs agreement of 1993 to control for endogeneity.

bility across multiple patent categories. The issue was whether certain patent claims for a computer-implemented scheme encompass abstract ideas, making the claims ineligible for patent protection. The Supreme Court decided that known ideas are abstract, and discussing the computer implementation of a known idea does not make it patentable.

The result of the case was quite uncertain, and it caused a debate among the judges. After a district court held the patents invalid, the case reached to the Court of Appeals for the Federal Circuit (CAFC). In this court, a randomly assigned three-judge panel could not unanimously decide on the case, and the panel reversed the district court decision with a majority opinion.[7] However, given the case's complexity and its importance for the whole industry, the CAFC vacated the panel's opinion and decided to the hear the case in a full session of all ten judges that then heard the case.[8][9]

The uncertainty in the en banc session was not any less than the one in the three-judge panel. Five of the ten judges upheld the district court's decision that Alice's systems claims were not patent-eligible, and five judges disagreed. Seven of the ten judges upheld the district court's decision that Alice's method claims were not patent-eligible. However, these seven judges reached their opinions for different reasons. Overall, the judges could not agree on a single standard to determine whether a computer-implemented invention is a patent-ineligible abstract idea.

After the deep division in the CAFC, the Supreme Court of the US granted certiorari and affirmed the en banc decision of the Federal Circuit Court of Appeals.[10] The Court held a two-step framework for determining the patent eligibility of applications that would be applied to claims of abstract ideas. The Court decided that the claims in Alice patents cover an abstract idea and the proposed method claims fail to transform the abstract idea into a patent-eligible invention. The Court also ruled that the media and systems claims are similar to the methods claim and that they are also patent ineligible.

[7]CLS Bank Int'l v. Alice Corp. Pty. Ltd., 685 F.3d 1341, 1356 (Fed. Cir. 2012)
[8]CLS Bank Int'l v. Alice Corp. Pty. Ltd., 484 Fed. Appx. 559 (Fed. Cir. 2012)
[9]CLS Bank Int'l v. Alice Corp. Pty. Ltd., 717 F.3d 1269, 1273 (2013).
[10]Alice Corp. Pty. Ltd. v. CLS Bank Int'l, 134 S. Ct. 2347, 2354 (2014).

The Alice decision had a significant stock market reaction. We compute abnormal returns at the time of the Alice decision by subtracting the equally weighted CRSP market return. We find a significant negative coefficient at the 1% level for the -1 to +1 event-day window surrounding the Alice decision for the most impacted firms. There is also substantial variation, as at the average of our treatment variable, the excess returns were close to zero at -0.1%. Yet for the top five percent of our treatment variable, this excess return is larger at -0.8%. Thus, while the Alice decision had a small stock market impact for most firms, it had a large impact for some firms.

2.2 Consequences of Lost IP protection

The Alice case had a large impact on ex post patenting. The process to eventually reject a patent first starts with a petition by a litigant or an office action that is filed by a USPTO examiner. In Table 1, we present statistics for the top 12 industries with patent applications that were rejected by USPTO patent examiners citing Alice as the reason for reject for patents applied for prior to the Alice decision. Over 33,700 distinct patent applications made prior to Alice have been rejected in the 3 years post-Alice by examiners citing the Alice precedent. These rejected patents cover 5,831 distinct CPC Subgroups (out of 126,540 total), 919 Groups, 283 Classes, and 8 CPC Sections.

This table reports annual statistics from USPTO patent application rejections based on the Supreme Court's Alice decision for the top 12 industries based on Alice rejections. We present the number of patent applications from 2008 to 2017, with the percentage of rejections in parentheses for these industries. We use rejection data provided by Lu et al. (2017) that extends until 2016; therefore the ratio of rejection is assigned NA for 2017. *Change* reports the percentage change from the number of patent applications in 2013 to the average number of patent applications for the 2015-2017 period. Corresponding CPCs for each industry are provided in Table 2.

Table 2 provides a description of the main CPC groups impacted by Alice. In Panel B, we provide the industry that contains these Alice impacted CPC groups. Kesan and Wang (2020) document that about 17.9% of office action final decisions were rejected based on section 101 for other reasons before Alice was decided. This rate increased to 72.4% of the rejections of applications filed before Alice but decided afterwards, and 72.8% of applications filed after Alice. Other categories including computer networks, GUI, document processing, and cryptography and security also had significant increases in section 101 rejections after Alice. The number of patent applications per month dropped significantly post-Alice by 12-31% across different categories. For example, patent applications in the business method area dropped 29.5%. Kesan and Wang (2020) show using a difference-in-difference regression that section 101 Alice rejections increased significantly in 11 different patent categories.

While Alice had a large impact on patenting, the Supreme Court left substantial ambiguity about whether an individual patent transformed abstract ideas sufficiently to make them patent-eligible. As legal scholars have noted, the court did not define "abstract" and the court did not define how to decide whether an abstract idea has been transformed sufficiently into an inventive concept by including additional limitations to the patent claim, the basis for rendering claims eligible for patent protection. Given this uncertainty about whether a given patent would be rejected because of Alice, we use a deep learning based language model called Longformer to predict the likelihood that each of the pre-Alice granted patents in our sample might become ineligible due to the Alice decision. We then study ex post firm decisions and outcomes based on this predicted likelihood.

We split firms by large and small size as we hypothesize that smaller firms might be hurt more by lost patent protection as they have fewer resources (managerial, financial and organizational) to defend their product spaces while large firms will be the leading firms with more resources who can defend their product area. The differential impact on innovation by leading vs. laggard firms has been modeled for example by Aghion et al. (2005). In our setting, given the large increase in competition post-Alice, we expect firms left behind will innovate more in order to escape competition from other small firms. Larger firms are predicted to behave more like the Schumpeterian model and will innovate less as most innovation will be preformed by laggard firms with low initial profits or by smaller firms looking to break into the market.

We thus examine firm R&D and performance outcomes including changes in sales, operating income, and market valuations and the impact on competition overall between firms. While we could conjecture that the impact of the loss of IP protection may be negative for affected firms, such an unconditional prediction is not clear given predicted differences in impact for firms with different abilities and innovative resources. We thus focus on testing predictions separately for large and small firms. Large firms might benefit from losses in IP protection in their sector, for example, as they may be able to adopt new ideas without paying the firms who originally created the ideas. These firms might see decreases in the competitive threats they face. We also test whether acquisitions decrease after Alice, as these larger firms might be able to copy ideas without buying the firms that created them. Finally, we predict that firms might seek alternative ways to increase secrecy and protect IP after patent protection is lost. We predict that afflicted firms will thus use more non-disclosure agreements and become more secretive to replace lost IP protection.

3 Data and Methods

We begin this section by creating and validating a model of Alice's impact on lost IP protection at the patent level by training and applying the Longformer transformer model on the text of each patent in our sample. We then aggregate patent-level impact to the firm level to derive our key firm-level Alice treatment variable used in our study. We then discuss our final sample and present summary statistics.

3.1 Experimental Challenges

Our experiment needs to identify patents that were granted in the pre-Alice period but that would lose protection if they are tested in a court in the post-Alice period. This identification is challenging as there are more than 3.8 million patents granted between 06/19/1994 and 06/19/2014. To make the experiment tractable, we focus on the patents having the same primary CPC as those that were rejected by the USPTO per the Supreme Court's Alice criteria. This filtering leaves us 642,678 patents that we need to score on the likelihood of losing protection. Since manual examination is not feasible, we consider automated models with reliable predictions in this context.

Basic text-based similarity techniques such as term frequency–inverse document frequency (TF-IDF) have two major shortcomings. First, as technology vocabulary evolves and patents use related but different terms, TF-IDF may have limited power to capture similarity between two patents. Second, even when two patents use a similar vocabulary, the Supreme Court's Alice decision might affect one but not the other. These challenges motivate us to use an automatized system such as the Longformer model, a transformer-based language model (TLM), which is capable of catching both syntactic and semantic information.[11]

The primary benefit of transformer models is that they process words in relation to all the other words in a sentence, rather than one-by-one or in a fixed-sized sliding window approach. Therefore, TLMs can examine the full context of a word by looking at the words that come before and after it. This mechanism is referred to as self-attention and it provides the capability to understand the intent behind a sentence. To illustrate, we examine two sentences with similar meaning: i) Symptoms of influenza include fever and nasal congestion; ii) A stuffy nose and elevated temperature are signs you may have the flu. While a TF-IDF

[11] A large number of empirical analyses document that TLMs are superior to the traditional NLP models such as Bag-of-Words (BOW), Term Frequency-Inverse Document Frequency (TF-IDF), Word Embedding models such as Word2Vec, FastText, GloVe, and other approaches that combine Word Embedding Models with Neural Networks for Text Classification tasks (Adhikari et al. (2020); Maltoudoglou et al. (2022); Esmaeilzadeh and Taghva (2021); Minaee et al. (2021); Roman et al. (2021)).

model that filters the stop words (such as “and”) has a similarity score of 0, the BERT and Longformer models find 0.86 and 0.98 similarity scores for these two sentences, respectively.

3.2 The Longformer Model

Transformer-based language models (TLMs) are neural network models that revolutionized the field of Natural Language Processing (NLP) (Kalyan et al. (2021)). BERT is an early example released by Google in 2019, and it is used in their search engine. BERT also achieves state-of-the-art performance on various NLP tasks (Devlin et al. (2019)). Even though BERT is powerful, however, a limitation is that it can only process 400 words from any given text. A new transformer-based language model (TLM) Longformer (Beltagy et al. (2020)), was specifically designed to overcome this limitation, and it can handle longer documents up to 3,200 words. Furthermore, for long documents, Longformer is superior to several TLMs, such as BERT and RoBERTa (Beltagy et al. (2020); Gutiérrez et al. (2020); Xiao et al. (2021); Gupta and Agrawal (2022)). Our results are consistent with these findings, as we find that Longformer outperforms these other models.[12] We use Longformer to predict the likelihood that each of the pre-Alice granted patents in our sample might lose IP protection per the Alice decision.

Since TLMs require high computational power, the Longformer model is pre-trained using text from the Wikipedia, Book Corpus, CC-News, Open Web Text, and Stories datasets. The pre-trained model is then fine-tuned for a specific NLP task using an additional deep-learning layer with labeled data. We use descriptions of patent applications to train the model. However, patent descriptions are generally longer than Longformer's 3,200 word limit. The average number of words is 9,173, and only 12.57% of the pre-Alice granted patents in the sample have descriptions less than 3,200 words. We thus use the TextRank

[12] We also compare Longformer's out-of-sample prediction performance, and its performance on economic validation tests, to TF-IDF and other computational linguistics methods such as Word2Vec and also to BERT. Overall, Longformer outperforms the other models on these validation tests and we adopt Longformer as our baseline model. Results using SciBERT, which is a version of BERT pretrained on a scientific corpus were in an earlier version of the paper and are available from the authors.

automatic summarization tool, which internally uses Google's popular PageRank algorithm, to reduce the text size to 3,200 words (Mihalcea and Tarau (2004), Upasani et al. (2020)).

3.3 Rejected Patent Applications

We first gather the list of patents that are rejected under 35 U.S.C. §101 from the USPTO website.[13] We then identify the set classified as Alice-rejections based on the method of Lu et al. (2017). This step leaves us with 56,709 rejected patent applications. However, some are reapplications with a minor change (i.e., a change of only one or two sentences). Therefore, we compute pairwise similarities between the applications using TF-IDF and tag those with 0.99 similarity score as duplicates. For duplicates, we only keep the application with the latest date. We are left with 33,734 unique rejected patent applications that have a document number and Cooperative Patent Classification (CPC) information. These Alice-rejected patents belong to 5,831 unique CPCs.

3.3.1 Training The Longformer Model

There are two phases of training and evaluation. First, we train the system using the text of Alice-rejected patent applications (positives) and texts of applications that were eventually granted (negatives). After training, we evaluate predictions using a hold-out testing sample.

From the set of 33,734 Alice-rejected patent applications, we randomly choose 10,000 for hold-out testing and use the remaining 23,734 as positives to train the system. Next, we create a sample of negatives from patents that are successfully granted after 06/19/2014 (i.e., the Supreme Court's Alice decision). To ensure robustness, we construct the sample of negatives in four different ways based on the granularity of a patent's CPC (which has five levels): i) section; ii) class; iii) subclass; iv) group; and v) main group or subgroup. For example, in CPC "B60K35/00"; B, 60, K, 35, 00 corresponds to the Section, Class, Subclass, Group, and Main Group, respectively.

[13] https://developer.uspto.gov/product/patent-application-office-actions-data-stata-dta-and-ms-excel-csv

In experiment A, for each of the 23,734 positives, we find a matching negative patent that is in the same CPC Group that was granted after 6/19/2014. In samples B, C, and D, we keep adding 23,734 more matching patents to the negatives pool based on CPC Subclass, Class, and Section, respectively. Therefore, from A to D, each sample has 23,734 more negatives with the newly added ones drawn from broader CPC codes.

3.3.2 Testing Longformer and Other Models

In this section, we evaluate how predictions from our trained Longformer Model compare to predictions from SciBERT, BERT, RoBERTa, TF-IDF, and Word2Vec. For TF-IDF and Word2Vec predictions, we combine the model with logistic regression, decision tree, and random forest. In Internet Appendix Section 1, we present a detailed technical comparison of the transformer language models.

For this testing phase, we start with the 10,000 hold-out positives noted above. For each positive, we choose two negatives, giving us 20,000 negatives. This 1:2 ratio balances the fact that the expected rejection ratio is lower than half, and we do not overestimate accuracy for models that do not learn and only predict negative results.[14] To obtain the 20,000 negatives from the 708,115 post-Alice granted patents, we first gather 50,000 negatives that are randomly selected based on the CPC frequency distribution used in our prediction sample (see next section). Finally, from this negative pool, we sample 20,000 negatives 1,000 times to boot-strap the performance of each machine learning model.

We then evaluate resuts using standard performance metrics from computer science: precision, recall, F1 score, and accuracy. These metrics can be calculated from a confusion matrix with the following elements: True Positives (TP), False Positives (FP), True Negatives

[14] For example, if the ratio of positives to negatives is 1:9, then a model that does not learn but only predicts that all patents are negative would have 90% accuracy.

(TN), and False Negatives (FN):

$$Precision = \frac{TP}{TP+FP}, \qquad Recall = \frac{TP}{TP+FN}, \tag{1}$$

$$F_1\ Score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}, \qquad Accuracy = \frac{TP+TN}{TP+TN+FP+FN}. \tag{2}$$

Table 3 reports the evaluation results for the machine learning models. In the last column, we use an ensemble of the two models that have the highest F_1 Score and Accuracy by taking the average of their prediction scores (A has the highest F_1 Score and D has the highest Accuracy). The results show that the ensembled Longformer model is superior to all other machine learning algorithms. It has the highest F1 Score (0.672) and Accuracy (0.804).

Insert Table 3 here

3.3.3 Longformer Model Predictions for Existing Granted Patents

Our set of patents "to be examined for lost protection" consists of patents that were granted between 06/19/1994 and 06/19/2014 and share the same primary CPC with at least one of the applications that were rejected by the USPTO based on the Alice decision. In total, there are 642,678 such patents representing 16.6% of the total granted patents over this period.

The results in Table 4 show that 111,420 of these 642,678 patents (or 17.34% of the sample) have a Longformer score (predicted Alice rejection score) of 0.5 or higher, our threshold of high likelihood of losing protection if these patents are challenged in a court.

Panel B of Table 4 provides the list of CPCs that have the highest number of patent applications that were rejected by the USPTO and the list of CPCs that belong to patents that have a Longformer score of 0.5 or higher. There is large overlap across these two lists. Eight of the top ten CPCs for Alice-rejected patents are also in the high Longformer score list.

In Table 5 we provide further detail by industry and year on the number of granted patents in impacted Alice industries. We present these by industry for the top 10 industries along with the percentage of patents our Longformer model predicts would be rejected (those with Longformer scores ≥ 0.5). The table shows that, among the granted patents in these industries, multiple industries have over 25% of granted patents with Longformer scores ≥ 0.5. These patents would likely lose protection post-Alice. Corresponding CPCs for each industry are provided in Table 2. These percentages are similar to those in Table 1 for patents that were actually rejected in post-Alice years.

To illustrate which key words are most informative for the Longformer model, we also report the top 15 informative words for the top CPC groups impacted by Alice in Table 6. The table lists the words that are used most frequently in patents with high Longformer Alice scores (≥ 0.5) compared to those with low Longformer scores (<0.5). These words "open the black box" and illustrate which words are important.

3.3.4 Longformer Alice Scores Pre- and Post-Alice

The nature of Alice is that it will not only impact current patents but also future patents in the technological areas impacted by Alice. While it is hard to estimate this impact, we provide some statistics to gauge its potential magnitude. In particular, we estimate differences in the patents granted in key technological areas both pre-Alice (2011-2013) and post-Alice in 2017.

Table 7 shows the distributional density of the Longformer Alice Score of granted patents before the Alice shock (2011 to 2013) and after the shock (2017) for the Top 20 technological areas impacted by Alice. To compute the density in a given year, we first identify, the set of patents granted in that year in the Top 20 technological areas. The number of patents in each year ranges from 21,404 in 2011 to 31,249 in 2013 to 32,662 in 2017 (of those patents granted in 2017, 17,643 were applied for after the Alice decision). For the year 2017, as our goal is to examine the patent distribution post-Alice, we restrict attention to these 17,643 patents applied for in the post-Alice period. We sort all patents each year into 10 bins based on each patent's Longformer Alice Score. Bins are defined as the ten equal segments in the interval (0,1), which is the range of the Alice Score. For each bin, the density is the number of patents in the given bin in the given year divided by the total number of patents in the given year.

Finally, to illustrate the impact of Alice on these areas, we compute the ratio in the final column as the density in 2017 (column 5) divided by the average pre-Alice density averaged over the years 2011 to 2013 (column 4). A ratio below unity indicates that the rate of patenting in the given bin declined post-Alice.

Column (6) of Table 7 shows that, for bins with Longformer Alice scores exceeding 0.5, patenting has declined sharply.[15] In decile 10, the decile with the highest Alice scores, patenting is only 38% of pre-Alice patenting. Overall, these numbers can be applied to the number of patents in 2013 to estimate the total number of patents that "likely would have been granted in 2017 if the Alice judgment had not occurred. In particular, for each bin having materially positive Alice Scores (those greater than 0.3 in Table 7), we multiply one minus the ratio in Column (6) by the number of patents in the given bin in 2013. We then add these "likely lost patents" over these bins, and the result is 3,237 patents. This calculation thus estimates that Alice resulted in 3,237 fewer patents per year by 2017 in these 20 technological areas. Because Alice is still in effect, this annual total is likely to accumulate every year, indicating an economically large impact.

The impact of Alice is also shown graphically in Figure 1, which shows the percentage of post-Alice patents granted in 2017 relative to the numbers in 2011-2013 (this is Column (6) of Table 7). The sharp drop-off on the RHS of the figure illustrates that firms greatly reduced patenting in technologies that had the most Alice exposure.

[15] We find similar results if we use patent applications instead of patent grants. The tables can be requested from authors.

3.4 Patent Sample and Treatment Measure

We create the treatment measure for each firm i that we use in our regressions as follows:

$$Treatment_i = \frac{\sum_{j=1}^{N_i} PatentValue_{i,j} \times AliceScore_{i,j}}{Sales_i} \tag{3}$$

In this equation, $Sales_i$ is firm i's total sales in 2013. $PatentValue_{i,j}$ refers to the dollar value of patent j for firm i obtained from the KPSS database (Kogan et al. (2017)). We compute an alternative treatment variable where we replace patent value with the number of citations the patent received (discussed below). The treatment variable is computed for each firm in 2014 using all valid granted patents prior to the Alice decision. The patent values in equation (3) are depreciated using an annual 20% rate relative to the base year 2014.[16] Figures are further adjusted for inflation. $AliceScore_{i,j}$ refers to Longformer's predicted probability that a patent j loses protection if it were to be challenged in a court. N_i is the total number of valid patents of firm i.

As an alternative to KPSS-based valuation, we create a citation-based metric to estimate $PatentValue_{i,j}$ in Equation (3). In this method, for each patent j, we count the number of granted patents that cite j and have an application date that is within five years of j's grant date. As we did for the KPSS-based methodology, we depreciate citation-based value using an annual 20% rate relative to the base year 2014.

3.5 Sample and Key Variables

We include public firms with at least one patent from a CPC category with a rejected patent. We link patents to public firms using the correspondence provided by Kogan et al. (2016), who extended the data until 2020. Our patent text data comes directly from the USPTO website. We also include the competitors of each firm in our sample using the TNIC-3

[16] We use a 20% depreciation rate following Hall and Li (2020)'s finding that depreciation rates are likely higher than the 15% typically used in the literature, especially in high technology sectors. Our results are fully robust to using a 15% rate.

competitor network of Hoberg and Phillips (2016). Our sample thus includes 3,444 unique firms: 1,035 Alice-affected firms and also 2,409 competitor firms.

Table 8 displays summary statistics for the sample of firms used in our analysis. Our sample contains 19,372 firm-year observations based on our sample screens noted above spanning the period from 2011 to 2017 (excluding 2014, the treatment year). We briefly describe the variables used in our analysis (full details and a variable list are in Appendix A). Table 8 presents summary statistics for firms both in the pre-Alice period of 2011-2013 and in the post-Alice period of 2015-2017.

Our goal is to examine firms with granted patents that were exposed to Alice as identified by our Longformer model. We examine their innovation decisions, their lawsuits and other legal consequences. We then examine the impact of Alice on their ex post profitability and the competition they face in their product markets. Lastly, we examine how they change their acquisitions in response to their Alice exposure.

Panel A of Table 8 presents accounting characteristics including the size of firms measured by assets and sales, sales growth, age, and profitability (Operating income / Sales) of firms. We also present firm Tobin's q ((market value of equity + book value of debt) / book value of assets). The table shows for example that overall operating margins and sales growth decline. Later, we explore these findings using rigorous models with firm fixed effects, and we explore if they differ for large vs. small firms.

Panel B presents the innovation and legal variables used in our study. The variable *Treatment Effect* measures the extent a firm's patent portfolio is impacted by the court decision as measured using the Longformer model in equation (3). It captures how much a firm is dependent on patents and also the percentage of patents' value that are impacted by the Alice court decision. R&D/Sales is Compustat R&D divided by total sales of the firm and is set to zero if R&D is missing for our base tests. Log(# of Patents) is the log of one plus the number of patent applications. We define Acquisitions/Sales as the dollar value of acquisitions scaled by sales. PatTargets/Sales is the dollar value of acquisitions where target has a patent scaled by sales. Acquisitions data is from the Securities Data Corp (SDC) database.

The legal variables we examine are *# Alleged*, *# NPE Alleged*, *# OC Alleged*,*# Accuser*, *IPrisk* and *PatInfringe*. We compute the first four using information in the Public Access to Court Electronic Records (PACER) database, which provides public access to all cases litigated in the U.S. District Courts, and Stanford Non-Practicing Entity (NPE) Litigation Database. For the last two, we use textual queries of each firm's 10-K statement filed with the SEC. *# Alleged* is the number of lawsuits that a firm was alleged for infringing on a patent in that year. *# NPE Alleged* (*# OC Alleged*) refer to the number of lawsuits that the firm was alleged infringing on a patent lawsuit by a Non-Practicing Entity (Operating Company) in that year. *# Accuser* is the number of lawsuits that the firm accused another party for infringing on a patent in that year. *IPrisk* is the total number of paragraphs mentioning "intellectual property" in the risk factor section of the firm's 10-K, scaled by the total number of paragraphs in the 10-K. *PatInfringe* is the total number of 10-K paragraphs containing both a patent word and a word that contains the word root "infringe", also scaled by the total number of paragraphs in the firm's 10-K. The table shows that patents decline and lawsuits and patent infringement all decline post-Alice while IP risk increases.

Lastly, Panel C of Table 8 presents the competition variables. *VCF/Sales*, is the a measure of VC entry in a given firm's product market and is the total first-round dollars raised by the 25 startups from Venture Expert whose Venture Expert business description most closely matches the 10-K business description of the focal firm (using cosine similarities), scaled by focal firm sales. *TSIMM* is the firm's TNIC-3 text-based total similarity to other public firm competitors from Hoberg and Phillips (2016). The next three variables are constructed using the metaHeuristica software package to run high speed queries on 10-Ks filed with the Securities and Exchange Commission. *Complaints* is the number of paragraphs in the firm's 10-K that complain about competition divided by the total number of paragraphs in the firm's 10-K. *Noncompete* is the number of paragraphs in a firm's 10K mentioning "non-compete" agreements, scaled by the total paragraphs in the 10-K. *Nondisclose* is the number of paragraphs mentioning "non-disclose" or NDA agreements in a firm's 10K, scaled by the total paragraphs in the 10-K. The table shows that competition overall increases post-Alice while nondisclosure agreements increase. We now turn to regressions that include firm fixed effects and explore the differences for large and small firms using above or below median total assets relative to TNIC-2 industry peers from Hoberg and Phillips (2016).

Our treatment variable is not binary as they represent the multiplication of percentage of a firm's patent portfolio value that is exposed to Alice scaled by sales. Each patent's Alice exposure score is the probability from our Longformer model that the patent will be ruled ineligible if it is challenged in court. For roughly half of the sample, the treatment Alice score is close to 0. The median and average scores of treatment in our sample are 0.001 and 0.062, and the 75^{th} percentile and 90^{th} percentiles are 0.034 and 0.224, respectively. The standard deviation of treatment variable is 0.025. We show the full distribution of firm-level treatment Alice scores for patenting firms in Figure 2. Panel A shows the histogram for our KPSS valuation-based treatment variable and Panel B for our citation-based treatment variable.

Both panels show that roughly 25% to 30% of our firms have zero Alice scores. Roughly 40% have positive scores that are close to zero. Thus, roughly 65% have Alice scores equal or slightly greater than zero (from 0.0 to .005). About 6% to 7% of firms have very high exposure to Alice with Alice scores in the rightmost bin. Using a continuous treatment score allows us to show how the ex post outcomes vary with the intensity of treatment. Going forward for parsimony, we refer to the Longformer Alice score just as "Alice Score".

4 The Impact and Outcomes of Alice

We now analyze the impact of Alice on innovation, firm performance and value, competition, lawsuits and legal risk, and acquisitions. Throughout this section, we present results separately for large and small firms, as we have found uniformly that there are key differences in treatment across these two groups.

The justification for examining heterogeneous effects based on firm size follows from Aghion et al. (2005) and is based on the fact that larger firms are more able to defend their product markets given their larger resources - both managerial and financial. We define firms as large or small, respectively, based on whether each firm's assets are above or below the median value among its TNIC-2 industry peers (see Hoberg and Phillips (2016)) in 2013.[17]

For all regression tables that follow, *Post* is an indicator variable that equals one if the year is after the Alice decision (2015 to 2017) and zero if before (2011 to 2013). We omit 2014 itself from our analysis as it is partially treated. *Treatment* throughout is a firm-level measure that combines information about the extent to which patents are important for the firm and the extent the firm's patent portfolio was affected by the Alice court decision. Throughout, we use the firm treatment value using each patent's Alice score weighted by the patent's importance to the firm. We present results using two different weights: (1.) using each patent's KPSS weighted value and also (2.) using each patent's citations weighted value. The mathematical notation for the estimation of this measure is provided in equation (3), with citations replacing patent value for the citation based measure. Inspection of the subsequent tables reveals that there is little difference in the results across these two different weighting methods for a patent's importance to the firm. All regressions include firm and year fixed effects. Standard errors are clustered at the firm level.

4.1 Alice and Innovation

We first examine the impact of Alice on firm innovation and we examine the number of patents scaled by sales, the log of 1 plus the number of patents, and R&D/Sales.

Insert Table 9 here

[17] In the appendix, we also present results that define large vs. small based on each firm's market share based on TNIC product-text industry peers. These results are fully robust.

The results for patents in columns (1)-(4) of Table 9 show that both large and small firms reduce patenting in the years after Alice. These results are highly significant at the 1% level, and these findings confirm the large importance of the Alice decision to reduce the incentives to patent through its weakening of IP protection. The effect is also larger for large firms in columns (3) and (4) consistent with large firms getting more patents in general. The economic effect of the decision is large. Using the coefficients in column (3), We calculate that large (small) firms patenting decreases by 9.4% (13.4%) relative to the mean patenting rates with a one standard deviation in the treatment variable. We show these results graphically in Figures 3 and 4 for small and large firms, respectively, to test for pre-trends. The graphical evidence shows no evidence of pre-trends and shows that patents discretely shifted downwards in the years following Alice.

Insert Figures 3 and 4 here

The results for R&D in columns (5)-(6) show that small firms increase R&D after Alice, while there is no significant change for large firms. Using the coefficient from column (6), we calculate that small firms R&D increase by 76.7% relative to the mean R&D of small firms with a one standard deviation in treatment relative to the mean pre-Alice.

The R&D results are consistent with small firms trying to increase R&D to make up for lost intellectual property, an interpretation more broadly supported in our later tables. In contrast, large firms do not increase R&D, indicating they were impacted by the shock in a fundamentally different way in which more R&D was not seen as a necessary response. This muted response by larger firms echoes results throughout our paper suggesting that larger firms (presumably due to their deep pockets and wider-array of knowledge capital) came out of the Alice shock as winners, whereas smaller firms experienced significant losses.

We also note that our regressions include firm fixed effects, and thus we do not report the lower interactions including the individual variables (Large, Small, and Treat) as these are absorbed by firm fixed effects given that they are defined in the treatment year.

We examine changes in competition in the next section, and relevant to the current discussion, we find that competition increases most for small firms in their local product markets. The results on innovation combined with these increases in competition are consistent with innovation spending increasing by small firms to "escape-competition" that materialized from Alice exposure. In contrast, large firms do not increase R&D as much. The results are consistent with the Schumpeterian effect where more innovation is done by small firms.

4.2 Alice and Competition

Unlike some existing studies, which focus on the impact of individual patent invalidations, our study examines the impact of a technology-area-wide loss in IP protection. Such a market-wide shock impacts both existing patents and also the incentives to patent more in the future. These shifts in patenting incentives furthermore affect incentives of potential competitors, and thus it is important to examine the impact of Alice on competition coming from either new VC funded entrants as well as from existing public firms.

We thus examine several different measures of changes in firm-level competition. We begin by examining entry by venture capital financed firms in each firm's product market, and we also examine changes in competition from existing public firms using product similarity from Hoberg and Phillips (2016). We also examine the most broad measure of competition as the intensity at which firms complain about competition in their 10-Ks. Finally, especially given the strong results we find in firm-year panel data analysis, we then examine measures of product market encroachment at the level of firm-pairs over time, to specifically examine if big firms or small firms move "closer" together in the product space post-Alice using firm-pair-level product similarity scores.

Columns (1) and (2) of Table 10 examine venture capital entry into a firm's local product market. The dependent variable, *VCF/Sales*, is the total first-round dollars raised by the 25 startups from Venture Expert whose Venture Expert business description most closely matches the 10-K business description of the focal firm (using cosine similarities), scaled by focal firm sales (see Hoberg, Phillips and Prabhala 2014). Columns (3) and (4) examine the firm's TNIC text-based total similarity (*TSIMM*) to public firm competitors (see Hoberg and Phillips 2016). We examine broad competition *Complaints* in columns (5) and (6). *Complaints* is the number of paragraphs in the firm's 10-K that complain about competition divided by the total number of paragraphs in the firm's 10-K.

Insert Table 10 here

Economically, Table 10's results indicate that entry by venture capital financed firms into the market of small firms increases by 99.6% with a one standard deviation increase in treatment relative to the average entry rate pre-Alice. This is significantly higher than that for large firms. Looking at direct measures of competition, both product similarity and complaints increase for small firms with no significant increases for large firms. Complaints by small firms increase by .33 and product similarity increases by 28.1% with a one standard deviation increase in treatment relative to the product similarity pre-Alice. We also present these results graphically for small firms where we allow each pre- and post-year to have its own indicator variable. These results are presented in Figure 5.

The results in Table 10 show, across all aspects of competition, that small firms face increased competition from myriad of sources post-Alice. In contrast, large firms face increased entry but do not experience changes in product similarity and complaints in any of the specifications and are generally unaffected. These results are consistent with our results on profitability decreases for small firms and decreased market values for small firms (discussed in the next section). The results reinforce our conclusion that small firms whose patent portfolios are exposed to Alice experience losses, while large firms experience a smaller amount of increased VC entry but no other increases in competition. Rather, large firms actually experience some gains in the form of increased sales and market valuations.

We now examine local pairwise product market encroachment post-Alice in Table 11. *Delta TNIC Score* is computed as the change in the TNIC similarity of the pair of firms from year $t-1$ to year t. Our panel database for this test is thus a large firm-pair-year panel.

A higher value of the *Delta TNIC Score* indicates that the firms in the pair encroached upon one another in the current year and lost pairwise product differentiation (becoming more intense competitors). For all RHS variables in Table 11, we use the tags “1” and “2” in each variable’s name to indicate whether the given variable is a trait of the first or second firm in the pair. For example, the variable Treat1 indicates the treatment intensity of firm 1.

The results in column (1) of Table 11 show that firms experiencing a larger treatment effect from Alice experience increased encroachment at the pair level. This is consistent with weaker IP protection resulting in rivals adopting patented technologies of rivals and in product offerings of the pair becoming more similar. These results are highly significant despite the inclusion of firm-pair fixed effects and clustering standard errors by firm-pair.

Column (2) of Table 11 illustrates our main result that outcomes are different for small and large firms. We find that small firms are particularly sensitive to encroachment when they lose their IP protection. This is consistent with these firms holding narrower advantages in the product market, and losses in protection can be catastrophic as rival firms would have free access to their technologies post-Alice. In contrast, larger firms actually experience increases in product differentiation when their markets are treated by Alice. This is consistent with these firms having broad patent portfolios that span technology areas, making it harder for rivals to enter their product markets.

The final column (3) in Table 11 interacts these results further to examine the four-way interactions of the sizes of firm1 and firm 2. The results indicate that positive encroachment only occurs when there is a small firm in the pair that is treated by Alice. Indeed, Small1×Big2×Treat1×Post has a positive coefficient as does Small1×Small2×Treat1×Post. However, once the treated firm is a large firm, the coefficient flips to negative, indicating that larger firms tend to experience radically different outcomes than small firms. Indeed many scholars argue that patent protection could either be harmful or helpful to incentivize innovation and growth.

4.3 Alice and Firm Performance

We now examine the profitability of firms post-Alice. Table 12 displays panel data regressions that examine whether the sales, profitability and market value of large vs. small firms were differently affected by the Alice decision. In columns (1)-(2), the dependent variable is *Sales Growth*, calculated as the natural logarithm of total sales in the current year t divided by total sales in the previous year $t - 1$. In columns (3) and (4), the dependent variable is *Operating Income/Sales*. In columns (5)-(6), the dependent variable is *Tobin's q*, calculated as the market to book ratio (market value of equity plus book debt and preferred stock, all divided by book value of assets).

Insert Table 12 Here

Table 12 shows that large firms whose patent portfolios are exposed to Alice experience sales growth and positive but insignificant gains in profitability and market value (measured using Tobin's q) post-Alice. Their sales go up by 1.3 percentage points (14.1% of their 2013 average growth rate). Thus, large firms appear to benefit some when they are operating in technology markets that experience market-wide losses in patent protection. As our later results will suggest, these gains at least partially come at the expense of small firms, as large firms would thus face weakening competition when their smaller rivals have to scale back.

Consistent with this view, Table 12 shows that small firms indeed experience losses after Alice. Small firms whose patent portfolio is exposed to Alice suffer decreased operating margins and also losses in their market valuations. These results persist when additionally controlling for firm age and also for firm size. Economically, small firms' operating margins go down by 27.5 percentage points (91% of their pre-Alice operating margin) and their Tobin's q declines by 0.21 which is 12 percent of their pre-Alice Tobin's q with a one standard deviation increase in treatment.

4.4 Legal Impact: Contractual Provisions and Lawsuits

The matter of intellectual property protection is inherently a matter of legal protection and a means of reducing the risk that rival firms will expropriate a focal firm's technological advantage. Thus we examine, across multiple legal metrics, how the legal situation changes for large and small firms post Alice.

We start with two important aspects of firm legal outcomes: the intensity at which they disclose risk of loss of IP (an important test of validity), and the extent to which firms use alternative "second best" contracts including non-compete and non-disclosure agreements to improve IP protection after IP protection through patents is lost following the Alice decision.

Table 13 displays the results. In columns (1)-(2), *IP Risk*, is the total number of paragraphs mentioning "intellectual property" in the risk factor section in the 10-K documents, scaled by the total number paragraphs in the 10-Ks. *Noncompete* is the total number of 10K paragraphs mentioning "non-compete" agreements, all scaled by the total paragraphs in the 10-K. *Nondisclosure* is the total number of 10-K paragraphs mentioning "non-disclose" or NDA agreements, all scaled by the total paragraphs in the 10-K.

The results in Table 13 show that small firms disclose significantly more information about increased intellectual property risk in the risk section of their 10-K, economically increasing by 7.34 percent of their pre-Alice mean. This provides important validation of the primary impact of the Alice case itself, and that the negative consequences were particularly felt by smaller firms. The table also shows that small firms also use more non-disclosure agreements - economically increasing the mention of these by 45.6 percent of their pre-Alice mentions (although they do not significantly increase non-compete agreements). Across all of these outcomes, we find no significant changes for large firms. Overall, the results show that small firms face greater IP risk and use alternative contracts to protect their IP after Alice.

In Table 14, we next examine whether patent lawsuits involving small and large firms were differentially affected by the Alice decision. We use the Stanford Non-Practicing Entity (NPE) Litigation Database to find NPE and operating company (OC) initiated lawsuits. In columns (1)-(2), the dependent variable, *# Alleged*, is the total number of lawsuits in which the firm was alleged to have infringed on a patent in the given year. In columns (3) to (4), *# NPE Alleged* is the number of such lawsuits in which the plaintiff is a non-practicing entity (NPE). In columns (5) to (6), *# OC Alleged* is the number of such lawsuits in which the plaintiff is an operating company. In columns (7)-(8), *PatInfringe* refers to the total number of paragraphs containing a patent word and infringe* in the firm's 10-K, scaled by the total number of paragraphs in the 10-Ks. The 10-K measure establishes robustness, as some infringement cases might be settled out of court, and might not appear in court records, but nevertheless might be discussed in a firm's 10-K. In columns (9)-(10), *# Accuser* is the number of lawsuits in which the firm accused any party in a patent lawsuit in the given year.

In contrast to earlier findings that illustrated strong results for small firms, Table 14 shows that small firms' lawsuit exposure changes less than it does for large firms post-Alice. We find that lawsuits including small firms increased in some specifications, which is opposite the widespread and highly significant decrease we observe for large firms. This result for small firms, especially when viewed alongside the greater IP risk and increased use of non-disclosure agreements post-Alice for these firms, is consistent with the broader losing-ground scenario we document. These firms are likely forced to test IP boundaries more (thus increasing their exposure to lawsuits), and they also attempt to replace lost patent protection using other contracts such as NDAs.

The results are different for large firms, whose lawsuit exposure significantly *decreases* across the board after Alice. Large firm are less likely to be alleged to infringe on other firms and their exposure to lawsuits also decreases for lawsuits by non-performing entities (patent trolls) post-Alice. These results are intuitively interpreted through two impacts of Alice. First, Alice reduced IP protection, resulting in lawsuits becoming less viable as a means to extract wealth from another party (one needs strong IP to successfully make a claim of infringement). Second, the gains associated with having fewer lawsuits, especially from patent trolls, accrued mostly to larger firms whose legal teams were able to internalize these gains. Smaller firms, whose ability to defend IP may be more limited, were less able to achieve this outcome as noted above. Overall our evidence again shows that large firms appear to benefit, and small firms experience losses, following the Alice ruling.

4.5 Alice and Acquisitions

We now examine the impact of Alice on firm acquisitions by small and large firms. Our hypothesis is that large-firm acquisitions will decline after Alice following the theory of Phillips and Zhdanov (2013) and empirical support of Caskurlu (2022). Phillips and Zhdanov (2013) show that large firms have strong incentives to buy small firms after small firms develop a new patentable innovation. Without patent protection, there is little incentive for large firms to continue paying to buy these small firms for their patents, as they can more cheaply copy the unprotected innovation. If they do purchase a small firm, the purchase price will be lower as the bargaining power of the small firms will have decreased post-Alice. In line with Phillips and Zhdanov (2013), Caskurlu (2022) shows that after a firm loses a patent lawsuit, it is more likely to acquire targets that have substitute patents. When patent rights are weaker, there will be fewer lawsuits and fewer patent-motivated acquisitions. We thus examine the impact of Alice on the dollars spent on acquisitions scaled by sales and the log of one plus the dollar value of acquisitions.

The results are displayed in Table 15. Across many specifications presented in Table 15, we indeed find that the amount spent by large firms on acquisitions post-Alice decreases significantly. Although we do not find results for the total unconditional number of acquisitions scaled by sales in columns (1) and (2), we do find the predicted results in columns (3) and (4) when we only consider acquisitions in which the target firm has at least one patent (as our hypothesis only applies to patented technologies). We also find results for the log of the number of acquisitions in columns (5) and (6). In contrast, we find no impact on small firms. Our results are overall consistent with the predictions of Phillips and Zhdanov (2013) and Caskurlu (2022) that decreased patent protection leads to decreased bargaining power for small firms that are targets, and thus large firms acquire less and pay less for any firms they do acquire. These results once again point to potential gains by large firms post-Alice (who save by spending less on acquisitions), and additional losses for smaller firms (who have fewer options for monetizing their IP through M&A).

4.6 Robustness Tests

Table 3 shows that the Longformer model outperforms other linguistic models in predicting out-of-sample predictions of a patent's likelihood of being rejected. In this section, we also assess the economic advantages of using the Longformer model.

We thus re-estimate our econometric tests using the SciBERT model. These results were in a previous version of the paper and are available from the authors. We also estimate our results separately using the TF-IDF method and also a simple binary dummy variable for the CPC category to identify a patent's exposure to Alice (see Section 3 in the Online Appendix). The SciBERT model came out second-best in our model performance validation tests, and these economic robustness tests indicate that its overall results are quite similar to Longformer. For the TF-IDF method, in the calculation of the treatment variable depicted in equation (3), we use a TF-IDF score instead of a Longformer Alice score. The binary CPC method sets exposure equal to one if the patent's primary CPC code belongs to one of the top-20 CPCs that have the most frequent Alice rejections. We then aggregate each over all of the firm's patents as before to get a total firm exposure. We expect and find material improvements using transformer models such as Longformer and SciBERT relative to using less-advanced methods such as TF-IDF or CPC dummies.

We present in the online appendix a large number of tables using these alternative models that estimate exposures to Alice. For example, Tables IA7 and IA8 display tests for the patenting and innovation results that are analogous to Table 9, but use the TF-IDF method and binary CPC category method, respectively.

Overall, in across all of our robustness tests, the signs are similar to our main results using the Longformer model. Yet the results also show the gains to using the more accurate Longformer model, as we lose some significance for patenting for small firms using TF-IDF and for R&D using the binary dummy variable. We also lose significance for several of the competition variables for small firms using either of these two less sophisticated methods. Given the higher out-of-sample prediction accuracy shown in Table 3 for the Longformer model versus the other methods, we conclude that the gains associated with using the deep learning neural network Longformer model are both statistically and economically important.

Our Alice Score treatment variable is estimated using machine learning techniques, and hence is measured with some noise. Although noise generally results in findings being understated, recent work in econometrics (de Chaisemartin and D'Haultfoeuille, 2018) indicates a new technique for estimating a fuzzy difference-in-differences model that is applicable in our setting. In Table IA9 and Table IA10, we report the local average treatment effect (LATE) indicated by the authors using the fuzzydid model (see de Chaisemartin et al. (2019) for implementation in Stata). Although power is somewhat reduced in this setting, we find that most of our results are robust in this specification.

5 Conclusions

We examine the impact of lost intellectual property protection on firm innovation, performance, competition, and mergers and acquisitions. We examine firms whose patents would likely lose protection if challenged following the Alice Corp v. CLS Bank International, 573 U.S. 208 (2014) Supreme Court decision. This decision revoked patent protection on patents whose fundamental idea is considered abstract with a transformation that is not novel. It impacted multiple areas including business methods, software, and bioinformatics. The outcome of this decision was very much in doubt and was not anticipated.

While the decision had an extremely large ex post impact on patenting, there was (and is) uncertainty about whether an existing or proposed patent transforms an idea sufficiently to be granted patent protection. Given the uncertainty about whether the Alice decision impacts individual patents, we apply an array of novel machine learning techniques on patent textual corpora to assess how much a given firm's patent portfolio is exposed to Alice.

We document that ex post patenting by firms whose patent stock portfolio is identified as being exposed to Alice significantly decreases for both large and small firms. We find a significant increase in R&D for small firms. These results are consistent with small firms' attempting to replenish their innovative portfolio as predicted by Aghion et al. (2005). Examining ex-post changes in sales growth and profitability along with firm value, we find an asymmetric impact of Alice on firms whose patent portfolio is exposed to Alice. Large firms gain and small firms lose. Exposed large firms gain in sales. Exposed small firms experience a decrease in operating margins and their market valuations decline.

We show that these differential losses by small firms can be explained by changes in competition and limited legal options to replace losses in IP protection. Small firms face increased competition using a number of different measures, while large firms are only minimally impacted. In the post-Alice period, small affected firms face increased venture capital financed entry into their product space, lost product differentiation relative to their existing competitors, and they complain more about increased competition. Consistent with trying to protect IP that has lost protection, small firms resort more to non-disclosure agreements with their employees post-Alice. In contrast, large firms once again appear to relatively gain as they face fewer lawsuits from both operating companies and non-producing entities ("patent trolls"), and decreased direct competition from smaller firms. They also acquire fewer target firms, especially those with patents, after Alice. Our results illustrate an uneven impact of lost IP protection across firms.

Our paper finds benefits of increased competition and fewer lawsuits from reduced patent protection but costs for existing small firms who most directly face the impact of increased competition from both large firms and new entrants. Our results overall show the costs and benefits of decreased IP protection.

References

Acemoglu, Daron, Phillipe Aghion, Rachel Griffith, and Fabrizio Zilbotti, 2010, Vertical integration and technology: Theory and evidence, *Journal of the European Economic Association* 989–1033.

Adhikari, Ashutosh, Achyudh Ram, Raphael Tang, William L Hamilton, and Jimmy Lin, 2020, Exploring the limits of simple learners in knowledge distillation for document classification with docbert, in *Proceedings of the 5th Workshop on Representation Learning for NLP*, 72–77.

Aghion, Philippe, Nicholas Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt, 2005, Competition and innovation: an inverted u relationship, *Quarterly Journal of Economics* 120, 701–28.

Ammar, Waleed, Dirk Groeneveld, Chandra Bhagavatula, Iz Beltagy, Miles Crawford, Doug Downey, Jason Dunkelberger, Ahmed Elgohary, Sergey Feldman, Vu Ha, et al., 2018, Construction of the literature graph in semantic scholar, in *Proceedings of NAACL-HLT*, 84–91.

Beltagy, Iz, Kyle Lo, and Arman Cohan, 2019, Scibert: A pretrained language model for scientific text, in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 3615–3620.

Beltagy, Iz, Matthew E Peters, and Arman Cohan, 2020, Longformer: The long-document transformer, *arXiv preprint arXiv:2004.05150* .

Boldrin, Michele, and David K Levine, 2013, The case against patents, *Journal of Economic Perspectives* 27, 3–22.

Budish, Eric, Benjamin N Roin, and Heidi Williams, 2015, Do firms underinvest in long-term research? evidence from cancer clinical trials, *American Economic Review* 105, 2044–85.

Caskurlu, Tolgal, 2022, Effects of patent rights on acquisitions and small firm r&d, *University of Amsterdam Working Papers* .

de Chaisemartin, C, and X D'Haultfoeuille, 2018, Fuzzy Differences-in-Differences, *The Review of Economic Studies* 85, 999–1028.

de Chaisemartin, Clement, Xavier D' Haultfoeuille, and Yannick Guyonvarch, 2019, Fuzzy differences-in-differences with stata, *The Stata Journal* 19, 435–458.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova, 2019, BERT: pre-training of deep bidirectional transformers for language understanding, in Jill Burstein, Christy Doran, and Thamar Solorio, eds., *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers)*, 4171–4186 (Association for Computational Linguistics).

Esmaeilzadeh, Armin, and Kazem Taghva, 2021, Text classification using neural network language model (nnlm) and bert: An empirical comparison, in *Intelligent Systems and Applications: Proceedings of the 2021 Intelligent Systems Conference (IntelliSys).*, volume 296, 175, Springer Nature.

Farre-Mensa, Joan, Deepak Hegde, and Alexander Ljungqvist, 2020, What is a patent worth? evidence from the us patent "lottery", *The Journal of Finance* 75, 639–682.

Galasso, Alberto, and Mark Schankerman, 2015, Patents and cumulative innovation: Causal evidence from the courts, *The Quarterly Journal of Economics* 130, 317–369.

Gokaslan, Aaron, and Vanya Cohen, 2019, Openwebtext corpus, `http://Skylion007.github.io/OpenWebTextCorpus`.

Gupta, Manish, and Puneet Agrawal, 2022, Compression of deep learning models for text: A survey, *ACM Transactions on Knowledge Discovery from Data (TKDD)* 16, 1–55.

Gutiérrez, Bernal Jiménez, Jucheng Zeng, Dongdong Zhang, Ping Zhang, and Yu Su, 2020, Document classification for covid-19 literature, in *Findings of the Association for Computational Linguistics: EMNLP 2020*, 3715–3722.

Hall, Bronwyn, and Wendy Li, 2020, Depreciatino of business r&d capital, *Review of Income and Wealth* 66, 161–180.

Hoberg, Gerard, and Gordon Phillips, 2016, Text-based network industries and endogenous product differentiation, *Journal of Political Economy* 124, 1423–1465.

Kalyan, Katikapalli Subramanyam, Ajit Rajasekharan, and Sivanesan Sangeetha, 2021, Ammu: a survey of transformer-based biomedical pretrained language models, *Journal of biomedical informatics* 103982.

Kesan, Jay, and Runhua Wang, 2020, Eligible subject matter at the patent office: An empirical study of the influence of alice on patent examiners and patent applicants., *Minnesota Law Review* 105, 527–617.

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

Kogan, Leonid, Dimtris Papanikolaou, Amit Seru, and Noah Stoffman, 2016, Technological innovation, resource allocation and growth, *Quarterly Journal of Economics* forthcoming.

Lerner, Josh, 2002, 150 years of patent protection, *American Economic Review* 92, 221–225.

Lim, Daryl, 2020, The influence of alice, *Minn. L. Rev. Headnotes* 105, 345.

Liu, Yinhan, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov, 2019, Roberta: A robustly optimized bert pretraining approach, *arXiv preprint arXiv:1907.11692* .

Lu, Qiang, Amanda Myers, and Scott Beliveau, 2017, Patent prosecution research data: Unlocking office action traits, USPTO Economic Working Paper No. 10, Available at SSRN: https://ssrn.com/abstract=3024621 or http://dx.doi.org/10.2139/ssrn.3024621.

Maltoudoglou, Lysimachos, Andreas Paisios, Ladislav Lenc, Jiří Martínek, Pavel Král, and Harris Papadopoulos, 2022, Well-calibrated confidence measures for multi-label text classification with a large number of labels, *Pattern Recognition* 122, 108271.

Mihalcea, Rada, and Paul Tarau, 2004, Textrank: Bringing order into text, in *Proceedings of the 2004 conference on empirical methods in natural language processing*, 404–411.

Mikolov, Tomás, Kai Chen, Greg Corrado, and Jeffrey Dean, 2013, Efficient estimation of word representations in vector space, in Yoshua Bengio, and Yann LeCun, eds., *1st International Conference on Learning Representations, ICLR 2013, Scottsdale, Arizona, USA, May 2-4, 2013, Workshop Track Proceedings*.

Minaee, Shervin, Nal Kalchbrenner, Erik Cambria, Narjes Nikzad, Meysam Chenaghlu, and Jianfeng Gao, 2021, Deep learning–based text classification: A comprehensive review, *ACM Computing Surveys (CSUR)* 54, 1–40.

Nagel, Sebastian, 2016, Cc-news.

Nordhaus, William D, 1969, An economic theory of technological change, *The American Economic Review* 59, 18–28.

Phillips, Gordon M., and Alexei Zhdanov, 2013, R&d and the incentives from merger and acquisition activity, *Review of Financial Studies* 34-78, 189–238.

Robertson, Stephen, 2004, Understanding inverse document frequency: on theoretical arguments for idf, *Journal of documentation* .

Roman, Muhammad, Abdul Shahid, Shafiullah Khan, Anis Koubaa, and Lisu Yu, 2021, Citation intent classification using word embedding, *IEEE Access* 9, 9982–9995.

Sampat, Bhaven, and Heidi L Williams, 2019, How do patents affect follow-on innovation? evidence from the human genome, *American Economic Review* 109, 203–36.

Trinh, Trieu H, and Quoc V Le, 2018, A simple method for commonsense reasoning, *arXiv preprint arXiv:1806.02847* .

Upasani, Siddhant, Noorul Amin, Sahil Damania, Ayush Jadhav, and A. M. Jagtap, 2020, Automatic summary generation using textrank based extractive text summarization technique, volume 07.

Xiao, Chaojun, Xueyu Hu, Zhiyuan Liu, Cunchao Tu, and Maosong Sun, 2021, Lawformer: A pre-trained language model for chinese legal long documents, *AI Open* 2, 79–84.

Zellers, Rowan, Ari Holtzman, Hannah Rashkin, Yonatan Bisk, Ali Farhadi, Franziska Roesner, and Yejin Choi, 2019, Defending against neural fake news, *Advances in neural information processing systems* 32.

Zhu, Yukun, Ryan Kiros, Rich Zemel, Ruslan Salakhutdinov, Raquel Urtasun, Antonio Torralba, and Sanja Fidler, 2015, Aligning books and movies: Towards story-like visual explanations by watching movies and reading books, in *Proceedings of the IEEE international conference on computer vision*, 19–27.

Figure 1: Ratio of Post-Alice Density to Pre-Alice Density

The figure illustrates whether there is a decrease in applications and grants of patents with high Alice scores after the Supreme Court decision (the final column of Table 7). We compute the density of the Alice score based on ten bins (increments of .1) from zero to unity both before Alice (2011 to 2013) and post Alice (2017). The figure reports the ratio of the density for each bin. The values below unity for the rightmost bins below indicate that many fewer patents with high Alice scores were applied for (Panel A) and granted (Panel B) after the Alice Supreme Court decision.

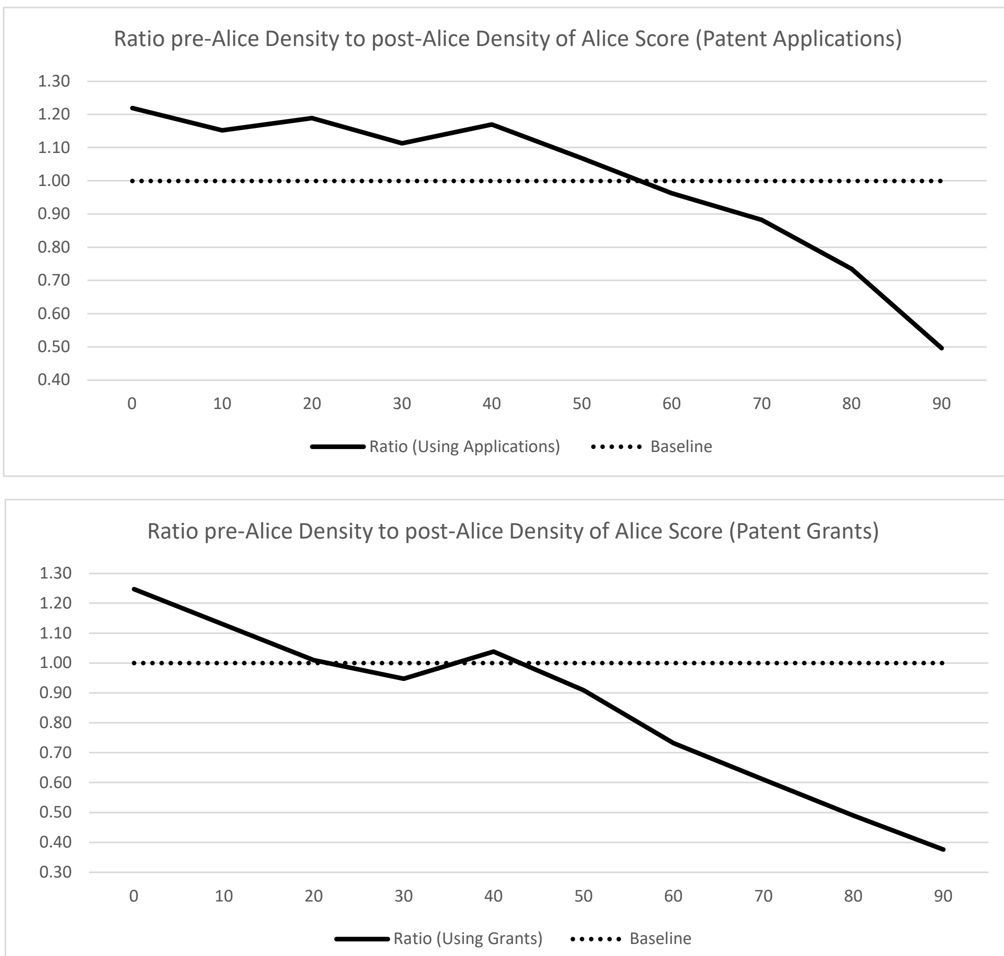
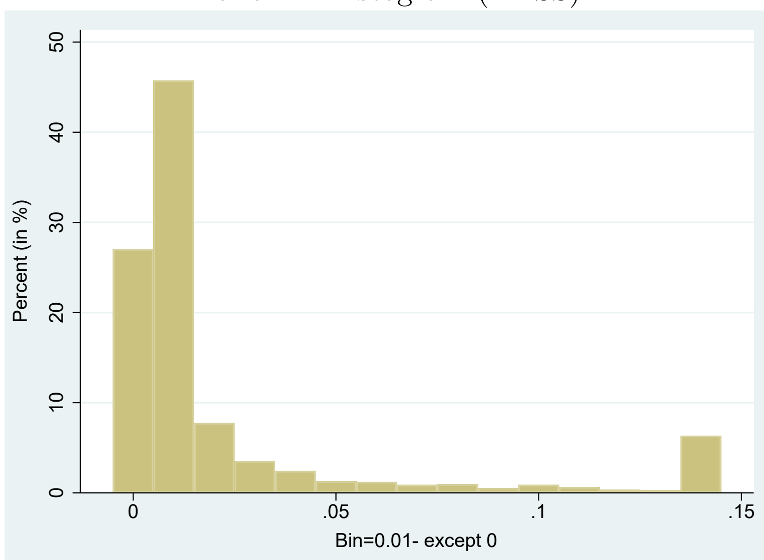


Figure 2: Histogram For Treatment

This figure shows the histogram for the treatment variables for firms with patents. In Panel A, the treatment is based on KPSS, and it is based on citation in Panel B. The bin width is 0.01 and y-axis is the percentage of treatment falls into the bin.

Panel A: Histogram (KPSS)



Panel B: Histogram (Cites)

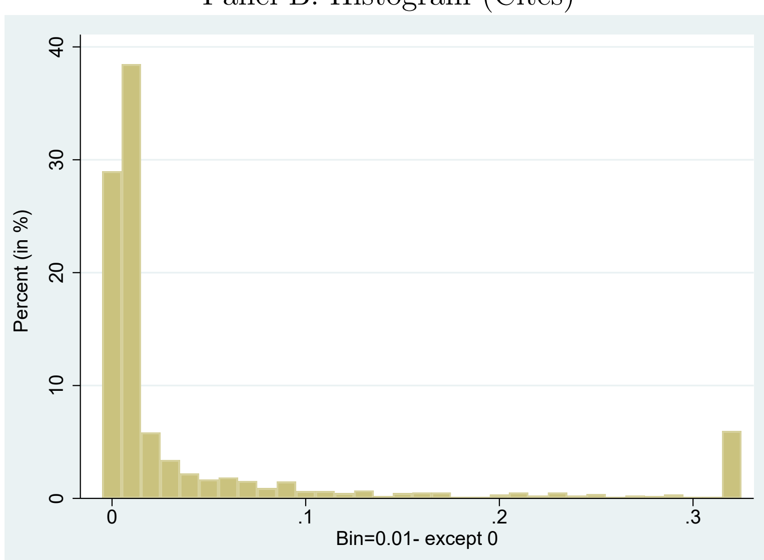
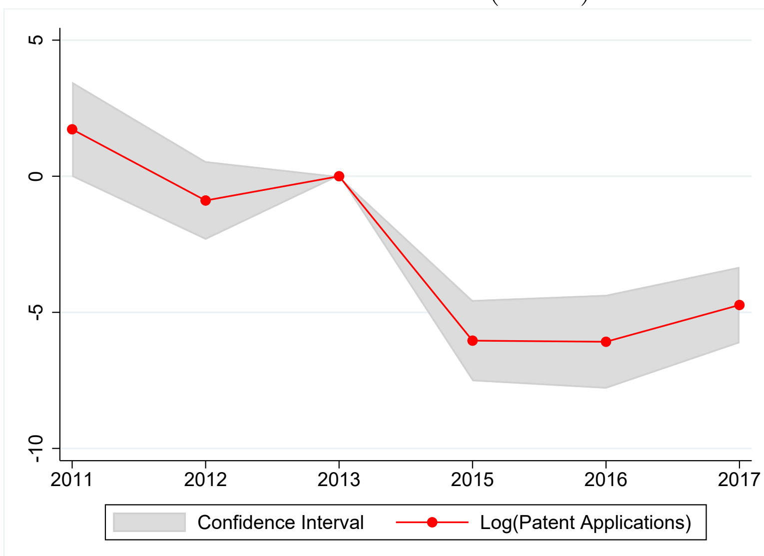


Figure 3: Patent Applications For Small Firms

This figure reports the point estimates per year for $Small \times Treatment$ from Table 9, where the dependent variable is Log(# of Patents). The regression specification is the same as those reported in columns (3) and (4) of Table 9, except that $Small \times Treatment$ is allowed to vary by year, and 2013 is chosen as the reference year. The treatment is calculated by using the KPSS values in Panel A, and it is calculated by citations in Panel B. The gray line indicates the 90% confidence interval.

Panel A: Small Firms (KPSS)



Panel B: Small Firms (Cites)

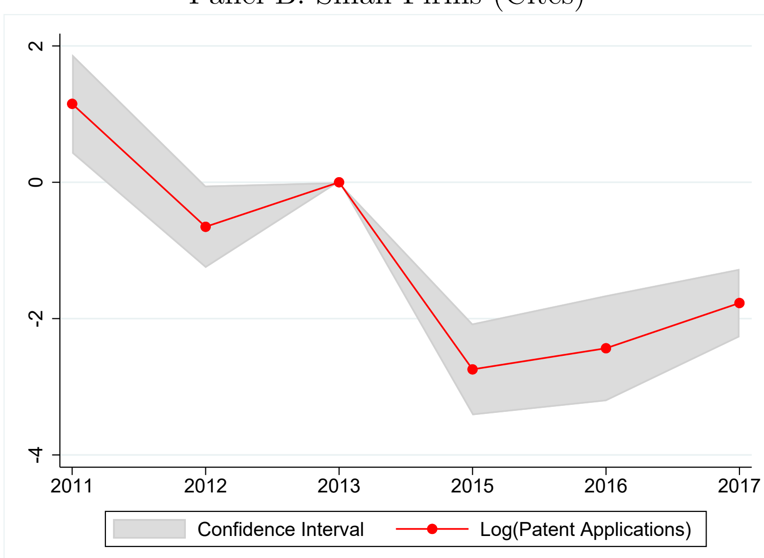
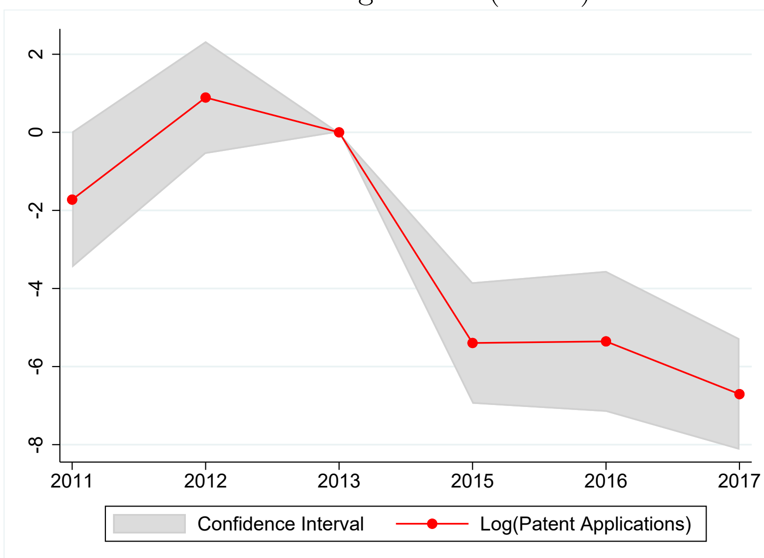


Figure 4: Patent Applications For Large Firms

This figure reports the point estimates per year for $Large \times Treatment$ from Table 9, where the dependent variable is Log(# of Patents). The regression specification is the same as those reported in columns (3) and (4) of Table 9, except that $Large \times Treatment$ is allowed to vary by year, and 2013 is chosen as the reference year. The treatment is calculated by using the KPSS values in Panel A, and it is calculated by citations in Panel B. The gray line indicates the 90% confidence interval.

Panel A: Large Firms (KPSS)



Panel B: Large Firms (Cites)

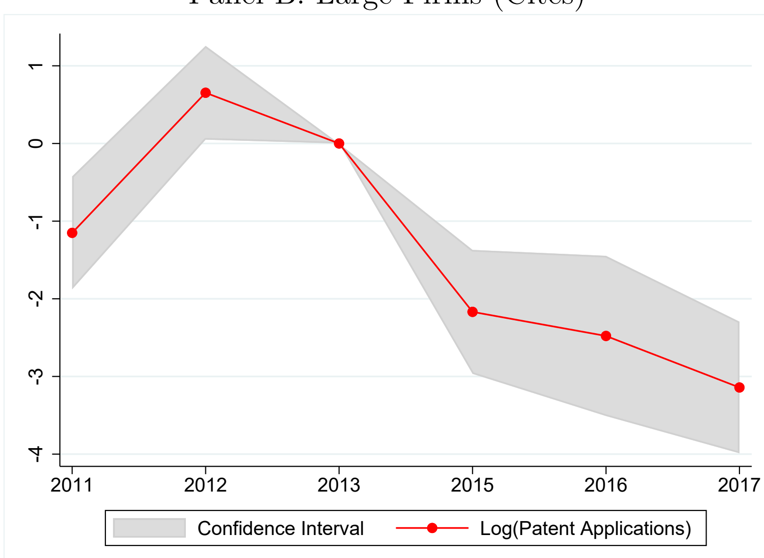
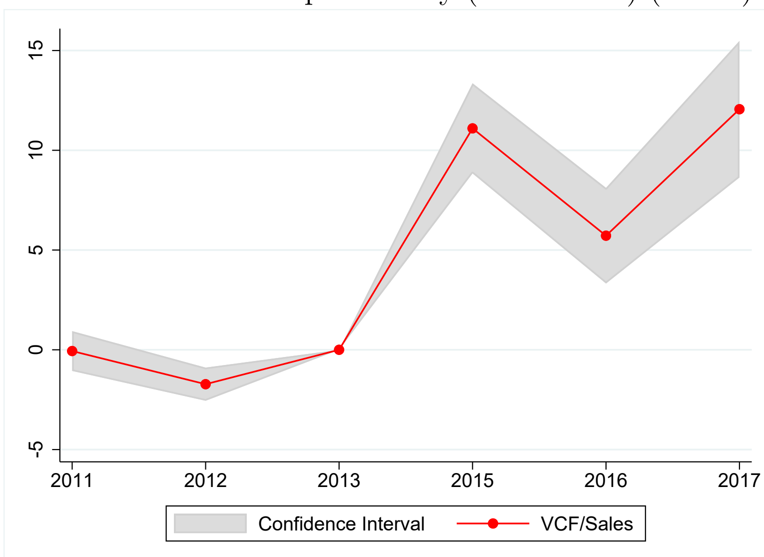


Figure 5: Competition For Small Firms

This figure reports the point estimates per year for $Small \times Treatment$ from Table 10 columns (1) and (2) where the dependent variable is VCF/Sales. The regression specifications are the same as those reported in columns (1) and (2) of Table 10, except that $Small \times Treatment$ is allowed to vary by year, and 2013 is chosen as the reference year. The treatment is calculated by using the KPSS values in Panel A, and it is calculated by citations in Panel B. The gray line indicates the 90% confidence interval.

Panel A: Venture Capital Entry (VCF Score) (KPSS)



Panel B: Venture Capital Entry (VCF Score) (Cites)

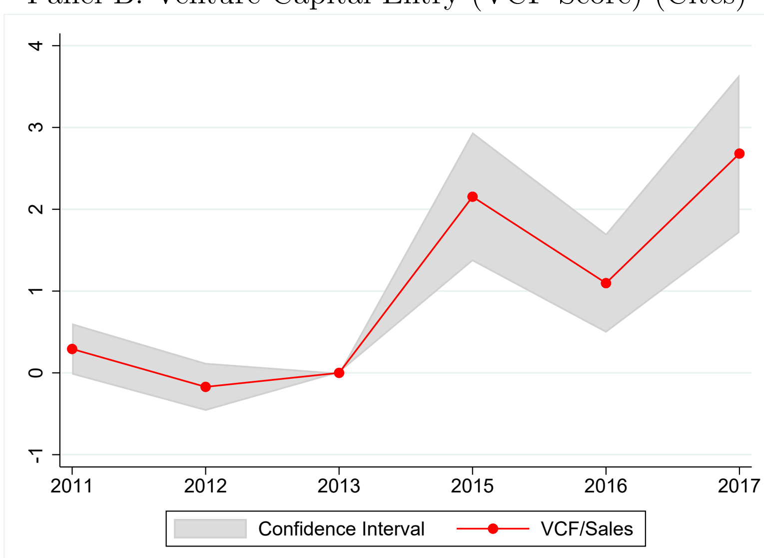
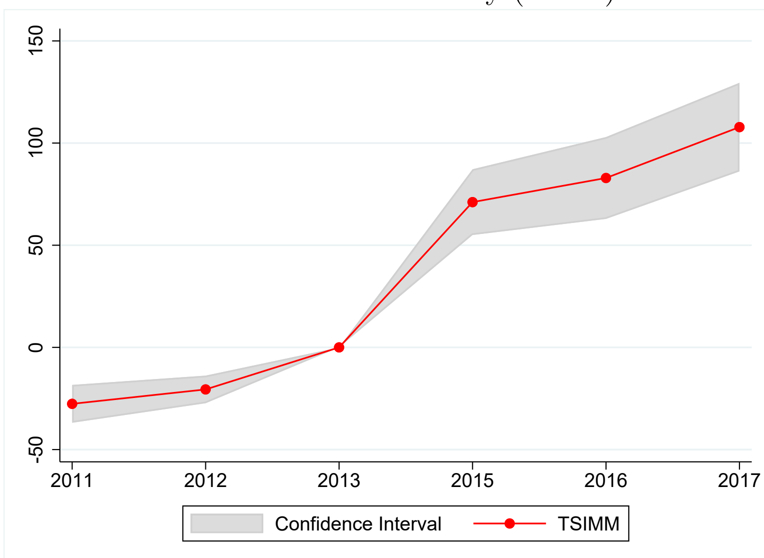


Figure 6: Competition For Small Firms

This figure reports the point estimates per year for $Small \times Treatment$ from Table 10 columns (3) and (4) where the dependent variable is Total Similarity (TSIMM). The regression specifications are the same as those reported in columns (3) and (4) of Table 10, except that $Small \times Treatment$ is allowed to vary by year, and 2013 is chosen as the reference year. The treatment is calculated by using the KPSS values in Panel A, and it is calculated by citations in Panel B. The gray line indicates the 90% confidence interval.

Panel A: Total Similarity (KPSS)



Panel B: Total Similarity (Cites)

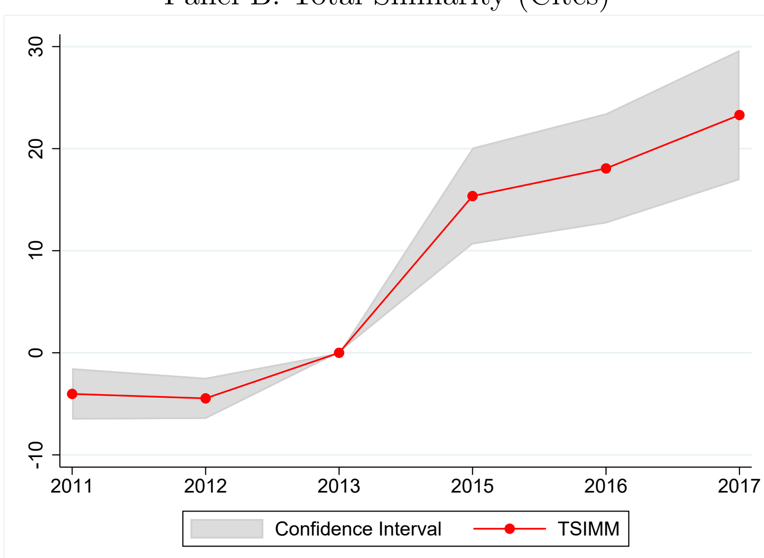


Table 1: Annual Patent Applications and Post-Alice Rejections By Industry

This table reports annual statistics from USPTO patent applications and the corresponding percentage that were rejected in parentheses based on the Supreme Court's Alice decision for the top 12 industries with patent rejections. The rejection data provided by Lu et al. (2017) extends until 2016; therefore ratio of rejection is assigned NA for 2017. *Change* reports the percentage change from the number of patent applications in 2013 to the average number of patent applications for the 2015-2017 period. Corresponding CPCs for each industry are provided in Table 2.

Patent Applications and USPTO Alice Rejections - Top 12 industries

	Number of Patent Applications & Rejection Percentage								
Industry	2008-2009	2010-2011	2012	2013	2014	2015	2016	2017	Change (2013 to 2015-2017)
Commerce	6582	7675	5033	5563	5223	4246	3405	3240	-34.7%
(Data Processing Methods)	(11.7%)	(17.9%)	(29.8%)	(36.2%)	(23.2%)	(6.6%)	(1.5%)	(NA)	
Administration	6681	6250	3658	2958	2970	2500	2527	2568	-14.4%
(Data Processing Methods)	(6.7%)	(11.1%)	(20.8%)	(31.3%)	(16.7%)	(3.6%)	(0.6%)	(NA)	
Finance	2297	2662	1545	1752	1512	1035	775	711	-52.0%
(Data Processing Methods)	(9.4%)	(13.2%)	(22.5%)	(42.1%)	(37.8%)	(8.7%)	(1.9%)	(NA)	
Payment Systems	1603	2043	1673	1946	2182	2157	2029	1895	4.2%
(Data Processing Methods)	(9.9%)	(12.9%)	(26.6%)	(36.7%)	(24.4%)	(5.8%)	(1.9%)	(NA)	
Coin-freed Facilities or Services	2385	1665	1221	1407	1134	980	939	937	-32.3%
(Coin-freed or Like Apparatus)	(3.9%)	(6.8%)	(17.0%)	(34.3%)	(31.2%)	(14.9%)	(6.7%)	(NA)	
Information Retrieval	7981	8451	5850	6566	6650	6339	6196	5816	-6.8%
(Digital Data Processing)	(0.5%)	(1.2%)	(2.4%)	(4.1%)	(5.1%)	(2.0%)	(1.0%)	(NA)	
Video Games	1414	1504	919	1045	1010	781	847	929	-18.4%
(Games)	(4.5%)	(7.0%)	(12.5%)	(27.4%)	(19.6%)	(7.8%)	(3.4%)	(NA)	
Specialized For Sectors	515	918	753	845	881	669	848	806	-8.4%
(Data Processing Methods)	(4.9%)	(10.9%)	(15.5%)	(32.1%)	(19.3%)	(4.5%)	(0.6%)	(NA)	
Computer Security	3886	3926	2617	2684	2641	2604	2675	2872	1.2%
(Digital Data Processing)	(1.6%)	(1.5%)	(2.6%)	(5.0%)	(5.2%)	(4.0%)	(0.7%)	(NA)	
Network Security	3522	3208	2206	2864	3433	4042	4124	3817	39.5%
(Transmission of Digital Information)	(0.8%)	(0.8%)	(1.9%)	(4.3%)	(5.5%)	(3.4%)	(0.8%)	(NA)	
Network Specific Applications	3389	3441	2282	2891	3174	3172	3098	2414	0.1%
(Transmission of Digital Information)	(0.8%)	(1.5%)	(3.2%)	(6.0%)	(4.7%)	(2.2%)	(0.6%)	(NA)	
Measuring or Testing Processes	3759	4311	2237	2356	2336	2105	2099	2082	-11.1%
(Microbiology & Enzymology)	(1.3%)	(2.5%)	(4.3%)	(4.9%)	(3.2%)	(2.6%)	(0.6%)	(NA)	

Table 2: CPC Descriptions by CPC group and Industry

This table provides descriptions for largest CPC patent subgroups for which we run the Longformer patent rejection models. We also give the larger industry correspondence for the main CPC groups impacted by the Alice decision.

Panel A: CPC Main/Sub Group Descriptions

CPC Main/Sub Group	Description
G06Q10/06	Administration; Management-Resources, workfSmalls, human or project management, e.g. organising, planning, scheduling or allocating time, human or machine resources; Enterprise planning; Organisational models
G06Q10/10	Administration; Management-Office automation, e.g. computer aided management of electronic mail or groupware ; Time management, e.g. calendars, reminders, meetings or time accounting
G06Q30/02	Commerce, e.g. shopping or e-commerce-Marketing, e.g. market research and analysis, surveying, promotions, advertising, buyer profiling, customer management or rewards; Price estimation or determination
G06Q30/06	Commerce, shopping or e-commerce-Buying, selling or leasing transactions
G06Q30/0631	Commerce, shopping or e-commerce-Buying, selling or leasing transactions-Electronic shopping-Item recommendations
G06Q30/08	Commerce, shopping or e-commerce-Buying, selling or leasing transactions Auctions; matching or brokerage
G06Q40/00	Finance; Insurance; Tax strategies; Processing of corporate or income taxes
G06Q40/02	Finance; Insurance; Tax strategies; Processing of corporate or income taxes-Banking, e.g. interest calculation, credit approval, mortgages, home banking or on-line banking
G06Q40/04	Finance; Insurance; Tax strategies; Processing of corporate or income taxes-Exchange, e.g. stocks, commodities, derivatives or currency
G06Q40/06	Finance; Insurance; Tax strategies; Processing of corporate or income taxes-Investment, e.g. financial instruments, portfolio management or fund management
G06Q40/08	Finance; Insurance; Tax strategies; Processing of corporate or income taxes-Insurance, e.g. risk analysis or pensions
G07F17/32	Coin-freed apparatus for hiring articles; Coin-freed facilities or games, toys, sports or amusements, casino games, online gambling

Panel B: Industries and Corresponding CPC Groups

Industry	CPC Group
Chemical & Physical Properties (Analyzing Materials)	G01N33
Coin-freed or Like Apparatus (Coin-freed Facilities or Services)	G07F17
Data Processing Methods (Administration)	G06Q10
Data Processing Methods (Commerce)	G06Q30
Data Processing Methods (Finance)	G06Q40
Data Processing Methods (Payment Systems)	G06Q20
Data Processing Methods (Specialized For Sectors)	G06Q50
Diagnosis, Surgery, Identification (Measuring for Diagnostic Purpose)	A61B5
Digital Data Processing (Arrangements for Program Control)	G06F9
Digital Data Processing (Computer Aided Design)	G06F30
Digital Data Processing (Computer Security)	G06F21
Digital Data Processing (I/O Arrangements for Data Transfer)	G06F3
Digital Data Processing (Information Retrieval)	G06F16
Digital Data Processing (Natural Language Processing)	G06F40
Games (Video Games)	A63F13
Graphical Data Reading (Recognizing Patterns)	G06K9
Microbiology & Enzymology (Measuring or Testing Processes)	C12Q1
Photogrammetry or Videogrammetry (Navigation)	G01C21
Pictorial Communication (Selective Content Distribution)	H04N21
Transmission of Digital Information (Network Security)	H04L63
Transmission of Digital Information (Network Specific Applications)	H04L67
Transmission of Digital Information (User-to-user Messaging)	H04L51

Source: https://patentsview.org/download/data-download-tables

Table 3: Comparison of Predictions For Longformer vs. Other Models

This table compares predictions of Longformer (Beltagy et al. (2020)), SciBERT (Beltagy et al. (2019)), BERT (Devlin et al. (2019)), RoBERTa (Liu et al. (2019)), TF-IDF (Robertson (2004)) and Word2Vec models (Mikolov et al. (2013)) based on F_1 Score and Accuracy. F_1 Score is the harmonic mean of recall and precision, which are defined in Equation (1). Accuracy is the ratio of correctly predicted observations to the total observations. For all models, we conduct four experiments in which the only difference is the way we create the training samples. In experiment A, for each of the 23,734 positives (rejected patent applications), we find a matching negative (a patent application that is granted) that is in the same CPC Group. In sample B, C, and D, without replacement, we keep adding 23,734 more matching granted patents to the negatives pool based on CPC Subclass, Class, and Section respectively. Therefore, from A to D, each sample has 23,734 more negatives but the newly added ones are selected from a broader CPC. In the last column of the table, we use an ensemble of the two models that have the highest F_1 Score and Accuracy by taking the average of their prediction scores. The use of this ensemble is motivated by the fact that A typically has the highest F_1 Score and D has the highest Accuracy. For the testing, we only use applications and granted patents not used in the training. In the testing, we have 10,000 positives in the sample of applications and for each positive, we choose two negatives. This 1:2 positives to negatives ratio is consistent with with the expected rejection ratio having more negatives than positives, while not overestimating accuracy for models that do not learn but only predict negative outcomes. From the negatives pool, we thus sample 20,000 negatives 1,000 times and boot-strap performance of the models. The table below then reports the average F_1 Score and Accuracy for each model.

	A		B		C		D		$\frac{A+D}{2}$	
Model Name	F_1 Score	Accuracy	F_1 Score	Accuracy	F_1 Score	Accuracy	F_1 Score	Accuracy	F_1 Score	Accuracy
Longformer Finetune	0.647	0.745	0.624	0.765	0.618	0.785	0.639	0.800	0.672	0.804
SciBERT Finetune	0.651	0.735	0.634	0.749	0.632	0.767	0.638	0.777	0.669	0.778
BERT Finetune	0.623	0.733	0.598	0.739	0.614	0.764	0.624	0.774	0.642	0.775
RoBERTa Finetune	0.600	0.716	0.555	0.740	0.540	0.756	0.515	0.758	0.592	0.765
TF-IDF + Logistic Regression	0.547	0.643	0.599	0.634	0.613	0.670	0.550	0.719	0.559	0.679
TF-IDF + Decision Tree	0.503	0.602	0.554	0.552	0.558	0.584	0.491	0.690	0.409	0.697
TF-IDF + Random Forest	0.628	0.743	0.368	0.717	0.263	0.696	0.209	0.689	0.387	0.723
Word2Vec + Logistic Regression	0.606	0.731	0.418	0.732	0.377	0.732	0.358	0.730	0.497	0.755
Word2Vec + Decision Tree	0.492	0.607	0.456	0.645	0.461	0.687	0.461	0.702	0.365	0.707
Word2Vec + Random Forest	0.619	0.747	0.439	0.746	0.387	0.739	0.365	0.735	0.500	0.766

Table 4: Summary of Longformer Prediction Statistics

This table reports statistics from Longformer model predictions for the set of patents that are examined for lost protection. A patent is included in the examination set if it is granted between 06/19/1994 and 06/19/2014 and shares the same primary CPC with at least one of the applications that were rejected by the USPTO based on the Alice decision. Panel A reports the frequency statistics from different thresholds for the 642,678 patents that fit to the examination criteria. In the default model, the threshold of 0.5 is used. The Panel B documents the most frequent primary CPCs for patent applications rejected by the USPTO, and for patents that have Larger than 0.5 as the Longformer score. In our sample, 111,420 patents that have Larger than Longformer Score of 0.5 have 114,885 primary CPCs. Panel C provides short descriptions for the most frequent CPCs.

Panel A: Longformer Predictions For Different Thresholds

Threshold	Percentage of Patents $\geq$ Threshold (%)	Number of Patents $\geq$ Threshold	Number of Unique CPCs
0.5	17.34	111,420	4,979
0.6	11.50	73,934	4,591
0.7	8.87	57,001	4,316
0.8	6.72	43,200	3,980
0.9	4.32	27,786	3,407

Panel B: Summary of CPCs For Alice Rejections and Longformer Predictions by CPC group

Alice Rejections (For Patent Applications)			**Longformer Predictions (For Granted Patents)**		
Most Frequent CPCs	Count	Percentage(%)	Most Frequent CPCs	Count	Percentage(%)
G06Q30/02	1185	3.49	G06Q30/02	2898	2.52
G06Q40/04	675	1.99	G06Q10/10	2133	1.86
G06Q10/06	486	1.43	G06Q10/06	1992	1.73
G06Q40/08	397	1.17	G06Q30/06	1638	1.43
G06Q40/06	383	1.13	G06Q40/04	1563	1.36
G06Q10/10	370	1.09	G06Q40/02	1381	1.20
G06Q30/06	343	1.01	G06Q40/06	865	0.75
G06Q40/02	293	0.86	G07F17/32	841	0.73
G06Q30/0631	248	0.73	G06Q40/00	753	0.66
G06Q30/08	247	0.73	G06Q40/08	717	0.62

Table 5: Patent Grants and Predicted Longformer Rejection Statistics By Industry

This table displays the total number of patents granted in each industry that have a high percentage of patents predicted to be rejected by our Longformer model. The years in the table start from 19th of June and end on 18th of June. The numbers in parentheses show the percentage of patents in that industry and period with a Longformer score of 0.5 or larger. Corresponding CPCs for each industry are provided in Table 2.

Patent Grants and Predicted Longformer Rejections

	Number of Patent Grants & Ratio of Longformer Cases ($\geq$ 0.5)			
Industry	1994-1999	1999-2004	2004-2009	2009-2014
Commerce (Data Processing Methods)	355 (52.7%)	1460 (53.2%)	3536 (52.3%)	10389 (50.5%)
Administration (Data Processing Methods)	665 (50.1%)	2001 (45.1%)	4447 (42.2%)	11467 (40.1%)
Finance (Data Processing Methods)	204 (68.1%)	473 (66.8%)	1253 (65.1%)	6387 (66.2%)
Payment Systems (Data Processing Methods)	263 (37.3%)	565 (37.7%)	1175 (35.0%)	3411 (43.8%)
Coin-freed Facilities or Services (Coin-freed or Like Apparatus)	445 (32.6%)	1126 (37.7%)	1483 (36.9%)	4486 (38.4%)
Information Retrieval (Digital Data Processing)	1238 (20.1%)	3823 (17.1%)	5894 (14.3%)	15811 (15.2%)
Video Games (Games)	336 (25.6%)	912 (33.6%)	708 (29.5%)	2598 (31.6%)
Specialized For Sectors (Data Processing Methods)	21 (61.9%)	72 (44.4%)	220 (33.6%)	936 (38.0%)
Computer Security (Digital Data Processing)	509 (27.5%)	1176 (24.7%)	2965 (22.0%)	8659 (22.0%)
Network Security (Transmission of Digital Information)	242 (26.4%)	1109 (23.2%)	3742 (20.5%)	9003 (22.7%)
Network Specific Applications (Transmission of Digital Information)	98 (28.6%)	950 (17.9%)	2943 (14.5%)	7565 (18.8%)
Measuring or Testing Processes (Microbiology & Enzymology)	1369 (9.2%)	2107 (13.4%)	1887 (16.1%)	3749 (24.2%)

Table 6: Most Frequently Used Words in Longformer Predictions

The table lists words that are used mostly frequently in patents with high Longformer scores (≥ 0.5) compared to those with low Longformer scores (<0.5). We first label patents with a Longformer score ≥ 0.5 as "high" and the remaining patents as "low". We remove non-alphabetic characters from patent texts, apply lemmatizing to each word, and calculate the number of high and low patents that each word appears in. We then filter out words that do not appear in at least 1% of the high patents. For each word w, we first assign it to the CPC Group with the Largest ratio of the number high patents in the CPC that contain the word to the total number of low patents in that CPC. Finally, we sort the words selected into each CPC according to their appearance ratio, defined as $\frac{Count_w^H}{1+Count_w^L}$, where $Count_w^H$ and $Count_w^L$ are high and low number of patents a word w appears in, respectively. The table reports the top 15 words sorted according to their appearance ratio.

Industry	Top Fifteen Words
Commerce (Digital Data Processing)	rebate, bidder, bidding, seller, auction, discounted, sponsor, referral, incentive, purchaser, solicitation, purchasing, solicit
Administration (Digital Data Processing)	interview, consultant, procurement, forecasting, accountability, contractor, consultation, planner, deadline, strategic, forecast, audit, objectively, finalized, logistics
Finance (Digital Data Processing)	underwriting, liquidity, lender, financing, equity, investor, treasury, debt, hedge, earnings, earning, owed, investing, insurer, mortgage
Payment Systems (Digital Data Processing)	settlement, refund, debited, credited, clearinghouse, transacting, approving, dispute, crediting, enroll, deducted, debiting, ach, paying, approves
Coin-freed Facilities or Services (Coin-freed or Like Apparatus)	rewarded, earn, payouts, payoff, redeem, earned, redeemed, redemption, awarding, betting, dealer, profitability, payout, wagered, wager
Information Retrieval (Digital Data Processing)	vowel,phoneme, docket, adjective, ranked, spelling, noun, categorizing, categorization, linguistic, verb, vocabulary, alphabetical, sentence, utterance, searchable
Video Games (Games)	opponent, contest, fun, participated, team, him, town, vote, himself, herself, war, personality, story, thinking, arena
Specialized For Sectors (Digital Data Processing)	hire, attorney, affiliate, reputation, qualification, prospective, pursue, teacher, education, affiliation, court, posting, historic, invited, submitting
Computer Security (Digital Data Processing)	netlist, royalty, licensing, confidential, licensed, license, denied, verilog, denies, creator, vhdl, unlimited, privilege, enforcing, granting
Network Security (Transmission of Digital Information)	certification, certified, certificate, confidentiality, signing, logon, expire, password, signed, violation, username, privacy, violate, someone, denial
Network Specific Applications (Transmission of Digital Information)	publish, subscription, cookie, subscribing, uploaded, uploads, publishes, apache, wap, subscribe, downloading, downloads, movie, activex, url
Measuring or Testing Processes (Microbiology & Enzymology)	institutional, enrolled, lifestyle, consent, smoking, logistic, percentile, multivariate, gender, emotional, exam, younger, college, whom, disability

Table 7: Comparison of Patent Grants Alice Longformer Scores in the Pre- and Post-Period

This table shows the distributional density of the Longformer Score before the Alice shock (2011 to 2013) and after the shock (2017) for the Top 20 technological areas impacted by Alice. To compute the density in a given year, we first identify, the set of patents granted in that year in the Top 20 technological areas. The number of granted patents in 2011, 2012, and 2013 are 21,404; 26,607; and 31,249, respectively. In 2017, we only consider the patents applied for after the Alice decision, and there are 17,643 granted patents that fit to this description. We sort all patents in each year into 10 bins based on each patent's Longformer Score. Bins are defined as the ten equal segments in the interval (0,1), which is the range of the Longformer Score. For each bin, the density is the number of patents in the given bin in the given year divided by the total number of patents in the given year. Finally, to illustrate the impact of Alice on these density distributions, we compute the Ratio in the final column as the density in 2017 divided by the average pre-Alice density from years 2011 to 2013. A ratio below unity indicates that the rate of patenting in the given bin declined post-Alice.

Longformer Score (LFS)	2011 (1)	2012 (2)	2013 (3)	2011- 2013 (4)	2017 (5)	Ratio (6)
$0.0 \leq$ LFS < 0.1	0.3627	0.3619	0.3775	0.3683	0.4595	1.2476
$0.1 \leq$ LFS < 0.2	0.1032	0.1017	0.1017	0.1021	0.1152	1.1283
$0.2 \leq$ LFS < 0.3	0.0749	0.0742	0.0715	0.0733	0.0740	1.0095
$0.3 \leq$ LFS < 0.4	0.0734	0.0714	0.0688	0.0709	0.0672	0.9478
$0.4 \leq$ LFS < 0.5	0.1048	0.1080	0.1062	0.1064	0.1105	1.0385
$0.5 \leq$ LFS < 0.6	0.0751	0.0757	0.0763	0.0758	0.0689	0.9090
$0.6 \leq$ LFS < 0.7	0.0413	0.0426	0.0406	0.0414	0.0303	0.7319
$0.7 \leq$ LFS < 0.8	0.0361	0.0374	0.0364	0.0367	0.0224	0.6104
$0.8 \leq$ LFS < 0.9	0.0434	0.0427	0.0421	0.0427	0.0209	0.4895
$0.9 \leq$ LFS < 1.0	0.0850	0.0843	0.0789	0.0824	0.0310	0.3762

Table 8: Firm Summary Statistics

This table provides summary statistics for our sample of public firms based on annual firm observations from 2011 to 2017 (excluding 2014, the treatment year). All variables are described in detail in the variable list in Appendix A and in Section 3 of the paper. In Panel D, firm characteristics are based on the values in 2013. Small Treatment and Large Treatment firms are the ones which have treatment scores that are below and above the median, respectively. *, **, and *** denote significant difference of the mean post-Alice vs. pre-Alice at the 10%, 5% and 1% level.

			Pre-Alice			**Post-Alice**			
Variable	N	# of Firms	Median	Mean	Std. Error	Median	Mean	Std. Error	Diff (Post-Pre)
Panel A: Firm Characteristics									
Assets (in mil.)	19372	3444	949.376	5351.245	214.347	1294.420	6842.200	268.894	***
Sales (in mil.)	19372	3444	443.011	2857.082	119.404	564.108	3012.597	115.613	
OI/Sales	18518	3296	0.144	-0.042	0.018	0.142	-0.143	0.027	***
Tobin's Q	19254	3436	1.120	1.536	0.023	1.163	1.520	0.021	
Sales Growth	19251	3403	0.065	0.088	0.003	0.038	0.034	0.003	***
Age	19372	3444	18.000	21.577	0.276	22.000	25.371	0.278	***
Panel B: Innovation, Acquisition & Lawsuit Characteristics									
R&D/Sales	19372	3444	0.000	0.118	0.007	0.000	0.158	0.012	***
Log(# of Patents)	19372	3444	0.000	0.585	0.020	0.000	0.518	0.019	**
Patents/Sales	19372	3444	0.000	0.014	0.001	0.000	0.008	0.000	***
Acquisitions/Sales	19372	3444	0.000	0.056	0.002	0.000	0.055	0.002	
PatTargets/Sales	19372	3444	0.000	0.002	0.000	0.000	0.001	0.000	***
Log(Amt. of Acq.)	19372	3444	0.000	0.788	0.022	0.000	0.762	0.023	
# Alleged	19372	3444	0.000	0.279	0.013	0.000	0.226	0.010	***
# NPE Alleged	19372	3444	0.000	0.163	0.008	0.000	0.154	0.007	
# OC Alleged	19372	3444	0.000	0.101	0.005	0.000	0.048	0.003	***
# Accuser	19372	3444	0.000	0.035	0.002	0.000	0.031	0.002	
IPrisk (10-K)	19289	3444	0.614	2.931	0.070	1.247	3.456	0.078	***
Patinfringe (10-K)	19289	3444	0.000	1.357	0.043	0.000	1.311	0.040	
Panel C: Competition Measures (Text-based measures from 10-Ks)									
VCF/Sales	19286	3444	0.014	0.135	0.006	0.012	0.225	0.014	***
TSIMM	19268	3442	2.231	9.405	0.268	2.134	9.619	0.262	
Complaints	19289	3444	13.521	14.583	0.116	13.622	14.729	0.115	
Noncompete	19289	3444	0.000	0.595	0.022	0.000	0.555	0.021	
Nondisclose	19289	3444	0.000	0.489	0.031	0.000	0.610	0.041	**

Table 9:
Patents and R&D (Longformer)

The table displays panel data regressions in which innovation and R&D measures are dependent variables. In columns (1)-(2), the dependent variable is the number of patent applications in that year divided by sales; and in columns (3) and (4), it is one plus log of the number of patent applications in the respective year. In columns (5)-(6), the dependent variable is R&D expenses scaled by sales. *Treatment* is the relative impact of the Alice decision on the firm's patent portfolio measured using KPSS dollar values or citations scaled by sales (see equation (3) for the formula). In the odd and even numbered columns, respectively, we use the KPSS and the number of citations approach to compute the *Patent Value* treatment. *Small* is a binary variable equal to one if a firm's total assets are smaller than the median total asset of its TNIC peers in 2013, and it is zero otherwise. *Large* is 1-*Small*. *Post* is a dummy variable that equals one if the year is after the Alice decision and zero otherwise. All variables are described in detail in the variable list in Appendix A. All regressions include firm and year fixed effects. Standard errors are clustered at the firm level. T-statistics are reported in parentheses; *, **, and *** denote significance at the 10%, 5% and 1% level.

Dependent Variable:	$\frac{\#\ of Patents}{Sales}$		Log(# of Patents)		$\frac{R\&D}{Sales}$	
	(1)	(2)	(3)	(4)	(5)	(6)
Small × Post × Treatment	-0.499***	-0.151***	-2.108***	-0.835***	6.646***	1.373***
	(-8.90)	(-7.70)	(-3.19)	(-3.36)	(4.51)	(3.69)
Large × Post × Treatment	-0.138***	-0.102***	-2.796***	-1.463***	-0.198	-0.233
	(-4.23)	(-3.94)	(-3.50)	(-2.83)	(-0.91)	(-1.19)
Log(Sales)	-0.010***	-0.010***	0.036***	0.036***	-0.092***	-0.095***
	(-7.75)	(-7.67)	(3.58)	(3.47)	(-6.25)	(-6.32)
Log(Age)	-0.000	-0.001	0.000	-0.002	0.140***	0.161***
	(-0.05)	(-0.50)	(0.00)	(-0.05)	(3.67)	(4.10)
Observations	19372	19372	19372	19372	19372	19372
Treatment Calculation	KPSS	Citation	KPSS	Citation	KPSS	Citation
Adj. R^2	0.101	0.098	0.049	0.048	0.078	0.050

Table 10:
Competition and Patent Protection (Longformer)

The table displays panel data regressions in which competition variables are the dependent variables. In columns (1)-(2), the dependent variable, *VCF/Sales*, is the a measure of VC entry in a given firm's product market and is the total first-round dollars raised by the 25 startups from Venture Expert whose Venture Expert business description most closely matches the 10-K business description of the focal firm (using cosine similarities), scaled by focal firm sales. *TSIMM* is the firm's TNIC text-based total similarity of the firm to public firm competitors. Complaints is the number of paragraphs in the firm's 10-K that complain about competition divided by the total number of paragraphs in the firm's 10-K. *Treatment* is the relative impact of the Alice decision on the firm's patent portfolio measured using KPSS dollar values or citations scaled by sales (see equation (3) for the formula). In the odd and even numbered columns, respectively, we use the KPSS and the number of citations approach to compute the *Patent Value* treatment. *Small* is a binary variable equal to one if a firm's total assets are smaller than the median total asset of its TNIC peers in 2013, and it is zero otherwise. *Large* is 1-*Small*. *Post* is a dummy variable that equals one if the year is after the Alice decision and zero otherwise. All variables are described in detail in the variable list in Appendix A. All regressions include firm and year fixed effects. Standard errors are clustered at the firm level. T-statistics are reported in parentheses; *, **, and *** denote significance at the 10%, 5% and 1% level.

Dependent Variable:	$\frac{VCF}{Sales}$		TSIMM		Complaints	
	(1)	(2)	(3)	(4)	(5)	(6)
Small × Post × Treatment	10.635***	2.173***	102.649***	21.852***	13.780***	3.840***
	(7.72)	(5.40)	(9.40)	(6.68)	(3.09)	(2.64)
Large × Post × Treatment	0.326*	0.152	8.951**	2.158	-7.025*	-2.102
	(1.89)	(1.48)	(2.19)	(0.93)	(-1.76)	(-0.90)
Log(Sales)	-0.451***	-0.455***	0.893***	0.854***	0.123	0.118
	(-15.17)	(-15.26)	(5.32)	(4.84)	(1.35)	(1.29)
Log(Age)	0.387***	0.420***	-0.147	0.160	-0.258	-0.217
	(6.43)	(6.89)	(-0.32)	(0.33)	(-0.81)	(-0.69)
Observations	19286	19286	19268	19268	19289	19289
Firm Fixed Effects	YES	YES	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES	YES	YES
Treatment Calculation	KPSS	Citation	KPSS	Citation	KPSS	Citation
Adj. R^2	0.265	0.232	0.096	0.057	0.008	0.007

Table 11:

Firm Level Competition and Encroachment (Longformer)

The table displays firm-pair-year panel data regressions in which pairwise product market encroachment (Delta TNIC Score) is the dependent variable. Delta TNIC Score is computed the change in pairwise TNIC similarity (see Hoberg and Phillips 2016) from year t-1 to year t. A large value indicates increased similarity and product market encroachment. To compute the RHS variables, we first sort firms into above and below median sales (relative to TNIC-2 peers) in 2013. We denote the two firms associated with each pairwise observation as 1 and 2. The variable Treat1 (Treat2) is the Alice Score for firm 1 (2). Analogously, Big1 is an indicator if firm 1's total assets are larger than the median total asset of its TNIC peers in 2013, and is zero otherwise. Small1 indicates firm 1 has assets that are below the median value. Relative size indicators are similarly defined for firm 2. Please note that all level effects and Smaller-order interactions are subsumed by the fixed effects and thus are not reported. All regressions include firm-pair and year fixed effects and standard errors are clustered at the firm-pair level. T-statistics are reported in parentheses; *, **, and *** denote significance at the 10%, 5% and 1% level.

Dependent Variable:	Delta TNIC Score		
	(1)	(2)	(3)
Treat1 × Post	0.477***		
	(13.57)		
Treat2 × Post	0.477***		
	(13.57)		
Big1 × Treat1 × Post		-0.439***	
		(-9.56)	
Small1 × Treat1 × Post		1.673***	
		(34.04)	
Big2 × Treat2 × Post		-0.439***	
		(-9.56)	
Small2 × Treat2 × Post		1.673***	
		(34.04)	
Big1 × Big2 × Treat1 × Post			-0.382***
			(-6.07)
Big1 × Small2 × Treat1 × Post			-0.534***
			(-8.31)
Small1 × Big2 × Treat1 × Post			1.776***
			(27.81)
Small1 × Small2 × Treat1 × Post			1.592***
			(21.66)
Big1 × Big2 × Treat2 × Post			-0.382***
			(-6.07)
Small1 × Big2 × Treat2 × Post			-0.534***
			(-8.31)
Big1 × Small2 × Treat2 × Post			1.776***
			(27.81)
Small1 × Small2 × Treat2 × Post			1.592***
			(21.66)
Observations	13,448,224	13,448,224	13,448,224
Pair Fixed Effects	YES	YES	YES
Year Fixed Effects	YES	YES	YES
R^2	0.92	0.092	0.092

Table 12:
Profitability (Longformer)

The table displays panel data regressions that examine whether the profitability of large and small firms were differently affected by the Alice decision. In columns (1)-(2), the dependent variable is sales growth, calculated as the natural logarithm of total sales in the current year t divided by total sales in the previous year t-1.; and in columns (3) and (4), it is operating income scaled by sales. In columns (5)-(6), the dependent variable is Tobin's Q, calculated as the market to book ratio (market value of equity plus book debt and preferred stock, all divided by book assets). *Treatment* is the relative impact of the Alice decision on the firm's patent portfolio measured using KPSS dollar values or citations scaled by sales (see equation (3) for the formula). In the odd and even numbered columns, respectively, we use the KPSS and the number of citations approach to compute the *Patent Value* treatment. *Small* is a binary variable equal to one if a firm's total assets are smaller than the median total asset of its TNIC peers in 2013, and it is zero otherwise. *Large* is 1-*Small*. *Post* is a dummy variable that equals one if the year is after the Alice decision and zero otherwise. All variables are described in detail in the variable list in Appendix A. All regressions include firm and year fixed effects. Standard errors are clustered at the firm level. T-statistics are reported in parentheses; *, **, and *** denote significance at the 10%, 5% and 1% level.

Dependent Variable:	Sales Growth		$\frac{Operating\ Income}{Sales}$		Tobin's Q	
	(1)	(2)	(3)	(4)	(5)	(6)
Small × Post × Treatment	0.513	0.041	-11.361***	-2.386***	-8.987***	-2.461***
	(1.20)	(0.38)	(-3.73)	(-3.10)	(-2.87)	(-2.83)
Large × Post × Treatment	0.475***	0.325***	0.655	0.602	0.957	1.229
	(3.02)	(3.07)	(1.26)	(1.27)	(0.77)	(1.48)
Log(Sales)	-0.208***	-0.208***	0.360***	0.365***	-0.344***	-0.342***
	(-27.33)	(-27.35)	(7.82)	(7.92)	(-6.57)	(-6.53)
Log(Age)	-0.017	-0.015	-0.538***	-0.577***	-1.093***	-1.116***
	(-0.81)	(-0.72)	(-5.10)	(-5.43)	(-7.01)	(-7.13)
Observations	19251	19251	18518	18518	18874	18874
Firm Fixed Effects	YES	YES	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES	YES	YES
Treatment Calculation	KPSS	Citation	KPSS	Citation	KPSS	Citation
Adj. R^2	0.172	0.172	0.067	0.055	0.074	0.073

Table 13:
Firm IP Risk and Legal Protections (Longformer)

The table displays panel data regressions examining the impact of Alice on intellectual property and noncompete and disclosure clauses. In columns (1)-(2), *IP Risk*, is the total number of paragraphs mentioning "intellectual property" in the risk factor section in the 10-K documents, scaled by the total number paragraphs in the 10-Ks. *Noncompete* is the total number of 10K paragraphs mentioning "non-compete" agreements, all scaled by the total paragraphs in the 10-K. *Nondisclosure* is the total number of 10-K paragraphs mentioning "non-disclose" or "NDA" agreements, all scaled by the total paragraphs in the 10-K. *Treatment* is the relative impact of the Alice decision on the firm's patent portfolio measured using KPSS dollar values or citations scaled by sales (see equation (3) for the formula). In the odd and even numbered columns, respectively, we use the KPSS and the number of citations approach to compute the *Patent Value* treatment. *Small* is a binary variable equal to one if a firm's total assets are smaller than the median total asset of its TNIC peers in 2013, and it is zero otherwise. *Large* is 1-*Small*. *Post* is a dummy variable that equals one if the year is after the Alice decision and zero otherwise. All variables are described in detail in the variable list in Appendix A. All regressions include firm and year fixed effects. Standard errors are clustered at the firm level. T-statistics are reported in parentheses; *, **, and *** denote significance at the 10%, 5% and 1% level.

Dependent Variable:	IP Risk		Noncompete		Nondisclosure	
	(1)	(2)	(3)	(4)	(5)	(6)
Small × Post × Treatment	11.675***	3.532***	1.083	0.398	14.783***	2.356***
	(4.07)	(4.21)	(1.57)	(1.30)	(4.73)	(3.38)
Large × Post × Treatment	2.611	1.161	-0.858	-0.307	-0.357	-0.537
	(1.02)	(0.70)	(-1.32)	(-0.98)	(-0.55)	(-1.22)
Log(Sales)	0.028	0.027	0.044**	0.044**	0.004	-0.004
	(0.55)	(0.53)	(2.21)	(2.19)	(0.09)	(-0.10)
Log(Age)	-0.282*	-0.257	-0.091	-0.088	0.188**	0.240**
	(-1.80)	(-1.64)	(-1.11)	(-1.08)	(2.01)	(2.46)
Observations	19289	19289	19289	19289	19289	19289
Firm Fixed Effects	YES	YES	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES	YES	YES
Treatment Calculation	KPSS	Citation	KPSS	Citation	KPSS	Citation
Adj. R^2	0.069	0.068	0.002	0.002	0.054	0.025

Table 14:

Lawsuits and Legal Protection (Longformer)

The table displays panel data regressions examining whether lawsuit metrics of large and small firms were differently affected by the Alice decision. In columns (1)-(2), the dependent variable, *# Alleged*, is the number of lawsuits that a firm was alleged in that year. In columns (3) to (4), *# NPE Alleged* is the total number of lawsuits that the firm was alleged by a non-practicing entity in that year. In columns (5) to (6), *# OC Alleged* is the number of lawsuits that the firm was alleged by a operating company in that year. In columns (7)-(8), *Patinfringe* refers to the total number of paragraphs containing both the word root "patent*" and "infringe*" in 10-K documents, scaled by the total number of paragraphs in the 10-Ks. In columns (9)-(10), *# Accuser* is the number of lawsuits that the firm accused any party in a patent lawsuit in that year. *Treatment* is the relative impact of the Alice decision on the firm's patent portfolio measured using KPSS dollar values or citations scaled by sales (see equation (3) for the formula). In the odd and even numbered columns, respectively, we use the KPSS and the number of citations approach to compute the *Patent Value* treatment. *Small* is a binary variable equal to one if a firm's total assets are smaller than the median total asset of its TNIC peers in 2013, and it is zero otherwise. *Large* is 1-*Small*. *Post* is a dummy variable that equals one if the year is after the Alice decision and zero otherwise. All variables are described in detail in the variable list in Appendix A. All regressions include firm and year fixed effects. Standard errors are clustered at the firm level. T-statistics are reported in parentheses; *, **, and *** denote significance at the 10%, 5% and 1% level.

Dependent Variable:	# Alleged		# NPE Alleged		# OC Alleged		Patinfringe		# Accuser	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Small × Post × Treatment	0.555**	0.034	0.371**	0.022	0.123	-0.012	-0.470	-0.683	0.276	0.184*
	(2.10)	(0.29)	(2.31)	(0.27)	(0.75)	(-0.19)	(-0.20)	(-0.93)	(0.75)	(1.71)
Large × Post × Treatment	-4.396***	-1.903***	-1.960***	-1.023**	-2.430***	-0.842***	-4.070***	-3.109***	-0.408	-0.368**
	(-4.56)	(-3.22)	(-2.95)	(-2.56)	(-5.09)	(-2.82)	(-2.81)	(-3.29)	(-1.49)	(-2.27)
Log(Sales)	0.038***	0.037***	0.023***	0.022***	0.011***	0.011**	0.033	0.031	-0.016***	-0.016***
	(4.66)	(4.51)	(3.99)	(3.87)	(2.72)	(2.54)	(1.01)	(0.94)	(-2.67)	(-2.65)
Log(Age)	0.148***	0.153***	0.067***	0.070***	0.068***	0.071***	-0.138	-0.134	-0.388***	-0.388***
	(3.85)	(3.97)	(2.59)	(2.69)	(3.41)	(3.53)	(-1.34)	(-1.30)	(-15.53)	(-15.62)
Observations	19372	19372	19372	19372	19372	19372	19289	19289	19372	19372
Firm Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Treatment Calculation	KPSS	Citation	KPSS	Citation	KPSS	Citation	KPSS	Citation	KPSS	Citation
Adj. R^2	0.021	0.017	0.015	0.014	0.028	0.022	0.004	0.006	0.056	0.057

Table 15:
Acquisitions and Legal Protection (Longformer)

The table displays panel data regressions in which acquisition variables are the dependent variables. In columns (1)-(2), the dependent variables are dollar value spent on acquisition scaled by sales. In columns (3)-(4), the dependent variables are dollar value spent on acquisition of targets that have at least patent scaled by sales. In columns (5)-(6), the dependent variables are log of one plus total value spent on acquisitions in that year. *Treatment* is the relative impact of the Alice decision on the firm's patent portfolio measured using KPSS dollar values or citations scaled by sales (see equation (3) for the formula). In the odd and even numbered columns, respectively, we use the KPSS and the number of citations approach to compute the *Patent Value* treatment. *Small* is a binary variable equal to one if a firm's total assets are smaller than the median total asset of its TNIC peers in 2013, and it is zero otherwise. *Large* is 1-*Small*. *Post* is a dummy variable that equals one if the year is after the Alice decision and zero otherwise. All regressions include firm and year fixed effects. Standard errors are clustered at the firm level. T-statistics are reported in parentheses; *, **, and *** denote significance at the 10%, 5% and 1% level.

Dependent Variable:	$\frac{Acquisitions}{Sales}$		$\frac{Targets\ With\ Patents}{Sales}$		Log(Acquisitions)	
	(1)	(2)	(3)	(4)	(5)	(6)
Small × Post × Treatment	0.186	0.019	0.003	-0.001	0.758	0.009
	(1.56)	(0.54)	(0.70)	(-0.64)	(1.03)	(0.04)
Large × Post × Treatment	0.012	-0.027	-0.019*	-0.015**	-4.275***	-2.739***
	(0.09)	(-0.32)	(-1.96)	(-2.18)	(-2.63)	(-2.96)
1/Sales	-0.036***	-0.036***	-0.001***	-0.001***	-0.126***	-0.127***
	(-6.61)	(-6.63)	(-3.75)	(-3.81)	(-3.71)	(-3.76)
Log(Age)	0.007	0.008	0.001**	0.001**	0.208*	0.213*
	(0.49)	(0.54)	(2.21)	(2.25)	(1.66)	(1.70)
Observations	19372	19372	19372	19372	19372	19372
Firm Fixed Effects	YES	YES	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES	YES	YES
Treatment Calculation	KPSS	Citation	KPSS	Citation	KPSS	Citation
Adj. R^2	0.008	0.008	0.006	0.006	0.004	0.004

Appendix A. Variable definitions

Table 16: Variable definitions Table 16

Variable	Definition
Panel A: Financial Characteristics	
Assets	Compustat item AT.
Sales	Compustat item SALE
OI/Sales	Compustat OIBDP divided by total sales.
Tobin's Q	Compustat sum of market equity (CSHO * $PRCC_F$), DLC, DLTT, PSTKL, all scaled by book assets.
Sales Growth	Natural logarithm of total sales in the current year t divided by total sales in the previous year t-1.
Log(Age)	Natural logarithm of one plus the current year of observation minus the first year the firm appears in the Compustat database.
Panel B: Innovation, Acquisition & Lawsuit Characteristics	
Treatment Effect	Treatment is a weighted average of a firm's patent values multiplied by the Alice Score and scaled by sales. For a firm's patent portfolio, we gather all patents that are valid by the third quarter of 2014. Firm's patent value is calculated in two ways: i) dollar amount provided by KPSS; ii) citations that the patent received. The mathematical notation is provided in equation (3).
R&D/Sales	Compustat XRD divided by total sales. This variable is set to zero if XRD is missing
Log(# of Patents)	Log of one plus number of patent applications.
Patents/Sales	The number of patent applications scaled by firm sales.
Acquisitions/Sales	The total amount of acquisitions divided by firm sales.
PatTargets/Sales	The dollar value of acquisitions where target has a patent scaled by sales.
Log(Acq. Amt.)	Log of one plus total amount of acquisitions.
# Alleged	It is the number of lawsuits that the firm was alleged for infringing a patent.
# NPE Alleged	It is the number of lawsuits that the firm was alleged by an NPE for infringing a patent.
# OC Alleged	It is the number of lawsuits that the firm was alleged by an OC for infringing a patent.
# of Accuser	It is the number of lawsuits that the firm alleged another party for infringing its patent..
IPrisk	The total number of paragraphs mentioning "intellectual property" in the risk factor section in the 10-K documents, scaled by the total number paragraphs in the 10-Ks.
Patinfringe	The total number of paragraphs containing both the word root "patent*" and "infringe*" in 10-K documents, scaled by the total number of paragraphs in the 10-Ks.
Panel C: Competition Measures	
VCF/Sales	A measure of VC entry in a given firm's product market computed as the total first-round dollars raised by the 25 startups from Venture Expert whose Venture Expert business description is most similar to the 10-K business description of the focal firm (using cosine similarities), scaled by focal firm sales.
TSIMM	Total similarity, sum of pairwise textual similarities between rivals as indicated by the TNIC-3 classification of Hoberg and Phillips (2016).
Complaints	The number of paragraphs in the firm's 10-K that mention competition divided by the total number of paragraphs in the firm's 10-K.

Continued on next page

Table 16 – *Continued from previous page*

Variable	**Definition**
Noncompete	#10K paragraphs mentioning "non-compete" agreements, all scaled by the total paragraphs in the 10-K.
Nondisclose	#10K paragraphs mentioning "non-disclose" or "NDA" agreements, all scaled by the total paragraphs in the 10-K.

Online Appendix B: Not for publication

1 Technical Comparison of Models

Among the transformer-based language models, the main differences are sourced from the attention mechanism, tokenization, pre-training task, and pre-training data. Table 17 documents these characteristics for BERT (Devlin et al. (2019)), SciBERT (Beltagy et al. (2019)), RoBERTa (Liu et al. (2019)), and Longformer (Beltagy et al. (2020)) models. In the text below, we explain the attention mechanisms, tokenization, and the pre-training procedure in detail.

BERT, SciBERT, and, RoBERTa process all words in a single iteration rather than one-by-one or in a fixed-sized sliding window approach. In these models, the context of a word depends not only on the words that come before it, but depends on the relative position to each other word in the text. The amount of attention given to each word is decided by the internal dynamics of these models. The mechanism where all words in the text have to be paid attention to is referred as the *full-attention* mechanism. In this mechanism, since there is a pairwise attention between words, memory usage is quadratic with respect to the number of words in the text, limiting the usage to 512 tokens (roughly 400 words).

In contrast, the Longformer model uses a sparse-attention mechanism. In this mechanism, for each word, the model does not use pairwise attention between each words in the text. Instead, for each token, the model pays attention only to the 256 tokens that come before and after it, and to a few special tokens. Therefore, memory usage is close to linear with respect to the number of words in a text. Overall, there is a trade-off between full-attention vs. sparse-attention models: the BERT, SciBERT, and, RoBERTa models have more precision for gathering context while the Longformer model can incorporate more tokens.

What is the relation between a word and a token and how the tokenization is different between the four models? While most of the words are converted to a single token, some words can be converted to more than one token. For instance, the word "embodiment" can be converted to the tokens "emb","-od", and "-iment", while the word "transistor" can be converted to the tokens "trans" and "-istor". The way the words will be tokenized depends on the model. The BERT and SciBERT models use WordPiece algorithm and the RoBERTa and Longformer models use the byte-level BPE algorithm for tokenization. It is worth noting that the resulting token of the same word may not be exactly the same even when the same algorithm is used since the pre-training data used for each model is different. Despite the differences in the tokenization algorithm and pre-training data, in general terms, 512 tokens correspond to 400-430 words.

Models	Attention Mechanism	Tokenization Algorithm	Pre-training Tasks	Pre-training Data
BERT	Full-attention	WordPiece	MLM, NSP	Wikipedia, Book Corpus
SciBERT	Full-attention	WordPiece	MLM, NSP	Scientific Articles
RoBERTa	Full-attention	Byte-level BPE	MLM	Wikipedia, Book Corpus, CC-News, Open Web Text, Stories
Longformer	Sparse-attention	Byte-level BPE	MLM	Wikipedia, Book Corpus, CC-News, Open Web Text, Stories, Realnews

Table 17: Comparison of the models

The pre-training procedure of BERT and SciBERT consists of two different tasks: Masked Language Modelling (MLM) and Next Sequence Prediction (NSP). In the Masked Language Modelling task, some randomly selected tokens are masked, and the models try to predict them. In the Next Sequence Prediction task, two sequences are given to the model, and the model predicts whether these two sentences follow each other. The pre-training procedure of RoBERTa and Longformer only use the Masked Language Modelling task. However, they are trained on a much larger dataset than the ones for BERT and SciBERT.

While the BERT is trained on a dataset that contains text from Wikipedia and Book Corpus (Zhu et al. (2015)), SciBERT is trained on a dataset that contains research articles obtained from Semantic Scholar (Ammar et al. (2018)). RoBERTa is trained on a dataset that contains the text used in the pre-training of BERT and some additional text, which is sourced from newsletters (Nagel (2016)), texts crawled from the URLs that are shared on Reddit and have at least three upvotes (Gokaslan and Cohen (2019)), and Stories dataset (Trinh and Le (2018)) in which every entry forms a story.

The Longformer model begins its pre-training from the already pre-trained RoBERTa model, and it is further pre-trained so that it can learn the new sparse-attention mechanism. The second-phase of the pre-training data incorporates additional text from Realnews dataset (Zellers et al. (2019))

As a separate note, the fine-tuning procedure is not dependent on the model, but it is dependent on the task. Therefore, in our paper, we use the same fine-tuning procedure for all of the models.

2 Longformer Model (Firms Categorized by Market Shares)

Table IA1:
Patents and R&D (Firms Categorized By Market Share) Longformer

The table displays panel data regressions in which innovation and R&D measures are dependent variables. In columns (1)-(2), the dependent variable is the number of patent applications in that year divided by sales; and in columns (3) and (4), it is one plus log of the number of patent applications in the respective year. In columns (5)-(6), the dependent variable is R&D expenses scaled by sales. *Treatment* is the relative impact of the Alice decision on the firm's patent portfolio measured using KPSS dollar values or citations scaled by sales (see equation (3) for the formula). In the odd and even numbered columns, respectively, we use the KPSS and the number of citations approach to compute the *Patent Value* treatment. *Low* is a binary variable equals one if a firm's TNIC market share is lower than the median industry-year market share in 2013 and zero otherwise. *High* is 1-*Low*. *Post* is a dummy variable that equals one if the year is after the Alice decision and zero otherwise. All variables are described in detail in the variable list in Appendix A. All regressions include firm and year fixed effects. Standard errors are clustered at the firm level. T-statistics are reported in parentheses; *, **, and *** denote significance at the 10%, 5% and 1% level.

Dependent Variable:	$\frac{\#\ of Patents}{Sales}$		Log(# of Patents)		$\frac{R\&D}{Sales}$	
	(1)	(2)	(3)	(4)	(5)	(6)
Low X Post X Treatment	-0.463***	-0.160***	-2.743***	-0.961***	5.467***	1.282***
	(-8.29)	(-8.13)	(-3.63)	(-3.68)	(4.10)	(3.42)
High X Post X Treatment	-0.113***	-0.076***	-2.273***	-1.095**	-0.030	-0.037
	(-4.42)	(-3.36)	(-2.95)	(-2.33)	(-0.27)	(-0.54)
1/Sales	-0.010***	-0.010***	0.035***	0.034***	-0.094***	-0.096***
	(-7.67)	(-7.68)	(3.50)	(3.36)	(-6.48)	(-6.48)
Log(Age)	-0.001	-0.002	0.005	0.002	0.144***	0.160***
	(-0.29)	(-0.66)	(0.14)	(0.06)	(3.77)	(4.10)
Observations	19504	19504	19504	19504	19504	19504
Firm Fixed Effects	YES	YES	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES	YES	YES
Treatment Calculation	KPSS	Citation	KPSS	Citation	KPSS	Citation
Adj. R^2	0.098	0.099	0.049	0.048	0.067	0.048

Table IA2:
Profitability (Firms Categorized By Market Share) Longformer

The table displays panel data regressions that examine whether the profitability of high and low market share firms were differently affected by the Alice decision. In columns (1)-(2), the dependent variable is sales growth, calculated as the natural logarithm of total sales in the current year t divided by total sales in the previous year t-1; and in columns (3) and (4), it is Operating Income scaled by sales. In columns (5)-(6), the dependent variable is Tobin's Q, calculated as the market to book ratio (market value of equity plus book debt and preferred stock, all divided by book assets). *Treatment* is the relative impact of the Alice decision on the firm's patent portfolio measured using KPSS dollar values or citations scaled by sales (see equation (3) for the formula). In the odd and even numbered columns, respectively, we use the KPSS and the number of citations approach to compute the *Patent Value* treatment. *Low* is a binary variable equals one if a firm's TNIC market share is lower than the median industry-year market share in 2013 and zero otherwise. *High* is 1-*Low*. *Post* is a dummy variable that equals one if the year is after the Alice decision and zero otherwise. All variables are described in detail in the variable list in Appendix A. All regressions include firm and year fixed effects. Standard errors are clustered at the firm level. T-statistics are reported in parentheses; *, **, and *** denote significance at the 10%, 5% and 1% level.

Dependent Variable:	Sales Growth		$\frac{OperatingIncome}{Sales}$		Market-to-Book	
	(1)	(2)	(3)	(4)	(5)	(6)
Low X Post X Treatment	0.587	0.095	-8.991***	-2.138***	-9.194***	-2.869***
	(1.56)	(0.88)	(-3.26)	(-2.74)	(-3.40)	(-3.43)
High X Post X Treatment	0.361**	0.119	0.056	0.076	2.953**	2.706***
	(2.36)	(1.45)	(0.36)	(0.66)	(2.21)	(2.89)
1/Sales	-0.208***	-0.208***	0.366***	0.368***	-0.344***	-0.342***
	(-27.88)	(-27.87)	(8.07)	(8.12)	(-6.65)	(-6.63)
Log(Age)	-0.027	-0.025	-0.544***	-0.572***	-1.061***	-1.089***
	(-1.25)	(-1.19)	(-5.19)	(-5.43)	(-6.72)	(-6.88)
Observations	19381	19381	18647	18647	18992	18992
Firm Fixed Effects	YES	YES	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES	YES	YES
Treatment Calculation	KPSS	Citation	KPSS	Citation	KPSS	Citation
Adj. R^2	0.173	0.172	0.061	0.054	0.076	0.076

Table IA3:
Competition and Patent Protection (Firms Categorized By Market Share) Long-former

The table displays panel data regressions in which competition variables are the dependent variables. In columns (1)-(2), the dependent variable, *VCF/Sales*, is the a measure of VC entry in a given firm's product market and is the total first-round dollars raised by the ten startups from Venture Expert whose Venture Expert business description most closely matches the 10-K business description of the focal firm (using cosine similarities), scaled by focal firm sales. *TSIMM* is the firm's TNIC text-based total similarity of the firm to public firm competitors. Complaints is the number of paragraphs in the firm's 10-K that complain about competition divided by the total number of paragraphs in the firm's 10-K. *Treatment* is the relative impact of the Alice decision on the firm's patent portfolio measured using KPSS dollar values or citations scaled by sales (see equation (3) for the formula). In the odd and even numbered columns, respectively, we use the KPSS and the number of citations approach to compute the *Patent Value* treatment. *Low* is a binary variable equals one if a firm's TNIC market share is lower than the median industry-year market share in 2013 and zero otherwise. *High* is 1-*Low*. *Post* is a dummy variable that equals one if the year is after the Alice decision and zero otherwise. All variables are described in detail in the variable list in Appendix A. All regressions include firm and year fixed effects. Standard errors are clustered at the firm level. T-statistics are reported in parentheses; *, **, and *** denote significance at the 10%, 5% and 1% level.

Dependent Variable:	$\frac{VCF}{Sales}$		TSIMM		Complaints	
	(1)	(2)	(3)	(4)	(5)	(6)
Low X Post X Treatment	9.116***	2.181***	83.836***	21.793***	8.928**	2.906**
	(7.18)	(5.54)	(7.86)	(6.57)	(2.18)	(2.06)
High X Post X Treatment	0.208	0.002	12.265***	3.971**	-5.356	0.002
	(1.27)	(0.02)	(2.70)	(2.03)	(-1.22)	(0.00)
Log(Sales)	-0.459***	-0.461***	0.803***	0.791***	0.113	0.111
	(-15.52)	(-15.53)	(4.68)	(4.47)	(1.25)	(1.22)
Log(Age)	0.385***	0.411***	0.334	0.548	-0.285	-0.253
	(6.41)	(6.79)	(0.71)	(1.13)	(-0.90)	(-0.80)
Observations	19379	19379	19362	19362	19388	19388
Firm Fixed Effects	YES	YES	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES	YES	YES
Treatment Calculation	KPSS	Citation	KPSS	Citation	KPSS	Citation
Adj. R^2	0.258	0.234	0.070	0.047	0.006	0.006

Table IA4:
Firm IP Risk and Legal Protections (Firms Categorized By Market Share) Longformer

The table displays panel data regressions examining the impact of Alice on intellectual property and noncompete and disclosure clauses. In columns (1)-(2), *IP Risk*, is the total number of paragraphs mentioning "intellectual property" in the risk factor section in the 10-K documents, scaled by the total number paragraphs in the 10-Ks. *Noncompete* is the total number of 10K paragraphs mentioning "non-compete" agreements, all scaled by the total paragraphs in the 10-K. *Nondisclosure* is the total number of 10-K paragraphs mentioning "non-disclose" or "NDA" agreements, all scaled by the total paragraphs in the 10-K. *Treatment* is the relative impact of the Alice decision on the firm's patent portfolio measured using KPSS dollar values or citations scaled by sales (see equation (3) for the formula). In the odd and even numbered columns, respectively, we use the KPSS and the number of citations approach to compute the *Patent Value* treatment. *Low* is a binary variable equals one if a firm's TNIC market share is lower than the median industry-year market share in 2013 and zero otherwise. *High* is 1-*Low*. *Post* is a dummy variable that equals one if the year is after the Alice decision and zero otherwise. All variables are described in detail in the variable list in Appendix A. All regressions include firm and year fixed effects. Standard errors are clustered at the firm level. T-statistics are reported in parentheses; *, **, and *** denote significance at the 10%, 5% and 1% level.

Dependent Variable:	IP Risk		Noncompete		Nondisclosure	
	(1)	(2)	(3)	(4)	(5)	(6)
Low X Post X Treatment	9.167***	3.445***	0.627	0.465	11.612***	2.254***
	(2.93)	(3.84)	(0.85)	(1.55)	(4.16)	(3.23)
High X Post X Treatment	3.311	1.820	-0.675	-0.440	0.461	-0.311
	(1.44)	(1.36)	(-1.09)	(-1.38)	(0.57)	(-0.84)
1/Sales	0.012	0.014	0.042**	0.042**	-0.006	-0.010
	(0.22)	(0.28)	(2.11)	(2.14)	(-0.14)	(-0.24)
Log(Age)	-0.201	-0.186	-0.070	-0.069	0.207**	0.242**
	(-1.26)	(-1.17)	(-0.87)	(-0.85)	(2.21)	(2.49)
Observations	19388	19388	19388	19388	19388	19388
Firm Fixed Effects	YES	YES	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES	YES	YES
Treatment Calculation	KPSS	Citation	KPSS	Citation	KPSS	Citation
Adj. R^2	0.064	0.065	0.002	0.002	0.042	0.023

Table IA5:

Lawsuits and Legal Protection (Firms Categorized By Market Share) Longformer

The table displays panel data regressions examining whether lawsuit metrics of high and low market share firms were differently affected by the Alice decision. In columns (1)-(2), the dependent variable, *is Alleged*, is a dummy variable that equals one if a firm was alleged in a patent lawsuit at least once in that year, and zero otherwise. In columns (3) to (4), *Alleged by NPE* is a dummy variable that equals one if a firm was alleged by a non-practicing entity in a patent lawsuit at least once in that year, and zero otherwise. In columns (5) to (6), *Alleged by OC* is a dummy variable that equals one if a firm was alleged by a operating company in a patent lawsuit at least once in that year, and zero otherwise. In columns (7)-(8), *Patinfringe* refers to the total number of paragraphs containing both the word root "patent*" and "infringe*" in 10-K documents, scaled by the total number of paragraphs in the 10-Ks. In columns (9)-(10), *Is Accuser* is a binary variable equals one if a firm accused any party in a patent lawsuit at least once in that year, and zero otherwise. *Treatment* is the relative impact of the Alice decision on the firm's patent portfolio measured using KPSS dollar values or citations scaled by sales (see equation (3) for the formula). In the odd and even numbered columns, respectively, we use the KPSS and the number of citations approach to compute the *Patent Value* treatment. *Low* is a binary variable equals one if a firm's TNIC market share is lower than the median industry-year market share in 2013 and zero otherwise. *High* is 1-*Low*. *Post* is a dummy variable that equals one if the year is after the Alice decision and zero otherwise. All variables are described in detail in the variable list in Appendix A. All regressions include firm and year fixed effects. Standard errors are clustered at the firm level. T-statistics are reported in parentheses; *, **, and *** denote significance at the 10%, 5% and 1% level.

Dependent Variable:	Alleged #		Alleged # by NPE		Alleged # by OC		Patinfringe		# of Sueing	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Low X Post X Treatment	0.883*	0.079	0.815**	0.097	-0.092	-0.064	-3.309	-1.408**	0.094	0.119
	(1.79)	(0.55)	(2.20)	(0.92)	(-0.41)	(-0.80)	(-1.60)	(-2.01)	(0.27)	(1.07)
High X Post X Treatment	-5.484***	-2.140***	-2.728***	-1.341***	-2.652***	-0.704**	-1.773	-0.936	-0.284	-0.153
	(-5.66)	(-3.99)	(-4.25)	(-4.09)	(-5.11)	(-2.52)	(-1.07)	(-0.80)	(-1.00)	(-0.96)
1/Sales	0.037***	0.034***	0.023***	0.022***	0.009**	0.008*	0.028	0.026	-0.017***	-0.017***
	(4.55)	(4.20)	(4.16)	(3.88)	(2.18)	(1.87)	(0.85)	(0.80)	(-2.82)	(-2.80)
Log(Age)	0.145***	0.158***	0.059**	0.066**	0.072***	0.076***	-0.105	-0.108	-0.398***	-0.397***
	(3.80)	(4.09)	(2.24)	(2.51)	(3.56)	(3.76)	(-1.02)	(-1.05)	(-15.75)	(-15.76)
Observations	19504	19504	19504	19504	19504	19504	19388	19388	19504	19504
Firm Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Treatment Calculation	KPSS	Citation	KPSS	Citation	KPSS	Citation	KPSS	Citation	KPSS	Citation
Adj. R^2	0.024	0.018	0.018	0.016	0.028	0.020	0.004	0.004	0.059	0.059

Table IA6:

Acquisitions and Legal Protection (Firms Categorized By Market Share) Longformer

The table displays panel data regressions in which acquisition variables are the dependent variables. In columns (1)-(2) and (3)-(4), the dependent variables are dollar value spent on acquisition scaled by sales and log of one plus total value spent on acquisitions in that year. *Treatment* is the relative impact of the Alice decision on the firm's patent portfolio measured using KPSS dollar values or citations scaled by sales (see equation (3) for the formula). In the odd and even numbered columns, respectively, we use the KPSS and the number of citations approach to compute the *Patent Value* treatment. *Low* is a binary variable equals one if a firm's TNIC market share is lower than the median industry-year market share in 2013 and zero otherwise. *High* is 1-*Low*. *Post* is a dummy variable that equals one if the year is after the Alice decision and zero otherwise. All regressions include firm and year fixed effects. Standard errors are clustered at the firm level. T-statistics are reported in parentheses; *, **, and *** denote significance at the 10%, 5% and 1% level.

Dependent Variable:	$\frac{Acquisitions}{Sales}$		$\frac{Targets\ With\ Patents}{Sales}$		Log(Acquisitions)	
	(1)	(2)	(3)	(4)	(5)	(6)
Low X Post X Treatment	0.220	0.057	-0.001	-0.002	1.070	0.376
	(1.55)	(1.36)	(-0.09)	(-0.83)	(1.13)	(1.38)
High X Post X Treatment	-0.068	-0.078	-0.019**	-0.011**	-5.638***	-3.398***
	(-0.65)	(-1.32)	(-2.01)	(-2.01)	(-3.32)	(-3.79)
1/Sales	-0.037***	-0.037***	-0.001***	-0.001***	-0.140***	-0.143***
	(-6.84)	(-6.86)	(-4.01)	(-4.10)	(-4.14)	(-4.21)
Log(Age)	0.019	0.019	0.002***	0.002***	0.299**	0.311**
	(1.29)	(1.35)	(2.72)	(2.79)	(2.36)	(2.45)
Observations	19504	19504	19504	19504	19504	19504
Firm Fixed Effects	YES	YES	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES	YES	YES
Treatment Calculation	KPSS	Citation	KPSS	Citation	KPSS	Citation
Adj. R^2	0.008	0.008	0.006	0.006	0.003	0.003

3 TF-IDF and CPC Models Instead Of Longformer

Table IA7:

Patents and R&D (Alice Scores Calculated by TF-IDF Instead of Longformer Model)

The table displays the robustness tests for the results in Table 9. In this table, in the calculation of the treatment variable depicted in equation (3), we use TF-IDF Instead of Longformer Model technique. In columns (1)-(2), the dependent variable is the number of patent applications in that year divided by sales; and in columns (3) and (4), it is one plus log of the number of patent applications in the respective year. In columns (5)-(6), the dependent variable is R&D expenses scaled by sales. In the odd and even numbered columns, KPSS and the number of citations that a patent received are used for *Patent Value* in calculation of the treatment, respectively. *Small* is a binary variable equal to one if a firm's total assets are smaller than the median total asset of its TNIC peers in 2013, and it is zero otherwise. *Large* is 1-*Small*. *Post* is a dummy variable that equals one if the year is after the Alice decision and zero otherwise. All variables are described in detail in the variable list in Appendix A. All regressions include firm and year fixed effects. Standard errors are clustered at the firm level. T-statistics are reported in parentheses; *, **, and *** denote significance at the 10%, 5% and 1% level.

Dependent Variable:	$\frac{\#\ of Patents}{Sales}$		Log(# of Patents)		$\frac{R\&D}{Sales}$	
	(1)	(2)	(3)	(4)	(5)	(6)
Small X Post X Treatment	-0.603***	-0.139***	-2.284	-0.579	6.173***	1.526***
	(-5.48)	(-4.40)	(-1.55)	(-1.38)	(2.90)	(2.73)
Large X Post X Treatment	-0.172***	-0.118***	-4.314***	-2.111***	0.019	-0.054
	(-4.34)	(-3.79)	(-3.54)	(-3.09)	(0.14)	(-0.54)
Log(Sales)	-0.009***	-0.010***	0.039***	0.037***	-0.098***	-0.097***
	(-7.42)	(-7.39)	(3.80)	(3.63)	(-6.45)	(-6.40)
Log(Age)	-0.001	-0.002	-0.006	-0.008	0.158***	0.166***
	(-0.43)	(-0.77)	(-0.18)	(-0.22)	(4.02)	(4.19)
Observations	19372	19372	19372	19372	19372	19372
Firm Fixed Effects	YES	YES	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES	YES	YES
Treatment Calculation	KPSS	Citation	KPSS	Citation	KPSS	Citation
Adj. R^2	0.081	0.074	0.047	0.044	0.044	0.039

Table IA8:
Patents and R&D (CPC Dummy Variable Instead of Alice Score)

The table displays the robustness tests for the results in Table 9. In this table, in the calculation of the treatment variable depicted in equation (3), we use CPC dummy instead of Alice Score. CPC dummy equals one if a patent's primary CPC belongs to one of the top-20 CPCs that have the most frequent Alice rejections and zero otherwise. In columns (1)-(2), the dependent variable is the number of patent applications in that year divided by sales; and in columns (3) and (4), it is one plus log of the number of patent applications in the respective year. In columns (5)-(6), the dependent variable is R&D expenses scaled by sales. *Treatment* is the relative impact of the Alice decision on the firm's patent portfolio measured using KPSS dollar values or citations scaled by sales (see equation (3) for the formula). In the odd and even numbered columns, respectively, we use the KPSS and the number of citations approach to compute the *Patent Value* treatment. *Small* is a binary variable equal to one if a firm's total assets are smaller than the median total asset of its TNIC peers in 2013, and it is zero otherwise. *Large* is 1-*Small*. *Post* is a dummy variable that equals one if the year is after the Alice decision and zero otherwise. All variables are described in detail in the variable list in Appendix A. All regressions include firm and year fixed effects. Standard errors are clustered at the firm level. T-statistics are reported in parentheses; *, **, and *** denote significance at the 10%, 5% and 1% level.

Dependent Variable:	$\frac{\#\ of Patents}{Sales}$		Log(# of Patents)		$\frac{R\&D}{Sales}$	
	(1)	(2)	(3)	(4)	(5)	(6)
Small *X* Post *X* Treatment	-0.323***	-0.081***	-1.731***	-0.456**	1.485*	0.196
	(-7.93)	(-6.12)	(-2.96)	(-2.52)	(1.68)	(1.06)
Large *X* Post *X* Treatment	-0.077***	-0.047***	-1.433***	-0.679**	-0.130	-0.102
	(-3.59)	(-3.39)	(-2.58)	(-2.29)	(-0.82)	(-0.94)
Log(Sales)	-0.009***	-0.010***	0.038***	0.037***	-0.098***	-0.098***
	(-7.32)	(-7.39)	(3.63)	(3.60)	(-6.40)	(-6.35)
Log(Age)	-0.001	-0.002	0.001	-0.006	0.166***	0.173***
	(-0.35)	(-0.89)	(0.02)	(-0.17)	(4.10)	(4.28)
Observations	19372	19372	19372	19372	19372	19372
Firm Fixed Effects	YES	YES	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES	YES	YES
Treatment Calculation	KPSS	Citation	KPSS	Citation	KPSS	Citation
Adj. R^2	0.090	0.084	0.046	0.045	0.035	0.032

4 FuzzyDID

Table IA9:
Fuzzy DID (Panel A) With Longformer Model

This table presents local average treatment effect (LATE) and t-statistics from the Fuzzy DID model developed by de Chaisemartin and D'Haultfoeuille (2018). In columns (1) and (2), the results are displayed for the subsample of Small Firms; and in (3) and (4), they are displayed for Large Firms. A firm is classified as small if its total assets are below the median of its peers in the TNIC database, and it is classified as large otherwise.

	Small		Large	
	(1)	(2)	(3)	(4)
$\frac{\#\ of Patents}{Sales}$	-1.313**	-0.512*	-1.574***	-1.056***
	(-2.369)	(-1.952)	(-4.759)	(-4.195)
Log(# of Patents)	-16.105**	-6.822*	-106.991***	-74.504***
	(-2.294)	(-1.909)	(-4.644)	(-5.108)
$\frac{R\&D}{Sales}$	36.040***	14.985**	3.230	2.139
	(2.892)	(2.186)	(1.465)	(1.593)
$\frac{VCF}{Sales}$	241.150*	87.239*	-23.073**	-14.828**
	(1.930)	(1.701)	(-2.158)	(-2.401)
TSIMM	743.933***	299.036***	205.197	143.202
	(4.144)	(2.629)	(1.369)	(1.578)
Complaints	143.088**	62.305*	30.320	2.341
	(2.264)	(1.658)	(0.250)	(0.032)
Sales Growth	6.592*	2.871	1.058	1.179
	(1.830)	(1.601)	(0.209)	(0.402)
$\frac{Operating Income}{Sales}$	-48.714***	-20.011***	7.701	5.225
	(-3.114)	(-3.167)	(1.419)	(1.635)
Tobin's Q	30.234	16.927	80.842**	50.022**
	(1.127)	(1.477)	(2.319)	(2.509)
Treatment Calculation	KPSS	Citation	KPSS	Citation

Table IA10:
Fuzzy DID Continued (Panel B) With Longformer Model

This table presents local average treatment effect (LATE) and t-statistics from the Fuzzy DID model developed by de Chaisemartin and D'Haultfoeuille (2018). In columns (1) and (2), the results are displayed for the subsample of Small Firms; and in (3) and (4), they are displayed for Large Firms. A firm is classified as small if its total assets are below the median of its peers in the TNIC database, and it is classified as large otherwise.

	Small		Large	
	(1)	(2)	(3)	(4)
# Alleged	-3.080 (-0.526)	-1.124 (-0.405)	-54.070** (-2.436)	-39.214** (-2.461)
# NPE Alleged	0.685 (0.180)	0.478 (0.290)	-3.984 (-0.324)	-4.652 (-0.526)
# OC Alleged	-5.515** (-2.099)	-2.351 (-1.611)	-46.407*** (-3.876)	-32.195*** (-3.475)
Patinfringe	-4.478 (-0.183)	-2.879 (-0.248)	-44.874 (-1.416)	-29.742 (-1.359)
# Accuser	11.622*** (3.415)	4.737** (2.498)	-9.026 (-1.316)	-5.666 (-1.390)
IP Risk	237.524*** (4.441)	94.853*** (2.835)	230.483*** (3.557)	148.853*** (2.895)
Noncompete	2.063 (0.192)	0.447 (0.082)	10.688 (0.808)	8.401 (0.959)
Nondisclosure	106.698*** (3.726)	45.111*** (2.606)	22.638* (1.722)	14.963* (1.707)
$\frac{Acquisitions}{Sales}$	-0.068 (-0.073)	0.051 (0.094)	-2.148 (-0.682)	-0.867 (-0.422)
$\frac{Targets\ With\ Patents}{Sales}$	-0.075 (-1.126)	-0.027 (-0.856)	-0.651** (-2.107)	-0.445** (-2.137)
Log(Acquisitions)	0.286 (0.028)	1.491 (0.299)	-47.417 (-1.140)	-32.151 (-1.175)
Treatment Calculation	KPSS	Citation	KPSS	Citation