RATIONALITY & COMPETITION
CRC TRR 190

Antitrust and (Foreign) Innovation: Evidence from the Xerox Case

Robin Mamrak (LMU Munich)

Discussion Paper No. 396

May 12, 2023

Collaborative Research Center Transregio 190 | www.rationality-and-competition.de
Ludwig-Maximilians-Universität München | Humboldt-Universität zu Berlin
Spokesperson: Prof. Georg Weizsäcker, Ph.D., Humboldt University Berlin, 10117 Berlin, Germany
info@rationality-and-competition.de

Antitrust and (Foreign) Innovation: Evidence from the Xerox Case

Robin Mamrak*

12 May 2023

Abstract

How does antitrust enforcement against patent-based monopolies affect innovation? I address this question by empirically studying the US antitrust case against Xerox, the monopolist in the market for plain-paper copiers. In 1975, Xerox was ordered to license all its copier-technology patents in the US and abroad. I show that this promoted innovation by other firms in the copier industry, measured by a disproportionate increase in patenting in technologies where Xerox patents became available for licensing. This positive effect is driven by increased innovation by Japanese competitors. They started developing smaller desktop copiers and their innovation became more diverse.

JEL Classification: O30, O34, L41, K21

Keywords: Antitrust, Innovation, Patents, Compulsory Licensing, Japan, Xerox

*Ludwig-Maximilians-Universität (LMU) Munich; robin.mamrak@econ.lmu.de. I thank my adviser, Monika Schnitzer, as well as Alessandra Allocca, Simon Block, Tim Bresnahan, Henry Chesbrough, Florian Ederer, Chiara Farronato, Fabian Gaessler, Shane Greenstein, Daniel Guhl, Anna Gumpert, Sebastian Hager, Hanna Hottenrott, Paul Pelzl, Kyle Myers, Markus Nagler, Felix Poege, Tim Simcoe, Fabian Waldinger, and Martin Watzinger for helpful comments. I also received valuable feedback from seminar and conference participants at Boston University, Harvard, LMU Munich, the European Commission, the NBER Productivity Seminar, the 10th CRC Retreat (Berlin), the 2022 Young Economist Symposium (Yale), the 2022 EPIP Conference (online), the 5th RISE Workshop (Munich), the 2023 MaCCI Annual Conference (Mannheim), the 16th RGS Doctoral Conference (Bochum), and the 2023 RES & SES Annual Conference (Glasgow). Financial support by the German Research Foundation through CRC TRR 190 (project number 280092119), by the Elite Network of Bavaria through the international doctoral programme 'Evidence-Based Economics', and by Fulbright Germany through a doctoral scholarship is gratefully acknowledged. I also thank the Laboratory for Innovation Science at Harvard for their hospitality while working on parts of this paper.

1 Introduction

Competition authorities in many countries have tightened their antitrust policy in recent years. In the US, this has raised concerns that stricter antitrust enforcement against domestic incumbents could undermine the American dominance of the high-technology sector, as it may particularly help foreign competitors to catch up.[1] This is an important policy concern, since promoting innovation is increasingly becoming a key objective of antitrust policy, especially in high-technology industries (Gilbert, 2022).

This paper investigates how antitrust enforcement against patent-based monopolies affects innovation by domestic and foreign firms. Patents – and intellectual property (IP) more broadly – are an important source of market power. On the one hand, this is the intended effect of patents. They incentivize innovation by granting patentees the right to exclude others from using the patented invention. On the other hand, this market power can be abused if patentees engage in exclusionary practices. For example, dominant firms may strategically use their patents to block entry by refusing to license their technology to potential competitors.[2] This can give rise to a conflict between patent and antitrust laws, which may warrant intervention by antitrust authorities (e.g., Carrier, 2002). However, there is still little evidence about the impact of antitrust enforcement against incumbents that strategically (ab)use their IP.

To fill this gap, I study the antitrust case against Xerox Corporation in the early 1970s. Xerox was the monopolist in the market for plain-paper copiers in the US throughout the 1960s. The American company, which had developed and commercialised a novel copier technology that is still widely used today, held more than 2,000 patents. It strictly refused to grant licenses to potential competitors and used patent infringement suits to block entry by competitors who developed their own patented technologies. In 1972, the Federal Trade Commission (FTC) charged Xerox with monopolization of the copier market through strategic abuse of the patent system. The case was settled by a consent decree in 1975 and Xerox was ordered to license all its domestic and foreign copier-related patents to any third parties at reasonable rates (FTC, 1975).

The case against Xerox is particularly well suited for addressing my research question. It 'defined what may have been a peak in antitrust prosecution directed toward patent-based monopolies' (Scherer, 2005, p. 300). Therefore, it was one of the most important American antitrust cases in the 20th century. The FTC's intervention was widely perceived as success and triggered a transition to competition in the market for plain-paper

[1]See, for example, 'Antitrust Can Hurt U.S. Competitiveness' in *The Wall Street Journal* (`https://www.wsj.com/articles/antitrust-can-hurt-u-s-competitiveness-11625520340`) or 'Big Tech: Breaking Us Up Will Only Help China' in *Wired* (`https://www.wired.com/story/big-tech-breaking-will-only-help-china`).

[2]For example, *The Economist* notes that 'patents should spur bursts of innovation; instead, they are used to lock in incumbents' advantages' (see `https://www.economist.com/leaders/2015/08/08/time-to-fix-patents`).

copiers (Bresnahan, 1985a; Tom, 2001). As many of the entrants were foreign firms, the case allows me to study the impact of antitrust enforcement on both domestic and foreign innovation.

To empirically estimate the effect of the antitrust case on innovation, I use data on patent applications and employ a difference-in-differences strategy across technology classes (Moser and Voena, 2012; Moser et al., 2014). My main approach uses a continuous treatment variable that exploits variation in the share of patents in a six-digit technology class (based on the Cooperative Patent Classification) that were subject to compulsory licensing. Specifically, I compare the annual number of US patent applications by applicants other than Xerox across differentially affected six-digit classes within the same four-digit class, controlling for a range of fixed effects.

I find that antitrust enforcement against Xerox had an overall positive effect on subsequent innovation by other firms in the copier industry. There was a disproportionate increase in patenting in technologies with a greater exposure to compulsory licensing of Xerox's patents. My estimates indicate that the antitrust case led to an additional 160 patent applications per year. This represents an economically meaningful increase in patenting in relevant technologies by around 1.4%. Event-study analyses illustrate that these estimates are not driven by any differences in pre-trends across groups. The result is also robust to a wide range of alternative model specifications and treatment definitions.

Moreover, I show that the number of forward citations received by compulsorily licensed Xerox patents increased disproportionately after 1975. This complementary analysis follows Watzinger et al. (2020) and employs a matching strategy to find a control patent for every compulsorily licensed Xerox patent. Overall, the findings support the conclusion that the antitrust case against Xerox spurred technological progress, as other firms used the newly available technology for follow-on innovation.

Interestingly, the main beneficiaries of increased access to Xerox's technology were competitors from Japan. In my main approach, when splitting the number of patent applications by applicant country, the positive effect of compulsory licensing is almost entirely driven by Japanese firms. In contrast, the estimated effect on patenting by US applicants is quantitatively small and statistically indistinguishable from zero. I further show that there was great heterogeneity in the effect of the antitrust case even among Japanese applicants. Only established firms increased their patenting, whereas small firms and start-ups did not benefit from access to Xerox's technology. Moreover, the positive innovation effect is driven by Japanese firms with extensive prior patenting experience in copier technologies – that is, (potential) competitors to Xerox in the copier market.

The finding that Japanese rivals increased their innovation following the antitrust case is in line with historical narratives about the development of the copier industry. Scherer (2005) notes that several Japanese copier producers (e.g., Canon, Konica, Ricoh) successfully entered the American market after 1975 and became important competitors to Xerox. Japanese entrants strategically focused on the lower end of the copier market. That is, they produced machines that were cheaper, smaller, and designed for lower volumes than existing plain-paper copiers. This business model was different from that of most American copier producers and is considered one of the key reasons for the Japanese success (e.g., Jacobson and Hillkirk, 1986).

Consistent with this narrative, I show that patents filed by Japanese competitors were more likely to contain words in their title or abstract that can be associated with smaller copiers. Moreover, I find that innovation became more novel and diverse following the antitrust case. These changes in the direction of innovation are again driven by Japanese patent applicants. Their innovation activity expanded to new technology fields, while there was no reduction in the quality of inventions. Therefore, the results are in line with the historical evidence suggesting that Japanese entrants focused on smaller desktop copiers, which were more differentiated from existing products.

Finally, I also investigate how Xerox's own patenting activities reacted to the removal of most of its IP. Relative to a synthetic control group (Abadie et al., 2010, 2015), Xerox and its subsidiaries filed around 16 to 30 patents less per year after 1975. This effect is much smaller than the increase in patenting by other firms, which I find in my main approach. Therefore, my estimates indicate that the overall impact of the antitrust case on subsequent innovation was positive.

The first important takeaway from this paper is that compulsory licensing can promote innovation by other firms in the target industry. This result is complementary to prior evidence by Watzinger et al. (2020), who empirically study the innovation effect of compulsory licensing following the antitrust case against Bell in the 1950s. Although the two cases bear certain similarities, there were important differences in the market structure of the target industry. Bell was a vertically integrated monopolist that could continue to foreclose its rivals even after the loss of most of its IP. In contrast, Xerox's monopoly was primarily based on the strategic use of its patent portfolio such that compulsory licensing removed the main barrier to entry. Accordingly, Watzinger et al. (2020) find no effect of compulsory licensing of Bell's patents in the target industry. My paper, conversely, finds the largest increase in patenting among firms whose prior experience overlaps with Xerox's technology. Therefore, it shows that compulsory licensing can be an effective antitrust remedy within the target industry if it removes the main entry barrier.

The second key takeaway from this paper is that the antitrust case against Xerox particularly benefited Japanese competitors. This result speaks to current debates about antitrust policy. For example, a comment in *The Wall Street Journal* warned that '[a]ntitrust action against leading U.S. tech companies would shrink American dominance of the world's fastest-growing industry'.[3] Based on the historical evidence from the Xerox case,

[3]See 'The Misguided Antitrust Attack on Big Tech' in *The Wall Street Journal* (`https://www.wsj.com/articles/the-misguided-antitrust-attack-on-big-tech-11600125182`).

concerns that antitrust could benefit foreign competitors may be justified. However, drawing appropriate policy conclusions from this finding requires a nuanced view. On the one hand, in the case of Xerox, most rents accrued to foreign competitors. On the other hand, American consumers benefited from increased competition and innovation in the copier industry through lower prices as well as a greater variety and higher quality in copiers.

My paper contributes to the literature on the effect of antitrust on innovation by estimating the differential impact of antitrust intervention on innovation by domestic and foreign firms. While most of the literature on antitrust and innovation is theoretical (Segal and Whinston, 2007; Cabral, 2018; Federico et al., 2020), there has been an increasing number of empirical contributions in recent years (Watzinger et al., 2020; Cunningham et al., 2021; Kang, 2021; Poege, 2022; Watzinger and Schnitzer, 2022). I further add to these studies by providing empirical evidence on the impact of one of the most important US antitrust cases in the 20th century.

The paper also complements prior studies on compulsory licensing and the protection of IP rights (Acemoglu and Akcigit, 2012; Moser and Voena, 2012; Galasso and Schankerman, 2015). Compulsory licensing is a frequently used remedy in competition cases (Delrahim, 2004). I contribute to the literature by studying the effectiveness of compulsory licensing in the specific case where the targeted monopoly is based on patents. My estimates show that Xerox's patents exerted a blocking effect on follow-on innovation by other firms. This effect is consistent with a rent dissipation theory (Arora and Fosfuri, 2003; Gaessler et al., 2019). Xerox likely refused to grant licenses to its competitors, because it feared that revenues from licensing would be lower than the loss in profits due to increased product market competition.

Finally, my paper adds to previous research on the case against Xerox (Bresnahan, 1985a,b; Tom, 2001). Most prominently, Bresnahan (1985a) describes the transition to competition in the copier market and discusses potential innovation effects.[4] However, to the best of my knowledge, I am the first to provide empirical evidence on the impact of the antitrust case on subsequent innovation.

The remainder of the paper is structured as follows. Section 2 explains the historical background on Xerox and the antitrust case. In section 3, I introduce the data and my empirical strategy. The main results are presented in section 4. In sections 5 and 6, I investigate which firms benefited from the antitrust case and study the underlying mechanism. Finally, section 7 analyses the effect on Xerox and section 8 concludes.

[4]Shorter discussions of the antitrust case against Xerox can also be found in contributions by Gomes-Casseres and McQuade (1991) or Scherer (2005, 2007). Moreover, Chesbrough and Rosenbloom (2002) as well as Vinokurova and Kapoor (2020) study innovation by Xerox from a management perspective.

2 Historical Background

Xerox Corporation was the de facto monopolist in the market for plain-paper copiers in the US until the early 1970s. At the start of the antitrust case in 1972, its share in the rapidly growing market was close to 95%. Xerox was the 17th most profitable American company based on return on stockholders' equity (FTC, 1975). This section provides a brief historical overview of Xerox's rise and the antitrust case.

2.1 Xerox's Monopoly in the Plain-Paper Copier Market

The foundations for Xerox's success were laid in 1938, when the American physicist Chester Carlson invented the process of electrophotography. This technology, which was later called 'xerography' (Greek for 'dry writing'), allowed to print images using an electrostatic process and a dry powder (i.e., toner).

Although xerography forms the basis of the technology used in copiers and laser printers still today, it took two decades to transform Carlson's invention into a marketable product. In 1946, the Haloid Photographic Company, a small manufacturer of photographic equipment from Rochester, NY, agreed to commercialise xerography.[5] Haloid introduced its first xerographic office copier in 1949. Despite the machine's limited initial success, the company continued investing in its xerographic technology and was later renamed Xerox Corporation.

Xerox achieved its major breakthrough in 1959 when it launched the Xerox 914 office copier. The fully automated machine could produce a xerographic copy within seconds, was easy to operate, and could be used with ordinary (plain) paper. This made Xerox's technology superior to that of competitors, as the 914 did not require using special (coated) paper. The Xerox 914 became an immediate commercial success and Xerox's annual sales increased by a factor of 25 from 1959 to 1968, making Xerox the fastest company to reach $1 billion in sales (Jacobson and Hillkirk, 1986; Gomes-Casseres and McQuade, 1991).

The success of the 914 rewarded Xerox for its large and risky investment into commercialising xerography. Despite widespread initial scepticism regarding the technology's potential, Xerox spent more than its total earnings throughout the 1950s on the development of the 914 (Jacobson and Hillkirk, 1986). The copier market also grew rapidly. The annual number of copies made in the US increased from approximately 20 million in the mid-1950s to around 10 billion a decade later (Jacobson and Hillkirk, 1986). In

[5] As Carlson lacked financial resources to develop xerography himself, he approached more than a dozen major US firms to commercialise his invention, but none of them was interested. In 1944, the Battelle Memorial Institute, a non-profit research organisation, agreed to invest in the technology's development and continued searching for a corporate partner. Battelle then entered into an agreement with Haloid. More details on the development of xerography and the origins of Xerox can be found in Kearns and Nadler (1992) and Owen (2005).

the 1960s, Xerox introduced several new copiers that could operate at higher speed or contained additional features.

Xerox had protected its copier technology through more than 2,000 patents and strictly refused to grant licenses to any other manufacturers of plain-paper copiers. Therefore, throughout the 1960s, no other firms could sell plain-paper copiers. Xerox also sold its plain-paper copiers and protected its IP abroad. To this end, the company had established two foreign subsidiaries, Rank Xerox in Europe and Fuji Xerox in Japan, which acted as Xerox's international sales organisations.[6]

In 1970, Xerox's monopoly in the market for plain-paper copiers was challenged for the first time when International Business Machines (IBM) introduced its first xerographic copier. Despite the inferior quality of IBM's copier, Xerox immediately sued for patent infringement, initiating a legal battle that lasted for several years (Jacobson and Hillkirk, 1986). Similarly, in 1972, Litton Industries launched a plain-paper copier and was sued by Xerox.

2.2 FTC Complaint and 1975 Consent Decree

In late 1972, the Federal Trade Commission (FTC) filed an antitrust complaint against Xerox that alleged monopolization of the plain-paper copier market.[7] Xerox was accused of violating Section 5(a) of the FTC Act by hindering effective competition in plain-paper copiers through its strategic use of the patent system and anti-competitive pricing policies (FTC, 1975). As outlined by Bresnahan (1985b), the theory of harm regarding Xerox's patent practices was twofold. On the one hand, entry into the market may have been inhibited by the size, complexity, and obscurity of Xerox's patent thicket in combination with the threat of infringement litigation. On the other hand, Xerox was accused of strategically building, maintaining, and using parts of its patent portfolio for the sole purpose of denying access to relevant technologies to its competitors. The second part of allegedly anti-competitive practices related to Xerox's pricing policies. In particular, Xerox pursued a lease-only policy with tied maintenance for its copiers and it used various ways of price discrimination between customers, based on the number of copies or the number of Xerox machines installed.

In 1975, Xerox settled with the FTC by signing a consent decree whose main remedy was compulsory licensing of all of Xerox's copier-technology patents.[8] Compulsory

[6]Rank Xerox was established in 1956 as joint venture between Xerox and the Rank Organization from the UK; Fuji Xerox was established in 1962 as joint venture between Rank Xerox and the Japanese company Fuji Photo Film (Gomes-Casseres and McQuade, 1991).

[7]The FTC defined the relevant market as the sale and lease of office copiers and supplies in the US (FTC, 1975). Plain-paper copiers represented the most important submarket. At the time of the complaint, there were 25 firms active in the American copier market, but only three firms (i.e., Xerox, IBM, and Litton) distributed plain-paper copiers. Amongst them, Xerox accounted for 95% of revenues.

[8]Xerox had initially denied all allegations but eventually decided to settle. In parallel to the FTC case, Xerox had also been involved in several other antitrust lawsuits brought forward by competitors.

licensing also applied to Xerox's foreign patents, including the ones held by Rank Xerox and Fuji Xerox. Xerox had to grant the first three licenses to each firm royalty-free and could then ask for reasonable royalties that were not to exceed 1.5% of the licensee's revenues. Future Xerox patents issued until 1981 were also covered by the licensing requirement (FTC, 1975). Moreover, Xerox was ordered to cease all patent infringement suits as well as one pricing policy that provided discounts on individual rental rates to very high-volume customers.

The features of the consent decree indicate that the FTC viewed Xerox's use of its patent portfolio as the main barrier to entry. The initial FTC complaint had also proposed to ban a long list of Xerox's pricing policies, out of which only one was prohibited by the final consent decree. In contrast, the FTC made no concessions on its claim that Xerox licenses its copier-related patents. Accordingly, Bresnahan (1985b, p. v) argues that 'the FTC's emphasis on patents over pricing practices was wise. [...] [T]he [pricing] practices were price discrimination devices – i.e., the fruits not the causes of monopoly power.'

The market for plain-paper copiers in the US became competitive in the 1970s. More than a dozen other firms – such as Eastman Kodak, Savin, or Smith Corona (SCM) from the US as well as Canon, Konishiroku (Konica), or Ricoh from Japan – entered the market. From 1972 to 1977, prices for plain-paper copiers declined by more than 30% and Xerox's market share in net new placements fell from close to 100% to less that 20% (Bresnahan, 1985b).

3 Data and Empirical Strategy

I use data on patent applications to empirically measure innovation. Patents are well-suited for my empirical approach for several reasons. First, patent data are consistently available throughout the relevant period. Second, as patents are assigned to hierarchical technology classes, I can compare patenting across different technologies within the same field. Third, patent citations allow me to measure follow-on innovation to Xerox's patents (Jaffe and Trajtenberg, 1996), since firms had to cite any prior art irrespective of the licensing order.[9]

Therefore, the company incurred high costs – both in terms of legal expenses and the time of its executives. As later explained by Peter McColough, then Xerox's chief executive officer, the reason for settling with the FTC was not to admit any wrongful acts but rather to find a way forward allowing Xerox to focus on its business again (Jacobson and Hillkirk, 1986).

[9]Despite these advantages, using patent data also has some drawbacks. On the one hand, patents may be an imperfect measure of innovation, because not all innovations are patentable and inventors may opt for secrecy as an alternative means of protection (e.g., Moser, 2012). However, this concern is mitigated by the fact that patent protection played an important role in the copier industry, as indicated by the historical background. On the other hand, patent citations may not accurately measure follow-on innovation, as citations may have been added by the examiner even in the absence of any knowledge flow (Alcácer and Gittelman, 2006).

My main data source is the Worldwide Patent Statistical Database (PATSTAT) of the European Patent Office. In addition, I use data from the 'HistPat' database (Petralia et al., 2016) as well as the 'patentCity' project by Bergeaud and Verluise (2022) to identify the applicants' country of origin, which is not yet consistently reported in PATSTAT throughout the 1970s.[10]

To identify which Xerox patents were subject to compulsory licensing, I use a list published in the Official Gazette of the United States Patent and Trademark Office (USPTO, 1975). This list, which was published in compliance with the consent decree, reports the publication number and title of more than 2,600 patents owned by Xerox and Fuji Xerox as of 1975. One caveat is that the list does not allow to exactly determine which patents were licensable. Instead, according to its description, 'the list [...] is believed to include all of the patents available for licensing', but 'there are several patents included in the list to which the consent order is not applicable' (USPTO, 1975, p. 1665). To approximate the set of licensable Xerox patents, therefore, I consider all patents on the 1975 list as if they were subject to compulsory licensing.

3.1 Class-Level Analysis of Cumulative Innovation

For my main empirical approach, I construct a panel dataset from 1970 until 1985 that counts the annual number of patent applications in the US on the level of six-digit technology classes based on the Cooperative Patent Classification (CPC). Patent applications by Xerox and its subsidiaries are excluded from the sample. I exploit the classes' differential exposure to compulsory licensing, depending on the share of patents that were licensable in each class. Since the classification system is hierarchical, I can compare patenting across differentially affected six-digit technology (sub)classes within the same four-digit class.[11]

I use the following difference-in-differences (DiD) regression model to estimate the effect of compulsory licensing on cumulative innovation:

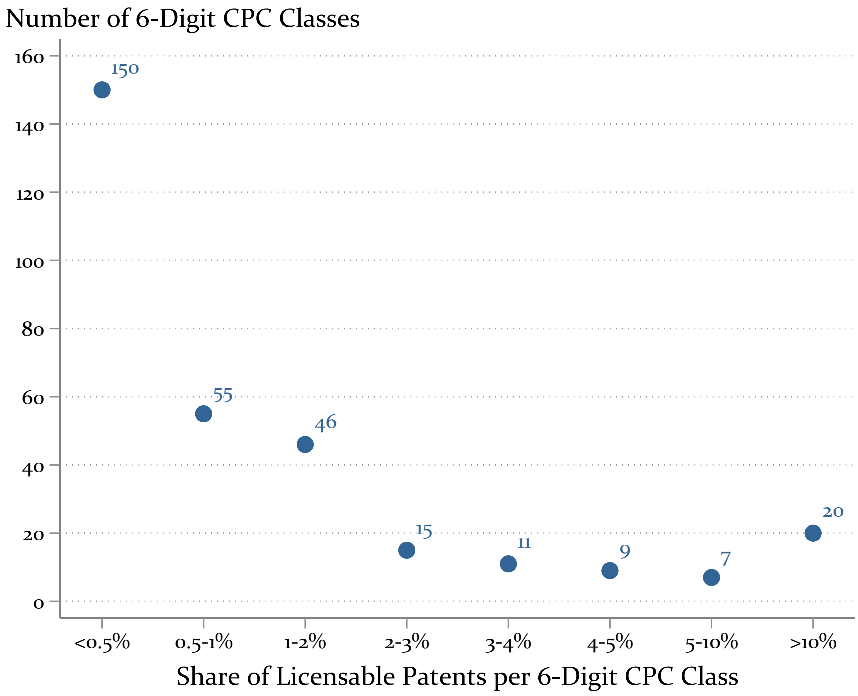
$$\text{Patents}_{c,s,t} = \beta \cdot \text{Share}_s \cdot \text{Post}_t + \alpha_s + \lambda_{c,t} + \epsilon_{c,s,t}, \quad (1)$$

where $\text{Patents}_{c,s,t}$ is the number of patent applications in year t assigned to six-digit subclass s within four-digit class c. Share_s is a continuous treatment variable that measures the exposure of a subclass to the antitrust intervention. It is defined as the share of unexpired patents per subclass (as of 1975) that were subject to compulsory licensing. That

[10] I thank Cyril Verluise for sharing their data.

[11] This class-level approach follows Moser and Voena (2012) and Moser et al. (2014). To highlight the hierarchical structure of the technology classes, I henceforth refer to a four-digit CPC class (e.g., G03G) as 'class' and to a six-digit CPC class (e.g., G03G 15) as 'subclass', although this slightly deviates from the official terminology. For an overview of the classification system, see `https://worldwide.espacenet.com/classification`.

Figure 1. Class-Level Analysis: Distribution of Treatment Variable (Share$_s$)



Notes: The figure depicts the distribution of the treatment variable (Share$_s$, i.e., the share of patents per six-digit subclass that were subject to compulsory licensing) across treated subclasses. It shows the number of subclasses in the sample (on the vertical axis) that have a given share of compulsorily licensed patents (on the horizontal axis). Example: there are 150 subclasses with a share of licensable patents below 0.5%.

is, for each subclass, the treatment variable captures the number of Xerox patents on the USPTO list relative to the overall size of the technology class. The variable Post$_t$ is a dummy that equals one in years after 1975. The regression also includes subclass fixed effects (α_s) as well as year $\times$ class fixed effects ($\lambda_{c,t}$). This controls for time-invariant differences across subclasses and allows the classes to experience idiosyncratic shocks over time. As a consequence, the DiD estimate $\hat{\beta}$ is identified only from variation over time across subclasses within the same class. Standard errors are clustered accordingly at the four-digit class level.

I estimate the DiD regression on a sample of 2,210 six-digit subclasses that belong to 141 four-digit classes. On average, there were 15 patent applications per year per subclass. 313 subclasses contain at least one Xerox patent that was subject to compulsory licensing – such that these subclasses are 'treated' (with Share$_s > 0$). The remaining 1,897 subclasses are 'untreated', as they had no exposure to compulsory licensing in the absence of spillovers.[12] I use the weights by Iacus et al. (2012) to adjust for the different number of untreated subclasses per treated subclass. Figure 1 shows the distribution of the treatment variable across subclasses. Among the 313 treated subclasses, there are 150 subclasses

[12] This sample is obtained after applying two restrictions. First, a six-digit subclass must have at least one patent application in the pre-treatment period. Second, every four-digit class must have at least one treated subclass.

with a share of licensable patents below 0.5%, while there are 20 subclasses where more than 10% of all patents were subject to compulsory licensing. Overall, the classes in the sample contain a total of 2,479 compulsorily licensed Xerox patents that were unexpired as of 1975.[13] Appendix A.1 presents additional summary statistics of the sample and for the treatment variable as well as more information on Xerox's patent portfolio.

There are several advantages to my approach of using the share of compulsorily licensed patents per subclass as treatment variable. First, the approach builds on the simple economic intuition that any effect of compulsory licensing on innovation should be more pronounced in technologies with a greater exposure to the antitrust measure. In contrast, using a binary treatment indicator would handle all treated subclasses equally. Second, using shares instead of the absolute number of licensable patents also takes into account the size of each subclass.[14] An additional nice feature of the approach is that the treatment variable can also be interpreted as capturing the propensity of a given subclass for containing xerography-related patents.

My empirical strategy identifies the causal effect of the antitrust measure under the assumption that, within the same four-digit technology class, patenting in subclasses with little or no exposure to compulsory licensing provides a valid counterfactual for patenting in subclasses with a greater share of licensable patents. In other words, the number of patent applications in subclasses with different values of the treatment variable must have followed a common trend in the absence of the antitrust intervention. The main concern with my identification strategy is that subclasses that were more exposed to compulsory licensing may have been different from less exposed subclasses in terms of unobserved characteristics, which may cause different patenting trends over time. For instance, Xerox may have chosen to patent in technology classes that had a higher likelihood of future innovation activity. To address this concern and assess the common-trend assumption, I also estimate the following event-study variation of equation (1):

$$\begin{aligned}\text{Patents}_{c,s,t} = &\sum_{\tau=1970}^{1974} \delta_\tau \cdot \text{Share}_s \cdot \mathbb{1}[\text{Year}_t = \tau] \\ &+ \sum_{\tau=1976}^{1985} \beta_\tau \cdot \text{Share}_s \cdot \mathbb{1}[\text{Year}_t = \tau] + \alpha_s + \lambda_{c,t} + \epsilon_{c,s,t},\end{aligned} \tag{2}$$

where the coefficients of interest are the lags (β_τ) that estimate the post-1975 treatment effects. In contrast, the leads (δ_τ) represent anticipatory effects and should not be statistically different from zero.

[13] This represents 95% of the patents on the USPTO list, because the list contains some patents that had expired by 1975, which I do not consider for the treatment definition.

[14] Nevertheless, I show in appendix A.4 that my results are robust to using a binary treatment variable or alternative continuous measures, where the treatment variable captures the number (as opposed to the share) of licensable Xerox patents per subclass.

3.2 Patent-Level Analysis of Follow-on Innovation

In a complementary approach, I use a different empirical set-up on the patent level to study direct follow-on innovation building on Xerox's technology. In particular, I investigate whether the number of forward citations to compulsorily licensed Xerox patents increased after 1975. Following Watzinger et al. (2020), I use exact matching to construct a control group for every licensable Xerox patent. I match on grant year, four-digit CPC class, and the aggregate number of citations until 1972 (i.e., the start of the antitrust case) or for at least two years.[15] Conditioning on the first two variables controls for differences in citations patterns over the patent term and across technologies. Conditioning on the number of previous citations additionally controls for how much a patent is used by other firms, although it does not necessarily imply that matched patents are of the same underlying quality (Watzinger et al., 2020).

My final sample consists of 1,311 compulsorily licensed Xerox patents matched to 25,899 control patents in 445 distinct strata. The number of matched Xerox patents is lower than the total number of compulsorily licensed patents for several reasons. Most importantly, I only match compulsorily licensed Xerox patents that were granted until 1972 (instead of 1975). This restriction circumvents concerns that Xerox may have strategically changed its patenting behaviour after publication of the FTC complaint. Similarly, matching only on citations until 1972 ensures that the matching procedure is unlikely to be confounded by endogenous changes in citation patterns. Moreover, I drop self-citations and only consider patents that received at least one citation during their term of validity. Among the 1,346 resulting Xerox patents that had not expired by 1975, 97% can be matched to at least one control patent. For the sample of matched patents, I then construct a panel that counts the number of forward citations received by every patent in every year from 1970 until 1985 or until patent expiry.

I estimate the following DiD regression model to estimate the effect of compulsory licensing on subsequent citations:

$$\text{Citations}_{i,t} = \beta \cdot \text{Xerox}_i \cdot \text{Post}_t + \alpha_i + \lambda_t + \epsilon_{i,t}, \tag{3}$$

where $\text{Citations}_{i,t}$ is the number of forward citations received by patent i in year t. Xerox_i is a dummy variable that equals one for compulsorily licensed Xerox patents and zero for matched control patents. The variable Post_t again is a dummy that equals one in years after 1975. The regression also includes patent fixed effects (α_i) and year fixed effects (λ_t). This controls for time-invariant differences in the number of citations across patents

[15] That is, for Xerox patents granted in 1971 and 1972, I match on the aggregate number of citations until 1973 and 1974, respectively. This is necessary because matching on citations becomes meaningless when using less than two years of data. However, my results are robust to only using Xerox patents granted until 1970, which I match on the aggregate number of citations until 1972.

and allows for year-specific shocks that affect all patents equally. I again use the weights by Iacus et al. (2012) to adjust for the different number of matched control patents per treated Xerox patent. Standard errors are clustered at the four-digit technology class level to account for potential serial correlation in citations within such a class.

The identifying assumption for this patent-level analysis is that Xerox patents would have received the same number of citations as their matched control patents in the absence of compulsory licensing. As pointed out by Watzinger et al. (2020), one concern regarding this common-trend assumption is that the authorities may have chosen to license patents with a high potential for follow-on innovation. However, official FTC documents on the Xerox case do not support this concern. The FTC did not attempt to promote innovation; instead, its main objective was to stop Xerox from using its patents to block entry into the product market for plain-paper copiers (FTC, 1975). To further address the concern, I also estimate an event-study variation of equation (3) with leads and lags.

4 Effect of the Antitrust Case on Innovation

This section introduces the main results of my empirical analysis. I first present estimates from the main approach on the class level, followed by estimates from the complementary approach on the patent level. The following section then investigates which firms benefited from access to Xerox's technology.

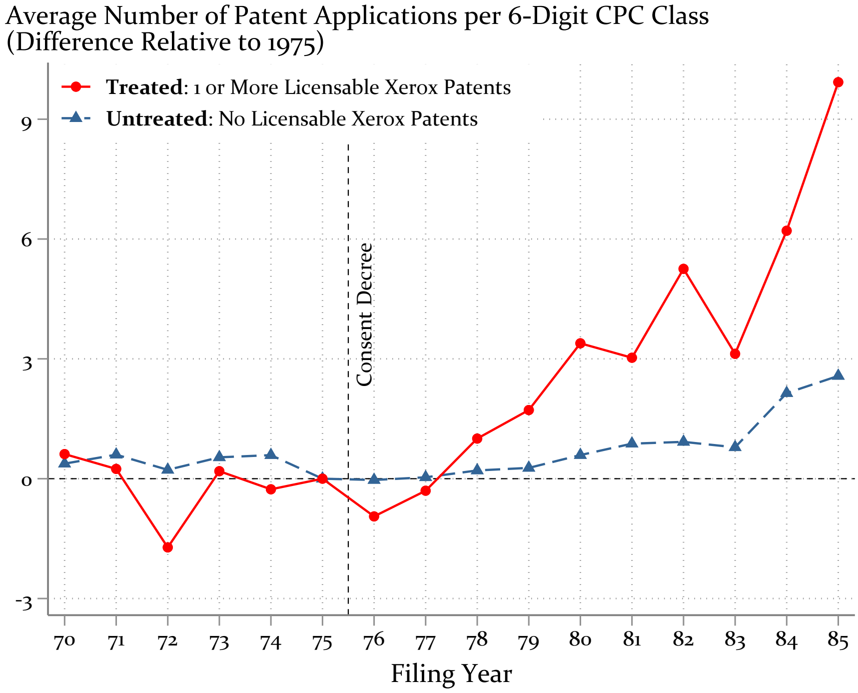
4.1 Cumulative Innovation

Panel (A) of Figure 2 depicts the average number of patent applications per six-digit technology subclass separately for (treated) subclasses with at least one compulsorily licensed Xerox patent and the remaining (untreated) subclasses in the sample. Although this figure solely presents averages, it shows that treated subclasses experienced a relative increase in the number of patent applications after 1975. In contrast, in the years preceding the consent decree, patenting in both groups followed a relatively common trend – consistent with the identifying assumption underlying my empirical strategy.

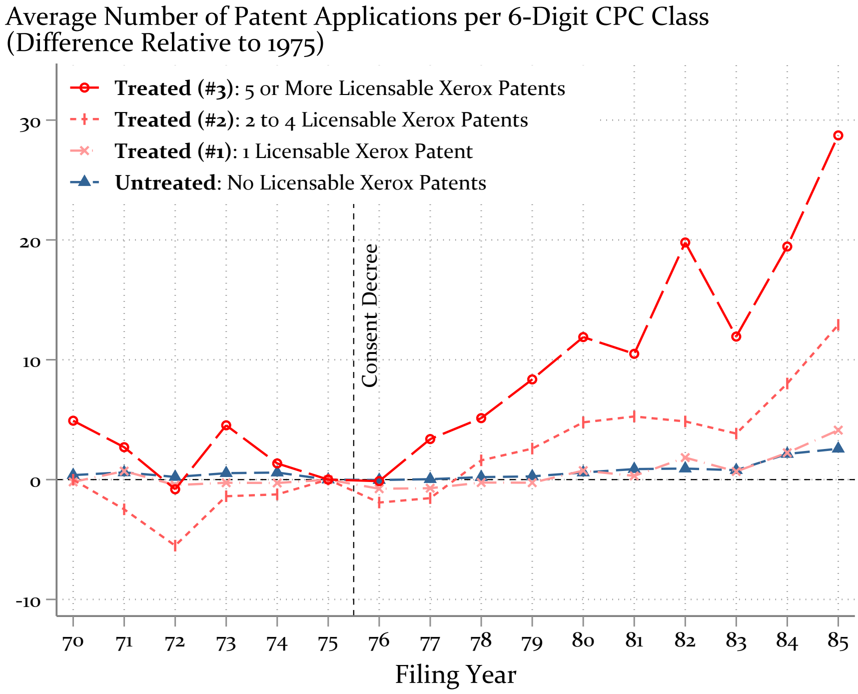
Next, I assess which treated subclasses were responsible for the relative increase in patenting following the antitrust intervention. To this end, in panel (B) of Figure 2, I split the treated subclasses into three subgroups, respectively containing one, two to four, and five or more licensable Xerox patents. As is evident from the figure, the relative increase in patenting is most pronounced in subclasses in which five or more Xerox patents were compulsorily licensed. It is smaller in subclasses with two to four licensable Xerox patents, whereas there is virtually no change in patenting around 1975 in subclasses with only one licensable Xerox patent. Finally, it is reassuring that pre-trends remain mostly

Figure 2. Class-Level Analysis: Patenting Trends Across Technology Classes

(A) Treated vs. Untreated Subclasses



(B) By Number of Licensable Patents



Notes: The figure depicts the average number of patent applications by firms other than Xerox per six-digit subclass relative to 1975. In panel (A), averages are computed separately for treated and untreated subclasses, where a subclass is defined as treated if it contained at least one compulsorily licensed Xerox patent. In panel (B), treated subclasses are further divided into three subgroups, containing (#1) one, (#2) two to four, and (#3) five or more licensed Xerox patents, respectively. In both panels, the subclasses are aggregated using the weights by Iacus et al. (2012). Note that the scale differs between panels (A) and (B).

Table 1. Class-Level Analysis: Baseline Estimates

	(1)	(2)	(3)
$\text{Share}_s \cdot \text{Post}_t$	0.210***	0.189**	
	(0.045)	(0.094)	
$\text{Share}_s \cdot \text{Post}_t \cdot \mathbb{1}[\text{Lic}_s = 1]$			-0.035
			(0.080)
$\text{Share}_s \cdot \text{Post}_t \cdot \mathbb{1}[2 \leq \text{Lic}_s \leq 4]$			0.085
			(0.097)
$\text{Share}_s \cdot \text{Post}_t \cdot \mathbb{1}[\text{Lic}_s \geq 5]$			0.377***
			(0.141)
Subclass FE	✓	✓	✓
Year FE	✓		
Year × Class FE		✓	✓
Mean of Outcome	15.13	15.13	15.13
4-Digit CPC Classes	141	141	141
Observations	35360	35360	35360

Notes: The table shows the results from difference-in-differences regressions following variations of equation (1). The outcome variable in all regressions is the number of patent applications by firms other than Xerox in a given six-digit CPC subclass and year. In column (3), the treatment variable is interacted with indicators for subclasses with one, two to four, and five or more compulsorily licensed Xerox patents, as indicated by the variable Lic_s. Standard errors clustered at the four-digit CPC technology class level are in parentheses. Significance levels: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

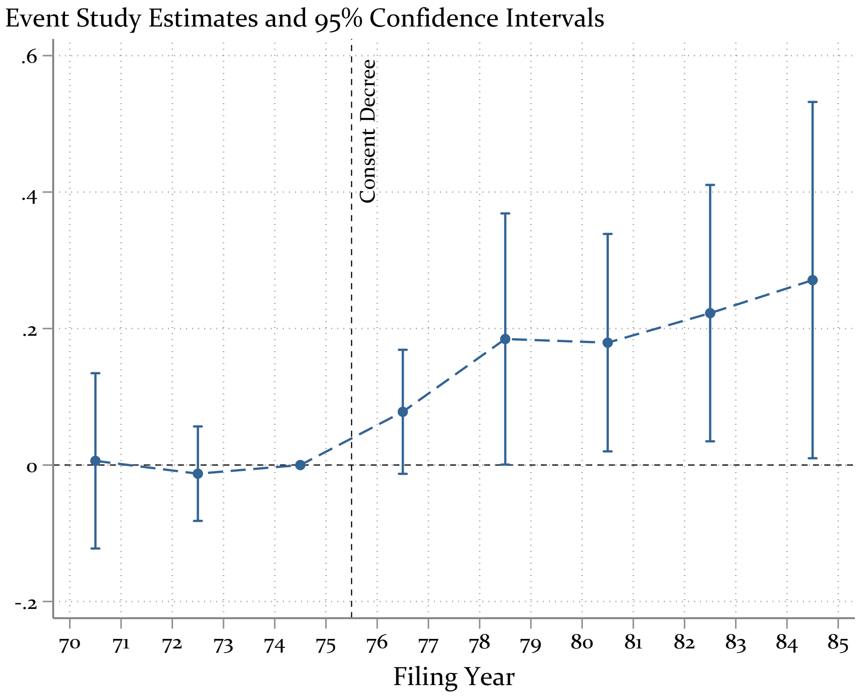
parallel even across the three treated subgroups with differential exposure to the antitrust measure.

I now turn to a regression framework to investigate the impact of compulsory licensing on innovation more systematically. As outlined above, my main approach compares patenting across six-digit subclasses with differential exposure to the antitrust measure within the same four-digit class. I use a continuous treatment specification, inspired by the pattern shown in panel (B) of Figure 2.

Table 1 presents the results from estimating different variations of the DiD regression model in equation (1). Column (1) shows the estimate when only controlling for subclass and year fixed effects. The point estimate is positive and highly statistically significant. My baseline specification is given by column (2). Adding year × class fixed effects slightly reduces the magnitude of the DiD estimate. The point estimate in column (2) indicates that, on average, a one percentage-point higher share of compulsorily licensed Xerox patents in a subclass is associated with 0.19 additional patent applications per year in that subclass after 1975.[16] This baseline estimate is statistically significant at the 5% level. In column (3), I again split the treated subclasses into three subgroups. Consistent with the pattern presented in panel (B) of Figure 2, the estimates show that the increase in patent

[16] I define the variable Share_s in percentage terms (i.e., × 100). Therefore, the estimate $\hat{\beta}$ in equation (1) can be interpreted as the average annual post-1975 increase in patenting per subclass that corresponds to a one percentage-point increase in the share of licensable patents.

Figure 3. Class-Level Analysis: Event-Study Estimates



Notes: The figure depicts point estimates and 95% confidence intervals from the event-study analysis in equation (2). Patent applications are binned in two-year groups to reduce noise in the estimates. All regressions use the weights by Iacus et al. (2012). Standard errors are clustered at the four-digit CPC technology class level.

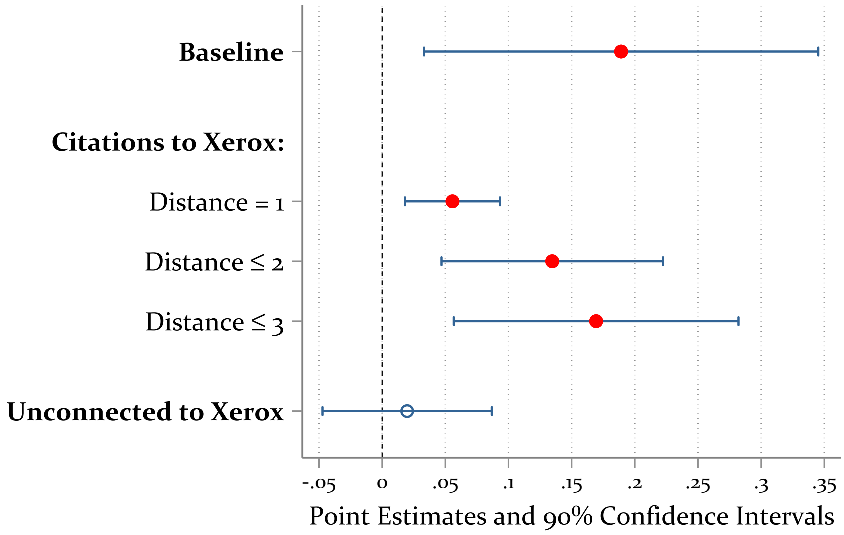
applications is driven by subclasses in which five or more Xerox patents were compulsorily licensed. This result is reassuring, as it implies that the share of compulsorily licensed patents in a given subclass only affected subsequent patenting if the absolute number of licensable patents in that subclass was sufficiently large.

The magnitude of my baseline estimate in column (2) of Table 1 corresponds to around 160 additional patents per year following the antitrust case.[17] This represents an economically meaningful effect, given that the average monetary value of a single patent in treated technologies was more than $20 million in today's dollars (Kogan et al., 2017). Relative to the overall number of patent applications in treated subclasses, my estimates imply that patenting increased by around 1.4%. Yet, it should be taken into account that patents in treated subclasses, which this percentage estimate is based on, also cover inventions for products other than copiers. If it were possible to identify copier-specific patents, the percentage increase would likely be much larger.

Figure 3 graphically depicts the point estimates and 95% confidence intervals from the event-study analysis in equation (2), corresponding to the simple DiD estimate in column (2) of Table 1. The figure shows that there was a disproportionate increase in patenting in treated technologies following the 1975 consent decree. That is, on average,

[17] This estimate is obtained by multiplying the point estimate with the share of licensable patents (i.e., the treatment variable) in every subclass and then aggregating.

Figure 4. Class-Level Analysis: (Indirect) Citations to Licensable Xerox Patents



Notes: The figure depicts point estimates and 90% confidence intervals from estimating the regression model in equation (1). In the baseline, the outcome variable is the overall number of patent applications by firms other than Xerox per subclass and year. In the remainder of the figure, the outcome variable only considers patent applications that built on compulsorily licensed Xerox patents through citations of different degrees, following the distance framework by Ahmadpoor and Jones (2017). I refer to patents as 'unconnected to Xerox' if they have a distance ≥ 4 to Xerox. Red dots (blue circles) indicate statistically significant (insignificant) point estimates. Standard errors are clustered at the four-digit CPC technology class level.

subclasses with a greater exposure to compulsory licensing experienced a greater relative increase in the number of patent applications. In contrast, before 1975, the number of patent applications across differentially treated subclasses followed a relatively common trend. As can be seen in Figure A1 in the appendix, pre-trends also remain parallel when extending the panel back to 1960. This supports the identifying assumption underlying my empirical strategy.

In the next step, I restrict the outcome variable to patent applications that built on compulsorily licensed Xerox patents through citations. This serves to investigate whether the estimated increase in patenting is, in fact, related to compulsory licensing of Xerox's patents. The corresponding DiD estimates are shown graphically in Figure 4. The first row repeats the baseline estimates from column (2) of Table 1. In the second row, the outcome variable is restricted to patent applications that directly cite a compulsorily licensed Xerox patent. Following the framework by Ahmadpoor and Jones (2017), these patents are defined to have a distance = 1 to Xerox. Although the point estimate remains statistically significant, its magnitude is only around 30% of that of the baseline. This highlights that only a fraction of the additional patents filed in treated subclasses after 1975 directly built on Xerox's technology. Therefore, in the following two rows, I include patent applications that are related to compulsorily licensed Xerox through higher-degree citations. Patents with distance = 2 (= 3) do not directly cite Xerox but cite a patent with distance = 1 (= 2). The estimated coefficients increase when higher-degree citations are included. When using patents with distance ≤ 3 as outcome variable, the magnitude of the point estimates is close to that of the baseline. This result suggests that only looking at direct citations to Xerox's patents may not capture the entire impact of the antitrust measure, highlighting the cumulative nature of innovation. Finally, the last row of Figure 4 shows the point estimate when restricting the outcome variable to patent applications that are unconnected to Xerox (i.e., with distance ≥ 4). This serves as a placebo check and, reassuringly, the estimate is close to zero and statistically insignificant.

I run a number of additional robustness checks to ensure that my main results are not driven by a specific model, treatment, or sample specification. The various results are reported in appendix A. For example, I find that my estimates are robust to estimating a Poisson pseudo-likelihood regression (instead of an ordinary least squares model) and to excluding the subclasses with the highest share of compulsorily licensed patents from the sample. In addition, I employ both a binary treatment measure and another continuous treatment specification that is based on the number (as opposed to the share) of compulsorily licensed Xerox patents per subclass. Using these alternative treatment definitions does not affect the main results. I also show that my findings are robust to aggregating treated and untreated six-digit subclasses within a four-digit class as in Watzinger et al. (2020) and Watzinger and Schnitzer (2022).

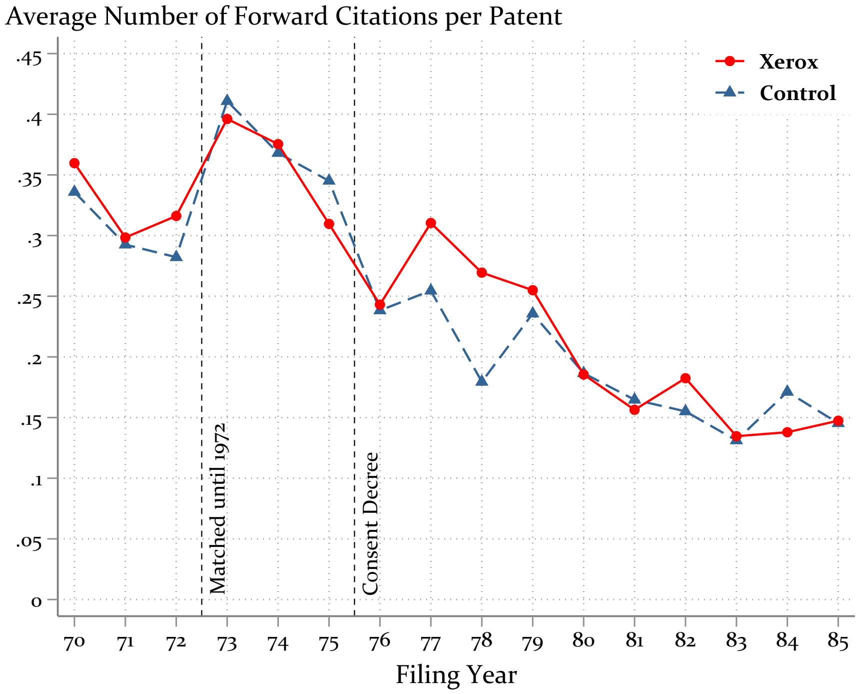
Overall, the empirical evidence suggests that compulsory licensing of Xerox's patents had a positive and statistically significant effect on subsequent innovation. There was a disproportionate increase in patenting in technologies where a larger share of Xerox patents became available for licensing. I now turn to the results of my complementary analysis on the patent level, which investigates direct follow-on innovation to compulsorily licensed Xerox patents.

4.2 Direct Follow-On Innovation

Figure 5 depicts the average number of forward citations per patent separately for compulsorily licensed Xerox patents and matched control patents. The figure suggests that Xerox patents experienced a relative increase in citations in the years 1977 to 1979. For instance, the difference in 1978 represents an increase in citations by roughly 50% and implies that, on average, there was one additional citation for every ten Xerox patents in that year. In most other years, the average number of citations across both groups closely tracks each other. The common trend until 1972 is mostly by construction due to the matching technique. Yet, the development until 1975 speaks in favour of the identifying assumption that the number of citations would have followed a common trend even in the absence of compulsory licensing.

This first visual impression of an increase in citations to Xerox patents is also confirmed by the event-study estimates shown in Figure 6. While the pre-treatment estimates are

Figure 5. Patent-Level Analysis: Citations to Xerox vs. Control Patents



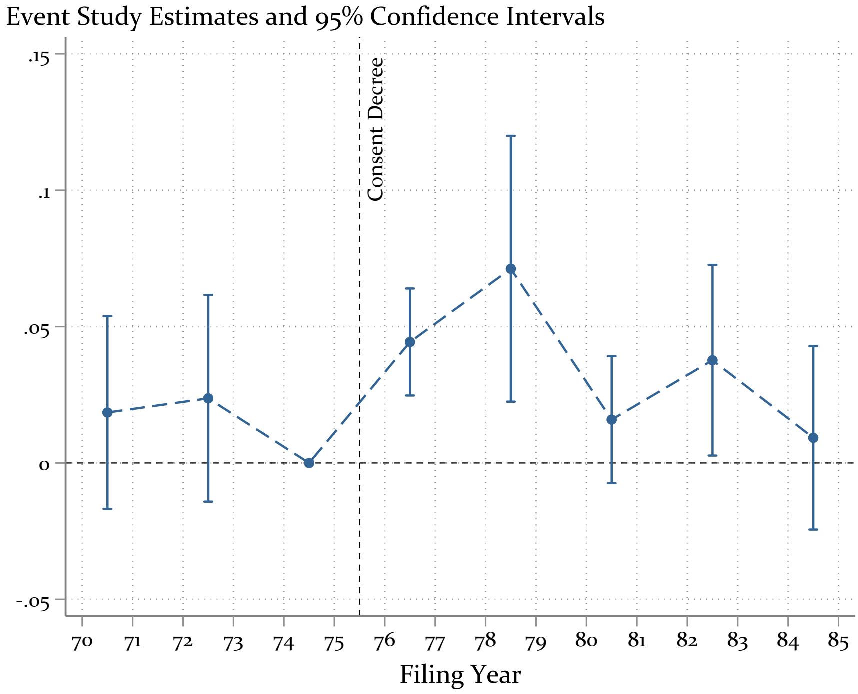
Notes: The figure depicts the average number of forward citations per patent. Self-citations are not taken into account. Averages are computed separately for compulsorily licensed Xerox patents and matched control patents. The patents are aggregated using the weights by Iacus et al. (2012).

not statistically distinguishable from zero, the estimates in the years following the consent decree show a positive effect of compulsory licensing on the number of forward citations received by Xerox patents. The effect then fades out again, indicating that the impact of compulsory licensing on direct follow-on innovation may be rather short-lived. This pattern is consistent with the findings by Watzinger et al. (2020). It also goes in line with my previous result from Figure 4, showing that direct citations to compulsorily licensed Xerox patents only explain a fraction of the overall effect on cumulative innovation.

The baseline DiD estimate from the patent-level analysis is shown in column (1) of Table B1 in the appendix. The point estimate indicates that, on average, every compulsorily licensed Xerox patent received an additional 0.02 citations per year after 1975 relative to matched control patents. The estimate is statistically significant at the 1% level.

Overall, the results from this complementary analysis on the patent level are in line with the findings from my main approach on the class level. After the 1975 consent decree, there was not only a disproportionate increase in patenting in fields where Xerox's technology became available for licensing to competitors; Xerox patents also experienced a relative increase in citations. I conclude that compulsory licensing of Xerox's patents had a positive effect on innovation, as other firms used the newly available technology for follow-on research and either directly or indirectly built on Xerox's patents.

Figure 6. Patent-Level Analysis: Event-Study Estimates



Notes: The figure depicts point estimates and 95% confidence intervals from an event-study variation of the model in equation (3). Citations are binned in two-year groups to reduce noise in the estimates. The regression uses the weights by Iacus et al. (2012). Standard errors are clustered at the four-digit CPC technology class level.

5 Which Firms Benefited?

I now investigate which firms benefited from access to Xerox's technology. In the first step, I split the number of patent applications by the applicant's country of origin, hence analysing where follow-on innovators were located. In the second step, I look at heterogeneity by firms' previous patenting behaviour. This allows me to asses whether the effect is driven by firms with prior experience in copier technologies – that is, firms that could become direct competitors to Xerox.

5.1 Effects by Applicant Country

Table 2 presents estimates of the effect of compulsory licensing on patenting by applicants from different countries. The table shows DiD estimates from my main class-level approach where the outcome variable (i.e., the number of patent applications) is split by applicant country. The results are striking: column (2) indicates that compulsory licensing of Xerox's patents had a quantitatively small and statistically insignificant effect on patenting by US applicants. In contrast, the positive and significant baseline estimate from column (1) is entirely driven by patenting by non-US firms, as shown in column (3). Amongst them, applicants from Japan were the driving force behind the increase in pat-

Table 2. Class-Level Analysis: Heterogeneity by Applicant Country

	Baseline	Applicant Country			
		USA	Non-USA	Among Non-USA	
				Japan	Others
	(1)	(2)	(3)	(4)	(5)
$\text{Share}_s \cdot \text{Post}_t$	0.189**	0.029	0.162**	0.143**	0.020
	(0.094)	(0.038)	(0.073)	(0.064)	(0.013)
Mean of Outcome	15.13	8.93	5.74	2.25	3.49
4-Digit CPC Classes	141	141	141	141	141
Observations	35360	35360	35360	35360	35360

Notes: The table shows the results from difference-in-differences regressions following equation (1). Column (1) repeats the baseline estimates from Table 1. In columns (2) to (5), the outcome variable is restricted to patent applications filed by assignees from selected countries. All regressions include subclass and year × class fixed effects. Standard errors clustered at the four-digit CPC technology class level are in parentheses. Significance levels: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

enting. The estimate in column (4) indicates that, among the foreign applicants, Japanese firms accounted for more than 85% of the additional patent applications after 1975. This effect is statistically significant at the 5% level. The magnitude of the effect implies that compulsory licensing of Xerox's patents increased Japanese patenting in treated technologies by around 5.4% per year. Column (5) shows that other non-US countries, if anything, experienced a quantitatively small increase in patenting.[18] The results are similar when repeating this exercise with my complementary approach on the patent level, as shown in Table B1 in the appendix.

These estimates suggest that, in terms of subsequent innovation, Japanese firms were the main beneficiaries of antitrust action against Xerox. Yet, one concern about these estimates could be that increased patenting by Japanese firms in the US may not necessarily represent novel innovation. It may also reflect that Japanese firms started to seek protection for their existing technologies in a foreign country where they previously did not enter the product market. If this were the case, I should not find any treatment effect when using data on domestic patent applications in Japan. To address this concern, therefore, I repeat my analysis with data on patent applications at the Japanese Patent Office (instead of at the USPTO). The data were obtained from the Japanese Institute of Intellectual Property (Goto and Motohashi, 2007). As shown in Table A6 in the appendix, the disproportionate increase in patenting in technologies where Xerox patents became available for licensing is also present in domestic patent applications in Japan. In addition, this effect is again driven by increased patenting by Japanese firms, whereas the coefficient on patenting by US firms is small in magnitude and statistically insignificant.

[18]The important role of Japanese applicants becomes even more apparent when comparing the DiD estimates to the average number of patent applications per country, which is even smaller for Japan than for other non-US countries. That is, in percentage terms, the effect on Japanese patenting is even more pronounced.

All in all, this robustness check confirms the interpretation of my results as increased innovation by Japanese firms.

Finally, it should be noted that interventions by the Japanese government – such as the Ministry of International Trade and Industry – played a very limited role in the development of the Japanese copier industry (Jacobson and Hillkirk, 1986). Therefore, Japanese copier producers equally had to rely on patents to protect their IP both domestically and abroad, even prior to the antitrust case against Xerox.

5.2 The Role of Prior Experience in Copier Technologies

Next, I investigate whether the firms that benefited from compulsory licensing were (potential) competitors to Xerox in the copier market. This need not necessarily be the case, as Xerox's patents covered some basic technologies that could also be used outside of the copier industry. For example, one application entirely unrelated to copiers is xeroradiography, a specific X-ray technique. Moreover, Watzinger et al. (2020) show that compulsory licensing in the case of Bell led to an increase in innovation only outside of telecommunications, which was Bell's core industry. As Bell was a vertically integrated company, it could still foreclose rivals in the telecommunications industry. However, the market structure in the case of Xerox was fundamentally different. Xerox's patents were the main entry barrier in the plain-paper copier market. Therefore, one would hypothesise that the removal of patent protection should allow Xerox's competitors to use the newly available technology for follow-on innovation.

To identify potential entrants into the copier market, I compute a measure of firms' closeness to Xerox based on their prior patenting experience. I define the variable

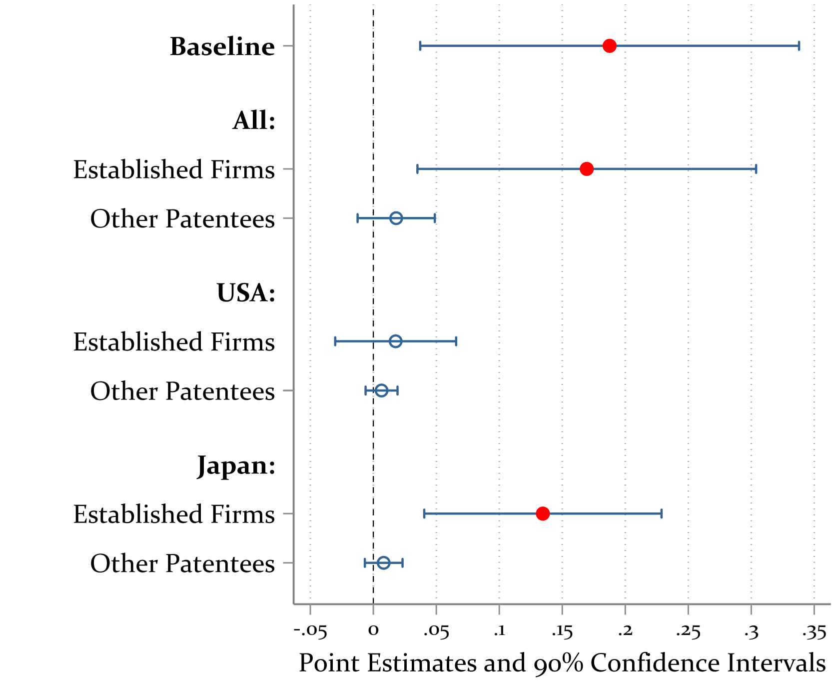
$$\text{Closeness}_i = \sum_s \text{w}_{is} \cdot \text{Share}_s, \tag{4}$$

where w_{is} is the share of firm i's unexpired patents (as of 1975) that are in subclass s. Share_s is the treatment variable from equation (1) and represents the share of patents in a given subclass that were subject to compulsory licensing. Therefore, the firm-level variable Closeness_i represents the degree to which a firm's existing patent portfolio overlaps with the set of compulsorily licensed Xerox patents. Summary statistics for the closeness measure are presented in Table A3 in the appendix. I make two sample restrictions to use this closeness measure. First, I only consider patents by firms, hence excluding patent applications filed by individuals, universities, or government bodies. Second, I only consider firms that filed at least ten patent applications from 1970 until 1975 – which I define as 'established' firms.[19] The resulting firm sample consists of 1,635 firms.

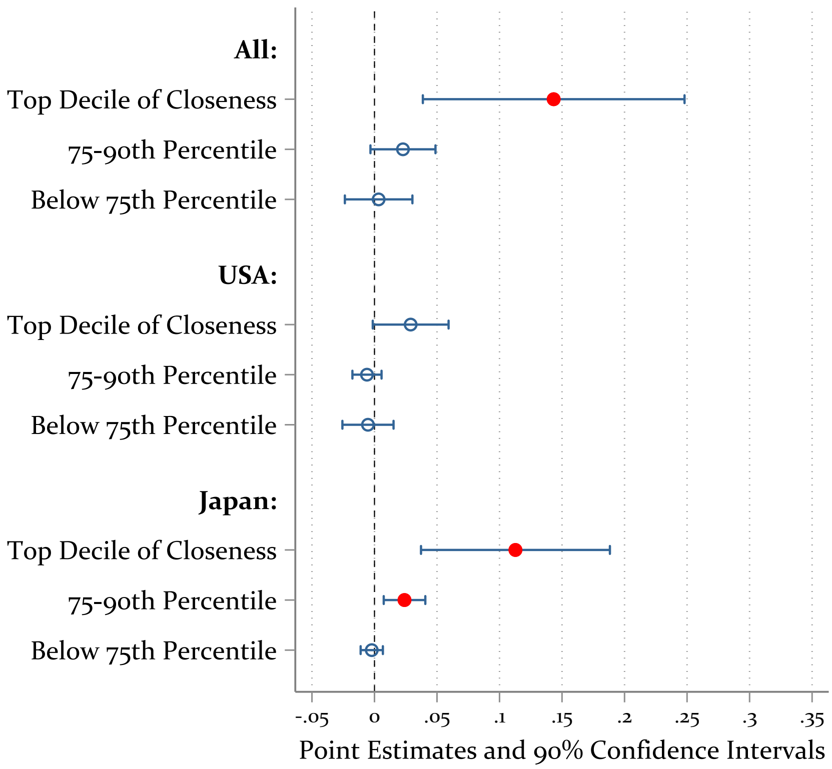
[19]For firms with very few patent applications in the pre-treatment period, the weighted sum in equation (4) becomes meaningless. However, my results do not hinge on this specific definition of the closeness

Figure 7. Class-Level Analysis: Heterogeneity by Firms' Patenting Experience

(A) Established Firms vs. Other Patentees



(B) Closeness to Xerox



Notes: The figure depicts point estimates and 90% confidence intervals from estimating the regression model in equation (1). In panel (A), the outcome variable (i.e., the number of patent applications) is split by applicant country as well as by whether the assignee is part of the restricted sample of 'established firms' with at least ten patent applications from 1970 until 1975. All remaining patent applications are labelled as coming from 'other patentees'. Panel (B) employs the closeness measure from equation (4). The outcome variable is split by applicant country and by the applicant's percentile in terms of the distribution of the closeness measure across all firms in the restricted sample. Red dots (blue circles) indicate statistically significant (insignificant) point estimates. Standard errors are clustered at the four-digit CPC technology class level.

Panel (A) of Figure 7, which depicts DiD estimates from my main class-level approach, shows that the sample of established firms accounts for almost the entire post-1975 increase in patenting. This is evident from the point estimate in the second row, where the outcome variable is restricted to patent applications by established firms. This result is not obvious, since established firms only account for around 57% of all patent applications in my baseline sample. The estimate in the third row of panel (A) further shows that all remaining patentees did not experience a significant increase in patenting in technologies exposed to compulsory licensing. This indicates that antitrust action against Xerox did not benefit start-ups or other small firms. The remaining estimates in panel (A) of Figure 7 highlight that the increase in patenting among established firms is almost exclusively driven by Japanese firms.

Panel (B) of Figure 7 now employs the closeness measure to further investigate which firms benefited from compulsory licensing. For that purpose, I split firms into three groups according to their percentile in the distribution of the closeness measure (across all countries). I then repeat my class-level DiD approach with the outcome variable restricted to patent applications by firms from each group. The estimates in the top three rows of panel (B) indicate that the observed increase in the number of patent applications after 1975 is due to firms with a high degree of prior experience in technologies related to Xerox. In other words, compulsory licensing of Xerox's patents promoted innovation primarily within the target industry.[20] The remainder of panel (B) of Figure 7 repeats the analysis separately for applicants from the US and Japan. Among the American firms, the point estimate is positive only for firms with the highest technology overlap with Xerox, but this effect is quantitatively small and not statistically significant. For Japan, the estimates highlight that there is great heterogeneity in the effect of the antitrust case even among Japanese patent applicants. The positive innovation effect is particularly driven by firms in the top decile of the distribution of the closeness measure. In contrast, firms below the top quartile of the closeness measure did not experience any change in their patenting in either country. Overall, Figure 7 highlights that the main beneficiaries of the antitrust case were those Japanese firms that had extensive prior knowledge in copier technologies. Accordingly, they could use the technology available for licensing from 1975 onwards for follow-on innovation.

measure or the sample. Moreover, in appendix A.7, I present an alternative approach to estimating heterogeneity by prior patenting experience, which leads to similar results as Figure 7.

[20] Strictly speaking, the results from Figure 7 are not indicative about whether the firms that benefited from compulsory licensing were, in fact, active in the same product market as Xerox. However, there are two reasons why I define closeness to Xerox based on prior technology experience rather than product market activity. First, I believe that firms' prior patenting better captures their potential to compete in the copier market. Second, and relatedly, empirically identifying Xerox's (potential) competitors based on industry assignment (e.g., Compustat segments data) is complicated. This is because entrants into the copier market originally operated in very different industries, ranging from photographic equipment (e.g., Canon, Kodak, Konica) to computing (e.g., IBM) to consumer electronics (e.g., Sharp, Toshiba).

The increase in innovation among Japanese competitors is in line with the historical events in the copier industry. Starting in the mid-1970s, several Japanese companies – including Canon, Konica, Minolta, Ricoh, Sharp, or Toshiba – entered the US copier market with great success and became important competitors to Xerox (e.g., Jacobson and Hillkirk, 1986; Gomes-Casseres and McQuade, 1991; Scherer, 2005). Reassuringly, these Japanese firms are all located in the top decile of the distribution of the closeness measure, hence supporting my interpretation of this measure as identifying potential competitors in the copier market.[21]

In summary, the heterogeneity analyses reveal two important results. First, compulsory licensing in the case of Xerox was effective at increasing innovation in the copier industry. That is, the antitrust intervention allowed established firms with prior experience in copier technologies to increase their patenting. However, second, this is true primarily for Japanese competitors, which were the main beneficiaries of the antitrust measure in terms of their subsequent innovation performance. To further corroborate this result, I present additional firm-level evidence in the next step.

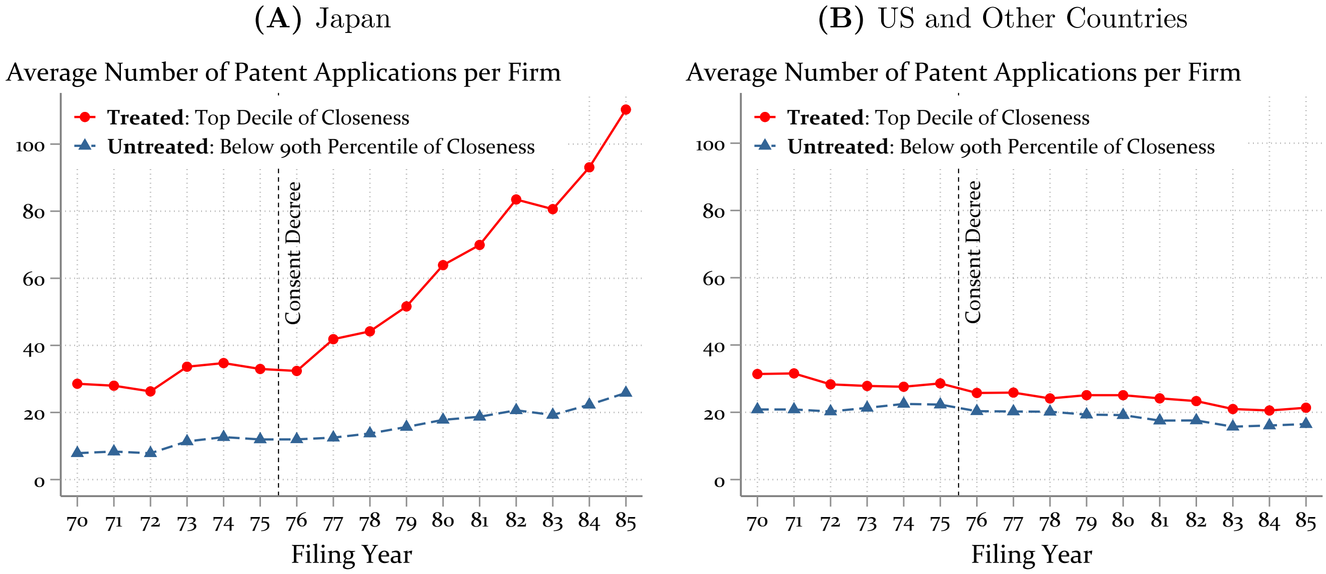
5.3 Firm-Level Results on the Effect on Japanese Competitors

I turn to the firm level to further investigate how Japanese competitors changed their patenting behaviour following the consent decree. Figure 8 depicts the average number of patent applications per firm separately for Japan and all other countries (including the US) in panels (A) and (B), respectively. Within each panel, I split firms into two groups and define a firm as treated if it is in the top decile of the distribution of the closeness measure. The figure reveals that Japanese firms were on a different patenting trend from the rest of the sample throughout the relevant period. While the average number of patent applications per firm in the US and other countries was mostly constant or even slightly decreased over time, patenting among Japanese firms steadily increased from 1970 until 1985. However, Japanese firms in the top decile of the distribution of the closeness measure experienced a much stronger increase in patenting after 1975 than the remaining Japanese firms in the sample. In other words, Japanese firms with a large amount of prior experience in copier technologies disproportionately increased their innovation activities after Xerox's patents became available for licensing. In other countries, in contrast, patenting by firms with a greater technology overlap with Xerox did not evolve differently over time.

These descriptive results demonstrate that the effect on Japanese competitors in my class-level approach does not simply reflect an aggregate increase in Japanese patenting. This is reassuring and corroborates my key empirical findings. In appendix C.2, I further investigate whether differences in observable firm characteristics may explain the heterogeneous effect across countries, but I find no evidence supporting this hypothesis.

[21] Among the 163 firms (i.e., 10%) that are in the top decile of the distribution of the closeness measure, 26 firms are from Japan. A list of these Japanese firms is provided in Table C1 in the appendix.

Figure 8. Firm-Level Analysis: Patenting Trends Across Firms



Notes: The figure depicts the average number of patent applications per firm. Panel (A) includes firms from Japan, whereas panel (B) includes all remaining firms in the sample of established firms. Averages are computed separately for treated and untreated firms, where a firm is defined as treated if it is located in the top decile of the distribution of the closeness measure defined in equation (4).

Therefore, my additional analyses on the firm level indicate that the positive effect on Japanese innovation represents, in fact, a phenomenon that is idiosyncratic to a specific group of Japanese firms that had extensive prior knowledge in copier technologies.

6 Mechanism

Why were Japanese copier producers more successful in building on Xerox's technology than their American counterparts? I now address this question by investigating the mechanisms underlying my results. First, I discuss historical narratives suggesting that Japanese entrants focused on producing smaller desktop copiers. I then study how compulsory licensing of Xerox's patents affected the quality and diversity of innovation.

6.1 The Japanese Focus on Smaller Desktop Copiers

Historical narratives indicate that American and Japanese firms entered the copier market with different strategies. On the one hand, American entrants such as IBM or Kodak started competing with Xerox in the same (high-volume) market segment where Xerox was dominant. For IBM, in particular, entering the high-volume segment may be explained by the potential to exploit economies of scale from the company's existing distribution network for mainframe computers. On the other hand, Japanese entrants strategically focused on the low end of the copier market by producing smaller and lower-volume plain-paper copiers than Xerox and most American competitors (e.g., Jacobson and Hillkirk, 1986; Porter, 1988; Gomes-Casseres and McQuade, 1991).[22] According to Jacobson and Hillkirk (1986, p. 105), key determinants of the Japanese success included entering the right market segment at the right time, standardising and externalising the production of inputs, and using automation to exploit economies of scale – all with the objective of building a 'value-added product that's simpler and cheaper to build and use'.

This indicates that the heterogeneous effect across countries is unlikely to be driven by some of the competitors being Japanese. Instead, in line with the arguments by Chesbrough and Rosenbloom (2002), Japanese entrants employed a different business model that expanded the market for plain-paper copiers to the lower-volume segment. This distinct competitive strategy is also well established in the management literature, where the Japanese entry into the American copier market is frequently named as an example of a successful attack on a dominant market leader (e.g., Porter, 1985; Paley, 2017).

Consistent with this possible mechanism, I show in Figure 9 that patents filed by Japanese competitors more frequently contained words associated with smaller desktop copiers. I focus on firms in the top decile of the distribution of the closeness measure and search the titles and abstracts of their patents for (variations of) one of the following words: compact, desktop, efficient, energy-saving, miniature, minuscule, portable, scale, small, simple, size, and tiny. As shown in Figure 9, the share of patent applications per firm with such 'small'-related words was roughly equal across countries in the early 1970s. Then, however, there was a divergence. The average share of Japanese patent filings that contained any 'small'-related word slightly increased up to around 12%, whereas the share steadily declined for competitors from the US and other countries.[23] Although this result is purely descriptive, it supports the narrative evidence about the role of the Japanese entrants' distinct business model – suggesting that a greater degree of product differentiation from existing copiers may (at least partly) explain the higher rate of innovation among Japanese competitors.

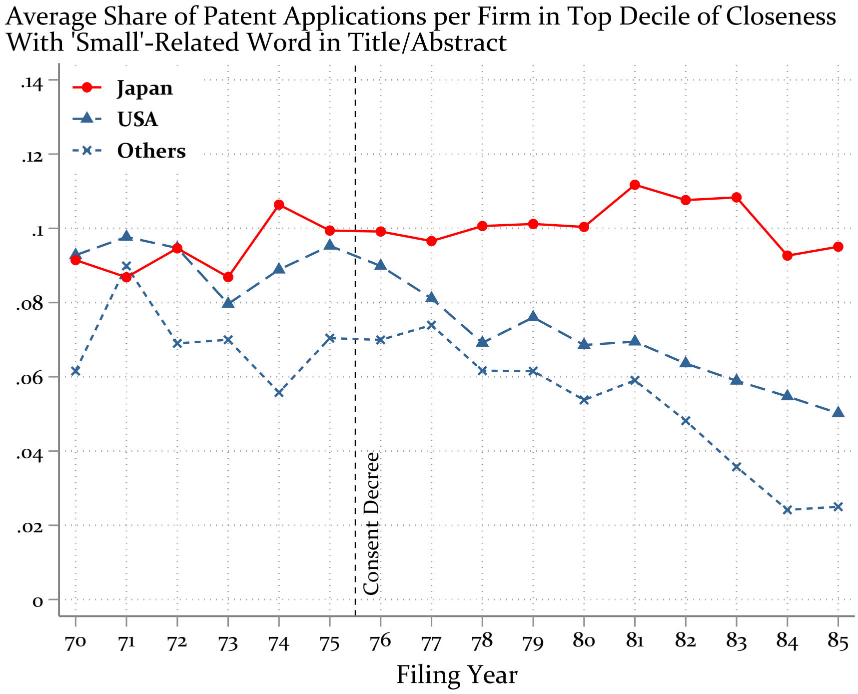
6.2 Effect on the Quality and Diversity of Innovation

Thus far, the paper focused on the effect of the antitrust case on the intensity of innovation. Yet, simple patenting numbers may not necessarily be informative about the quality and content of the underlying inventions. Therefore, I now analyse how compulsory licensing of Xerox's patents affected the quality and diversity of innovation.

[22] See also, 'Small Is Better, Xerox-san' in *The Washington Post* from 1981 (`https://www.washingtonpost.com/archive/business/1981/02/15/small-is-better-xerox-san/bb081ecb-1061-4492-8bd2-29aa5f91212d/`).

[23] Even if the Japanese's focus on lower-volume copies fully explained the heterogeneous effect across countries, it is unclear how this should affect the pattern in Figure 9. After all, building smaller copiers required inventions far beyond miniaturisation of existing technologies. Therefore, I consider Figure 9 as indicative that the Japanese entrant's distinct business model likely played a role; but it is inconclusive regarding the importance of this channel.

Figure 9. Firm-Level Analysis: Use of 'Small'-Related Words in Patents



Notes: The figure depicts the average share of patent applications per firm that contain a 'small'-related word in the patent title or abstract. The following words (and variations of them) are considered: compact, desktop, efficient, energy-saving, miniature, minuscule, portable, scale, small, simple, size, and tiny. The sample is restricted to firms in the top decile of the distribution of the closeness to Xerox. Averages are computed separately by country. A three-year moving average is applied to these averages to reduce noise.

To study patent quality, Table 3 makes use of two different patent measures: the number of forward citations that a patent has received and the ten-year quality measure by Kelly et al. (2021, KPST). The latter is constructed using textual analysis and based on the idea that high-quality patents should be novel relative to prior art but must have a high impact on future inventions. Citations, in contrast, only capture how much a given patent is used by subsequent patents. In columns (2) and (3) of Table 3, I restrict the outcome variable to patents in the 10% of the distribution of these two patent measures. The DiD estimates indicate that the newly filed patents after 1975 were of higher quality in terms of the KPST measure but not in terms of forward citations. In columns (4) and (5), the outcome variables are the average number of forward citations per patent and the average KPST quality measure per patent, respectively. This set-up leads to a similar result, finding a small positive effect only on the average KPST measure.

The estimates suggest that the additional patents filed after 1975 were more dissimilar from prior art but did not have a greater impact on subsequent patents, relative to the average patent. Both citations and the KPST measure capture how much a patent is used in subsequent patent filings. While citations rely on explicit references, the KPST measure identifies a forward similarity through textual analysis (Kelly et al., 2021). However, KPST then divide this impact measure by a patent's backward similarity (i.e., an indicator

Table 3. Class-Level Analysis: Patent Quality

		Patents in Top 10%		Mean	
	Baseline	Forward Citations	Quality (KPST)	Forward Citations	Quality (KPST)
	(1)	(2)	(3)	(4)	(5)
$\text{Share}_s \cdot \text{Post}_t$	0.189**	0.004	0.142**	0.106	0.002*
	(0.094)	(0.009)	(0.057)	(0.126)	(0.001)
Mean of Outcome	15.13	1.60	2.00	14.14	-0.00
4-Digit CPC Classes	141	141	141	141	141
Observations	35360	35360	35360	35360	35360

Notes: The table shows the results from difference-in-differences regressions following equation (1). Column (1) repeats the baseline estimates from Table 1. In columns (2) and (3), the outcome is the number of patent applications in the top 10% of the distribution of forward citations and the quality measure by Kelly et al. (2021, KPST), respectively. In columns (4) and (5), the outcome variables are the average number of forward citations per patent and the average KPST measure per patent, respectively. All regressions include subclass and year × class fixed effects. Standard errors clustered at the four-digit CPC technology class level are in parentheses. Significance levels: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

for novelty). A high KPST quality measure can, therefore, result from a low backward similarity or a high forward similarity (or both). As I do not find any effect on citations, it is plausible to assume that the higher KPST measure among the additional post-1975 patents stems from their low backward similarity, suggesting that these patents were more novel.

Next, I turn to the diversity of innovation and analyse whether overall patenting activity expanded to new technologies. Watzinger and Schnitzer (2022) propose to measure the diversity of innovation by looking at the number of technology subgroups with at least some patenting activity. In the hierarchical patent classification system, subgroups are the level below six-digit technology classes, which I use in my main approach. Therefore, they represent the most disaggregate technology classification of a patent.[24] Following Watzinger and Schnitzer (2022), I refer to a subgroup as 'active' if it contained at least one patent application in a given year. To estimate the effect of the antitrust case on the diversity of innovation, I then use the same DiD approach as in my baseline analysis. Now, however, the outcome variables counts the annual number of active subgroups within each six-digit technology class.

Table 4 presents the resulting DiD estimates. Column (1) shows that compulsory licensing of Xerox's patents had an overall positive effect on the number of active subgroups. This increase in the diversity of innovation is again driven by patent applications from Japan. The estimate in column (4) indicates that, on average, a one percentage-point higher share of compulsorily licensed Xerox patents in a subclass is associated with 0.04 additional subgroups with at least one Japanese patent per year in that subclass

[24] In fact, subgroups have their own hierarchical order that is referred to as 'dot'-hierarchy. I follow Watzinger and Schnitzer (2022) and aggregate subgroups to the two-dot level, which is the second highest level. On average, there are around 4.8 subgroups at the two-dot level per six-digit technology class.

Table 4. Class-Level Analysis: Active Subgroups

	All	Applicant Country			
		USA	Non-USA	Among Non-USA	
				Japan	Others
	(1)	(2)	(3)	(4)	(5)
$\text{Share}_s \cdot \text{Post}_t$	0.037^{**}	0.006	0.044^{**}	0.043^{**}	0.006
	(0.015)	(0.007)	(0.020)	(0.021)	(0.006)
Mean of Outcome	4.79	3.58	2.62	1.14	1.96
4-Digit CPC Classes	141	141	141	141	141
Observations	35360	35360	35360	35360	35360

Notes: The table shows the results from difference-in-differences regressions following a variation of equation (1). Unlike in the main approach, the outcome variable now is the number of 'active' subgroups (aggregated to two dots) within a six-digit technology class. Column (1) reports the baseline estimates. In columns (2) to (5), only patent applications by assignees from selected countries are counted to determine whether a subgroup was 'active'. All regressions include subclass and year × class fixed effects. Standard errors clustered at the four-digit IPC technology class level are in parentheses. Significance levels: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

after 1975. This effect is statistically significant at the 5% level. It corresponds to around 37 additional active subgroups among Japanese applicants per year, which represents an increase by around 4%. The corresponding event-study estimates are presented in Figure A8 in the appendix. Reassuringly, there were no significant pre-trends in the number of active subgroups with at least one Japanese patent application.

These results demonstrate that (Japanese) innovation became more diverse after 1975, as patenting activity expanded to a larger number of distinct technologies. In appendix A.8, I further investigate whether firms also changed the direction of their innovation, following the approach by Kang (2021). I find that firms patented relatively more outside of their primary technology fields. That is, firms shifted the focus of their innovation activities towards previously peripheral technologies.

Overall, my analyses indicate that antitrust enforcement against Xerox made innovation more novel and diverse. Patenting activity expanded to new technology fields, while there was no reduction in the quality of inventions. These changes in the direction of innovation are again driven by Japanese patent applicants. Therefore, the results are in line with narrative evidence suggesting that Japanese entrants focused on smaller and lower-volume desktop copiers, which were more differentiated from existing products.

Finally, I show in appendix D that there was also a disproportionate increase in Japanese copier exports to the US after 1975, relative to exports in other industries and by other countries. This indicates that Japanese competitors benefited from the antitrust case not only in terms of increased innovation; they were able to generate higher revenues in the product market as well.

7 Effect on Xerox

While the paper so far focused on innovation by firms other than Xerox and its subsidiaries, I now study how Xerox's own patenting activities reacted to the removal of most of its IP. Estimating the effect on Xerox is challenging, because it requires to find a good counterfactual, indicating how much Xerox would have innovated in the absence of compulsory licensing.

I address this challenge by using the synthetic control method by Abadie et al. (2010, 2015) to find a control group for patenting by Xerox and its subsidiaries. I match on the yearly number of patent applications from 1960 until 1972 (i.e., until the start of the antitrust case). As donor pool, I use the sample of established firms. I exclude firms in the top decile of the distribution of the closeness measure, as these potential competitors may have been affected by the antitrust case themselves. Therefore, they do not represent a suitable counterfactual. The resulting synthetic control group consists of 66.6% Siemens, 17.4% Bell, and 16.1% Westinghouse.[25] That is, it represents a weighted average of other companies active in high-technology sectors but not in the copier industry.

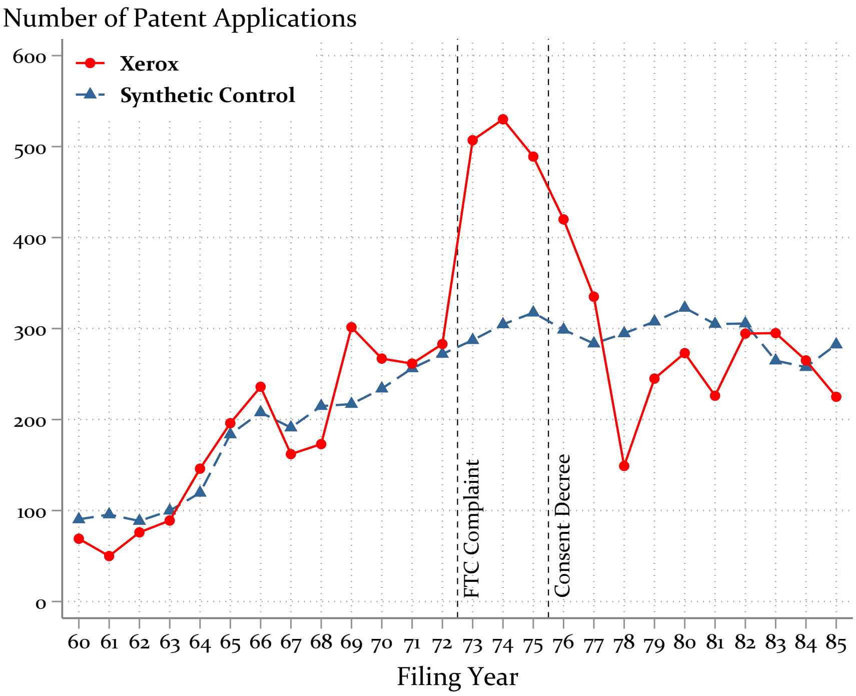
Figure 10 depicts the annual number of patent applications by Xerox and the synthetic control group. Mechanically, patenting by Xerox and its synthetic control closely tracks each other until 1972, when the FTC published its antitrust complaint. Then, in 1973, Xerox increased its patenting by around two thirds relative to the previous year, whereas there was no such trend break among the synthetic control. It is unclear what explains this sudden rise in Xerox's patenting, although the timing suggests that it may be linked to the antitrust case. Finally, following the consent decree, Xerox's patenting declined until 1978 and then slowly increased again. In contrast, patenting by the synthetic control remained relatively constant after 1975.

A simple (hand-computed) DiD estimate indicates that patenting by Xerox and its subsidiaries decreased by around 16 patents per year on average after 1975, relative to the synthetic control and relative to the pre-1972 period. Excluding Bell from the donor pool leads to a slightly larger reduction in Xerox's patenting by around 30 patents, as shown in appendix E. However, both of these numbers are much smaller in magnitude than the increase in innovation by other firms of around 160 patents per year, which I find in my main approach. Therefore, the results indicate that the overall effect of the antitrust case on innovation remains largely positive. However, the increase in patenting among competitors was partly attenuated by a decrease in Xerox's own patenting.

There is one important caveat regarding these results on the effect on Xerox. After 1975, a decline in Xerox's patenting may not necessarily represent less innovation. This

[25] Bell also faced an antitrust lawsuit in the late 1970s and was ultimately broken up in 1984 (Watzinger and Schnitzer, 2022). In appendix E, I show that excluding Bell from the donor pool leads to quantitatively similar albeit slightly more negative effects on Xerox.

Figure 10. Effect on Xerox: Patenting by Xerox vs. Synthetic Control



Notes: The figure depicts the number of patent applications per year for Xerox and its subsidiaries (in red) and a synthetic control group (in blue). The synthetic control group is computed using the algorithm by Abadie et al. (2010, 2015) and consists of 66.6% Siemens, 17.4% Bell, 16.1% Westinghouse.

is because the consent decree also imposed compulsory licensing on future Xerox patents issued until 1981. Therefore, unlike for other firms, the incentives for Xerox to file for patent protection were significantly reduced after 1975. As a consequence, the estimated decrease in Xerox's patenting likely represents an upper bound of the actual reduction in Xerox's innovation activities.

8 Conclusion

In this paper, I study how antitrust enforcement against patent-based monopolies affects subsequent innovation by domestic and foreign firms. To answer this question, I analyse the impact of compulsory patent licensing in the context of the antitrust case against Xerox in the 1970s. I find a positive effect of compulsory licensing on innovation by other firms, measured by a disproportionate increase in patenting in technologies where Xerox patents became available for licensing. Moreover, Xerox patents were cited more frequently after 1975 relative to a set of matched control patents. Therefore, antitrust action against Xerox was successful not only by promoting competition on the product market (Bresnahan, 1985a). My results indicate that it also significantly spurred innovation and promoted technological progress in copier technologies. Yet, this positive effect is mainly driven by increased innovation by Japanese competitors. They started developing smaller desktop copiers and their innovation became more diverse.

The positive innovation effect raises the question why Xerox did not license its technologies to other firms in return for royalties prior to the consent decree. In theory, efficient bargaining between the owner of an upstream technology and downstream innovators leads to ex-ante licensing such that any surplus-enhancing follow-on innovation should be developed (Green and Scotchmer, 1995). However, patents may exert a blocking effect on follow-on innovation if upstream and downstream parties fail to reach a licensing agreement. In the case of Xerox, the most plausible mechanism explaining the absence of ex-ante licensing is a rent dissipation theory (Arora and Fosfuri, 2003; Gaessler et al., 2019). That is, licensing likely would have been unprofitable for Xerox, as its loss in profits due to increased product market competition would have exceeded possible licensing revenues. Appendix F presents a brief conceptual framework that outlines this point in greater detail. In particular, I point out how several specificities of the Xerox case make the rent dissipation theory more likely than alternative explanations for the absence of ex-ante licensing such as asymmetric information (Bessen and Maskin, 2009) or coordination failure (Galasso and Schankerman, 2015).

While this paper identifies the blocking effect of Xerox's patents on follow-on innovation, it remains an open question whether compulsory licensing affected firms' incentives to innovate. The Xerox case was one out of more than 100 compulsory licensing cases since the 1940s (Scherer and Watal, 2014). Therefore, it is unlikely that there were large disincentive effects, as the case probably did not alter expectations about the probability of future compulsory licensing orders. Moreover, I show that Xerox itself only modestly reduced its patenting in response to the antitrust case.

There are two important policy implications from my paper. First, I show that compulsory licensing can be a suitable antitrust measure to increase innovation in the target industry if it removes the main barrier to entry. This result complements prior evidence from the antitrust case against Bell (Watzinger et al., 2020). Since Bell was a vertically integrated monopolist that could continue to foreclose its competitors after the antitrust case, Watzinger et al. (2020) find no effect on innovation in the target industry. In contrast, compulsory licensing of Xerox's patents removed the main barrier to entry and allowed competitors to build on its technology. Therefore, my paper provides a more complete picture of the effectiveness of compulsory licensing as an antitrust remedy.

Second, the finding that Japanese competitors particularly benefited from antitrust action against Xerox is relevant in light of recent concerns that stricter antitrust enforcement by US authorities may weaken American competitiveness in high-technology sectors. The evidence from the case against Xerox indicates that such concerns may partly be warranted, as compulsory licensing allowed foreign competitors to build on Xerox's copier technology. However, on the upside, the antitrust case brought enormous benefits to American consumers. They experienced lower prices as well as a greater variety and higher quality in copiers, irrespective of whether the innovators were Japanese. I leave the quantifications of the overall welfare effects of antitrust enforcement as a promising topic for future research.

References

Abadie, A., A. Diamond, and J. Hainmueller (2010): "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program," *Journal of the American Statistical Association*, 105, 493–505.

——— (2015): "Comparative Politics and the Synthetic Control Method," *American Journal of Political Science*, 59, 495–510.

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

Ahmadpoor, M. and B. F. Jones (2017): "The Dual Frontier: Patented Inventions and Prior Scientific Advance," *Science*, 357, 583–587.

Alcácer, J. and M. Gittelman (2006): "Patent Citations as a Measure of Knowledge Flows: The Influence of Examiner Citations," *Review of Economics and Statistics*, 88, 774–779.

Arora, A. and A. Fosfuri (2003): "Licensing the Market for Technology," *Journal of Economic Behavior and Organization*, 52, 277–295.

Bellemare, M. F. and C. J. Wichman (2020): "Elasticities and the Inverse Hyperbolic Sine Transformation," *Oxford Bulletin of Economics and Statistics*, 82, 50–61.

Bergeaud, A. and C. Verluise (2022): "A New Dataset to Study a Century of Innovation in Europe and in the US," *CEP Discussion Paper No. 1850.*

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

Bresnahan, T. F. (1985a): "Post-Entry Competition in the Plain Paper Copier Market," *American Economic Review*, 75, 15–19.

——— (1985b): "The Transition to Competition In The Plain Paper Copier Market," Federal Trade Commission Report.

Cabral, L. M. B. (2018): "Standing on the Shoulders of Dwarfs: Dominant Firms and Innovation Incentives," *CEPR Discussion Paper No. 13115.*

Carrier, M. A. (2002): "Unraveling the Patent-Antitrust Paradox," *University of Pennsylvania Law Review*, 150, 761–854.

Chesbrough, H. and R. S. Rosenbloom (2002): "The Role of the Business Model in Capturing Value from Innovation: Evidence from Xerox Corporation's Technology Spin-Off Companies," *Industrial and Corporate Change*, 11, 529–555.

Correia, S., P. Guimarães, and T. Zylkin (2020): "Fast Poisson Estimation with High-Dimensional Fixed Effects," *Stata Journal*, 20, 95–115.

Cunningham, C., F. Ederer, and S. Ma (2021): "Killer Acquisitions," *Journal of Political Economy*, 129, 649–702.

Delrahim, M. (2004): "Forcing Firms to Share the Sandbox: Compulsory Licensing of Intellectual Property Rights and Antitrust," *European Business Law Review*, 15, 1059–1069.

Federico, G., F. Scott Morton, and C. Shapiro (2020): "Antitrust and Innovation: Welcoming and Protecting Disruption," *Innovation Policy and the Economy*, 20, 125–190.

Feenstra, R. C. (1996): "U.S. Imports, 1972-1994: Data and Concordances," *NBER Working Paper No. 5515.*

Federal Trade Commission (1975): "Xerox Corporation: Consent Order, etc., in Regard to Alleged Violation of the Federal Trade Commission Act," in *Federal Trade Commission Decisions*, vol. 86, 364–386.

Gaessler, F., D. Harhoff, and S. Sorg (2019): "Bargaining Failure and Freedom to Operate: Re-Evaluating the Effect of Patents on Cumulative Innovation," *CEPR Discussion Paper No. 13969.*

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

Gilbert, R. J. (2022): *Innovation Matters: Competition Policy for the High-Technology Economy*, MIT Press.

Gomes-Casseres, B. and K. McQuade (1991): "Xerox and Fuji Xerox," *Harvard Business School Case 9-391-156.*

Goto, A. and K. Motohashi (2007): "Construction of a Japanese Patent Database and a First Look at Japanese Patenting Activities," *Research Policy*, 36, 1431–1442.

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

Iacus, S. M., G. King, and G. Porro (2012): "Causal Inference without Balance Checking: Coarsened Exact Matching," *Political Analysis*, 20, 1–24.

Jacobson, G. and J. Hillkirk (1986): *Xerox: American Samurai*, Macmillan.

Jaffe, A. B. and M. Trajtenberg (1996): "Flows of Knowledge from Universities and Federal Laboratories: Modeling the Flow of Patent Citations over Time and across Institutional and Geographic Boundaries," *Proceedings of the National Academy of Sciences*, 93, 12671–12677.

Kang, H. (2021): "How Does Competition Affect Innovation? Evidence from U.S. Antitrust Cases," *USC Marshall School of Business Research Paper*.

Kearns, D. T. and D. A. Nadler (1992): *Prophets in the Dark: How Xerox Reinvented Itself and Beat Back the Japanese*, Harper Business.

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

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

Moser, P. (2012): "Innovation without Patents: Evidence from World's Fairs," *Journal of Law and Economics*, 55, 43–74.

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

Moser, P., A. Voena, and F. Waldinger (2014): "German Jewish Émigrés and US Invention," *American Economic Review*, 104, 3222–3255.

Owen, D. (2005): *Copies in Seconds: How a Lone Inventor and an Unknown Company Created the Biggest Communication Breakthrough Since Gutenberg – Chester Carlson and the Birth of the Xerox Machine*, Simon & Schuster.

Paley, N. (2017): *The Manager's Guide to Competitive Marketing Strategies*, Routledge.

Petralia, S., P.-A. Balland, and D. Rigby (2016): "HistPat Dataset," Harvard Dataverse.

Poege, F. (2022): "Competition and Innovation: The Breakup of IG Farben," *Boston University School of Law Research Paper No. 22-24*.

Porter, M. E. (1985): *Competitive Advantage: Creating and Sustaining Superior Performance*, Free Press.

——— (1988): "Canon, Inc.: Worldwide Copier Strategy," *Harvard Business School Case 9-384-151*.

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

Scherer, F. M. (2005): "The Role of Patents in Two US Monopolization Cases," *International Journal of the Economics of Business*, 12, 297–305.

——— (2007): "Technological Innovation and Monopolization," *HKS Working Paper No. RWP07-043*.

Scherer, F. M. and J. Watal (2014): "Competition Policy and Intellectual Property: Insights from Developed Country Experience," *HKS Working Paper No. RWP14-013*.

Segal, I. and M. D. Whinston (2007): "Antitrust in Innovative Industries," *American Economic Review*, 97, 1703–1730.

Tom, W. K. (2001): "The 1975 Xerox Consent Decree: Ancient Artifacts and Current Tensions," *Antitrust Law Journal*, 68, 967–990.

United States Patent and Trademark Office (1975): "Xerox License Offer," in *Official Gazette of the United States Patent and Trademark Office*, vol. 939, 1665–1739.

Vinokurova, N. and R. Kapoor (2020): "Converting Inventions into Innovations in Large Firms: How Inventors at Xerox Navigated the Innovation Process to Commercialize Their Ideas," *Strategic Management Journal*, 41, 2372–2399.

Watzinger, M., T. A. Fackler, M. Nagler, and M. Schnitzer (2020): "How Antitrust Enforcement Can Spur Innovation: Bell Labs and the 1956 Consent Decree," *American Economic Journal: Economic Policy*, 12, 328–359.

Watzinger, M. and M. Schnitzer (2022): "The Breakup of the Bell System and its Impact on US Innovation," *CEPR Discussion Paper No. 17635*.

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

——— (2017): "How Do Patents Affect Research Investments?" *Annual Review of Economics*, 9, 441–469.

Appendix

The appendix presents further details on the dataset, robustness checks, additional empirical results, and a conceptual framework.

- Appendix A presents supplementary results for the main class-level approach to estimate the impact of the Xerox case on cumulative innovation. This includes descriptives of the sample, a broad range of robustness checks, and additional results.
- Appendix B contains supplementary results for the complementary analysis on the patent level to estimate the effect of compulsory licensing on direct follow-on innovation to Xerox's patents.
- Appendix C presents results from an auxiliary analysis on the firm level and reports additional details on the sample of established firms used in the main part of the paper.
- Appendix D descriptively analyses trade data to assess whether Japanese competitors benefited from the antitrust case in terms of subsequent exports to the US.
- Appendix E reports supplementary results on the effect on Xerox by computing an alternative synthetic control group.
- Appendix F introduces a brief conceptual framework explaining the paper's key findings.

A Supplementary Results for Class-Level Approach

This appendix contains supplementary results for my main class-level approach to estimate the effect of compulsory licensing of Xerox's patents on cumulative innovation. First, I present additional descriptives of the dataset. Then, I show robustness checks for my main results and heterogeneity analyses.

A.1 Data Description

Table A1 presents additional information on Xerox's portfolio of compulsorily licensed patents by showing the top ten four-digit CPC technology classes. Additional summary statistics for the outcome variables used in the main part of the paper are reported in Table A2.

Table A1. Xerox's Patent Portfolio Subject to Compulsory Licensing

4-Digit Classes		6-Digit Subclasses			
Top 10		Largest		Other	
Code	Title	Code	Weight	Number	Weight
G03G	ELECTROGRAPHY; ELECTROPHOTOGRAPHY; MAGNETOGRAPHY	G03G 15	40.6%	10	29.0%
G03B	APPARATUS OR ARRANGEMENTS FOR TAKING PHOTOGRAPHS OR FOR PROJECTING OR VIEWING THEM	G03B 27	2.9%	6	0.4%
B65H	HANDLING THIN OR FILAMENTARY MATERIAL, e.g. SHEETS, WEBS, CABLES	B65H 3	0.6%	12	1.6%
H04N	PICTORIAL COMMUNICATION, e.g. TELEVISION	H04N 1	1.3%	3	0.4%
H01J	ELECTRIC DISCHARGE TUBES OR DISCHARGE LAMPS	H01J 29	0.4%	10	1.0%
G02B	OPTICAL ELEMENTS, SYSTEMS OR APPARATUS	G02B 27	0.3%	9	1.0%
G06K	GRAPHICAL DATA READING	G06K 15	0.5%	4	0.4%
H01L	SEMICONDUCTOR DEVICES; ELECTRIC SOLID STATE DEVICES NOT OTHERWISE PROVIDED FOR	H01L 31	0.4%	6	0.5%
B41L	APPARATUS OR DEVICES FOR MANIFOLDING, DUPLICATING OR PRINTING FOR OFFICE OR OTHER COMMERCIAL PURPOSES	B41L 39	0.4%	4	0.4%
B41M	PRINTING, DUPLICATING, MARKING, OR COPYING PROCESSES; COLOUR PRINTING	B41M 5	0.5%	3	0.3%

Notes: The table presents additional information on the top ten four-digit CPC technology classes in Xerox's portfolio of compulsorily licensed patents. The first and second column indicate the code and title of the ten largest four-digit classes, measured by the number of patents. The third column provides the code of the largest six-digit subclass and the fourth column shows the weight of this subclass in Xerox's patent portfolio. The last two columns indicate the number and weight of the remaining six-digit subclasses within the given four-digit class.

Table A2. Class-Level Analysis: Summary Statistics

	Mean	SD
(A) Main Outcome		
# Patents	15.132	28.153
(B) Citations to Xerox		
Distance = 1	0.213	2.909
Distance $\leq$ 2	0.582	5.071
Distance $\leq$ 3	1.124	6.834
Unconnected to Xerox	14.008	25.348
(C) Applicant Country		
USA	8.934	16.889
Non-USA	5.742	12.252
Japan	2.251	7.708
Others	3.491	6.576
(D) Firm Sample		
Established Firms	8.572	19.249
Other Patentees	6.670	12.064
(E) Closeness to Xerox		
Top Decile of Closeness	1.607	7.568
75-90th Percentile	2.167	6.476
Below 75th Percentile	4.736	11.394
(F) Quality		
Top 10% of Forward Citations	1.603	4.649
Top 10% of Quality (KPST)	2.000	8.921
Mean of Forward Citations	14.144	12.392
Mean of Quality (KPST)	-0.005	0.114
(G) Active Subgroups		
All	4.790	6.196
USA	3.576	4.920
Non-USA	2.618	3.804
Japan	1.137	2.443
Others	1.964	2.811

Notes: The table presents summary statistics for the outcome variables used in the main part of the paper. Panel (A) refers to the baseline specification, where the outcome is the number of patent applications by firms other than Xerox per six-digit CPC subclass per year. The variables depicted in panels (B) to (G) correspond to the results from Figure 4, Table 2, Figure 7, Table 3, and Table 4 respectively. There are 35,360 observations for each variable (i.e., 2,210 six-digit subclasses observed in 16 years).

Table A3 reports additional summary statistics for the treatment variable (Share_s) and the measure of closeness to Xerox (Closeness_i) defined in equation (4). Summary statistics for Share_s are shown for the 313 treated six-digit subclasses in the main sample, whereas those for Closeness_i are shown for the 1,635 firms in the sample of established firms. In the average treated subclass, 2.7% of the unexpired patents (as of 1975) were subject to compulsory licensing. In contrast, in the subclass most exposed to the antitrust case, 58.5% of the unexpired patents became available for licensing. As Share_s is defined

Table A3. Class-Level Analysis: Summary Statistics for Share_s and Closeness_i

	Observations	Mean	SD	Min	Max
Share_s	313	2.746	7.440	0.027	58.519
Closeness_i	1635	0.339	1.904	0.000	33.483

Notes: The table presents summary statistics for the main treatment variable (Share_s) and the measure of closeness to Xerox (Closeness_i) defined in equation (4). As Share_s is defined in percentage terms (i.e., $\times$ 100), both variables can, in theory, take on values between 0 and 100.

in percentage terms, both variables can, in theory, take on values between 0 and 100. For the closeness measure, a value of 100 would indicate that all of a firm's unexpired patents were in a subclass in which all patents were compulsorily licensed. As no such technology class exists, the closeness measure has a maximum of 33.5. The mean value is only around 0.3, indicating that most firms in the sample of established firms had very little exposure to copier technologies.

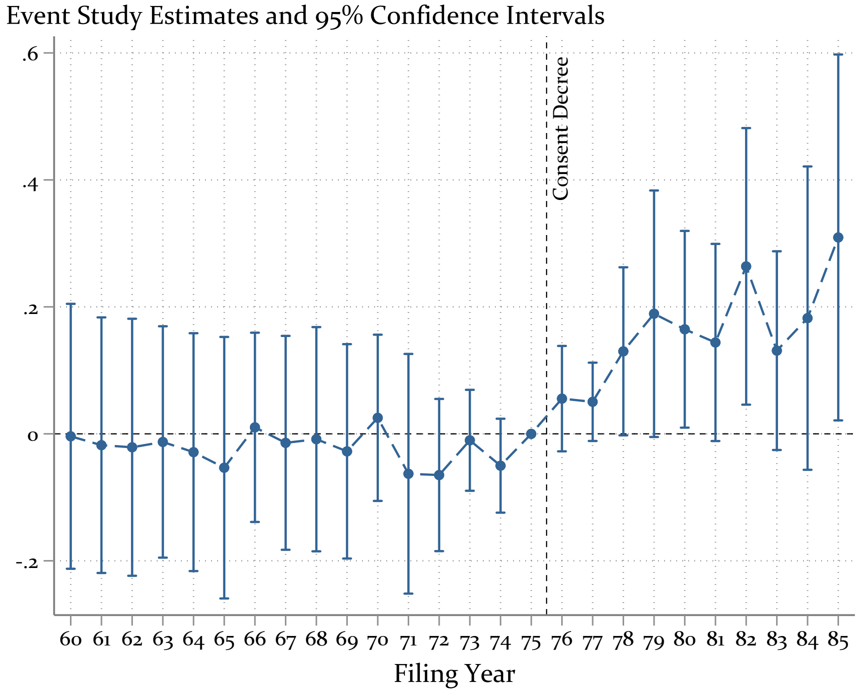
A.2 Extended Pre-Treatment Period

Figure A1 reproduces the event-study analysis from my main class-level approach (see Figure 3 in the main part of the paper), but it presents annual estimates and covers an extended pre-treatment period back to 1960. The estimates show that, on average, there were no significant differences in patenting across differentially exposed six-digit subclass in the 15 years prior to the consent decree. This further supports the identifying assumption underlying my class-level approach. After compulsory licensing of Xerox's patents in 1975, however, there was a disproportionate increase in patenting in technologies where most of Xerox's patents became available for licensing.

A.3 Model Specification

In Table A4, I analyse the robustness of my baseline estimate by employing alternative model specifications. Column (1) repeats the baseline estimate from the main part of the paper. In column (2), I estimate the model without using the coarsened exact matching (CEM) weights by Iacus et al. (2012), which does not change the point estimate by much. In column (3), I estimate a Poisson pseudo-likelihood regression instead of an ordinary least squares model, using the algorithm by Correia et al. (2020). This alternative non-linear estimation model takes into account that patent applications represent count data, but it comes at the sacrifice of the point estimate's simple (linear) interpretation. Reassuringly, the point estimate is still positive and highly statistically significant. Columns (4) to (6) employ alternative treatment definitions. In column (4), I apply the inverse hyperbolic sine (IHS) transformation to the share of compulsorily licensed patents per subclass. Again, the estimate remains positive and statistically significant. This addresses potential

Figure A1. Class-Level Analysis: Event-Study Estimates with Extended Pre-Period



Notes: The figure depicts point estimates and 95% confidence intervals from the event-study analysis in equation (2), using an extended pre-treatment period that covers years since 1960. All regressions use the weights by Iacus et al. (2012). Standard errors are clustered at the four-digit CPC technology class level.

concerns that the underlying relationship between the treatment variable and the number of patent applications may be non-linear. The estimates in columns (5) and (6) further show that my results are robust to using a binary treatment measure or defining treatment based on the number (as opposed to the share) of compulsorily licensed patents. These alternative treatment measures are further discussed below in appendix A.4. Finally, in column (7) of Table A4, I exclude the top ten subclasses with the greatest share of compulsorily licensed patents (i.e., the treatment variable) from the sample. This is to verify that the positive baseline estimate is not purely driven by patenting in these subclasses that were highly exposed to compulsory licensing. Again, the point estimate remains statistically significant at the 5% level. Additional robustness checks on the sample definition are reported below in appendix A.5.

A.4 Treatment Definition

Next, I turn to assessing the robustness of the treatment definition. In the first step, I employ a simple binary specification and show the corresponding event-study estimates in Panel (A) of Figure A2. The estimates reveal no statistically significant pre-trends and indicate that there was a disproportionate increase in patenting after 1975 in technologies where at least one Xerox patent became available for licensing. The DiD estimate in

Table A4. Class-Level Analysis: Alternative Model Specifications

	Model			Treatment			Outliers
	Baseline	No CEM Weights	Poisson	IHS of Share	Binary	IHS of Abs. No.	Excl. Top 10 Subcl.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\text{Share}_s \cdot \text{Post}_t$	0.189**	0.169*	0.033***				0.666**
	(0.094)	(0.087)	(0.011)				(0.318)
$\text{IHS}(\text{Share}_s) \cdot \text{Post}_t$				1.760**			
				(0.806)			
$\mathbb{1}[\text{Lic}_s \geq 1] \cdot \text{Post}_t$					2.797*		
					(1.527)		
$\text{IHS}(\text{Lic}_s) \cdot \text{Post}_t$						2.974**	
						(1.430)	
Subclass FE	✓	✓	✓	✓	✓	✓	✓
Year × Class FE	✓	✓	✓	✓	✓	✓	✓
CEM Weights	✓			✓	✓	✓	✓
Mean of Outcome	15.13	15.13	15.13	15.13	15.13	15.13	15.06
4-Digit CPC Classes	141	141	141	141	141	141	140
Observations	35360	35360	34585	35360	35360	35360	35088

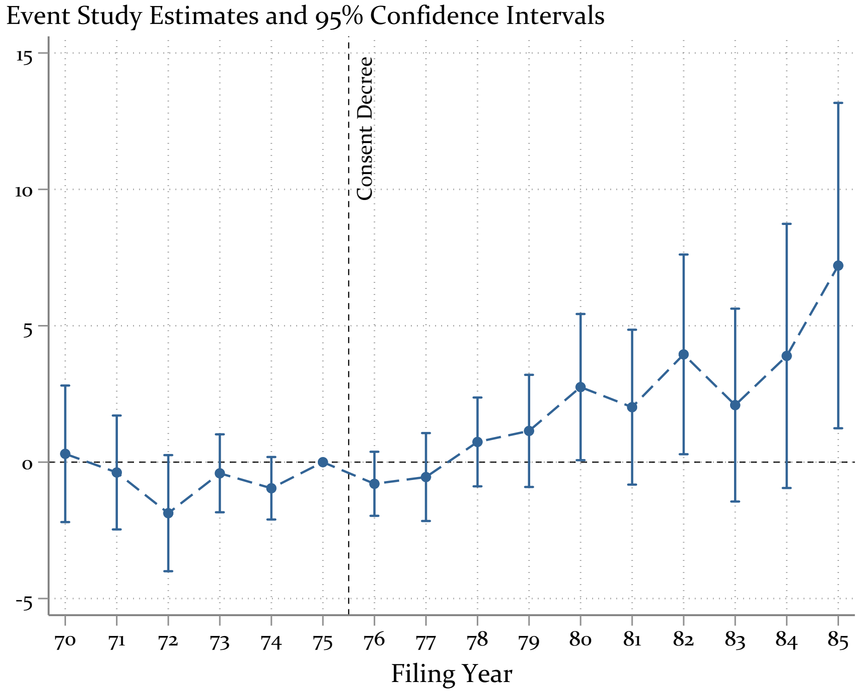
Notes: The table shows the results from difference-in-differences regressions following equation (1). The outcome variable in all regressions is the number of patent applications by firms other than Xerox in a given six-digit CPC subclass and year. Column (1) replicates the baseline estimate from Table 1 in the main part of the paper. In column (2), the model is estimating without using the weights by Iacus et al. (2012). In column (3), I estimate a Poisson pseudo-likelihood regression instead of a linear ordinary least squares model. In column (4), I apply the inverse hyperbolic sine (IHS) transformation to the share of compulsorily licensed patents per class (i.e., the treatment variable). Column (5) employs a binary treatment indicator that equals one for subclasses with at least one licensable Xerox patent. In column (6), I use a an alternative continuous treatment variable defined as the inverse hyperbolic sine of the number of compulsorily licensed Xerox patents, denoted by the variable Lic_s. Finally, column (7) employs the baseline model specification but excludes the top ten six-digit subclasses with the highest share of compulsorily licensed patents from the sample. Standard errors clustered at the four-digit CPC technology class level are in parentheses. Significance levels: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

column (5) of Table A4 confirms this visual impression. The results are equally robust to defining an alternative binary treatment that considers all subclasses as treated that contain at least two (as opposed to one) compulsorily licensed Xerox patents. This is not surprising in light of the pattern shown in panel (B) of Figure 2 in the main part of the paper, indicating that subclasses with exactly one compulsorily licensed Xerox patent experienced a similar patenting trend as subclasses with no exposure to compulsorily licensing. However, one drawback of the alternative binary specification is that the number of treated subclasses is reduced to 122 and, as a consequence, there are only 66 clusters on the four-digit class level.

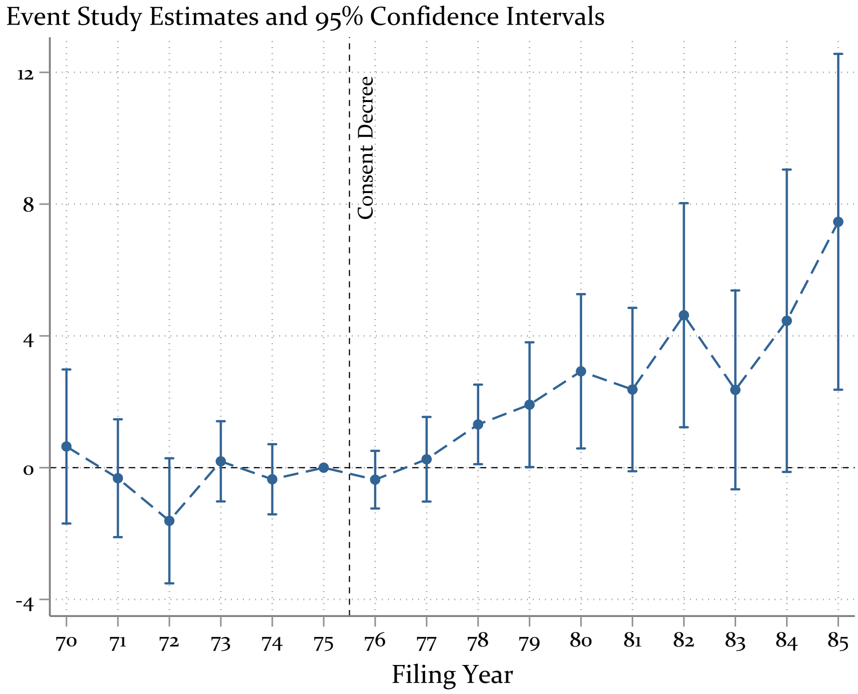
To avoid making an arbitrary choice which subclasses should be considered as treated, my main empirical approach employs a continuous treatment specification. As explained in the main part of the paper, I use the variable Share_s as treatment measure, which captures the share of unexpired patents per subclass (as of 1975) that were subject to compulsorily licensing. In Panel (B) of Figure A2, I further show that my results do not hinge on the specific definition of the continuous treatment variable. I now define the treatment variable as the inverse hyperbolic sine of the absolute number of compulsorily

Figure A2. Class-Level Analysis: Event-Study Estimates with Alternative Treatments

(A) Binary Treatment

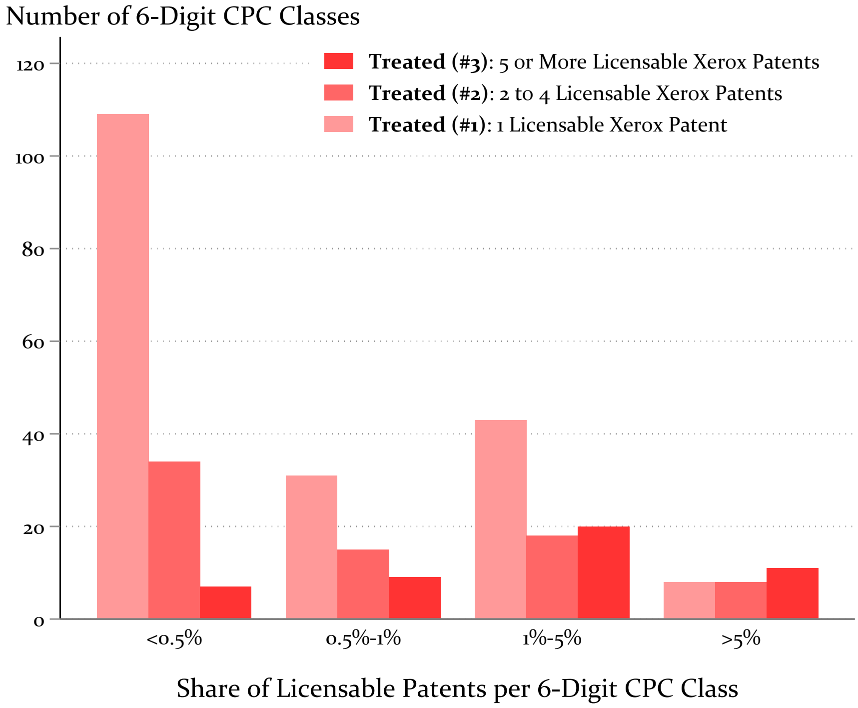


(B) Number of Licensable Patents



Notes: The figure depicts point estimates and 95% confidence intervals from two variations of the event-study analysis in equation (2). Panel (A) employs a binary treatment specification, where a six-digit subclass is treated if it contains at least one compulsorily licensed Xerox patents. Panel (B) uses an alternative continuous treatment specification, where the treatment variable is defined as the inverse hyperbolic sine of the number of compulsorily licensed Xerox patents per subclass. Both regression models use the weights by Iacus et al. (2012). Standard errors are clustered at the four-digit CPC technology class level.

Figure A3. Class-Level Analysis: Share vs. Number of Licensable Patents



Notes: The figure depicts the relationship between the share of unexpired patents per subclass that were subject to compulsorily licensing (i.e., the treatment variable) and the absolute number of licensable patents per subclass. It shows the number of subclasses in the sample (on the vertical axis) that have a given share variable (on the horizontal axis), which is discretized into four groups. The figure presents separate bars for subclasses with (#1) one, (#2) two to four, and (#3) five of more compulsorily licensed Xerox patents. Example: there are approximately 110 subclasses with one compulsorily licensed Xerox patent whose share variable is below 0.5%.

licensed Xerox patents per subclass (Bellemare and Wichman, 2020). This non-linear transformation is necessary to ensure that the DiD estimate is not driven by the single subclass G03G 15, which represents an extreme outlier in terms of the number (but not in terms of the share) of licensable Xerox patents. As can be seen in Panel (B) of Figure A2, the pattern of increased patenting in those subclasses that were more strongly exposed to compulsory licensing of Xerox's patents is robust to this alternative treatment definition. Similarly, the point estimate in column (6) of Table A4 remains positive and statistically significant.

Figure A3 shows the relationship between the main treatment variable and the absolute number of compulsorily licensed Xerox patents per subclass. The figure uses the same subgroups of subclasses as in Table 1 or panel (B) of Figure 2 in the main part of the paper. According to Figure A3, on average, subclasses with a larger absolute number of licensable Xerox patents also have a higher share of compulsorily licensed patents. However, there is important variation in the share variable even within a subgroup, which stems from differences in the total number of unexpired patents across subclasses (i.e., the denominator of the share variable). Arguably, these differences contain mean-

Table A5. Class-Level Analysis: IPC Classes

	(1)	(2)	(3)
$\text{Share}_s \cdot \text{Post}_t$	0.156**	0.204***	
	(0.063)	(0.059)	
$\text{Share}_s \cdot \text{Post}_t \cdot \mathbb{1}[\text{Lic}_s = 1]$			0.009
			(0.029)
$\text{Share}_s \cdot \text{Post}_t \cdot \mathbb{1}[2 \leq \text{Lic}_s \leq 4]$			0.052
			(0.032)
$\text{Share}_s \cdot \text{Post}_t \cdot \mathbb{1}[5 \leq \text{Lic}_s]$			0.288***
			(0.063)
Subclass FE	✓	✓	✓
Year FE	✓		
Year × Class FE		✓	✓
Mean of Outcome	16.28	16.28	16.28
4-Digit ipc Classes	137	137	137
Observations	34432	34432	34432

Notes: The table shows the results from difference-in-differences regressions following equation (1). The outcome variable in all regressions is the number of patent applications by firms other than Xerox in a given six-digit IPC (instead of CPC) subclass and year. In column (3), the treatment variable is interacted with indicators for subclasses with one, two to four, and five or more compulsorily licensed Xerox patents, as indicated by the variable Lic_s. Standard errors clustered at the four-digit CPC technology class level are in parentheses. Significance levels: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

ingful information about the exposure of a subclass to compulsory licensing. Therefore, I employ the variable Share_s as my main treatment measure.

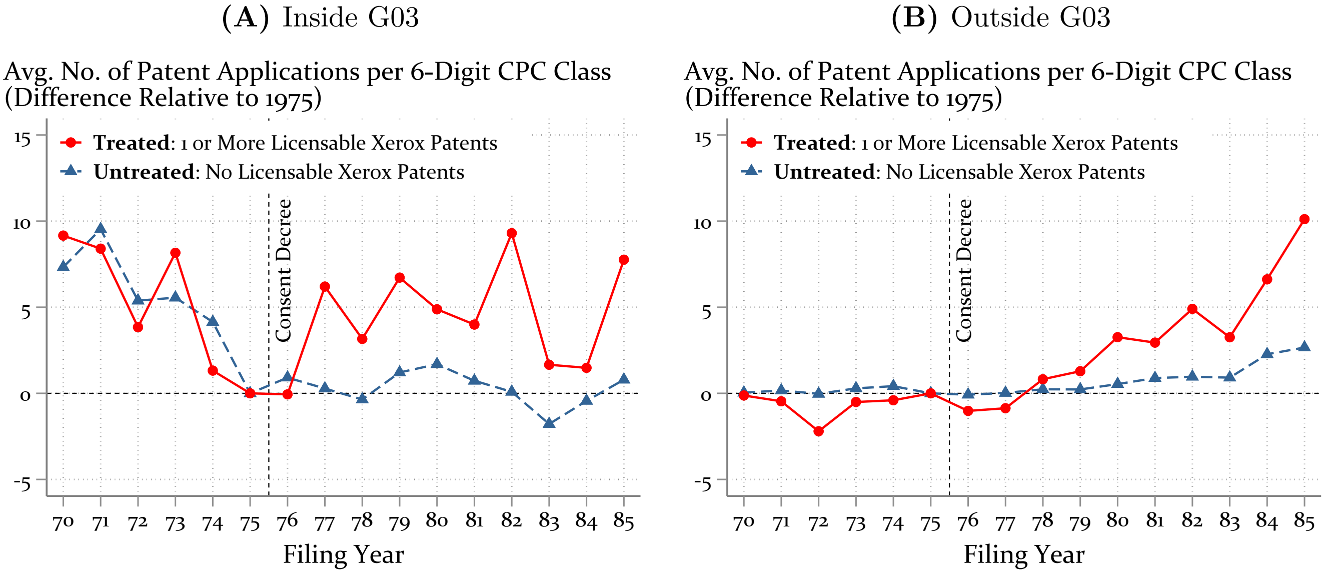
Finally, Table A5 presents estimation results when using technology classes based on the International Patent Classification (IPC). The table shows that using IPC (instead of CPC) classes yields very similar estimates both in terms of magnitude and statistical significance.

A.5 Sample Definition

I now analyse whether my estimates are robust to different sample definitions. As noted above in appendix A.3, the estimate in column (7) of Table A4 indicates that my estimates remain positive and statistically significant even after excluding the top ten subclasses with the greatest share of compulsorily licensed patents. That is, the positive effect is not driven by outliers with high values of the treatment variable.

In the next step, I more closely examine the role of the three-digit class G03, which covers photography and electrography among other technologies. As shown in Table A1, a large fraction of Xerox's compulsorily licensed patents is clustered in this technology class. Figure A4 presents descriptive evidence indicating that the post-1975 increase in patenting in subclasses exposed to compulsory licensing is present both inside and outside the three-digit class G03. Panel (A) of Figure A4 depicts the average number of patent applications

Figure A4. Class-Level Analysis: Patenting Inside vs. Outside Class G03



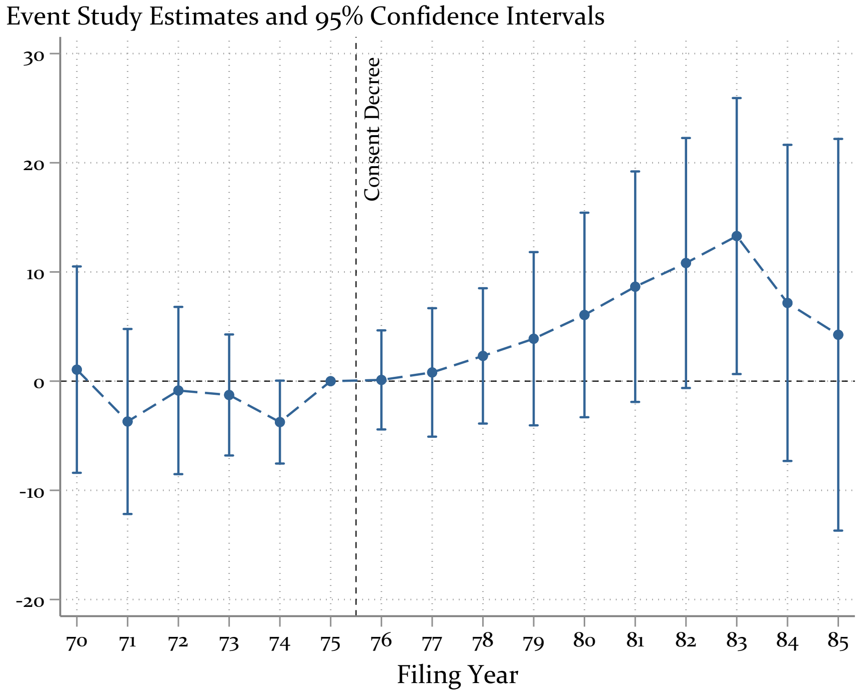
Notes: The figure depicts the average number of patent applications per six-digit subclass relative to 1975. Averages are computed separately for treated and untreated subclasses, where a subclass is defined as treated if it contained at least one compulsorily licensed Xerox patent. Panel (A) only includes subclasses that are part of the three-digit CPC class G03. Panel (B) includes all remaining subclasses in the sample. In both panels, the subclasses are aggregated using the weights by Iacus et al. (2012).

separately for treated and untreated subclasses inside class G03, while Panel (B) reports the averages for all remaining technologies. As is evident from the figure, even inside class G03, there was a relative increase in patenting in subclasses exposed to compulsory licensing of Xerox's patents. However, the averages in Panel (A) are relatively noisy due to the small number of observations. As shown in Panel (B), patenting also increased after 1975 in treated subclasses belonging to three-digit classes other than G03. However, it is noteworthy that the increase inside G03 is quite pronounced as soon as 1977, whereas the effect in the remaining technology fields evolves more progressively. This may point to the presence of spillovers across technologies over time.

In a second robustness check, I follow Watzinger et al. (2020) and aggregate subclasses into one treated and one untreated subclass per four-digit class. The advantage of this aggregation is that the resulting regression model gives equal weight to every four-digit class. In contrast, in my baseline approach, by using the weights of Iacus et al. (2012), the regression gives equal weight to every treated six-digit subclass, implying that classes with a larger number of treated subclasses get a higher weight.

Figure A5 depicts the event-study estimates with aggregated subclasses. The estimates are based on a binary treatment specification, where a subclass is treated if it contained at least one compulsorily licensed Xerox patent. Consistent with my baseline results, there was a disproportionate increase in the number of patent application after 1975 in (aggregated) subclasses where Xerox patents became available for licensing. However, this positive effect is statistically significant at the 5% level only in 1983 and then fades

Figure A5. Class-Level Analysis: Event-Study Estimates with Aggregated Subclasses



Notes: The figure depicts point estimates and 95% confidence intervals from an event-study analysis akin to equation (2), but where subclasses are aggregated into one treated and one untreated subclass per class. A six-digit subclass is defined as treated if it contained at least one compulsorily licensed Xerox patent. Standard errors are clustered at the four-digit CPC technology class level.

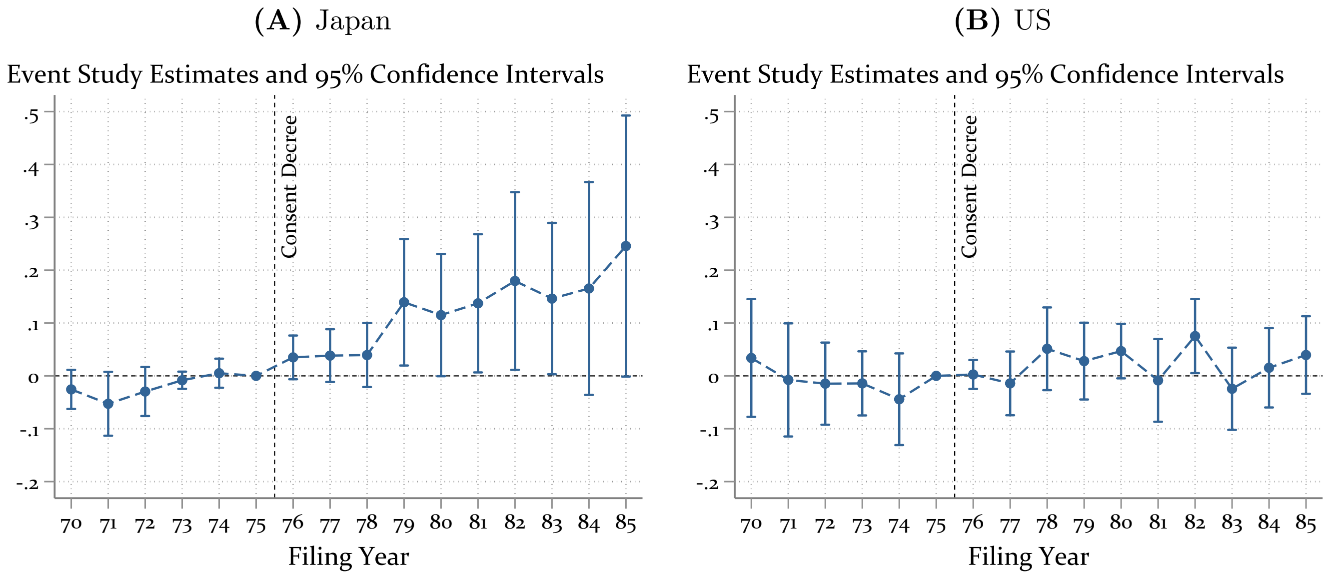
out again. One reason for the lower precision of these estimates could be the binary treatment specification, which does not exploit all the available information in terms of the subclasses' exposure to compulsory licensing. Yet, reassuringly, the figure shows no meaningful pre-trends, which gives further credibility to my identification strategy.

A.6 Heterogeneity by Applicant Country

This section presents additional results of my heterogeneity analysis by applicant country. Figure A6 depicts event-study estimates separately for US and Japanese patent applicants. These estimates correspond to the simple DiD estimates in columns (2) and (4) of Table 2. The figure shows that there were no significant pre-trends in either country, which again supports the identification strategy underlying my empirical approach. After 1975, technology classes with a greater exposure to compulsory licensing of Xerox's patents experienced an increase in the number of patent applications only by Japanese applicants. The point estimates for Japan are larger than those for the US throughout the post-treatment period. However, standard errors are also larger, because the overall number of Japanese patent applications was lower.

In Table A6, I show results from a robustness check that uses data on patent applications at the Japanese Patent Office (JPO), using data from the Japanese Institute of

Figure A6. Class-Level Analysis: Event-Study Estimates by Applicant Country



Notes: The figure depicts point estimates and 95% confidence intervals from the event-study analysis in equation (2). In panels (A) and (B), the outcome variable is restricted to patent applications by assignees from Japan and the US, respectively. All regressions use the weights by Iacus et al. (2012). Standard errors are clustered at the four-digit CPC technology class level.

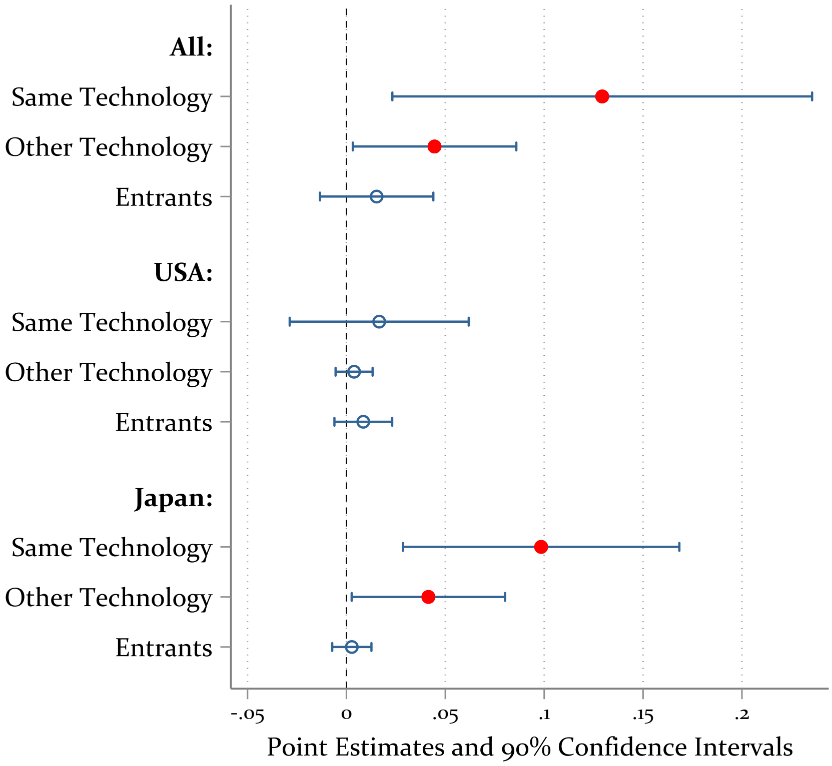
Table A6. Class-Level Analysis: Data from Japanese Patent Office

	Baseline	Applicant Country			
		USA	Non-USA	Among Non-USA	
				Japan	Others
	(1)	(2)	(3)	(4)	(5)
$\text{Share}_s \cdot \text{Post}_t$	0.195*	-0.004	0.235*	0.234*	0.002
	(0.105)	(0.008)	(0.136)	(0.136)	(0.002)
Mean of Outcome	24.46	1.60	22.88	21.11	1.77
4-Digit IPC Classes	117	117	117	117	117
Observations	22272	22272	22272	22272	22272

Notes: The table shows the results from difference-in-differences regressions following equation (1). The underlying data in this table are patent applications at the Japanese Patent Office (Goto and Motohashi, 2007). The treatment definition is based on IPC (instead of CPC) classes. Column (1) reports the baseline estimates. In columns (2) to (5), the outcome variable is restricted to patent applications filed by assignees from selected countries. All regressions include subclass and year × class fixed effects. Standard errors clustered at the four-digit IPC technology class level are in parentheses. Significance levels: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Intellectual Property (Goto and Motohashi, 2007). In contrast, my main results are based on patenting at the USPTO. This robustness check addresses the concern that increased patenting by Japanese firms in the US may not necessarily represent novel innovation; it may also reflect Japanese copier producers seeking protection for existing technologies abroad. The set-up of Table A6 is analogous to that of Table 2 in the main part of the paper. The only difference is that the analysis with JPO data uses a treatment definition based on IPC (instead of CPC) classes, as the Japanese patent data only report IPC

Figure A7. Class-Level Analysis: Alternative Heterogeneity by Patenting Experience



Notes: The figure depicts point estimates and 90% confidence intervals from estimating the regression model in equation (1). The outcome variable (i.e., the number of patent application) is split by applicant country as well as by whether the assignee had already filed a patent by 1970 and to which technology class this patent belonged. A previous patent is considered to be of the 'same' ('other') technology if it belongs to the same (another) three-digit CPC class as the focal subclass. Patents by assignees without any patenting experience until 1970 are labelled as 'entrant' patents. Red dots (blue circles) indicate statistically significant (insignificant) point estimates. Standard errors are clustered at the four-digit CPC technology class level.

classes. However, as shown above in appendix A.4, my main results are also robust to using IPC classes.

Overall, the DiD estimates in Table A6 support the interpretation that the increase in Japanese patenting after 1975 represents novel innovation. The baseline estimate in column (1) is positive and statistically significant at the 10% level. Again, this positive effect is almost entirely driven by increased patenting by Japanese firms, as indicated by column (4). In contrast, the estimates for patenting by applicants from the US or other countries are small in magnitude and statistically indistinguishable from zero.

A.7 Heterogeneity by Prior Patenting Experience

Figure 7 in the main part of the paper shows that the observed increase in patenting after 1975 is primarily driven by firms with previous patenting experience in technologies related to Xerox. Figure A7 shows estimates from a complementary heterogeneity analysis that considers whether firms had prior patenting experience and in which technologies they patented. The first row shows the effect on patenting by firms with previous experience in the same technology. More specifically, in the first row, the outcome variable only counts those patents whose assignee had filed a patent until 1970 in the same-three digit technology class as the focal subclass. In contrast, the second and third row depict estimates of the effect on patenting by firms with previous experience in other fields and by firms without any patenting experience (i.e., entrants), respectively.

The results again indicate that the observed increase in the number of patent applications after 1975 is due to firms with previous patenting experience in technologies related to Xerox. The remainder of Figure 7 repeats the analysis separately for applicants from the US and Japan. The estimates highlight that, even within Japan, the positive innovation effect is driven by patent applications by firms that previously patented in treated technologies. Overall, these results are consistent with those from Figure 7 in the main part of the paper, where patents are split by firms.

A.8 Diversity and Direction of Innovation

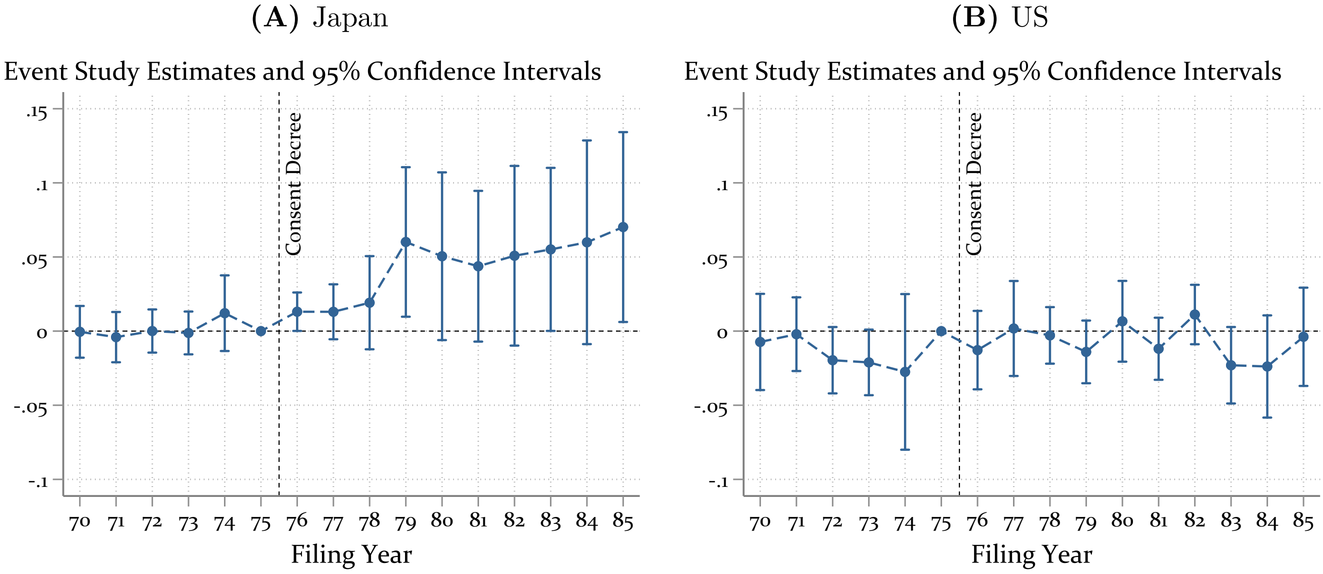
This section presents additional results on the diversity and direction of innovation. In the main part of the paper, I follow Watzinger and Schnitzer (2022) and look at the number of active technology subgroups to study the effect of compulsory licensing on the diversity of innovation. Figure A8 presents the corresponding event-study estimates separately for patents by applicants from Japan and the US in panels (A) and (B), respectively. The figure shows that the number of active subgroups with at least one Japanese patent application increased after 1975, but there were no significant pre-trends.

In the next step, I investigate whether firms also changed the direction of their innovation. To this end, similar to Kang (2021), I identify the primary and peripheral technology fields of every firm in my sample of established firms. I define a firm's primary field as its top four-digit CPC technology classes, where the firm filed at least 50% of its patents in the pre-treatment period from 1970 until 1975. All remaining four-digit technology classes represent a firm's peripheral technology fields.[26] This approach has the advantage that the number of patent applications is roughly equal across primary and peripheral technology fields in the pre-treatment period. I then use my main class-level approach to estimate how compulsory licensing of Xerox's patents affected patenting across firms' technology fields.

Figure A9 shows that the increase in innovation after 1975 is driven disproportionately by patenting in firms' peripheral technology fields. Panel (A) shows the estimates for all established firms, whereas panel (B) only considers patents filed by firms in the top decile of the distribution of the closeness measure from equation (4). Both panels confirm the pattern from previous figures and tables that the overall positive effect of the antitrust

[26]In contrast, Kang (2021) defines a firm's primary technology field as the firm's top three technology classes. The results reported below remain robust to this alternative approach.

Figure A8. Class-Level Analysis: Event-Study Estimates of Effect on Active Subgroups



Notes: The figure depicts point estimates and 95% confidence intervals from a variation of the event-study analysis in equation (2). Unlike in the main approach, the outcome variable now is the number of 'active' subgroups (aggregated to two dots) within a six-digit technology class. In panels (A) and (B), only patent applications by assignees from Japan and the US, respectively, are counted to determine whether a subgroup was 'active'. All regressions use the weights by Iacus et al. (2012). Standard errors are clustered at the four-digit CPC technology class level.

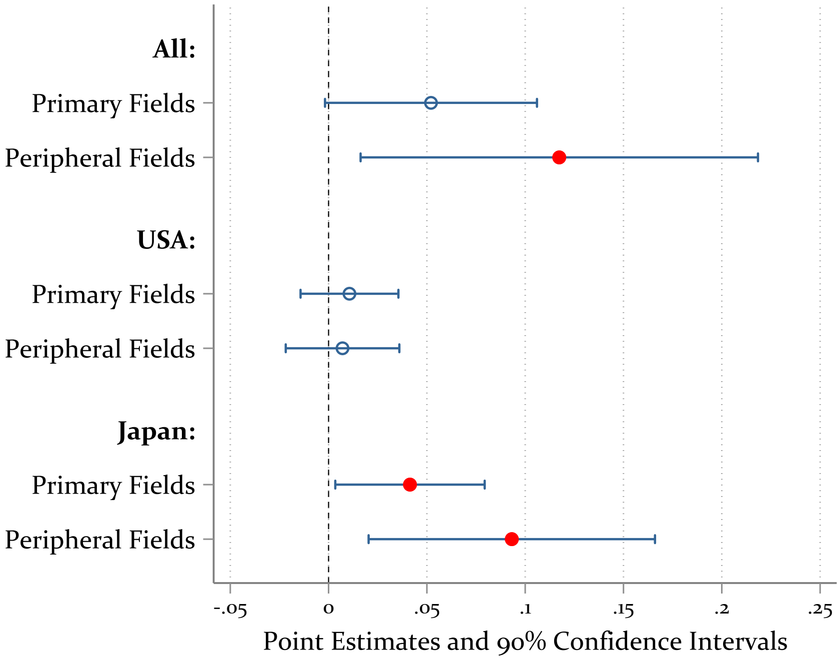
case is primarily driven by increased innovation by Japanese applicants. When focusing on Japanese firms in the top decile of the distribution of the closeness measure, the DiD estimates are positive and statistically significant both for patenting in primary and peripheral fields. Yet, the point estimate for peripheral fields is about double the size of that for primary fields. This suggests that Japanese competitors shifted the focus of their innovation activities towards different technologies. That is, their research and development became more explorative (and less exploitative). In contrast, if firms had continued innovating in the same technology fields as before, one would expect the DiD estimates to be of roughly equal size.

B Supplementary Results for Patent-Level Approach

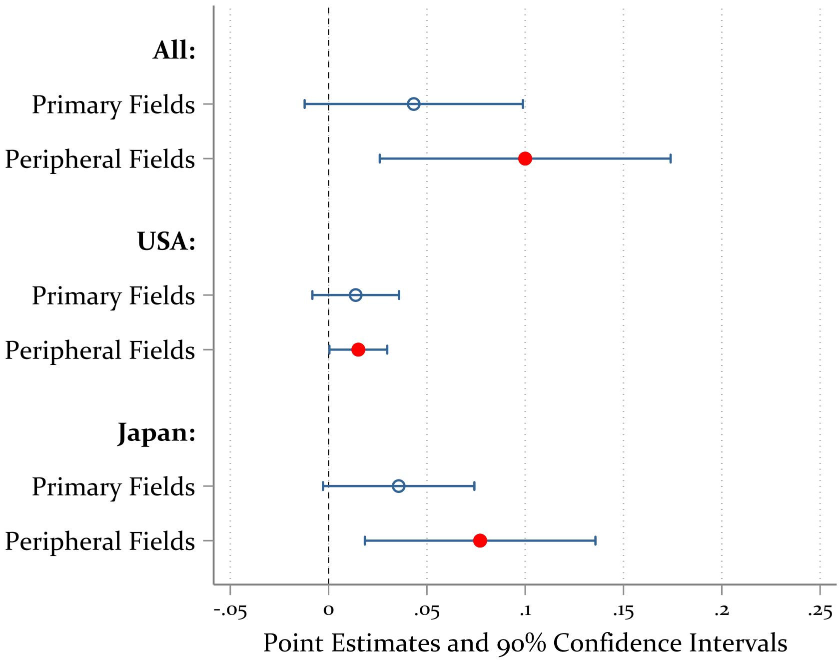
This appendix contains supplementary results for my complementary analysis on the patent level to estimate the effect of compulsory licensing on direct follow-on innovation to Xerox's patents. Column (1) of Table B1 presents the baseline DiD estimate from the patent-level analysis. It indicates that, on average, every compulsorily licensed Xerox patent received an additional 0.02 citations per year after 1975 relative to the matched control patents. In the remaining columns of Table B1, I split the number of forward citations by the applicant country of the citing patent. Unlike in the results of the class-level analysis in Table 2, the estimate for the US is now statistically significant.

Figure A9. Class-Level Analysis: Patenting in Firms' Primary vs. Peripheral Fields

(A) All Established Firms



(B) Firms in Top Decile of Closeness



Notes: The figure depicts point estimates and 90% confidence intervals from estimating the regression model in equation (1). In both panels, the outcome variable (i.e., the number of patent applications) is split by applicant country as well as by whether the focal patent belongs to its assignee's primary or peripheral technology field. A firm's primary (peripheral) field is defined as those four-digit CPC technology classes where that firm filed at least 50% (the remainder) of its patent applications from 1970 until 1975. Panel (A) considers patent applications by all 'established firms', whereas panel (B) restricts attention to patents filed by firms located in the top decile of the distribution of the closeness measure defined in equation (4). Red dots (blue circles) indicate statistically significant (insignificant) point estimates. Standard errors are clustered at the four-digit CPC technology class level.

Table B1. Patent-Level Analysis: Heterogeneity by Citing Country

	Baseline	Citing Country			
		USA	Non-USA	Among Non-USA	
				Japan	Others
	(1)	(2)	(3)	(4)	(5)
$\text{Xerox}_i \cdot \text{Post}_t$	0.024***	0.017**	0.009**	0.013***	-0.003
	(0.008)	(0.008)	(0.004)	(0.003)	(0.003)
Mean of Outcome	0.20	0.12	0.08	0.03	0.04
4-Digit CPC Classes	108	108	108	108	108
Observations	409050	409050	409050	409050	409050

Notes: The table shows the results from difference-in-differences regressions following equation (3). Column (1) repeats the baseline estimates. In columns (2) to (5), the outcome variable is restricted to forward citations from citing patents filed by assignees from selected countries. All regressions include patent and year fixed effects. Standard errors clustered at the four-digit CPC technology class level are in parentheses. Significance levels: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

It accounts for around half of the baseline estimate, implying that part of the increase in direct citations to compulsorily licensed Xerox patents came from patents filed by American firms. However, citations by non-US and especially Japanese applicants again play a key role in explaining the increase in citations to Xerox's patents. The estimate in column (4) is not only highly statistically significant; it is also quantitatively large compared to the average number of forward citations from Japanese patents. Overall, therefore, the estimates from Table B1 confirm the finding from my main approach that Japanese competitors particularly benefited from access to Xerox's technology.

C Supplementary Results for Firm-Level Approach

As discussed in section 5, I also construct a sample of established firms that filed at least ten patent applications from 1970 until 1975. This appendix contains additional information on the firm sample and presents results from an auxiliary analysis on the firm level.

C.1 Japanese Firms in Top Decile of Closeness Measure

Figure 7 in the main part of the paper indicates that the positive innovation effect is primarily driven by Japanese firms in the top decile of the distribution of the closeness measure. Table C1 presents the names of this set of Japanese firms. Reassuringly, the main Japanese competitors that, ex post, entered the US copier market (i.e., Canon, Konica, Minolta, Ricoh, Sharp, and Toshiba) are all included in this list. The list also contains a number of relatively unknown companies whose name suggests that their business may be related to printing or paper manufacturing. It is also important to note that other

Table C1. Firm-Level Analysis: List of Japanese Firms in Top Decile of Closeness

Name
CANON
CASIO COMPUTER COMPANY
DENKI ONKYO COMPANY
DIC CORPORATION
DAINIPPON PRINTING COMPANY
FUJIFILM
FUJITSU
HITACHI METALS
ISE ELECTRONICS CORPORATION
IWASAKI TSUSHINKI
IWATSU ELECTRIC COMPANY
JAPAN SYNTHETIC RUBBER
KANSAI PAINT COMPANY
KANZAKI PAPER MANUFACTURING COMPANY
KATSURAGAWA DENKI
KONICA CORPORATION
MINOLTA CAMERA COMPANY
MITA INDUSTRIAL COMPANY
MITSUBISHI PAPER MILLS
NIPPON TELEGRAPH AND TELEPHONE CORPORATION
PANASONIC CORPORATION
PILOT CORPORATION
RICOH COMPANY
SHARP CORPORATION
TOSHIBA CORPORATION
WEST ELECTRIC COMPANY

Notes: The table reports the names of Japanese companies in the sample of established firms that are in the top decile in terms of the distribution of the closeness measure defined in equation (4).

well-known Japanese companies in the high-technology sector (e.g., Hitachi, NEC, Sony), which were extremely successful at the time, are not included in the list.

C.2 Auxiliary Firm-Level Analysis

I now present an auxiliary analysis on the firm level that allows me to rule out alternative explanations for the heterogeneous effect across countries. In particular, I use the firm sample to test whether the increase in patenting by Japanese competitors may be explained by either of the following possibilities.

First, it could be that differences in observable firm characteristics across countries explain the heterogeneity. For example, Japanese firms may have had more previous experience in copier technologies than American firms, hence giving them higher values of the closeness measure. I address this possibility by running simple firm-level DiD

Table C2. Firm-Level Analysis: Regression Estimates

	(1)	(2)	(3)	(4)
$\text{Closeness}_i \cdot \text{Post}_t$	0.262			-0.247
	(0.338)			(0.359)
$\text{Stock}_i \cdot \text{Post}_t$	0.001			-0.001
	(0.008)			(0.007)
$\mathbb{1}[\text{Closeness}_i \geq p90] \cdot \text{Post}_t$		3.547	-0.441	0.247
		(2.376)	(1.590)	(1.847)
$\mathbb{1}[\text{Closeness}_i \geq p90] \cdot \text{Post}_t \cdot \text{Japan}_i$			29.046^{**}	29.589^{**}
			(13.411)	(13.520)
$\mathbb{1}[\text{Closeness}_i \geq p90] \cdot \text{Post}_t \cdot \text{Other}_i$			-1.632	-1.853
			(2.325)	(2.315)
Firm FE	✓	✓	✓	✓
Year × Country FE	✓	✓	✓	✓
Mean of Outcome	11.50	11.50	11.50	11.50
Firms	1635	1635	1635	1635
Observations	26160	26160	26160	26160

Notes: The table shows the results from several firm-level difference-in-differences regressions, where the outcome is the number of patent applications per year and firm. Closeness_i is the measure of closeness to Xerox defined in equation (4) and $\mathbb{1}[\text{Closeness}_i \geq p90]$ denotes firms in the top decile of the distribution of the closeness measure. The variable Stock_i represents the number of unexpired patents that a firm held as of 1975. The indicator Post_t equals one in years after 1975. The variables Japan_i and Other_i are country indicators that equal one for firms from Japan and countries other than the US and Japan, respectively. Standard errors clustered at the firm level are in parentheses. Significance levels: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

regressions, as shown in columns (1) and (2) of Table C2. Specifically, in column (1), I regress the number of patent applications per firm and year on the interaction of the closeness measure (Closenss_i) and a post-1975 indicator (Post_t). I add a second interaction that uses a firm's stock of unexpired patents (Stock_i) as an additional firm-level covariate. The regression also includes firm and year × country fixed effects, which may absorb any country-specific patenting trends. If the explanation above were correct and the heterogeneity across countries were driven by differences in firm observables, one would expect a significantly positive relationship between either of the interaction terms and the outcome variable. However, as shown in column (1), this is not the case: both estimates are quantitatively small and very imprecisely estimated. In column (2), I convert the closeness measure into a binary indicator that takes on the value of one for firms in the top decile of the distribution of the closeness measure ($\mathbb{1}[\text{Closeness}_i \geq p90]$). Although the estimate is much larger now, it is still statistically insignificant. This result indicates that, on aggregate, differences in the closeness to Xerox cannot explain the increase in Japanese patenting after 1975.

Second, one may be worried that the disproportionate increase in the number of patent applications among Japanese competitors could reflect a more aggregate country-specific pattern of increasing Japanese patenting. In the regression estimates in Table C2, such an effect should be absorbed by the year × country fixed effects. I address this second possibility in columns (3) and (4) of Table C2 by adding a triple interaction of an indicator for firms in the top decile of the distribution of the closeness measure ($\mathbb{1}[\text{Closeness}_i \geq p90]$), the post-treatment indicator (Post_t), and a country indicator for Japan (Japan_i). I include a similar interaction term with an indicator for all other countries other than the US and Japan (Other_i). If the increase in the number of patent applications were indeed driven by an aggregate increase in Japanese patenting, then the triple interaction with the Japan indicator should not be statistically significant, as the entire effect would be absorbed by the year $\times$ country fixed effects. Columns (3) shows that, again, this is not the case. The coefficient on the interaction with the Japan indicator is large and statistically significant at the 5% level. The point estimate implies that, following the consent decree, Japanese firms in the top decile of the distribution of the closeness measure filed an additional 29 patents per year on average, relative to American firms in the top decile and conditional on firm and year $\times$ country fixed effects. This estimate remains largely unchanged when controlling for the continuous closeness measure and the stock of a firm's patents, as shown in column (4). Intuitively, this result reflects the pattern visible in panel (A) of Figure 8 in the main part of the paper – that is, there was a disproportionate increase in patenting among Japanese competitors, relative to their American counterparts but also relative to other Japanese firms.

Overall, this auxiliary analysis on the firm level suggests that the increase in patenting among Japanese competitors can neither be explained by differences in terms of observable firm characteristics across countries nor by an aggregate patenting trend that is specific to Japan. Yet, an important caveat regarding the first result is that my analysis is limited to firm characteristics related to firms' patenting, because I do not have any other firm-level data that cover all firms in the sample. However, if the heterogeneity were explained by other common firm characteristics, one would expect such differences to be reflected (at least partly) in the patent data as well.

D Effect on Trade Values

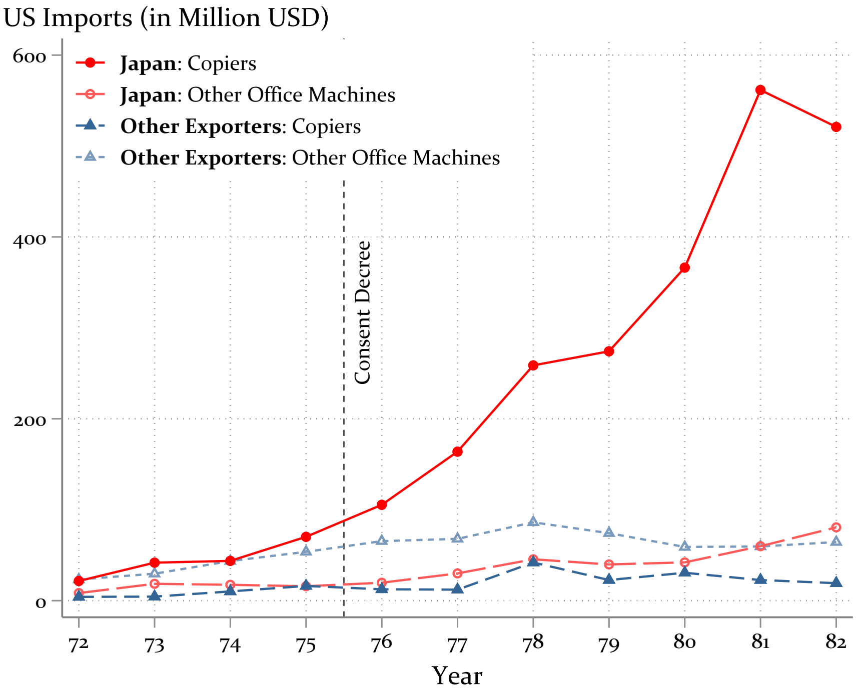
The empirical results of this paper show that compulsory licensing of Xerox's patents promoted innovation particularly among Japanese competitors. Yet, this finding does not imply that Japanese firms also increased their revenues or profits. Therefore, this appendix analyses trade data to assess whether Japanese competitors benefited from the antitrust case in terms of subsequent exports to the US.

I use data from Feenstra (1996) that report the annual value of US imports by commodity and exporting country.[27] US imports were originally classified based on the Tariff Schedule of the United States Annotated (TSUSA). The data from Feenstra (1996) report these highly disaggregated TSUSA classes, which allows me to precisely identify the value

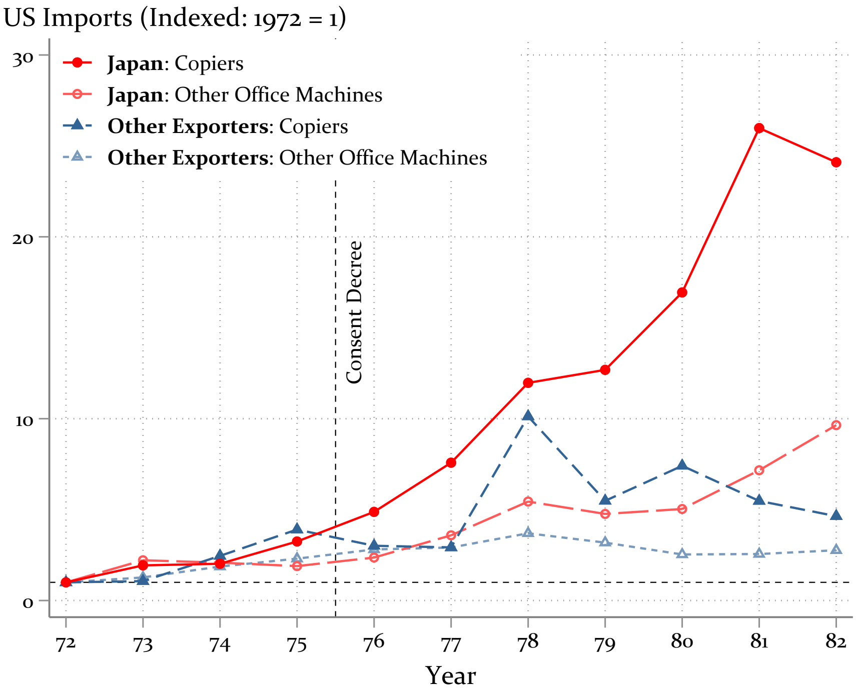
[27] These data are publicly available at `https://cid.econ.ucdavis.edu/usix.html`.

Figure D1. Effect on Trade Values: US Imports of Copiers and Other Office Machines

(A) Dollar Value



(B) Indexed



Notes: The figure depicts the value of US imports in the four-digit SITC class 7518, using data from Feenstra (1996). Trade values are shown separately for copiers and the remaining office machines in SITC class 7518. Copier imports are identified based on the more disaggregated TSUSA classes, which are also reported in the data. The figure also distinguishes trade values by exporting country. Panel (A) shows absolute dollar values. In panel (B), the time series are indexed to reflect growth rates relative to 1972.

of US copier imports. The data also contain a concordance to industry classes based on the Standard International Trade Classification (SITC, Rev. 2). Copiers are classified under the four-digit SITC class 7518, which covers 'office machines, n.e.s.' (i.e., not elsewhere specified). However, this class also contains goods other than copiers such as 'duplicating machines' or 'number, dating, and check-writing machines'.[28] One drawback of the data from Feenstra (1996) is that their coverage only starts in 1972, which is why I focus on the ten-year period until 1982.

Figure D1 depicts the value of US imports in the four-digit SITC class 7518. Panel (A) shows that Japanese copier exports to the US skyrocketed in the late 1970s and amounted to a value of $560 million (in current dollars) in 1981. In contrast, all other countries jointly exported copiers worth less than $60 million to the US in the same year. As further shown in panel (B), Japanese copier exports to the US grew at similar rates as the aggregate copier exports by all other countries from 1972 until 1975. Then, however, copier exports from Japan increased disproportionately and reached up to 26 times of their 1972 value. Figure D1 also shows trade values of other office machines that are classified in the same four-digit SITC class as copiers. These other office machines represent a natural control group for copiers. As is evident from the figure, Japanese exports of other office machines also increased after 1975, but this increase was less pronounced than that of Japanese copier exports. In addition, the total annual value of US imports of other office machines remained at a much lower value of around $140 million even in the early 1980s.

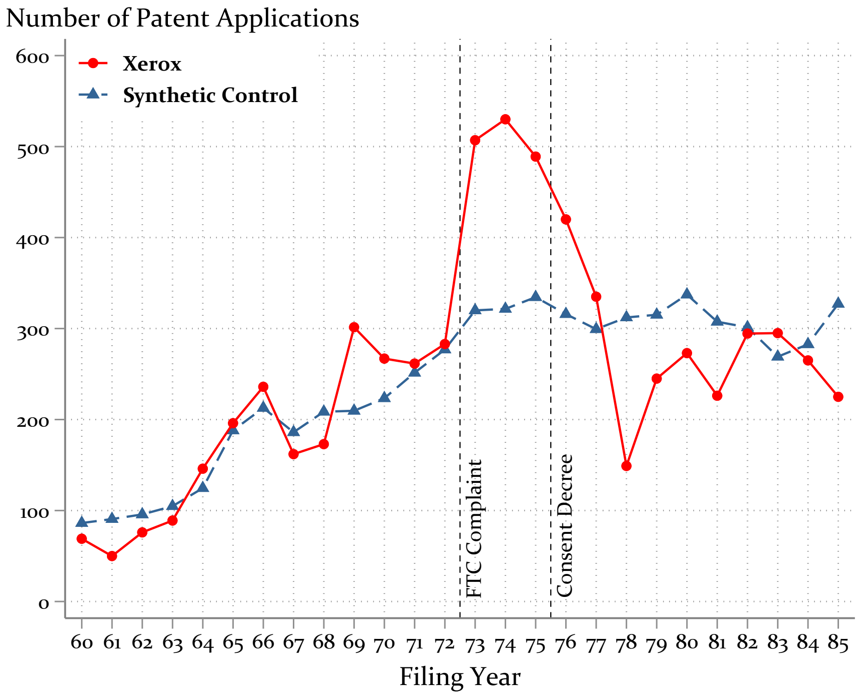
In summary, there are two key takeaways from my analysis of trade values. First, Japanese copier exports to the US increased after 1975, relative to the aggregate copier exports by all other countries. Second, there was no such disproportionate increase in Japanese exports for other office machines that are classified in the same four-digit SITC class. This descriptive evidence is consistent with my findings regarding the innovation effect of the antitrust case against Xerox. Therefore, the results on trade values support the conclusion that Japanese copier producers did not only benefit from access to Xerox's technology in terms of innovation; they also generated revenues in the product market by exporting copiers to the US.

E Supplementary Results for Effect on Xerox

In this appendix, I report supplementary results on the effect on Xerox by computing an alternative synthetic control group. As discussed in the main part of the paper, including Bell into the synthetic control may not represent a good counterfactual for Xerox's patenting, as Bell itself faced an antitrust lawsuit in the late 1970s and was

[28]For more information on SITC, Rev. 2, see `https://unstats.un.org/unsd/publication/SeriesM/SeriesM_34rev2E.pdf`.

Figure E1. Effect on Xerox: Patenting by Xerox vs. Alternative Synthetic Control



Notes: The figure depicts the number of patent applications per year for Xerox and its subsidiaries (in red) and an alternative synthetic control group (in blue). The synthetic control group is computed using the algorithm by Abadie et al. (2010, 2015) and consists of 71.4% Siemens, 21.9% Westinghouse, 4.9% General Electric, and 1.8% Ciba-Geigy.

broken up in 1984 (Watzinger and Schnitzer, 2022). Therefore, I compute an alternative synthetic control, where Bell is excluded from the donor pool. The resulting control group consists of 71.4% Siemens, 21.9% Westinghouse, 4.9% General Electric, and 1.8% Ciba-Geigy.

The patenting trends depicted in Figure E1 show an overall similar picture to Figure 10 in the main part of the paper. However, the synthetic control group now has a slightly higher average number of patent applications after 1975. As a consequence, the hand-computed DiD estimate now indicates a decline in Xerox's patenting by around 30 patents per year on average.

F Conceptual Framework

Why, in theory, should the removal of patent protection on Xerox's technology affect innovation by other firms? In this appendix, I introduce a brief conceptual framework explaining my key findings. In principle, patent rights should not hinder follow-on innovation, because efficient bargaining between the owner of an upstream technology and downstream innovators leads to ex-ante licensing (Green and Scotchmer, 1995). Consequently, any surplus-enhancing downstream innovation should be developed, irrespective of whether the upstream technology has patent protection or not.

However, patents may exert a blocking effect on follow-on innovation if upstream and downstream parties fail to reach a licensing agreement.[29] The economic literature has identified several reasons that may explain the absence of ex-ante licensing. On the one hand, there may be bargaining failure between the parties. This may arise, for example, due to asymmetric information (Bessen and Maskin, 2009) or coordination failure among downstream innovators (Galasso and Schankerman, 2015). On the other hand, rent dissipation may make licensing unprofitable for the upstream firm (Arora and Fosfuri, 2003). This is the case if the upstream firm's revenues from licensing are lower than its expected loss in profits due to increased product market competition.

In the case of Xerox, I argue that rent dissipation is the most likely reason for the absence of ex-ante licensing. Following the theoretical framework by Gaessler et al. (2019), the prevalence of rent dissipation versus bargaining failure can be assessed by considering (i) the degree of market overlap between upstream and downstream firms, (ii) the number of firms that control relevant technologies, and (iii) the size of the parties. In the case of Xerox, first, market overlap between Xerox and potential licensees was high. If Xerox had granted other firms unrestricted licenses, they may have become its direct competitors in the plain-paper copier market. Consequently, Xerox's monopoly rents would likely have dissipated. Second, Xerox was the single owner of the relevant technology, as it held almost all xerography patents. There was no need for downstream innovators to negotiate with several upstream parties, which speaks against bargaining failure. Third, both Xerox and potential licensees (e.g., Canon, IBM, Kodak) were large companies. This should have allowed face-to-face negotiations and facilitated cross-licensing agreements, again making bargaining failure unlikely (Gaessler et al., 2019).

The rent dissipation theory is also most consistent with narrative evidence about Xerox's patenting strategy. Xerox executives believed that, by giving a license to one competitor, they would also have to license everybody else (Jacobson and Hillkirk, 1986). However, this would likely have created product market competition by reducing Xerox's technological advantage. In hindsight, this concern seems justified, given that I find that compulsory licensing increased subsequent innovation primarily in the copier industry. A further aspect that speaks against bargaining failure is that, in fact, Xerox did enter a number of licensing agreements with other firms. However, all licenses were restricted to manufacturing products other than plain-paper copiers.

Overall, a rent dissipation effect can plausibly explain the absence of ex-ante licensing between Xerox and potential follow-on innovators. Consequently, Xerox's patents may have blocked follow-on innovation during the monopoly period. This changed in 1975

[29] This blocking effect has also been identified empirically in several studies (e.g., Williams, 2013; Galasso and Schankerman, 2015; Gaessler et al., 2019). However, the evidence is still mixed, as other studies find no effect of patent protection on follow-on innovation (e.g., Sampat and Williams, 2019). For a broader overview of how patents affect research investments, see Williams (2017).

when antitrust intervention gave other firms access to Xerox's technology and, therefore, may have enabled subsequent innovation.