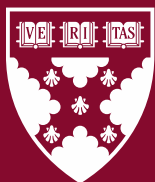# You've Got Mail! The Late 19th Century US Postal Service Expansion, Entrepreneurship, and Firm Performance

Astrid Marinoni
Maria P. Roche

Harvard Business School

You've Got Mail! The Late 19th Century US Postal Service Expansion, Entrepreneurship, and Firm Performance

Astrid Marinoni
Georgia Institute of Technology

Maria P. Roche
Harvard Business School

Working Paper 23-022

You've Got Mail! The Late 19th Century US Postal Service Expansion, Entrepreneurship, and Firm Performance[*]

Astrid Marinoni[†]   Maria P. Roche[‡]

October 11, 2022

preliminary results - please do not cite or circulate

Abstract

We analyze the impact of the US Postal Service's western expansion in the late 19th century, a major institutional innovation, on entrepreneurship and its heterogeneous influence on firm performance. Exploiting a novel dataset constructed from digitized archival information on historic business establishments, post office locations, and road network characteristics in the state of California, we find a positive relationship between the expansion of the postal service and entrepreneurship. To address omitted-variable concerns we use information on early settlement pathways and weather variation and further examine feasible channels through which the US Mail may have promoted firm entry: by providing access to monetary distribution centers, government services, communication channels, and specialized knowledge. Our results support the latter. The implications for firm performance are nuanced. While increasing competition given new entry exerts downward pressure on many incumbents, actors relying on specialized knowledge inputs are able to benefit from access to the US Mail.

Keywords: *Entrepreneurship, Institutional Innovation, Knowledge Exchange, US Postal Service, Firm Performance*

1 Introduction

Scholars of innovation have long stressed the critical role of institutions and suggest that the institutional environment fundamentally influences technological progress through a variety of factors (Sokoloff, 1988; Almeida & Kogut, 1999; Khan, 2005). These include, for example, the security of property rights, patent laws, the educational system, and competition policy. Although some scholars have documented the fundamental role of institutional structures in the economic development of the late 19th century United States (Acemoglu et al., 2016), we still lack an understanding of how these important institutional foundations are connected to the pace and nature of entrepreneurship.

One such structure is the US Postal Service, which was the dominant means of communication during the late 19th century. To establish a national system of communications, the US government wove together a so called "gossamer network" across the West – a new and unfamiliar model for a government institution (Blevins, 2021). The federal government did not own or operate most of the postal infrastructure, but crafted temporary and part-time arrangements to transport and distribute the mail. Several thousand of the offices formed in the second half of the nineteenth century only operated for a few years or even a few months before closing. Openings and closings often occurred contemporaneously, creating much churn in the stock of regional post offices over this period. Sprawling and fast-moving, this web of post offices connected the region's settlements by operating in far more places than any other government institution. Between 1880 and 1900 alone, 82,500 new post offices were established across the country.

The question we strive to address in this paper is, empirically, what the impact of the US Postal Service was on entrepreneurship and to unveil potential implications for firm performance. We expect that being connected to the rest of the country and world via the US Post may have been one of the most pressing necessary conditions for firms to set up shop, for example through access to money, the regulatory system, communication or knowledge. But, it is not clear if such connection –on net– would lead to better performance outcomes provided potentially higher competitive pressures (Alcácer & Chung, 2007). To shine light on this line of inquiry, we make use of new data capturing the precise location and time of operation of post offices in the state of California between 1880 and 1890 (Blevins, 2021): a young state that was rapidly developing at the time. We complement these data with historic R.G. Dun & Co. credit reports that we retrieve from their original fragile paper format in library archives and transcribe. These records, unique and not accessible to the public, provide a roster of businesses and fine-grained information on the financial performance of firms as well as rich detail on the nature of the local business and financing environment.

To assess entry using these unique data, we estimate a first-differenced model with the settlement as our unit of analysis. Our main dependent variable is the change in stock of Post Offices within 20km from a settlement, measured using a novel digitized map encompassing the entire California street network of 1880. Our model includes controls for population, town-level economic activity, and county and industry fixed effects. To control for potential omitted-variable bias, we instrument post offices using information on historic settlement routes used by California's early pioneers. We also leverage an alternative instrument based on extreme weather conditions. Our findings suggest that an increase in the number of post offices within 20 kilometers of the settlement center is associated with higher firm entry. In particular, instrumented specifications suggest that the presence of a post office increased entry growth rates by 66-percent on average across industries over the course of a decade. Moreover, we find that those settlements in which post offices were less spatially concentrated experienced the largest boost from an increase in the stock of post offices.

Additional analyses provide further insight into potential mechanisms driving these results. Ex ante, there are several candidate channels through which the post office may have operated. We concentrate our efforts on the following four: 1) as a financial service, 2) as an arm connecting to other government services, 3) as mass communication infrastructure, or 4) as carrier of specialized

knowledge.[1] To empirically investigate this, we complement our data with historic post office reports and directories to retrieve information on the extent to which a post office also served as a money order touch point, where individuals could send and receive money. For the role as an arm of the legal system, we gather information on the presence of courts and assess whether post offices serve as substitutes or not. To explore the role of communication, we exploit the fact that operations in some industries may be more affected than operations in others by quicker and more reliable information flow. To understand the role of knowledge in this framework, we retrieve detailed summaries of specialized trade journals by industry and check whether industries in which trade journals were circulating abundantly out-of-state[2] experience higher growth in entry. As a result of this exercise, we provide support for the explanation that access to specialized knowledge in the form of periodicals and magazines may have been a critical mechanism through which post offices contributed to increasing entrepreneurship during our study period (Haveman, 2015). Our results therefore suggest that the post office expansion in the late nineteenth century may have very likely contributed substantially to entrepreneurship via the offices' role as "Carrier of (...) Knowledge." Extrapolating our results, the enhanced flow of knowledge between distant individuals and communities that the rapid development of the post office network enabled may well have contributed to making America a developed economy (Rogowski et al., 2021; Aneja & Xu, 2022).

Regarding firm performance, our results are more nuanced. Estimating a similar model, but on the firm-level and controlling for firm-specific characteristics such as if the founder had established multiple establishments, if the company was family-based, and the nationality of the founder, we find that increases in competition given new entry exert downward pressure on many incumbents. As in the case for entry, we detect that firms in industries with out-of-state trade journals experience better performance with increases in post offices and further find that actors in those industries relying on a specific type of government distributed knowledge and technology (Cohen et al., 2002; Nagaraj, 2022) –seeds and food safety– experienced large boosts in performance. Taken together, this suggests that the presence of a post office may have been particularly crucial for those firms that could leverage specialized knowledge and technology inputs most in their day-to-day operations.

---

[1]In the period we consider, the postal system carried only letters and envelopes, so we need not take into consideration its role as a carrier of larger goods.

[2]Since the number of local trade journals might be endogenous to performance and, in general, to the industry composition of a given location (i.e., we might expect to find more journals dedicated to a specific industry published in locations where this industry is developing and growing), we consider only trade journals published *outside* California, while controlling for all the journals published in California.

Many Americans at the time of our study "considered the role of the postal system in facilitating the transmission of information [. . .] to be its most important 'national service' " (John, 1995: 13). Building on this notion, we provide empirical evidence on the Postal Services' influence on entrepreneurship and firms via its impact on knowledge diffusion. Taken together, our findings speak to the literature on the importance of a region's capacity to connect people and ideas for firm growth and innovation (Jacobs, 1969; Gaspar & Glaeser, 1998; Feldman & Link, 2001; Rosenthal & Strange, 2001; Storper & Venables, 2004). Furthermore, we provide insights on the diffusion of knowledge and technologies (Jaffe et al., 1993; Singh, 2005; Thompson & Fox-Kean, 2005; Agrawal et al., 2006; Singh & Marx, 2013; Agrawal et al., 2017), the gains from inward knowledge spillovers but also the possible cost of local competitive pressures (Alcácer & Chung, 2007), and highlight the role of an institutional innovation[3] in infrastructure expansion in promoting firm entry and boosting regional performance. Our results stress the importance of the Postal Service, and such institutional innovations more broadly, in providing an environment that fosters knowledge exchange. In light of recent discussion on infrastructure improvement in the US (and across the globe), our findings may provide a useful bases for debate. Moreover, our results on the importance of wide access to specialized knowledge from other geographies for entrepreneurship and the role of broad distribution of public technology, provide critical managerial understanding with regards to the relationship between knowledge sourcing and firm outcomes.

This paper proceeds as follows. Section 2 briefly describes the US Postal Service expansion and outlines some potential mechanisms by which it might have affected businesses. Sections 3 describes our data in detail. Section 4 uses city-level analyses to estimate the impact of post offices on firm entry and delves into possible mechanisms that can explain our results. Section 5 investigates the impact of post offices on firm performance highlighting important competition-, quality, and knowledge-based nuances. Section 6 concludes.

2   Literal Paper Trails: The US Postal Service Expansion and Theoretical Considerations

Between the 1870s and the early 1900s, the United States – especially the Western part – underwent a dramatic reorganization of people, land, capital, and resources. Major drivers suggested

---

[3]Following (Li et al., 2020), we think of institutional innovation as "...the creation of a new and more effective system to encourage people's behaviour, and the realisation of social sustainable development and innovation under the existing production and living environment" (p.115801).

to have supported this migration and catalyzed economic growth are rooted in infrastructure (Démurger, 2001). For instance, empirical evidence directly links growth in manufacturing productivity throughout the nineteenth century to the expansion of the railroad (Redding & Venables, 2004; Hanson, 2005; Redding et al., 2007; Head & Mayer, 2011; Donaldson & Hornbeck, 2016; Balboni, 2019; Hornbeck & Rotemberg, 2019; Jaworski & Kitchens, 2019) and of canals and portages (Bleakley & Lin, 2012). Creating necessary infrastructure substantially improved the ability to move people and materials.

One particular type of infrastructure, which expanded in a very novel way for its time was the US Postal Service. This institutional innovation has been far less examined in the entrepreneurship and strategy literatures but may have had far-reaching consequences for economic growth (Rogowski et al., 2021), innovation (Aneja & Xu, 2022), and in laying the groundwork for present-day entrepreneurial conditions. As millions of settlers moved in the post–Civil War period, they relied heavily on letters, newspapers, magazines, and money orders to stay connected to the wider world. During this period, the US postal system was the nation's largest communications network, with no other public institution as omnipresent and as central to everyday life. Over the course of the second half of the 19$^{\text{th}}$ century, the US Post extended its reach into nearly every American city, town, and village following a novel approach to how it grew. To establish a national system of communications, the US government wove together a so called "gossamer network" across the West. This western postal network was very different from the typical civil service bureaucracies associated with government institutions. Early on, the US Mail contracted with stagecoach companies (e.g., the infamous Pony Express, which was only in service for a year and a half) to carry the mail and paid local merchants to distribute letters from their stores (Blevins, 2021). In this way, the US Post rapidly spun out a vast web of postal infrastructure to thousands of remote places, connecting widely dispersed settlements into a national system of communications. The rapid and branching expansion of the postal system was facilitated by the relative ease of requesting a post office, typically granted after a written request from a few citizens was filed and sent to the Postmaster General. The US Post operated in far more places than any other government institution, having established roughly 73,300 post offices between 1880 and 1890 alone. Though expanding rapidly, the postal infrastructure was highly unstable. Several thousand of the offices established in the second half of the nineteenth century only operated for a few years or even a few months before closing. Openings and closings often occurred contemporaneously, creating much churn in

the stock of regional post offices over this period.[4]

We propose that the flexible and transient manner of the expansion of the US Postal Service – above and beyond the traditional infrastructure factors suggested to have driven economic progress in this period such as railroads and canals – set up important necessary conditions enabling firm entry and potentially impacting firm performance. Being connected to others via the services the post office afforded – sending and receiving letters, newspapers, periodicals, and money orders – may have had a substantial influence on the decision to start a firm and on firm performance. Yet, while a positive relationship between the expansion of the postal system and business entry seems intuitive, it is unclear exactly how the postal service expansion might have contributed to entrepreneurship. Moreover, ex ante, it is not obvious what the relationship with firm performance may be.

Entrepreneurship: Regarding firm entry we explore four candidate channels through which the post office may have operated that we will detail in the following: 1) as a financial service, 2) as an arm of the legal system, 3) as mass communication infrastructure, or 4) as carrier of specialized knowledge.[5]

The first possibility is that the US Post served as a monetary distribution center, enabling actors to send money orders from one certified post office to another. Individuals could thus pay bills, receive payments, and help friends and family who needed money quickly. One could exchange cash for a replaceable piece of paper that could only be used by a specific person or company, very much like a check. Although money orders were introduced in 1864, not all post offices provided this kind of service.

It is also feasible that the US Post served as a branch of the legal system. Legal institutions have been found to serve an important role by reducing uncertainties and establishing a stable economy and social relations. Through formal rules, legal institutions can determine what is acceptable and unacceptable and thus determine the rules of the game (Cernea, 1987). This may be what enabled firms to be established in the first place.

---

[4]Note that at the same time, new laws promoting settlement and investment in the western United States were passed disowning many first natives. Much of this legislation centered on the distribution of plundered Native land: to individual settlers (the Homestead Act), state colleges (the Morrill Act), and railroad companies (the Pacific Railway Acts).

[5]One might argue that the post office served as carrier of goods, facilitating exchange of good between suppliers, businesses, and customers. Mailing packages bigger than an envelope via the postal service was only possible starting in 1913, and thus falls outside the time scope of this paper.

Another major contribution of the post office is related to communication. In particular, post offices enabled connection to a widespread communication grid and rapidly revealed itself as substantially faster, cheaper, and more reliable than the existing means of communication, such as private arrangements to deliver mail. The post office offered businesses a fast and cheap method to communicate with customers, suppliers, and other stakeholders. Note that in the period of our analysis, the telegraph was also a communication option, but message length was extremely limited and it required physical wires to connect telegram stations. The wires generally ran along railroad tracks; as a result, by the end of the nineteenth century, telegram stations were only present in major cities.

Related to this last mechanism, the post office might have had a direct impact on knowledge flows through the dissemination of knowledge. With the postal service expansion, trade journals and newsletters – which have been attributed a fundamental role as carriers of specialized knowledge in the late nineteenth century (Haveman, 2015) – began to circulate more widely via the Post Office granting businesses greater access to more knowledge. Two changes around mid-century contributed to the steady rise of newsletters and trade journals as conduits of specialized knowledge. The first was the availability of codified specialized knowledge; for example, in the form of patents and other schematics. Second, the postage price for magazines was substantially reduced, becoming similar to the prices for letters and newspapers. These factors, in conjunction with the expansion of the post office, were crucial in reducing spatial barriers to interactions between magazines and their audiences and contributors. A statement by a reader in a leading agriculture journal of the time highlights the critical role of the ability to disseminate scientific knowledge: "It is to our agricultural paper, most emphatically, that we owe the awakening which has taken place in this State, on the subject of agriculture" (Haveman, 2015, p.265). This awakening was not restricted to agriculture, but also took place in other industries and professions. Trade journals were not only disseminating specialized knowledge by providing information about best practices and new innovations, but were also providing domain-specific information about suppliers and customers.

Performance: The net relationship between the US postal service expansion and the performance of firms is not obvious, since there are possible countervailing forces at work. For one, via it's impact on firm entry, incumbents may have experienced higher competitive pressures leading to a reduction in financial outcomes for many. For another, certain firms, such as the more sophisticated ones (with higher quality and growth mentality), or those who can leverage access to the mail the

most, may have experienced a boost in performance (Alcácer & Chung, 2007). If this is the case, then we are likely to find either a null or even negative relationship on average, but interesting heterogeneity along dimensions such as quality, experience, and industry.

3 Data

For the purpose of this paper, we focus our attention on one of the fastest-growing states in the West during the post–Civil War era – California – which achieved statehood in 1850. Settlers were attracted by the abundant natural resources and California quickly became one of the most economically important states in the West. Besides its economic importance, we also focus on this state given the extensive digitization efforts that this study required.

3.1 Post Offices

Information about post offices was provided by (Blevins, 2021) and is based on historical documents of the Post Office Department originally collected by R. Helbock and now in the National Archives. Data on post office locations was carefully geocoded, using historical names and landmarks of the time, through the Geographic Names Information System. Moreover, the data include information about the establishment and closing year of each post office, so that the life of each post office can be precisely determined. We capture post offices by counting the change in the number of active post offices in the vicinity of a city center from 1880 to 1890.

To accurately measure distance between city centers and post offices, we leverage the existing network of trails and roads existing in California in 1882. Given the relative scarcity of infrastructure in 19th-century California and the geographical conformation of the land, rich in mountainous and impervious terrains, the use of Euclidean distance might produce incorrect measurements as relatively close post offices might have actually been very distant (Figure A1 in the Appendix shows an example). We manually georeferenced and digitized what is believed to be one of the most accurate maps of California (Bancroft, 1882) of the time, which includes the full network of trails and roads existent in 1882 (Figure A2 in the Appendix). Further details on the georeferencing and digitization process of this map can be found in Section A of the Appendix. We supplement the full network of roads and trails with information on active railways, navigable rivers, and water canals active in that period Atack (2015, 2016, 2017). We then calculate the distance from each city and post office by leveraging this detailed infrastructural network, counting the number of post offices

within 20km from each city center. This distance corresponds to approximately four hours walking and an hour and a half trotting on horseback.[6]

3.2 Entrepreneurship and firm performance

The primary source we use to measure firm entry and performance is the R.G. Dun & Co. credit report collection we were granted access to in library archives. This unique, historic collection consists of 2,522 volumes of handwritten credit reports on individuals and firms from the United States, its western territories, Canada, and several foreign countries starting in the 1840s. Entries include information on the business's worth, life span, industry, sources of financing, and the character and reputation of its owners, their partners and successors. The credit report volumes are arranged geographically by state (or territory). Each state (or territory) is then subdivided alphabetically by county and/or city.

3.2.1 Entrepreneurship

We measure entrepreneurship as the growth in the number of new firms between 1880 and 1890 at city level by sector.[7] We include all cities, towns, and villages mentioned at some point during our period in the R.G. Dun & Co. credit report collections. We geolocate the historical position of this cities using different sources, including the above-mentioned georeferenced 1882 map of California (Bancroft, 1882), the position of historical towns included in Blevins (2021) and by manually looking up the remaining cities on websites providing information on the locations of historical towns. Our final sample is composed of 614 historic cities in California.

3.3 Other controls

California city-level population data for 1880 come from the *Historical Census Populations of California, Counties, and Incorporated Cities, 1850–2010*. Moreover, we obtain data containing

---

[6]One has to keep in mind that in 1890 most cities in California were not as extensive as they are nowadays. From a rough calculation using maps of Los Angeles dating in the 1890s, the city could be encompassed by a 15km diameter circle.

[7]R.G. Dun & Co. credit report collection classifies businesses in 31 different sectors: barber shops, books & publishing, breweries & distilleries, butchers & fisheries, builders & contractors, wagon-makers & locksmiths, tailors & clothing, mining, commission businesses, crockery & pottery, physicians & doctors, dry goods & cannery, fancy goods & varieties, florists, flour mills & bakers, foundries & heavy machinery, furniture & decorations, general hardware & ironworks, general stores, groceries, jewelry & watches, lumber mills, musical instruments, paints & oils dealers, plumbing & wire works, stables & harnesses, hospitality, financial services, shoemakers & leather goods, cigars & tobacco, warehouses & storage.

population information from Gibson (2007), based on decennial census reports on townships and places, and from the R.G. Dun & Co. credit reports, which provide relatively comprehensive information on the population of the smallest cities.[8] We supplement this data with information about industry, which is specified for each firm in the R.G. Dun & Co. reports and employ the original classification used in the data.[9]

We also construct a set of controls that capture access to alternative communication and transportation systems, i.e., railways, rivers, and canals. We build the corresponding variables indicating whether a city was in proximity (i.e., within a 20km radius) to any railways, rivers, or canals in each decade using the data provided by Atack (2015, 2016, 2017). Finally we add county information by leveraging 1880 counties' boundaries (Logan et al., 2011).

Table 1 shows summary statistics for our main variables and controls. In general, entry by industry was very skewed in both 1880 and 1890, with on average one firm entering each industry in 1890. Some combinations of sector-city tend to experience high entry in both periods. For instance, San Francisco scores the highest in terms of entry of hospitality-related businesses, with 726 and 1491 companies established in 1880 and 1890, respectively. When looking at entrepreneurial growth in these two periods, entry appears to be more scattered and heterogeneous. Figure A3 in the Appendix shows the entry growth (calculated as the inverse hyperbolic sine) of the 20 most populous cities in California between 1880 and 1890. Some industries are growing in almost every major city: this is the case for barber shops, cigars and tobacco, jewelry and watches, and medical instruments. At the same time, some cities experience growth in almost every industry. One example is Alameda, Los Angeles, and San Jose. Other cities, such as Bodie, experience a strong drop when it comes to new entry in 1890 compared to the previous decade. When examining the top 10 growing and declining cities in California (see Figure A4 in the Appendix) it is striking to note that no major cities make the list. Instead, it is cities such as San Jacinto, Delano, and Greenwich that score the

---

[8]If the information for one city was missing in 1880 we imputed the value using the average growth of the population between the 1870-1880 for California. Despite the various sources employed to find information about population, we were still unable to find population data for some minor settlements. For these cities we imputed the population value corresponding to the 25th percentile as these were most likely less populated settlements.

[9]The categories are: barber shops, books and publishing, breweries and distilleries, butchers and fisheries, builders and contractors, wagon-makers and locksmiths, tailors and clothing, mining, commission businesses, crockery and pottery, physicians and doctors, dry goods and cannery, fancy goods and varieties, florists, flour mills and bakers, foundries and heavy machinery, furniture and decorations, general hardware, tinware, and iron-works, general stores, groceries, jewelry and watches, lumber mills, musical instruments, paints and oils dealers, plumbing and wire works, stables and harnesses, hospitality, financial services, shoemakers and leather goods, cigars and tobacco, warehouses and storage.

highest when it comes to aggregate businesses across all industries. Most of the firms that enter in 1880 and 1890 (27%) are active in the hospitality sector, which primarily consists of saloons and hotels. Wagon-makers & locksmiths represent the second category with 9% of new firm entry in 1880 and 1890. This sector is followed by groceries and general stores. Figure A5 in the Appendix shows a detailed breakdown by sector of the firms in our sample.

Figure 1 shows the expansion of the US Postal Service from 1880 to 1890, highlighting its growth throughout California. While there were 841 post offices in 1880, there were 1,254 by 1890. During this time, 565 new post offices were created, and 152 closed. The growth in the number of post offices can be explained by how easy it was to establish a new post office. All it took was a letter from a few citizens to the Postmaster General; requests were rarely rejected (Blevins, 2021). The ease of establishing post offices allowed for their capillary expansion across the US territory. For our empirical analysis, this provides some reassurance that the establishment of most post offices was not a consequence of long-lasting economic development in a specific area. When looking at cities, roughly 65% of them do not gain any new post offices between 1880 and 1890 (53% gain no post office, and 12% lost 1 or 2 post offices), while 35.3% get at least one. 20% of cities gained one post office, 8% gained two, 4% gained 3. The remaining gained between 4 to 12 post offices. Figure A6 in the Appendix shows the change in the number of post offices by city from 1880 to 1890. Figure A7 displays patterns across the entire country.

4 Post Office Location and Entrepreneurship

4.1 Empirical strategy

To estimate the relationship between post offices and entrepreneurship, we use a first-differences model (1880-1890) at the city (c) and sector (s) level which abstract from yearly variation, as follows:

$$\Delta^{80-90} Entrepreneurship_{i,s} = \alpha + \beta \ \Delta^{80-90} PostOffices_i + X_i\Gamma + \lambda_c + \theta_s + e_{i,s} \tag{1}$$

where i, s, c indexes cities, sectors, and counties, respectively. $\Delta^{80-90} Entrepreneurship_{i,s}$ is measured as the change in the inverse hyperbolic sine transformation ($arcsin$) of the new number

of businesses in sector s between 1880 and 1890 in city i.[10] $\Delta^{80-90}PostOffices_i$ is a dummy indicating whether the city has gained at least one new post office between 1880-1890. X_i is a set of socio-economic and transportation-related controls at the city-level. The first set of controls include population and the initial stock of businesses. In general, these controls help us account for a possible alternative explanation based simply on the agglomeration of people regardless of communication infrastructure (Carlino et al., 2007). To these controls we add variables accounting for alternative transportation methods, i.e., a variable indicating if a city had access to a railway line, a waterway, a navigable river within a radius of 20km. λ_c and θ_s represent county- and sector-fixed effects, respectively. $e_{i,s}$ is the error term. We cluster our standard errors at the city level (Abadie et al., 2017).

4.1.1 Addressing identification

In an ideal experiment, post offices would be randomly assigned to locations. Post offices would thus not be located where there is population and economic growth, nor would firms choose to be close to a post office based on their characteristics, nor would simultaneously occurring events influence location decisions, allowing the researcher to cleanly estimate the effect of post offices on establishment growth and firm performance. Although such an experiment is not within our reach, the thought experiment highlights two major threats to identification that we must address as best we can: omitted variable bias and reverse causality.

With regard to omitted variable bias, urban growth, economic activity, and communication infrastructure may be simultaneously co-determined and regions that were developed earlier may have attracted more people and created more employment than younger regions. These locations may then also have continuously attracted firm founders, who then need employees (and customers). Some areas in a region will have been more suitable for development than others due to, for example, access to water (Duranton & Turner, 2012) or other natural resources. To partially address the issues related to omitted variable bias, we include county fixed-effects, which keep unobservable features of a county, such as its natural resources, constant; we also include industry fixed effects which help us control for industry specific trends and features.

---

[10]The inverse hyperbolic sine transformation has recently become very popular given it exhibits the usual properties of the natural logarithm. At the same time, it allows for zero- and negative-valued observations, which tend to convey important information that a simple log- or the log+1-transformations ignore (Bellemare & Wichman, 2020).

Another threat to our identification strategy is that firm entry may have attracted the postal services, rather than the other way around. If that were the case, we would be measuring the pull relationship of economic activity on firm creation. This type of reverse causality is unlikely to be the driver of our results, given how easy it was to request a new post office. As a consequence, counties with low population density were often the areas that received new post offices, especially as post offices expanded into the West.[11] We also empirically assess whether reverse causality plays a major role in our analyses in Section 4.2.2. The results suggest it is unlikely.

In addition to the standard first-difference model described in equation 1, we apply an instrumental variable (IV) estimation approach to address endogeneity concerns. In this case, an appropriate instrument to detect the causal relationship between the stock of post offices and establishment growth must be strongly correlated to the number of post offices, but have little influence on the number of new firms other than through its effect on post offices.

For our city-level analysis, we use an instrumental variable to address omitted variable bias that arguably meets the necessary conditions. One advantage of focusing our analysis on California is the presence of the historic trails used in the early 1800s in the region. These trails were the first exploratory trails that individuals from the East embarked on to get to the West and were discontinued as main passage routes by 1869.[12] Many of them were not very suitable for passage in the first place and brought great perils and danger (e.g., such as the South Pass crossing established by General William H. Ashley's and Major Andrew Henry's "Enterprising Young Men"(Morgan, 1953)). Part of the trails were initially used for the infamous Pony Express, which lasted for less than two years (1860-1861), and as early gold discovery routes. As displayed in Figure 2, the trails were often created following the path of least resistance, heavily dictated by the natural environment, such as by mountain ranges and passes. Our rationale for using this instrument is that cities close to these historic trails might have experienced more growth in the stock of post offices because they could already rely on these trails as a well-developed communication and transportation system. The exclusion restriction of this instrument relies on the fact that since they were established over 50 years before the period of our study, and were dictated by physical constraints rather than suitability, their presence should not be correlated with changes in firm entry between 1880 and 1890. We gather data about the exact location of the historic trails from

---

[11]For a graphic display of the relationship between county population density and new post offices, please refer to Figure A8 in the Appendix.

[12]For more information, see https://www.nps.gov/cali/learn/historyculture/index.htm

the California Natural Resources Agency and build our instrument calculating the distance (in km) between each city in our sample and the closest point on the trails.

Using this instrument, we estimate the following first-stage and instrumented first-differenced models:

$$
\begin{cases}
\Delta^{80-90} PostOffices_i = \gamma + \theta\ Z_i + X_i\Lambda + \lambda_c + \delta_s + \epsilon_{i,s} \\
\Delta^{80-90} Entrepreneurship_{i,s} = \alpha + \beta\ \Delta^{80-90} \widehat{PostOffices}_i + X_i\Gamma + \lambda_c + \theta_s + e_{i,s}
\end{cases}
\tag{2}
$$

where the first equation is our first stage, and where Z_i is the instrument for the growth of post offices in the 1880-1890 period.

In order to boost the credibility of our causal estimates we also consider another instrument, which exploits high temperature and precipitation variability (i.e., extreme weather) experienced by cities from 1880 to 1890. In general, temperature and aridity had a substantial impact on the stock of post offices through their impact on farming and ranching productivity (Blevins, 2021). Given the high population mobility in that period, extreme weather might have pushed farmers and ranchers to move to more favorable environments. We capture extreme weather in the 1880s using the Palmer Drought Severity Index (PDSI) (Palmer, 1965; Cook et al., 2010), widely used in the meteorological literature to gauge soil moisture and aridity. It is available in a spatial resolution of 0.5 x 0.5 degrees. We define extreme weather as the difference between the minimum and the maximum PDSI value for the 1880 decade. We argue that extreme variations in precipitation and aridity might have forced individuals involved in the farm business to move elsewhere in order to continue their operations, which might in turn have negatively impacted the growth of local post offices. Figure 3 shows a map depicting extreme weather in California in the 1880s.

The exclusion restriction of this instrument relies on the fact that extreme weather in terms of precipitation and temperature will influence farming and ranching but not non-agricultural firms. One main way in which this condition might be violated is that it could be that the migration of farmers affects the agglomeration economies in a given city, which might result in a lower local demand for products and a depleted labor pool. We argue that the inclusion of population as a control should be able to roughly capture these dynamics.

4.2 Results

4.2.1 OLS

Table 2, Column (1) presents the results only including county and industry fixed effects. In Column (2), we introduce population, in Column (3) we add the initial stock of businesses, in Column (4), we show our preferred specification, where we add transportation controls (i.e., proximity to railways, waterways, and the sea). This specification indicates a strong positive relationship between having a new post office and entrepreneurial growth. In particular, the coefficient suggests that gaining at least one new post office is associated with an average 3.8-percent increase in the growth rate of new firms on across industries. The coefficient across all models are similar and stable.

4.2.2 Robustness Checks and Alternative Explanations

In Table A2 we perform several robustness checks by employing different specifications and controls. First, we control for alternative economic factors that might have influenced the entry decision of companies. Of notable importance at the time were natural resources, and especially gold deposits, which might have had a great influence on economic growth of cities. We add information about the location of historical gold mines (McFaul et al., 2000) and we build a dummy capturing high concentrations[13] of gold sites within 20km of each cities. Our main results remain unchanged, and the post office coefficient is statistically equal to the one found in our preferred specification (Table 2, Column (1)). Another factor that might influence the economic pull of a location is the distance to major city centers. Since settlements might be located close to major cities which might have provided better economic opportunities, we also calculate the distances (in km) between each city and the two major cities existent in California at that time, i.e., Sacramento and San Francisco. Column (2), Table A2, shows that the results are robust to the inclusion of these controls.

We might also be worried of the presence of other amenities and geographical features that might influence entry. In order to make sure that our results are not driven by geographical differences that might exist across cities, we include as variable the latitude and longitude of the centroid of each city, following Donaldson and Hornbeck (2016). The results, reported in Column (3), remain robust.

---

[13]Given the vast presence of gold deposits in California, we use the median value of gold sites to build our dummy.

Another factor that might have impacted entry is the presence of financial institutions in a city. In particular, the presence of banks might be an important determinant of entrepreneurial entry (Bates, 1990). Once again we rely on the R.G. Dun & Co. credit report collection data to build a variable capturing the presence of at least a bank in each city. Results in Column (4) show how the post office coefficient remains unchanged.

In Column (5) we add the initial stock of post offices present in proximity to a city (within 20km) in 1880, while in Column (6) we add the share of businesses in each sector to account for specific cluster-like dynamics. We might also be worried that our results are driven by a few major cities, since some of them (e.g., Los Angeles) tend to experience high growth in entrepreneurship and the number of post offices. In Column (7), we run our baseline regression while excluding the three main cities in California at the time, i.e., Los Angeles, San Francisco, and Sacramento. Results are robust and the coefficient remains statistically unchanged. In Column (8) we employ the growth rate of post offices (as inverse hyperbolic sine) between 1880-1890 as the main independent variable instead of the usual dummy variable. The post office coefficient is positive and significant.

Another threat in our empirical strategy is the presence of reverse causality, i.e., business entry is attracting post offices and not vice-versa. Besides employing an IV strategy that will be described in more detail below, we perform a simple test to check whether future presence of post offices can predict present growth of entrepreneurship. We add the lead value of the number of post offices in 1890 to our preferred specification (Eq. 1). The coefficient of future post offices is not significant and close to zero, while the coefficient capturing the present value of post offices is still positive, significant, and not statistically different from the one we find in Table A2 (Column 9). These results suggest that it is unlikely that an increase in entrepreneurship led to an increase in post offices. Our findings are in line with recent work suggesting that postal offices preceded development (Rogowski et al., 2021) and contributed substantially to advancing the notion of nationhood (John, 1998). A further concern is that more populated cities may have attracted more post offices, on average. To address this, we control for population and also specifically examine whether cities with above- or below-average population experience different rates of new firm growth. It is reassuring to see that the interaction coefficient in Column (8) and the main effect on population is not significant on conventional levels.

4.2.3 Instrumented Results

In Table 3, Columns (1) and (2) present our main IV results, making use of proximity to historic trails to predict the change in post offices. As displayed in Column (1), our first-stage results are positive and precisely estimated, suggesting that, as predicted, cities close to historic trails were more likely to receive a post office. The F-statistics are sufficiently large, with a value of 30.20. When looking at the instrumented coefficient for post office, its magnitude is larger than the one presented in the OLS model (Table 2, Column 5), which may indicate that the IV is shifting the behavior of some industry in some cities where the returns to a post office are higher than average. If the local average treatment effect is larger than the average treatment effect, it is plausible that IV estimates are larger than OLS estimates because of heterogeneity in the sample we are analyzing. Our instrumented specification suggests that the creation of one new post office is associated with a 66% increase in the growth rate of new firms in a given city-industry pair; a city with an average entry of roughly 24 businesses in 1880 would experience the entry of circa 7 new companies a decade later.

In Columns (3) and (4), we display the results from our second instrument. As expected, extreme weather negatively predicts the likelihood that a new post office is formed. The instrumented coefficient in Column (4) is still positive and significant. In Column (5) and (6) we show an instrumented model which includes both proposed instruments. The second stage results in Column (6) confirm the previous results. This model passes the tests for validity of the overidentifying restrictions providing additional support for our model specification.

4.2.4 Heterogeneity Analyses

Next, we explore potential heterogeneity in the relationship between post offices and entrepreneurship. First, we are interested in uncovering which industries experience more entrepreneurial activity as a consequence of the establishment of at least one new post office nearby. Figure 4 displays the coefficients of our baseline regression while filtering for industry. The sectors experiencing positive entry are wagon-makers and locksmiths, mining, fancy goods and varieties, general hardware and ironworks, paints and oil dealers, stables and harnesses, and shoemakers and leather goods. More "unsophisticated" industries, relying on local demand and dealing with non-complex prod-

17

ucts, respectively, did not experience an increase in entrepreneurship. Among these industries we find for instance hospitality, and warehouse and storage (Porter, 2003).

Second, we investigate the distribution of post offices and we examine whether the relative concentration of post offices within an area moderates the effects on entrepreneurship. To do so, we build a measure to capture the spatial concentration of post offices within the vicinity of a city by creating a circle with a 20km radius around each city centroid and dividing land into hexagons with a side length of four kilometers.[14] For each city, we check the distribution of post offices by examining the number of post offices that fall in each hexagon. Hexagons have been widely used in the urban economic literature as they exhibit some favorable traits: in particular, they can be arranged so as to cover any contiguous land, and, since the hexagon is the closest polygon to a circle, it minimize the distance between two or more points of interest (Oberfield et al., 2020). Leveraging a common measure used in point pattern analysis, we define post office concentration as the variance of the distribution of post offices considering the hexagons lying within each city radius. Figure A9 shows hypothetical low- and high-variance counties: when most hexagons have about the same number of post offices, spatial concentration will be low; if most post offices are instead clustered in a few hexagons within a county, spatial concentration will be high. Column (1) in Table A1 shows that cities with high concentration experience lower growth in entrepreneurship than those below the average. In general, this result suggests that the Postal Service expansion was especially beneficial in increasing the number of new firms in places that were otherwise not well connected to communication infrastructure.

4.3   Towards Understanding Potential Mechanisms

To better understand which potential mechanisms underlie our main results, we examine four channels by which the US Postal Service expansion might have affected firm entry. First, as post offices also allowed individuals to send and receive money, we analyze their role as a monetary distribution centers, which may have lowered the costs and increased the convenience of money exchanges. Second, the expansion of the postal service may have help government extend its reach, particularly that of the legal system reducing uncertainties and establishing a stable economy and social relations. Third, access to the post office may have streamlined communication with customers, suppliers, and other stakeholders, enabling entry. Finally, post offices might have facilitated and

---

[14]These hexagons have an area of 41.6 square kilometers.

encouraged knowledge flows; specifically, flows of specialized knowledge that might have affected businesses' best practices, promoted adoption of innovation, and lowered search costs in general.[15]

In order to shed light on the first mechanism, we leverage the fact that not all post offices offered money order services and that post offices with and without that capability might have had heterogeneous effects on performance. We use data from the *United States Official Postal Guide* (of the Post Office Department, 1880), which lists post office locations offering money order services. For each city, we count how many post offices offering money orders are within 20 kilometers of the city center and create a dummy to indicate cities at least one post offices offering money ordering capabilities.

To understand the relationship with government institutions, and the legal system in particular, we create a variable that locates all courts existing in 1880 in the state of California. We leverage data coming from the Bureau of the Census and United States Civil Service Commission (1877) which lists all courts existing by city in California and we create a dummy indicating whether the city has at least one court.

To capture increases in firm entry due to improved communications (that is, access to a faster, cheaper, and more reliable way to communicate), we exploit the fact that different industries will be differently exposed to such improvement. For instance, a business in the banking or finance industries might benefit more than a business in the hardware sector. In general, it is very challenging to quantify how industries might benefit from improved communication. We use the data coming from the work of Forman et al. (2003), where they assess the extent to which different industries have adopted another major innovation that substantially improved communication: the internet. Though this approach has limitations[16], we argue that there is, nonetheless, a strong correlation between the industries that might have benefited most from the US postal system and those that benefited most from the internet. We use the percentage of businesses within an industry that have adopted the internet to proxy for the importance of communication in an industry. We divide our businesses into two categories: above- and below-median internet adoption. It is worth

---

[15]The role of the postal service as a carrier of goods might seem to be a fourth possible mechanism. But we can exclude it because mailing packages larger than an envelope was only possible starting in 1913, after our period of analysis.

[16]In particular, there might be a difference in how the postal system and the internet have affected various industries since the phenomena are somewhat distinct and also because businesses operations have changed substantially since the nineteenth century, possibly making the impact of the internet more pervasive than that of the postal service. In addition, it is generally hard to map nineteenth-century industries to modern, NAICS-based industries.

recalling that we already control for an important contemporary technology that also affected the cost of communication in general—the telegraph. Telegraph lines ran along railroad tracks, which we control for by including measures for the proximity to railroads in all our specifications.

To capture knowledge flows—in particular, flows of specialized knowledge—we leverage data on trade journals and the states in which they were published. By leveraging the specialized nature of trade journals—which targeted precise industries and sectors and provided information about innovations and best practices and other industry-related news—we hope to understand if knowledge flows influenced entry. We collect information about trade journals and their industries from *N.W. Ayer & Son's American Newspaper Annual* (Ayer & Son, 1884), which includes the name, number of issues, location of publication, and—most importantly—the target industry for every trade journal published in the early 1880s. The rationale behind the use of these data is that, if knowledge flows across the country played a role in economic growth, firms in industries that experienced a surge of available knowledge through a high number of trade journals should exhibit higher growth. Since the number of local trade journals might be endogenous to entry and in general to the industry composition of a given location (that is, we might expect to find more journals dedicated to a specific industry being published in locations where this industry is developing and growing), we consider only trade journals published *outside* California, while controlling for all the journals published in California. In this way, we aim to capture knowledge flows originating from outside the state, which had to flow through the postal system in order to be diffused in California and, most importantly, were less likely to be correlated with the endogenous productivity growth in the state. We use the number of trade journals in each industry (published outside California) as a proxy for knowledge flows.

Using these data, we estimate Equation 1 and include interaction variables that best capture each mechanism: financial service, government institutions, communication, and specialized knowledge flow. Table 4, displays the results. Those for money orders are in Column (1), those related to government institutions in column (2), those for communication are in Column (3), and those for trade journals are in Column (4). In all specifications, we include the usual controls from earlier regressions as well as county and industry fixed effects. In Column (1), the interaction term is not statistically significant at the 10-percent level suggesting that access to money orders is not the primary driver of our results. In Column (2), we find a similar result: the interaction term is not statistically significant, suggesting that access to government institutions is not the

primary driver of our results although the positive coefficient suggests that the relationship may be reinforcing. Column (3) considers the role of the US Post as a facilitator of communication: again, the interaction is insignificant, while the main effect of post offices remains positive and significant, suggesting that improvement in communication does not fully explain our result. In Column (4), we find the interaction between post offices and specialized knowledge to be positive and significant, suggesting that access to specialized knowledge may indeed explain the relationship between increases in the stock of post offices and firm financial performance. In particular the presence of one more trade journal catering to a given sector is associated with an increase in entry rate of 0.2% for that sector.

5 Post office location and firm performance

5.1 Data and Empirically Strategy: Firm Performance

In the previous sections, our goal was to establish a plausible causal connection between the expansion of post offices and firm entry. In order to delve deeper into the potential contributions of the Postal Service expansion, we consider a different outcome variable—firm performance—and a different level of analysis—the firm. With the following analyses, we aim to understand if post offices had an impact on firm performance of existing firms as well, besides entry. To measure the performance of existing businesses across decades, we continue to exploit the R.G. Dun & Co. credit report collection. A unique feature of the R.G. Dun & Co. data is a measure capturing the net worth of each firm over time, which is our main performance outcome. This measure falls into 12 categories ranging from less than $1,000 to more than $1 million. [17]

Given data limitations[18] and the substantial collection effort, we focus on tracking the performance of firms between two decades, i.e., 1880 and 1890.[19] Our final sample is composed of 2,459 firms.

Our original performance variable consists of ordered categories capturing firms' net worth. Given the non-linearity of the categories, we transform our categories into a dummy variable capturing

---

[17]The categories are: AA: over $1 million; A+: $750,000–$1million; A: $500,000–$750,000; B+: $300,000–$500,000; B,C+,C: $75,000–$300,000; D: $40,000–$75,000; E: $20,000–$40,000; F: $10,000–$20,000; G: $5,000–$10,000 H: $2,000–$5,000; K: $1,000–$2,000; L-M: less than $1,000. Note that some of the categories had to be aggregated together to keep consistency between the two periods.

[18]Data about firm performance before 1880 is often not included in the R.G. Dun & Co. reports and in general very few firms survived for more than a decade.

[19]Our sample is then composed of businesses that exist both in 1880 and 1890. We match the name of the business and its location in order to flag businesses observed in both years.

whether the firm's net worth increased between the decades of 1880 and 1890 to increase the interpretability of the results.

We then estimate the following first-differenced Probit model:

$$\Delta^{1880-1890} FinancialGrowth_{j,i,s,c} = \alpha + \beta \, \Delta^{1880-1890} PostOffices_i + X_j\Gamma + W_i\Lambda + \phi_s + \lambda_c + e_{i,j,s,c} \qquad (3)$$

where j indexes firms, s sectors, i cities, and c counties. $\Delta^{1880-1890} FinancialGrowth_{i,j}$ is a dummy capturing the growth in financial performance measured as an increase in the firms' net worth of firm j in city i between the beginning and the end of the decade we consider (1880-1890).[20] $\Delta^{1880-1890} PostOffices_j$ is a dummy equal to 1 if the city has experience the birth of a new post office within 20km from its centroid (using the full infrastructure network in 1880). X_j is a set of controls at firm-level. Most businesses in that period were named after their owners, which helps us infer several characteristics about the founders and the company. We control whether the firm is part of multi-establishment (i.e., the entrepreneur has connected businesses in other location)[21], whether there are multiple founders, i.e., we observe the presence of multiple names in the company name or the company name includes words that suggest the presence of multiple co-founders, (the presence of the words " co." and similar terms), whether the company is a family-based business, i.e., whether the word "brothers" or "sons" and their respective abbreviation terms are present in the name of the company, and the ethnic origin of the main founder.[22] W_i is a set of controls at city-level which include population, the close proximity of railroads, waterways, and the sea (within 20km) as well as main cities. Controlling for railways is not only important as they were a significant driver of economic growth per se (Donaldson & Hornbeck, 2016), but they held telegraph lines, which ran along the existing railroad tracks. λ_c are county dummies, ϕ_s are industry dummies, and $e_{i,j,s,c}$ is the error term. We cluster the standard errors at the county level.[23]

For summary statistics of variables at the firm-level, please refer to Table 5. Dun & Co. reports

---

[20]Source: R.G. Dun & Co. credit report collection. Since the original measure of performance is a categorical variable including 12 non-linear categories of net worth, which had to be aggregated in order to provide consistency over time, it is not possible to provide a clear economic interpretation of the underlying coefficient. We avoid these issues by building a dummy capturing growth and running a Probit model that can provide a measure for the probability of growth.

[21]This information is provided directly in the Dun & Co. Credit Reports.

[22]We leverage an AI-based service (Namsor) to infer the origin of the name of the main founder based on the first and last name. As individuals in that period were primarily first or second-generation immigrant, it is fairly straightforward to infer the country of origin by analyzing one's first and last name.

[23]Note that results are robust to the use of standard errors clustered at different level, including robust and at city level.

financial information for roughly 70% of the firms who survive from 1880 to 1890 (2,501 out of 3728). Of the 2,501 firms in our sample, 43.8% experience a performance increase. The median firm in 1880 has a net worth of 3.32 (corresponding to the category $2,000–$5,000). In 1890, this value rises to 3.95. Moreover, 5.1% of the firms are part of a multi-establishment, 8% have multiple founders listed, and 2.7% are family business. Regarding the ethnicity of founders, 4% of the firms were created by an individual of African origin, 1.6% of Asian, 92.7% of European, and 1.8% of Hispanic.

5.2 Results

We estimate Equation 3 and present the results in Table 6. The first column contains the results from a Probit model where the outcome variable is equal to 1 if the firm's net worth increased from 1880 to 1890. Columns (2) and (3) present analogous models that use a continuous measure of financial growth as main dependent variable (and specifically the delta in the value of the financial categories between 1880-1890). Column (2) presents the result of a simple OLS, while Column (3) presents the results from a Ordered Logit specification. In general, all results suggest that an increase in the number of post offices near the city of the focal firm (i.e., within 20 kilometers) is not associated with a clear change in firms' financial performance. This null result is interesting but not surprising, since the net relationship between the US postal service expansion and the performance of firms is not obvious. There are possible countervailing forces at work.

For one, via its impact on firm entry, incumbents may have experienced higher competitive pressures leading to a reduction in financial outcomes for many. For another, certain firms, such as those of the highest quality, with most skin in the game, or those who can leverage access to the mail the most, may have experienced a boost in performance. If this is the case, then we are likely to find either a null or even negative relationship on average, but interesting heterogeneity along dimensions such as quality, experience, and industry. In particular, above and beyond the negative pressures of competition and given our previous results, those firms that benefit the most may be able to leverage specialized knowledge the most.

To shed more light on the potential sources of heterogeneity, we examine the interaction of having a new post office (*1(New Post Office)*) with different competition-, quality- and knowledge-based measures. The results of our preferred Probit model are displayed on Table 7. In Columns (1) and (2), our findings suggest that firms in industries that have a large representation within their

city and those firms in cities with increasing numbers of new entrants, experience large drops in performance. Columns (3) and (4) suggest quality alone, i.e., family businesses and firms started by founders who have already existing establishments is not a sufficient condition to benefit from the opening of a post office. In Columns (5) and (6) we explore the relationship with specific knowledge and technology channels. First, we use our trade journal data and we interact having an out-of-state trade journal in a given sector with our post office measure. The interaction is positive and significant on conventional levels. The main effect turns negative, suggesting that in absence of such a knowledge channel post offices may dampen performance.

Then, we provide more evidence for our hypothesis that knowledge is the fundamental channel through which post offices impact performance by examining a special type of technology of the time. Based on the importance of public knowledge and data for private sector outcomes (Cohen et al., 2002; Nagaraj, 2022), our second measure for knowledge exploits a very specific technology that was distributed by a branch of the government – the newly founded USDA – via mail: seeds. In 1839 the Agricultural Division was established by Congress within the Patent Office. Two decades and a half later, Abraham Lincoln established the independent Department of Agriculture through the Morrill Act, which received Cabinet status in 1889.[24] From its establishment within the Patent Office until 1923, an integral task allotted to what would become the USDA and then was, was to send seeds across the country to the public.[25] The goal was to improve food safety and experiment with the country's crops by relying on an agricultural depository to collect seeds and plants. Handling seeds (and bulbs) was not a straightforward task. It required extensive knowledge to produce successfully (Lyon-Jenness, 2004).

From this, seeds appear to have been an important "technology" throughout the end of the 19th century. We exploit this fact and create a measure that equals to one for those industries that directly handled seeds and their derivatives. These industries are "florists" who dealt with seeds and ran nurseries as well as "grocers" who distributed them. Using this measure, we find suggestive evidence that those firms acting in industries that could benefit from this specific technology experienced boosts in performance (Table 7, Column 6) more than the potential negative competitive pressures. Taken together with our previous results, the findings from our performance heterogeneity analysis suggest that the US Postal Services may, indeed, have played an impactful role both for

---

[24]For more information, see: https://www.usda.gov/our-agency/about-usda.
[25]https://history.house.gov/Blog/2016/May/5-23-photo-seed-distribution/

firm entry and performance in the late 19th century as a channel spreading specialized knowledge and technologies.

6 Discussion and Conclusion

In this paper, we analyze the impact of the US Postal Service expansion of 1880–1900 – an institutional innovation – on regional firm entry and firm performance outcomes. To do so, we exploit a novel dataset and estimate both region- and firm-level outcomes. We first examine the role of the US Postal Service expansion on the rate of new firm entry at the city level. To deal with concerns of omitted-variable bias, we saturate our models with county and industry fixed effects and a host of demographic and transportation-related controls and include instruments based on historic trails and extreme weather conditions. Our findings suggest that an increase in the number of post offices in a city is associated with an increase in entrepreneurship. We further find that those places where existing post offices were less concentrated benefited most from increases in the stock of post offices. As such, post offices may have helped connect more remote areas to—and may also have helped establish—a denser communication grid.

Additional analyses provide further insight into potential mechanisms driving these results. By testing several candidate channels through which the post office may have operated – as a financial service, as an arm of the legal system, as mass communication infrastructure, or as carrier of specialized knowledge – we unveil that access to specialized knowledge in the form of periodicals and magazines is feasibly the most dominant explanation and may have been a critical mechanism through which post offices contributed to increasing entrepreneurship during our study period (Haveman, 2015). Our results therefore suggest that the post office expansion in the late nineteenth century may have very likely contributed substantially to entrepreneurship via the offices' role as "Carrier of (...) Knowledge."

Regarding firm performance, our results are more nuanced. Estimating a similar model, but on the firm-level and controlling for firm- and founder-specific characteristics, we find that increases in competition given new entry exert downward pressure on many incumbents. Although we detect a negative or null relationship for most, we still find positive association with businesses that can rely on knowledge flows coming from out of state. We further detect that actors in those industries relying on a specific type of government distributed knowledge and technology –seeds

and food safety– experienced a boosts in performance. Overall, this suggests that the presence of a post office may have been particularly crucial for those firms who could leverage such specialized knowledge and technology inputs most in their day-to-day operations.

Taken together, our results highlight the critical role of knowledge exchange for entrepreneurship and firm performance. We thereby speak to the literature that stresses the importance of a region's capacity to connect people and ideas for firm growth and innovation (Jacobs, 1969; Gaspar & Glaeser, 1998; Feldman & Link, 2001; Rosenthal & Strange, 2001; Storper & Venables, 2004) and provide important insights on the diffusion of knowledge (Jaffe et al., 1993; Singh, 2005; Thompson & Fox-Kean, 2005; Agrawal et al., 2006; Agrawal et al., 2017; Dutta et al., 2022). Moreover, we shine light on the role of broadly used communication infrastructure in promoting firm entry and boosting firm performance and find suggestive evidence for the importance of wide access to specialized knowledge from other geographies for firm performance. Especially this finding, furthers our understanding critically with regards to the relationship between knowledge sourcing and firm outcomes.

This paper has important implications for managers and policymakers. We present evidence that one reason why firms benefit from co-locating near infrastructure is superior knowledge flows. Our results provide critical information regarding infrastructure investment decisions, which may be especially crucial in the current debate on the pervasive roll-out of broadband, which now serves, as the Post Office once did, as an essential communication and knowledge vehicle. Our findings also suggest that investments made to connect remote places are worthwhile and may have a considerable effect on entrepreneurship and firm performance. Such connection may be particularly valuable to overcome the increasing urban–rural and even intra-urban divide in the US and many other developed or developing economies.

Naturally, our work has limitations. The expansion of the US Postal Service was in many ways a unique endeavor, taking place at a time when the US was undergoing many changes and with much violence and injustice. This is not the focus of our paper, but may require further attention by the literature. In addition, our focus on California, though critical for our empirical strategy, may be too narrow. Follow-on work may provide more detailed evidence pertaining to other states. Finally, we offer only a small window into the role of the Post Office – "Carrier of News and Knowledge" – for entrepreneurship and firm performance. More work will be required—for example, in the form

of a general equilibrium model—in order to precisely estimate the gains that post offices brought to the economic development of regions.

References

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. (2017). *When should you adjust standard errors for clustering?* (Tech. rep.). National Bureau of Economic Research.

Acemoglu, D., Moscona, J., & Robinson, J. A. (2016). State capacity and American technology: Evidence from the nineteenth century. *American Economic Review*, *106*(5), 61–67.

Agrawal, A., Cockburn, I., & McHale, J. (2006). Gone but Not Forgotten: Knowledge Flows, Labor Mobility, and Enduring Social Relationships. *Journal of Economic Geography*, *6*(5).

Agrawal, A., Galasso, A., & Oettl, A. (2017). Roads and Innovation. *Review of Economics and Statistics*, *99*(3), 417–434.

Alcácer, J., & Chung, W. (2007). Location Strategies and Knowledge Spillovers. *Management Science*, *53*(5), 760–776.

Almeida, P., & Kogut, B. (1999). Localization of knowledge and the mobility of engineers in regional networks. *Management Science*, *45*(7), 905–917.

Aneja, A., & Xu, G. (2022). *Strengthening state capacity: Postal reform and innovation during the gilded age* (Working Paper No. 29852). National Bureau of Economic Research. https://doi.org/10.3386/w29852

Atack, J. (2015). Historical Geographic Information Systems (GIS) database of Steamboat-Navigated Rivers During the Nineteenth Century in the United States.

Atack, J. (2016). Historical Geographic Information Systems (GIS) database of U.S. Railroads.

Atack, J. (2017). Historical Geographic Information Systems (GIS) database of Nineteenth Century U.S. Canals.

Ayer, N., & Son. (1884). *American Newspaper Annual and Directory*. NW Ayer; Son.

Balboni, C. A. (2019). *In harm's way? Infrastructure investments and the persistence of coastal cities* (Doctoral dissertation). The London School of Economics and Political Science (LSE).

Bancroft, A. (1882). Map of California and Nevada. *A.L. Bancroft Co.*

Bates, T. (1990). Entrepreneur human capital inputs and small business longevity. *The Review of Economics and Statistics*, 551–559.

Bazzi, S., Fiszbein, M., & Gebresilasse, M. (2020). Frontier culture: The roots and persistence of "rugged individualism" in the United States. *Econometrica*, *88*(6), 2329–2368.

Bellemare, M. F., & Wichman, C. J. (2020). Elasticities and the inverse hyperbolic sine transformation. *Oxford Bulletin of Economics and Statistics*, *82*(1), 50–61.

Bleakley, H., & Lin, J. (2012). Portage and Path Dependence *. *The Quarterly Journal of Economics*, *127*(2), 587–644.

Blevins, C. (2021). *Paper Trails: The US Post and the Making of the American West.* Oxford University Press.

Carlino, G. A., Chatterjee, S., & Hunt, R. M. (2007). Urban Density and the Rate of Invention. *Journal of Urban Economics, 61*(3), 389–419.

Cernea, M. (1987). Farmer organizations and institution building for sustainable development.

Cohen, W. M., Nelson, R. R., & Walsh, J. P. (2002). Links and Impacts: The Influence of Public Research on Industrial R&D. *Management Science, 48*(1), 1–23.

Cook, E. R., Seager, R., Heim Jr, R. R., Vose, R. S., Herweijer, C., & Woodhouse, C. (2010). Megadroughts in North America: Placing IPCC projections of hydroclimatic change in a long-term palaeoclimate context. *Journal of Quaternary Science, 25*(1), 48–61.

Démurger, S. (2001). Infrastructure development and economic growth: An explanation for regional disparities in China? *Journal of Comparative economics, 29*(1), 95–117.

Donaldson, D., & Hornbeck, R. (2016). Railroads and American economic growth: A "market access" approach. *The Quarterly Journal of Economics, 131*(2), 799–858.

Duranton, G., & Turner, M. A. (2012). Urban growth and transportation. *Review of Economic Studies, 79*(4), 1407–1440.

Dutta, S., Armanios, D. E., & Desai, J. D. (2022). Beyond spatial proximity: The impact of enhanced spatial connectedness from new bridges on entrepreneurship. *Organization Science, 33*(4), 1620–1644.

Feldman, M. P., & Link, A. N. (2001). *Innovation policy in the knowledge-based economy* (Vol. 23). Springer Science & Business Media.

Forman, C., Goldfarb, A., & Greenstein, S. (2003). Which industries use the Internet? *Organizing the new industrial economy.* Emerald Group Publishing Limited.

Gaspar, J., & Glaeser, E. L. (1998). Information Technology and the Future of Cities. *Journal of Urban Economics, 43*(1), 136–156.

Gibson, C. (2007). Population total by township and place for California counties: 1860 to 1950.

Hanson, G. H. (2005). Market potential, increasing returns and geographic concentration. *Journal of International Economics, 67*(1), 1–24.

Haveman, H. A. (2015). *Magazines and the Making of America.* Princeton University Press.

Head, K., & Mayer, T. (2011). Gravity, market potential and economic development. *Journal of Economic Geography, 11*(2), 281–294.

Hornbeck, R., & Rotemberg, M. (2019). *Railroads, reallocation, and the rise of American manufacturing* (tech. rep.). National Bureau of Economic Research.

Jacobs, J. (1969). *The Economy of Cities.* New York: Vintage Books.

Jaffe, A. B., Trajtenberg, M., & Hall, B. (1993). Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *Quarterly Journal of Economics*, *108*(3), 577–598.

Jaworski, T., & Kitchens, C. T. (2019). National policy for regional development: Historical evidence from appalachian highways. *Review of Economics and Statistics*, *101*(5), 777–790.

John, R. (1998). *Spreading the News: The American Postal System from Franklin to Morse*. Harvard University Press. https://books.google.com/books?id=EUooEAAAQBAJ

Khan, B. Z. (2005). *The Democratization of Invention: Patents and copyrights in American economic development, 1790-1920*. Cambridge University Press.

Li, H., Sheng, Y., & Yan, X. (2020). Empirical research on the level of institutional innovation in the development of china's high-tech industry. *IEEE Access*, *8*, 115800–115811.

Logan, J. R., Jindrich, J., Shin, H., & Zhang, W. (2011). Mapping America in 1880: The urban transition historical GIS project. *Historical Methods*, *44*(1), 49–60.

Lyon-Jenness, C. (2004). Planting a Seed: The Nineteenth-Century Horticultural Boom in America. *The Business History Review*, *78*(3), 381–421.

McFaul, E., Mason, G., Ferguson, W., & Lipin, B. (2000). *US Geological Survey mineral databases; MRDS and MAS/MILS* (tech. rep.). US Geological Survey.

Morgan, D. L. (1953). *Jedediah Smith and the Opening of the West* (Vol. 184). U of Nebraska Press.

Nagaraj, A. (2022). The private impact of public data: Landsat satellite maps increased gold discoveries and encouraged entry. *Management Science*, *68*(1), 564–582.

Oberfield, E., Rossi-Hansberg, E., Sarte, P.-D., & Trachter, N. (2020). *Plants in space* (Working Paper No. 27303). National Bureau of Economic Research. https://doi.org/10.3386/w27303

of the Post Office Department, A. (1880). *the United States Official Postal Guide*. Boston: Houghton, Osgood, Company.

Palmer, W. C. (1965). *Meteorological drought* (Vol. 30). US Department of Commerce, Weather Bureau.

Porter, M. (2003). The economic performance of regions. *Regional Studies*, *37*(6-7), 549–578.

Redding, S., & Venables, A. J. (2004). Economic geography and international inequality. *Journal of International Economics*, *62*(1), 53–82.

Redding, S. J., Sturm, D. M., & Wolf, N. (2007). *History and industrial location: Evidence from german airports* (tech. rep.). CEPR Discussion Paper, 6345.

Rogowski, J. C., Gerring, J., Maguire, M., & Cojocaru, L. (2021). Public infrastructure and economic development: Evidence from postal systems. *American Journal of Political Science*.

Rosenthal, S. S., & Strange, W. C. (2001). The determinants of agglomeration. *Journal of Urban Economics*, *50*(2), 191–229.

Singh, J. (2005). Collaborative Networks as Determinants of Knowledge Diffusion Patterns. *Management Science*, *51*(5), 756–770.

Singh, J., & Marx, M. (2013). Geographic constraints on knowledge spillovers: Political borders vs. spatial proximity. *Management Science*, *59*(9), 2056–2078.

Sokoloff, K. L. (1988). Inventive Activity in Early Industrial America: Evidence From Patent Records, 1790-1846. *The Journal of Economic History*, *48*(4), 813–850.

Storper, M., & Venables, A. J. (2004). Buzz: Face-to-face contact and the urban economy. *Journal of Economic Geography*, *4*(4), 351–370.

the Bureau of the Census and United States Civil Service Commission. (1877). *Official Register of the United States: Persons in the Civil, Military, and Naval Service, Exclusive of the Postal Service*. U.S. Government Printing Office.

Thompson, P., & Fox-Kean, M. (2005). Patent citations and the geography of knowledge spillovers: A reassessment. *American Economic Review*, *95*(1), 450–460.

Tables and Figures

Figure 1: Post Office Expansion in California (1880-1890)
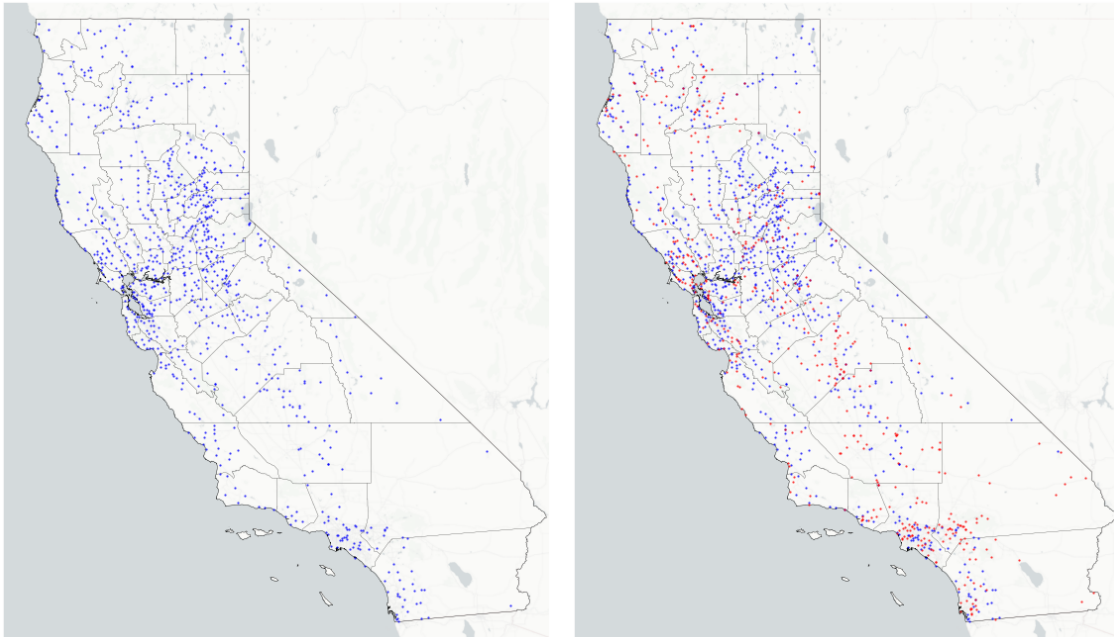
Notes: This figure displays the location of US post offices in 1880 (left) and 1890 (right). The blue dots represent post office locations in 1880 and the red dots represent new post offices that were formed between 1880 and 1890.

Figure 2: California Historical Trails

Notes: In this figure, we map out the original Historic Trails in California. On the left, we depict the trails spanning from Nevada and reaching to Oregon. The small green triangles represents cities in California in 1880. On the right, we present a zoomed in version with topographic detail. The bold red lines depict the trails.

Figure 3: Extreme Weather in California (1880-1890)
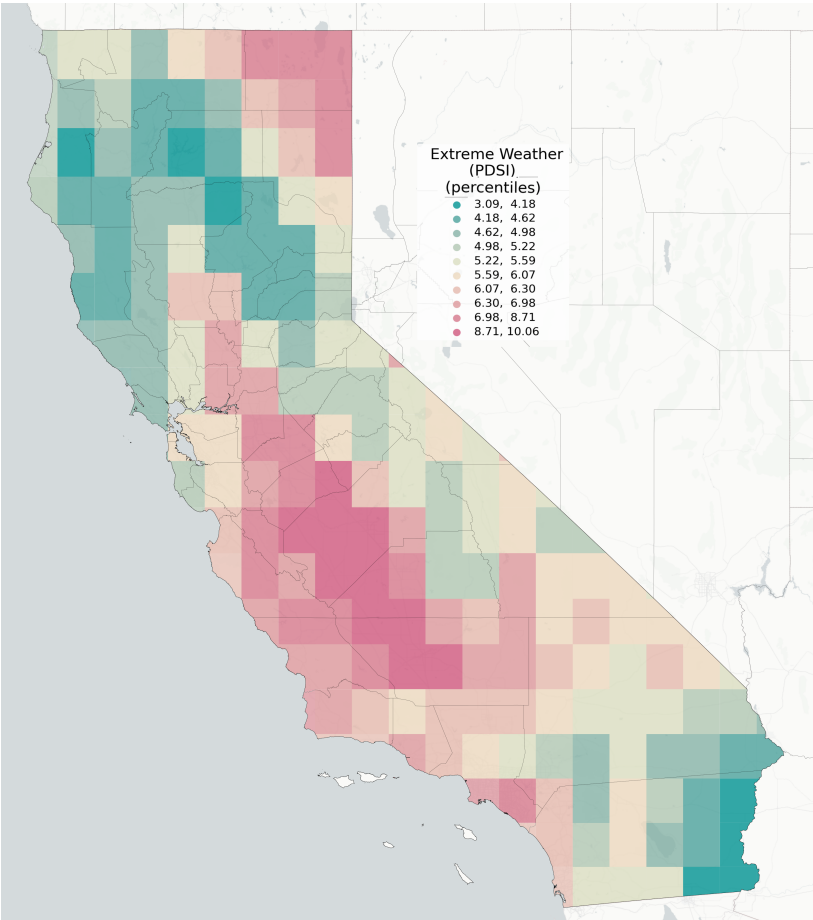
Notes: This figure displays extreme weather related to precipitation events. It is calculated as the difference between the maximum and minimum value of the Palmer Drought Severity Index (PDSI) in the period 1880-1890.

Figure 4: Entry by industry (California, 1880-1890)
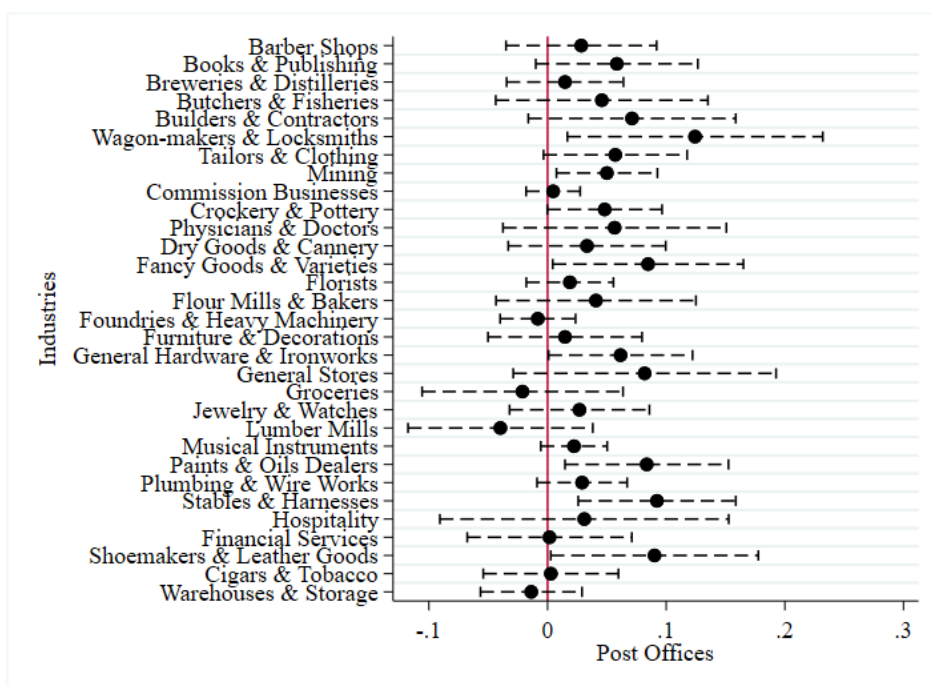
Notes: This figure shows the entry coefficients from estimating Equation 1 by industry (excluding industry fixed effects). The x-axis denotes the magnitude of the coefficient and the y-axis displays the industry names. In all regressions we include controls related to population, the initial stock of businesses in 1880, and the proximity to railroads and waterways. In each model, we add county fixed effects and cluster standard errors (in parentheses) at the city level.

Table 1: Summary Statistics - California sample

City-Sector	min	p25	mean	p50	p95	max
New businesses 1880	0.00	0.00	0.78	0.00	3.00	726.00
New businesses 1890	0.00	0.00	1.00	0.00	3.00	1491.00
1(Post Offices)	0.00	0.00	0.35	0.00	1.00	1.00
Population 1880	6	90	1,177.34	239	3000	179,040
Stock of Firms 1880	1	2	25.73	6	79	3882
1(Railroads)	0.00	0.00	0.50	1.00	1.00	1.00
1(Waterways)	0.00	0.00	0.11	0.00	1.00	1.00
Observations	19,065					

Notes: This table displays summary statistics on the city-sector level for the main variables in our model.

Table 2: Post office location and firm entry: OLS (Δ 1880-1890)

Dep.Var.: Entrepreneurship	(1)	(2)	(3)	(4)
Model	OLS	OLS	OLS	OLS
1(Post Offices)	0.037**	0.034*	0.037**	0.038**
	(0.018)	(0.018)	(0.017)	(0.017)
Population	No	Yes	Yes	Yes
Initial Stock of Businesses	No	No	Yes	Yes
Rail & Waterways	No	No	No	Yes
Observations	19065	19065	19065	19065

Notes: This table displays the results from estimating Equation 1. *1(Post Offices)* equals to one if the stock of post office within a 20km distance from the center of each city increased. Column (1) to (4) progressively include controls. We show some results with only county and industry fixed effects (Column 1). In column (2) we further add population, the initial stock of businesses in 1880 (column 3), and the proximity to railroads and waterways (column 4). We cluster standard errors (in parentheses) at the city level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3: Post office location and entrepreneurship: Instrumented Results (Δ 1880-1890)

Dep.Var.	1(Post Offices)	Entrepr.	1(Post Offices)	Entrepr.	1(Post Offices)	Entrepr.
Model	First Stage	IV	First Stage	IV	First Stage	IV
Instrument	Trails	Trails	Weather	Weather	Weather & Trails	Weather & Trails
	(1)	(2)	(3)	(4)	(5)	(6)
Historic Trails	0.001***				0.001***	
	(0.000)				(0.000)	
Extreme Weather			-0.085***		-0.075***	
			(0.000)		(0.011)	
1(Post Offices)		0.661***		0.404***		0.456***
		(0.229)		(0.142)		(0.125)
Observations	19065	19065	19065	19065	19065	19065
F-Stat	30.20		65.99		43.99	
Hansen's J						
Stat. P-value						0.527

Notes: This table displays some robustness checks and the results from estimating Equations 2, using two different instruments (i.e., historical trails and extreme weather). *1(Post Offices)* equals one if the stock of post office within a 20km distance from the center of each city increased. As controls at city-level we include population, the stock of businesses, and the proximity to railroads, and waterways. To each model, we also add county and industry fixed effects. Columns (1)-(2) show the first stage and instrumented results of our preferred instrument, i.e., historical trails. Column (3)-(4) show results considering an alternative instrument, extreme weather. Finally in Column (5)-(6) we present some results including both of our instruments. We report the F-statistics where applicable and the p-value obtained from the Hansen J Statistic, which tests the validity of the overidentifying restrictions in column (6). Standard errors (in parentheses) are at the city level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Potential Mechanisms

New Firm Growth	(1)	(2)	(3)	(4)
1(Post Offices)	0.040*	0.029	0.047**	0.021
	(0.022)	(0.018)	(0.021)	(0.015)
Financial Services	-0.002			
	(0.040)			
1(Post Offices) × Financial Services	-0.008			
	(0.029)			
Government Institution		-0.042		
		(0.042)		
1(Post Offices) × Gov. Inst.		0.118		
		(0.081)		
Communication			0.012	
			(0.009)	
1(Post Offices)× Communication			-0.017	
			(0.014)	
Out-of-State Trade Journals				0.001
				(0.001)
1(Post Offices) × Out-of-State Trade Journals				0.002**
				(0.001)
Observations	19065	19065	18450	19065

Notes: This table displays the results from estimating Equation 3. *New PO* equals to one if the stock of post office within a 20km radius around the center of each city increased. The controls at city-level we include are population, the stock of businesses, and the proximity to railroads and waterways. The results including and interaction with money orders are presented in Column (1), those related to government institutions are in Column (2), the results pertaining to industries that benefited from internet enhancement are presented in Column (3), and the results pertaining to trade journals are presented in Column (4). In Column (4) we further control for in-state trade journals. In each model, we add county and industry fixed effects and cluster standard errors (in parentheses) at the county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Summary Statistics - Firm Level

Firm	min	p25	mean	p50	p95	max
1(Financial Growth)	0	0	0.44	0	1	1.00
Net Worth 1880 (categ.)	1	1	3.32	3	8	12.00
Net Worth 1890 (categ.)	1	2	3.95	3	8	12.00
1(Post Office)	0	0	0.59	1	1	1.00
Population	6	896	60,145	5,	179,040	179,040
1(Railroads)	0	1	0.81	1	1	1.00
1(Waterways)	0	0	0.14	0	1	1.00
Multiple Establishments	0	0	0.05	0	1	1.00
Multiple Founders	0	0	0.08	0	1	1.00
Family Business	0	0	0.03	0	0	1.00
Founder						
- African	0	0	0.04	0	0	1
- Asian	0	0	0.02	0	0	1
- European	0	1	0.93	1	1	1
- Hispanic	0	0	0.02	0	0	1
Observations	2,459					

Notes: This table displays summary statistics at the firm-level.

Table 6: Post office location and performance: OLS and Probit estimates (Δ 1880-1890)

Financial Growth	(1) 0/1 Probit	(2) Δ OLS	(32) Δ OLogit
1(Post Office)	-0.186	-0.0827	-0.312
	(0.122)	(0.0934)	(0.199)
Firm Controls	Yes	Yes	Yes
City Controls	Yes	Yes	Yes
County FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Observations	2450	2459	2459

Notes: This table displays the results from estimat-
ing Equations 3. *1(Post Office)* equals to one if the
stock of post office within a 20km radius around
the center of each city increased. We include a
set of company- and city-specific controls. We cap-
ture if the company has multiple establishment, was
founded by multiple individuals, is a family business
and the ethnicity of the founders. At city level, we
control for population and the proximity to railroads
and waterways. In each model, we add county and
industry fixed effects and cluster standard errors (in
parentheses) at the county level. * $p < 0.10$, **
$p < 0.05$, *** $p < 0.01$.

Table 7: Performance Heterogeneity: Probit Models

1(Financial Growth)	Competition		Quality/Reput.		Knowl./Techn.	
	(1)	(2)	(3)	(4)	(5)	(6)
1(New PO)	-0.124	-0.0608	-0.157	-0.157	-0.383***	-0.172
	(0.118)	(0.121)	(0.118)	(0.119)	(0.139)	(0.116)
Industry Concentration	-0.642**					
	(0.313)					
1(New PO) × Industry Concen.	-0.782*					
	(0.460)					
>Median # New Entrants		0.005*				
		(0.003				
1(New PO) × >Median # New Ent.		-0.005*				
		(0.003)				
Family Business			-0.0407			
			(0.446)			
1(New PO) × Family Business			0.504			
			(0.474)			
Multiple Establishments				-0.108		
				(0.288)		
1(New PO) × Multiple Establ.				0.229		
				(0.300)		
Out-of-State Trade Journal					-0.429***	
					(0.105)	
1(New PO) × Out-of-State Trade Journ.					0.318**	
					(0.139)	
Seeds						0.797
						(0.613)
1(New PO)× Seeds						0.549**
						(0.223)
Firm-level controls	Yes	Yes	Yes	Yes	Yes	Yes
City-level controls	Yes	Yes	Yes	Yes	Yes	Yes
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	No	Yes
Observations	2450	2450	2450	2450	2450	2450

Notes: This table displays the results from estimating the Probit Model displayed in Table 6, but with a battery of interactions. *1(Post Office)* equals to one if the stock of post office within a 20km radius around the center of each city increased. We include a set of company- and city-specific controls. We capture if the company has multiple establishment, was founded by multiple individuals, is a family business and the ethnicity of the founders. At city level, we control for population and the proximity to railroads and waterways. In each model, we add county and industry fixed effects (except for the journal-related regression in column 5, which relies on industry variation) and cluster standard errors (in parentheses) at the county level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Online Appendix for:
You've Got Mail! The Late 19th Century US Postal Service Expansion, Entrepreneurship, and Firm Performance

A   Digitization of 1880 road networks

In order to measure the distance between each city center and each post office, we rely on distance calculated using the full network of trails and roads (defined as roads where stagecoaches could circulate) existent in 1882 as well as railways and waterways of the time (Atack, 2015, 2016, 2017). This type of measurement more accurately captures the actual distance that individuals had to travel to reach a post office, as it takes into account geographical features of the land, such as mountains, gorges and other water bodies, which could not be easily crossed (see Figure A1 for a graphical example).

In order to include this network-based distance in our paper, we had to digitize what is thought to be one of the most comprehensive map of California in 1880 (Bancroft, 1882) from scratch. One of the many obstacles in digitizing relatively old map is the fact that most of them do not conform to any common modern world projections, given the means to produce a perfectly to scale map were limited. We geo-reference the map by using points that identify land features and borders which did not change from 1880 to today, and reassigning their latitude and longitude values. For instance, we leverage several points located on the California coast, given that the shape and position of the shoreline has remain unchanged over time. We geo-reference more than 100 points to shift the whole map from the old to the new projection (Mercator). This step is necessary to ensure this map is aligned with the other set of spatial data employed in the paper (e.g., post offices) and that distances (in km) are correctly calculated.

Once the new map gets transformed and "distorted" in position to fit the new projection system, we proceed to carefully draw each road and trail present on the map by hand. Besides roads and trails, we also pinpoint the position of main cities, which was sometimes slightly different than their current position, mainly because their spatial extent was much more limited in 1882. After digitizing every road and trail present in California in 1882, we build a network of roads, railways, and waterways by joining all the individual spatial lines, which we can then use to calculate distances between points. Given not all points of interests are exactly located on a road, we calculate distance by taking the Euclidean distance from a given origin point to the closest point located on the road network and calculate the network distance from there to the closest point on the network to the destination point. Finally, we sum the Euclidean distance between this point and the location of the destination point.

Figure A1: Euclidean vs. Street Network-based distance
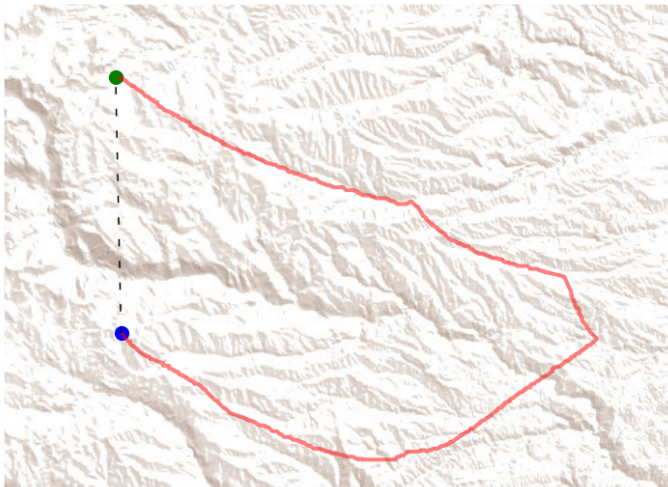

Notes: The figure shows the distance between two points using two different methods. The first is Euclidean distance (black dotted line) while the second is a distance calculating using the full network of streets, railroads, and waterways in 1880 (red line). The figures uses a post office and a point located in Alpine county. The figures demonstrates that in some regions, especially the ones rich in mountainous terrains and bays, the Euclidean distance can severely underestimate the distance between points, as some geographical features impede to simply reach the desired destination through a straight line.

**Figure A2:** Digitized 1880 network of roads, railways, and waterways

Notes: This figure displays the main network of roads, trails, railways, and waterways existent in California in 1882, which has been digitized manually from Bancroft (1882).

**Figure A3:** Entrepreneurial Growth by city and industry (California, 1880-1890)
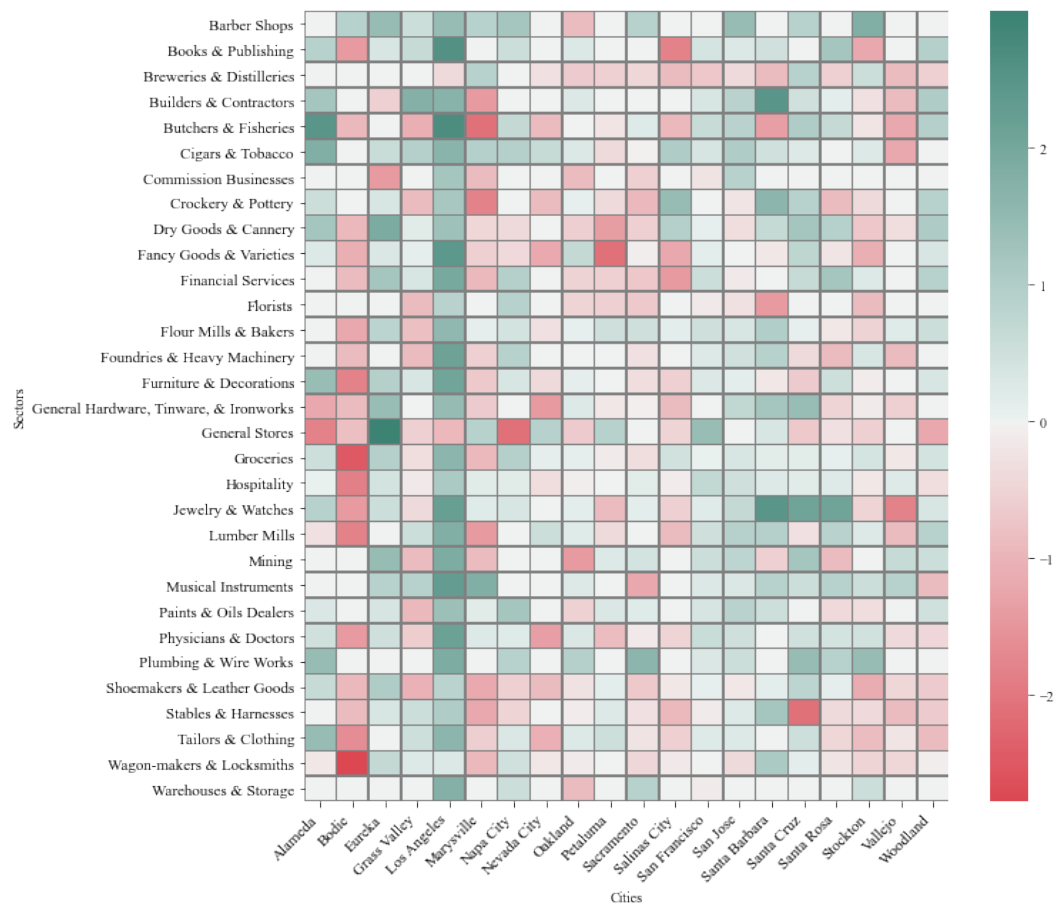
Notes: This figure shows entrepreneurial growth by city and industry between the years 1880-1890. Growth is calculated using an inverse hyperbolic sine transformation. White squares indicate no growth, while green and red squared indicate positive and negative growth, respectively.

Figure A4: Top 10 growing and declining cities in terms of entrepreneurial entry (California, 1880-1890)
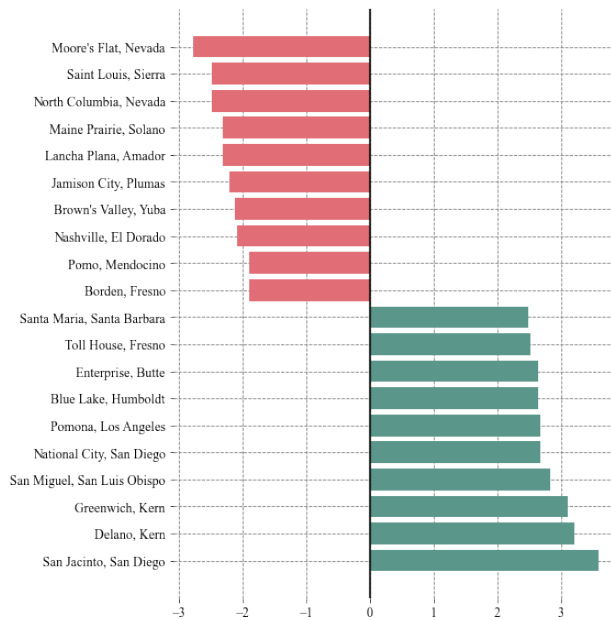

Notes: This figure shows the top 10 cities that experienced the highest (in green) and lowest (in red) growth in entrepreneurship

**Figure A5:** New businesses by industry (California, 1880-1890)
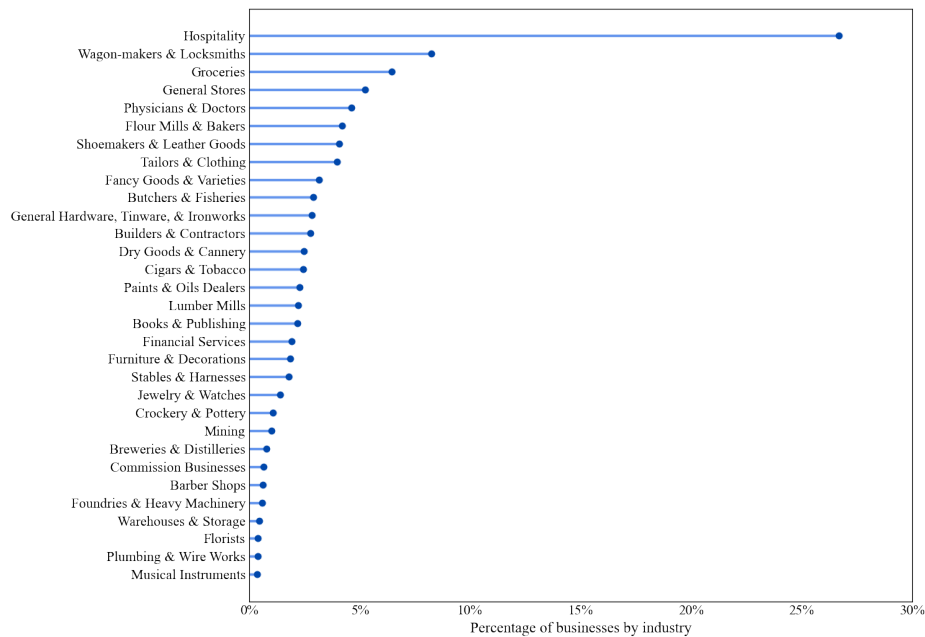
Notes: This figure shows the distribution of firms by industry for companies born in California in the years between 1880 and 1890.

**Figure A6:** Delta in number of Post Offices (1880-1890)
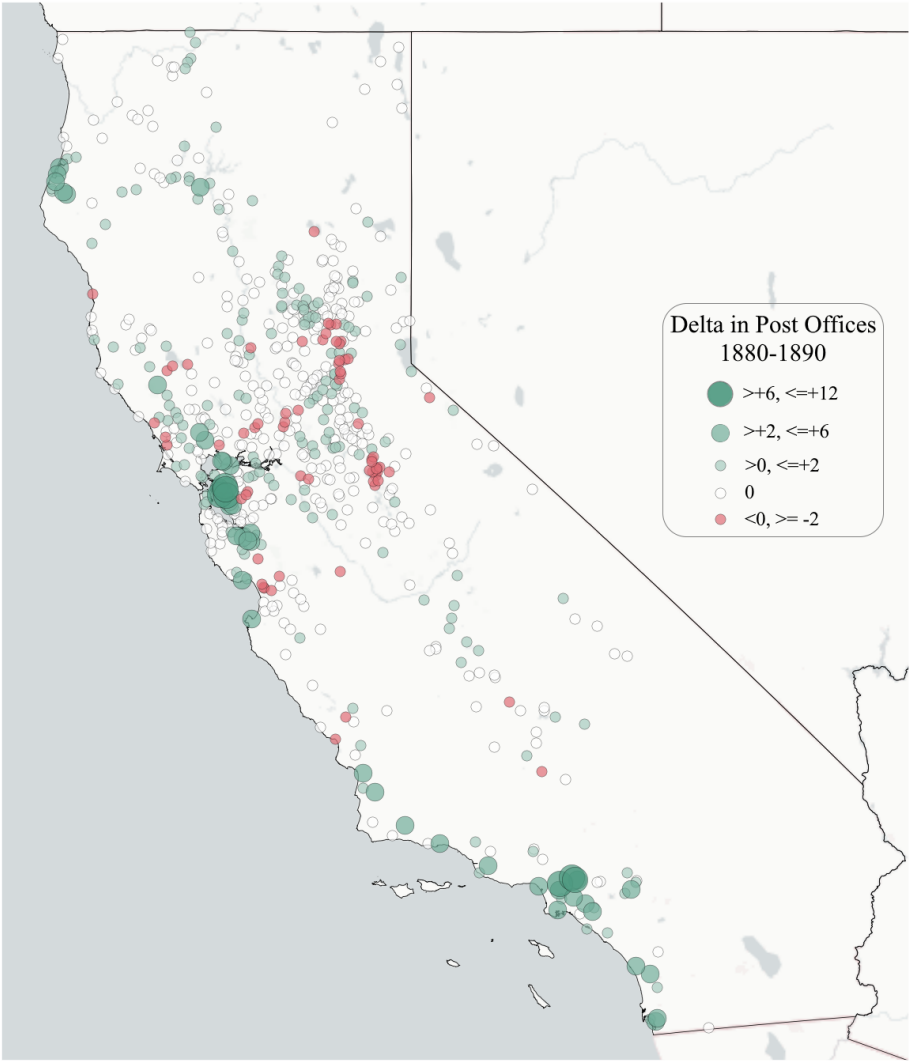
Notes: This figure displays the change in the number of post offices between 1880 and 1890. Green and red dots indicate cities that have lost or gained post offices, respectively. White dots indicate cities that did not experience a change in the number post offices. The size of the dots and their color are proportional to the number of post offices gained and lost.

**Figure A7:** Post Office Expansion Over Time in the US
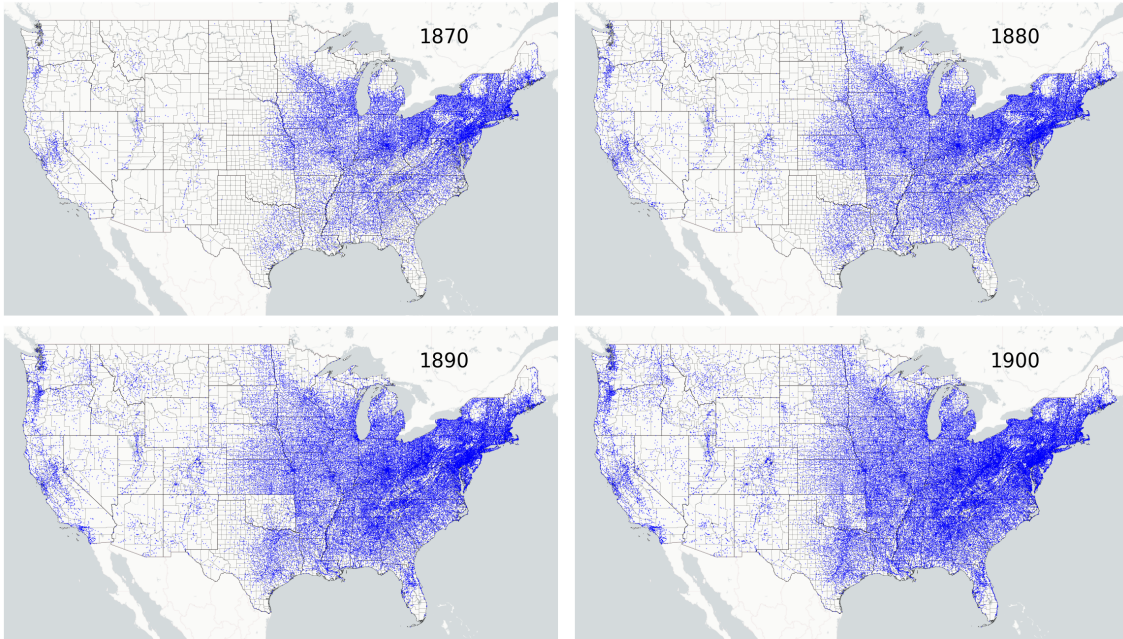
Notes: This figure displays the location of US post offices for each decade spanning from 1880 to 1900 (clockwise from top left). Each dot represents a post office location.

Figure A8: Population Density and new post offices in the West: Binned Scatter Plot
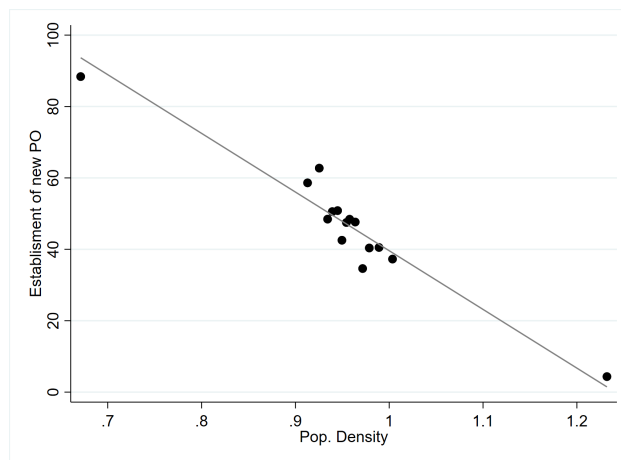
Notes: This figure displays the relationship between population density and the establishment of new post offices at a county level for the period 1880-1900. The results are displayed in the form of a binned scatter plot which includes fixed effects and controls. In particular, we include a set of controls capturing the share of foreign born individuals, the share of white individuals on the total population, the value of farms, the ratio of female to male population, the number of employees in manufacturing on population, the share of republican votes, whether the county was at the "frontier" (Bazzi et al., 2020), and the proximity to railways and waterways. Year and county fixed-effects are also included. Western states include: California, Washington, Oregon, Nevada, Utah, Arizona, New Mexico, Montana, Wyoming, and Colorado.

Figure A9: Concentration of post offices by city
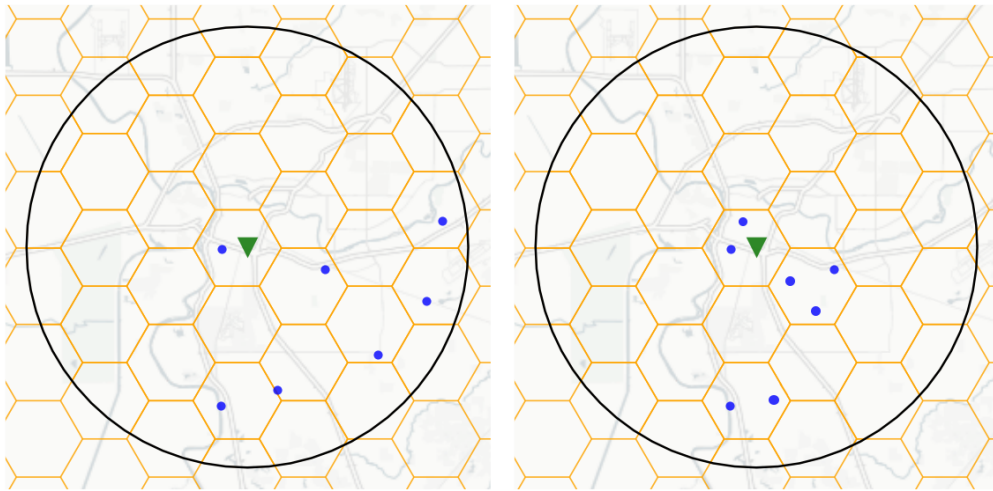
Notes: This figure displays two hypothetical examples with different post office concentration for the city of Sacramento. The figure on the left displays the actual distribution of post offices, which is spread out in multiple hexagons and has high variance of post office within the 20km radius surrounding the city center. The figure on the right shows an hypothetical more concentrated distribution, where variance will be high. Hexagons have areas of roughly 42km with a side length of 4km.

Table A1: Post office location and firm entry: Interaction with concentration measures

New Firm Growth	(1)
1(Post Offices)	0.059***
	(0.020)
Spatial Conc.	0.115
	(0.076)
1(Post Offices) × Spatial Conc.	-0.284*
	(0.156)
Observations	17,887

Notes: This table displays the results from estimating Equations 1 and including interactions that considers the spatial concentration of post offices within a 20km from the city center. *1(Post Offices)* equals to one if the stock of post office within a 20km radius around the center of each city increased. The controls at city-level include population, the stock of businesses, and the proximity to railroads and waterways. In each model, we add county and industry fixed effects and cluster standard errors (in parentheses) at the city level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A2: Post office location and firm entry: Robustness Checks (Δ 1880-1890)

Spec. Dep.Var.: Entrepren.	Baseline (1)	Baseline (2)	Baseline (3)	Baseline (4)	Baseline (5)	Baseline (6)	Baseline (7)	Log-Log (8)	Rev. Causality (9)
1(Post Office)	0.037** (0.017)	0.037** (0.017)	0.040** (0.017)	0.039** (0.017)	0.038** (0.017)	0.038** (0.017)	0.036** (0.017)		0.039** (0.018)
Growth of POs								0.027*** (0.007)	
Lead Post Office (1890)									-0.001 (0.006)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Gold Mines	Yes	No	No	No	No	No	No	No	No
Distance to main cities	No	Yes	No	No	No	No	No	No	No
Latitude & Longitude	No	No	Yes	No	No	No	No	No	No
Bank City	No	No	No	Yes	No	No	No	No	No
Initial Stock of POs	No	No	No	No	Yes	No	No	No	No
Cluster Dynamics	No	No	No	No	No	Yes	No	No	No
Major Cities Excluded	No	No	No	No	No	No	Yes	No	No
Observations	19,065	19,065	19,065	19,065	19,065	19,065	18,972	19,065	19,065

Notes: This table displays the results from estimating Equations 3 including further robustness checks. *1(Post Office)* equals to one if the stock of post office within a 20km distance from the center of each city increased. Columns (1) to (7) include a variety of controls that might be driving our results. In Column (*) we present a log-log model where both the dependent and independent variable are in log form. In column (9), we show the results from a reverse causality test. In all models, we control for population, the stock of businesses, and the proximity to railroads and waterways. We further add county and industry fixed effects and cluster standard errors (in parentheses) at the city level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.