# ‘Novel’ Ideas: The Effects of Carnegie Libraries on Innovative Activities

**Preliminary: please do not circulate**

Enrico Berkes Peter Nencka *

February 22, 2019

**Abstract**

We show that the historical rollout of public libraries increased the innovation output of recipient towns. Between 1886 and 1919, Andrew Carnegie donated $34.5 million (approximately $1 billion in 2019 dollars) to fund the construction of more than 1,500 public libraries across the United States. Drawing on a new data set based on original historical records, we identify cities that qualified to receive a library grant, applied for the program, received preliminary construction approval, but ultimately rejected Carnegie's offer. Using the rejecting cities as a control group, we estimate the effects of Carnegie library formation on patenting activity. We provide evidence that the trends in the patenting activity in the two groups are indistinguishable before the construction of the libraries and then diverge. Cities that accepted grants experienced both short- and long-run gains in patenting activity. We also describe ongoing work to estimate how library exposure during childhood affects long-run innovative potential.

*Department of Economics, The Ohio State University. Thanks to Joshua Hall, Lauren Jones, Kurt Lavetti, Dean Lillard, Trevon Logan, and Bruce Weinberg for their helpful comments on earlier versions of this work. We also thank participants at Ohio State and West Virginia University seminars for their insightful comments.

*And daily in the papers thou shalt read,*
*Of the new libraries, in cities vast,*
*In villages, and Indian wigwams too,*
*In Texas ranches and Esquimaux huts,*
*In Heaven, Hell, and stations in between*
– Upton Sinclair, undated poem

# 1 Introduction

There is a widespread consensus that differences in the rate of innovation and technological progress play an important role in explaining gaps in productivity, economic growth, and inter-generational mobility both across and within countries (e.g., Romer, 1990; 1996; Aghion and Howitt 1992). However, there is much to learn about the specific policies and institutions that promote innovative activity. In this paper, we study the effects of one of the largest local public institution rollouts in U.S. history: the construction of 1,687 public libraries in the late 19th and early 20th centuries. Andrew Carnegie financed this public library construction boom, inspired by his childhood experiences and his belief that libraries promote self-sufficiency. In 1920, nearly half of all public libraries in the United States had been built with funds donated by Carnegie (Jones, 1997).

We show that the historical spread of libraries plausibly affected innovative activity through both information and community-building mechanisms. For recipient cities, a Carnegie library was a new source of scientific, technical, and practical knowledge. Crucially, public libraries were accessible by most citizens regardless of income or social status.[1] Evidence suggests that access to technical information increases innovative activity (e.g., Graham and Hegde, 2015; Hegde and Luo, 2017; Furman et. al 2018; Biasi and Moser 2018). Beyond providing access to knowledge, libraries also served an important social role. They were—and in many cities remain—the only indoor space freely available for community gatherings. Previous work highlights the importance of open, often informal, communication in the innovative production process (Allen, 1983, Campos et. al, 2018; Andrews, 2017).

Research also shows that the characteristics of childhood neighborhoods matter for long-run

[1] An important exception—to which we return later—was in some segregated cities in the US South.

outcomes (Chetty, 2018; Chyn, 2018) and that the presence of a local college during childhood may change youth trajectories (Currie and Moretti, 2003). Our focus is public libraries, which both today and in the past explicitly aim to positively affect childhood development. Motivated by this intuition, we test whether library exposure in childhood increases individual-level propensities to innovate or become a scientist, including among children who later move away from their hometowns.

Our analysis proceeds in two steps, focusing on city and individual-level outcomes, respectively. First, we estimate how public library entry affected innovative activity, proxied by patenting output, in a given town. We construct a unique historical data set of geo-located, city-level patents linked to the rollout of public libraries. We identify the causal effect of the arrival of a new library on innovative activity and human capital development using a historically motivated, novel control group. More than 200 cities that qualified to receive a library grant, applied for the program, and received preliminary construction approvals ultimately rejected the grant. Rejecting cities follow statistically indistinguishable patenting trends prior to library entry. We provide historical and statistical evidence that cities that rejected Carnegie libraries often did so for reasons orthogonal to the economic activity of the town, including the political unpopularity of Carnegie himself.

Our main results, illustrated in Figure 1, show that the patenting activity in towns that accepted Carnegie libraries sharply increased after the library was granted relative to control cities. After following statistically indistinguishable patenting trends for many years, the trends start diverging shortly after the receipt of a library. Patenting differences between cities that accepted and rejected libraries peak between 20 and 30 years after acceptance and persist for decades. Patenting in cities that accepted Carnegie libraries increased by approximately 2 patents per year, or about 20% relative to the sample mean. The results are robust to controlling for a wide range of socio-economic factors, including city size.

We study the differential impact of access to libraries on innovative activities by separately estimating effects across technology classes. Our estimates suggest that libraries mainly affected patenting in the practical trades, such as farming, construction, and mechanical engineering. This is in line with historical records of the books that libraries commonly carried. Throughout this time-period, the American Library Association published a list of the books recommended to libraries. The list provided an authoritative starting point for new libraries, as many were established in towns with limited library experience and staffing. Consistent with our results, many of the titles in the list were technical/reference books (e.g., *Modern Machine Shop Tools, Their Construction, Operation and Manipulation*) or magazines (e.g., *Scientific American*). If information is a mechanism through which patenting increases in treated cities, patents in treated cities may be more likely to cite books. To test this, we compile a list of keywords commonly used to cite books and use full-text patent search to identify patents that contain these keywords. Our preliminary results indicate that library entry induces more book citations, both in absolute number and conditional on overall patenting activity. The later result is particularly strong in the first 20 years after library entry. This suggests that increased information is a plausible mechanism for the overall increase in patents that we observe.

We also test whether the average patent quality changed after library entry. Patents are an imperfect measure of innovative activity, in part because some patents do not represent new knowledge.[2] Following Kuhn, Younge and Marco (2017), we construct a text-based measure of patent scope using the wordcount of the first patent claim. Unlike other measures of patent quality (e.g., forward citations), this scope-based proxy can be consistently applied throughout our estimation sample. Our results indicate that the average quality of patents did not decrease after the introduction of the libraries. If anything, the average patent scope was broader in cities with Carnegie libraries, suggesting that patents became more valuable.

In the second part of this project—which is still a work in progress and not included in this draft—we estimate the effects of library exposure during childhood on the probability of becoming an inventor or scientist. Because innovators are often mobile, focusing on city-level outcomes alone may understate the effect of public libraries on innovation outputs. Linking the 1900, 1910, 1920, and 1940 full count U.S Census data across decades, a dataset of historical patents, and historical records of scientists and scientific production (e.g., doctoral dissertations), we will estimate the long-term effects of library formation on human capital accumulation.

[2] It is also true that not all innovative activity is patented. In future work, we anticipate estimating the effect of libraries on other proxies for local innovative activity.

This paper contributes to the literature on the determinants of scientific and innovative output. We provide new evidence on how access to additional information affects the production of innovation. Graham and Hedge (2015) and Hedge and Luo (2017) study how innovation is affected after recent policy mandated that patent applications be disclosed at an earlier date after filing. They find that earlier disclosed patents are more likely to be cited and licensed, and that increased disclosure appears to facilitate intellectual property transactions. Furman et. al (2018) find that local patenting behavior increases after the opening of nearby patent depositories. Relative to this work, we focus on informational effects that are generated from non-patent sources.

We also contribute to a series of papers that show how information shocks in the early to mid-20th century affected innovative activity. Biasi and Moser (2018) find that price declines generated from stripping copyrights from German books in 1943 led to an increase in the probability of citing these books in scientific articles, patents, and PhD theses. Gross (2019) finds that patents that were made secret for national security reasons during World War II for a lengthy period of time are cited less than patents that were secret for shorter periods of time. He also finds that new concepts introduced in secret patents were less likely to be referenced in the text of books after the end of World War II. Iaria et al. (2018) find that shocks to international scientific cooperation during World War I led to productivity declines among domestic scientists who relied on foreign research. We contribute to this literature by exploiting a distinct shock to information access that operates at the local level.

Because libraries serve local communities, our results relate to prior work on local institutions and innovative activity. Previous studies in this literature have mainly focused on colleges (e.g., Moretti, 2003; Furman and MacGarvie, 2007; Aghion et. al., 2009; Andrews, 2018). These papers consistently find that after the establishment of a college, innovative activity rises, although they disagree on the channels through which this effect manifests itself, as well as its magnitude. To the best of our knowledge, we are the first to estimate the effect of public libraries on local innovation and the propensity to become an innovator or scientist. We view this as a particularly important topic because public libraries plausibly played a distinct role in disseminating information in the early 20th century. Most inventors had no access to universities and worked alone or in small teams without a formal affiliation. Libraries provided low-cost, state-of-the-art information across disciplines, as well as information on the patenting process itself.

Last, we contribute to a literature on the history of public libraries in the United States. Beyond the work of library historians (e.g., Bobinksi, 1969; Jones, 2007), we build on the literature on the political economy and development of libraries in the early 20th century. Kevane and Sundstrom (2014) outline the characteristics that predicted local library entry, including the positive impact of state library associations (Kevane and Sundstrom, 2016a). Kevane and Sundstrom (2016b) estimate the effect of library entry during the 20th century on short-run political participation using an interrupted time-series strategy. They find no clear relationship between library entry and participation in the following election. We expand this work by focusing on Carnegie libraries and studying long and short-run human capital and innovation outcomes. In ongoing work that began contemporaneously with this project, Karger (2019) estimates the effects of public library expansion on human capital and occupation choices. We view our projects as complements given our focus on the right tail of the human capital distribution: inventors and scientists.

The remainder of the paper is organized as follows. Section 2 provides historical context for the rise of public libraries in the U.S, describes the details of Carnegie's library program, and explains why some cities rejected library offers. Section 3 describes the construction of the library and patent data. Section 4 describes the empirical strategy and presents preliminary city-level results. Section 5 presents discusses potential mechanisms. Section 6 concludes.

# 2 The rise of public libraries and Andrew Carnegie

While libraries have ancient roots, their current form as local, public spaces open to all residents and supported by taxpayers is relatively new. The library as a space to store written material emerged simultaneously with the discovery of writing during early civilizations. The ancient Greeks, Romans, and Byzantines all built and commonly used libraries as archival storage (Murray, 2009). The Great Library of Alexandria in Egypt reportedly held scrolls equivalent to 100,000 books in 100 BCE

(MacLeod, 2004). During the Middle Ages, smaller libraries in monasteries served an important book-preserving role. At the same time, royalty and the the upper class cultivated extensive private book collections. Academic libraries—which have distinct missions as university-supported entities—existed in Europe from at least 1500.

It was not until the 18th century that non-academic libraries, focused on providing service for nearby residents, emerged as town staples in Europe and the U.S. These early local libraries differed from their modern counterparts. They were often not free nor open to all (Shera, 1949). These early, local libraries can be categorized into two broad types: social and subscription libraries. Social libraries were extensions of pre-existing associations. These included women's leagues, churches, or even drinking clubs at taverns (Weigand, 2015). Social libraries stocked books that appealed to their membership and were often not open to the public. Subscription libraries were explicitly commercial, open to most with a fee. To attract consumers, subscription libraries mostly stocked popular fiction, including the relatively new and increasingly popular novel format. The growth of social and subscription libraries was rapid in both the U.S. and the UK and paved the way for publicly funded endeavors. But access costs and commercial pressures limited their reach and effectiveness at disseminating scientific knowledge (Weigand, 2015).

In 1833, the small town of Petersborough, New Hampshire established the first US library open to all citizens and supported by town tax dollars. New Hampshire, Maine, and Massachusetts passed laws in the early 1850s authorizing local library taxation. These laws helped spread public libraries to Northwest cities. The first large US city to form a municipal library was Boston in 1852. Despite a growing movement in favor of public libraries, their diffusion during the rest on the 19th century was slow. Both the financial pressures of the Civil War and the enduring popularity of the subscription model slowed the demand for publicly funded libraries. An important turning point was the 1893 World Fair in Chicago. There, the American Library Association (ALA)—an interest group of librarians founded in 1876 (ALA, 2019) that advocates the spread of municipally funded libraries—showcased a demonstration public library with 5,000 books. Spearheaded by ALA president Melvil Dewey, of decimal system fame, the exhibit attracted national attention (Sharp, 1893; Weigand, 2015). The ALA subsequently published a list of 5,000 recommended books for new libraries, which they updated throughout the 1940s. The popularity of this exhibit helped fuel the local demand for public libraries.

In 1893, at the time of the ALA exhibit and at the cusp of the library revolution, the US had approximately 800 public libraries, almost exclusively located in the Northeast and California (Jones, 1997). Many of these libraries were not free-standing, but instead located in the basements or attics of pre-existing buildings. In Malta, Montana, the library was located on the balcony of a drugstore; in Dunkirk, New York, in the basement of a hospital; in Marysville, Ohio, in the horse stall of a fire department (Bobinski, 1968). Less than 30 years later in 1919, the US would have 3,500 public libraries in rural and urban locations across the US, most in free-standing buildings and many occupying the largest building in town. Over half of the construction of these new libraries were entirely funded by one person: Andrew Carnegie.

## 2.1 The Carnegie construction program

Andrew Carnegie's library construction program is one of the largest acts of philanthropy in U.S. history. From his first grant in 1886 (Allegeny, New York) to his last grant in 1919 (a branch library in Philadelphia, Pennsylvania), Carnegie fully funded the construction of 1,687 public libraries across the US at the cost of $1.2 billion in 2018 dollars. Carnegie himself was a complicated character, a successful and controversial steel tycoon and the single largest philanthropist of his time. The different perception of these two sides of his character—and particularly whether they could be consistent—affected how the public viewed his gifts. We discuss this in further detail in the next section when we discuss the reasons behind the refusal of Carnegie's grant by some cities. In this section, we describe the basic outlines of Carnegie's library philanthropy.

Carnegie's stated motivation for the library grants is consistent with his larger views on philanthropy. He viewed public libraries as a way citizens could improve themselves if they had the drive to do so. This motivation is evident in his famous 1899 essay on philanthropy, "The Gospel of Wealth:"

> In bestowing charity, the main consideration should be to help those who will help themselves; to provide part of the means by which those who desire to improve may do

so; to give those who desire to rise the aids by which they may rise; to assist, but rarely or never to do all.

Carnegie's library program started small. Carnegie himself referred to two distinct periods of his library philanthropy, the “retail" and “wholesale" phases. In the “retail" phase, Carnegie gave money to build eight libraries in selected communities. These were located mostly in his adopted homes of Pennsylvania and New York, where many of his business interests were located (Bobinski, 1969). By 1898, Carnegie shifted his priorities to providing library access for as many people as possible. He opened the library application process to essentially all cities that did not already possess a stand-alone, self-sufficient library. This opened the floodgates to city-applicants, a trend well-noticed in the newspapers of the time (see Figure 2).

Given the scale of the eventual construction, the Carnegie library application process itself was surprisingly informal. It was entirely conducted by mail. James Bertram, Carnegie's private secretary, managed the application process throughout the program. After the “retail" phase, Carnegie himself was rarely involved in the details of each granted library. Bertram and other staff would pre-approve library applications, presenting bundles to Carnegie for his signature. (Bobinski, 1969) Almost everything we know about the library program comes from Bertram's letters to cities, which were catalogued and maintained by the Carnegie Corporation and New York University.[3]

Typically, the application process started with a letter of interest from the town. Often initial letters to Bertram came from everyday citizens or leaders of civic groups. Bertram would reply and note that he was happy to hear about their interest in libraries but that future correspondence should occur with elected city representatives. Bertram instructed cities to fill out a short form, which asked for information on the city population, the names of city officials, whether the city already had a public library, and if so whether it had a freestanding building, its expenses, and its circulation. This step was mostly to ensure that cities understood what Carnegie was willing to supply (money to construct a new public library) and identify whether such a library already existed. Almost all cities that applied had

[3]We have reviewed a selection of this letters, mainly related to cities that rejected Carnegie libraries. We cite them as the Carnegie correspondence archives throughout the remainder of the discussion

need and progressed to the next stage.[4]

If Bertram and his assistants assessed that the request was legitimate, the application moved to the next stage of the process. Accepted libraries received a short letter from Bertram, like this one to Charleston, Illinois, reported in Jones, (1997):

> Dear Sir:
>
> Mr. Carnegie has considered yours of Aug. 23, and if Charleston will furnish a suitable site and pledge not less than twelve hundred dollars a year for support of library, Mr. Carnegie will be glad to give twelve thousand dollars for a Free Library Building.

The letter highlights the award amount that Carnegie judged needed to fully construct the library. It also outlines the main requirements of Carnegie's grant that can be summarized in four main points:[5]

1. **The granted amount was determined by Carnegie and Bertram.** With very limited exceptions, Carnegie and Bertram decided the exact grant amount based on reported population, at approximately $2-3 per person. The exact grant amount could be controversial. Cities often felt that they were entitled to additional funds and tried to press this point to Bertram. A common tactic was noting that the census figure was out of date. A few cities noted that they expected that their library would draw attendees from beyond their city limits. These protests rarely succeeded (Bobinski, 1969).

2. **Carnegie libraries needed to be free and public.** Carnegie's grants were for "free" public libraries. He did not want to give money for social or subscription libraries.

3. **The land needed to be provided by the city.** The cities needed to supply a construction site for the library. Carnegie required that either the city purchased a site or re-purpose existing city

[4]As noted in Bobinski (1969), Carnegie did reject requests at this stage for state, subscription, and historical society libraries. Carnegie also funded the construction of 108 academic libraries during this period. We do not focus on these libraries in this paper, as they have a distinct mission from public libraries.

[5]For more details on the structure of the program, we recommend Bobinski (1969) and Jones (1997). After 1908, Carnegie began to impose more requirements on the specific construction techniques and floor plans that libraries could use. This occurred after a number of towns tried to combine libraries with other civic buildings that Carnegie was not interested in funding, like gyms or city halls.

property. Bertram asked cities to send proof of site ownership before the funds were dispensed. Bobinski (1969) estimates from his review of the Carnegie correspondence that one-in-three cities had some sort of controversy about the site locations. Because libraries often became town centerpieces, it is unsurprising that citizens fiercely argued in favor of their preferred locations.

4. **Cities were required to commit funds for ongoing maintenance of the libraries.** Carnegie required that cities themselves also contribute to the ongoing maintenance of their libraries. Carnegie knew that building the library was not enough. He provided funding for construction, but wanted to make sure that the cities could fill the libraries with books, pay the staff, and mantain the building. His solution, as illustrated in his letter, was to require that cities pledge to spend 10 percent of the initial construction grant on annual library upkeep. Practically, this 10 percent maintenance requirement was at the lower end of what would be required to staff and maintain a library in the early 20th century. Cities often had to allocate additional funds beyond the 10 percent to keep their libraries running, particularly as average city-sizes grew throughout the 1920s and 1930s (Bobinksi, 1969).

   Despite the written pledge, once a library was built, Carnegie had little ability to enforce the 10% contribution requirement. This appears to have been well-known to applying cities, and there is ample historical evidence of cities failing to meet the 10% requirement. In 1917, the Carnegie Corporation—which by this point had been entrusted to manage the library program and related philanthropy—sent a survey to investigate reports that the pledge was not being met. The results were stark: in Ohio, for example, 23 out of 77 cities were not meeting the pledge (Bobinski, 1969). This led to a brief suspension of library grant-giving to Ohio, but no direct action against the offending libraries themselves.

Despite his desire for libraries to be free and open, Carnegie made no requirement that Southern communities integrate their library services. Letters suggest that Bertram occasionally communicated and asked if the town planned to build a separate library for blacks, but our review of the Carnegie communication indicates that Carnegie never refused library services to cities based on plans to construct segregated libraries or whites-exclusive libraries.[6] Despite this permissive stance, Southern communities were less likely to request libraries. This is likely due, at least in part, to a mistaken belief that Carnegie would force them to integrate.

## 2.2 Reactions to Carnegie and his libraries

Carnegie libraries were often met with praise. Indeed, communities that were happy with their libraries could create a cascading effect within a state, as nearby cities rushed to join in on the construction. But some Carnegie grants were controversial. A large number of cities that would have qualified to build a library never applied. Even among cities that applied and were "accepted", 209 ultimately rejected Carnegie's offers and did not build a library. Representing approximately 15% of cities that were offered grants, this was a notable rejection of Carnegie's generosity. Throughout the rest of the paper, we refer to these 209 cities as "rejecting cities". This section describes the cities that rejected Carnegie's library and the reasons that motivated their rejection. We argue that these cities are a good counterfactual for accepting cities in our empirical analysis.

To understand why cities would reject a library in the first place, it is important to understand Carnegie's reputation. Before his library philanthropy began, Carnegie was already well-known in households across the United States. As chairman and founder of Carnegie Steel, Carnegie was one of the richest men in America. In this context, Carnegie was reviled for his labor practices in many households.

The key event that generated Carnegie's long-term negative reputation was the steel worker strike at Homestead, PA, in June 1892. After months of rising tensions in the face of increasing production demands by Carnegie's managers, workers finally struck. Carnegie was determined to defeat the union. He locked the union workers out, but soon the rest of the plant workforce agreed to join the strike. Carnegie's managers hired a private militia to break the strike and take back the town. The resulting battle lead to the deaths of nine strikers, ten members of the militia, and scores of wounded. The battle made national news (see Figure 3). Carnegie's actions were never forgotten by the labor movement,

[6] The correspondence between Bertram and Richmond, VA is an excellent example of Bertram asking but not appearing to require an additional library for blacks. (Carnegie Correspondence)

many of whom later became involved in the fight against libraries.

A 1909 editorial in the *Pittsburg Kansan* reproduced in Jones (1997) illustrates the long-lasting impact of Carnege's unpopularity among labor in library debates:

> A library that is built on money wrung from the hearts and homes of Homestead miners who were shot down in cold blood...is not fitting monument for the kind of men that built Pittsburg. If Mr. Carnegie wants to be charitable, let him commence with the widows and orphans of the murdered

In Wheeling, WV, which ultimately rejected Carnegie's offer of funding, a union leader declared that "[i]n view of Mr. Carnegie's attitude toward labor it is the duty of organized labor to adopt stringent measures to defeat the erection of this disgraceful monument." (1901 *Electrical Worker* via Jones, 1997) In Detroit, opposition to Carnegie was fierce, with the city treasurer quoted as saying "We ought to take care of ourselves...[not] accept a big chunk of money as a gift from a man who has made his money the way Carnegie did" (Krass, 2011). Such opposition was not limited to local officials. Prominent national politicians and writers, including socialist and recurring Presidential candidate Eugene V. Debs and *The Jungle* author Upton Sinclar, spoke out against accepting libraries. Even Samuel Clemens (better known today as Mark Twain), weighed in, noting that Carnegie's strive for personal recognition might be behind his generosity: "He bought fame and paid cash for it" (Bobinksi, 1969).

Opposition from the political left and labor unions was only one obstacle standing between cities and their libraries. The 10% yearly pledge—despite being approximately the minimum investment needed to support a library and largely unenforceable—was unpopular. In order to avoid the 10% pledge, some cities rejected Carnegie in favor of offers from local philanthropists.[7] Some cities could not secure or decide on a library site, eventually forfeiting their application. For a number of rejecting cities, the historical records do not offer clear guidance on why they rejected a library despite being accepted in the first place. In the data section, we describe rejecting cities in more detail and we

[7] In the empirical analysis, we report the estimates obtained both including and excluding these cities in the control group. The results are robust. While many of these cities did build a library, the free and public requirements of the Carnegie grant and the 10% funding pledge create a plausibly distinct treatment.

compare their observable characteristics with those of their library-accepting peers.

Carnegie's grant giving program ended in 1917. In 1915, Alvin Johnson, a Cornell University economics professor, was tasked by the Carnegie Corporation with conducting a cost-benefit analysis of the program. Johnson visited over 100 Carnegie libraries and wrote a detailed report that he presented to the Carnegie trustees (Johnson, 1915). Johnson concluded that while aspects of the Carnegie program had worked well, the lack of a centralized Carnegie organization and planning body had hindered library development. He recommended that the Carnegie Corporation hire grant-managers and local field staff across the country to help manage local affairs. Perhaps partially based on this recommendation—which the Corporation deemed too intrusive and expensive at the peak of World War I—the program was halted in 1917 (Bobinski, 1969).

# 3 Data

## 3.1 Location of accepted and rejected libraries

We construct a dataset of all Carnegie libraries using historical records collected by Bobinski (1969) and Jones (1997). Both Bobinksi and Jones compile their lists from the original Carnegie library program correspondence and surveys of libraries. In the handful of cases in which these two sources disagree, the locations identified in Jones are given the priority. We have also completed a careful review of the original Carnegie correspondence and news archives to search for any records that these authors might have missed—we did not identify any additional libraries. We assign each library to the city where the library was built. In some cases, this assignment might understate the reach of library services. In particular, Carnegie encouraged small towns ($<$1,000 people) to submit joint bids. We focus on outcomes in construction cities to simplify the analysis, with the understanding that the effects that we observe might understate the actual impact of libraries on the local innovative activity. Besides the construction city, we also record the grant amount and the grant date for each Carnegie offer.[8]

[8] A small number of cities (approximately 50) that received or rejected a Carnegie library cannot be matched to either patent data or demographic variables. We currently exclude those cities from the analysis. In the future, access to decennial

To cities that needed them, Carnegie funded multiple libraries. These awards typically paid for the construction of the main library in a city, as well as branch libraries. Multi-library grants occurred in approximately 5% of recipient cities and accounted for roughly 200 of the 1,687 constructed libraries observed in our time period. In multi-library cities that received libraries at different points in time, we assign the city-level grant year as the first year that a Carnegie library was granted.[9] We use grant years in our analysis since they can be uniformly interpreted across cities. Research on a selection of libraries indicates that the time required to build a library after being awarded grant varied, but that construction was typically completed within one year. We exclude a small number of library grants near areas that already had a well-developed library system, such as Cook County in Illinois. Carnegie grants in these areas were often distinct and came with additional requirements not found in the rest of the "wholesale" grants given at this time.[10]

To identify cities that rejected Carnegie library grants once they were approved, we rely on Bobinksi (1969). Bobinski identified over 200 libraries "that never materialized" in Table 13 of his history of Carnegie libraries. His primary source is the original Carnegie library correspondence between Bertham and rejecting cities, which we have also requested and personally reviewed. In addition to the locations of the rejecting cities, Bobinksi identifies the grant amount and the date of the offer. He also identifies a handful of cities that rejected the library because they had competing offers from local philanthropists. Our results are robust when we exclude these cities from the control group.

## 3.2 Patent data

We use patents data as a proxy for innovative activity in a certain city and year. Although patents have been shown to be an imperfect measure of innovation, they offer multiple advantages that make them an appealing data source for our analysis. First, patents provides a coherent snapshot of technological output at a certain point in time spanning more than two centuries. Second, the U.S. Patent and Trademark Office (USPTO) associates to each patent a set of technology classes that are consistent over

Censuses, as well as reviewed patents data might allow us to bring these cities back into the analysis.

[9]Our results are robust to alternative choices, including assigning the median granted year.

[10]Our patenting results are not sensitive to excluding these cities.

time and that provide important information about the knowledge content of the underlying invention. Third, text analysis allows us to extract information about the patent's scope, as well as the references cited. Finally, patents offer a wide range of information that is missing from other sources used in the literature (e.g., manuals) such as the name and city of residence of each inventor who worked on the invention and the assignee(s) who sponsored it. This piece of information is particularly important for our paper, since we are interested in studying the impact of public libraries on the innovation output at the local level.

Data on historical patents are collected from the Comprehensive Universe of U.S. Patents (CUSP).[11] In particular, we collect information on the technology class, inventor names, filing year, inventors city of residence, the number of forward citations, and the word count of the first patent claim for the close-to-universe of patents filed between 1870 and 2015.

We create a longitudinal dataset of the count of patents by filing year in each city that accepted or rejected a Carnegie library. We assign each patent to the city of the first listed inventor, as is common practice in the literature. Each patent is associated to a single International Patent Classification (IPC) technology class based on the main class assigned by the USPTO and a frequentistic mapping between the U.S. and international classification systems. For each patent, we calculate two measures of patent quality, namely the average number of forward citations and the word count of the first patent claim. The first is a measure of patent's impact, while the latter is a proxy for the scope of the patent. The idea is that a longer claim on average reflects a more specific—and therefore less valuable—patent (Kuhn, Younge, and Marco, 2017).

To test for potential mechanisms, we identify a list of keywords commonly associated with book citations in patents using a training sample of patents that cite books. We then search the corpus of patents for those phrases to identify other patents that likely cited books.

[11]We refer the reader to Berkes (2018) for a full description of the data and Andrews (2018) for a comparison of the CUSP with other existing data sources.

### 3.3 City and county covariates

We construct city and county-level covariates from historical Census data and related sources. We use each city's time-varying population collected in a consistent format by Erik Steiner and Jason Heppler.[12] For other standard covariates, including sex shares, race shares, average ages, share of the population enrolled in school, and the occupation and industry of employed workers, we use the 1900 Census micro-data aggregated at the county level. This is the finest level of geographical aggregation that allows us to construct covariates for our sample. We are currently processing and standardizing restricted-access historical census data, which will allow us to use additional city-level covariates in future versions of this work. To proxy for places that might be particularly hostile to Carnegie and his library grants, we calculate the fraction of each county's laborers in the mining industry in 1900 using Census occupation responses. The census did not ask for earnings until 1950. To calculate a proxy for county-level earnings we use Saavedra and Twinam's (2018) predicted earnings based on state, sex, age, race, occupation, and industry. We aggregate these values to create a county-level income score in 1900.[13]

### 3.4 Summary statistics

Figure 4 shows the cumulative count of Carnegie grants by grant year. As discussed in the previous section, Carnegie's library granting program began before 1890 but did not fund a significant number of libraries until the "wholesale" period near 1900. The number of granted libraries rapidly increased between 1899 and 1903 before levelling off into a more steady rate of grants through 1917, when the last grant was disbursed.

Figure 5 shows the geographic distribution of Carnegie library libraries across states. As the map indicates, Carnegie's program was national. Almost every state received at least one Carnegie library.

[12]See https://github.com/cestastanford/historical-us-city-populations/ for a full description of the data. A number of cities do not have consistent population data. This is particularly an issue for smaller cities. We anticipate augmenting this data via the historical census in our next revision.

[13]We use the 1900 data since it is the most complete set of Census records available before the majority of libraries were built (see Figure 4). The 1890 Census records were lost in a fire. Results using 1880 Census covariates are similar and available upon request.

Despite the national outreach, some geographical patterns can be identified. Figure 5 illustrates that granted Carnegie libraries were mostly located in the Midwest and Northern states.[14] Figure 6 shows a similar map for rejected Carnegie libraries by state. As with cities that received a library, rejecting cities are located across the United States. Cities in Southern states were more likely to reject libraries conditional on receiving an offer. For that reason, throughout the rest of the paper, we confine most of our discussion to *within*-state comparisons of accepting and rejecting cities.

Figure 7 illustrates a comparison of rejected and accepted Carnegie libraries across observable county characteristics in 1900. We plot the coefficient associated to an indicator variable for building a Carnegie library from a regression on standardized versions of each indicated covariate conditional on state fixed effects. The covariates include population, the share of women, the average age, the share of blacks, the average predicted earnings score, the share of the population currently enrolled at school, and the share of the workforce currently working in the mining industry.[15] Figure 7 indicates that conditional on state, places that accepted and rejected Carnegie libraries were similar on observable characteristics in 1900—the estimated magnitude of differences across rejecting and accepting areas is small, and statistically indistinguishable from zero.[16] This preliminary evidence suggests that rejecting and accepting locations were similar to each other across many dimensions.

# 4 Empirical analysis

For the empirical analysis, we employ a standard difference-in-difference approach to estimate the effect of libraries on patenting behavior. We compare changes in patenting before and after Carnegie grants in cities that did build a library and those that were deemed eligible but eventually declined the grant. Our identification assumption is that in the absence of library construction, cities that accepted and rejected libraries would have followed similar patenting trends after the approval dates. As in most difference-in-difference settings, we cannot directly test this assumption, since we do not observe

[14]While Southern states had fewer people, this result also holds on a per capita basis.

[15]Each covariate is standardized to have mean zero and standard deviation 1 so that they can appear on the same scale. This standardization does not affect the interpretation of the results.

[16]Across states, places that accepted Carnegie libraries have slightly larger populations, lower shares of blacks, and higher occupational income scores in 1900.

counterfactual patenting in cities that accepted libraries nor rejected them.

In Figure 1, we plot the mean patenting output within our treatment and control samples forty years before and eighty years after library grants. More precisely, we show the average count of patents per city in each year before and after library grants with the granting year normalized to zero.[17] These results indicate that cities that built and rejected Carnegie libraries followed statistically indistinguishable parallel trends prior to library grants. The trends start diverging shortly after library receipt. Patenting differences between cities that accepted and rejected libraries peak between 20 and 30 years and slowly converge thereafter.

## 4.1 Difference-in-difference estimation

We estimate the magnitude of the effect of a new library on the local patenting activity of a city, with a difference-in-difference regression model. Formally, we estimate:

$$PatentCount_{ist} = \beta_1 Library_i + \beta_2 Post_{t,i} + \beta_3 Library_i * Post_{t,i} + X_i + \delta_s + \gamma_t + \epsilon_{ist} \quad (1)$$

where $PatentCount_{ist}$ is the count of patents in city $i$, state $s$, and year $t$; $Library_i$ indicates cities that were approved for a library and actually constructed the library building; $Post_{t,i}$ is a dummy variable that takes value 1 in the years after a city received a library offer; $X_i$ is a vector of covariates collected from the 1900 Census; and $\delta_s$ and $\gamma_t$ are state and year fixed effects, respectively. The error term is $\epsilon_{ist}$. The coefficient of interest is $\beta_3$ which identifies the incremental increase in patenting in cities that were offered and built a library relative to cities that were offered but did not build a library in years after grants were made. We estimate (1) using a balanced panel of cities 30 years before and 80 years after library grants. Therefore, $\beta_3$ can be interpreted as the weighted treatment effect over all post-grant years. Here and throughout the remainder of the paper we cluster standard errors at a city level.

Table 1 reports the results from estimating equation (1). Column 1 includes 1900 Census covariates and both state and year fixed effects. The remaining columns present estimates for less saturated

[17]We adjust these predictions to remove year effects to make the graph easier to interpret.

models.[18] The estimates in Table 1 indicate that patenting in Carnegie libraries cities increased by about 1.7 patents more per city-year with respect to cities that did not accept Carnegie libraries. Estimates are stable and statistically indistinguishable when adding fixed effects and 1900 covariates.[19]

## 4.2 Results by patent classes

Access to library information might not affect all types of patenting behavior equally across technological classes. In fact, the ALA's book guide (1904) for new libraries included many practical "how-to" books for trades and agriculture. Intuitively, it seems plausible that public libraries have a smaller impact on the most technical inventions that require a large amount of human (and possibly physical) capital. Even in the early 1900s, the most updated scientific material was likely to be only available in research libraries and at universities. However, some of the most critical inventions in the 19th and 20th century were built on fairly well-known basic science.

To study the differential effect of access to libraries on patenting behavior, we separately estimate effects across patent classes. To do so, we focus on patent classifications, which identify the primary industrial application of each patent.

Our estimates, reported in Table 2, suggest that libraries mainly affected patenting in the practical trades, such as farming, construction, and mechanical engineering. This is in line with historical records of the books that libraries commonly carried. In particular, Table 2 shows that the largest patenting effects are found in human necessities, performing operations/transport, constructions, and mechanical engineering. Human necessities includes agriculture, clothing, and domestic innovations like brushes and kitchen appliances.[20] This provides suggestive evidence that patenting in classes that overlap with library collections increased differentially more in cities that accepted Carnegie libraries.

[18]Results for the non-shown combinations of fixed effects and Census covariates are similar and available upon request.

[19]In results not presented here but available upon request, we have also estimated similar models that control for city-level fixed effects. The interaction term Got Library x Post is identified in these models. The results are statistically indistinguishable from the coefficients in Table 1. For example, the equivalent Got Library x Post coefficient estimate is 1.69 with a standard error of 0.79.

[20]A full taxonomy of each of the eight categories can be found at https://www.wipo.int/classifications/ipc/en/

## 4.3 Patent quality

Not all patents have the same innovative content and even among the most innovative ones, their real-life impact and value can differ significantly. To analyze whether library had any impact on the ultimate value of the inventions produced in a certain city, we estimate the effect of library entry on average, city-level patent quality. To measure quality, we rely on the claim-based word count measure described in the data section. In particular, we extract the word count of the first claim of each patent. Claims are the legally binding statement of the monopoly restrictions granted by the patent. Patent examiners review claims closely. Extensive claim revisions are commonly required to distinguish proposed claims from prior art. Often these revisions include adding additional information to the claim; a short, very broad claim is unlikely to pass the novelty test. Shorter claims therefore reflect on average broader and more valuable patents. We estimate an analogue to equation (1) with average claim word count in city *i*, state *s*, and year *t* as the dependent variable:

$$AvgClaimLength_{ist} = \beta_1 Library_i + \beta_2 Post_{t,i} + \beta_3 Library_i * Post_{t,i} + X_i + \delta_s + \gamma_t + \epsilon_{ist} \quad (2)$$

Table 3 includes the parameters estimates from this regression model. Our results indicate that the introduction of a Carnegie library did not cause large changes in the average first claim word count of published patents. If anything, libraries appear to be associated with shorter—and therefore broader and more valuable—patents.

## 4.4 Alternative control groups

Carnegie grants were rejected for multiple reasons. One important reason was that another philanthropist could step in and offer to build a library. These were often local donors with a more direct connection to the city than Carnegie. While Carnegie's libraries are arguably a distinct treatment, in this subsection we report results that exclude a small number of cities from our control group that Bobinksi (1969) identifies as rejecting Carnegie libraries because a local philanthropist offered a library. The results from this alternative control group are presented in Table 4. The point estimates are very similar to the main findings in Table 1.

# 5 Potential mechanisms

We consider three possible mechanisms behind our estimated patenting increases: (1) Libraries may serve as “anchor” institutions, drawing additional population—and therefore patenting—to cities; (2) Libraries may have increased the stock of available knowledge in cities; and (3) Libraries might have increased collaborative opportunities between inventors.

## 5.1 Population growth

First, we consider whether population growth alone can explain our results. Andrews (2018) finds that population changes explain most of the increase in patenting observed after colleges are built. Libraries are less likely than colleges or other large institutions like prisons or factories to mechanically attract large numbers of people because of new employment opportunities. To formally test for this possibility, we estimate the following model, which augments equation (1) with time-varying controls for city population:[21]

$$PatentCount_{ist} = \beta_1 Library_i + \beta_2 Post_{t,i} + \beta_3 Library_i * Post_{t,i} + X_i + \beta_4 LnPop_{it} + \delta_s + \gamma_t + \epsilon_{ist} \quad (3)$$

The results, reported in Table 5, indicate that population growth alone is unlikely to explain all the growth in patenting behavior that we observe.

## 5.2 Information

Next, we test whether patent citations to books increased after Carnegie library construction. This is a proxy for exposure to new ideas following library entry. We construct a measure of patents that cite books by hand-collecting a set of words that are associated with book citing in a training data set and searching the corpus of remaining patents for similar phrases. We focus on short-run book-citing effects before 1930. As illustrated in Figure 1, this is the period where we observe the largest difference

[21]We exclude any city from the analysis sample that is missing population records. Because our population data is bottom-coded, we exclude very small cities that have more than two decennial census years worth of bottom-coded data.

in patenting behavior between Carnegie and control group cites. Moreover, at this point in time, citations to prior patents were not required, and it was more common to cite non-patent materials (e.g., books, scientific magazines). Most importantly, to the extent that libraries had immediate and direct effects on the information available to inventors, it likely occurred early in our sample period when there were fewer outside information options.

To test this mechanism, we estimate analogues of our baseline difference-in-difference regression from equation (1). The outcome variable is the number of patents that we observe citing books in a particular city-year. We estimate the raw count of patents that cite books unconditionally and also conditional on a flexible function of total patent counts. Our results, in Table 6, indicate the patents that cite books increased more in cities that built rather than rejected Carnegie libraries. In particular, unconditional on patent volume, building a library is associated with approximately 0.05 patents that cite books per city-year before 1930. This is small in absolute terms but large relative to the sample mean of 0.07 book-citing patents per city-year. Conditioning on patent counts reduces the magnitude of the effect—as expected—but cited books still increase. This suggests that access to information through library books is a plausible mechanism generating increased patenting.

## 5.3 Collaboration

Finally, we test whether libraries affect innovative collaboration. We calculate the total number of inventors in each city-year by summing over all patents filed. If collaboration has become more common after library entry, we would expect the number of inventors to increase more in cities that received a Carnegie library. We estimate analogues of our baseline difference-in-difference regression with $ln(inventorcount + 1)$ as the dependent variable. As with the book citation mechanism results discussed above, we focus on short-run library effects before 1930 for similar reasons.

Table 7 shows preliminary results from this analysis. These results indicate that the number of inventors increased after library entry. The large unconditional increase in inventors (approximately 15 percent) is unsurprising given that we find that libraries increased patenting behavior, particularly before 1930 (See Figure 1). More importantly, the second row of estimates shows that *even conditional* *on the number of patents* the number of inventors increased. This implies that patents became more collaborative after the introduction of libraries. Given the importance of collaboration in the innovative process, these results suggest that increased collaboration in cities with libraries generated additional patents.

# 6 Conclusion

In this paper, we study the rollout of one of the most common public institutions in local communities: the public library. Leveraging the expansion in library services generated by Andrew Carnegie's grants in the late 19th and early 20th centuries, we test whether cities that accepted libraries increased innovative activity, proxied by patenting. We find that patenting behavior increased in cities that accepted libraries relative to a novel control group of cities that did not build a library despite being deemed eligible to receive a grant. We show that increased access to information and collaboration are potential mechanisms driving this trend. Future work will investigate the long-run innovative output of individuals exposed to libraries during their youth.

Our results motivate a need for more research on the historical and contemporary effects of libraries. In particular, more work on the impact of library exposure on literacy, human capital accumulation, and other innovative activities is warranted. Our findings also suggest that access to information is a key input to the innovation process.

# References

[1] Aghion, Philippe, and Peter Howitt. "A Model of Growth Through Creative Destruction." *Econometrica* 60, no. 2 (1992): 323-51.

[2] Aghion, Philippe, Leah Boustan, Caroline Hoxby, and Jerome Vandenbussche. "The causal impact of education on economic growth: evidence from US." *Brookings Papers on Economic Activity* 1 (2009): 1-73.

[3] Allen, Robert C. "Collective invention." *Journal of Economic Behavior & Organization* 4, no. 1 (1983): 1-24.

[4] Andrews, Michael. "Bar Talk: Informal Social Interactions, Alcohol Prohibition, and Invention" (2017)

[5] Andrews, Michael. "How Do Institutions of Higher Education Affect Local Invention? Evidence from the Establishment of U.S. Colleges" (2018)

[6] Andrews, Michael. "Comparing Historical Patent Datasets" (2018)

[7] American Library Association. "About us." http://www.ala.org/aboutala/ (accessed 2/15/2019)

[8] Biasi, Barbara, and Petra Moser. "Effects of Copyrights on Science-Evidence from the US Book Republication Program." No. w24255. National Bureau of Economic Research, 2018.

[9] Berkes, Enrico. "Comprehensive Universe of U.S Patents (CUSP): Data and Facts." (2018).

[10] Bobinski, George S. "Carnegie ibraries: Their history and impact on American public library development." ALA Bulletin 62, no. 11 (1968): 1361-1367.

[11] Bobinski, George S. *Carnegie Libraries: Their History and Impact on American Public Library Development*. American Library Association, 1969.

[12] Campos, Raquel, Fernanda Leon, and Ben McQuillin. "Lost in the Storm: The Academic Collaborations That Went Missing in Hurricane Issac." *The Economic Journal* 128, no. 610 (2018): 995-1018.

[13] Chetty, Raj, and Nathaniel Hendren. "The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects." *The Quarterly Journal of Economics* (2018) 133: 1107-1162.

[14] Chyn, Eric. "Moved to opportunity: The long-run effects of public housing demolition on children." *American Economic Review* 108, no. 10 (2018): 3028-56.

[15] Currie, Janet, and Enrico Moretti. "Mother's education and the intergenerational transmission of human capital: Evidence from college openings." *The Quarterly Journal of Economics* 118, no. 4 (2003): 1495-1532.

[16] Furman, Jeffrey L., and Megan J. MacGarvie. "Academic science and the birth of industrial research laboratories in the US pharmaceutical industry." *Journal of Economic Behavior & Organization* 63, no. 4 (2007): 756-776.

[17] Furman, Jeffrey L., Markus Nagler, and Martin Watzinger. "Disclosure and Subsequent Innovation: Evidence from the Patent Depository Library Program." No. w24660. National Bureau of Economic Research, 2018.

[18] Graham, Stuart, and Deepak Hegde. "Disclosing patents' secrets." *Science* 347, no. 6219 (2015): 236-237.

[19] Gross, Daniel P. "The Consequences of Invention Secrecy: Evidence from the USPTO Patent Secrecy Program in World War II." Harvard Business School Strategy Unit Working Paper 19-090 (2019).

[20] Hegde, Deepak, and Hong Luo. "Patent publication and the market for ideas." *Management Science* 64, no. 2 (2017): 652-672.

[21] Johnson, Alvin Saunders. *A report to Carnegie Corporation of New York on the policy of donations to free public libraries*. 1915.

[22] Jones, Theodore. *Carnegie Libraries Across America: A Public Legacy* . John Wiley Sons Inc, 1997.

[23] Karger, Ezra. "The Long-Run Effect of Public Libraries on Children: Evidence from the Early 1900s" (2019)

[24] Kevane, Michael, and William A. Sundstrom. "Public libraries and political participation, 1870-1940." (2016b).

[25] Kevane, Michael, and William A. Sundstrom. "State promotion of local public goods: The case of public libraries." (2016a).

[26] Kevane, Michael, and William A. Sundstrom. "The development of public libraries in the United States, 1870 to 1930: A quantitative assessment." *Information & Culture* 49, no. 2 (2014): 117-144.

[27] Krass, P. *Carnegie*. John Wiley Sons. 2011.

[28] Kuhn, Jeffery; Younge, Kenneth; and Alan Macro "Patent Citations and Empirical Analysis." (2017)

[29] MacLeod, Roy, ed. *The Library of Alexandria: Centre of Learning in the Ancient World*. IB Tauris, 2004.

[30] Murray, A. P. *The library: An Illustrated History.* Skyhorse, 2009.

[31] Romer, Paul M. "Endogenous technological change." *Journal of Political Economy* 98, no. 5, Part 2 (1990): S71-S102.

[32] Romer, Paul M. "Why, Indeed, in America? Theory, History, and the Origins of Modern Economic Growth." *American Economic Review* 86, no. 2 (1996): 202-206.

[33] Saavedra, Martin Hugo, and Tate Twinam. "A Machine Learning Approach to Improving Occupational Income Scores." (2018).

[34] Sharp, Katharine L. "The ALA Library Exhibit at the WorldâĂŹs Fair." *Library Journal* 18, no. 8 (1893): 280-84.

[35] Shera, Jesse Hauk. *Foundations of the Public Library*. University of Chicago Press, 1949.

[36] Wiegand, W.A. *Part of Our Lives: A People's History of the American Public Library*. Oxford University Press. 2015

Figure 1: Patents per city-year in cities that accepted and rejected libraries

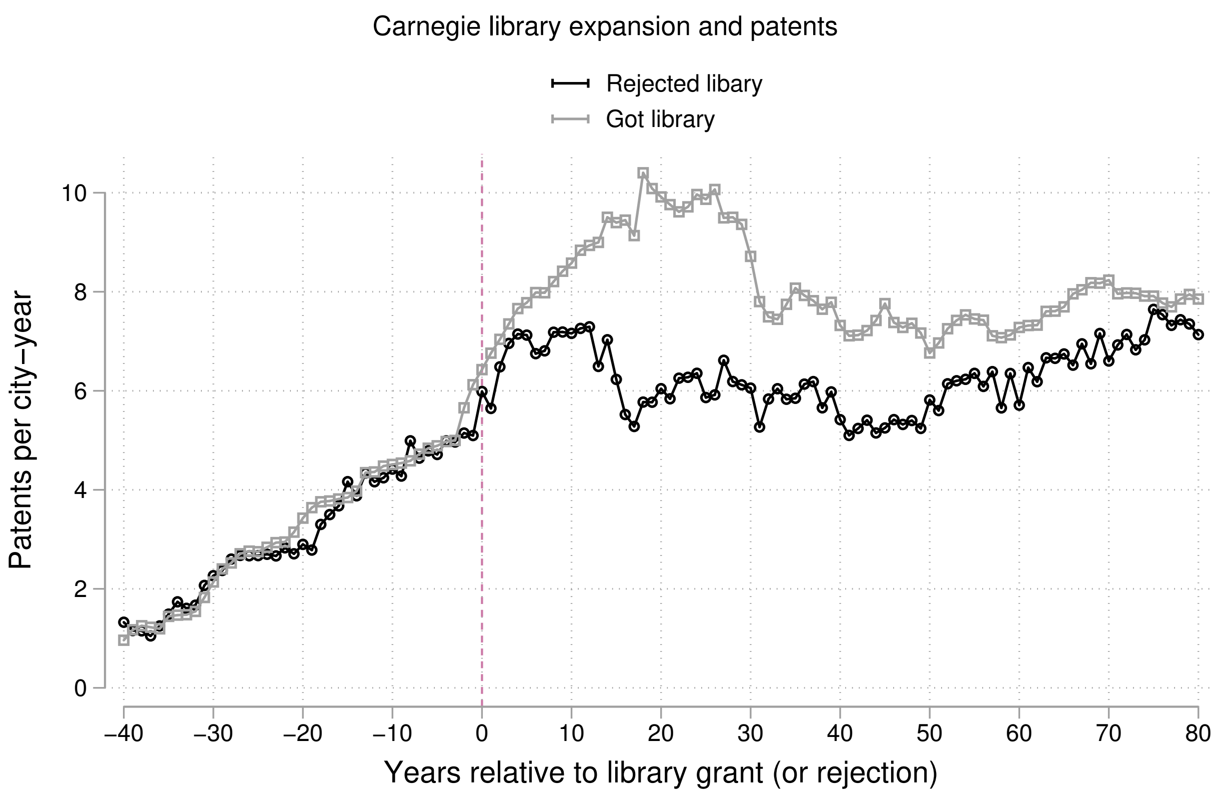

Figure 2: Carnegie's mail (from Jones, 1997)


**SPEAKING OF LIBRARIES, HOW WOULD YOU LIKE TO BE THE MAIL MAN?**

Figure 3: Front page headline of the July 7th, 1892 New York Times

# MOB LAW AT HOMESTEAD

*PROVOKED BY AN ATTACK OF PINKERTON DETECTIVES.*

TEN MEN KILLED AND AT LEAST FIFTY WOUNDED.

**FIERCE BATTLES FOUGHT AT THE STEEL WORKS—THE DETECTIVES ATTEMPT TO LAND FROM BOATS AND ARE DRIVEN BACK AND HELD UNTIL THEY SURRENDER.**

PITTSBURG, July 6.—Mob law has prevailed at Homestead to-day. The Carnegie Steel Works

Figure 4: Cumulative Carnegie library grants by year

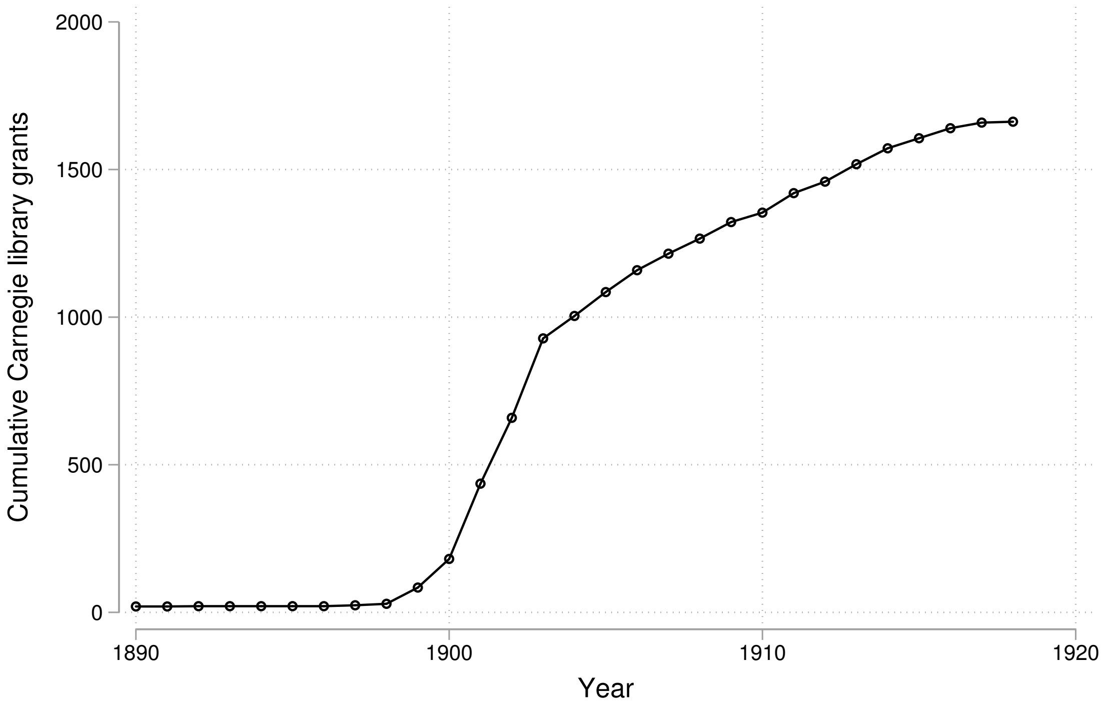


Figure 5: Granted Carnegie libraries per state

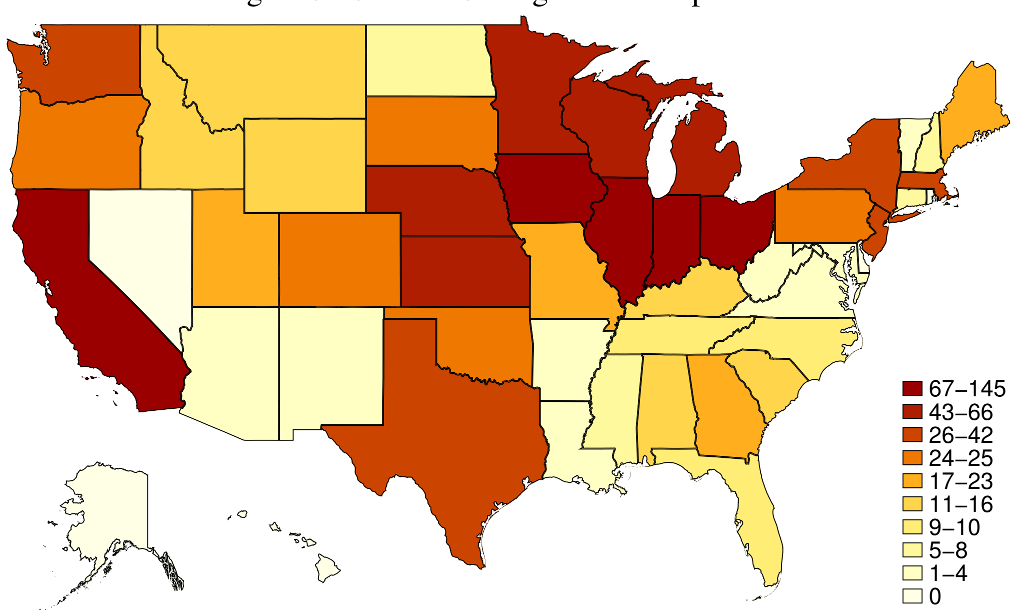

Figure 6: Rejected Carnegie libraries per state

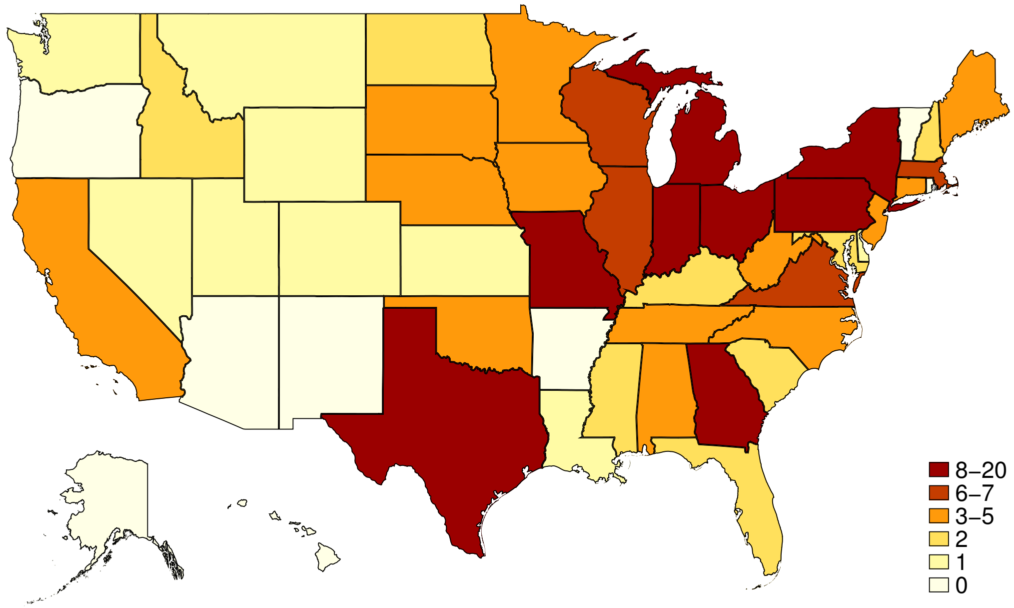


Figure 7: Covariate balance across counties that rejected and accepted libraries

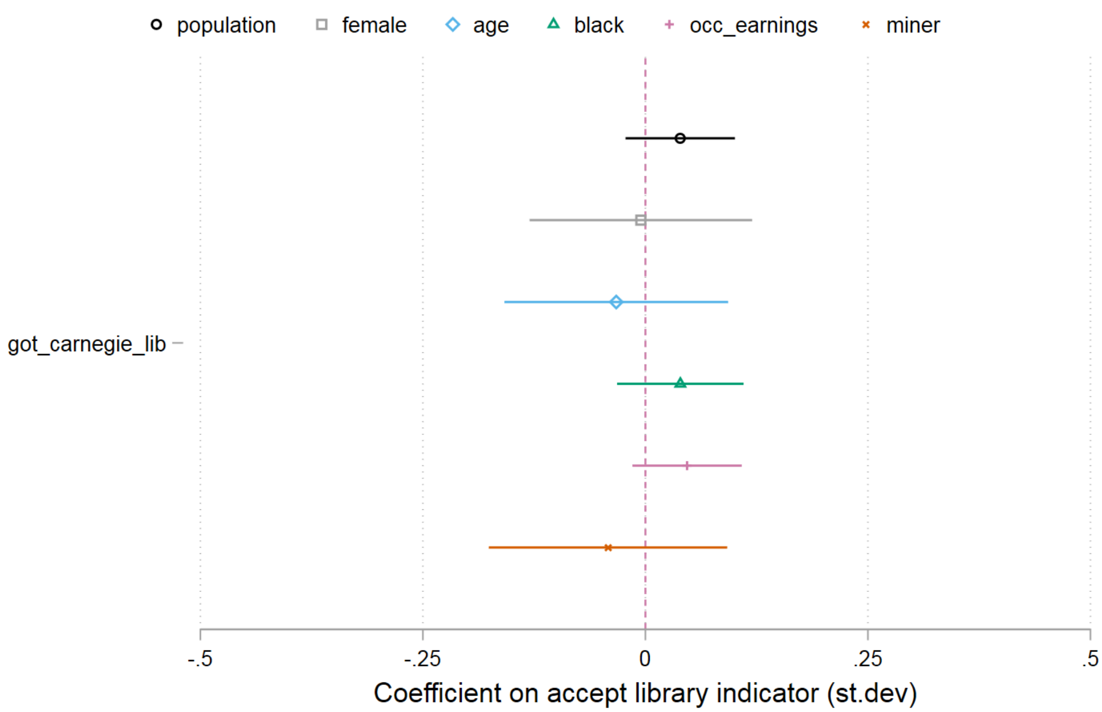

Table 1: Effect of library formation on patenting behavior

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Post | -0.01 | 2.70 | 3.91 | 2.34 |
| | (1.38) | (1.31) | (1.35) | (0.57) |
| Got library x post | 1.69 | 1.72 | 1.73 | 1.76 |
| | (0.80) | (0.80) | (0.80) | (0.81) |
| 1900 covariates | Y | N | N | N |
| Year FE | Y | Y | N | N |
| State FE | Y | Y | Y | N |
| Obs (city-years) | 165,001 | | | |
| Cities | 1,514 | | | |
| Mean patents per city-year | 7.83 | | | |

Standard errors in parentheses, clustered by city
Estimates from model (1)
The dependent variable is the count of patents
Post indicates years after getting or rejecting a library grant
Got library indicates cities that built a library

Table 2: Heterogeneity in library difference-in-difference effects across International Patent Classifications

| IPC patent group | Library x Post coefficient | Mean patents per city |
|---|---|---|
| Human necessities | 0.31*** | 1.20 |
| Performing ops/transport | 0.48* | 2.12 |
| Chemistry | 0.07 | 0.73 |
| Textiles | -0.03 | 0.21 |
| Constructions | 0.19** | 0.52 |
| Mech. engineering | 0.30* | 1.06 |
| Physics | 0.13 | 0.69 |
| Electricity | 0.09 | 0.62 |

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$
Estimates are from separate versions of model (1) across classifications
Includes state and year fixed effects and 1900 covariates
Post indicates years after getting or rejecting a library grant
Got library indicates cities that built a library

Table 3: Effect of library formation on average city-level patent first claim word count

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Post | 3.56 | 5.22 | 5.73 | 253.62 |
| | (3.33) | (3.27) | (3.27) | (9.88) |
| Got library x post | -6.04 | -5.69 | -5.67 | -28.45 |
| | (3.40) | (3.38) | (3.40) | (10.45) |
| 1900 covariates | Y | N | N | N |
| Year FE | Y | Y | N | N |
| State FE | Y | Y | Y | N |
| Obs (city-years) | 89,016 | | | |
| Cities | 1,514 | | | |
| Mean first claim word count | 285.38 | | | |

Standard errors in parentheses, clustered by city
Estimates from model (2)
Sample limited to cities-years with observed patents
The dependent variable is the mean word count in patent first claims
Post indicates years after getting or rejecting a library grant
Got library indicates cities that built a library

Table 4: Effect of library formation on patents, excluding non-Carnegie philanthropist cities

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Got library x post | 1.73 | 1.76 | 1.77 | 1.82 |
| | (0.83) | (0.83) | (0.83) | (0.84) |
| 1900 covariates | Y | N | N | N |
| Year FE | Y | Y | N | N |
| State FE | Y | Y | Y | N |
| Obs (city-years) | 162,821 | | | |
| Cities | 1,494 | | | |
| Mean patents per city-year | 7.83 | | | |

Standard errors in parentheses, clustered by city
Estimates from model (1)
Control group excludes cities that rejected Carnegie because of local donors
The dependent variable is the count of patents
Post indicates years after getting or rejecting a library grant
Got library indicates cities that built a library

Table 5: Effect of library formation on patents conditional on city population

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Got library x post | 3.40 | 3.42 | 3.59 | 3.33 |
| | (1.46) | (1.84) | (1.87) | (1.73) |
| 1900 covariates | Y | N | N | N |
| Year FE | Y | Y | N | N |
| State FE | Y | Y | Y | N |
| Obs (city-years) | 9,714 | | | |
| Cities | 891 | | | |
| Mean patents per city-year | 13.38 | | | |

Standard errors in parentheses, clustered by city
Estimates from model (3)
Sample limited to cities with consistent population data
The dependent variable is the count of patents
Post indicates years after getting or rejecting a library grant
Got library indicates cities that built a library

Table 6: Effect of library formation on patents that cite books

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| **Unconditional on patent counts** | | | | |
| Got library x post | 0.04 | 0.05 | 0.05 | 0.05 |
| | (0.01) | (0.01) | (0.01) | (0.01) |
| **Conditional on patent counts** | | | | |
| Got library x post | 0.02 | 0.02 | 0.02 | 0.02 |
| | (0.01) | (0.01) | (0.01) | (0.01) |
| 1900 covariates | Y | N | N | N |
| Year FE | Y | Y | N | N |
| State FE | Y | Y | Y | N |
| Obs (city-years) | 76,590 | | | |
| Cities | 1,514 | | | |
| Mean book-citing patents | 0.07 | | | |

Standard errors in parentheses, clustered by city
Sample limited to 1930 and before
The dependent variable is the count of patents that cite books
Post indicates years after getting or rejecting a library grant
Got library indicates cities that built a library

Table 7: Effect of library formation on total inventors

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| **Unconditional on patent counts** | | | | |
| Got library x post | 0.15 | 0.17 | 0.15 | 0.15 |
| | (0.04) | (0.04) | (0.05) | (0.05) |
| **Conditional on patent counts** | | | | |
| Got library x post | 0.08 | 0.07 | 0.06 | 0.06 |
| | (0.03) | (0.03) | (0.03) | (0.03) |
| 1900 covariates | Y | N | N | N |
| Year FE | Y | Y | N | N |
| State FE | Y | Y | Y | N |
| Obs (city-years) | 76,590 | | | |
| Cities | 1,514 | | | |
| Mean log inventors | 1.21 | | | |

Standard errors in parentheses, clustered by city

Sample limited to 1930 and before

The dependent variable is $ln(inventors + 1)$

Post indicates years after getting or rejecting a library grant

Got library indicates cities that built a library