Fostering the Diffusion of General Purpose Technologies: Evidence from the Licensing of the Transistor Patents

Markus Nagler (FAU Erlangen-Nuremberg, CESifo and LASER)
Monika Schnitzer (LMU Munich, CESifo and CEPR)
Martin Watzinger (University of Muenster, CESifo and CEPR)

Discussion Paper No. 297

November 15, 2021

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

Fostering the Diffusion of General Purpose Technologies: Evidence from the Licensing of the Transistor Patents[*]

Markus Nagler[†], Monika Schnitzer[‡], and Martin Watzinger[§]

How do licensing and technology transfer influence the spread of General Purpose Technologies? To answer this question, we analyze the diffusion of the transistor, one of the most important technologies of our time. We show that the transistor diffusion and cross-technology spillovers increased dramatically after AT&T began licensing its transistor patents along with symposia to educate follow-on inventors in 1952. Both these symposia and the licensing of the patents itself played important roles in the diffusion. A subsequent reduction in royalties did not lead to further increases, suggesting that licensing and technology transfer were more important than specific royalty rates.

---

[*] We thank the editor, Ryan McDevitt, as well as Oliver Falck, Katrin Hussinger, Thomas Misa, Markus Trunschke, two very helpful referees, and conference participants at the annual conferences of the EEA 2020, the Verein für Socialpolitik 2020, the ZEW Conference on the Economics of ICT 2021, and the Munich Summer Institute 2021 for helpful comments and suggestions. We thank Dr. Sheldon Hochheiser for sharing relevant documents with us. Financial support by Deutsche Forschungsgemeinschaft through CRC TRR 190 is gratefully acknowledged. Nagler thanks the Joachim Herz Foundation for financial support. Watzinger thanks the REACH - EUREGIO Start-Up Center for their support.

[†] Author's affiliation: Friedrich-Alexander-University Erlangen-Nuremberg, CESifo Munich, and LASER, Lange Gasse 20, D-90403 Nuremberg, Germany, *e-mail: markus.nagler@fau.de*

[‡] Author's affiliation: Ludwig-Maximilians-University Munich, Germany, CESifo Munich, and CEPR London, Akademiestr. 1, D-80799 Munich, Germany, *e-mail: schnitzer@econ.lmu.de*

[§] Author's affiliation: University of Muenster, Germany, CESifo Munich, and CEPR London, Am Stadtgraben 9, D-48143 Muenster, Germany, *e-mail: martin.watzinger@wiwi.uni-muenster.de*

I. Introduction

Historical accounts suggest that the diffusion of General Purpose Technologies (GPTs) and thus technological progress and economic growth can be hampered by patent protection. The best known example is James Watt's steam engine patent. Mokyr [1994], among others, writes that 'because [Watt] held a wide-ranging patent, he succeeded in blocking [the development of high-pressure steam engines] for many years' (p. 24, quoted in Selgin and Turner [2011]). According to Boldrin and Levine [2008], 'by keeping prices high and preventing others from producing cheaper or better steam engines, Boulton and Watt hampered capital accumulation and slowed economic growth' (p.4).[1] Similarly, the Wright brothers' patent war is blamed for stalling the development of the U.S. aviation industry, and Selden's patent on an internal combustion engine allegedly slowed automobile development in the early 20th century (e.g., Merges and Nelson [1990, 1994]). These narratives of harmful patents on key technologies are often used as prime examples for the 'case against patents', suggesting that patenting rights should be weakened or abolished altogether.

---

[1] Boldrin and Levine [2008] also recount the story that improvements to Watt's inventions were blocked by patents of rival inventors, highlighting the mutual spillovers between earlier and subsequent developments prevalent in General Purpose Technologies. For a more positive view on Watt's patent and more context on its alleged blocking effects, see Selgin and Turner [2011]. For another critical view of the alleged hold up by the Wright brothers, see Katznelson and Howells [2015]. As another example, Edison's patent on the incandescent lamp allegedly led inventors to invent around Edison's key technology (Katznelson and Howells [2012]).

Patents on GPTs might be particularly harmful because they can impede positive feedback loops, the key characteristic of General Purpose Technologies. Improvements in the GPT stimulate innovations in the application sector, which in turn give incentives to improve the GPT. But this feedback loop is only possible if patents on the GPT do not block follow-on inventions, either in the application sector or for the GPT itself. Patents have been shown to block follow-on invention in various settings (Moser and Voena [2012]; Williams [2013]; Sampat and Williams [2019]; Gaessler et al. [2019]; Watzinger et al. [2020]). But it is not clear whether this is a relevant concern for GPTs as their potential benefits are so large that they might provide sufficient incentives for efficient technology licensing (Green and Scotchmer [1995]; Galasso and Schankerman [2015]). In addition, patent licensing per se may not help follow-on inventors if tacit knowledge is important in making use of the patent. This is especially true since patent disclosure is often not complete (see, e.g., Roin [2005]; Ouellette [2012]). Understanding whether patents block the diffusion of GPTs is important because while GPTs are rare, they are credited with driving sustained economic growth since the industrial revolution (e.g., Helpman [1998]).

In this paper, we study the effects of patent licensing and active knowledge transfer on follow-on inventions to the transistor, the defining General Purpose Technology of the 21st century (Bresnahan and Trajtenberg [1995]; Helpman [1998]).[2] From early applications such as hearing aids and pocket radios to modern technology like fast computer chips and smartphones, the transistor and its subse-

---

[2] Note that while the concept of GPTs has been popular to characterize important technologies that influence broad parts of the economy since the seminal paper on the topic by Bresnahan and Trajtenberg [1995], there is some debate on the use of the term. See, e.g., Field [2008] for an overview. For some common definitions, see, e.g. Jovanovic and Rousseau [2005]; Bresnahan [2010]. For example, Jovanovic and Rousseau [2005] state that GPTs (i) spread to most sectors of the economy, (ii) improve vastly over time and (iii) generate substantial spillovers by allowing the invention of new products. We believe that the transistor fulfills these criteria.

quent developments spread to almost all sectors of the economy.[3] The first working transistor was invented in 1947 by American physicists John Bardeen, Walter Brattain, and William Shockley at the Bell Laboratories. The three shared the 1956 Nobel Prize in Physics for their achievement. The Solid State Physics Group at Bell responsible for the transistor filed 166 patents, of which 110 were published by 1952. We refer to these patents collectively as 'transistor patents.'

In 1952, the Bell System decided to license the transistor patents at a standardized rate of $25,000 and provided training programs for all firms who bought such licenses (Holbrook et al. [2000]; Reid [2001]). Commentators saw this generous licensing regime as a calculated political move to appease the authorities in an ongoing antitrust case against the Bell System that sought to break up the company (Mowery [2011]; Gertner [2012], p.111). But according to internal memos at the Bell Labs written a decade later, engineers at the Bell Labs also understood that 'by involving engineers around the world in the evolution of the device - making it better, cheaper, more reliable - the hope was that everyone would profit from the advances, especially the Bell System' (Gertner [2012], p. 375). The standardized licensing opened the transistor technology, reducing the entry barriers to the industry as one commentator vividly described: 'If you were going to be a player in semiconductors in the early 1950s, you'd wish you knew the AT&T patent lawyer just as you wish you knew your rich uncle' (Carrick [1982],p. 33).

There are many stories of how a diverse set of entrepreneurs and inventors benefited from the easily accessible license and the training. Jack Kilby, the eventual co-inventor of the integrated circuit, got his start with the transistor technology when he attended Bell's ten-day crash course that came with buying a

---

[3] Holbrook et al. [2000] tells four case studies of companies that build on the transistor patents of Bell.

license (Reid [2001], p. 71-72). Masaru Ibuka licensed the transistor patents in 1953 to build a transistor radio at SONY, at the time a young company that he had co-founded and that was struggling to stay in business. By 1957, SONY had issued a pocket transistor radio that sold over 1.5m units and had become an internationally known company (Nathan [2001]; Flamm [2010]). The licensing and technology transfer of the transistor technology arguably also led Pete Haggerty of Texas Instruments to hire Gordon Teal to build the first transistor pocket radio in the U.S., starting the rise of Texas Instruments to become one of the biggest technology companies in the world (Reid [2001], p. 73).

To see whether the licensing and technology transfer increased follow-on invention to the transistor, we compare the number of follow-on innovations building on the transistor patents with the number of follow-on innovations building on control patents before and after standardized licensing was implemented. We measure follow-on innovations using patent citations. As control group we use exactly matched non-Bell patents with the same filing year, the same technology class, and the same number of citations until 1952, i.e., before the standardized licensing started. We provide extensive evidence that our empirical strategy is robust using a variety of alternative identification strategies. Most importantly, we show that an alternative identification strategy not based on matching and using within-patent variation yields qualitatively identical results.

We find that the standardized licensing of the transistor technology led to a jump in patents building on Bell's transistor patents. In particular, it increased cross-technology spillovers. As cross-technology spillovers are a defining characteristic of General Purpose Technologies, this suggests that patents on GPTs might indeed be more harmful. We find that follow-on invention by the attendees of the Transistor Symposia was particularly affected relative to baseline patenting. However, in absolute terms, the effect is driven by inventors that did not participate in these training sessions. This suggests that both information transfer through the Transistor Symposia and the standardized licensing per se played important roles in increasing the diffusion of the transistor. As hoped by the engineers of the Bell System, the licensing led to the involvement of a larger number and a more diverse set of inventors. The impacts are driven by inventors unrelated to the Bell System, working in unconcentrated markets. A disproportionate share of the increase is driven by young and small companies, suggesting that licensing can promote the entry of small firms (Lanjouw and Schankerman [2004]; Galasso [2012]).

Closest to our paper is Watzinger et al. [2020], which studies the innovation effects of the 1956 compulsory licensing of Bell's patents on follow-on innovation. Our paper goes beyond that in two important ways: First, focusing on General Purpose Technology patents allows us to uncover that GPT patents differ from regular patents in their impact on cross-technology spillovers. Second, using a different treatment, namely voluntary standardized licensing for significant royalties instead of compulsory licensing with zero royalties, we can shed light on the relevance of royalties for follow-on innovation. Third, while we borrow our main identification strategy from Watzinger et al. [2020], we also introduce a to our knowledge entirely novel identification strategy in the robustness section comparing follow-on innovation building on the *same* patent across differentially affected fields ("within-patent identification"). This alternative identification strategy addresses potential concerns about the suitability of our matching strategy for an extraordinary technology such as the transistor.

Our study adds empirical evidence to case studies on the effect of patents on important technologies as recounted in Boldrin and Levine [2008]. It shows that the effect of patents on technologies with significant potential for cross-technology spillovers might be particularly harmful. This calls for tailor-made solutions for such technologies, for example compulsory or incentivized licensing or patent buyouts (Kremer [1998]). Some firms may even have an incentive to openly license their patents to learn from competitors, as suggested by the internal memos at Bell. As a recent example, Tesla has pledged to not enforce their patent rights.[4] Licensing may also be fruitful for GPT inventors since they may benefit from complementary follow-on innovation or complementary assets in downstream firms (e.g., Arora and Ceccagnoli [2006]; Lerner et al. [2007]).

This study also contributes to the literature on the impacts of patents on follow-on innovation.[5] Galasso and Schankerman [2015] study the effect of patent invalidation on follow-on innovation as measured through patent citations and find an average increase of 50%. Sampat and Williams [2019] study whether patents on genes reduce follow-on innovation, but find no effect. Murray and Stern [2007] and Moser and Voena [2012] study patent removals and find increases in follow-on innovation of 10-20% in biotech and chemistry. Gaessler et al. [2019] study patent invalidation at the European Patent Office and find sizable effects on innovation. We add to this literature by showing that the impact of patent licensing on follow-on innovation is substantially stronger when patents cover a GPT. We also provide evidence that the type of follow-on innovation that is blocked by these patents differs from follow-on innovations blocked by less exceptional patents. In addition, we show that the role of royalties in blocking follow-on innovation is limited.

Finally, our paper contributes to the literature on the history of U.S. innova-

---

[4] See https://www.tesla.com/blog/all-our-patent-are-belong-you, last accessed May 12, 2021.

[5] For a recent survey, see Williams [2017].

tion with the first in-depth analysis of the diffusion of the transistor technology. Already in 1962, Richard Nelson highlighted that the transistor 'has stimulated growth, including the invention and innovation on a considerable scale of products which can profitably use transistors as components' (Nelson [1962], p. 553). Although the enormous significance of the transistor technology is widely recognized and the importance of the non-discriminatory licensing by Bell has been suspected to have played a crucial role for its diffusion (e.g., Levin [1982], quoted in Merges and Nelson [1994]), this paper is the first to provide an empirical analysis of how important the licensing decision of the technology by Bell was for the inventions in the semiconductor industry.

II. The Bell System and the Transistor

In the early 1950s, American Telephone & Telegraph (AT&T) was the dominant provider of telecommunications services in the U.S, owning or controlling 98% of all facilities providing long distance telephone services and 85% of those providing short distance telephone services. Together, the Bell system employed around 750,000 people. It generated total revenues of $5.3 billion or 1.9% of the U.S. GDP in 1950 (Antitrust Subcommittee [1959]; Temin and Galambos [1987]; Watzinger et al. [2020]).[6] Its R&D subsidiary, the Bell Laboratories (Bell Labs), were arguably the most innovative industrial laboratory of the time. The Bell Labs produced path-breaking research in applied and in basic science. Several of the scientists employed by Bell Labs in the 1950s were subsequently awarded prestigious research prizes, such as the Nobel Prize, the Turing Award, and the IEEE Medal

---

[6] More details on the Bell System can be found in web appendix A. See the Journal's editorial web site for further details and all web appendices.

of Honor. Their inventions included the development of radio astronomy (1932), cellular telephone technology (1947), information theory (1948), solar cells (1954), the laser (1957), and the Unix operating system (1969).

The most important invention of the Bell Labs was the transistor in 1947. Bell filed for patents on the first transistor in June 1948 and announced the invention on July 1 of the same year. The patents were published in 1950 and 1951. Bell, the military, and the research community at large immediately understood the importance of the transistor. The Nobel Prize in physics for the original inventors followed in 1956. The public was enthusiastic about the workings of the new technology. TIME Magazine ran a story concluding that 'to all industrial needs, and most human physical needs, the electronics magicians are sure they know the key' (quoted in Reid [2001], p. 61).

Transistors switch and amplify electric current, a skill that almost all electric devices require. As an example for the switching function of transistor, modern microchips have billions of transistors printed on them and work by switching on and off combinations of these, which can then be interpreted by software through logic combinations. As an example for amplification, hearing aids translate currents picked up via microphones to bigger currents via tiny loudspeakers, increasing the sound by basically just amplifying electric current. Before the transistor, devices that required the switching and amplification of electric current relied on vacuum tubes. These tubes were however quite large as well as relatively sensitive. For example, they often burned out. In comparison to vacuum tubes, transistors were much smaller, more efficient, more reliable, more durable, safer, and more economical. The transistor consequently revolutionized the way in which electric current was switched on and off as well as amplified in nearly all applications that required this. More importantly, transistors allowed entirely new products to be manufactured, for example hearings aids, pocket radios, or microchips. Richard Nelson gave a vivid illustration of the importance of the transistor in 1962:

> 'The transistor has had its most significant impact not as a component replacing vacuum tubes in established products, but as a component of products which were uneconomical before the development of the transistor. Very compact computers are the most striking example. Without transistors, computers of a given capability would have to be much larger both because vacuum tubes are larger than equivalent transistors and because cooling requirements are much greater for vacuum tubes. Almost all of our new airborne navigation, bombing, and fire control systems, for example, are transistorized. So are all of our satellite computers. And without transistors our large computers [...] undoubtedly would be much more expensive - probably so much so that many of their present uses would not be economically sound.' (Nelson [1962], p. 553)

In 1952, Bell started to license all its transistor patents in an open and standardized way to private companies. Commentators at the time thought that this was a political move to appease the regulator in an ongoing antitrust trial. In 1949, US Government filed an antitrust lawsuit with the aim to split up the company (Antitrust Subcommittee [1959],p.31). According to experts, because of the ongoing antitrust lawsuit, Bell's management was reluctant to draw attention to its market power by charging high prices for transistor licenses (Reid [2001]; Mowery [2011]). As a consequence, Bell's top managers agreed to share and license the transistor device with standardized non-discriminatory licensing contracts (Gertner [2012], p.111). Bell's management also decided to actively promote the transistor by organizing conferences, the Transistor Symposia, to explain the technology (Holbrook et al. [2000]). Among Bell's engineers, there was the perception that standardized licensing would help Bell technology-wise. For example, Bell's Jack Morton, the inventor of the microwave tube, advocated the licensing of transistor-related patents as he saw an opportunity to learn from other companies.

In September 1951, Bell held a first meeting at Bell Labs for scientists and engineers to visit the lab and learn about the technology. This meeting was designed specifically for inventors working on military applications as well as the technical and procurement arms of the U.S. military (Holbrook et al. [2000]). In addition, Bell waived all patent royalties on the first important transistor product, the miniature hearing aid in homage to Alexander Graham Bell's work on these devices (Reid [2001], p. 60). In April 1952, a second nine-day conference with over 100 representatives from almost 40 private companies gathered for the 'Transistor Technology Symposium'. The conference conferred information about manufacturing techniques as well as the workings of the transistor, including substantial informal and tacit knowledge (Holbrook et al. [2000]). After the conference, over 30 companies decided to license the transistor technology for a nonrefundable advance payment of \$25,000 ($\sim$\$245,000 in today's dollars) that was credited against future royalty payments (Antitrust Subcommittee [1958], p.2957). Royalty rates amounted to 5% of the net selling price of the transistor in 1950, which were reduced to 2% in 1953 (Antitrust Subcommittee [1959], p. 117).

Various now well-known companies made use of this offer. Centralab licensed the transistor and made Jack Kilby, the eventual co-inventor of the integrated circuit, go to the transistor conference to use the technology in his inventions (Reid [2001], p. 71). Masaru Ibuka led SONY to license the transistor in 1953 and developed pocketable transistor radios that were huge commercial successes (Nathan [2001]). And Texas Instruments hired Bell's Gordon Teal in 1952 to scale transistors to mass production, eventually leading to U.S. manufactured pocket radios (Reid [2001], p. 73). Bell was also successful in continuing to invent new technologies around the transistor. For example, in 1959 two researchers at Bell Labs, Mohamed Atalla and Dawon Kahng, invented the metal-oxide-semiconductor field-effect transistor (the 'MOSFET'), the most widely manufactured device in history.

Whether the political move of the Bell System to license the transistor openly made a difference to the antitrust case is unclear. The antitrust lawsuit went back and forth over several years, ending in the 1956 consent decree that required Bell to share all its granted patents royalty-free and all subsequently published patents for reasonable royalties (Watzinger et al. [2020]). This was perceived as a major win for the Bell System that continued to be the monopolistic provider of telecommunication in the US until it was finally broken up through another antitrust suit in 1984 (Watzinger and Schnitzer [2021]).

III. Estimation Framework and Data

A. Data and Summary Statistics

To be able to analyze the effects of the transistor licensing, we identify all patents related to the Solid State Physics Group at the Bell Labs.[7] There are two main transistor patents: Patent #2,524,035 with the title 'Three-Electrode Circuit Element Utilizing Semiconductive Materials' granted in 1950 to John Bardeen and Walter

---

[7] Researchers whom we classify to have participated in this group and thus to have actively contributed to the transistor at Bell Labs were in alphabetical order Bardeen, Becker, Brattain, Buehler, Gomperez, Green, Haynes, Little, Morgan, Ohl, Pearson, Pfann, Scaff, Shive, Shockley, Sparks, Storks, Teal, Theurer, and Zinc (Nelson [1962]; Buehler [1983]).

Brattain and Patent #2,569,347 with the title 'Circuit Element Utilizing Semiconductive Material' issued to William Shockley in 1951. To these two patents, we add all patents of all researchers who actively worked towards the development of the transistor at Bell Labs. We identify 164 'transistor' patents held by Bell Labs (i.e., affected by 1956 the consent decree).[8] 110 of those were published up to 1952. We also delete the 27 patents that were published with delay due to secrecy orders during World War II (Gross [2019]). This sample is most likely a super-set of all transistor patents. For example, it also includes patent #2,402,662 with the title 'Light Sensitive Device' granted to Russell Ohl, the original patent of the solar cell, a semiconductor but not a transistor patent in the narrow sense.[9]

Table 1 shows summary statistics of the unweighted raw data. All patent data is from the Worldwide Patent Statistical Database (PATSTAT) of the European Patent Office. Column (1) reports summary statistics for the patents in our estimation sample that are not transistor or other Bell patents but that are part of our control group, i.e., are in the same technology classes as affected Bell patents, have the same number of citations in the five years before 1952 than some Bell patent in our treatment group, and were published in the same year as some Bell patent. Columns (2) and (3) split these control group patents according to their use in telecommunications. We classify a patent as a telecommunications-related patent if in its patent class patents have a probability of more than 15% of being used in the production of telecommunications equipment according to the data

---

[8] We identify all patents owned by the Bell System with the help of a list of patent numbers published in the 'Hearings before the Antitrust Subcommittee' of the U.S. Congress on the 1956 consent decree of Bell in May 1958 (Antitrust Subcommittee [1958]). The list is the complete list of all patents owned by the Bell System in January 1956. Of these patents, we drop all that have assignee names other than companies of the Bell System. The list also includes patents of Typesetter Corp., which were explicitly excluded from compulsory licensing in Section X of the consent decree. We assume that these patents are not part of the Bell System.

[9] In web appendix B, we show that our results are robust to using text-based or co-citation based definitions of the transistor.

Table 1: Summary Statistics

	(1)	(2) Control Group	(3)	(4)	(5) Transistor Patents	(6)
	All	Telecommunications	Others	All	Telecommunications	Others
	mean	mean	mean	mean	mean	mean
Filing Year	1941.52	1942.32	1941.45	1941.76	1943.88	1941.11
Publication Year	1944.69	1945.46	1944.62	1944.71	1947.00	1944.00
# Years in patent protection after 1952	9.69	10.46	9.62	9.71	12.00	9.00
Total cites	2.28	2.96	2.94	6.21	9.53	7.04
Citations by others	2.10	2.83	2.73	5.03	6.88	6.18
Self Citations	0.18	0.13	0.22	1.18	2.65	0.85
Citations by others prior to 1952	0.55	0.90	0.52	1.17	1.35	1.11
Observations	2287	184	2103	72	17	55

Notes: The table reports the average filing and publication year, the average number of years until patent expiration and citation statistics for published patents as follows: Column (1) describes all patents published until 1952 of non-Bell System companies that were published in the same year, in the same technology class, and with the same number of citations than our transistor patents (i.e., our control group patents). Columns (2) and (3) split all control patents published until 1952 according to their use in telecommunications. We classify a patent as a telecommunications-related patent if in its patent class patents have more than a 15% likelihood of being used in the production of telecommunications equipment according to the data of Kerr [2008]. Columns (4) through (6) repeat the same figures for Bell's transistor patents. A citation is identified as a self-cite if the applicant of the cited and citing patent is the same. The data are from the Worldwide Patent Statistical Database (PATSTAT) of the European Patent Office.

of Kerr [2008]. Columns (4) through (6) repeat the same summary statistics for transistor patents, i.e., patents in our treatment group. The average control group patent in our data set receives 2.1 citations per patent from other inventors while our transistor patents receive on average 5 citations by others. Before the second Transistor Symposium in 1952, the average non-Bell patent receives 0.5 citations by others while the average Bell transistor patent receives 1.2 citations.

B. Estimation Framework

To measure follow-on innovations building on Bell transistor patents, we use patent citations in our main specification (Williams [2015]). Citations give us a direct link between follow-on innovations and Bell's transistor patents.[10] To construct a measure of what would have happened to the follow-on innovation building on Bell's transistor patents in the absence of the licensing, we use as control group *all* other patents that are published in the same year, that have the same total number of citations as the Bell transistor patents in the five years before 1952, and that are in the same USPC technology class. We condition on the publication year because young patents are cited more often on average. We condition on prior citations to control for a patent's potential for follow-on inventions. We also match on the same technology class to control for the number of potential follow-on inventors and for technology-specific citation differences.

To quantify the difference in the number of follow-on innovation to Bell transistor patents and to control patents we use the following specification:

---

[10] Citations are also consistently available from 1947 onward, in contrast to most alternative measures such as new products or R&D spending. Citations have the additional advantage that they have a high frequency, which allows a precise measurement of effects. The caveat is that some citations might have been added by the patent examiner, which adds noise to the measure (Alcacer and Gittelman [2006]; Alcacer et al. [2009]).

$$(1) \quad \#Citations_{i,t} = \beta_1 \cdot Transistor_i + \beta_2 \cdot Post_t + \beta_3 \cdot Transistor_i \cdot Post_t + \varepsilon_{i,t}$$

where $\#Citations_{i,t}$ is the number of citations of other companies to patent i from 1953 until patent expiration (the treatment period). Note that this implies that most of our effects are driven by citations in the 1950s, shortly after Bell's decision. $Transistor_i$ indicates whether patent i is a transistor patent owned by the Bell System and is therefore treated. The coefficient of interest is β_3, which reflects the difference in follow-on citations to Bell's transistor patents relative to patents in the control group.

We can interpret our results as causal if, in the absence of the licensing, the number of citations to control patents have the same trend as the Bell's transistor patents would have had in absence of licensing (parallel trends). This assumption *does not* require that transistor and control patents necessarily have the same underlying quality or value, which would be doubtful in our setting. We only assume that in the absence of the licensing both treatment and control patents would have continued to receive the same number of follow-on citations.

There are three main limitations of this study. First, the transistor, similar to the steam engine or electricity, was a once-in-a-century invention. Therefore, finding a suitable control group of patents is challenging. The key assignees of control group patents are General Electric, RCA, Westinghouse, the key competitors of Bell Labs that were exempted from the 1956 consent decree (see Watzinger et al. [2020]). The patent in the control group that received the highest number of lifetime citations is by RCA, namely patent #2,354,591 on the 'Television apparatus'. This is followed by Wright Aeronautical patent #2,255,203 ('Fuel injection spark plug’), the two General Electric patents #2,536,805 (‘Hall effect telemetering transmitter’) and #2,569,345 (‘Transistor multivibrator circuit’), and Edwin Vonada’s patent #2,556,017 (‘Electrolytic method and apparatus for cleaning strip’). While all of these patents were important and experienced substantial follow-on innovation, even the RCA television patent is not similar in its generality to the transistor patents of Bell. We address this concern by using several different identification strategies to show that our result is robust. Among others, we construct a ‘within-patent’ control group that does not depend on matching patents. To do this, we compare citations to transistor patents from technologies close and far from telecommunications, holding the patent under consideration fixed. This draws on the insights in Watzinger et al. [2020] that Bell continued to foreclose the market in telecommunications, making it impossible for competitors to enter the market. Thus, we would not expect impacts of standardized licensing on citations in telecommunications while we would expect an effect outside of telecommunications. This is what we find.

A second limitation of this study is that we cannot conclusively say whether the resulting follow-on inventions increased or whether they just happened earlier. For example, it seems doubtful that no one would have thought of the integrated circuit eventually. But given our results it seems unlikely that Jack Kilby would have invented it as early as 1959. A third limitation is that with every license of the transistor an extensive training course in the production of transistor devices came along. We leverage the list of attendees of the Transistor Symposia to assess how much of the effects are driven by information transfer through the Symposia versus the licensing in itself. While this provides suggestive evidence for the relative importance of each effect, we cannot ultimately disentangle the two.

IV. The Diffusion of the Transistor

The Impact of Licensing on Subsequent Innovation

We first compare citations to Bell's transistor patents to citations to exactly matched non-Bell patents in the same technology class, published in the same year, and with the same number of citations up to (but excluding) 1952, the year of the main Transistor Symposium.

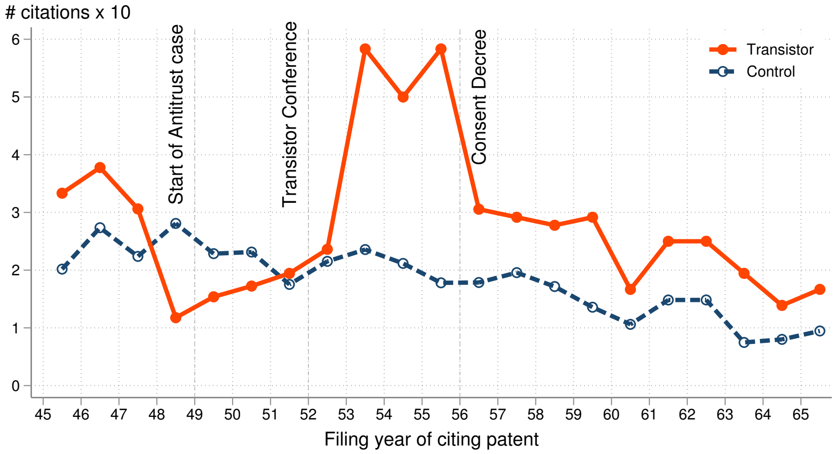
Panel (a) of Figure 1 plots the average number of citations to transistor patents relative to control patents over time. While these rates are similar before the second Transistor Symposium in 1952, citations to Bell's transistor patents spike after the conference, reverting a bit after Bell's consent decree in 1956. Most of the effect of the licensing and technology transfer thus is visible in the 1950s, shortly after the decision by Bell and the Symposia. However, citations to transistor patents remain higher until at least 1965. Panel (b) of Figure 1 shows two-yearly excess citations to Bell's transistor patents relative to the control group, adapting equation 1. The impact of licensing is again visible right after the Transistor Symposium. This suggests that patent licensing and active knowledge transfer had a positive impact on follow-on innovation. The fact that the impact does not increase further after 1956, when the consent decree that settled the antitrust lawsuit against AT&T reduced licensing fees to zero, suggests instead that the subsequent price reduction had little further impact. What mattered was the access to Bell's transistor patents.[11]

---

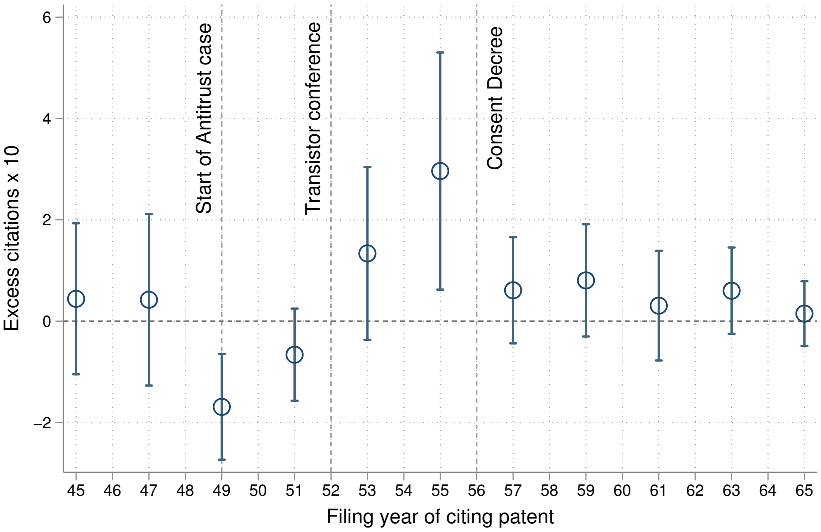
[11] Note that the *decrease* is due our empirical strategy that requires both treatment and control patents to have been published by 1952. Because of the fast pace of technological change in these areas, citations drop relatively soon. In alternative empirical strategies that do not make this requirement, we do not observe a decrease in the effect. We however do also not see further increases in treatment effects after 1956, again suggesting that it was the compulsory licensing decision and not the price reduction which mattered for the diffusion of the transistor technology. See, e.g., Section G in the Appendix.

Figure 1: The Impact of Standardized Licensing on Excess Citations to Transistor Patents

(a) Means



(b) Two-Yearly



Notes: Panel (a) shows the average number of citations to Bell's transistor patents in every year after publication. The red line with solid circles shows patent citations of the treated patents (Bell transistor patents) and the blue line with empty circles shows patent citations of control patents, with the same publication year and the same three-digit technology class as the Bell transistor patents. For aggregation, we use the weights of Iacus et al. [2009] to adjust for a different number of control patents for each Bell patent. Panel (b) shows the number of two-yearly excess citations to transistor patents published before 1952 relative to patents with the same publication year, in the same three-digit U.S. Patent Classification (USPC) primary class and with the same number of citations up to (and including) 1951, estimated adjusting the specification in equation (1). We correct for self-citations. The blue lines represent the 95% confidence bands calculated from standard errors clustered on the three-digit technology class level. To adjust for the different number of control patents per treatment patent in each stratum, we use the weights suggested by Iacus et al. [2009]. The sample under consideration contains 110 transistor patents, 83 of which were not affected by the secrecy program. We can match 72 transistor patents. All coefficients are multiplied by 10 for better readability.

We quantify this in Table 2. In column (1), we report the results from our baseline regression, equation 1. The treatment period is defined to start in 1953 (the year after the second Transistor Symposium) and to last until the expiration of the patent. We find that yearly excess citations to the transistor patents increase by around 135% relative to the control group mean in the treatment period.[12] General Purpose Technologies are typically applied in a variety of downstream innovations. Thus, the blocking effects of patents on other technologies than the patent's own may be particularly large. The next two columns therefore show the impacts of licensing on the breadth of use of the transistor technology. In columns (2) and (3), we split our dependent variable by whether the citations accrued in the same technology class as the underlying patent or in a different technology class.[13] The effects seem somewhat larger in the same technology class than the licensed patent, but are also strong in technology classes different than the one of the underlying patent.[14]

Is this specific to GPTs or would we expect similar results in other cases of licensing? To provide evidence on this, in columns (4) through (6) we show the same results for the compulsory licensing of Bell's patents in the 1956 consent decree (Watzinger et al. [2020]). We drop all transistor-related patents from this specification. Two points become evident. First, the impact of the transistor licensing on follow-on innovation was substantially higher than the impact of compulsory licensing on regular Bell patents. This is in line with historical accounts that suggest a particularly harmful role of patents for the diffusion of GPTs. Sec-

---

[12] To arrive at this number, we relate the coefficient of 2.02 to the control mean of 1.48 in the treatment period.

[13] We use IPC categories to disentangle same and different technology since these reflect intended use more than the USPC classification does (Lerner [1994]).

[14] This pattern is also true when using value-weighted citations as the dependent variable, for example when using Dollar-weighted citations using the values of Kogan et al. [2017] or when using citation-weighted forward citations (not shown).

ond, the results show that the increases in citations following the compulsory licensing of Bell's patents in the consent decree were concentrated in the same technology classes as the underlying patents. This is in contrast with the results from columns (1) through (3). These results are consistent with a more important role of patents on general purpose technologies for cross-technology spillovers.

Table 2: The Effect of Standardized Licensing on Follow-on Innovation

Dep. Var.:	Citations					
	1952 Transistor Licensing			1956 Consent Decree		
	Baseline	Same Tech.	Diff. Tech.	Baseline	Same Tech.	Diff. Tech.
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	0.08	-0.30**	0.38***	-0.05	-0.15***	0.10***
	(0.14)	(0.14)	(0.12)	(0.04)	(0.03)	(0.04)
Post	-0.43	-0.32	-0.11	-0.48***	-0.28***	-0.21***
	(0.43)	(0.20)	(0.30)	(0.05)	(0.04)	(0.03)
Treated x Post	2.02***	1.14***	0.88**	0.18***	0.15***	0.03
	(0.55)	(0.32)	(0.37)	(0.06)	(0.04)	(0.04)
Control Mean	1.48	0.73	0.75	1.09	0.50	0.60
# treated	72	72	72	3556	3556	3556
Clusters	30	30	30	206	206	206
Obs.	35629	35629	35629	657126	657126	657126

Notes: This table shows the results from difference-in-differences estimations with Bell Labs transistor (Columns 1-3) and other Bell Labs (Columns 4-6) patents as treatment groups. We define patents as transistor patents if they were filed by one of the researchers who actively worked towards the development of the transistor at Bell Labs. In columns (1) to (3), we define the treatment period as starting in 1953. In these columns, treated is an indicator variable equal to one if a patent is a transistor patent as defined above and a patent of the Bell system. As control patents, we use all patents that were published in the U.S., matched by publication year, primary three-digit USPC technology class, and the number of citations up to 1952. To adjust for the different number of control patents per treatment patent in each stratum, we use the weights suggested by Iacus et al. [2009]. Column (1) is our baseline specification and uses all citations by other companies as the dependent variable. In columns (2) and (3), the dependent variable is citations by patents in the same field (4-digit IPC) as the patent and in different fields, respectively. Columns (4) to (6) repeat the regressions using the same measures but using the empirical setting of the paper by Watzinger et al. [2020] in which the licensed patents do not cover General Purpose Technologies. In this specification, we drop all transistor-related patents from their sample. In these columns, treated is an indicator variable equal to one if a patent is a patent of the Bell System. The control group consists of all patents that were published in the U.S., matched by publication year, primary three-digit USPC technology class, and the number of citations up to the start of the antitrust case in 1949, as in Watzinger et al. [2020]. To adjust for the different number of control patents per treatment patent in each stratum, we use the weights suggested by Iacus et al. [2009]. The treatment period in these specifications starts in 1956 until patent expiration. "Control Mean" is the mean value of the dependent variable for control group observations in the treatment period. Standard errors are clustered on the primary three-digit USPC technology class level. *, **, *** denote statistical significance on 10%, 5% and 1% level, respectively.

Investigating Mechanisms

Given the variety of activities by Bell to diffuse the transistor technology, we now assess the plausibility of different mechanisms behind the increase in follow-on innovation following the licensing decision. There are two potential explanations. First, common recounts of blocking effects suggest that standardized licensing may benefit subsequent inventors directly (e.g., Galasso and Schankerman [2015]). This would imply that the codified knowledge shared by Bell was sufficient for follow-on innovation and that the transfer of tacit knowledge through the Transistor Symposia was not necessary to produce follow-on invention. One should note that Bell made great efforts to transfer knowledge. For example, Bell published books on the contents of the Transistor Symposia (Bell Telephone Laboratories [1952a,b]) that, anecdotally, were very useful in the transfer of knowledge (to the extent that they collectively became known as 'Ma Bell's cookbook'). Second, the transfer of information via the participation of firms in one of the Transistor Symposia may have been the driving force behind the increased diffusion of the transistor technology. The key argument for this explanation is that while patents disclose useful information (e.g., Furman et al. [2021]), many observers argue that codified knowledge, such as the knowledge disclosed in patents, is insufficient to produce follow-on innovation (e.g., Roin [2005]).

To investigate the relative merits of these two explanations, we study the relevance of information transfer by Bell for our effects using two approaches. First, we investigate whether the original attendees of the Transistor Symposia in 1951 and 1952 show a different response in terms of follow-on invention than non-attendees. Second, we do the same for the original set of licensees of the transistor patents.[15] Both the participation in the symposia and being among the first batch of companies to receive a license might indicate that these companies had preferential access to the tacit knowledge of Bell. We received data on the attendees and the original licensees directly from the AT&T Archives and History Center (AT&T Archives, and History Center [1951, 1952]).[16]

We match these lists to patent assignees in our patent data by hand. We then split the dependent variable by whether the citations came from a firm that was among the attendees of the Transistor Symposia or among the original licensees or not. Table 3 shows the results of this analysis. Column (1) shows our baseline result for comparison. Columns (2) and (3) split citations by whether staff of the citing patent's assignee attended one of the Transistor Symposia in 1951 and 1952. Relative to baseline levels, i.e., the average number of citations in the control group in the post period, the effects are substantially larger for attendees than for non-attendees. Attendees increased their citations to the transistor by around six times the baseline mean in the treatment period.[17] Figure 2 shows that for both groups, the effects only show up after the transistor licensing. Further investigating this, Columns (4) shows that assignees with one of the first licenses of the transistor increased their citations to transistor patents by over nine times the control group mean in the treatment period.[18] In comparison to Column (5), relative to baseline patenting levels, this is a disproportional increase in citations. These results suggest that the information transferred by Bell through the

---

[15] While the set of licensing firms has a strong overlap with the set of participants of the second Transistor Symposium, it also has some overlap with the set of participants of the first Transistor Symposium. Also, there is not a full overlap with the set of participants of the second Transistor Symposium.

[16] We thank Dr. Sheldon Hochheiser for sharing this data with us.

[17] The mean number of citations in the control group in the treatment period ('Control group mean') is 0.108, the differences-in-differences coefficient is 0.659.

[18] The mean number of citations in the control group in the treatment period ('Control group mean') is 0.043, the differences-in-differences coefficient ('Treated x Post') is 0.400.

Table 3: The Effect of Standardized Licensing on Follow-on Innovation by Participation in Transistor Symposia

Dep. Var.:	Citations					
	Baseline	Symposia Attendee		First Round Licensee		Neither
		Yes	No	Yes	No	
	(1)	(2)	(3)	(4)	(5)	(6)
Treated x Post	2.02***	0.66**	1.36***	0.40*	1.62***	1.21***
	(0.55)	(0.27)	(0.47)	(0.22)	(0.48)	(0.42)
Control Mean	1.48	0.11	1.37	0.04	1.44	1.08
Percent Change	136	612	99	925	113	112
# treated	72	72	72	72	72	72
Clusters	30	30	30	30	30	30
Obs.	35629	35629	35629	35629	35629	35629

Notes: This table shows the results from difference-in-differences estimations with Bell Labs transistor patents as treatment groups. We define patents as transistor patents if they were filed by one of the researchers who actively worked towards the development of the transistor at Bell Labs. We define the treatment period as starting in 1953. Treated is an indicator variable equal to one if a patent is a transistor patent as defined above and a patent of the Bell system. As control patents, we use all patents that were published in the U.S., matched by publication year, primary three-digit USPC technology class, and the number of citations up to 1952. Column (1) is our baseline specification and uses all citations by other companies as the dependent variable. In columns (2) and (3), the dependent variable from column (1) is split into citations by assignees that participated or did not participate in one of the two Transistor Symposia, respectively, as evidenced by the lists of attendees (AT&T Archives, and History Center [1951, 1952]). In columns (4) and (5), the dependent variable from column (1) is split into citations by assignees that held one of the original licenses of the transistor or not, as evidenced by the list from the AT&T Archives, and History Center [1982]. Column (6) uses as dependent variable citations by assignees that neither participated in one of the Symposia nor held one of the original licenses. To adjust for the different number of control patents per treatment patent in each stratum, we use the weights suggested by Iacus et al. [2009]. "Control Mean" is the mean value of the dependent variable for control group observations in the treatment period. "Percent Change" is the percent change in citations to the transistor patents for the treatment group, relative to the control mean from the prior sentence. Standard errors are clustered on the primary three-digit USPC technology class level. *, **, *** denote statistical significance on 10%, 5% and 1% level, respectively.

Transistor Symposia helped participating firms in producing follow-on innovation. Note, however, that these firms selected themselves or were selected by Bell into attending and/or being among the original licensees. One should thus expect their follow-on innovation to increase more than the follow-on innovation of other firms. In absolute terms, most of the effect is driven by non-attendees and firms that were not among the first licensees since both sets of firms are small. That is, symposium attendees and early licensees increased their citations to transistor patents much more *per assignee*, but since both sets of assignees are small, this is not driving the large effect we see on subsequent citations to the transistor. In Column (6), we show that citations from firms that neither attended one of the Symposia nor held one of the original licenses increased their patenting significantly. Thus, the information transferred through the Symposia (or being among the first licensees) does not seem to have been a necessary ingredient for follow-on innovation to the transistor. Instead, the codified knowledge transferred through the patents and Bell's transistor books seem to have allowed inventors to leverage the liberal licensing regime by Bell for their follow-on invention (Bell Telephone Laboratories [1952a,b]).[19]

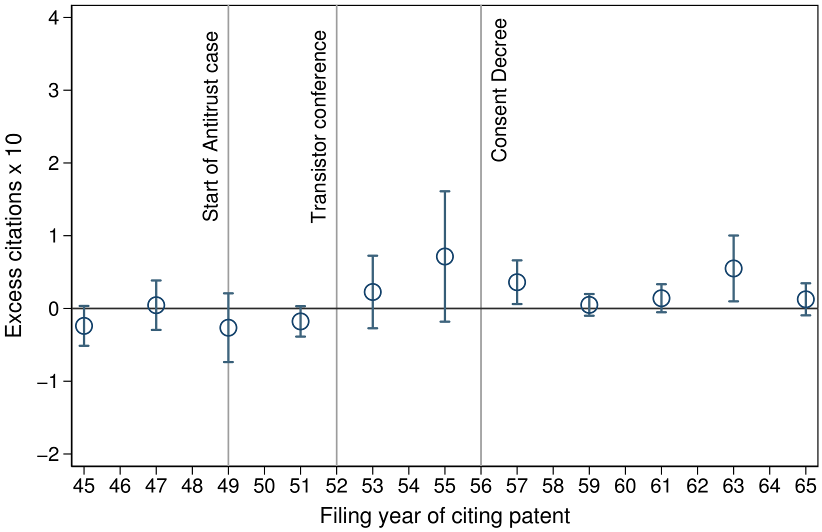
To summarize, while assignees with access to the information from the Symposia increased their patenting more in relative terms, in absolute terms the effect is driven by those firms that did not participate in the Symposia. In our reading of the evidence, both mechanisms therefore seem to have been important.

---

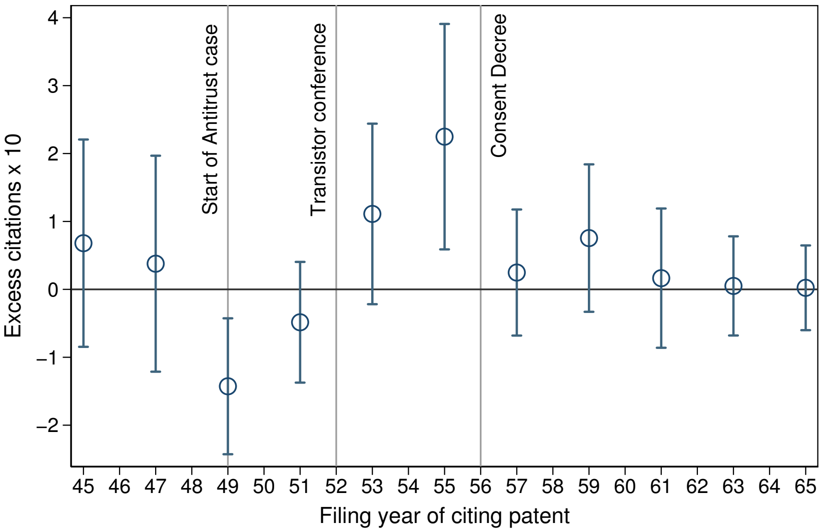
[19] Note, however, that we do not have exact information on further transistor licensees after the first round of licensing. For example, SONY only applied for a license for the transistor in 1953 and was only awarded one in 1954 (Flamm [2010]).

Figure 2: The Impact of Standardized Licensing on Excess Citations to Transistor Patents by Symposia Attendance

(a) Attendees



(b) Non-attendees



Notes: This figure shows the number of two-yearly excess citations to transistor patents published before 1952 relative to patents with the same publication year, in the same three-digit U.S. Patent Classification (USPC) primary class and with the same number of citations up to (and including) 1951, estimated adjusting the specification in equation (1). The figure splits these citations by whether the citing patent is from an assignee that sent staff to one of the Transistor Symposia, identified through attendance lists (Bell Telephone Laboratories [1952a,b]). Panel (a) shows citations by Symposia attendees. Panel (b) shows citations by non-attendees. The blue lines represent the 95% confidence bands calculated from standard errors clustered on the three-digit technology class level. To adjust for the different number of control patents per treatment patent in each stratum, we use the weights suggested by Iacus et al. [2009]. All coefficients are multiplied by 10 for better readability.

Who Benefited From the Licensing?

Historical accounts report an exodus of Bell researchers from Bell Labs in the early 1950s. In 1953, for example, Pete Haggerty from the then small Texas Instruments Inc. convinced Gordon Teal, the inventor of a method to improve transistor performance, to join the company. Similarly, William Shockley, one of the inventors of the transistor, left Bell in 1956 to start Shockley Semiconductors Laboratory. On the one hand, one possible channel is thus that former Bell employees account for many of the new patented inventions following the standardized licensing of the transistor patents. On the other hand, historical accounts on the impacts of the licensing suggest that researchers outside the Bell System who previously did not have the chance to work with the transistor benefited the most.

Table 4 investigates this empirically. In column (1), we replicate our baseline result. Columns (2) and (3) split the dependent variable into citations of companies that we can link to known licensing deals with the Bell system until 1956 through the information of the Antitrust Subcommittee [1958, p.2957]. Relative to baseline citation rates, the effect is substantially larger for inventors with a known license than for those without. Columns (4) and (5) split the citations by the inventor's relationship to Bell. We distinguish between those related to Bell, i.e., Bell employees (those who patented for Bell but are not at Bell anymore) and their first- and second-order co-inventors, and unrelated inventors. The effect is driven by unrelated inventors, suggesting that the standardized licensing was especially important for inventors without connections to the Bell system. The impact of the licensing relative to baseline patenting is slightly larger for young and small assignees than for other inventors (column 6). This indicates that standardized licensing allowed young and small firms to enter the market and develop new technologies building on the groundbreaking invention of the transistor, as suggested by historical accounts such as the story of SONY's pocket transistor radio. In columns (7) and (8) we split the dependent variable by whether the citing patent is in a highly concentrated market or not, using the concordance of Kerr [2008].[20] We find that the increase is driven by citations from patents in markets with low concentration.

Finally, we analyze in Figure 3 whether closeness to telecommunications was a determinant of excess citations. As described in Watzinger et al. [2020], Bell foreclosed the telecommunications market and continued to do so after the 1956 consent decree. This made entry in the field difficult, so we would not expect an effect of standardized licensing in these technologies. We define telecommunications technologies as all patent classes that have a more than 15% likelihood to be used in the production of telecommunications equipment according to the classification of Kerr [2008]. In line with results on the impact of the 1956 consent decree on follow-on innovation, we find that all citations come from patents that are unrelated to telecommunications.[21]

In summary, the effects of Bell's patent licensing and active knowledge transfer seem to have mainly materialized outside the Bell system. The effect stems from unrelated inventors, is large for young and small companies, and stems from unconcentrated markets and markets outside telecommunications.

---

[20] This gives us for each patent class and each industry classified by four-digit SIC code a likelihood that a patent in this class is used in this industry. We multiply this likelihood with the 8-firm market share in an industry that we get from the U.S. Census and aggregate the product on the patent class level. Thus, we get for each patent class the weighted average 8-firm market share in the industry in which the patent is used. In the last step, we classify a citing patent as being used in a highly concentrated industry if the average 8-firm market share is above 60%, which is the 75th percentile.

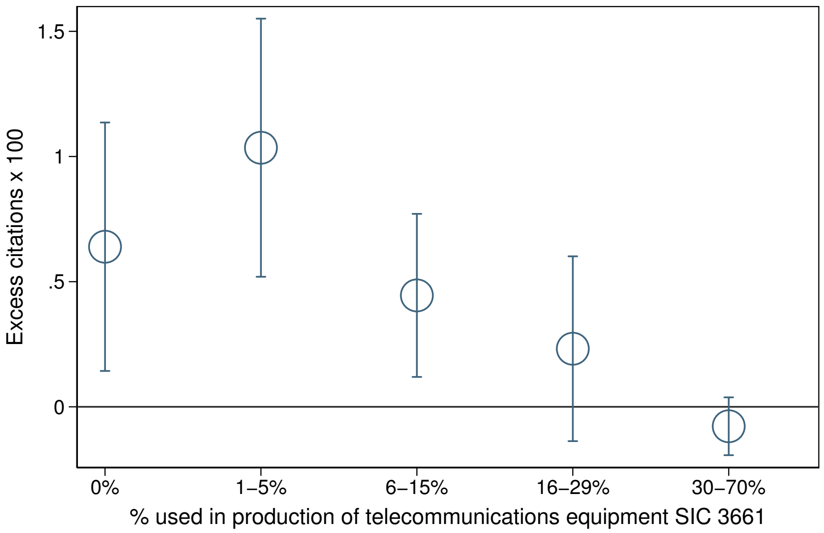
[21] In web appendix C, we also find no time-varying effects on excess citations for transistor patents closely related to the telecommunications industry. Among these patents, this null-result also holds true for young and small assignees.

Table 4: The Effect of Standardized Licensing on Follow-on Innovation by Type of Citing Party

Dep. Var.:	Citations							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Base	Pre-1956 Bell Licensee		Bell		Young &	Concentration	
	line	Yes	No	Related	Unrelated	Small	High	Low
Treated	0.08	0.03	0.05	0.07	0.01	-0.19^{**}	0.95^{*}	-0.87^{*}
	(0.14)	(0.12)	(0.13)	(0.05)	(0.14)	(0.08)	(0.52)	(0.51)
Post	-0.43	0.08	-0.51	-0.08^{**}	-0.35	0.24^{*}	-0.19	-0.24
	(0.43)	(0.15)	(0.35)	(0.04)	(0.42)	(0.13)	(0.39)	(0.20)
Treated x Post	2.02^{***}	0.82^{**}	1.20^{***}	-0.03	2.05^{***}	0.61^{**}	-0.09	2.11^{***}
	(0.55)	(0.37)	(0.37)	(0.07)	(0.55)	(0.28)	(0.38)	(0.61)
Control Mean	1.48	0.27	1.22	0.04	1.44	0.65	0.50	0.98
# treated	72	72	72	72	72	72	72	72
Clusters	30	30	30	30	30	30	30	30
Obs.	35629	35629	35629	35629	35629	35629	35629	35629

Notes: This table shows the results from difference-in-differences estimations with transistor patents. We define patents as transistor patents if they were filed by one of the researchers who actively worked towards the development of the transistor at Bell Labs. As the dependent variable, we use all citations by companies other than the filing company. We define the treatment period as starting in 1953. Treated is an indicator variable equal to one if a patent is a transistor patent as defined above and a patent of the Bell system. As control patents, we use all patents with the same publication year, primary three-digit USPC technology class, and the same number of citations up to 1952 as Bell transistor patents. To adjust for the different number of control patents per treated patent, we use the weights suggested by Iacus et al. [2009]. We repeat our baseline specification in column (1). In columns (2) and (3), we split the citations according to whether the assignee on the citing patent held at least one license for Bell patents before the consent decree 1956 (not necessarily for transistor patents). In columns (4) and (5), we split the dependent variable according to whether the citing patent's inventors are related to Bell, meaning they ever patented for Bell or ever were (first- or second-order) co-authors with Bell inventors, or whether they are unrelated. Column (6) uses citations by young and small companies. We define an assignee as young if its first patent was filed less than ten years before it cited the Bell patent and as small if it had less than ten patents before 1949. In columns (7) and (8), we classify citing patents as belonging to a market with high or low concentration. To this end, we use the concordance of Kerr [2008], which gives us for each patent class and each industry classified by four-digit SIC code a likelihood that a patent in this class is used in this industry. We multiply this likelihood with the average 8-firm market share in an industry that we get from the U.S. Census (Federal Trade Commission [1992]) and aggregate the product on the patent class level. In the last step, we classify a citing patent as being used in a highly concentrated industry if the average 8-firm market share is above 60%, which is the 75th percentile. All coefficients are multiplied by 100 for better readability. "Control Mean" is the mean value of the dependent variable for control group observations in the treatment period. Standard errors are clustered on the primary three-digit USPC technology class level. *, **, *** denote statistical significance on 10%, 5% and 1% level, respectively.

Figure 3: Impacts by Distance to Telecommunications



Note: This figure shows results on follow-on citations by varying likelihood of Bell's transistor patent to be used in telecommunications. Relevance is measured by the likelihood that a patent is used in industry SIC 3661, using the data of Kerr [2008].The figure shows results from the difference-in-differences specification of the licensing on follow-on patent citations by closeness to telecommunications, with 1953 until patent expiration as the treatment period. We report the treatment effect along with 95% confidence intervals separately for citations from patents with differing relevance for the production of telecommunications equipment (SIC 3661 - "Telephone and Telegraph Apparatus"). The bins labeled 0%, 1-5%, 6-15%, 16-29%, 30-70% aggregate citations of 367, 75, 28, 17 and 7 technology classes, respectively.

Robustness: Within-Patent Identification

A potential caveat of our matching approach is that it is inherently difficult to find suitable patents to match an extraordinary invention such as the transistor. To address such concerns, Figure 4 shows results of a different identification strategy that does not depend on matching transistor patents.

In this figure, we compare the average number of citations from a treated group of patents to Bell's transistor patents to citations from various control groups, comparing citations from different groups to the *same patent*. These comparisons thus hold the patent under consideration fixed. We define all citations from non-telecommunications patents as treated, as the transistor as a GPT had a large influence on a wide range of technologies. The red solid line shows the average number of citations from non-telecommunications patents to Bell's transistor patents.

As our first control group, we use patents in telecommunications. The green dashed line shows the number of citations to patents in our matched control group. As described above, Bell foreclosed the telecommunications market and continued to do so after the 1956 consent decree (Watzinger et al. [2020]). Thus, we would not expect follow-on innovation in these areas.

As another control group, we use citations to transistor patents from less affected companies (IBM, RCA, and GE) that had existing cross-licensing agreements with Bell, represented by a solid blue line. If there had been a concurrent technology shock of concern to our identification strategy, we would expect a reaction of these high-tech companies. While there is an increase in citations from these companies, it is by no means comparable to the effect on more affected inventors. Finally, we show self-citations by Bell to Bell's transistor patents , represented by a grey dashed line.

No matter which control group we use, only citations from non-telecommunications to Bell's transistor patents show a strong increase after the Transistor Symposia. In contrast, citations from less affected companies, from markets that continued to be foreclosed, and from Bell itself seem far less affected or unaffected. Patent citations to our matched control group develop similarly to citations to Bell's transistor patents by less affected groups, in line with the identification assumption.

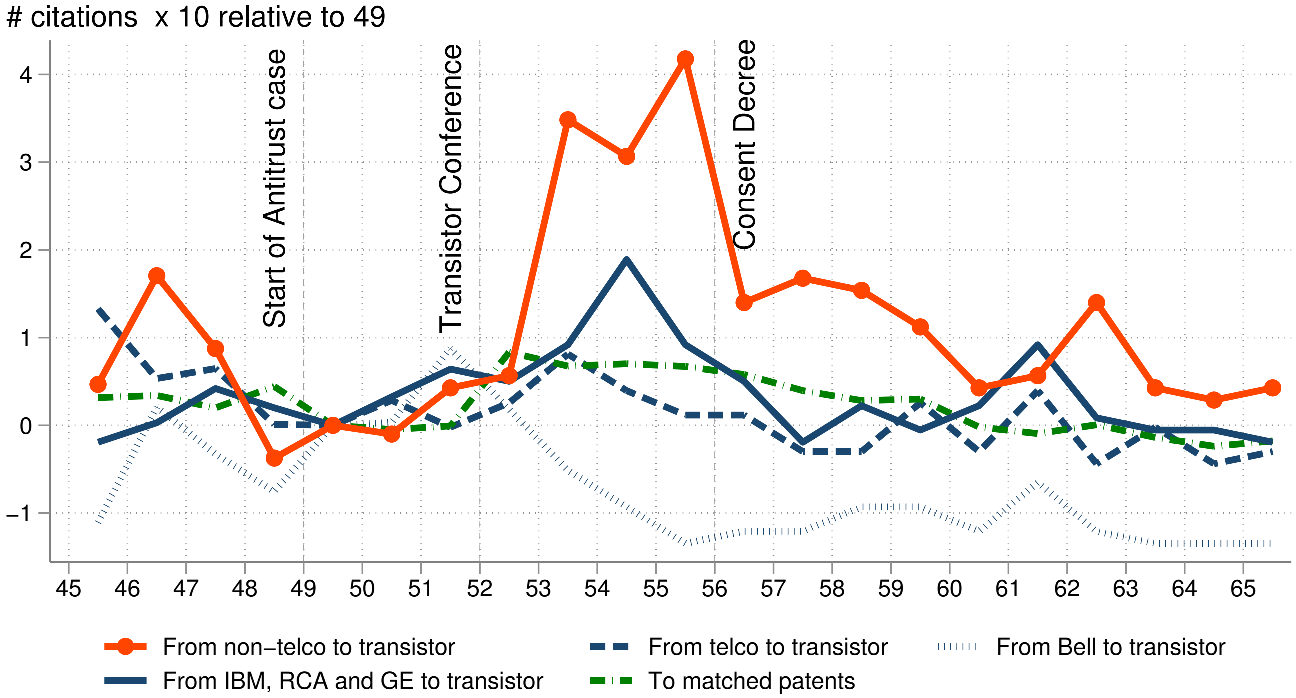
In Table 5, we quantify the results from our within patent analysis. In column (1) of this table, we show our baseline specification for comparison. In the remaining columns, we show results from within-patent analyses using the following specification. :

$$\#Citations_{i,j,t} = \beta_1 \cdot Treated_j + \beta_2 \cdot Post_t + \beta_3 \cdot Treated_j \cdot Post_t + \alpha_i + \varepsilon_{i,j,t} \tag{2}$$

where $\#Citations_{i,j,t}$ is the number of citations that patent i receives from group j in year t and $Treated_j$ is an indicator variable, which takes the value of one for citations by our treatment group, i.e., non-telecommunications patents. The treatment period starts in 1953, as before. We use the full set of non-secret transistor patents in these columns. Note that in this analysis, we do not include matched control patents but compare citations to transistor patents by different groups.

The odd-numbered columns (except column (1)) include patent fixed effects (α_i), while the even-numbered columns do not. Columns (2) and (3) compare citations from patents in the treatment group to citations from patents close to

Figure 4: The Impact of Standardized Licensing on Excess Citations to Transistor Patents: Within Patent Identification



Notes: This figure shows the average number of citations to bell's transistor patents from non-telecommunication patents and to the matched control group as well as from other control groups. The green dash line shows citations to patents that are in the control group in Panel (a). The blue solid line shows citations from IBM, RCA, and GE to Bell's transistor patents. These companies had existing licensing agreements with Bell and were thus affected to a lesser extent. The blue dashed line shows citations from patent classes close to telecommunications, where Bell continued to foreclose the market (Watzinger et al. [2020]). The blue dotted line shows self-citations by Bell. We normalize all time series to their level in 1949, before the start of the antitrust case against the Bell System. The sample under consideration contains 110 transistor patents, 83 of which were not affected by the secrecy program. We can match 72 transistor patents. All coefficients are multiplied by 10 for better readability.

telecommunications, where Bell foreclosed the market (Watzinger et al. [2020]). Columns (4) and (5) compare citations from patents in the treatment group to citations from patents of less affected companies (IBM, RCA, and GE) that had existing cross-licensing agreements with Bell. The final two columns use self-citations as the comparison group. This table again shows that no matter which control group we use, our results are the same qualitatively. In comparison to Table D.3 in the web appendix that only uses transistor patents that are matched in our main approach, using the full set of transistor patents leads to estimating larger treatment effects. This is in line with non-matched transistor patents being more affected by the standardized licensing. We also see higher average citations to these patents, in line with high-quality transistor patents not finding a proper match in our main empirical approach.

Additional Robustness Tests in the Appendix

In web Appendix E, we show that our main effect is not driven by citation substitution, i.e., we do not see decreases in citations to similar, but not licensed technologies outside the Bell System. We also show results for alternative control groups based on IPC instead of CPC. In web Appendix F, we show that our matching is robust to matching transistor patents to control patents with higher citation counts up to 1952 or to control patents that have the same number of citations *after* the licensing of the transistor technology. Our results remain robust.

In web Appendix G, we complement our main empirical analysis and show that the patent licensing and active knowledge transfer of the transistor led to an increase in the number of patents in affected technology subclasses relative to similarly sized subclasses of the same technology class that did not experience the

Table 5: The Effect of Standardized Licensing: Within Patent Approaches

Dep. Var.:	Citations						
Approach:	Baseline	Within Patents					
Treated:	Transistor Patents	Citations from non-telecommunications patents					
Control Group:	Matched	Telecomm. Cit.		Cit. by B3 Comp.		Self-Cites	
Patent FE:	No	No	Yes	No	Yes	No	Yes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treated x Post	2.02***	2.03***	2.13***	1.26***	0.99**	2.93***	2.70***
	(0.55)	(0.68)	(0.72)	(0.45)	(0.43)	(0.78)	(0.81)
Control Mean	1.48	1.47	1.47	1.31	1.31	0.56	0.56
# treated	72	83	83	83	83	83	83
Clusters	30	83	83	83	83	83	83
Obs.	35629	3202	3202	3202	3202	3202	3202

Notes: This table shows the results from difference-in-differences estimations with transistor and Bell Labs patents following equation 2. We define patents as transistor patents if they were filed by one of the researchers who actively worked towards the development of the transistor at Bell Labs. In all columns, we define the treatment period as starting in 1953. In Column (1), we use all patents that were published in the U.S., matched by publication year, primary three-digit USPC technology class, and the number of citations as control patents. To adjust for the different number of control patents per treatment patent in each stratum, we use the weights suggested by Iacus et al. [2009] in column (1). Column (1) is our baseline specification and uses all citations by other companies as the dependent variable. In all remaining columns, the sample only consists of transistor patents and the estimation is within patent. In columns (2) and (3), the treatment group consists of citations by non-telecommunications patents while the control group are citations by telecommunications patents where Bell foreclosed the market (Watzinger et al. [2020]). In columns (4) and (5), the control group are citations by the so-called B-3 companies (IBM, RCA, and GE) that had existing cross-licensing agreements with Bell. In columns (6) and (7), the control group are self-citations. "Control Mean" is the mean value of the dependent variable for control group observations in the treatment period. Standard errors are clustered on the primary three-digit USPC technology class level in column (1) and on the patent level in all other columns. *, **, *** denote statistical significance on 10%, 5% and 1% level, respectively.

licensing of a transistor patent. To avoid a confounding effect, we drop subclasses affected by Bell's 1956 consent decree. In summary, the results of this analysis mirror the results when using our main strategy: The effects are not present in telecommunications and are driven by technology classes with low levels of concentration. And again, the contribution of young and small companies is higher than expected given their share in total patenting.

Finally, in web Appendix H, we show that the patent licensing and active knowledge transfer led to an increase in patents in technology subclasses that cited the transistor but that did not contain transistor patents or other Bell patents themselves. Our control group comprises similarly sized non-citing subclasses within the same technology classes. As in our main result, the number of patents in treated subclasses start to increase relative to the number in untreated subclasses only after the Transistor Conference. Our analysis suggests that the spillovers of the transistor licensing were substantial.

V. Conclusion

Historical accounts suggest that the diffusion of General Purpose Technologies (GPTs) and thus technological progress and economic growth can be hampered by patent protection. The key reason is that improvements in downstream technologies benefit the GPT and vice versa. Since these technologies are rare, most are historical, and because the patents on most GPTs were never revoked, evidence on the role of patents and patent licensing for follow-on innovation in these technologies is difficult to provide.

In this study, we leverage the licensing of the transistor by the Bell Labs in 1952 that came with the transfer of information through the Transistor Symposia and that took place in defense of antitrust lawsuits to investigate the blocking effects of patents for General Purpose Technologies. Our results show that this licensing decision was an important factor in the diffusion of the transistor. In particular, we show that cross-technology spillovers were large. Both information transfer via the Transistor Symposia and the licensing of patents in itself seem to have played important roles in this. Our results suggest that patent licensing in key technologies can induce more market entry since unrelated inventors, as well as young and small firms, particularly benefited from the licensing. These results may inform the current debate about the role of intellectual property rights in the global slowdown of business dynamism (Andrews et al. [2016]; Akcigit and Ates [2019, 2021]).

References

Akcigit, U. and S. Ates (2019): "What Happened to U.S. Business Dynamism?" *NBER Working Paper No. 25756*.

——— (2021): "Ten Facts on Declining Business Dynamism and Lessons from Endogenous Growth Theory," *American Economic Journal: Macroeconomics*, 13, 257–298.

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

Alcacer, J., M. Gittelman, and B. Sampat (2009): "Applicant and Examiner Citations in US Patents: An Overview and Analysis," *Research Policy*, 38, 415–427.

Andrews, D., C. Criscuolo, and P. N. Gal (2016): "The best versus the rest: The global productivity slowdown, divergence across firms, and the role of public policy," *OECD Productivity Working Paper*.

Antitrust Subcommittee (1958): *Consent Decree Program of the Department of Justice. American Telephone & Tel. Subcommittee - Hearings before the Antitrust Subcommittee.*

——— (1959): *Consent Decree Program of the Department of Justice. Committee on the Judiciary, House of Representatives, Eighty-Sixth Congress, First Session.*

Arora, A., S. Belenzon, and A. Patacconi (2017): "The Decline of Science in Corporate R&D," *Strategic Management Journal*, 39, 3–32.

Arora, A. and M. Ceccagnoli (2006): "Patent Protection, Complementary Assets, and Firms' Incentives for Technology Licensing," *Management Science*, 52, 293–308.

AT&T Archives, and History Center (1951): "Transistor Symposium: List of Guests," Item/Box Number 11-04-02-02, AT&T Archives, and History Center, Warren NJ.

——— (1952): "Transistor Technology Symposium," Item/Box Number 11-04-02-03, AT&T Archives, and History Center, Warren NJ.

——— (1982): "Letter to Mr. J. Millenstein," Item/Box Number 11-04-02-01, AT&T Archives, and History Center, Warren NJ.

Bell Telephone Laboratories (1952a): *Transistor Technology*, vol. 1, Princeton, N.J.: Van Nostrand.

——— (1952b): *Transistor Technology*, vol. 2, Princeton, N.J.: Van Nostrand.

Boldrin, M. and D. K. Levine (2008): *Against intellectual monopoly*.

Bresnahan, T. F. (2010): "General Purpose Technologies," in *Handbook of the Economics of Innovation*, ed. by B. H. Hall and N. Rosenberg, Elsevier, vol. 2, chap. 18, 761–791.

Bresnahan, T. F. and M. Trajtenberg (1995): "General purpose technologies 'Engines of growth'?" *Journal of Econometrics*, 5, 83–108.

Buehler, E. (1983): "The first Czochralski silicon," *50 years Progress in Crystal Growth*, 13, 105.

Carrick, R. (1982): "AT&T Likely to Limit License," *InfoWorld*, 32–33.

Federal Trade Commission (1992): "Economic Census: Concentration Ratios for the U.S." .

Field, A. J. (2008): "Does Economic History Need GPTs?" *Available at SSRN: https://ssrn.com/abstract=1275023.*

Flamm, K. (2010): *Mismanaged Trade? Strategic Policy and the Semiconductor Industry*, Brookings Institution Press.

Furman, J. L., M. Nagler, and M. Watzinger (2021): "Disclosure and Subsequent Innovation: Evidence from the Patent Depository Library Program," *American Economic Journal: Economic Policy*, forthcoming.

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. (2012): "Broad Cross-license Negotiations," *Journal of Economics & Management Strategy*, 21, 873–911.

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

Gertner, J. (2012): *The Idea Factory: Bell Labs and the Great Age of American Innovation*, Penguin.

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

Gross, D. P. (2019): "The Consequences of Invention Secrecy: Evidence from the USPTO Patent Secrecy Program in World War II," Working Paper 25545, National Bureau of Economic Research.

Helpman, E. (1998): *General Purpose Technologies and Economic Growth*, MIT Press.

Holbrook, D., W. M. Cohen, D. A. Hounshell, and S. Klepper (2000): "The Nature, Sources, and Consequences of Firm Differences in the Early History of the Semiconductor Industry," *Strategic Management Journal*, 21, 1017–1041.

Iacus, S. M., G. King, and G. Porro (2009): "CEM: Software for Coarsened Exact Matching," *Journal of Statistical Software*, 30, 1–27.

——— (2012): "Causal Inference Without Balance Checking: Coarsened Exact Matching," *Political Analysis*, 20, 1–24.

Jovanovic, B. and P. L. Rousseau (2005): "General Purpose Technologies," in *Handbook of Economic Growth*, ed. by P. Aghion and S. N. Durlauf, Elsevier, vol. 1B, 1181–1224.

Katznelson, R. D. and J. Howells (2012): "Inventing-around Edison's incandescent lamp patent: evidence of patents' role in stimulating downstream development," *Northwestern University Pritzker School of Law Working Paper*.

——— (2015): "The myth of the early aviation patent hold-up-how a US government monopsony commandeered pioneer airplane patents," *Industrial and Corporate Change*, 24, 1–64.

Kerr, W. R. (2008): "Ethnic Scientific Communities and International Technology Diffusion," *The Review of Economics and Statistics*, 90, 518–537.

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

Kremer, M. (1998): "Patent buyouts: A mechanism for encouraging innovation," *The Quarterly Journal of Economics*, 113, 1137–1167.

Lanjouw, J. O. and M. Schankerman (2004): "Protecting Intellectual Property Rights: Are Small Firms Handicapped?" *Journal of Law and Economics*, 47, 45–74.

Lerner, J. (1994): "The Importance of Patent Scope: An Empirical Analysis," *The RAND Journal of Economics*, 25, 319–333.

Lerner, J., M. Strojwas, and J. Tirole (2007): "The design of patent pools: the determinants of licensing rules," *The RAND Journal of Economics*, 38, 610–625.

Levin, R. (1982): "The semiconductor industry," in *Government and Technical Progress: A Cross-Industry Analysis*, ed. by R. R. Nelson, New York: Pergamon Press.

Merges, R. P. and R. R. Nelson (1990): "On the Complex Economics of Patent Scope," *Columbia Law Review*, 90, 839–916.

——— (1994): "On limiting or encouraging rivalry in technical progress: The effect of patent scope decisions," *Journal of Economic Behavior and Organization*, 25, 1–24.

Mokyr, J. (1994): "Technological Change,1700-1810," in *An Economic History of Britain since 1700*, ed. by R. Floud and D. N. McCloskey, Cambridge University Press, vol. 1, 12–43.

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

Mowery, D. C. (2011): "Federal Policy and the Development of Semiconductors, Computer Hardware, and Computer Software: A Policy Model for Climate Change R&D?" in *Accelerating Energy Innovation: Insights from Multiple Sectors*, University of Chicago Press, 159–188.

Murray, F. and S. Stern (2007): "Do Formal Intellectual Property Rights Hinder the Free Flow of Scientific Knowledge?: An Empirical Test of the Anti-commons Hypothesis," *Journal of Economic Behavior & Organization*, 63, 648–687.

Nathan, J. (2001): *Sony: The Private Life*, Harpercollins.

Nelson, R. (1962): "The Link between Science and Invention: The Case of the Transistor," in *The Rate and Direction of Inventive Activity: Economic and Social Factors*, Princeton University Press, 549–584.

Nelson, R. R. (1959): "The Simple Economics of Basic Scientific Research," *Journal of Political Economy*, 67, 297–306.

Ouellette, L. L. (2012): "Do patents disclose useful information?" *Harvard Journal of Law & Technology*, 25, 546–602.

Reid, T. R. (2001): *The Chip: How Two Americans Invented the Microchip and Launched a Revolution*, Random House Trade Paperbacks.

Roin, B. (2005): "The disclosure function of the patent system (or lack thereof)," *Harvard Law Review*, 118, 2007–2028.

Rosenberg, N. (1990): "Why do Firms do Basic Research (with their own Money)?" *Research Policy*, 19, 165–174.

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.

SELGIN, G. AND J. L. TURNER (2011): "Strong Steam, Weak Patents, or the Myth of Watt's Innovation-Blocking Monopoly, Exploded," *Journal of Law and Economics*, 54, 841–861.

TEMIN, P. AND L. GALAMBOS (1987): *The Fall of the Bell System: A Study in Prices and Politics*, Cambridge University Press.

WATZINGER, M., T. 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 (2021): "The Break-Up of the Bell System," *Mimeo*.

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

——— (2015): "Intellectual Property Rights and Innovation: Evidence from Health Care Markets," in *Innovation Policy and the Economy, Volume 16*, ed. by J. Lerner and S. Stern, University of Chicago Press, 15, 53–87.

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

A Appendix

A. The Bell System

As described in the main text, American Telephone & Telegraph (AT&T) was the dominant provider of telecommunications services in the U.S. in the early 1950s.[22] Its operating companies bought more than 90% of their equipment from Western Electric, AT&T's manufacturing subsidiary. Western Electric produced telecommunications equipment based on the research done by the Bell Laboratories, the research subsidiary of AT&T and Western Electric. All these companies together were known as the Bell System, stressing its vertical integration.

The Bell System was also an innovation powerhouse. Its Bell Labs were unique in their commitment to basic research. When the Bell Labs were founded in 1925, no one knew which part of science might yield insights into the problems of electric communication [Rosenberg, 1990; Nelson, 1962, p.31]. As a result, the Bell System decided that - besides supporting the day-to-day need of the System - the Bell Labs would engage in basic science, assuming it would eventually yield products for some part of the large Bell System [Gertner, 2012; Nelson, 1959; Arora et al., 2017]. According to the first head of basic and applied research at Bell Labs, Harold Arnold, his department would include "the field of physical and organical chemistry, of metallurgy, of magnetism, of electrical conduction, of radiation, of electronics, of acoustics, of phonetics, of optics, of mathematics, of mechanics, and even of physiology, of psychology and meteorology." This broad focus led to major advances in basic science, but also to a large number of unused patents. For example, an investigation of the FCC in 1934 reported that Bell owned or controlled 9,255 patents, but actively used only 4,225 patent covered inventions

---

[22] This section is largely based on Watzinger et al. [2020].

[Antitrust Subcommittee, 1958, p.3842]. The 1950 staff of Bell Labs alone consisted of four future Nobel Laureates in physics, one Turing Award winner, five future U.S. National Medals of Science recipients, and ten future IEEE Medals of Honor recipients.

B. Robustness to alternative transistor definitions

An exact definition of transistor patents is unfortunately unavailable [AT&T Archives, and History Center, 1982]. We therefore take a set of core patents around the transistor technology from historical accounts (see `http://www.patents4technologies.com/assetsp4t/textsp4t/PioneeringPatents.htm`, last accesses 17 May 2021). Beyond the two key patents #2,524,035 ("Three-electrode circuit element utilizing semiconductive materials") by Bardeen and Brittain and #2,569,347 ("Circuit element utilizing semiconductive material") by Shockley mentioned in the text, this list also contains the Shockley patents #2,623,102 ("Circuit element utilizing semiconductive materials"), #2,666,818 ("Transistor amplifier"), #2,672,528 ("Semiconductor translating device"), and #2,744,970 ("Semiconductor signal translating devices"). This list is certainly too short as we know from historical accounts that many patents were necessary to use the core transistor technology and that came with a license.

Starting from these baseline patents, we use several alternative definitions of the transistor. First, we determine transistor patents text-based by predicting their likelihood of being transistor-related through the words contained in their title and abstract. For our classification algorithm, we leverage the words used in the very narrow set of original transistor patents cited above. We use all words from the abstracts and titles of these patents and their counts as predictors for whether any given patent is a transistor patent. We then define a patent as transistor-based if the patent has a probability of more than one percent, returning 80 transistor patents. Our results are robust to alternative plausible cutoffs. The idea is that patents that use similar words as the original transistor patents are likely to be transistor patents themselves (or at least necessary to use the transistor technology).

Second, we use co-citation patterns, following the idea that Bell patents that were co-cited with the original transistor patents are likely to be necessary to use the transistor technology. We again use the above list of core transistor patents and determine which patents cite them. We then classify as transistor patents all other patents assigned to Bell companies that are cited along with the narrow set of core transistor patents mentioned above. This method yields 38 transistor patents.

Third, we use a class-based citation approach to define transistor patents. To this end, we work backwards from the main transistor technology classes, namely USPC classes 257 ("Active Solid-State Devices"), 326 ("Electronic Digital Logic Circuitry"), and 438 ("Semiconductor Device Manufacturing"). We then define transistor patents as those patents cited by patents in these technologies. This method generates 199 transistor patents. The idea is that Bell patents that were frequently cited by subsequent transistor patents are likely to be necessary to use the transistor technology.

We show the results of this analysis in Table A.1. The first three columns repeat our baseline result from Table II. In columns (4) to (6), we show the results from the same analysis using our text-based definition of the transistor. The results are qualitatively identical to the results presented in the first three columns. While the effects are somewhat smaller, they show an even larger relative effect on citations in technology classes that differ from the one of the underlying patent, reinforcing our conclusion that GPTs may have particularly strong spillovers effects on other technologies. In columns (7) to (9), we use our co-citation approach. The results are similar again. While the effect in Column (8) is statistically not different form zero on the 10% level, the p-value is .101. In columns (10) to (12), we then use our class-based citations to define treated transistor patents. Again, the results are very similar to the results using our baseline definition of the transistor.

C. Effects by distance to telecommunications

In Figure A.1 we distinguish the effects of standardized licensing for patents with a different likelihood of being used in the production of telecommunications equipment further. In Figure 3 in the main text, we found a negative relation between the closeness to telecommunications and excess citations, in line with results on the impact of the 1956 consent decree on follow-on innovation [Watzinger et al., 2020]. All excess citations come from patents that have no or little relation to telecommunications. In Panels (a) and (b) of Figure A.1, we show that these effects again only show up after the Transistor Conference.

D. Within Patent Comparison: Transistor Patents from Main Analysis

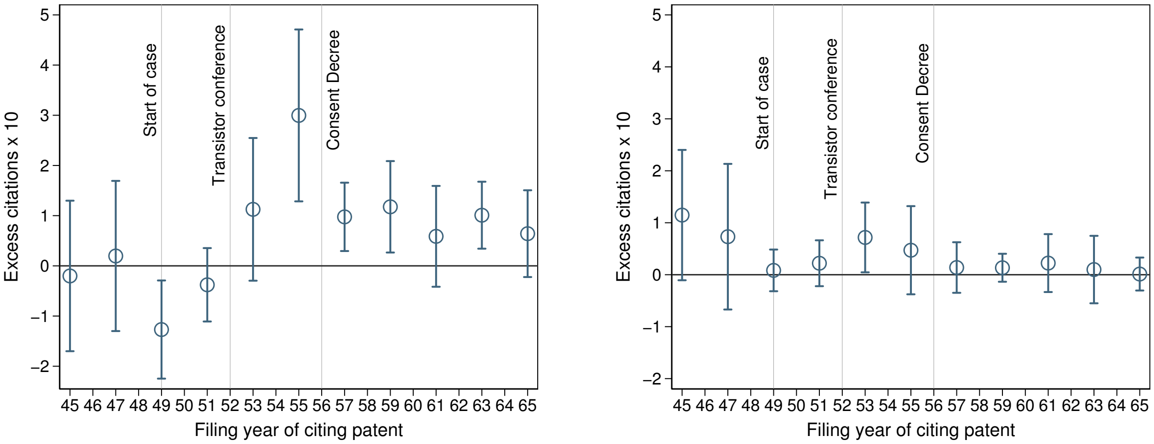
In Table A.3, we show results from our within patent analysis using the set of non-secret transistor patents from our main analysis. The table is analogous to Table V in the main text. For comparability, we however keep the transistor patents constant to our baseline specification. In column (1) of this table, we show our baseline specification. In the remainder, we show results from within-patent analyses. We only use transistor patents as our sample. The odd-numbered columns

Table A.1: The Effect of Standardized Licensing on Follow-on Innovation using Alternative Transistor Definitions

Dep. Var.:	Citations											
Definition	Baseline			Text-based			Co-cited			Cited by transistor class		
	Overall	Same Tech.	Diff. Tech.	Overall	Same Tech.	Diff. Tech.	Overall	Same Tech.	Diff. Tech.	Overall	Same Tech.	Diff. Tech.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Treated x Post	2.02***	1.14***	0.88**	1.06**	0.36*	0.74*	3.46***	1.17	2.29**	2.06***	0.95***	1.11***
	(0.55)	(0.32)	(0.37)	(0.52)	(0.21)	(0.42)	(0.87)	(0.66)	(0.86)	(0.31)	(0.16)	(0.33)
Control Mean	1.48	0.73	0.75	1.73	0.63	0.86	1.83	0.74	1.10	1.58	0.68	0.90
# treated	72	72	72	80	80	80	38	38	38	199	199	199
Clusters	30	30	30	45	45	45	14	14	14	63	63	63
Obs.	35629	35629	35629	50637	50637	50637	18341	18341	18341	114013	114013	114013

Notes: This table shows the results from difference-in-differences estimations with Bell Labs transistor patents as treatment groups. We define the treatment period as starting in 1953. In these columns, treated is an indicator variable equal to one if a patent is a transistor patent as defined below and a patent of the Bell system. In the first three columns, we define patents as transistor patents if they were filed by a member of the original transistor team. Column (1) is our baseline specification and uses all citations by other companies as the dependent variable. In columns (2) and (3), the dependent variable is citations by patents in the same field (4-digit IPC) as the patent and in different fields, respectively. Columns (4) to (6) repeat the regressions using a text-based definition of transistor patents as described in the text. Columns (7) to (9) use all Bell patents as treated patents that were cited along with the core transistor patents. Columns (10) to (12) use all Bell patents cited by other patents in transistor technology classes as treated patents. As control patents, we use all patents that were published in the U.S., matched by publication year, primary three-digit USPC technology class, and the number of citations up to 1952 as the respective treated patents. To adjust for the different number of control patents per treatment patent in each stratum, we use the weights suggested by Iacus et al. [2009] in all columns. "Control Mean" is the mean value of the dependent variable for control group observations in the treatment period. Standard errors are clustered on the primary three-digit USPC technology class level. *, **, *** denote statistical significance on 10%, 5% and 1% level, respectively.

Figure A.1: The Impacts of the Consent Decree on Follow-on Innovation by Closeness to Telecommunications



(a) Time-Varying Impacts: Fields outside Telecommunications (<15%)

(b) Time-Varying Impacts: Telecommunications (>15%)

Notes: This figure shows results on follow-on citations by varying likelihood of Bell's transistor patent to be used in telecommunications. Relevance is measured by the likelihood that a patent is used in industry SIC 3661, using the data of Kerr [2008]. . Figure (a) shows the average number of excess citations from telecommunications patents over a two-year period of Bell's transistor patents ("Bell patents") relative to their control patents. Figure (b) shows the average number of excess citations from patents in other fields over a two-year period of patents affected by the consent decree ("Bell patents") relative to their control patents. We classify a patent as a telecommunications patent if it has more than a 15% likelihood to be used in the production of telecommunications equipment (SIC 3661) according to the data of Kerr [2008]. In all panels, the blue lines represent the 95% confidence bands calculated from standard errors clustered on the three-digit technology class level. All coefficients are multiplied by 10 for better readability.

Table A.2: The Effect of Standardized Licensing on Subsequent Citations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Baseline	Telecommunications		Other Fields						
		All	Young & Small	All	Dollar weighted	Young & Small	High Concentration	Low Concentration	Y&S High Concentration	Y&S Low Concentration
Bell	0.08	0.50	0.06	-0.42	0.43	-0.25***	-0.07	-0.35	-0.07**	-0.17**
	(0.14)	(0.31)	(0.06)	(0.37)	(0.40)	(0.09)	(0.28)	(0.24)	(0.03)	(0.07)
Post	-0.43	0.07	0.11	-0.50	2.95***	0.13	-0.26	-0.24	0.06	0.07
	(0.43)	(0.23)	(0.09)	(0.31)	(1.03)	(0.11)	(0.21)	(0.20)	(0.04)	(0.13)
Bell x Post	2.02***	-0.07	-0.08	2.09***	7.24***	0.68**	0.34	1.75***	0.06	0.62**
	(0.55)	(0.19)	(0.07)	(0.55)	(1.55)	(0.28)	(0.31)	(0.48)	(0.06)	(0.25)
Control Mean	1.48	0.10	0.03	1.38	7.27	0.62	0.39	0.99	0.18	0.44
# treated	72	72	72	72	72	72	72	72	72	72
Clusters	30	30	30	30	30	30	30	30	30	30
Obs.	35629	35629	35629	35629	35629	35629	35629	35629	35629	35629

Notes: This table shows the results from difference-in-differences estimations with transistor patents. We define patents as transistor patents if they were filed by a member of the original transistor team. As the dependent variable, we use all citations by companies other than the filing company. We define the treatment period as starting in 1953. Bell is an indicator variable equal to one if a patent is a transistor patent as defined above and a patent of the Bell system. As control patents, we use all patents that were published in the U.S., matched by publication year, primary three-digit USPC technology class, and the number of citations. We use all patents with a publication year before 1952, and we match all citations up to and including 1951. To adjust for the different number of control patents per treatment patent in each stratum, we use the weights suggested by Iacus et al. [2009]. Column (1) is our baseline specification. In columns (2) and (3), the dependent variable is citations by patents in fields related to telecommunications. We classify a citing patent as a telecommunications-related patent if in its patent class patents have more than a 15% likelihood of being used in the production of telecommunications equipment, using the data of Kerr [2008]. Column (3) shows results for this dependent variable using the citations by young and small companies. We define an assignee as young if its first patent was filed less than ten years before it cited the Bell patent and as small if it had less than ten patents before 1949. All remaining columns show results for citations from patents in all other fields. Column (4) repeats the baseline specification for Other Fields. In column (5), we weight each citation by the average dollar value of a patent in the same publication year and technology class derived from the values provided by Kogan et al. [2017]. In column (6), we use citations by young and small companies (as defined above) as the dependent variable. In columns (7) and (8), we classify citing patents as belonging to a market with high or low concentration. To this end, we use the concordance of Kerr [2008], which gives us for each patent class and each industry classified by four-digit SIC code a likelihood that a patent in this class is used in this industry. We multiply this likelihood with the average 8-firm market share in an industry that we get from the U.S. Census [Federal Trade Commission, 1992] and aggregate the product on the patent class level. In the last step, we classify a citing patent as being used in a highly concentrated industry if the average 8-firm market share is above 60%, which is the 75th percentile. Columns (8) and (9) repeat this split using patents by young and small inventors (as defined before) as the dependent variable. All coefficients are multiplied by 10 for better readability. Standard errors are clustered on the primary three-digit USPC technology class level. *, **, *** denote statistical significance on 10%, 5% and 1% level, respectively.

include patent fixed effects, while the even-numbered columns do not. The treatment group are citations from non-telecommunications patents. Columns (2) and (3) compare these to citations from patents close to telecommunications, where Bell foreclosed the market [Watzinger et al., 2020]. Columns (4) and (5) compare them to citations by less affected companies (IBM, RCA, and GE) that had cross-licensing agreements with Bell. The final two columns use self-citations as the comparison. The table again shows that no matter which control group we use, our results are the same qualitatively.

E. Pseudo Treatment: Citation Substitution is Small

One potential concern might be that our estimates do not capture an increase in follow-on innovation, but merely reflect a substitution effect. Due to the free availability of Bell technology, companies might have substituted away from other, potentially more expensive technologies. To assess this, we exploit the fact that a patent's technology is classified twice: once in the USPC system, which has a technical focus, and once in the IPC system, which reflects more closely the intended industry or profession ("usage") [Lerner, 1994]. In columns (2) and (3) of Table A.4, we assign a pseudo-treatment to all patents that have the same USPC class and the same IPC class as the Bell patents. As control group, we use in column (2) patents with the same USPC, but a different IPC classification as Bell patents. In column (3), we use as a control group patents with the same IPC, but a different USPC classification as Bell patents. Thus, we compare patents that are arguably more similar to the Bell patents to two different control groups. We find a small and statistically insignificant effect. Again, this speaks in favor of limited citation substitution or - alternatively - a homogeneous citation substitution to all

Table A.3: The Effect of Standardized Licensing on Follow-on Innovation: Within Patent Identification

Dep. Var.:	Citations						
Approach:	Baseline	Within Patents					
Treated:	Transistor Patents	Citations from non-telecommunications patents					
Control Group:	Matched	Telecomm. Cit.		Cit. by B3 Comp.		Self-Cites	
Patent FE:	No	No	Yes	No	Yes	No	Yes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treated x Post	2.02***	1.38***	1.22***	1.00***	0.71**	2.04***	1.91***
	(0.55)	(0.43)	(0.39)	(0.34)	(0.30)	(0.56)	(0.61)
Control Mean	1.48	0.58	0.58	0.59	0.59	0.26	0.26
# treated	72	72	72	72	72	72	72
Clusters	30	72	72	72	72	72	72
Obs.	35629	2836	2836	2836	2836	2836	2836

Notes: This table shows the results from difference-in-differences estimations with transistor and Bell Labs patents. We define patents as transistor patents if they were filed by a member of the original transistor team. In all columns, we define the treatment period as starting in 1953. Treated is an indicator variable equal to one if a patent is a transistor patent as defined above and a patent of the Bell system. As control patents, we use all patents that were published in the U.S., matched by publication year, primary three-digit USPC technology class, and the number of citations in column (1). To adjust for the different number of control patents per treatment patent in each stratum, we use the weights suggested by Iacus et al. [2009] in column (1). Column (1) is our baseline specification and uses all citations by other companies as the dependent variable. In all remaining columns, the sample only consists of transistor patents and the estimation is within patent. We restrict the analysis to those transistor patents that are part of our main analysis sample. In columns (2) and (3), the treatment group consists of citations by non-telecommunications patents while the control group are citations by telecommunications patents where Bell foreclosed the market [Watzinger et al., 2020]. In columns (4) and (5), the control group are citations by the so-called B-3 companies (IBM, RCA, and GE) that had existing cross-licensing agreements with Bell. In columns (6) and (7), the control group are self-citations. "Control Mean" is the mean value of the dependent variable for control group observations in the treatment period. Standard errors are clustered on the primary three-digit USPC technology class level in column (1) and on the patent level in all other columns. *, **, *** denote statistical significance on 10%, 5% and 1% level, respectively.

Table A.4: Auxiliary Regressions

	(1)	(2)	(3)	(4)	(5)	(6)
		Pseudo Treatment		Diff. Control Group		
	Baseline	Control: Same USPC diff IPC	Control: Same IPC diff USPC	Control: Same IPC diff USPC	Control: Same IPC	Loose
Treatment	0.08	-0.08	0.04	0.63	0.29	0.57
	(0.14)	(0.06)	(0.15)	(1.24)	(0.55)	(0.35)
Post	-0.43	0.15	0.21^{***}	-0.85^{***}	-0.64^{*}	-0.51
	(0.43)	(0.23)	(0.06)	(0.24)	(0.34)	(0.40)
Treat x Post	2.02^{***}	0.20	0.06	3.04^{***}	2.60^{***}	2.70^{***}
	(0.55)	(0.17)	(0.23)	(0.65)	(0.72)	(0.68)
Control Mean	1.48	1.39	1.51	1.54	1.43	1.48
# treated	72	1465	1511	81	77	74
Clusters	30	23	154	164	105	30
Obs.	35629	32668	208998	204104	41413	58146

Notes: This table shows the results from difference-in-differences estimations with transistor patents. We define patents as transistor patents if they were filed by a member of the original transistor team. As the dependent variable, we use all citations by companies other than the filing company. We define the treatment period as starting in 1953. Bell is an indicator variable equal to one if a patent is a transistor patent as defined above and a patent of the Bell system. As control patents, we use all patents that were published in the U.S., matched by publication year, primary three-digit USPC technology class, and the number of citations. We use all patents with a publication year before 1952, and we match all citations up to and including 1951. To adjust for the different number of control patents per treatment patent in each stratum, we use the weights suggested by Iacus et al. [2009]. In columns (2) and (3), we assign pseudo treatments. In column (2), we assign all patents that have the same USPC and different 3-digit IPC technology class than transistor patents of Bell Labs as treated, and in column (3), we assign patents with the same IPC and different USPC classification than transistor patents of Bell Labs as treated. In column (4), we use as controls patents in the same IPC 3 class but in a different USPC class than the Bell patents. In column (5), we use as controls patents with the same 4-digit IPC class as the Bell patents. In column (6), we coarsen the publication year to two-year windows and sort all pre-citations into ten equally sized bins to match a larger number of patents. All coefficients are multiplied by 10 for better readability. The data are from the Worldwide Patent Statistical Database (PATSTAT) of the European Patent Office. All coefficients are multiplied by 10 for better readability. "Control Mean" is the mean value of the dependent variable for control group observations in the treatment period. Standard errors are clustered at the three-digit USPC technology class level. *, **, *** denote statistical significance on 10%, 5% and 1% level, respectively.

control groups.

F. Robustness: Using Different Matching Procedures

In columns (5) - (7) of Table A.4, we report results from using several alternative matching variables. In the main specification, we use the age (measured by the publication year), the technology (measured by USPC class) and the quality of a patent (measured by the number of citations up to 1952). In column (4), we use patents in the same IPC but different USPC class instead of using those in the same USPC class. In column (5), we match on the IPC classification, independent of the USPC class. Finally, in column (6), we do a coarsened exact matching in order to match all Bell patents.[23] In all three cases, the size of the effects is similar to the one in the main specification.

In Table A.5, we show the corresponding regression results using a matching on higher citation counts. Column (1) repeats our baseline specification. In columns (2) and (3), we match Bell's transistor patents to control patents that have one and two more citations up to 1952, respectively. Matching on higher citation counts increases the magnitude of the measured effect relative to the baseline. In column (4), we match Bell patents to patents having the same citation rates *after* the consent decree in 1956. If, in contrast to our identification assumption, Bell patents indeed have a higher counterfactual citation trend than the matched patents with the same number of pre-citations, then it seems plausible that these Bell patents should have the same citation trend before the consent decree as control patents with the same number of post-citations. Yet, this is not the case.

---

[23] Coarsened exact matching was proposed by Iacus et al. [2012]. In this specification, we match on one of five publication-year categories that contain two years each and one of ten prior-citation categories.

Similar to our main result, the size of the estimated effect is around two using this alternative matching strategy.

G. Complementary Empirical Strategy: Impact on Number of Patents in Affected Technology Subclasses

To complement our main analysis using patent citations, in this section we additionally analyze the impact of standardized licensing on the number of innovations. More specifically, we compare the change in the total number of patents in a USPC technology subclass with a licensed transistor patent to the change in the total number of patents in subclasses without before and after the standardized licensing. To do this, we employ the regression model:

$$\#Patents_{s,c,t} = \beta \cdot Treat_s \cdot I[1953 - 1970] + YearFE_t + SubclassFE_s + \varepsilon_{s,c,t} \tag{3}$$

where the dependent variable $\#Patents_{s,c,t}$ is the number of non-Bell patents in subclass s in class c in year t. $Treat_s$ is an indicator function that is equal to one if there is at least one transistor patent by Bell in subclass s, and Post is an indicator function for the years 1953 to 1970. β measures the number of excess patents in treated relative to untreated classes. We control for technology subclass fixed effects to account for permanent differences in patenting rates between technology subclasses and for year fixed effects to account for developments in patenting activity in the U.S. that are common across technology subclasses. Because transistor classes patent at a higher level than control classes but are few in nature, we additionally match on the size of the subclass as measured by the

Table A.5: The Effect of Standardized Licensing on Subsequent Citations using Different Matching Procedures

	(1)	(2)	(3)	(4)
Matching:	Baseline	Higher Citations		Post-1952
		+1	+2	
Bell	0.08	-1.46***	-3.09***	-0.85**
	(0.14)	(0.17)	(0.27)	(0.40)
Post	-0.43	-1.78***	-2.96***	-0.37
	(0.43)	(0.44)	(0.47)	(0.53)
Bell x Post	2.02***	2.87***	3.89***	0.85**
	(0.55)	(0.52)	(0.50)	(0.40)
Control Mean	1.48	1.88	2.24	2.25
# treated	72	69	67	64
Clusters	30	30	30	28
Obs.	35629	17825	9734	25065

Notes: This table shows the results from difference-in-differences estimations with transistor patents. We define patents as transistor patents if they were filed by a member of the original transistor team. As the dependent variable, we use all citations by companies other than the filing company. We define the treatment period as starting in 1953. Bell is an indicator variable equal to one if a patent is transistor patent of the Bell system. As control patents, we use all patents with the same publication year, primary three-digit USPC technology class, and the same number of citations up to 1952 as Bell patents in column (1). In column (2), we use the same procedure, but add one citation to the number of citations up to 1952 for Bell patents. In column (3), we again use the same procedure, but add two citations to the number of citations up to 1952 for Bell patents. In columns (2) and (3), we add an additional fixed effect for the period 1950 to 1952 for Bell patents to the estimation equation to account for mechanical changes in citation rates due to the different matching before 1950. In column (4), we use the same procedure, but use the number of citations *after* the transistor licensing in 1952 instead of the number of citations up to 1949. To adjust for the different number of control patents per treatment patent, we use the weights suggested by Iacus et al. [2009]. As dependent variable, we use all citations by companies other than the filing company. Column (1) is our baseline specification. All coefficients are multiplied by 10 for better readability. "Control Mean" is the mean value of the dependent variable for control group observations in the treatment period. Standard errors are clustered on the three-digit USPC technology class level. *, **, and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

number of patents between 1945 and 1951. We use the weights suggested by Iacus et al. [2009] to account for the different number of control subclasses per treated subclass. The standard errors allow for clustering on the patent class level.

We can interpret the estimates from this specification as causal if, in the absence of the consent decree, treated and untreated classes would have developed the same in terms of patenting rates (parallel trends assumption). One potential concern could be that these technology classes would have also grown in the absence of the standardized licensing agreement. This assumption is untestable and may not be met since the transistor technology is one of the most important general-purpose technologies of the post-World War II period and thus likely exhibits very different trends than other technology classes without transistor patents. Below, we however find that the trends of treated and untreated classes are parallel up to the beginning of standardized licensing.[24]

This empirical strategy based on patent counts complements our main empirical strategy based on patent citations in several ways. First, it does not rely on matching Bell patents to other patents. Thus, there is no concern that results might be driven by inventors who might have strategically under-cited Bell's transistor patents before the consent decree. Second, we do not need to worry about potential substitution between similar patents, e.g. due to the salience of Bell patents after the 1956 consent decree, since we estimate the net growth in the number of patents. Lastly, citations capture only the immediate impact of standardized licensing; i.e., first-round effects on follow-on innovations citing Bell patents. Patent counts might give us a more comprehensive picture as they also include second-round effects. However, we acknowledge that the identifying as-

---

[24] We also drop subclasses that had less than five patents from 1945 to 1952 to avoid changes in the composition of classes over time.

sumption behind this analysis is stronger than for our main analysis.

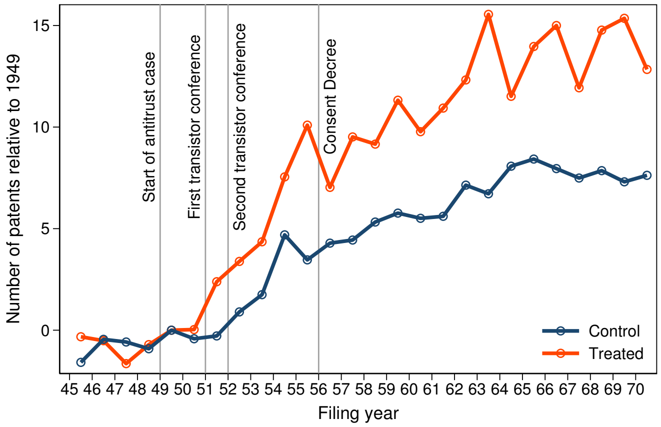
Panel (a) of Figure A.2 shows the comparison of mean patenting rates across affected and unaffected technology subclasses relative to patenting in 1950. For this figure, we drop all technology classes that do not contain any transistor patent. We also drop all subclasses that do not contain a transistor patent, but contain another Bell patent that was part of the 1956 consent decree. As becomes clear, up to the first transistor conference in 1951, subclasses with and without transistor patents developed similarly in terms of patenting. However, in subclasses with licensed transistor patents, there is a strong increase in patenting after the transistor conferences, especially after the second conference in 1952. After the consent decree in 1956, there is a more steady increase in the number of publications in these technology classes. Thus, standardized licensing seems to have affected the timing of the increase in new patenting in affected technology classes.

Panel (b) uses time-varying regressions on the subclass level to analyze this result further. Relative to 1949, there was no difference between affected and unaffected subclasses up until 1952. Following the transistor conference, however, technology classes with transistor patents grew. Again, the 1956 consent decree does not seem to have changed the course of these technologies. While the two-yearly treatment effects become more noisy, the point estimate is barely affected.

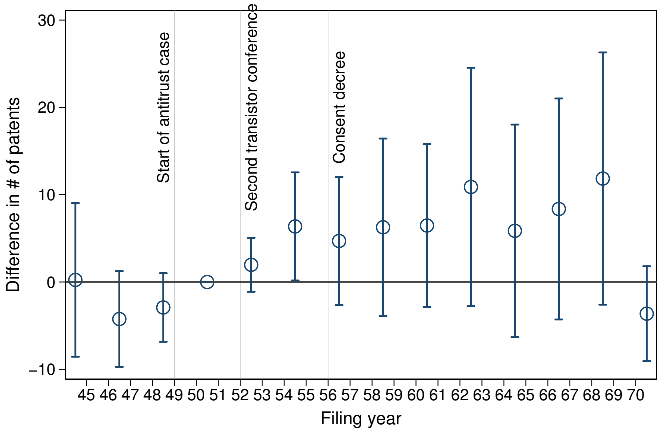
We quantify these results in Table A.6. Following our results in Appendix C., we split our results by distance to telecommunications. For this analysis, Column (1) gives the baseline estimate using this identification strategy for fields not related to telecommunications. It shows that on average, patent subclasses with transistor patents filed around 4.5 patents more per year than subclasses without such patents, showing the strong increase of over 30% in innovation after

Figure A.2: The Effect of Standardized Licensing on the Number of Patents in Transistor Classes

(a) Change in Total Patenting



(b) Annual Treatment Effects on the Number of Patent Applications



Notes: These graphs show the impacts of the licensing on the number of patents in subclasses with a Bell transistor patent relative to subclasses without such patents. We drop patent classes without any transistor patents for the analysis. We also drop patent subclasses without transistor patents that were directly affected by the 1956 consent decree. We only consider patents from subclasses outside telecommunications. We classify a subclass as telecommunications-related if in its patent class patents have more than a 15% likelihood of being used in the production of telecommunications equipment, using the data of Kerr [2008]. Finally, because transistor classes patent at a higher level than control classes but are few in nature, we additionally match on the size of the subclass as measured by the number of patents between 1945 and 1951. We use the weights suggested by Iacus et al. [2009] to account for the different number of control subclasses per treated subclass. We define patents as transistor patents if they were filed by a member of the original transistor team. Panel (a) shows the impact on the total number of yearly patents in affected relative to unaffected technology subclasses relative to 1949. Panel (b) shows the two-yearly average difference in total patenting between affected and unaffected classes adapting equation (3), along with 95% confidence bands. Standard errors allow for clustering within three-digit technology classes.

the transistor conference. Columns (2) and (3) split technologies by market concentration using data from the Federal Trade Commission [1992]. The effects are again driven by markets with low concentration levels, although the effects are measured with substantial noise. Column (4) again shows that young and small inventors contributed disproportionally to the increase in patenting in transistor technology classes, in line with historical accounts. Around 75% of the effect is due to patenting by these inventors, who only account for around 31% of baseline patenting. The remaining 25% of the effect stem from other inventors. Note that the increase in patenting does not apply to all technology fields: again, in subclasses close to telecommunications, we do not observe an increase in patenting after the standardized licensing of the transistor patents (column 6). This casts doubt on the suspicion that the rise in patenting observed for other technology classes after the licensing of the transistor is mechanical.

H. Investigating Spillovers: The Impact on Number of Patents in Technology Subclasses Citing the Transistor

To investigate the spillover effects of the licensing of the transistor patents, we now investigate its impacts on subclasses that cited transistor patents, but that did not contain transistor patents themselves. To this end, we adopt equation 3. The treatment group now contains subclasses that cite transistor patents at some point in time. The control group contains subclasses that do not but that are in the same technology class.[25] We drop all subclasses that contain transistor patents or patents that were affected by Bell's 1956 consent decree to avoid direct impacts of Bell's licensing on our results. Figure A.3a shows mean patenting rates in treated

---

[25] We again additionally match on the size of the subclass as measured through the number of patents between 1945 and 1951.

Table A.6: Treatment Effects on Patent Applications per Class and Year

	(1)	(2)	(3)	(4)	(5)	(6)
	Fields outside Telecommunications					Telecommunications
	All	High Concentration	Low Concentration	Young & Small	Others	All
Treated x I(52-70)	4.46*	1.93	6.06	3.36**	1.10	1.36
	(2.52)	(1.68)	(4.13)	(1.59)	(1.26)	(8.34)
Mean Dep.	6.72	9.31	6.12	2.79	3.93	3.62
N Cluster	22	9	13	22	22	6
Observations	7640	3012	4628	7640	7640	2488

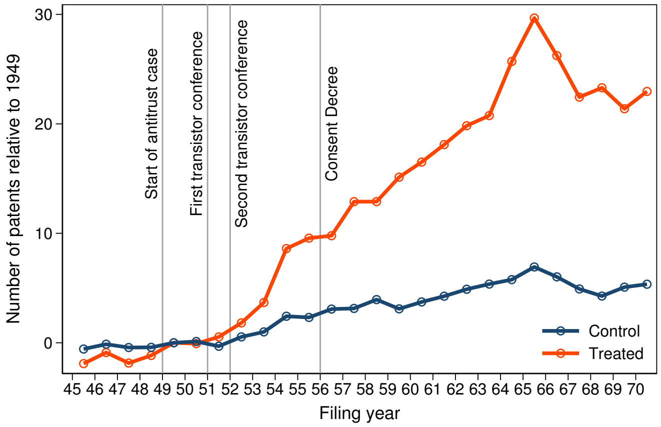
Notes: This table shows the results from estimating equation (3). The dependent variable is the total number of patent applications per year that are either in a treated or untreated USPC technology subclass. A subclass is in the treatment group if it contains at least one Bell transistor patent. Because transistor classes patent at a higher level than control classes but are few in nature, we additionally match on the size of the subclass as measured by the number of patents between 1945 and 1951. We use the weights suggested by Iacus et al. [2009] to account for the different number of control subclasses per treated subclass. We define patents as transistor patents if they were filed by a member of the original transistor team. The treatment variable is interacted with an indicator that is equal to one for the period after 1953 to 1970. Column (1) shows the baseline estimates. Column (2) through (5) restrict the dependent variable to patents outside telecommunications-related subclasses. We classify a patent as a telecommunications-related if in its patent class patents have more than a 15% likelihood of being used in the production of telecommunications equipment, using the data of Kerr [2008]. Columns (2) and (3) classify patents as belonging to a market with high or low concentration. For this classification, we use again the concordance of Kerr [2008] that gives us for each patent class and each industry classified by four-digit SIC code a likelihood that a patent in this class is used in this industry. We multiply this likelihood with the 8-firm market share in an industry that we get from the U.S. Census [Federal Trade Commission, 1992] and aggregate the product on the patent class level. In the last step, we classify a patent as being used in a highly concentrated industry if the 8-firm market share is above 60%, which is the 75th percentile. Column (4) uses patents from young and small assignees; i.e., assignees whose first patent was granted less than ten years ago and who had less than ten patents in 1949. Column (5) uses patents from all other assignees. Column (6) restricts the dependent variable to patents from telecommunications-related subclasses as defined before. The regressions include subclass and year fixed effects. "Control Mean" is the mean value of the dependent variable for control group observations in the treatment period. Standard errors are clustered on the class level. *, **, *** denote statistical significance at the 10%, 5%, and 1% level, respectively.

vs. untreated subclasses, relative to 1949. Up until the first transistor conference, there is no clear difference between treatment and control group. Starting after the first transistor conference and especially after the second transistor conference, subclasses that cite transistor patents grow substantially more than subclasses in the same technology classes that do not. Figure A.3b analyzes this pattern using two-yearly treatment effects. Again, there is no significant difference and no different pre-trend between patenting in the treatment and the control group up until the first and second transistor conferences (combined in the 1952 estimate). Starting with the transistor conferences, the subclasses that relied on transistor patents as inputs grew substantially faster than those who did not.

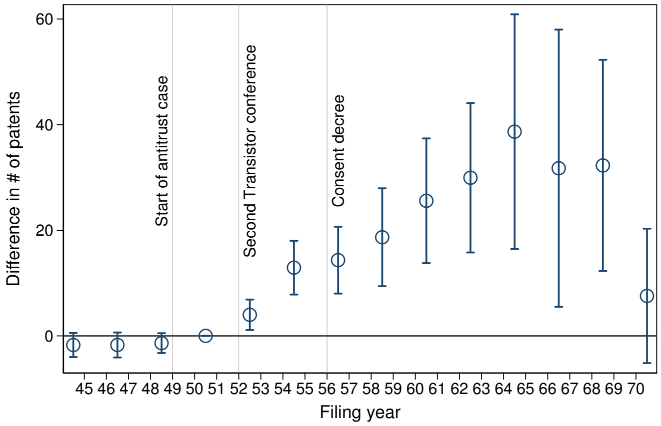
We quantify these results in table A.7, estimating a model analogous to equation 1. Following our results in Appendix C., we again split our results by distance to telecommunications. We classify a patent class as telecommunications-related if its patents have more than a 15% likelihood of being used in the production of telecommunications equipment, using the data of Kerr [2008]. Column (1) shows that subclasses that cite transistor patents (but do not contain transistor patents themselves) see around 11 more patent applications per year than those that do not cite transistors, but are in the same technology class. This is a more than 100% increase in patenting relative to the mean. Columns (2) and (3) again split technologies by market concentration using data from the Federal Trade Commission [1992]. The results show that the effects are similarly large in markets with high and low concentration, in contrast to our citation results. Columns (4) and (5) show that the effect stems both from young and small and from other assignees. However, the effects on patenting for young and small companies are substantially higher relative to the baseline. Finally, Column (6) shows that in telecommunications, the overall effect is smaller than the effect outside telecom-

Figure A.3: The Effect of Standardized Licensing on the Number of Patents in Subclasses Citing the Transistor

(a) Change in Total Patenting



(b) Annual Treatment Effects on the Number of Patent Applications



Notes: These graphs show the impacts of the licensing on the number of patents in subclasses that cited a Bell transistor patent relative to subclasses that did not. We drop patent classes that include transistor patents for the analysis. We define patents as transistor patents if they were filed by a member of the original transistor team. We also drop patent subclasses without transistor patents that were directly affected by the 1956 consent decree. We only consider patents from subclasses outside telecommunications. We classify a subclass as telecommunications-related if in its patent class patents have more than a 15% likelihood of being used in the production of telecommunications equipment, using the data of Kerr [2008]. Panel (a) shows the impact on the total number of yearly patents in affected relative to unaffected technology subclasses relative to 1949. Panel (b) shows the two-yearly average difference in total patenting between affected and unaffected classes adapting equation (3), along with 95% confidence bands. Standard errors allow for clustering within three-digit technology classes.

munications. It is also statistically insignificantly different from zero.

Table A.7: Treatment Effects on Patent Applications per Class and Year in Subclasses Citing the Transistor

	(1)	(2)	(3)	(4)	(5)	(6)
	Fields outside Telecommunications					Telecommunications
	All	High	Low	Young & Small	Others	All
		Concentration				
Treated x I(52-70)	10.70***	10.18***	11.02***	4.54***	6.15***	4.25
	(1.88)	(2.30)	(2.67)	(0.96)	(1.10)	(3.08)
Mean Dep.	8.45	10.74	7.93	3.55	4.91	4.00
N Cluster	61	26	35	61	61	11
Observations	12849	3435	9414	12849	12849	2024

Notes: This table shows the results from estimating equation (3). The dependent variable is the total number of patent applications per year that are either in a treated or untreated USPC technology subclass. A subclass is in the treatment group if it *cited* at least one Bell transistor patent. Because classes citing transistor patents patent at a higher level than control classes but are fewer, we additionally match on the size of the subclass as measured by the number of patents between 1945 and 1951. We use the weights suggested by Iacus et al. [2009] to account for the different number of control subclasses per treated subclass. We define patents as transistor patents if they were filed by a member of the original transistor team. The treatment variable is interacted with an indicator that is equal to one for the period after 1953 to 1970. Column (1) shows the baseline estimates. Column (2) through (5) restrict the dependent variable to patents outside telecommunications-related subclasses. We classify a patent as a telecommunications-related if in its patent class patents have more than a 15% likelihood of being used in the production of telecommunications equipment, using the data of Kerr [2008]. Columns (2) and (3) classify patents as belonging to a market with high or low concentration. For this classification, we use again the concordance of Kerr [2008] that gives us for each patent class and each industry classified by four-digit SIC code a likelihood that a patent in this class is used in this industry. We multiply this likelihood with the 8-firm market share in an industry that we get from the U.S. Census [Federal Trade Commission, 1992] and aggregate the product on the patent class level. In the last step, we classify a patent as being used in a highly concentrated industry if the 8-firm market share is above 60%, which is the 75th percentile. Column (4) uses patents from young and small assignees; i.e., assignees whose first patent was granted less than ten years ago and who had less than ten patents in 1949. Column (5) uses patents from all other assignees. Column (6) restricts the dependent variable to patents from telecommunications-related subclasses as defined before. The regressions include subclass and year fixed effects. "Control Mean" is the mean value of the dependent variable for control group observations in the treatment period. Standard errors are clustered on the class level. *, **, *** denote statistical significance at the 10%, 5%, and 1% level, respectively.