Disclosure and Subsequent Innovation: Evidence from the Patent Depository Library Program

By Jeffrey L. Furman, Markus Nagler, and Martin Watzinger*

How important is access to patent documents for subsequent innovation? We examine the expansion of the USPTO Patent Library system after 1975. Patent libraries provided access to patents before the Internet. We find that after patent library opening, local patenting increases by 8-20% relative to similar regions. Additional analyses suggest that disclosure of technical information drives this effect: inventors increasingly take up ideas from outside their region and the effect is strongest in technologies where patents are more informative. We thus provide evidence that disclosure plays an important role in cumulative innovation.

"Patent law requires disclosure for the same reason that innovators dislike it: it is the vehicle by which technical knowledge is passed from the patenting firm to its competitors."
Scotchmer (1991, p.39)

I. Introduction

The disclosure of technical information is one of the patent system's central economic functions (e.g., Machlup and Penrose 1950; Scotchmer and Green 1990; Scotchmer 1991; Romer 1990). In legal debate, the U.S. Supreme Court has labeled disclosure the, "quid pro quo of the right to exclude."[1] Ideally, patent disclosure will facilitate follow-on innovation by transmitting useful knowledge and by avoiding unnecessary duplication of investment in innovation. In practice,

* Jeff Furman: Boston University Questrom School of Business and NBER, 595 Commonwealth Ave.-#653a, Boston, MA 02215, USA; Markus Nagler: FAU Erlangen-Nuremberg, Lange Gasse 20, 90403 Nuremberg, Germany; Martin Watzinger: LMU Munich, Department of Economics, Akademiestrasse 1, 80799 Munich, Germany; Furman (corresponding author): furman@bu.edu. We thank Jonathan Kolstad and three anonymous referees for very helpful suggestions. We also thank Sharon Belenzon, Iain Cockburn, Ben Jones, Abhishek Nagaraj, Ariel Dora Stern, Scott Stern, Toby Stuart, Rosemarie Ziedonis, Bo Zhao, and participants at the NBER-AIEA Hong Kong, NBER-PIE program meeting, NBER Productivity Seminar, REER Atlanta, TILEC Tilburg, and UT-Austin Empirical Patent Law conferences, and numerous seminars for helpful comments and suggestions. We thank Regina Seibel for excellent research assistance. We also thank Robert Berry, Christopher Dillon, Eileen Fischlschweiger, and Larayne Dallas for their helpful and patient replies to all our questions. Markus Nagler and Martin Watzinger gratefully acknowledge financial support from the DFG through CRC TR 190. Nagler gratefully acknowledges financial support by the Elite Network of Bavaria through Evidence-Based Economics and by the DAAD through a scholarship for doctoral students. Jeff Furman gratefully acknowledges financial support from NSF SciSIP grant, SES-1564368. Parts of this paper were written while Markus Nagler was visiting MIT Economics and while Martin Watzinger was visiting the Questrom School of Business at Boston University. We thank both departments for their hospitality.

[1]Kewanee Oil Co. v. Bicron Corp., 416 U.S. 470 (1974).

however, scholars have expressed doubts about whether patent disclosure is, indeed, effective in fostering cumulative innovation (e.g. Roin, 2005; Lemley, 2012). For example, legal scholars argue that strategic behavior in crafting patent documents can make it difficult for follow-on inventors to extract key information from prior art searches and that many inventors do not even read patents because this increases the legal risk of "willful infringement." As a result, such researchers are doubtful that, in practice, patent disclosure transmits truly valuable information to potential future inventors.

Understanding the extent to which information disclosure plays a role in inducing subsequent innovation is of first order importance for the design of the patent system and for the design of levers to increase cumulative innovation. Empirical evidence on this question is scarce, however, because of a fundamental challenge for causal analysis: The patent system makes the right to exclude competitors dependent on disclosing technical information. This leaves little variation to measure the "enablement effect" of disclosure, i.e., the value of information provision on subsequent innovation separately from the effects of exclusion (Graham and Hegde, 2015; Hegde and Luo, 2017; Williams, 2017; Gross, 2019; Baruffaldi and Simeth, 2018).

In this paper, we analyze the large scale expansion of the USPTO Patent and Trademark Depository Library (PDL or patent libraries) system from 1975 to 1997 to investigate the effect of access to patent information on regional innovation. Before the Internet, such patent libraries were the only places outside of USPTO headquarters in Washington DC that provided public access to the full range of technical information available on patent documents and that provided tools to search for prior art. Opening a patent depository library in a particular region may, therefore, have reduced by a substantial degree the costs local inventors had to bear to access prior art. While exclusion rights remain national (and constant across region), the opening of patent libraries yields effective variation across regions in the extent of 'patent disclosure,' i.e., the costs of accessing the information contained in patents, during the pre-Internet era.

With the aim of information diffusion in mind, the patent library system was founded in the 1870s to provide patents and innovation-related resources for inventors, entrepreneurs, and incumbent firms. By 1975, 20 libraries had been established, primarily in New England and East of the Mississippi. Beginning in 1975, the USPTO embarked on an effort to open at least one patent library in each of the U.S. states to increase the percentage of U.S. citizens with a patent collection in their commuting zone. This goal was achieved in 1997. We focus our analysis on this period of library system expansion. Although we refer to patent library "opening," establishing Patent Depository Libraries did not require the construction of new facilities. Instead, opening required that existing libraries dedicated sufficient space, staff, and resources for patent library materials and received official designation as a USPTO Patent Depository Library.

To estimate the impact of opening up a patent library on regional innovation, we compare the change in the number of ultimately-granted patents (by year of application) filed in the region proximate to the newly-opened library after opening with the change in the number of patents in the regions around a matched control sample of Federal Depository Libraries (FDLs). Federal Depository Libraries make government documents, such as laws and Acts of Congress, freely available to the public. As the missions of patent libraries and FDLs are similar, i.e., providing the public with official documents, nearly all patent libraries are also Federal Depository Libraries. According to one librarian we inquired with, "a factor that would influence a library in becoming a patent library is whether they had been involved with government documents in another capacity." Patent libraries typically served initially as FDLs and only later became patent libraries, making FDLs in the same state a natural control group.

FDLs are a control group that will provide a valid counterfactual if, in the absence of the opening of the patent library, the number of patents in the region would have followed the same trend as those of the regions around the control libraries. One potential concern about this identification approach is that in the period from 1975 to 1997, libraries could apply to become Patent Depository Libraries. As a consequence, our results could be spurious if the local librarian or the USPTO based their decisions regarding library opening on expected future patenting. While we cannot exclude this threat completely, we document that this is only a concern if the librarian submitting the application (or the USPTO) is able to correctly predict the exact year and place of an increase in future patenting in technologies where patent disclosure is important.

In our main specification, we find that the number of patents within 15 miles of newly-opened patent libraries increased by between 8 and 20% relative to baseline patenting rates. In our preferred estimate, we find that local patenting increases by 18%, an average of around 3.2 patents per 100,000 persons per year. We do not find a negative effect on patent quality, which suggests that the additional patents induced by PDL opening are not of lesser economic value than those produced prior to library opening. This effect of library opening is, however, localized and becomes substantially weaker outside of 50 miles, a plausible commuting distance. Our results suggest that easier access to prior art increases patenting and contributes to like-minded inventors building on each others' ideas.

We demonstrate that it is unlikely that concurrent shocks drive these effects. In the years before library opening, the number of patents per capita are similar in the regions around the control and to-be-treated libraries. This is consistent with the parallel trends assumption of differences-in-differences analyses. There is also no differential trend between control libraries, suggesting that the libraries do not simply relocate innovative activities from nearby regions. We also show that our results are robust to alternative specifications, for example using regions of future, not-yet-opened, patent libraries as an alternative control group.

In additional analyses, we find three pieces of evidence consistent with the idea that improved access to patented technical information is the most likely mechanism driving the effect. First, the effect is most pronounced in chemical technologies. This is consistent with prior survey research that documents that patents report valuable and specific knowledge for these technologies and that such information is, indeed, read by follow-on innovators (Mansfield, 1986; Levin et al., 1987; Cohen, Nelson and Walsh, 2000; Gambardella, Harhoff and Nagaoka, 2011). Second, the increase in patenting is mainly driven by patents using words that are new to the region around the library, but not to U.S. patenting overall. This effect also shows up only after patent library opening and is consistent with knowledge transfer as playing a key role as a mechanism underlying our findings. By contrast, there is no significant increase in the number of patents that do not include new words. Third, we find that new and existing inventors and inventor teams are similarly impacted by the opening of patent libraries, which speaks against the alternative explanation that the results arise from libraries facilitating matching between like-minded inventors.

Our study demonstrates that access to patent documents contributes to subsequent innovation. We therefore contribute to the discussion on the benefits of patent disclosure and the merits of the patent system as a whole (Williams, 2017). Our evidence is consistent with the argument that a patent, "serves to disseminate technological information, and that this accelerates the growth of productivity in the economy" (Machlup, 1958, p.76).[2] Not only can disclosure facilitate the market for ideas (Hegde and Luo, 2017), but we find that it can actually help inventors in producing new ideas. By finding particularly strong effects in chemical innovations but weaker in other areas, a technological area in which disclosure is thought to be particularly effective, our study also offers qualified support for critics that argue that the usefulness of disclosure through patents for subsequent innovation is limited (e.g., Roin, 2005; Lemley, 2012; Chien, 2016) and that the benefits of reading patents are mixed (Arora, Ceccagnoli and Cohen, 2008; Gambardella, Harhoff and Nagaoka, 2011; Cohen et al., 2002; Hall and Harhoff, 2012; Ouellette, 2012, 2017).[3]

A small but emerging literature is beginning to find effects in line with our results. Gross (2019) studies the USPTO's secrecy program during World War II and finds that patents that were under secrecy order for longer were less likely to be cited than patents that were under secrecy for shorter time periods. In addition, words appearing in chemical patents under secrecy were less likely to be included in chemical product catalogs, showing that secrecy matters for product commercialization.[4] Hegde, Herkenhoff and Zhu (2019) use the introduction of the

[2]In a similar vein, Romer (1990, p.84) writes that patent disclosure increases economic growth because, "other inventors are free to spend time studying the patent application for the widget and learn knowledge that helps in the design of a widget."

[3]Newer studies on the American Inventor Protection Act show that many inventors voluntarily disclose their inventions, leading to earlier licensing deals (Graham and Hegde, 2015).

[4]De Rassenfosse, Pellegrino and Raiteri (2019) examine the impact of a secrecy order during the period 1982-2002. They find that patents cited by such 'secret patents' receive fewer citations during the period of secrecy enforcement, which suggests that information disclosure plays a role in the ability to build on knowledge generated by patented innovations.

American Inventor's Protection Act in 2011 to show that after earlier disclosure of patent applications, patents in the U.S. are cited faster and more often. Our results suggest that by enhancing the ease with which potential inventors can access and deploy that information, policies could increase innovation, though we should take care to note the important general equilibrium caveat that such policies might also depress R&D incentives.

More generally, our study contributes to the literature on research-enhancing institutions by showing that investments in patent libraries helped to fuel regional innovation. Research enhancing institutions lower the cost of access to useful knowledge and thus help to foster geographical and intertemporal spillovers on which economic growth is based (Mokyr, 2002). For example, Furman and Stern (2011) demonstrate that biological resource centers, libraries of living organisms, can foster follow-on innovation by providing open and low cost access to life sciences research materials. In recent work, Biasi and Moser (2016) show that reducing the access costs to science books during World War I increased scientific output particularly in those regions in which libraries bought these books. Berkes and Nencka (2019) demonstrate the opening of the Carnegie Libraries, the very first public libraries in many towns, increased patenting. Andrews (2019) documents the impact of universities on local innovation, finding that the establishment of a sample of U.S. universities between 1839-1954 for which 'winner' and 'runner-up' locations could be identified resulted in a boost in patenting of 45 percent among winner locations relative to runner-up locations. Having access to the actual runners-up is information that we unfortunately lack in our analysis. His result is however primarily driven by increases in population density and does not seem to be driven by knowledge transfers. Our research contributes to this literature by showing that patent libraries increased innovation across U.S. states by improving access to patent documents. In addition, through our various additional analyses, we provide evidence that access to prior patents is the most plausible mechanism.

Historical analyses of the U.S. patent system have noted its role in democratizing innovation, i.e., in enabling innovation to take root in various geographic regions, across socioeconomic groups, and among different types of enterprises (Machlup, 1958; Scotchmer and Green, 1990; Lamoreaux and Sokoloff, 1999; Khan, 2005). Our results provide evidence that, through its 1975-1997 Patent Depository Library program, the USPTO continued to play a role in spreading innovation across regions. In an age before the Internet, searching patent documents at a close-by patent library provided an accessible way to study such prior art.[5]

The remainder of this paper is organized as follows: Section II describes the U.S. Patent Depository Library Program. In Section III we describe the data

[5]For example, avid patent library user and hearing aid innovator Geoffrey Ball stresses the importance of technical information in prior patents to his own work in his autobiography, in which he lauded the Sunnyvale CA patent library as the "only place to research patents" available to him (Ball, 2012).

and the empirical strategy. In Section IV, we show that opening a patent library increased innovation in its close vicinity and present robustness checks. In Section V, we present evidence on the underlying mechanism. Section VI concludes.

II. The U.S. Patent Depository Library Program[6]

The Establishment of the Patent Deposit Library ProgramIn the years following the Civil War, the U.S. Congress acknowledged that the increasingly industrial and innovation-focused country could benefit from expanding access to technical information contained in patent documents. In the early 1870s, Congress therefore enabled the creation of a nationwide network of Patent and Trademark Depository Libraries. Prior to 1871, official patent documents were housed and available for widespread perusal only at the Patent Office in Washington DC. In that year, federal statute 35 USC 12 officially entitled the Patent Office to distribute copies of patents to designated libraries outside the capital. In addition, the Patent Office began in 1872 to publish and disseminate weekly the *Official Gazette*, which reported a brief abstract and a representative drawing of each patented invention.[7] According to Lamoreaux and Sokoloff (1999), patent agents and solicitors emerged in the 1800s as an important institution that inventors and firms outside of Washington could use in order to obtain information about and build upon new inventions. The patent depository libraries were another key institution aiming to support innovation via information diffusion.

The first set of patent depository libraries were established in the 1870s at The New York State Library, the Boston Public Library, The Public Library of Cincinnati and Hamilton County, the Science and Engineering Library at Ohio State University, the Detroit Public Library, the Los Angeles Public Library, the New York Public Library, and The St. Louis Public Library. These locations were chosen because of their potential for innovation and their demand for information about patented inventions. New libraries were slowly added over the next few decades and, by 1975, the number of patent libraries had grown to twenty, most of which were located in the industrial Midwest and eastern seaboard.[8] By the 1970s, each library received weekly shipments of unbound paper patents, the *Official Gazette* of the U.S. Patent and Trademark Office, and two search indices.

[6]This section draws upon historical reviews in Sneed (1998) and Jenda (2005).

[7]The *Gazette* was one of a number of publications that provided limited information about patented inventions. For example, the journal *Chemical Abstracts* began publishing abstracts of chemical patents in 1907 and the periodical *Scientific American*, which began publication in 1845, featured patent summaries throughout its history. Like the *Gazette*, however, each of these sources published only patent abstracts and up to one drawing and, did not, therefore, provide the rich source of technical information available in original patent documents or in patent depository libraries.

[8]There is less information documenting the reasons for the establishment of the small number of libraries (six) opened between 1902 and 1975. Around 11% of ultimately-granted utility patents applied for in US in 1975 derive from regions that already had patent libraries in 1975, which is a substantial share of total patenting.

The Expansion of USPTO Patent Library System 1975-1997Because access to patent documents remained limited to paper-based methods, individual inventors and small and medium-sized enterprises engaged with patent prior art in the early 1970s in much the same way as they had in the 1870s, i.e., through intermediaries, including patent agents, or via travel to locations with complete patent records. Recognizing the need for expanded facilities after one hundred years of relative inactivity, the USPTO began an aggressive expansion of the patent library system beginning in 1975 to increase the percentage of the US population within commuting distance of a patent collection.[9] The revived program established goals of increasing the number of patent libraries by at least three per year and, ultimately, operating at least one patent library in each state.[10] This latter aim was achieved in 1997.

The map in Figure 1a identifies the twenty libraries in operation before 1975, while Figure 1b lists all patent libraries opened after 1975. Figure 2 shows the expansion of the patent depository library system over time.[11] Currently, about half of the PDLs are based in academic libraries and nearly as many are affiliated with public libraries.[12] After 1997, the patent library system adopted a new goal of controlled growth in areas with high population combined and high patent and trademark activity (Sneed, 2000).

How (and why) Libraries Joined the Patent Depository Library Program Beginning in 1975, existing library facilities became eligible to apply to become PDLs if they fulfilled a set of requirements. First, libraries had to demonstrate that they had the physical capacity (space) to acquire and make available for use a collection of *all* U.S. utility patents issued in the twenty years prior to the date of library designation (i.e., *all* patents in force). Second, each patent library had to commit to employing and training sufficient staff to assist the public in the search for prior art. To ensure adequate training, each patent library had to send a representative to the annual PDL Training Seminar in Washington DC.[13] Third, they had to provide free access and a collection of search tools for the public. According to the USPTO, the first library in each state that (a) applied for PDL status and (b) successfully fulfilled these criteria would receive a designation as a patent depository library.

These criteria implied that larger libraries, such as university libraries and city public libraries, were able to fulfill the resource requirements of becoming a patent

[9]This effort was initiated by USPTO Assistant Commissioner William I. Merkin, beginning with an assessment of the patent library system in 1974.

[10]The goals were stated in the testimony before Congress by the Commissioner of Patents and Trademarks, Gerald J. Mossinghoff in 1983 (Mossinghoff, 1984).

[11]Table A-1 and Table A-2 in the online appendix list patent libraries up to 2002.

[12]Since 1871, six PDLs have withdrawn for various reasons, including library closing, no funding for the back file, and a change in institutional priority creating a lack of ability to perform required services.

[13]Indeed, several of the librarians we interviewed mentioned that the opportunity to participate in the annual training was a nontrivial reason for their association with the Patent Deposit Library program.

depository library.[14] The fact that the process of becoming a patent library during the 1975-1997 period was initiated by the library itself rather than solicited by the USPTO, may explain, in part, why patent libraries were not opened in the sequence one might expect *a priori*. For example, Honolulu HI and Big Rapids MI each received patent libraries before either New Haven CT or San Francisco CA (Figure 2). The librarians that we interviewed reported a number of reasons that their institutions applied to join the PDL community during this period. Some librarians we interviewed suggested that their libraries applied to join as a result of their institutional missions or out of a sense of duty to their patrons. These factors may reflect local demand for patent information. However, interviewees also mentioned factors more idiosyncratic and less predictable in driving their library's participation, including the perceived attractiveness of annual PDL trainings in Washington DC and the professional benefits of participating in the PDL librarian community.[15]

Most patent libraries had prior experience handling government documents as Federal Depository Libraries before applying to become patent depository libraries. Federal Depository Libraries make U.S. federal government publications available to the public at no cost. As of 2008, there were 1,252 Federal Depository Libraries, at least two in each of the 435 U.S. Congressional Districts.[16] Because of this structure and the requirements associated with serving as in either library program, we believe that Federal Depository Libraries constitute a natural control group for patent depository libraries.[17]

What Services did Patent Depository Libraries Provide (and to whom)?The main aim of the patent deposit libraries, both in the modern era and in the 1800s, was to provide access to technical information to potential users and to help them with prior art searches. Throughout the program's history, however, patent librarians have been embargoed from providing legal advice or other legal services. Thus, their services have focused on information provision. The records of the annual conference document the exceptional dedication of the library professionals to these tasks (see, e.g., Sneed, 1998, and Oliver, 2002). Surveys of patent

[14]Over time, however, the space requirement became less a concern after the introduction of microfilm. Indeed, the conversion from paper to microfilm distribution has been cited as a reason why many new libraries joined the program after 1982.

[15]The annual PDL trainings in Washington DC appear to have been both highly valued professionally and personally enjoyable to the PDL librarian community. Both the professional lessons and personal reflections are documented in the Patent and Trademark Resource Center Association Newsletters, which are available for review at http://ptrca.org/newsletters.

[16]There are two ways in which a library may qualify for FDL status: First, each member of Congress may delegate two qualified libraries or a library may be designated. Second, all libraries at land-grant colleges and universities, libraries of federal agencies, the highest appellate court of a state, and accredited law schools automatically qualify for the status of Federal Depository Library.

[17]The USPTO continues to operate the patent library program even after the advent of freely available patent document search engines, like Google Patents. The librarians we interviewed suggested that the current libraries, now called Patent and Trademark Resource Centers, aim to create value for the communities they serve by assisting with the search for prior art and by helping users negotiate databases that offer more sophisticated prior art search capabilities than publicly-available resources.

depository library users in the 1990s suggest that the libraries served mainly local users and inventors. Specifically, the 1991/1992, 1997, and 2002 surveys suggests that the majority of users were inventors or entrepreneurs, while fewer than 16% were attorneys, patent agents, or other legal staffers. The 1997 and 2002 surveys report that the median user estimated traveling between 11-20 miles to use the library, with 49% of 2002 users and 38% of 1997 users traveling fewer than ten miles (Brown and Arshem, 1993; United States Patent and Trademark Office, 1999, 2003).

(a) Opened from 1870 to 1974

(b) Opened from 1975 to today

Figure 1. : Location of all Patent Libraries in the U.S.

Note: Figure 1a shows the position of patent libraries in the continental United States opened before the major expansion from 1975 to 1997. Figure 1b shows the patent libraries that were opened during the expansion phase, the basis of our estimation sample. The gradation scheme informs about the time of patent library opening. The darker the shading, the earlier the library was opened. The four categories are 1975-1980, 1981-1985, 1986-1990 and after 1990

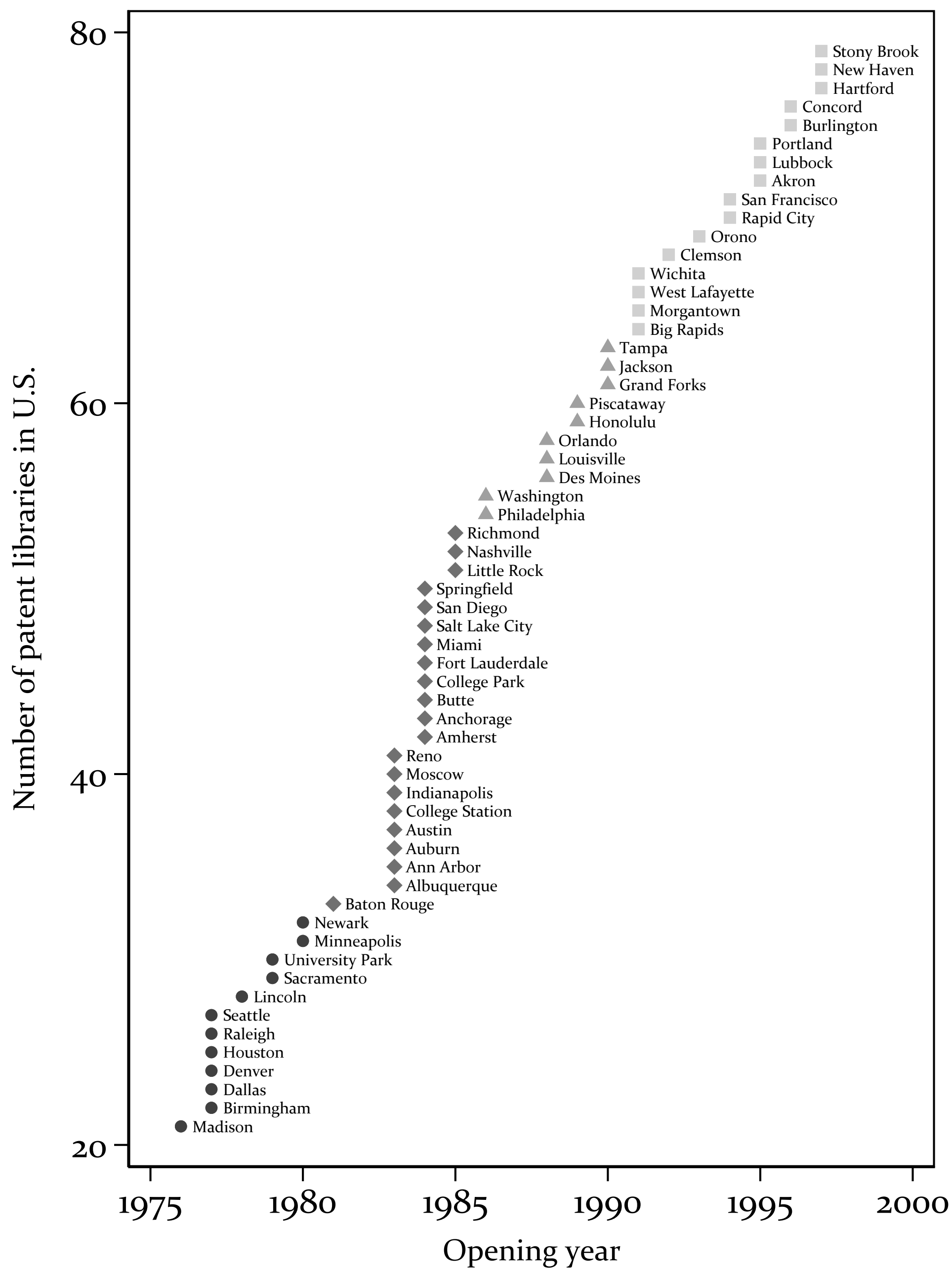


Figure 2. : The Expansion of the Patent Depository Library Program

Note: This figure shows the opening years of patent libraries that were opened during the expansion phase, the basis of our estimation sample, and the total number of patent libraries in the United States. The gradation scheme informs about the time of patent library opening. The darker the shading, the earlier the library was opened. The four categories are 1975-1980, 1981-1985, 1986-1990 and after 1990

III. Empirical Setup

Identification: Federal Depository Libraries as Control Group

To measure the impact patent library opening on regional innovation, we need a counterfactual estimate of what would have happened to patenting in region if the library had not opened. To do this, we use a control group that includes regions that are geographically proximate to newly-opened patent depository libraries, that have medium or large Federal Depository Libraries (FDLs), and that are within the same state as the treated patent library.[18]

Regions around Federal Depository Libraries are attractive as control group for three reasons: First, FDLs already handle government documents, which is one of the key criteria for becoming a patent library. Second, there are at least two FDLs in each congressional district, which ensures that FDLs would be able to serve as control group in each state. Third, medium and large FDLs satisfy the formal requirement for becoming patent libraries. They are therefore likely to possess the space, human capital, and library infrastructure required to become patent libraries. Indeed, 83% (53 of 64) of new patent libraries after 1975 are FDLs.

In our main specification, we focus on patent libraries that are also Federal Depository Libraries, as this enables us to work with a well-matched control group. In our principal estimation sample, we drop (a) all patent libraries that were opened before 1975 or after 1997, as the motivation for opening libraries in those periods was different than it was during the period of dedicated program expansion as explained above, (b) all patent libraries that are not Federal Depository Libraries or that only became Federal Depository Libraries after becoming a patent library, and (c) all patent depository libraries for which we cannot identify an FDL control library between 30 miles and 250 miles. We also drop the library in Burlington VT, because a suitable control region does not exist within its state. Burlington hosted a primary research facility of IBM during the sample period and its patents per capita ratio vastly exceeded that of other regions.[19] Our primary estimation sample, thus, includes 45 patent libraries that opened after 1975, along with 267 control libraries. Figure 3 shows the position of all patent libraries and all Federal Depository Libraries in our sample.

FDLs constitute a valid control group if the number of patents in patent library regions would have followed the same trend as the number of patents around the control libraries had patent libraries not been opened. Our identification assump-

[18]The Federal Depository Library program classifies libraries as small if they contain fewer than 250,000 volumes in the library, medium if they contain 250,000 to one million volumes, and large if they possess more than one million volumes of public materials.

[19]Indeed, no suitable control region exists for Burlington anywhere in the United States. During the sample period, the Burlington library region averaged more than 295 patents per 100,000 persons, while the sample average was approximately 17 and the region with the second highest number of patents per capita was Newark DE (home of a DuPont primary research facility) with around one hundred patents per 100,000 population.

tion would be threatened if librarians or administrators had applied to become Patent Depository Libraries (or were selected by the USPTO) in the expectation that innovative activities were about to burgeon in their regions. Although we cannot rule out this possibility, three factors suggest that this does not constitute a substantial threat to causal interpretation. First, in order to condition their application to become a patent library on future innovation, local librarians must be able to predict accurately near-term changes in local private sector innovation that diverge from those of broader trends in their state. While librarians likely have local insights, they would need to anticipate coming boosts in local patenting that do not involve pre-trends, would need to prepare their applications to become PDLs, and would need to have them approved at precisely the time that local patenting is about to increase. Based on our understanding of the mix of reasons that librarians report as having played important roles in libraries' application decisions, we believe that this is unlikely. Second, the program expansion from 1975 to 1997 aimed at opening one library in each state, with the goal of supporting equal access to patent materials across the country. Thus, it seems less likely that the USPTO accepted library applications based on changes in expected future patenting. Indeed, according to the USPTO, patent library status was supposed to be conferred upon the first qualified library in each state that applied for program participation. After 1997, however, the USPTO did switch to favoring regions with high patenting per capita. Third, while most centers of innovation ultimately receive patent libraries, the key to our identification strategy is that the timing with which the libraries are opened must be random with respect to innovation trends. The particular dates on which regions receive libraries does not follow a pattern of increasing or decreasing innovation importance, either in levels or in changes. To explore whether the assumptions underlying our identification strategy are reasonable, we conduct several robustness checks in Subsection IV.B.

Data Sources

For our empirical analysis, we combine data on libraries with geolocated patent data and population data from the U.S. Census (Census Bureau, 2010). The data on the opening dates of each patent library is from Jenda (2005) and the complete list of Federal Depository Libraries is from the online Federal Depository Library Directory.[20] We obtain patent data from PATSTAT European Patent Office (2016). To identify the geographic location of the inventors and inventor disambiguation we rely on the data of Balsmeier et al. (2018) and of Morrison, Riccaboni and Pammolli (2017). If there are several inventors on a patent, we allocate each location a share of the patent.[21] In our main specification, we use the

[20]The Federal Depository Library Directory is available on https://catalog.gpo.gov/fdlpdir/FDLPdir.jsp (last accessed 2017-07-30).

[21]Using patents as an indicator for innovative output is standard but not uncontroversial. In our particular case, patent libraries also might increase patenting without increasing innovation because they

15 miles around libraries as the unit of analysis.[22] Appendix A gives a complete list of patent libraries with opening dates and provides a description of the data processing.

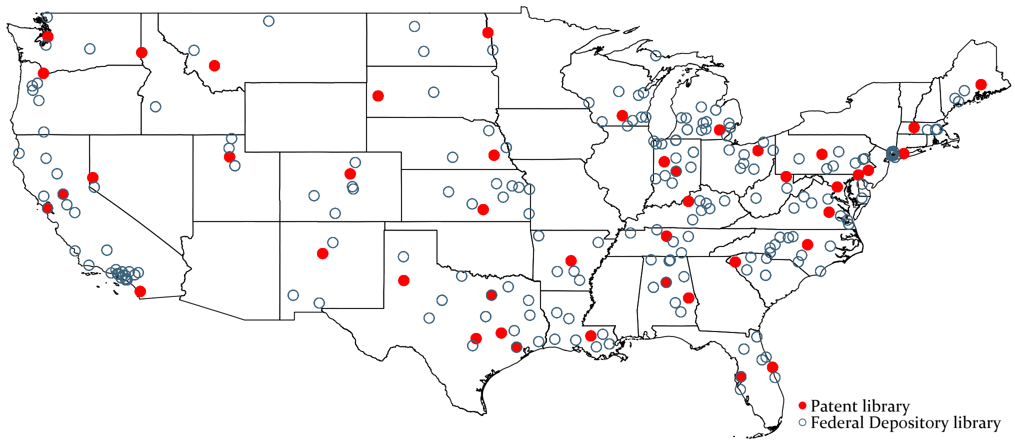


Figure 3. : Locations of Patent and Control Libraries

Note: The red dots show the position of patent libraries. The hollow dots show the positions of control libraries.

Table 1 shows summary statistics for patent libraries and matched Federal Depository Libraries in the year before the opening of the patent library in levels. Some of the means of the treated observations are (close to being) statistically significantly different from those of the control group. This is due to some outlier regions that do not have patenting activity in some years. In Section B-1 in the online appendix, we document that balancing improves when we drop these observations.

might make it easier to file a patent or because the librarians might give advice on how to structure a patent. Yet, this seems unlikely because a U.S. patent application can be mailed from any post office and the employees of patent libraries are only allowed to help with the search for prior art but not with the preparation of a patent filing.

[22]The 15 miles come close to the average commuting distance of workers in the United States (Rapino and Fields, 2013). In establishing the 15 mile radius, we have relied on surveys administered by the USPTO in 1997 and 2002 in which respondents indicated that the median user traveled 11-20 miles to use the library, while the modal user (38% in 1997, 49% in 2002) traveled fewer than 10 miles.

Table 1—: Summary Statistics in the Year before Opening

Main sample	Patent Libraries	Control Libraries	Diff	P-Value
Population in 100k	7.25	4.49	-2.77	0.10
Uni Library	0.69	0.69	0.00	1.00
# Patents	135.13	74.26	-60.87	0.08
# Patents/100k	17.71	13.65	-4.06	0.20
Citation-weighted patents	259.45	196.81	-62.63	0.27
# Pat. small firms/100k	7.94	6.90	-1.04	0.39
# Pat. big firms/100k	9.77	6.75	-3.02	0.21
# Pat. young firms/100k	5.71	4.89	-0.82	0.35
# Patents old firms/100k	12.00	8.76	-3.24	0.21
Number of libraries	45	267		

Patents by field	Patent Libraries	Control Libraries	Diff	P-Value
Electrical Engineering	2.67	2.35	-0.32	0.66
Instruments	2.74	2.10	-0.65	0.23
Chemistry	4.82	2.49	-2.33	0.12
Process Engineering	2.53	2.57	0.04	0.94
Mechanical Engineering	2.82	2.33	-0.49	0.41
Other Fields	2.11	1.80	-0.31	0.37

Note: This table shows the averages of the data for patent libraries and associated control libraries in the year prior to patent library opening. The last two columns show differences with the associated significance levels. A firm is defined as young if its first patent was filed less than three years before the opening of the patent library. Otherwise it is old. A firm is defined as small if it has no more than 20 patents before the opening of the patent library. Otherwise it is large. The p-values result from a t-test with unequal variances.

IV. Do Patent Libraries Increase Local Innovation?

It is unclear *a priori* whether the opening of a patent library will have an impact on innovation in the library's geographic region. On the one hand, improved access to patent literature could induce local innovation if inventors were to read the patent literature, draw valuable information from it, and, as a result, innovate at lower cost, or with greater effect than would have been the case in the absence of the library (Machlup and Penrose, 1950; Scotchmer and Green, 1990; Scotchmer, 1991; Landes and Posner, 2003). On the other hand, there are reasons to think that patent libraries may have no impact on local innovation. In addition to the possibility that such institutions are simply ineffective, it is also possible that they are rendered ineffective in disseminating technical information because of the opacity with which patents are written, the prospect of avoiding patent disclosure through the use of trade secrecy (Levin et al., 1987; Moser, 2011, 2013), the potential that follow-on inventors avoid reading patents to avoid willful infringement (Roin, 2005; Lee and Cogswell III, 2004), or even the possibility that disclosure decreases incentives to innovate by inhibiting duplicate inventive efforts.

A. *Primary Analysis: Patenting Increases After Library Opening*

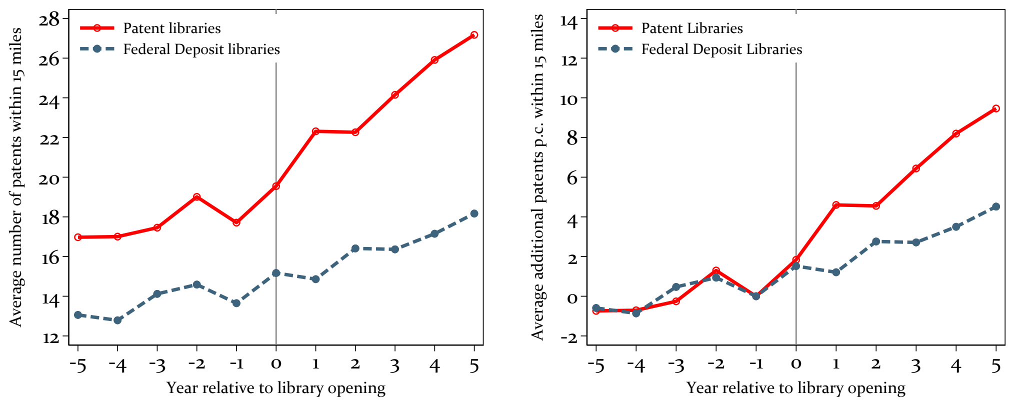
We begin by asking whether patent library opening impacts patenting within 15 miles around the new library. In Figure 4a we compare the raw difference in the average number of patents per 100,000 persons around treatment and control libraries. In Figure 4b we subtract from each series its value in the year before the opening of the library to account for different levels. In both cases, the two series begin to diverge in the period after the patent library opened.[23]

In Figure 5, we plot the yearly difference in the number of newly filed patents per 100,000 population in the 15 miles radius around the control and the patent libraries. For each library region, we consider the five years before and the five years after the library opening and we normalize the number of new patents to zero in the year before opening. We use weights to adjust for the different number control libraries per patent library to arrive at the average treatment effect on the treated (Iacus, King and Porro, 2012).[24]

The data suggest that the number of newly-filed patents around the patent library increases significantly after opening. Prior to the opening of the patent library, the number of patents per capita is similar for treatment and control libraries, with no visible trend in patenting. This is consistent with the parallel trends assumption and provides some confidence that the estimates represent a causal effect.

To quantify the size of the effect of opening a patent library on subsequent

[23] In online appendix B.2 we show these averages for different potential control groups.
[24] As we show in Table B-3 in the online appendix, our results do not depend on using these weights.



(a) Raw (b) Normalized in the Year before Opening

Figure 4. : Compare Averages

Note: This figure plots the average number of patents within 15 miles of the patent library (red solid line) and around Federal Depository Libraries (blue dashed line) in the five years before and after the opening of the library. Figure 4a shows the raw average and in Figure 4b we normalize the average relative to its value in the year before the opening.

patenting, we estimate the following difference-in-differences specification:

$$(1) \qquad \frac{\#Patents_{ijt}}{Population_{ij}} = \beta_1 \cdot Post_t + \beta_2 \cdot PatLib_{ij} \cdot Post_{it} + \alpha_{ij} + \gamma_t + \varepsilon_{ijt}$$

where j indexes each library (i.e., the 15 miles region around each library), $PatLib_{ij}$ is an indicator equal to one if the library in that region is a patent library or if it is not, but belongs to the control observations of patent library i, and $Post_{it}$ is an indicator equal to one in the years following patent library opening. We incorporate both individual library and year fixed effects as controls.[25] The coefficient of interest, β_2, measures the average yearly increase in the number of patents around a patent library in the five years after it was opened relative to the period before it was opened and relative to the controls in that period.

We report the results for estimating Equation (1) in Table 2. Column (1) documents that the number of patents per capita (patents per 100,000 persons) in the region around of the patent library increased on average by 3.2 relative to the control group. This implies an increase of around 18% relative to the average. This estimate is the primary result in the paper. If we can interpret the regression as causal, it implies that patent library opening induces local innovation in the area proximate to the libraries.

The increase in newly-filed patents we find in (1) is not associated with a

[25]The baseline effect of $PatLib_{ij}$ is taken up by the α_{ij}library fixed effects.

Table 2—: Patent Libraries and Local Innovation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Baseline			Age		Size		University	
	Patents	Citations	Av. Citations	Young	Old	Small	Large	Yes	No
Post	0.2	-10.2	-2.0	-0.4	0.6	0.3	-0.1	0.5	-1.0
	(0.7)	(15.2)	(1.6)	(0.4)	(0.5)	(0.5)	(0.4)	(0.9)	(0.9)
Pat Lib x Post	3.2**	60.7**	0.8	1.4	1.8**	1.7**	1.5	2.7	4.3
	(1.5)	(29.5)	(0.8)	(0.8)	(0.9)	(0.8)	(0.9)	(1.6)	(3.3)
Mean Dep.	17.8	238.8	13.0	6.2	11.6	9.4	8.4	17.8	17.7
R2 (within)	0.12	0.16	0.08	0.24	0.10	0.12	0.05	0.09	0.23
Obs.	3432	3432	3294	3432	3432	3432	3432	2266	1166

Note: This table shows the results from a difference-in-differences estimation with five years before opening as pre-period and five years after opening as post-period. The estimation equation is:

$$\frac{\#Patents_{ijt}}{Population_{ij}} = \beta_1 \cdot Post_t + \beta_2 \cdot PatLib_{ij} \cdot Post_{it} + \alpha_{ij} + \gamma_t + \varepsilon_{ijt}$$

where $PatLib_{ij}$ is an indicator if the library j is a patent library or if it is not, but belongs to the control observations of patent library i,and $Post_{it}$ is an indicator for all years after the opening of the patent library. As controls we use library and year fixed effects. In column (1) we use Federal Depository Libraries (FDLs) within 250 miles as controls. In column (2) we weight each patent with its forward citations. In column (3) we use average forward citations per patent as an outcome. In columns (4) and (5) we split the dependent variable by young and old assignees. An assignee is young if it filed its first patent no more than three years before the opening of the library and old otherwise. In the following two columns we split the dependent variable by the size of assignee. An assignee is defined as large if it has more than 20 patents before the opening of the patent library. In column (8) and (9) we consider the subsample where the patent library is also a university library and where it is not. In all regressions, we use the weights suggested by Iacus, King and Porro (2012) to identify the average treatment effect on the treated. Standard errors are clustered on the (assigned) patent library level. *, **, and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

decrease in patent quality. We test this in column (2) by estimating the impact of library opening on citation-weighted patents. Using this common approach to account for patent quality, we do not observe a decrease in citations received, which would be consistent with a decline in quality, but rather a boost of around 60 citations. This implies an increase of around 25 percent relative to the mean number of citations received. In column (3), we then use average citations per patent as the outcome to assess how patents changed on average after patent library opening. The impact is positive, but not significantly different from zero. This speaks against an explanation where the additional patents are just marginal.

In columns (4) to (7) we split the dependent variable by the type of assignee. In our analysis, we consider firms to be young if their first patent was filed less than three years prior to the library opening. These young firms may be entrepreneurial ventures, but they may also be existing firms that had not previously applied for patents. The opening of the average patent library increases the number of patents by young companies by 1.4 patents per 100,000 residents, an increase of around 23 percent relative to the mean. The effect for old companies, reported in column

(5) is larger in magnitude than that of young firms relative to the mean but is smaller in relative terms. Columns (6) and (7) demonstrate that the impact is similarly large for small and large companies. An assignee is defined as large if it has more than 20 patents before the opening of the patent library. In columns (8) and (9), we split the result by the type of library. We investigate the impact of library opening at universities and at other public libraries and find that the impact on innovation is similar and generally noisy.

These results suggest that the opening of a patent libraries had substantial advantages in terms of local patenting. Considering that patent collections constitute only a small fraction of the total operating expenses of each patent deposit library makes it likely that the boost in patenting induced by access to patent technical documents is, across the program, justified by the cost. We do not have information on the operating costs of patent libraries during this time period. We have, however, information on the costs and revenues of patent libraries in the 1991/92 patent library survey conducted by the USPTO (all numbers in 1991 Dollars). Although the survey is small and not all libraries responded to each question, the average library indicated that they had around $14,000 in (patent) library expenses. In addition, the average salaries paid in connection to patent library services was $43,000. However, note that the costs and benefits accrue to different stakeholders. The benefits for the libraries were rather small. The average annual revenue from providing patent library services was only around $7,000 per year. Thus, while the overall benefits to the region seem to have been high, the benefits for the individual libraries seem to have been low in comparison to the costs.

B. Auxiliary Analysis

The principal concern for our estimation strategy is whether regions that receive a new patent library would have experienced equal boosts in patenting even if they had not received patent depository library facilities. This could occur if the patent librarians were to have accurately anticipated the timing of local innovation bursts and, in this expectation, applied to become a patent library. An analysis of patenting under such circumstances might yield observationally equivalent results, though these additional local patents would not have been induced by the library opening. In this section we report the results from auxiliary analyses that shed light on this possibility. Our conclusion is that such a coincidence is not likely.

Alternative Control Group: Not-yet-opened Patent Libraries Our main analysis relies on the assumption that treatment and control regions differ only as a result of treated regions receiving patent libraries and, hence, the control regions enable us to estimate a counterfactual for the patenting that would have occurred in the absence of the patent libraries. We explore the robustness of our results to the relaxation of this assumption. In the following, we report the results of a model that identifies the impact of patent library opening using regions around not yet opened patent libraries as controls for patent libraries that were opened earlier ("Within sample"). In some sense, this is our strictest test since the identification here relies only on the timing of the treatment and not on any differences between libraries that receive patent collections and those that do not.

Specifically, we estimate Equation (1) using not-yet-opened patent libraries as controls. The coefficient of interest now measures the average yearly increase in the number of patents around a patent library in the five years after it was opened relative to the period before it was opened and relative to the not-yet-opened patent libraries in that period. There are 45 library openings during the sample period that we use for our baseline analysis. In the within analysis, we cannot find a control library for the patent library in Stony Brook NY, since it is the last library opening in our sample (in 1997), leaving 44 library openings. On average, there are 25 control libraries per treated library and as our sample includes 11 years, this results in 12,100 treatment $\times$ control $\times$ year combinations. As we drop control libraries as soon as they open, the sample ends up including 9,040 observations.

We explore the robustness of the principal results to the use of this control group in Table 3. Column (1) shows the effect on the number of patents per capita, our main measure of innovation. We again see a similar, but somewhat smaller effect than in our baseline specification (around 13% relative to baseline). Column (2) demonstrates that the effect on citation-weighted patents is smaller and not statistically significant, though it remains positive. Column (3) shows the impact on the average number of forward citations per patent, which again is small and insignificant. Together, these analyses suggest that, even in this very restrictive sample set, the results hold. Columns (4) and (5) shows the effect for young and for old companies. The effects are only statistically significant for young assignees and are again larger relative to their mean patenting rate. Columns (6) and (7) show that in the within-sample, the effect is positive and statistically significant for small companies, and is insignificant for large companies. These results suggest that young and small companies may have particularly responded to patent library openings.

More Alternative Control Groups In Figure 6 we continue to explore the robustness of the baseline results to different control groups.[26] In line 1) we repeat our baseline specification. In line 2) in the figure, we match on both state affiliation and on the status of being a university library. In line 3) we match on state and university and employ coarsened exact matching to ensure similarity in patenting per capita in the year before the opening, using five bins for patents per capita. In line 4) we use the same matching approach as in line 3) and also match on population within 15 miles of the library using five bins. In each case,

[26] Table B-2 in the online appendix shows the results in table format.

Table 3—: Using Future Patent Libraries as Controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Baseline			Age		Size	
	Patents	Citations	Av. Citations	Young	Old	Small	Large
Post	-0.0	-0.5	0.1	0.1	-0.1	-0.1	0.0
	(0.2)	(5.8)	(0.3)	(0.1)	(0.2)	(0.1)	(0.1)
Pat Lib x Post	2.2**	23.0	-0.2	1.6***	0.6	1.6***	0.6
	(1.0)	(21.3)	(0.6)	(0.6)	(0.7)	(0.4)	(0.7)
Mean Dep.	16.3	214.8	13.0	1.6	14.7	7.4	8.9
R2 (within)	0.27	0.21	0.15	0.46	0.14	0.30	0.11
Obs.	9040	9040	8975	9040	9040	9040	9040

Note: This table shows the results from a difference-in-differences estimation with five years before opening as pre-period and five years after opening as post-period. We use not-yet opened patent libraries as control for opened patent libraries. The estimation equation is:

$$\frac{\#Patents_{ijt}}{Population_{ij}} = \beta_1 \cdot Post_t + \beta_2 \cdot PatLib_{ij} \cdot Post_{it} + \alpha_{ij} + \gamma_t + \varepsilon_{ijt}$$

where i indexes each patent library (i.e., the 15 miles region around each library), $PatLib_{ij}$ is an indicator equal to one if the library in that region is the focal patent library or if it is not, but belongs to the control observations of patent library i (i.e., not-yet-opened patent libraries at time t), and $Post_{it}$ is an indicator equal to one in the years following focal patent library opening. We incorporate both year fixed effects and fixed effects for all combinations of focal patent libraries and not-yet-opened patent libraries as controls. In column (2) we weight each patent with its forward citations. In column (3) we use average forward citations per patent as an outcome. In columns (4) and (5) we split the dependent variable by young and old assignees. An assignee is young if it filed its first patent no more than three years before the opening of the library and old otherwise. In the following two columns we split the dependent variable by the size of assignee. An assignee is defined as large if it has more than 20 patents before the opening of the patent library. In all regressions, we use the weights suggested by Iacus, King and Porro (2012) to identify the average treatment effect on the treated. Standard errors are clustered on the (assigned) patent library level. *, **, and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

the narrower control sample yields a reduction in the number of observations but a similarly-sized mean effect. In lines 5) and 6) we do not condition on state but rely on a control group of FDLs defined by their distance to the patent library. In line 5) we use the closest FDLs that are within 100 miles. In line 6) we use the (up to) five closest control libraries within 250 miles. In line 7), we construct the counterfactual by computing the “synthetic development” of patent library regions holding their share of a region among all U.S. patents constant to the pre-opening level. The results are quantitatively similar to our main estimates throughout. In relative terms, they range from around 8% (line 5) to 20% (line 4) relative to the mean patenting rate of the respective estimation sample (see Table B-2 in the online appendix for details). Taken together, these additional analysis document the robustness of the findings to the choice of control group.[27]

[27]In Appendix B.2 we report average patenting per capita around patent library opening for all alternative control groups.

No Effect of Pseudo Treatment of Closest Control Library A prospect that threatens the interpretation that patent library opening induces an innovation response is the possibility that library opening may cause inventors to move geographically, but achieve no net effect on innovation. Were inventors to relocate to patent library regions, our results could reflect a simple change in the spatial distribution of patenting rather than an increase in innovation. If this were the case, the treatment assignment would then violate the stable unit treatment value assumption (SUTVA), as the opening up of the patent library would decrease innovation in control library regions. To test this possibility, we re-run our analysis assuming that the inventors most likely to relocate to the patent library regions are those in the geographically closest control regions. We therefore replicate our baseline analysis omitting patent libraries and, instead, assigning a fake treatment indicator to the Federal Depository Library closest to each dropped patent library. We report the result in line 8) of Figure 6. The fact that the coefficient is slightly negative, not statistically significant, and not of a comparable magnitude suggests that there is no differential trend between closer control libraries and libraries that are further away.

No Effect Outside of Local Region Figure 5 demonstrates that in the five years before the patent library opening, there are no systematic differences in patenting between regions with Federal Depository Libraries that are about to receive a patent library and those regions with Federal Depository Libraries that do not obtain patent libraries. The number of patent applications that are ultimately granted increases in the years after patent library opening. In addition to not being present prior to the arrival of patent deposit libraries, the effects in our main specification are also substantially smaller in regions outside of the patent library's commuting radius. In Figure 6, we also show the robustness of our results to this change in the dependent variable. Our preferred specification is replicated in line 9). Line 10) of the figure demonstrates that the increase in patents is localized in the geographic region most proximate to the arriving patent library. For patents filed between 15 and 50 miles the effect is smaller than in the region close to the library, and is statistically insignificant. In line 11), we show that the effect between 50 and 100 miles is insignificant and slightly negative.

Additional Analyses in the Appendix In Online appendix B we conduct further sample splits and robustness analysis. Online appendix B.3 shows the time-varying treatment effects for longer time window. Online appendix B.4 shows, among others, that our baseline effect is mostly driven by patents assigned to firms and that there is only a small effect associated with patents assigned to universities. In addition, the effect is about the same size in regions with historically high than in regions with historically low patenting levels, even though this is measured with substantial noise.

In Table B-3 described in online appendices B.4 and B.5 we show that our results are robust to not using the CEM weights suggested by (Iacus, King and Porro, 2012), to including patent library-specific trends, to only using co-located inventors, to using other circles around the libraries (for example 25 or 50 miles), and to using the number and the log number of patents as the dependent variable while controlling for (time-varying) population.

In Online appendix B.6 we drop each library in turn and find that the effect does not depend on any particular library in our main sample. Appendix B.7 provides evidence that the structure of patent documents is also affected by the opening of patent libraries. Patents of young companies experience an increase in median backward citation distance, consistent with the idea that library opening eased access to prior art. These results are consistent with patent libraries improving the access to distant and therefore less likely to be known patents.

One could be worried that patent libraries substitute for patent attorney activity. In Appendix B.8, we show that the number of local registered patent attorneys did not change in response to patent library openings. This also speaks against an alternative explanation where local economic activity rises irrespective of disclosure due to other local developments (Andrews, 2019).

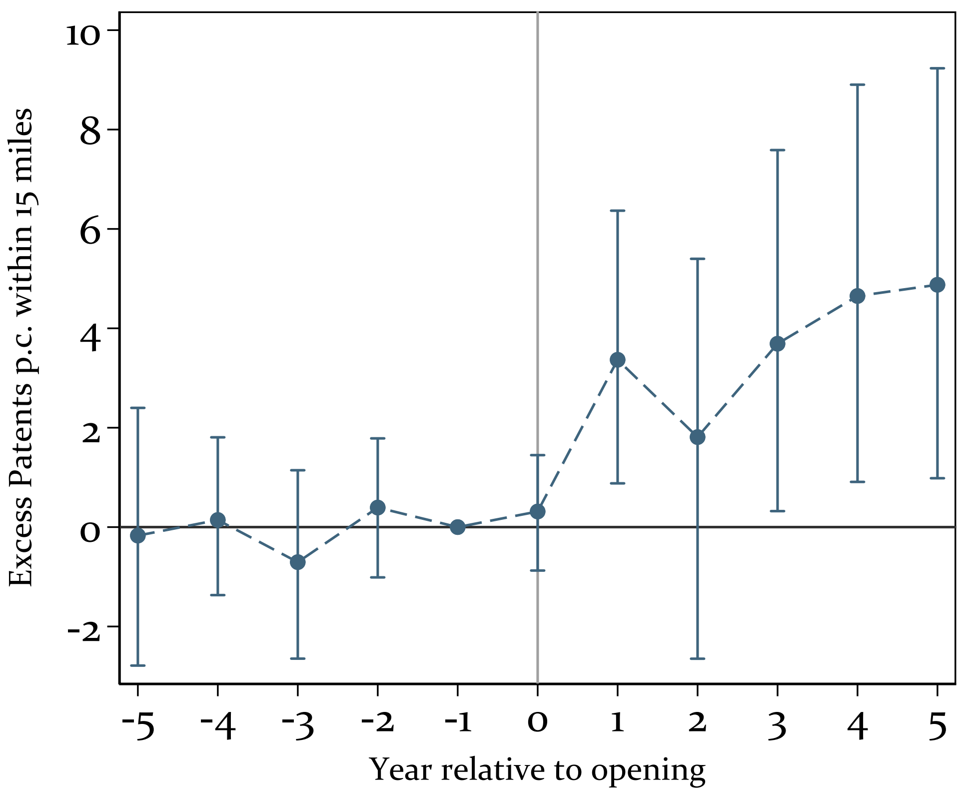


Figure 5. : Event-study

Note: This figure shows the yearly average treatment effects on the treated of opening up a patent library on the average number of patents within 15 miles of patent libraries relative to the average number of patents around matched federal depository libraries. The 95% confidence intervals are based on bootstrapped standard errors. We use the weights of Iacus, King and Porro (2012) to arrive at the average treatment effect on the treated. We assign each patent library and all Federal Depository Libraries within the same state and within 250 miles as control group. We exclude the patent library of Burlington VT.

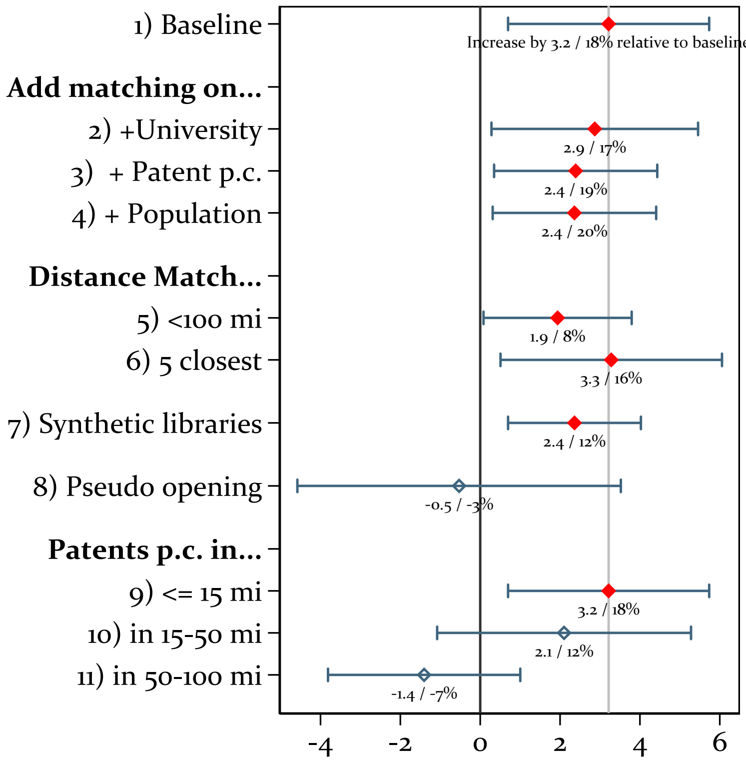


Figure 6. : Auxiliary Analysis

Note: This figure shows the results from difference-in-differences estimations with five years before opening as pre-period and five years after opening as post-period using equation 1. As controls we use library and year fixed effects. In line 1) we use our baseline specification, i.e., Federal Depository Libraries (FDLs) within the same state as controls. In the next line 2), we use only FDLs that are also university libraries as controls if the patent library is also a university library. If the patent library is not a university library we only use FDLs that are not university libraries as controls. In line 3) we match on state, being a university, and in addition we do a coarsened exact matching (CEM) with 5 bins on patent per capita in the year before the opening. In line 4) we repeat the analysis of line 3) but also employ coarsened exact matching with 5 bins on population. Thus only FDLs in similar sized cities, with similar number of patent per capita, and the same type of library are used as controls. In line 5) we do not match on state but take all FDLs within 100 miles as controls that do not have a closer different patent library. In line 6) we use the 5 closest FDLs as controls. If more than one FDL is at the same distance, we keep all of these. In line 7) we construct a counterfactual by keeping the share of patents around patent libraries among U.S. patents constant to pre-opening levels. In line 8) we assign a treatment indicator to the FDL closest to the patent library and drop all patent libraries from the sample. In line 9) we use the number of patents below 15 miles as outcomes (the baseline). In line 10) we use the number of patents between 15 and 50 miles as outcomes. In line 11) we use the number of patents between 50 and 100 miles as outcomes. We use the weights suggested by Iacus, King and Porro (2012) to identify the average treatment effect on the treated. Standard errors are clustered on the (assigned) patent library level. Red full dots denote statistical significance at the 10% level. Below the coefficient and the error bars, we show the point estimate as well as the magnitude of the coefficient relative to mean patenting rates. Table B-2 in the online appendix shows these results in table format.

V. Mechanism: Better Access to Patented Knowledge

Our prior analyses document that the opening of patent libraries induces an increase in local patenting. These analyses suggest that reducing the costs of accessing the patent record facilitates inventions. This might be the case because of different types of information that the patent record – via these libraries – provides to inventors. First, they may learn about new technical knowledge, the main idea behind many arguments advocating the importance of disclosure. Second, inventors may learn about what sorts of inventions are already patented, avoiding duplicative efforts. Third, inventors may learn about the patent system, e.g., what kinds of inventions are patentable. Fourth, inventors may learn about which other inventors are working on which topics, improving the matching of inventors into new teams.

In this section, we explore the empirical merit of these potential explanations. First, we investigate whether access to technical knowledge might drive our results. To illustrate this mechanism, we recount the story of the development of Zithromax, which suggests that if patents are informative, easier access to technical information in patents might improve the ability to build new technologies based on prior patented knowledge. Thereafter, we examine variation in the strength of the library effect across technology areas. We then find that the effect is strongest in the field of chemistry, in which patents provide substantially better disclosure of technical information than in all other fields. Next, we analyze the patent text to understand whether the boost in patenting induced by library opening is the result of inventors building on knowledge that existed in the region before the library was opened or whether the libraries expand the knowledge base of local inventors. We find that library opening leads to an increase in patents with words that are new in the library region but were previously used in patents outside the region. This is consistent with an explanation that libraries expand the knowledge base of local inventors and disseminate information valuable for subsequent patenting. Our final analysis in this section finds that library opening has a similar impact on new and existing inventors and inventor pairs, which suggests that the patent library boost is not the result of improving inventor matching.

Like any observational study, ours is limited in pinning down the exact mechanism. While we thus cannot entirely rule out that all of these mechanisms play a role, these results suggest to us that access to prior art, and especially access to technical information in patents, is a primary driver of the effects we observe.

A. An Illustrative Example: Patent Disclosure and the Development of Zithromax

Although academic analyses of patent libraries are scant, inventor accounts suggest that the disclosure of prior art via patent libraries facilitated inventive activity. Historical reports suggest that, in the early days of patent libraries, Thomas Edison made use of them to search for prior art (Sneed, 1998). Jack

Kilby, the co-inventor of the integrated circuit, was expansive in his efforts to read patent documents issued by the U.S. government: "You read everything-that's part of the job. You accumulate all this trivia, and you hope that someday maybe a millionth of it will be useful" (Reid, 1985, p. 65).

The case of Zithromax provides an example of how this mechanism might work in practice (Idris, 2002; Li, 2009). Beginning in 1974, Pfizer had undertaken a program to develop a new macrolide, an antibiotic of the same type as erythromycin, but with greater antimicrobial effect. Despite significant investment, more than 2000 tested compounds and eight human trials, the firm did not make material progress. As a result, the firm was on the verge of closing down the program in 1980. While reviewing patent documents at the United States Patent and Trademark Office, Pfizer's chemists stumbled upon a patent for a molecule with precisely the features they desired, which had been granted to the then-Yugoslavian firm Pliva. Building directly upon the initial patent, two Pfizer chemists, G. Michael Bright and Dick Watrous, methylated an amine of Pliva's drug, thus creating a slightly-modified version of the molecule, which Pfizer then patented. The time between the publication date of Pliva's patent and the application date of Pfizer's patent was only six months. Subsequent to its patent filing, Pfizer reached a licensing agreement with Pliva and in 1991 received FDA approval to offer Azithromycin, for sale in the United States under the branded name Zithromax. During the 1990s, it became one of the best selling branded antibiotics in the United States and worldwide, with annual sales peaking at US$2 billion at the time its patent expired in 2005.

B. Technology: Effect is Concentrated in Chemistry

If the effect in our principal analyses is, indeed, driven by improved access to patent prior art, we would expect it to be concentrated in technologies where patents are particularly informative. Evidence on this is presented by Gambardella, Harhoff and Nagaoka (2011), who report the results of surveys asking inventors how much time they saved by reading patents in various fields. They find that the average inventor claims to save twenty-five hours by reading patents in Chemistry, eight hours in Process Engineering, five hours in Instruments and Mechanical Engineering, three hours in Electrical Engineering, and eight hours in Other Fields. These results suggest that patents are most informative in chemistry. One reason may be that patents on chemical compounds display the specific molecular formations, thus fully disclosing the invention covered by the patent. To illustrate this point, online appendix C.1 shows the patent on Acetyl Salicylic Acid, commonly known by its trade name, Aspirin, and displays the formula for the molecule. Due to the clarity of chemical disclosure and the clarity of the associated patent rights, chemistry is a field in which patents have been documented to be valuable and important for appropriability (Cohen, Nelson and Walsh, 2000).

In Figure 7 we report the results of estimating Equation (1) using patents by technological fields as outcome variable.[28] To define technology fields we use a classification that aggregates IPC technology classes to larger sub fields (Schmoch, 2008). This is the same technology classification as in Gambardella, Harhoff and Nagaoka (2011). We find that the increase in patents is most pronounced in the field of Chemistry.[29] There is also smaller effect in Instruments. Consistent with the prospect that patent disclosure is a key mechanism driving the results, the effect is largest for inventors in fields where patents are most informative.[30]

C. New Words: Effect Driven by Knowledge Transfer

One potential mechanism that might drive the increase in innovation is that reading the technical information in existing patents gives inventors ideas for new patents or helps them understand the technical bases for subsequent innovations. To see if this is the case, we analyze how the text of patents changes after a patent library opens. We use the data of Arts, Cassiman and Gomez (2018) that gives us the set of words used in the abstract and title of each U.S. patent from 1976 to 2013 and add to this data all words of the first independent claim from the PatentsView database (PatentsView, 2020).[31] If an inventor reads technical information in a patent and uses this information for her own invention, it seems plausible that she might use the words from the prior art she read. So if reading prior art helps an inventor, we would expect that the number of patents increases that use words that are new to the region but were used previously in patents in other regions. In contrast, if the opening of a library changes the number of inventors or improves their productivity for any other reason we would expect to find an increase in new words across all patents independent of their content. The maintained assumption here is that access to other channels of learning of new words, such as access to scientific literature, does not follow differential trends in treatment and control regions. This is plausible as the libraries housing patent collections did not open as entirely new libraries, but expanded existing library collections to incorporate full-text patent documents and search technologies.

In Figure 8 we show the time-varying treatment effects of our analyses. In the four panels we split the total number of patents into four groups: In panel (a) we use the number of patents that include at least one word that is new to the region but not new to the world. These are words that were used in U.S. patents before, but not in the region around the library under consideration. In panel (b) we use the number of patents that use a word that only appeared in the region after patent library opening, but that is not entirely new to the region. This captures

[28]Note that this specification splits the main dependent variable by field. Thus, the coefficients represent impacts on subsets of the number of patents per capita and therefore add up to the main effect.

[29]In online appendix C.2, we show that this effect only arises after the opening of the patent library.

[30]We replicate these results in online appendix C.3, using a different technology classification scheme that is more detailed and includes a larger number of fields.

[31]We only have data on claims for the time period after 1974 and hence can only use this time period for our analysis. We provide details on how we process the data in online appendix A.

second-round effects of newly introduced words from the first category. In panel (c), we use the number of patents with a word that is new to the world, i.e., words that had not previously appeared in the USPTO patent corpus as the outcome. This captures innovations with genuinely new terms. In Panel (d), we use the number of patents containing only words already used in the region before patent library opening as outcome, i.e., innovations incorporating only 'old' words. All patent counts are standardized by population as in our main analysis.

If technical disclosure were an important mechanism inducing post-library patenting, we would expect patent library opening to increase the local diffusion of words that had been previously used in patents generated outside the local region. Therefore, we would expect an increase in patents containing words that are new to the region but not the world (panel a). On the other hand, since local knowledge is already available at low cost, we would not expect that easier access to technical information via patent libraries would lead to a significant increase in patents containing words used in the region prior to patent library opening (panel d).

This is, indeed, what we find. In line with our main results, treatment and control regions display parallel trends before the patent library opens. After patent library opening, the number of patents per capita that contain words that are new to the region, but not to the world, increases. In contrast, we do not find a significant increase in the number of patents with words that were known in the region before patent library opening, although we should note that the coefficients are imprecisely estimated. We also find the expected increase in the number of patents that use a word that only appeared in the region after patent library opening, but that is not entirely new to the region (panel b). This suggests that inventors either keep on inventing using the same novel words or others also learn about the new words from the patent record or from interpersonal contact. We do not have a clear prediction of what should happen with the number of patents with words that are new to the world as disclosure might either help or hinder the creation of truly new ideas (panel c). We find no effect of library opening for this class of words.

Taken together, these results speak in favor of the hypothesis that access to patent documents provides access to technical information, the key function of patent disclosure. They speak against the hypothesis that patent libraries predominantly provide information about what inventions are already patented or about what kind of inventions are patentable. In these cases, we would expect that the number of patents without new words in panel (d) also changes.

In Table 4, we quantify these results using the difference-in-differences specification from equation (1). In column (1) we repeat the main specification with patents per capita as the outcome variable. In column (2), we use the same measure but count only those patents where we can classify words using their text. In this and all remaining columns, we restrict our analysis to patents filed in or after 1975 as the patent text is only available from then. In column (3) we use the number of patents that use a word new to region but not the world standardized by population in the region as the outcome. In column (4), we measure second-round effects by using the number of patents that use a word that only appeared in the region after patent library opening, but that is not entirely new to the region, standardized by population as the outcome. In column (5) we use the number of patents with a word that is new to the world as the outcome. In column (6) we use the number of patents without new words as the outcome. As suggested by the time-varying treatment effects, the only statistically and economically significant increase is in the number of patents per capita using words that are new to the region but not new to the world. The increase of 1.9 patents is around 56% of the overall increase of 3.4 patents. Together with second-round effects of newly introduced words, patents transferring knowledge from outside the region account for almost 80% of the total effect. The mean estimates for the number of patents with new to the world words (columns 5) do not rise. Despite being around 51% of the total number of all patents, the mean increase in patents without new words is only 24% of the overall effect (0.8 of 3.4) and is statistically not different from zero (column 6).

Table 4—: Text analysis

	(1)	(2)	(3)	(4)	(5)	(6)
			Patents p.c. with words...			
	Baseline	Classified	New to region but not world	Words that appeared in region after opening	New to the world	Words already used in the region
Pat Lib x Post	3.2**	3.4**	1.9***	0.8**	-0.2	0.8
	(1.5)	(1.6)	(0.7)	(0.4)	(0.4)	(0.8)
Mean Dep.	17.8	18.3	6.5	0.5	1.7	9.5
R2 (within)	0.12	0.10	0.10	0.23	0.09	0.06
Obs.	3432	3250	3250	3250	3250	3250

Note: This table shows the results from a difference-in-differences estimation with five years before opening as pre-period and five years after opening as post-period analogous to equation 1. As dependent variable, we use the number of patents per 100,000 population as in the baseline regression in column (1). In column (2), we repeat this regression using only the subset of patents where we can classify words using our text analysis. Because of data availability, we restrict our analysis to patents filed in or after 1975 in all remaining columns, which leads to a lower number of observations. In column (3), we use the number of patents that contain a word that is new to the region around the library, but not to U.S. patenting overall as the numerator of the dependent variable. In column (4), we use the number of patents that contain words that are not entirely new to the region, but only appeared in the region after patent library opening. In column (5), we use the number of patents that contain a word that is new to U.S. patenting overall. In column (6), we use the number of patents that do not contain a word from the previous three categories. In all regressions, we use the weights suggested by Iacus, King and Porro (2012) to identify the average treatment effect on the treated. Standard errors are clustered on the (assigned) patent library level. *, **, and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

D. Results are not Driven by Improved Inventor Matching

One potential benefit of a patent library might be that potential inventors meet there to form up new inventor teams. This improved matching may have increased innovation in a region independently of patent disclosure. To examine whether this mechanism is at work we determine for each patent whether the combination of inventors on this team is new ("new team") or whether the patent is filed by an existing inventor team ("old team") by using the inventor disambiguation of Li et al. (2014). We also classify patents with only one inventor into "first time inventors" and "repeat inventors." If improved matching were the main driver of our results, we should see a rise in new teams and no effect for existing teams. In contrast, if a patent library makes all inventors more productive, e.g., by providing technical information of value for teams regardless of composition, we should see a similar boost in patenting among new teams, old teams, and repeat inventors. It is unclear which influence we would expect for new individual inventors. On the one hand, a patent library might enable would-be inventors to become actual inventors. On the other hand, improved matching might enable them match to new teams.

Table 5—: Inventor matching

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Baseline	Single inventors		Multiple inventors		Combined	
	Patents	First time	Repeat	New team	Old team	All new	All old
Post	0.2	0.4	-0.3	-0.2	0.2	0.2	-0.1
	(0.7)	(0.3)	(0.3)	(0.3)	(0.2)	(0.5)	(0.4)
Pat Lib x Post	3.2**	0.4	0.7*	1.5**	0.6*	1.9*	1.3**
	(1.5)	(0.4)	(0.4)	(0.7)	(0.3)	(1.0)	(0.6)
Mean Dep.	17.8	5.7	4.1	6.0	1.7	11.7	5.9
R2 (within)	0.12	0.04	0.03	0.19	0.13	0.10	0.07
Obs.	3432	3432	3432	3432	3432	3432	3432

Note: This table shows the results from a difference-in-differences estimation with five years before opening as pre-period and five years after opening as post-period analogous to equation 1. As dependent variable, we use the number of patents (per 100,000 population) as in the baseline regression in column (1). In column (2), we repeat this regression using only patents filed by single inventors that had not patented before. In column (3), we use patents by single inventors who already had patented before. In column (4), we use patents by inventor teams that had not patented before in this group composition. In column (5), we use patents by inventor teams that had patented before in this group composition. In column (6), we use patents by inventors or inventor teams that had not patented before, summing the dependent variables of columns (2) and (4). In column (7), we use patents by all inventors or inventor teams that had patented before and thus use the sum of the dependent variables in columns (3) and (5). In all regressions, we use the weights suggested by Iacus, King and Porro (2012) to identify the average treatment effect on the treated. Standard errors are clustered on the (assigned) patent library level. *, **, and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

In column (1) of Table 5 we repeat our main specification. In columns (2) to (5) we show the results for new and repeat single inventors and new and old teams. While improved matching as the main mechanism would suggest that we see disproportionally more patenting by new inventor teams, this does not appear to the be the case in the data: Repeat single inventors and inventor teams that we observed in the data before the patent library opening increase their patenting by roughly the same rate as new inventor teams. Patents by new teams increase by 25% (column 4; 1.5 relative to a baseline of 6). Patents by old teams increase by 35% (column 5; 0.6 relative to a baseline of 1.7) and patents by repeat inventors by 17% (column 3; 0.7 relative to a baseline of 4.1). Taking the last two together this implies that repeat inventors and existing inventor teams increase their patents by 22% (column 7; 1.3 relative to a baseline of 5.9). This is similar to the 25% increase by new inventor teams.

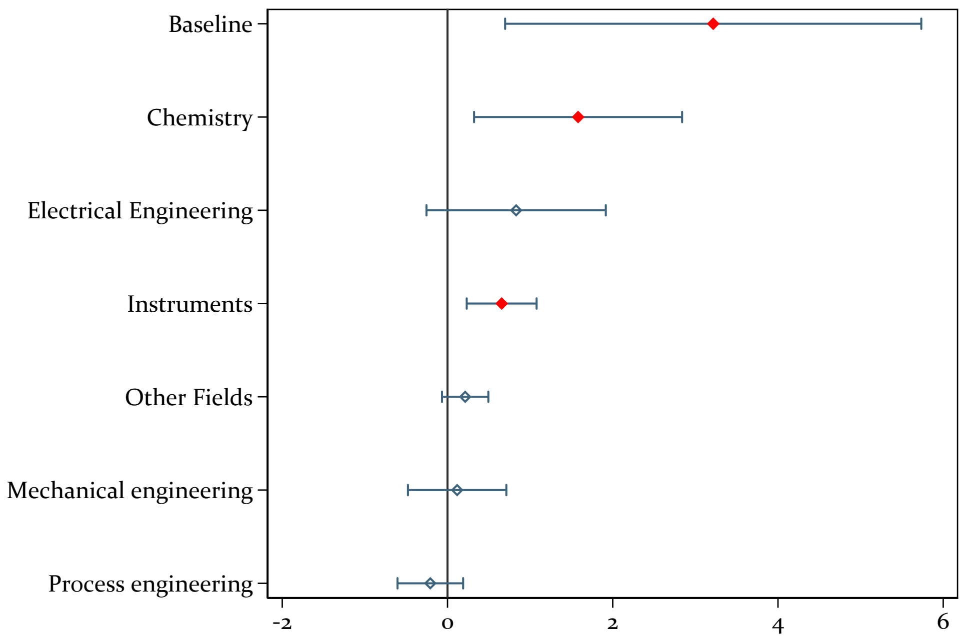
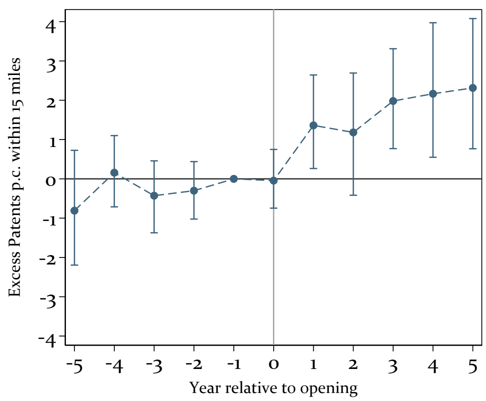


Figure 7. : Effect by Technology Category

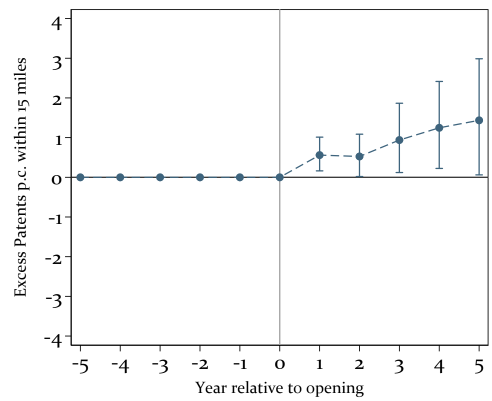
Note: This figure shows the results from a difference-in-differences regressions with five years before opening as pre-period and five years after opening as post-period. The estimation equation is:

$$\frac{\#Patents_{ijt}}{Population_{ij}} = \beta_1 \cdot Post_t + \beta_2 \cdot PatLib_{ij} \cdot Post_{it} + \alpha_{ij} + \gamma_t + \varepsilon_{ijt}$$

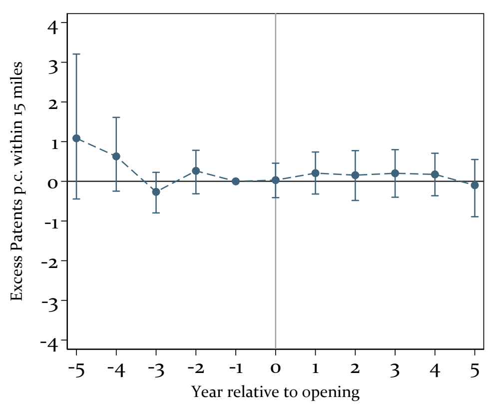
where $PatLib_{ij}$ is an indicator if the library j is a patent library or if it is not, but belongs to the control observations of patent library i,and $Post_{it}$ is an indicator for all years after the opening of the patent library. As controls we use library and year fixed effects. Each coefficient is from a separate regression where we split the dependent variable of our baseline regression by field as indicated in the figure. Thus, we only use patents in a specific field per 100,000 population as the dependent variable. The technological fields follow the ISI-OST-INPI classification of 1995 as defined in Schmoch (2008). The range plots indicate the 90% confidence intervals for the coefficient that are plotted with a hollow diamond if the coefficient is not significantly different from zero or a full diamond if the coefficient is significantly different from zero. In online appendix C.3 we report the results for more detailed and alternative classifications of technological sub fields.



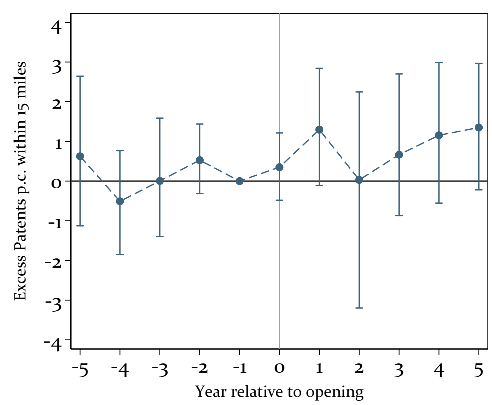
(a) New to region but not world



(b) Words that appeared in region after opening



(c) New to the world



(d) Words already used in the region

Figure 8. : Text analysis: Patents p.c. with words...

Note: Note: These figures show the yearly average treatment effects on the treated of opening up a patent library on the different measures of the average number of patents per 100,000 population within 15 miles of patent libraries relative to the average number of patents around matched federal depository libraries. In panel (a), we use the number of patents that contain a word that is new to the region around the library, but not to U.S. patenting overall as the numerator of the dependent variable. In panel (b), we use the number of patents that contain words that are not entirely new to the region, but only appeared in the region after patent library opening. In panel (c), we use the number of patents that contain a word that is new to U.S. patenting overall. In panel (d), we use the number of patents that do not contain a word from the previous three categories. Because of data availability, we restrict our analysis to patents filed in or after 1975 in all panels. The 95% confidence intervals are based on bootstrapped standard errors. We use the weights of Iacus, King and Porro (2012) to arrive at the average treatment effect on the treated. We assign each patent library and all Federal Depository Libraries within the same state and within 250 miles as control group. We exclude the patent library of Burlington VT.

VI. Conclusion

The grand bargain in the patent system is that inventors disclose their ideas in exchange for exclusive rights to market their inventions for a limited period. Courts and scholars argue that disclosure constitutes a significant benefit of the patent system, as it helps inventors to avoid duplication and gives them new ideas to recombine with their own. However, there is little evidence regarding whether the disclosure mandated by the patent system affects subsequent innovation. By leveraging geographic variation in access to patent prior art during a key period of expansion in the USPTO Patent Depository Library Program, we are able to shed light on this issue. Specifically, we document that the opening of patent libraries from 1975 to 1997 increased innovation in the regions in which those libraries opened. Consistent with the expectations of Machlup (1958) and Scotchmer and Green (1990), we provide evidence that the enablement effect resulting from the disclosure of knowledge contained in patents is quantitatively important for subsequent innovation.

While the literature on the role of disclosure in the patent system is growing, this is still a topic that is poorly understood. This paper adds to the available evidence and suggests that access to prior art is an important determinant of inventors' productivity. These findings complement those of other emerging research projects examining the impact of disclosure (including Gross, 2019; De Rassenfosse, Pellegrino and Raiteri, 2019; Hegde, Herkenhoff and Zhu, 2019), each of whose findings, though in quite different circumstances, are consistent with those we present here.

Considering the ease with which current researchers can review and distill patent prior art, the modern implications of our results speak more to the importance of ensuring that patent documents impart technical information and that it is possible to effectively search these documents. While modern information technology has made huge strides, investment in improved text analysis might make it possible to better classify the topic or measure the importance of a patent for a particular problem. Nevertheless, future research is necessary to isolate the role of disclosure further. In particular, understanding what exactly disclosure changes for future inventors is an important question that the current literature can only partially address.

In addition to providing evidence regarding a key question in the economics of intellectual property, our study contributes to the literature on research enhancing institutions. While economists generally agree that institutions that lower the costs of access to useful knowledge may support innovation (Mokyr, 2002), empirical research has provided few examples (Furman and Stern, 2011; Biasi and Moser, 2016; Waldinger, 2016; Andrews, 2019; Berkes and Nencka, 2019). In this work, we document the value of patent libraries as knowledge hubs whose operation contributes to follow-on innovation.

REFERENCES

Andrews, Michael. 2019. "How Do Institutions of Higher Education Affect Local Invention? Evidence from the Establishment of U.S. Colleges."

Arora, Ashish, Marco Ceccagnoli, and Wesley M Cohen. 2008. "R&D and the patent premium." *International Journal of Industrial Organization*, 26(5): 1153–1179.

Arts, Sam, Bruno Cassiman, and Juan Carlos Gomez. 2018. "Text matching to measure patent similarity." *Strategic Management Journal*, 39(1): 62–84.

Ball, Geoffrey. 2012. *No more laughing at the bad boy: A technological adventure between Silicon Valley and the Alps.* Haymon.

Balsmeier, Benjamin, Guan-Cheng Li, Tyler Chesebro, Guangzheng Zang, Gabe Fierro, Kevin Johnson, Aditya Kaulagi, Sonja Lueck, Doug O'Reagan, Bill Yeh, et al. 2018. "Machine learning and natural language processing on the patent corpus: Data, tools, and new measures." *Journal of Economics & Management Strategy*, 27(3): 535–553.

Baruffaldi, Stefano, and Markus Simeth. 2018. "Patents and Knowledge Diffusion: The Effect of Early Disclosure."

Berkes, Enrico, and Peter Nencka. 2019. "Knowledge Access: The Effects of Carnegie Libraries on Innovation." *working paper.*

Biasi, Barbara, and Petra Moser. 2016. "Effects of copyrights on science: Evidence from the WWII Book Republication Program." *Available at SSRN 2542879.*

Brown, William H., and James A. Arshem. 1993. *Survey of patent and trademark depository libraries 1991/92.* U.S. Patent and Trademark Office, Washington D.C.

Census Bureau. 2010. "Gazzetteer Files: Places File."

Chien, Colleen V. 2016. "Contextualizing Patent Disclosure." *Vanderbilt Law Review*, 69(6): 1849–1890.

Cohen, Wesley M, Akira Goto, Akiya Nagata, Richard R Nelson, and John P Walsh. 2002. "R&D spillovers, patents and the incentives to innovate in Japan and the United States." *Research Policy*, 31(8): 1349–1367.

Cohen, Wesley M., Richard R. Nelson, and John P. Walsh. 2000. "Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not)." *NBER Working Paper 7552.*

De Rassenfosse, Gaétan, Gabriele Pellegrino, and Emilio Raiteri. 2019. "The Disclosure Effect of Patents: Evidence from the 1951 Invention Secrecy Act." *Working paper.*

European Patent Office. 2016. "PATSTAT Global, Autumn Edition."

Furman, Jeffrey, and Scott Stern. 2011. "Climbing atop the shoulders of giants: The impact of institutions on cumulative research." *American Economic Review*, 101: 1933–1963.

Gambardella, Alfonso, Dietmar Harhoff, and Sadao Nagaoka. 2011. "The social value of patent disclosure." Mimeo.

Graham, Stuart, and Deepak Hegde. 2015. "Disclosing patents' secrets." *Science*, 347(6219): 236–237.

Gross, Daniel P. 2019. "The Consequences of Invention Secrecy: Evidence from the USPTO Patent Secrecy Program in World War II." *Working Paper.*

Hall, B. H., and D. Harhoff. 2012. "Recent research on the economics of patents." *Annual Review of Economics*, 4: 541–565.

Hegde, Deepak, and Hong Luo. 2017. "Patent publication and the market for ideas." *Management Science*, 64(2): 652–672.

Hegde, Deepak, Kyle Herkenhoff, and Chenqi Zhu. 2019. "Patent Publication and Innovation."

Iacus, Stefano M, Gary King, and Giuseppe Porro. 2012. "Causal inference without balance checking: Coarsened exact matching." *Political Analysis*, 20(1): 1–24.

Idris, Kamil. 2002. *Intellectual property: a power tool for economic growth.* Vol. 888, WIPO.

Jenda, Claudine Arnold. 2005. "Patent and trademark depository libraries and the United States Patent and Trademark Office: A model for information dissemination." *Resource Sharing & Information Networks*, 18(1-2): 183–201.

Khan, B. Zorina. 2005. *Democratizing Innovation.* Cambridge University Press: Cambridge UK.

Lamoreaux, Naomi R., and Kenneth L. Sokoloff. 1999. "Inventors, firms, and the market for technology in the late nineteenth and early twentieth centuries." , ed. Daniel M. G. Raff Naomi R. Lamoreaux and Peter Temin, Chapter Learning by Doing in Markets, Firms and Countries, 19 – 60. University of Chicago Press.

Landes, William M, and Richard A Posner. 2003. "Indefinitely renewable copyright." *The University of Chicago Law Review*, 70(2): 471–518.

Lee, William F, and Lawrence P Cogswell III. 2004. "Understanding and addressing the unfair dilemma created by the doctrine of willful patent infringement." *Houston Law Review*, 41: 393.

Lemley, Mark A. 2012. "The myth of the sole inventor." *Michigan Law Review*, 110(5): 709–760.

Levin, Richard C, Alvin K Klevorick, Richard R Nelson, Sidney G Winter, Richard Gilbert, and Zvi Griliches. 1987. "Appropriating the returns from industrial research and development." *Brookings Papers on Economic Activity*, 1987(3): 783–831.

Li, Guan-Cheng, Ronald Lai, Alexander D'Amour, David M. Doolin, Ye Sun, Vetle I. Torvik, Amy Z. Yug, and Lee Fleming. 2014. "Disambiguation and co-authorship networks of the U.S. patent inventor database (1975-2010)." *Research Policy*, 43(6): 941–955.

Li, Jie Jack. 2009. *Triumph of the heart: the story of statins.* Oxford University Press.

Machlup, Fritz. 1958. *An economic review of the patent system.* US Government Printing Office.

Machlup, Fritz, and Edith Penrose. 1950. "The patent controversy in the nineteenth century." *The Journal of Economic History*, 10(1): 1–29.

Mansfield, Edwin. 1986. "Patents and innovation: An empirical study." *Management Science*, 32: 173–181.

Mokyr, Joel. 2002. *The gifts of Athena: Historical origins of the knowledge economy.* Princeton University Press.

Morrison, Greg, Massimo Riccaboni, and Fabio Pammolli. 2017. "Disambiguation of patent inventors and assignees using high-resolution geolocation data." *Scientific Data*, 4.

Moser, P. 2011. "Do patents weaken the localization of innovations? Evidence from World's Fairs." *The Journal of Economic History*, 71(2): 43–74.

Moser, P. 2013. "Patents and innovation: Evidence from economic history." *Journal of Economic Perspectives*, 27(1): 23–44.

Mossinghoff, Gerald J. 1984. *Hearing before the Subcommittee on Courts, Civil Liberties and the Administration of Justice on the Committee on the Judiciary House of Representatives, Ninety-Eigth Congress, 1983.* US Government Printing Office.

Oliver, Jeanne, ed. 2002. *Celebrating 25 years of libraries and USPTO cooperation: 1977-2002.* United States Patent & Trademark Office.

Ouellette, Lisa Larrimore. 2012. "Do patents disclose useful information?" *Harvard Journal of Law & Technology*, 25(2): 546–602.

Ouellette, Lisa Larrimore. 2017. "Who reads patents?" *Nature Biotechnology*, 35(5): 421–424.

PatentsView. 2020. "PatentsView Database."

Rapino, Melanie A., and Alison K. Fields. 2013. "Mega Commuters in the U.S. Time and Distance in Defining the Long Commute using the American Community Survey." *Working Paper.*

Reid, T.R. 1985. *The chip: How two Americans invented the microchip and launched a revolution.* Random House.

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

Romer, Paul M. 1990. "Endogenous technological change." *Journal of Political Economy*, 98(5, Part 2): S71–S102.

Schmoch, Ulrich. 2008. "Concept of a technology classification for country comparisons." *Final report to the world intellectual property organisation (wipo), WIPO.*

Scotchmer, Suzanne. 1991. "Standing on the shoulders of giants: Cumulative research and the patent law." *Journal of Economic Perspectives*, 5(1): 29–41.

Scotchmer, Suzanne, and Jerry Green. 1990. "Novelty and disclosure in patent law." *RAND Journal of Economics*, 21(1): 131–146.

Sneed, Martha Crockett. 2000. "Fully disclosed yet merely descriptive: Intricacies of training the patent and trademark information professional." *Journal of Library Administration*, 29(1): 59–78.

Sneed, M Crockett. 1998. "125 years of patent information to the people: the US Patent and Trademark Depository Library Program."

United States Patent and Trademark Office. 1999. *Survey of Patent and Trademark Depository Libraries: 1997 Customer Satisfaction Survey. Survey of Patent and Trademark Depository Libraries: 1997 Customer Satisfaction Survey*, U.S. Patent and Trademark Office.

United States Patent and Trademark Office. 2003. *Survey of Patent and Trademark Depository Libraries: 2002 Customer Satisfaction Survey. Survey of Patent and Trademark Depository Libraries: 2002 Customer Satisfaction Survey*, U.S. Patent and Trademark Office.

Waldinger, Fabian. 2016. "Bombs, Brains, and Science: The Role of Human and Physical Capital for the Production of Scientific Knowledge." *The Review of Economics and Statistics*, 98(5): 811–831.

Williams, Heidi L. 2017. "How do patents affect research investments?" *Annual Review of Economics*, 9: 441–479.