State Capacity as an Organizational Problem

State Capacity as an Organizational Problem*

Nicola Mastrorocco
University of Bologna

Edoardo Teso
Bocconi, Northwestern Kellogg, NBER, CEPR

May 20, 2026

Abstract

We investigate how technologies that reduce the costs of monitoring by central authorities have shaped the historical transition from small patrimonial states to large bureaucratic organizations. Our analysis is based on a novel dataset that traces changes in the organizational structure and geographic presence of the U.S. federal government over the nineteenth century. To identify causal effects, we develop a new identification strategy that exploits the expansion of the railroad network as a source of variation in the travel time—and thus monitoring costs—between Washington D.C. and other locations. We present three main findings. First, reductions in travel time to Washington D.C. significantly increased the likelihood of federal government presence in a location. Second, this effect is stronger for occupations and tasks characterized by more severe agency problems. Third, decreases in travel time to Washington D.C. are associated with a decline in patrimonial features of the federal government in the location, in line with enhanced monitoring capacity reducing dependence on personal trust and connections.

*Mastrorocco: University of Bologna (email: nicola.mastrorocco@unibo.it); Teso: Bocconi University, and Kellogg School of Management, Northwestern University (email: edoardo.teso@unibocconi.it). For comments we are indebted to Charles Angelucci, Oriana Bandiera, Robert Gibbons, Simon Hix, Massimo Morelli, Elias Papaioannou, Torsten Persson, Nancy Qian, Martin Rotemberg, Shanker Satyanath, Jim Snyder, Stephane Wolton, Guido Tabellini, Michael Ting, Tianyi Wang, our discussants Erika Deserranno, Emeric Henry, Ben Marx, and Ruixue Jia, and seminar participants at Berkeley, Bocconi, Cattolica University of Milan, Chicago Booth, Columbia University, EUI, Harvard, IIES Stockholm, ITAM, Northwestern, NYU, Rochester, Trinity College Dublin, UCSD, the University of New South Wales, the Applied Economics Workshop in Petralia Sottana, the Stanford Call to Service Conference, the LSE ODDO conference, the Nottingham NICEP Conference, the NBER DAE Summer Meeting, the NBER Organizational Economics meeting, and the Political Economy workshop in Brunico. We are grateful to the coeditor and the anonymous referees for their valuable feedback and insightful suggestions, which have significantly improved this paper. Siddhant Agarwal, Michael Giordano, Alessio Giorgetta, Jimin Han, Benjamin Lualdi, and Navika Mehta provided outstanding research assistance. All errors are solely our responsibility.

1 Introduction

The structure and functions of modern states differ significantly from the ways in which states were organized throughout the first several millennia of human civilization (Weber, 1978; Finer, 1997). Historically, states employed relatively few officials directly and faced substantial challenges in extending their authority to areas distant from the central seat of power. The organization of these early states was typically patrimonial: officials were appointed based on personal loyalty and networks of patronage, and served at the rulers' discretion. In contrast, modern states administer vast territories and directly employ large numbers of officials. Their organizational structure is bureaucratic, characterized by formal rules that govern the careers of officials. Recruitment and promotion are based on merit rather than personal connections to superiors. While historians have proposed a range of explanations for this institutional evolution, systematic empirical evidence remains limited.

This paper examines the role of technologies that reduce agency costs between rulers and officials as a relevant driver of this historical transformation. When transportation and communication costs are high, central authorities face significant difficulties in monitoring remote officials and ensuring their compliance with the state's interests. Under such conditions, patrimonial arrangements can help mitigate agency problems by reducing the need for monitoring through the appointment of trusted individuals from personal networks. However, this ultimately constraints the expansion of the state. Technological innovations that lower monitoring costs make bureaucratic organization more feasible and enable states to expand in size and territorial reach.[1]

Empirically testing this hypothesis presents two major challenges. First, it requires detailed data on the presence of the state across space, covering a long stretch of time, and within a historical context where monitoring technologies were undergoing significant change. Second, it requires exogenous variation in technologies that reduce monitoring costs—variation that is not confounded by other factors influencing state development.

We address both challenges by focusing on the expansion of the U.S. federal government during the nineteenth century.[2] Our analysis draws on a novel dataset documenting changes in the presence and organizational structure of the federal government over time and across the U.S. territory. To identify the effect of monitoring capacity, we develop a methodology

[1]This hypothesis dates back to Weber (1978) and has been further developed in the sociological literature through agency theory approaches to state formation (e.g., Coleman (1990); Kiser and Schneider (1994); Kiser and Karceski (2017)). Section 2.1 reviews this literature and presents the conceptual framework in greater detail.

[2]Section 2.2 discusses historians' accounts of the challenges faced by politicians in Washington D.C. in monitoring the performance of federal officers employed throughout the territory, and how loyalty-based hiring and patronage networks were central to the federal government for most of the nineteenth century.

that leverages the expansion of the railroad network as a source of plausibly exogenous variation in monitoring costs between Washington D.C. (henceforth, simply D.C.) and individual locations.

Underpinning this study is a large data collection effort to construct a novel micro-database based on personnel records of federal employees spanning the period from 1817 to 1905. We augment these records with additional hand-collected information that enables the geo-location of each employee and allows to measure several features of the internal organization of the federal government. We start by showing a series of aggregate descriptive facts consistent with a gradual transformation of the U.S. federal government over the nineteenth century. In the first half of the century the federal government employed few officers, had limited territorial reach, and presented several features typical of a patrimonial state. Over the second half of the century, the federal government grew in size and was present in more localities across the U.S. In addition, it started to resemble a more bureaucratic organization, with a more stable workforce and lower reliance on personal connections. While there are certainly several drivers of this historical evolution, we hypothesize that technologies improving monitoring capacity were an important mechanism behind this transformation.

To test our hypothesis, we exploit the growth of the railroad network, which dramatically reduced communication and transportation costs over the course of the nineteenth century. Using data from Donaldson and Hornbeck (2016), we calculate travel times between D.C. and each U.S. county for every decade between 1820 and 1900. We then investigate whether a decrease in travel time between a county and D.C. increases the presence of the federal government in the county. The key empirical challenge in testing our hypothesis is the endogeneity in the expansion of the railroad network across space. Specifically, changes in travel times from D.C. to specific counties are arguably correlated with unobservable shocks (e.g. in economic growth) affecting the federal government's incentives to establish a presence in those counties.

A key contribution of our paper is to develop an empirical strategy to overcome this identification challenge. We do so in three steps. First, we flexibly control for local railroad construction in a county. By doing so, our estimates of the impact of travel distance to D.C. are identified from more-distant changes in the railroad network. Second, since more distant changes in the railroad network are associated with changes in market access, which in turn affected county's agricultural land values (Donaldson and Hornbeck, 2016) and manufacturing activity (Hornbeck and Rotemberg, 2024), we control for a measure of a county's market access. By doing so, we identify our effect of interest by comparing counties that, following an expansion of the railroad network, experience similar railroad construction in their territory and similar changes in market access, but experience different changes in their

travel time to D.C. *specifically*. Third, we address the remaining concern that lower travel time to D.C. per se could increase economic growth given the relevance of this specific city. Although this concern is arguably reduced by the fact that, at this time, D.C. was not an important economic center, we conduct placebo experiments to address this: we show that, conditional on our other controls, lowering travel time between a county and other large cities has no influence on state presence in the county. This lends further support to our identification strategy.

Our estimates show that decreasing travel time between D.C. and a county substantially increased the probability of presence of the federal government in the county.[3] The estimates imply that, had the distance of each county to D.C. not changed during the nineteenth century, the share of counties with state presence would have been lower by 7.3 percentage points in 1860. By 1900, this share would have been lower by 14.5 percentage points, corresponding to a 24 percent decrease relative to the share of counties with state presence observed in 1905.

We then provide indirect evidence in favor of our monitoring mechanism. If increased monitoring capacity is a relevant mechanism, we expect to find evidence for the following predictions: (i) the relationship between travel time and state presence should be particularly strong for occupations and tasks for which agency problems are more severe; (ii) a lower travel time between D.C. and a given county should reduce the patrimonial features of the state organization in the county. We find support for both of these predictions.

We show that the impact of a reduced travel time on state presence is particularly strong for occupations that are expected to be more characterized by agency problems. Additionally, we find that the effects are driven by bureaus of the government in charge of taxation, in line with these tasks being difficult to monitor and especially prone to corruption. We also show that, holding fixed the size of the county's federal workforce, reduced travel time increases the probability of employing officers with managerial tasks in the county, consistent with stricter monitoring facilitating delegation of decision power away from the center of power.

In the last part of the paper, we show that a lower travel time to D.C. contributed to the transition from a patrimonial to a bureaucratic organization of the state. First, we show that faster travel time to D.C. lowers federal employee turnover in the location. This is in line with higher monitoring capacity decreasing the need to replace existing officials with new, trusted ones following political turnover. Second, we show that faster travel time to

[3]We also show that our results are robust to the inclusion of an extensive set of economic, geographic, and demographic controls. Additionally, our estimates are very similar when we focus exclusively on states that were already part of the U.S. at the beginning of our sample period: this suggests that our results are not significantly driven by different dynamics of state development on the frontier (Bazzi et al., 2020), whose westward expansion was facilitated by the development of the railroad network.

D.C. reduces the likelihood that workers in the same local office share the same last name. This is in line with higher monitoring capacity decreasing patronage and personal ties in the organization, which might ameliorate agency problems when monitoring costs are high. Third, leveraging information on workers' states of birth, we document the emergence of a substantial employment gap between Southern-born workers and those from other states following the Civil War. This pattern is consistent with a decline in the federal government's trust in individuals from former Confederate states. Supporting the view that improved monitoring capacity reduces reliance on personal trust, we show that lower travel time to D.C. is associated with a higher share of Southern-born workers employed in a county after the Civil War—a relationship that did not exist prior to the conflict.

Our findings speak to a number of literatures. First, we speak to a growing literature on state formation and the development of state capacity. Recent contributions document the relationship between state capacity and economic development (Besley and Persson, 2011, 2013; Dincecco and Katz, 2016; Besley et al., 2025), as well as the impact of administrative reforms improving state capacity (Chambru et al., 2024; Chiovelli et al., 2023; Cantoni et al., 2024). While most previous studies focus on rulers' incentives to invest in state capacity, we study the problem of concretely organizing the state to perform its functions.[4] Similarly to us, Angelucci et al. (2024) also study the organizational challenges faced by central rulers. They focus on a period of time preceding centralized bureaucractic apparatuses, and examine the communication and coordination challenges when the ruler administer the territory through delegation to local elites. We instead focus on agency problems, and on the role of technologies improving monitoring capacity, as drivers of the growth and organizational change of centralized states.

Second, we speak to a burgeoning literature studying the personnel economics of the public sector (see Finan et al. (2017) and Besley et al. (2022) for recent reviews). An important strand of this literature provides micro-level evidence on how to best select (Dal Bo et al. (2013); Deserranno (2019); Ashraf et al. (2020); Weaver (2021)) and incentivize (Ashraf et al. (2014); Muralidharan and Sundararaman (2011); Duflo et al. (2012); Khan (2023); Bandiera et al. (2021)) bureaucrats to reduce agency problems within these organizations. Our paper underlines how the ability to reduce agency problems is crucial not only to explain the functioning of bureaucracies at a given point in time, but also their growth and organizational evolution.

A related strand of this literature examines the costs (Iyer and Mani (2012); Xu (2018); Colonnelli et al. (2020); Akhtari et al. (2022); Riaño (2021)) and potential benefits (Voth and Xu (2020); Spenkuch et al. (2021)) of political discretion in the selection of bureaucrats,

[4]We review the literature on the role of incentives to invest in state capacity in Section 2.1.

and the effects of transitioning to a merit-based bureaucratic organization (Evans and Rauch (1999, 2000); Folke et al. (2012); Ujhelyi (2014); Ornaghi (2016); Moreira and Pérez (2020, 2022); Aneja and Xu (2023)). A key advantage of our study is the ability to observe the internal organization of a bureaucracy over a long period of time. This allows us to describe how different systems for organizing the state might be optimal depending on the extent of government's monitoring capacity.[5]

Third, our paper is related to the literature on the relationship between delegation of tasks, monitoring ability, and the growth of firms, dating back to Penrose (1959), Chandler (1962), and Lucas Jr (1978).[6] In particular, our findings are in line with the model by Dessein (2002), in which delegation increases as the incentive conflict between the principal and the agent decreases.[7] While we focus on the role of monitoring capacity, cognitive models of organizations without agency problems (Garicano, 2000; Snowberg and Ting, 2019; Gumpert et al., 2022) are also consistent with the relationship between travel time to D.C. and state presence that we estimate: as communication costs decrease, central managers in D.C can better address complex problems that local officials cannot solve. We acknowledge that this mechanism can also lead to the expansion of state capacity as travel time decreases, although some of our additional results are difficult to reconcile with models without agency costs.[8]

Finally, in exploiting the introduction of the railroads as a shock to the government's monitoring capacity, we also contribute to the rich economic history literature on the expansion of the railroads in the U.S. (Fogel (1965); Nerlove (1966); Atack et al. (2010); Atack and Margo (2011); Donaldson and Hornbeck (2016); Hornbeck and Rotemberg (2024)).

The rest of the paper is organized as follows. In section 2, we outline a simple conceptual framework and we discuss the historical context of our study. In section 3, we describe our data collection. In section 4, we present aggregate descriptive facts on the evolution of the U.S. federal government over the nineteenth century. Section 5 presents our identification strategy and our results linking travel time to D.C. to state presence. Section 6 explores the role of monitoring capacity as mechanism. Section 7 concludes.

[5]This is not inconsistent with an increase in efficiency after the introduction of objective and meritocratic selection procedures, since reforms might accelerate the transition to a more efficient system in presence of resistance from entrenched groups.

[6]Jayachandran et al. (2020) and Shahe Emran et al. (2021) argue that many firms, especially in developing countries, have a limited scale because of the high costs of monitoring their workforce.

[7]Our results are also in line with Chen (2017)'s insight that better monitoring capacity increases organizational growth in a general equilibrium monitoring-based hierarchy model. On the empirical side, Kelley et al. (2021) show that technologies that improve owners' monitoring ability lead them to expand the size of their firm, and Giroud (2013) show that decreasing travel time between headquarters and plants increases plant-level investment, by facilitating monitoring and access to information.

[8]For instance, higher monitoring capacity predicts greater decentralization of managerial power in presence of agency costs, while a cognitive model that abstracts from agency problems predicts *less* decentralization as communication becomes cheaper (Bloom et al., 2014).

2 Conceptual Framework and Historical Context

This section outlines the conceptual framework behind the central hypothesis that we test in this paper: that technological innovations reducing transportation and communication costs between a state's center of power and its periphery influence both the capacity and the organizational structure of the state. We then situate this framework in the historical context of the development of the U.S. federal bureaucracy, which offers a unique empirical setting to test this hypothesis.

2.1 Conceptual Framework

What drives the expansion of state organizations? We conceptualize a ruler's decision to invest in state capacity in a particular location as being shaped by two primary considerations: the expected return from state presence, and the associated costs of ensuring the effective operation of the organization in that location.

The return from state presence. Most of the literature on state development has focused on the expected returns from state presence, which create incentives for rulers to invest in state capacity.[9] A prominent theory dating back to Tilly (1975, 1990) argues that incentives to invest in state capacity stem from the fiscal demands of interstate warfare, a claim that has received substantial empirical support (Besley and Persson, 2008; Gennaioli and Voth, 2015; Becker et al., 2022). Other perspectives highlight the state's role in coordinating the provision of public goods (Allen et al., 2023), or the incentives of elites to coerce and extract resources from the population (Carneiro, 1970; Olson, 1993; Scott, 2017).[10]

Once a state chooses to expand, which locations within its territory attract greater investment in capacity? First, we expect higher levels of state capacity in areas where public goods provision yields greater economic returns (Besley and Persson, 2009). In contexts where private provision of public goods is hindered by collective action problems, state intervention becomes essential. For instance, public canals that support irrigation and agriculture can generate high economic returns (Wittfogel, 1976). Bates (1987) similarly suggests that state capacity tends to emerge along ecological boundaries, where the gains from trade are substantial.[11] Second, we expect greater investment in locations with higher returns from taxation. Historically, state presence has been more prevalent in areas with storable economic

[9] Some of the works cited in this section examine the origins of states, seeking to answer the question "Why did states emerge in certain places at particular points in time?" Others focus on the extensive margin, exploring the conditions under which existing states choose to expand their capacity. Both strands emphasize the importance of anticipated benefits as a key driver of state capacity.

[10] See Heldring (2025) for a review of "cooperative" and "extractive" studies on state formation.

[11] Fenske (2014) finds evidence in support of this hypothesis.

surpluses (Mayshar et al., 2022) and in regions with high land productivity surrounded by unproductive areas (Allen, 1997; Mayoral and Olsson, 2019).[12]

The costs of state presence. Realizing the potential returns from state presence requires the ability "to implement logistically political decisions throughout the realm." (Mann (1986), p. 113). While there are many costs associated with this ability, we focus on those associated with agency problems. The employment of officials in locations distant from the center of power implies that the ruler (the principal) cannot easily observe the actions taken by officials (the agents). These officials may shirk responsibilities or engage in corruption, undermining the effectiveness of state investment. This problem is at the center of the difficulties faced by rulers of vast territories in growing their state apparatus.[13]

A long tradition in sociology highlights how a *bureaucratic* organization of the state – characterized by a fixed hierarchy of officials whose recruitment and promotion is based on merit – was for most of history an inefficient solution to these agency problems (Coleman, 1990; Kiser and Schneider, 1994; White, 2004). This form of organization requires high monitoring capacity, which was hindered by historically high communication and transportation costs. In such contexts, it was often more effective for rulers to appoint trusted individuals through personal networks. Personal loyalty reduced the need for formal monitoring, with rulers retaining the option to remove officials when trust broke down. Since these features—appointments grounded in loyalty rather than rules, and arbitrary dismissal—are central to Weber's concept of patrimonialism (Weber, 1978), we define this model of organizing the state as *patrimonial*.[14]

Despite its advantages when formal monitoring is difficult, this model has two important limitations. First, it constrains the expansion of the state apparatus to the availability of loyal and trustworthy individuals. Second, it makes the state apparatus vulnerable to turnover, as new leaders routinely replace officials with their own loyalists, eroding expertise in the organization.

[12] Schonholzer and Francois (2023) argue that these locations are characterized by both higher possibility for extraction and greater need to coordinate public goods provision.

[13] In the words of Finer (1997), a transition from a city-state to a country-state implies that "the central government begins to require outposts for its officials to take care of the localities. [...] The problem was to devise means of securing their continued loyalty and accountability. Both proved very hard to ensure, and in pre-industrial times few large country-states – let alone empires – ever found satisfactory long-term solutions." (Finer (1997), p.155).

[14] Weber defines patrimonialism primarily in terms of personal loyalty to the ruler, as opposed to obedience to abstract norms. Our focus here is on recruitment and dismissal based on loyalty and trust, but Weber's conception of a patrimonial state is broader: as noted by Kiser and Karceski (2017), it includes multiple administrative arrangements, all of which share the characteristic of lower reliance on formal monitoring of officials. For instance, historically rulers used extreme forms of decentralization, e.g. through tax farming, selling the right to collect taxes to local elites. The system guaranteed some revenue but led to excessive extraction and limited central control (Kiser, 1994).

The viability of bureaucratic organization increases as monitoring costs decline. The development of technologies reducing communication and transportation costs significantly lowered the cost of monitoring officials across dispersed territories. As Weber observed, *"a certain degree of development of the means of communication [...] is one of the most important prerequisites for the possibility of bureaucratic administration"* (Weber (1978), p. 973). More recent work in sociology echoes this view (Kiser and Karceski, 2017).[15]

An additional way in which improved communication and transportation infrastructure can lower the costs of ensuring performance away from the center of power is by facilitating information flows between the central managers and local officials. This mechanism is central to cognitive models of organizations (Garicano, 2000; Snowberg and Ting, 2019), where lower communication costs allow central managers to better address complex problems arising in the local offices.[16]

The ruler's decision. In deciding whether to invest in state capacity in a particular location, a ruler weighs the anticipated return against the associated costs. On one hand, investment is more likely in locations where taxable economic surplus or returns from public goods provision are high.[17] On the other hand, investment is only worthwhile if these returns are higher than the cost of ensuring effective implementation by officials. If transportation and communication costs between the center and a given location are high, the cost of monitoring officials is more likely to exceed the return from state presence.

In such cases, rulers face a choice: either forgo investment in the location, or rely on a patrimonial organizational form. While officials in both bureaucratic and in patrimonial organizations receive wages, personal loyalty decreases monitoring costs in the latter organizational form.[18] However, this option is limited by the available supply of trusted individuals. Investment might not take place in some locations with high returns and high monitoring costs, due to the lack of available trusted officials to staff the location. In sum, the ruler faces three options for each location: no state presence, a patrimonial state organization, or a bureaucratic state organization.

[15]Officials in turn face costs related to the need of obtaining information about the population and the state of the economy (Scott, 1998; Lee and Zhang, 2017; Mayshar et al., 2017; Garfias and Sellars, 2021; Sanchez de la Sierra, 2020). A number of studies have linked the development of modern tax systems to the availability of hard (Dzansi et al., 2025) or soft (Jensen, 2022) information about taxpayers.

[16]Bloom et al. (2012) present an extension of a cognitive model of the firm, in which principals and employees have misaligned incentives. The presence of trust between principals and employees make the former more willing to delegate to the latter, allowing the organization to grow.

[17]For instance, state capacity investments along trade routes can yield significant tax revenues, while defense expenditures are particularly valuable in border regions where security concerns are more acute.

[18]Patrimonial arrangements have the additional cost of being vulnerable to turnover. However, this is a fixed cost that does not depend on transportation and communication costs between the center and the location.

Technological innovations, monitoring costs, and state capacity. Technological innovations that improve communication and transportation between the center and a given location will decrease the cost of monitoring the location. If monitoring was not initially cost effective, the return might now outweigh the cost of monitoring. Therefore, locations that initially lacked state presence, because monitoring costs were higher than the return and because there was no available trusted official to staff the location, might now see investments in state capacity. This leads to the central hypothesis that we test in the paper.

Hypothesis 1: a reduction in communication and transportation costs between the center and a location should increase the probability of state investment in that location.

We provide evidence in support of this hypothesis in Section 5. This hypothesis is also consistent with a decrease in the cost of solving problems arising in the location following lower communication costs, in line with cognitive models of organizations without agency problems. We show evidence in support of monitoring as an important mechanism by providing tests of two additional hypotheses.[19] A prediction that follows from our conceptual framework is that, once monitoring costs decrease, a monitoring-based bureaucratic form might become less costly than a patrimonial organizational form. In addition, improvements in monitoring capacity will be particularly crucial for functions and tasks characterized by high agency problems. These leads to the following two additional hypotheses, that we test in Section 6.

Hypothesis 2: if monitoring is an important mechanism, the state should be organized in a less patrimonial way in locations with lower communication and transportation costs with the center: officials should be more likely to have stable positions and less likely to be embedded in personalistic, trust-based relationships with their superiors.

Hypothesis 3: if monitoring is an important mechanism, the impact of lower communication and transportation costs on state presence should be particularly strong for functions and tasks characterized by high agency problems.

Identification challenge. The aim of this paper is to empirically test the hypotheses outlined above. A central challenge in doing so is identifying the causal impact of technological innovations: advances in communication and transportation not only reduce the costs of state presence, but also tend to increase its returns. For example, improved transportation infrastructure can enhance a location's trade potential, thereby increasing its taxable surplus. Because expected returns are not directly observable, and thus cannot be perfectly controlled for, isolating the effect of technological innovations on the cost side of state capacity poses

[19] Importantly, our goal is to show evidence in support of monitoring as an important mechanism. We do not claim that better problem-solving following a decrease in communication costs is not partially responsible for the higher likelihood of investment in state capacity.

a significant empirical challenge.

The central contribution of this paper is to leverage a historical setting and an identification strategy that allows us to isolate the effect of these technological innovations on the *costs* side of the ruler's decision, while controlling for their potential impact on the ruler's anticipated *returns* from investing in state capacity.

2.2 Historical Context

In this section, we describe the historical setting of the paper: the development of the U.S. federal bureaucracy during the nineteenth century. In our setting, the "rulers" are the politicians in D.C., and the "officials" are the federal bureaucrats employed throughout the territory. First, we discuss the challenges in monitoring the performance of federal bureaucrats, and how a patrimonial organization of the state apparatus was used to ameliorate agency problems.[20] We then describe the expansion of the railroad network during the nineteenth century.

Delegation and monitoring problems in the U.S. federal bureaucracy. The expansion of the federal bureaucracy throughout the nineteenth century necessitated the increasing delegation of responsibilities from department heads to subordinates. As early as 1778, Treasury Secretary Alexander Hamilton recognized the need for delegation, writing that *"it is essential to the success of the minister of a great department that he subdivide the objects of his care, distribute them among competent assistants, and content himself with a general but vigilant superintendence."* (Hamilton, 1795, p. 484)

While delegation within D.C. permitted relatively close supervision, monitoring the activities of field officials operating far from the center was substantially more difficult. This was true both for workers involved in simple tasks, and for agents that were delegated significant decision powers. For example, administrators in the Land Office, who operated on the frontier, were *"relatively secure from the prying eyes of Washington bureaucrats,"* resulting in frequent episodes of fraud, corruption, and neglect of duty (Crenson, 1975, pp. 86-87). Similar cases of corruption and lack of effort in the performance of duties were common among employees in the custom houses (Prince and Keller, 1989).[21] Custom collectors, for instance, sometimes undervalued imports to attract shipping traffic to their ports, thereby increasing their own fees. (White, 1954, p. 179).

[20] A quantitative description of the development of the U.S. federal bureaucracy is presented in Section 4, where we provide a series of novel descriptive facts based on our newly collected data.

[21] One of the most famous cases of corruption involved the collector of the New York custom office between 1829 and 1838. Through embezzlement and bribes, it is estimated that he appropriated more than $1 million (Prince and Keller (1989), p. 102).

In the early decades of the nineteenth century, the high cost of communication and transportation severely limited the center's ability to supervise remote agents. Since the 1840s, special agents were employed by the Secretary of the Treasury to monitor the operations of officials at the custom houses.[22] In 1851, the Treasury Department attempted to strengthen oversight by appointing eight district appraisers tasked with inspecting customs offices and reporting to D.C. These appraisers were authorized to investigate misconduct and intervene directly when necessary (U.S. Customs Service, 1988, pp. 43-47). Similarly, the Land Office instituted periodic inspections of local branches. However, the large distances between D.C. and the various field offices made these tools insufficient to ensure adequate monitoring. For example, Land Office inspectors typically visited each office only once a year, providing ample opportunity for malfeasance to go undetected (Crenson, 1975, pp. 92)).

Patrimonial organization as a response to high monitoring costs. Given the limited feasibility of formal monitoring, department heads in D.C. relied heavily on the personal character and political loyalty of field officials to ensure compliance with bureaucratic responsibilities. Political leaders routinely emphasized moral rectitude and allegiance to the ruling party as primary criteria for selection (Fish, 1905).[23] As the First Comptroller noted in 1837, *"the only safeguard for the public security against fraud and embezzlement upon which entire reliance can be placed is to be found in the heart and conscience of the individual intrusted with the receipt and disbursement of the public funds"* (Senate Doc. 1 25th Congress, 1837).

Personal networks played a crucial role in identifying trustworthy candidates.[24] For example, Thomas Jefferson defended one of his appointments by citing "private sources" from which he learned that *"his understanding was sound, his integrity pure, his character unstained."* (Jefferson, 1854). Similarly, Secretary of State Daniel Webster in 1851 asked a correspondent for *"the name of a man, the fittest, within your knowledge, to be Naval Offi-*

[22]Secretary Guthrie praised the system of inspection in his 1854 Annual Report, writing: *"The department has caused the collector's offices in all the ports to be examined within the year, with but few exceptions, by agents of the department, in order to ascertain how their books and accounts were kept, and, by personal inspection, how the official corps discharged their duties. These examinations have enabled this department to correct errors and omissions, and to see that the official corps devote themselves, in person, to the duties confided to them [...]"* (Report of the Secretary of the Treasury on the State of the Finances for the Year Ending June 30, 1854).

[23]Writing about his goals in selecting federal bureaucrats, President Washington noted: *"[I have tried] as far as my own knowledge extended, or information could be obtained, to make fitness of character my primary object"* (Washington, 1855, pp. 57).

[24]In a letter written in 1801, Thomas Jefferson remarked that *"Of the various executive duties, no one excites more anxious concern than that of placing the interests of our fellow citizens in the hands of honest men, with understandings sufficient for their stations. No duty, at the same time, is more difficult to fulfill. The knowledge of the characters possessed by a single individual is, of necessity, limited"* (Jefferson, 1854, pp. 402).

cer. He must be a firm and energetic friend to the present Administration; not too old, all together trustworthy and enjoying public confidence" (Webster, 1904). Members of Congress were also involved in the appointment process, regularly recommending constituents deemed loyal and competent (White, 1954, p. 116). In turn, within field offices, supervisory relationships were often characterized by strong interpersonal ties: subordinates were linked to their chief *"by personal loyalty, friendship, and, not infrequently, kinship."* (Crenson, 1975, p. 72). Upon assuming his post as Surveyor for the Salem and Beverly Customs District in 1846, Nathaniel Hawthorne petitioned the Treasury Secretary to replace two *"incompetent"* inspectors, and suggesting the names of two *"capable and efficient men"* and *"firm friends of the administration"* (Woodson et al., 1985).[25]

In sum, for much of the nineteenth century, the U.S. federal bureaucracy exhibited key features of the patrimonial organizational form outlined in our conceptual framework: appointments were frequently based on loyalty and the personal trust leaders placed in subordinates, and leaders retained the discretion to dismiss officials. While we use the term "patrimonial" to describe these organizational features of the state, other terms—such as "personal organization" (Crenson, 1975) or "patronage system" (Fish, 1905)—are also commonly used to characterize the nineteenth-century U.S. federal government.

Expansion of the railroad network. The dramatic expansion of the railroad network across the U.S. during the nineteenth century revolutionized transportation. Originating in the 1820s, railroads rapidly superseded canals and other waterways, becoming the dominant and most efficient means of moving people and goods across the country. By 1850, railroad mileage more than doubled that of canals, and by 1860, it exceeded the rest of the world's combined total. (Atack et al., 2010). The resulting decreased transportation costs and increased speed were critical drivers of economic growth (Fogel, 1965; Donaldson and Hornbeck, 2016; Hornbeck and Rotemberg, 2024).

In this paper, we leverage the evolving railroad network as a source of variation in the ability of D.C. to visit and communicate with locations across the country. To illustrate how this development materially affected monitoring capacity, we examine its impact on the district appraiser system established in 1851 to inspect custom offices. Our analysis, utilizing 1850 railroad network data, indicates that the median district required approximately 148

[25]While we emphasize the possible benefits of political discretion on appointment and removals, which allowed to establish trust relationships in presence of low monitoring capacity, staffing of the federal bureaucracy also responded to electoral considerations. Federal employment was an important tool to build support for the party, and political support could substitute for qualifications as a determinant of appointment (Fish, 1905; Hoogenboom, 1968). Meritocratic hiring in the federal bureaucracy was introduced *de jure* in 1883 by the Pendleton Act. Initially, only a small share of positions were affected by the reform, with most of the positions transitioning to meritocracy only in subsequent decades. Our results will show that turnover and trust-based hiring were *de facto* more limited in locations that could be more easily monitored by D.C.

hours of uninterrupted travel for appraisers to visit each office and report back to D.C.[26] For the two most remote districts, this travel time extended to 440 hours. A counterfactual simulation suggests that, had the 1890 railroad network been in place, this would have reduced these travel times by approximately 50 percent.[27]

3 Data Collection

Our study relies on a novel micro-database combining federal employees' personnel records and hand-collected information on the internal organization of the U.S. federal government between 1817 and 1905. We additionally use data on the development of the railroad network to calculate travel times across different locations. In this section, we describe our data collection effort and the sources of the data.[28]

3.1 Personnel records from the U.S. Official Registers

Personnel records of the U.S. federal bureaucracy come from the Official Registers of the United States (Registers henceforth). The Registers were compiled and published biennially, in every odd year from 1817 until 1959.[29] We digitized all issues of the Registers between 1817 and 1905, for a total of 15,801 pages.

We focus on civilian employees of the executive branch of government. That is, we drop the names of members of the army, of the judiciary, and of offices that were under the direct control of Congress (e.g., the government printing office, or the library of Congress). Importantly, we have digitized information for employees working in all executive departments except the Postal Office. Our choice is motivated by the size of this department, which would have significantly increased our data collection effort, and by the more limited information on these employees.[30] Finally, we drop employees in navy yards and in the engineer department

[26] We defer a detailed description of the data on the railroad network, and how we calculate travel times across different locations to section 3.2.

[27] We present details on how we performed these calculations in Appendix section E.

[28] In our analysis, we also use a variety of demographic, geographic, and historical controls. We discuss them as they are introduced. We provide the exact definition and the source for each variable in Online Appendix C. Full details on the data collection from the Official Registers can be found in a supplementary appendix on the authors' website.

[29] The Registers were initially compiled and published under the direction of the Secretary of the Department of State, and since 1861 of the Secretary of the Department of Interior. The head of each department was responsible to compile the list for his own department. The first book, for 1817, is 33 pages long and it contains 1056 employees. The last book of our sample period, for 1905, is 1254 pages long and it contains more than 120,000 employees. Online Appendix Figure A1 shows the cover page of the 1817's Register, and the first page of the Treasury Department in the 1875's Register.

[30] Employees in the Postal Office span 97 pages in the 1817 Register, and 1922 pages in the 1905 Register. The Registers usually exclude information on place of birth and appointment of postal office employees, and

at large. We impose this data restriction since employees rosters from these offices seem to be missing from the Registers before 1881 and between 1845 and 1879, respectively.[31] Our final panel includes a total of 304,441 unique employees, and 810,958 employee-year observations.

This data source allows us to observe a rich set of characteristics of all the individuals employed by the federal government. For each employee, the Register reports their full name, state (or foreign country) of birth, and state of appointment (i.e., of residence at the time of appointment). It also provides detailed information on the job that each employee performs in the bureaucracy: we observe information on employees' occupation, location of employment, and compensation. In addition, the layout of the Registers allows us to observe the hierarchical division of this organization into departments, offices, and divisions, and to assign each employee to the specific organizational unit in which they are employed. We link employees over time, in order to track their careers in the federal bureaucracy.[32]

Geo-location of places of employment. The Register contains information on each worker's location of employment.[33] We harmonize the names of the locations across years, and we manually collect information on the geographic coordinates of each location. This allows us to assign each location to its county and state. Since county boundaries change over time, we maintain consistent geographic units over time by holding constant county boundaries in 1890 throughout our sample period. We follow the procedure in Hornbeck (2010) and we harmonize all the county-level covariates used in the analysis to reflect 1890 county boundaries.[34]

Of the 810,958 observations in our dataset, 800,432 (or 98.7%) have non missing information on the location of employment. Of these, 32,497 (or 4%) correspond to workers employed in a foreign country. Of the remaining 767,935 observations that are located within the United States, we can recover information on the county of employment for 95% of observations.[35] In total, the data include 9,666 unique geo-located places of employment.

Construction of the hierarchical structure. From 1817 to 1905, the Official Register was arranged in a tabular format. This layout provides us with a picture of the organizational structure of the federal bureaucracy at each point in time. Relying on a series of publications on the history of the U.S. federal government, we construct a consistent hierarchy of the

often report only the initials of the first names.

[31] None of the central results of the paper is affected by this choice.

[32] We match employees using several steps of matching, based on their full name, place of birth, state of residence at time of appointment, gender, and department of employment.

[33] Online Appendix Figure A2 shows an extract from the 1875 Register, highlighting the locations under the "where employed" column.

[34] This procedure uses area-based weights to harmonize county boundaries across years.

[35] For the remaining 5% of observations, either the Register reports only the State of employment, or it reports vague geographic information (such as "on a river" or "along the coast"), which prevents us from assigning precise coordinates.

organization by following the evolution of its units over time.[36] This step is crucial, since units were often added, deleted, or transferred within the organization, or experienced changes in their name.

We identify, and divide the organization into, four hierarchical layers. The first layer is composed of the departments (e.g., Treasury, War, Navy, Interior). The second layer is composed of the bureaus within each department. The third layer is composed of the divisions within each office. The fourth layer is composed of the different local offices within each division.[37]

The reconstruction of this hierarchy allows us to assign each worker to an organizational unit, which we identify as a local office of a specific bureau. We can also recover the chain of command in the organization, assigning all workers to their direct supervisor.

Categorization of job positions. The Registers contain information on the specific occupation of each employee. After standardizing the names of the job titles in the data, we obtain a total of 11,923 unique occupation codes. We group occupations into five categories based on the type of task performed. The first category includes the top managers of the organization: the heads of department, deputy heads of department, and heads of bureaus. The second category includes workers employed in a managerial capacity (for example: chief of divisions, chief clerks, chief of local offices). The third category includes clerical occupations (for example: clerk, copyist). The fourth category includes non-clerical occupations requiring high skills (for example: engineer, doctor). Finally, the fifth category includes non-clerical occupations requiring a relatively low level of skills (for example: laborer, messenger). Of the 11,923 occupation codes in the data, 2.1% are categorized as top managers, 11.8% as managers, 26.6% as clerical workers, 34.6% as high skills non-clerical workers, and 25% as low skills non-clerical workers.

[36] Specifically, we mostly relied on *"The Development of National Administrative Organization in the United States"* (Short, 1923); *"The Executive Departments of The United States at Washington"* (Elmes, 1879); *"The United States Government: Its Organization and Practical Workings"* (Lamphere, 1881)

[37] Some examples of bureaus within the Treasury department are the Office of the Secretary, the First Comptroller Office, and the Customs Office; some examples within the Interior Department are the General Land Office and the Indian Office. We use the generic term "division" to refer to the different sub-units in which some bureaus are further divided. For example, the Customs Office is composed of several customs districts; the General Land Office is composed of several surveyor districts. At the bottom of the hierarchy there are local offices. For example, the Providence customs district in 1853 has three local offices (Providence, Pawtucket, and East Greenwich). Online Appendix Figure A3 provides a partial graphical representation of the hierarchy in 1853.

3.2 Railroad network data

In order to measure how the expansion of the railroad network decreased the travel time between D.C. and different counties, our starting point is the transportation network database by Donaldson and Hornbeck (2016), based on initial GIS railroad files by Atack (2013). The database contains both the location of the time-varying railroad network in each decade from 1830 to 1900, and the time-invariant locations of canals, navigable rivers, and other natural waterways. The database is then overlaid to a map of 1890 county boundaries. Online Appendix Figure A4 shows the expansion of the railroad network over time.

Following Donaldson and Hornbeck (2016), we calculate the shortest path between D.C. and the centroid of each county. While Donaldson and Hornbeck (2016) are interested in the lowest-cost freight routes and thus need to specify transportation cost parameters, we need to specify travel time parameters. Travel times (measured in minutes) are calculated using a combination of stagecoach, waterway, and railroad transportation modes. We set time-varying speed estimates based on historical sources documenting nineteenth-century transportation conditions. Specifically, we assume railway speeds increase from 13 miles per hour (mph) in 1830 to 37 mph by 1900, and navigable waterway speeds from 4 mph in 1830 to 17 mph in 1900. Stagecoach speeds are held constant at 7.5 mph throughout the period.[38] Since travel conditions and infrastructure varied substantially across regions, there are inevitable simplifications embedded in these parameters. Importantly, we show that our results are robust to using alternative specifications with fixed (time-invariant) speeds over the entire period. The resulting measure, $Log\ Time\ to\ DC_{ct}$, is the log travel time (in minutes) in year t, between D.C. and the centroid of county c.[39]

Using the same procedure, we calculate the travel time distance between each pair of counties in each decade. As we explain in section 5.2, we use these measures to control for travel time between a county and other important economic centers, and to compute an alternative measure of market access based on travel time.

4 Descriptive Facts on the Development of the U.S. Federal Government

In this section, we show a series of novel descriptive facts on the growth and organizational development of the U.S. federal government over the nineteenth century. First, we document

[38] See Online Appendix, section D, for details on our choice of these parameters.

[39] Given the absence of railroads in 1820, we calculate the travel times between D.C. and each county in 1820 assuming that travel could take place through either stagecoach or navigation.

its growth in size and territorial reach. Second, we document how the presence of the federal government is related to local economic growth. Third, we document the evolution in the way in which the federal government was organized. The picture that emerges is one of a transition from a small, patrimonial organization, to a larger, more bureaucratic organization.

4.1 Evolution of state presence

Growth in size. In Panel A of Figure 1 we plot the total number of federal employees in each year between 1817 and 1905. The figure shows that the U.S. federal government experienced limited growth prior to the 1860s, followed by a marked expansion afterward.[40] In Panel B, we show that the growth of the bureaucracy spiked in D.C. in the years of the Civil War, while the increase outside the capital in the second half of the century was more gradual, both in the South and outside the South.[41]

Growth in geographic reach. A closely analogous trajectory characterizes the federal government's geographic reach across the territory. As shown in Figure 1, Panel C, the share of counties with a federal presence (i.e. with at least one individual employed within the county borders) remained roughly constant at around 15 percent between 1817 and 1859. In the second half of the nineteenth century, the state begins to increase its presence across the territory.[42] Online Appendix Figure A7 shows the presence of the state across space at four points in time.[43]

Federal, state, and local governments. In Online Appendix Table A1, we explore the association between the presence of the federal government and that of lower levels

[40]There are 917 employees in 1817, 5,856 in 1859, and 79,835 in 1905. This significant growth in the second part of the nineteenth century was not unique to the United States. In France, estimates from Turquan (1899) suggest that the number of individuals employed by the central state rose from approximately 200,000 in 1846 to 410,000 by 1886, a doubling over forty years. Ruiz (2013)'s more detailed analysis revises these figures but confirms a broadly similar growth trajectory. In Italy, official figures reported by Cassese (1982) show an increase from 50,000 in 1861 to 376,777 in 1910—a sevenfold expansion. More generally, while cross-nationally comparable counts of bureaucrats are rare, Armstrong (2015) emphasizes that state-building and bureaucratic growth were widespread in nineteenth century Europe.

[41]We normalize the three series in Panel B by the population in D.C., in Southern states and outside the South, respectively.

[42]For each year, the number of counties with potential state presence (i.e., the denominator of this share) is the number of counties (following the 1890 county borders) that were included in the most recent census. In Online Appendix Figures A5 and A6 we show that we see similar trends if we limit the sample only to counties in states that were already part of the U.S. in 1817, or if we weight each county by the fraction of the U.S. population living in the county in a specific year.

[43]In Appendix B, we implement a growth decomposition exercise showing how the (modest) growth of the U.S. federal government between 1817 and 1859 was entirely driven by a higher number of functions performed by the state and by an increase of workers per office. In the post-1859 period, the geographic expansion of the state was also an important driver of growth in the size of the U.S. federal government.

of government in a specific county and year, conditional on county fixed effects and year fixed effects. Interestingly, the results suggest that the presence of federal bureaucrats in a location does not substitute for state capacity by other levels of government. If anything, the presence of federal bureaucrats is positively associated with the presence of employees of local and state governments in a county.[44] This is in line with the largely different activities performed by different levels of government during the nineteenth century (Wallis, 2000).

Growth by department. Online Appendix Figure A8 breaks down the growth of the state by department. Throughout the entire 1817-1905 period, the Treasury was the largest department, consistent with the relevance of its primary tasks – raising revenues and supervising their expenditure by other departments. Until the 1880s, the only other sizable departments were War, Navy, and Interior. By the 1880s, a large number of additional, smaller departments started to employ a large number of employees. Online Appendix Table A2 shows the main tasks of the federal government outside D.C., listing the bureaus with the largest overall number of employees between 1817 and 1905. The two largest bureaus – customs and internal revenue – were responsible for the collection of the two main sources of federal revenue in the nineteenth century, namely custom duties on imports and excise taxes on goods such as tobacco and liquor.

Growth by occupation. Figure 2, Panel D, shows the occupational breakdown of federal employees from 1817 to 1905. Managerial positions (left axis) remained flat until the 1850s, with growth accelerating only after 1860. In contrast, the number of clerical workers and of high- and low-skilled non-clerical workers (right axis) increased slowly but steadily between 1817 and the end of the 1850s. By 1905, the bureaucracy had a pyramidal structure: a broad base of low-skill workers, a middle tier of higher-skills non-clerical and clerical staff, and a smaller managerial layer.

4.2 Local economic development and state presence

We investigate whether the federal government was more likely to be present in areas characterized by higher economic growth. We create a country-year panel and we proxy for economic development with the following variables: the logarithm of a county's population, its urbanization share, the share of its population employed in manufacturing and an indicator to capture the presence of a railroad in a county. Our outcome is state presence: a binary indicator for federal government presence in a county. Table 1 reports the results

[44] We observe the number of individuals employed as local or state government employees in each county in the full-count censuses between 1850 and 1900. Unfortunately, data on individuals' occupation is not available before the 1850 census.

of this analysis.[45] The central takeaway is that, across all specifications, economic development is consistently and significantly associated with greater federal government presence. These results are consistent with theories linking the expansion of the state to both greater opportunities to tax the economic surplus and to higher citizens' demand for public goods.

4.3 Evolution of the State organization

Employee Turnover. Our data allow us to provide the first full quantification of this phenomenon throughout the nineteenth century and for the entire U.S. federal bureaucracy. We compute the share of employees who leave the organization in each year t from 1819 to 1905, defined as the share of employees who were present in the Official Register in year $t-2$ but not anymore in year t.[46] Figure 2, panel A, plots the evolution of turnover rates over the nineteenth century, together with a local polynomial fit with 95 percent confidence bands. The red vertical lines indicate years with a change in the party controlling the federal government. Two patterns emerge from the data. First, turnover exhibits large spikes in the years of a presidential transition. Second, the rate of turnover steadily increases until the end of the 1850s, and is on a declining trend thereafter.[47]

In Online Appendix Figure A9 we separately plot turnover rates in D.C. and outside of D.C. (i.e."in the field"). Turnover rates are consistently lower in D.C. than in the field.[48]

Link between employees' and supervisors' careers. The second organizational dimension that we analyze is the link between an employee's career and that of her supervisor. Specifically, we assign employees in each year and local office to their direct supervisor (or supervisors), and we ask whether the turnover of a supervisor leads also the direct subordi-

[45]Importantly, all specifications include county fixed effects, year fixed effects and a second-order polynomial in the straight line distance between the country and D.C. interacted with year fixed effects. We exclude D.C. from the sample, given that we are interested in the presence of the federal government outside of its center of power. We also drop the two administrative divisions of the Alaska Territory (the Northern and the Southern Districts), which account for 26 county-year observations, and have zero employees throughout the sample period. Including the Alaska Territory leaves the results virtually identical.

[46]Since the Register does not list the reason for an employee's exit, we do not know whether departing employees were fired, resigned, or died. While we would ideally only focus on exits because of firing or resignation, it is important to note that U.S. life expectancy at age twenty did not significantly increase over the nineteenth century (Hacker, 2010). Thus, the rate of employees' exit because of death can be assumed roughly constant over our sample period.

[47]Specifically, during the 1861 transition 72 percent of employees left the organization, up from 60-63 percent during the 1849 and 1853 transition and from 52-53 percent during the 1841 and 1845 transitions; the turnover rate dropped to 55 percent during the 1869 transition, to 44-48 percent during the 1885, 1889, and 1893 transitions, and to 35 percent during the 1897 transition.

[48]This is not due to the different nature of jobs and bureaus between D.C. and the field: when we regress an indicator equal to one if the employee leaves the organization on an indicator for D.C., including a set of year-bureau-position type fixed effects, being employed in D.C. is associated with a 42 percent reduction in turnover probability (Online Appendix Table A3).

nates to leave.[49] Figure 2, panel B, presents the standardized effects, namely the coefficient β normalized by the mean sample probability that an employee leaves when none of her supervisors do. Before 1841, moving from none to all supervisors leaving the organizational unit increases turnover probability among subordinates by 37 percent. This effect is similar between 1841 and 1859. In the subsequent twenty years period, the effect drops substantially, to 22 percent, and remains roughly constant after 1881. In summary, there exists a tight link between supervisors' career and the career of their subordinates, but this link is significantly less pronounced in the second part of the century.

Personal connections among co-workers. Finally, we examine how the extent of personal connections within the federal bureaucracy evolved over the nineteenth century. We proxy personal connections using information on the last name of workers. Specifically, for each worker-year observation in our data, we construct an indicator equal to one if the worker shares the last name with another worker in the same local office. Figure 2, panel C, shows that the probability that an employee shares the same last name with a co-worker decreased over time:[50] relative to the period 1817-1829, it was about 30 percent lower in the 1830s and 1840s, between 40 and 50 percent lower between the 1850s and the 1870s, and about 60 percent lower from the 1880s onwards. This trend suggests that personal connections among workers in the federal bureaucracy became less and less prevalent over the nineteenth century.

Taken together, these findings indicate that the internal organization of the federal bureaucracy changed meaningfully during the nineteenth century. Both turnover among employees and the link between the careers of workers and their supervisors became less significant during the second part of the century. In addition, the data suggest that personal connections among co-workers became less common over time. These descriptive patterns are consis-

[49]Specifically, we regress an indicator for worker turnover on the share of their supervisors who leave, including year, bureau, and location fixed effects.

$$\text{Turnover}_{it} = \alpha_t + \gamma_{b(it)} + \delta_{l(it)} + \sum_{\tau} \beta_\tau \text{Share Supervisor Turnover}_{it} + \epsilon_{it} \quad (1)$$

where the variable Turnover_{it} is an indicator equal to one if worker i leaves the local office in year t and $\text{Share Supervisor Turnover}_{it}$ is the share of i's supervisors who leave in year t (an employee has a median of 3 supervisors). We allow the relationship between $\text{Share Supervisor Turnover}_{it}$ and Turnover_{it} to vary over time, estimating its effect for four periods of roughly the same length: before 1841, between 1841 and 1859, between 1861 and 1881, and after 1881. We cluster standard errors by local office times year.

[50]Specifically, in a sample of workers not located in D.C. and employed in a local office with more than one worker, we regress the indicator equal to one if the worker shares the same last name with a co-worker on a set of indicators for specific decades. We include bureau fixed effects, county fixed effects, fixed effects for the number of workers in the office (to account for the mechanical correlation between office size and the probability of having workers sharing a last name) and last name fixed effects (to account for the fact that some last names are more common. We cluster standard errors by local office times year.

tent with a gradual transition of the federal government from a patrimonial to a modern bureaucratic organization.

5 Railroad Expansion, Travel time to D.C., and State Presence

Our goal is to estimate how the decrease in travel time between a county and D.C. resulting from the expansion of the railroad network affected the presence of the federal government. We first present our estimating equation. We then discuss how our identification strategy allows to isolate the cost-side of the decision to invest in state capacity, before presenting the results.

5.1 Estimating equation

We estimate the following regression model on a county-year panel between 1821 and 1905:[51]

$$y_{ct} = \alpha_c + \gamma_t + \beta Log\ Time\ to\ DC_{ct} + \delta_t f(Distance_c) + X_{ct}\theta + \epsilon_{it} \quad (2)$$

where y_{ct} is one of our outcomes of interest for county c and year t. $Log\ Time\ to\ DC_{ct}$ is the log travel time (measured in minutes) in year t, between D.C. and the centroid of county c.[52] We include a set of county fixed effects (α_c), which absorb time-invariant county-specific unobservables which affect the development of the state. Year fixed effects (γ_t) absorb aggregate time-varying shocks in state development (e.g., the increased need for tax revenue during the Civil War). We also control for a second-order polynomial in the straight line distance between county c and D.C., interacted with year fixed effects, allowing for differential changes over time in the outcome variables in counties with different geographic distances from D.C. The matrix X_{ct} includes a set of controls that we discuss in the next section. We report standard errors clustered by county.

The coefficient β estimates the impact of travel time to D.C. on outcome y_{ct}, comparing changes in counties that experience a decrease in travel time to changes in other counties, and accounting for changes flexibly associated with geographic distance to D.C. Panel (a)

[51] Our sample in each decade includes all counties whose territory was covered by the most recent U.S. census, with the exception of D.C. and Alaska. We drop Confederate states between 1861 and 1865, since the federal government did not control the territory during the Civil War. The final sample includes a total of 1,554 counties in 1821 and a total of 2,796 counties in 1905.

[52] Each county-year ct is assigned the value of $Log\ Time\ to\ DC_{ct}$ at the beginning of t's decade. Results in which the sample is restricted to one year of each decade give similar results (see Online Appendix Table A6).

of Figure 3 shows the estimated average decline in travel time between D.C. and counties throughout the U.S. between 1830 and 1900. Relative to 1820, travel time was on average 17 percent lower by 1850, 60 percent lower by 1860, and 120 percent lower by the end of our sample period.

5.2 Identification

The key threat to identification is that counties that, following the expansion of the railroads network, experienced larger reductions in travel time to D.C. could have experienced a change in investment in state capacity for reasons unrelated to lower communication and transportation costs between D.C. and the county. In particular, identification is challenging since the development of the railroads network affected not only travel time between county c and D.C. but also, through its economic and demographic impacts, the return from investing in state capacity in county c. We address this issue in three steps.

First, we control for railroads construction in county c in year t. This is crucial, since railroad promoters and investors sought locations with high profitability, and were more likely to direct construction toward counties with higher growth in population density and agricultural productivity (Atack et al., 2010). In addition, new railroad construction might have increased local manufacturing activity through higher demand for construction materials (Fishlow, 1965). This implies that railroads construction in a county is not only associated with a reduction in travel time to D.C., but is also likely correlated with changes in the federal government's incentives to invest in state capacity in that county. To address these concerns, X_{ct} includes an indicator taking value one if the county contains any railroad track, and a cubic polynomial function of the railroad track mileage in the county. In doing so, we exploit the fact that changes in travel time between county c and D.C. are driven by two sources of variation: (i) railroad construction in county c, and (ii) changes in more distant portions of the railroad network and how the railroad network interacts with other components of the transportation network (Donaldson and Hornbeck, 2016). After the inclusion of these controls, β is identified solely from the second source of variation.

Second, we control for a measure of market access as in Hornbeck and Rotemberg (2024). More distant changes in the railroad network which reduced travel time between county c and D.C. are also associated with an increase in county c's ability to trade with the rest of the country. Since this in turn led to an increase in county agricultural land values (Donaldson and Hornbeck, 2016) and manufacturing activity (Hornbeck and Rotemberg, 2024), this might once again create a spurious correlation between $Log\ Time\ to\ DC_{ct}$ and

state presence.[53] We exploit the fact that expansions of the network which create similar changes in county c's market access (and similar changes in local railroad construction in county c) do not necessarily result in equal changes in travel time between county c and D.C. To illustrate this point, consider the example of the way in which the expansion of the railroad network between 1840 and 1860 affected Adams county and Medina county, both in Ohio. In this period, no railroad track was built within the area of these two counties, and more distant changes in the network affected their market access in the same way.[54] Nevertheless, travel time to D.C. changed in a different way for the two counties: between 1840 and 1850, travel time to D.C. decreased by 20 percent in Medina county and by only 4 percent in Adams county; between 1850 and 1860, the decrease was instead significantly larger in Adams county (48 percent versus 33 percent). The bottom panel of Figure 3 graphically illustrates the intuition behind our identification strategy. We plot the relationship between a county's change in $Log\ Time\ to\ DC_{ct}$ between 1870 and 1880 and the county's change in log market access over the same period, after partialing out changes in the controls for local railroad construction and a second order polynomial in the straight line distance between the county and D.C. While we observe a significant negative relationship between these changes, the correlation is far from perfect: counties with similar increases in market access can exhibit markedly different reductions in travel time to D.C. Our ability to identify the causal effect of reduced travel time between a county and D.C. relies on this residual variation, for every decade between 1820 and 1900.[55]

We take a third step to further strengthen the credibility of our identification strategy. Our identification relies on the assumption that, conditional on changes in local railroad construction and in market access, changes in travel time to D.C. *specifically* do not affect the returns from investment in state capacity in a county. Essentially, we are assuming that a better connection to D.C. is important only because D.C. is the headquarter of the federal

[53]Market access captures how easily county c can trade with all other U.S. counties, assigning higher weights to counties with greater population. Formally, we control for log market access, with market access of county c at time t defined as $MA_{ct} = \sum_{d \neq c} (\tau_{cdt})^{-\theta} L_{dt}$, where τ_{cdt} is the cost of transporting goods between county c and county d, θ is a measure of trade elasticity, and L_{dt} is the population of county d in year t. τ_{cdt} is measured as $(1 + t_{cdt}/P)$, where t_{cdt} is the per ton county-to-county transportation costs (as in Donaldson and Hornbeck (2016)), and P is the average price per ton of transported goods between counties c and d at time t. We follow Hornbeck and Rotemberg (2024) and use a value for θ of 3.05 and a value for P of 38.7.

[54]In Medina county, log market access increased by 0.024 between 1840 and 1850, and by 0.026 between 1850 and 1860. The corresponding increases for Adams county were 0.024 and 0.025.

[55]We show that our results are robust to using an alternative market access measure that relies on county-to-county travel times in minutes, rather than on costs of transportation per ton-mile (as, for instance, in Chiovelli et al. (2024)). Namely, in $MA_{ct} = \sum_{d \neq c} (\tau_{cdt})^{-\theta} L_{dt}$, we measure τ_{cdt} as the travel time in minutes between county c and county d in decade t. Although the two measures specify different parameters in the calculation of market access, they are highly correlated.

government, and this allows better monitoring and communication. This assumption would not be met if proximity to D.C. per se contributed to spur economic growth. Although at this time D.C. was not an important manufacturing and trading center (Abbott, 1989), its proximity to Baltimore might invalidate this identification assumption and bias our estimates. We can address this concern by conducting placebo experiments, relating travel time between a county and other important economic centers to state presence. As we will show, conditional on our other controls, lowering travel time between a county and these other cities has no influence on state presence in the county.

Finally, we show that our results are unaffected by the inclusion of an extensive set of economic, geographic, and demographic controls. This further assuages concerns about the presence of time-varying unobservable shocks correlated with both $Log\ Time\ to\ DC_{ct}$ and state presence that are not captured by local railroad construction and market access. We describe these controls in the next section.[56]

5.3 Estimated Impact on State Presence

Main estimates. Table 2 presents results from the estimation of equation 2, using as dependent variable an indicator taking value one if the federal government has a presence in the county.[57] Column 1 reports estimates from the simple specification including only year fixed effects, county fixed effects, and the second-order polynomial in the straight line distance between the county and D.C. interacted with year fixed effects. In column 2 we additionally control for local railroad construction and for a county's market access. The estimated coefficient on $Log\ Time\ to\ DC_{ct}$ decreases by about 50 percent with the inclusion of these additional controls. This is consistent with their importance in order to account for time-varying shocks in the incentive to invest in state capacity generated by the development of the railroad network. However, the coefficient remains statistically significant and large in magnitude. A one standard deviation decrease in $Log\ Time\ to\ DC_{ct}$ is associated with an increase in the probability of state presence of 10 percentage points. This represents a 32 percent increase relative to the mean probability of observing state presence in a county.

[56] A common approach to identify the effect of railroad network expansions is the "inconsequential units approach," which identifies the effect for economically small units lying between large cities. The intuition behind this approach is that these units will be connected to a railroad only because they lie along a convenient route between two large cities (Redding and Turner, 2015). This approach is infeasible with the data that we use: the GIS network database from Donaldson and Hornbeck (2016) and Atack (2013) does not include detailed information that can allow to identify which cities were meant to be connected by the construction of new lines.

[57] In column 1 of Online Appendix Table A4, we show a pure correlation between state presence and travel time to D.C., estimating a specification with only year fixed effects. At any point in time, places that are further away from D.C. are less likely to see a presence of the federal government.

In the following columns we include an extensive set of additional controls. In column 3, we include a number of controls that capture a county's economic growth: the log of the county's total population, the share of the county's population that is employed in manufacturing, and urbanization rate (measured as the share of a county's population that lives in places with more than 2,500 inhabitants). The expansion of the railroad network might have made counties better connected to their state capital, which might affect the federal government's incentives to invest in state capacity if the federal government complements or substitutes state governments. In order to account for this possibility, we control for a second-order polynomial in the (log) travel time distance between the county and the state capital. In column 4, we additionally control for immigration (measured as the share of inhabitants that are foreign born), to account for any effect of the railroad network on immigrant flows to specific counties (Sequeira et al., 2020). In column 5, we account for possible differential trends in state presence in counties that were more affected by the Civil War. Specifically, we control for the number of casualties, the number of battles, an indicator equal to one if a major battle (defined as one with more than 500 recorded casualties) took place in the county, and an indicator equal to one if the county is within 5 miles of the General Sherman's march, all interacted with year fixed effects.[58] Finally, in column 6 we account for possible differential impacts of the railroad expansion on economic growth depending on a county's natural endowments.[59] Specifically, we control for average agricultural suitability, average rainfall, waterpower potential, and the number of mineral deposits, all interacted with year fixed effects.[60] The estimated effect of $Log\ Time\ to\ DC_{ct}$ is not affected by the inclusion of this extensive set of controls.[61]

In the most demanding specification, in column 6 of the table, the estimate implies a 31 percent increase in the probability of state presence if log travel time is reduced by one standard deviation. To further illustrate the economic significance of this effect, consider that the average log travel time decreased by 0.5 between 1850 and 1870. This decrease translates into an estimated 19 percent increase in the likelihood of state presence at the county level. To put this magnitude in perspective, the estimate on log market access implies that a one standard deviation increase in market access is associated with a 14 percent increase in the

[58] Feigenbaum et al. (2022) show that General Sherman's march was associated with significant capital destruction and declines in agricultural investment and manufacturing growth.

[59] As we move from column 2 to column 3, we drop 1,968 county-year observations with missing information on manufacturing employment. As we move from column 3 to column 4, we drop 720 county-year observations with missing information on the number of foreign born individuals.

[60] Hornbeck et al. (2024) show evidence that manufacturing activity grew faster in counties with less waterpower potential between 1850 and 1880, as the costs of steam power declined.

[61] In Online Appendix Figure A10 we present a graphical representation of the relationship between state presence and $Log\ Time\ to\ DC_{ct}$, after partialing out the full list of controls from the specification in column 6 of Table 2.

probability of state presence.[62]

Robustness. In Online Appendix Table A6, we show that our results are robust to a long battery of robustness checks. First, we drop from the sample counties that were affected by the Civil War.[63] Second, we limit the sample only to counties in states that were already part of the U.S. by 1860, or by 1820 (at the beginning of the sample period).[64] Third, we use an alternative control for market access that relies on county-to-county travel times in minutes, rather than on costs of transportation per ton-mile. Fourth, we show that our results are robust to alternative choices of the travel time parameters used to compute $Log\ Time\ to\ DC_{ct}$. Fifth, we keep only the middle year for each decade. Sixth, we show that travel time to D.C. at time t is not associated with differential state presence in the previous decade. Finally, in Online Appendix Figure A11, we show that the results are not driven by any specific decade, or by any specific state.

Distance to other important cities. Since our identification relies on the assumption that proximity to D.C. per se was not a significant driver of economic growth, we conduct a series of placebo tests to lend credibility to this assumption. We re-estimate equation 2 with the inclusion of a set of additional variables ($Log\ Time\ to\ P_{ct}$) measuring the travel time in year t between county c and placebo city P.[65] We include a set of "placebo cities" (New York, Boston, Philadelphia, New Orleans, Chicago, San Francisco, Saint Louis and Cincinnati) that were important economic centers but, differently from D.C., not the center of power of the federal government.[66] We report the results in Figure 4. In Panel (a), we plot the coefficients for travel time to D.C. (in red) and to each other city (in black) from a series of regressions for each city P, while in Panel (b) we plot results from a single regression with all cities included at the same time.[67] Panels (c) and (d) report results from similar regressions, but with the control for market access calculated using travel times instead of transportation costs between pairs of counties. Consistent with our interpretation, after

[62] Online Appendix Table A5 shows the estimated coefficients on some of the controls included in our most demanding specification.

[63] Results are robust to different ways of defining counties affected by the Civil War: all counties with battles with over 500 casualties, all counties affected by General Sherman's march, or all counties with any battle. Results are also robust to dropping all Confederate states in the 1860s and 1870s.

[64] This suggests that the results are not driven by the ability of the railroads to extend westward the American frontier (Bazzi et al., 2020).

[65] We include the full set of controls in column 6 of Table 2. Additionally, as we do for D.C., we control for a second order polynomial in the straight line distance between the county and each city P, interacted with year fixed effects.

[66] New York, Boston, and Philadelphia were the largest cities in the North-East, New Orleans was the largest city in the South, and San Francisco was the largest city in the West, over the sample period. Chicago, Saint Louis and Cincinnati were the largest cities in the Midwest at some point during our sample period.

[67] Since distances to New York, Boston, and Philadelphia are highly correlated, with include only distance to New York in the specification which includes all cities. Substituting travel time to New York City with travel time to Boston or to Philadelphia produces similar results.

accounting for local railroad construction and for overall market access, only travel time to D.C. explains the presence of the state in a county, while travel time to other important cities is not associated with differences in state presence.

Counterfactual growth in state presence. We can use our estimates to provide a counterfactual evolution of the presence of the federal government across the territory, had the travel time from D.C. not changed over the nineteenth century. To do so, we first compute, for each decade and for each county, the probability of state presence predicted by the estimation of equation 2. We then compute the counterfactual probability of state presence, had the travel times between each county and D.C. remained fixed at their values in 1820. We plot the average of these probabilities for each decade between 1820 and 1900 in Online Appendix Figure A12, together with the actual share of counties with state presence that we observe in the data. The trends in the three series are flat until the 1850s, when they start to grow. The post-1850s growth in the counterfactual scenario is significantly lower than what predicted by the model (and what we observe in the data). Relative to the share predicted by the model, the share of counties with state presence in the counterfactual scenario would have been lower by 7.3 percentage points in the 1860s, by 11.5 percentage points in the 1880s, and by 14.5 percentage points between 1900 and 1905. Had the travel time from D.C. not changed over the nineteenth century, the estimates imply that in the 1900-1905 period there would have been 404 fewer counties with state presence (i.e. $0.145 \times 2{,}784$ counties in the sample). This corresponds to a 24 percent decrease relative to the number of counties with state presence observed in 1905.

6 Monitoring Capacity as Mechanism

This section presents a series of results that support our monitoring mechanism. First, we show that the effect of lower travel time to D.C. is stronger for occupations and tasks where agency problems are expected to be more severe. Second, we show that lower travel time between D.C. and a given county is associated with a state organization with less patrimonial features.[68]

[68] Columns 2-7 of Online Appendix Table A4 show a correlation between the outcomes analyzed in this section and travel time to D.C., after conditioning only on year fixed effects. The estimates show that, at any point in time, places that are further away from D.C. have less workers employed in managerial and clerical occupations or in bureaus in charge of taxation, and are less likely to have a bureaucratic organization.

6.1 Impact on specific tasks

Occupations. If increased monitoring capacity is an important mechanism explaining the relationship between travel time to D.C. and state capacity, we expect our estimates to be larger for occupations where agency problems are more severe. Theoretically, we expect agency problems to be more severe for workers employed as managers of field offices or in other clerical positions, relative to non-clerical positions. Managers and clerks had greater discretion and autonomy in their tasks, and their performance was less directly tied to observable production metrics. The performance of blue-collar workers was instead more easily observable, and their compensation was often tied to simple metrics.[69] Consistent with this intuition, Table 3, panel A, shows that the effect of a reduced travel time between a county and D.C. is concentrated among managerial and clerical occupations: a one standard deviation decrease in travel time is associated with a 0.22 standard deviation increase in the presence of managerial workers and a 0.32 standard deviation increase in the presence of clerical workers. By contrast, there is no significant effect on the employment of non-clerical workers.

Local delegation of managerial power. Since delegating managerial tasks away from the center of power is particularly challenging when monitoring costs are high, we expect lower travel time to D.C. to be associated with increased managerial delegation, holding workforce size constant. To test for this hypothesis, we limit the sample to county-year observations with state presence, and we estimate a version of equation 2 in which the dependent variable is an indicator equal to one if there is at least one worker employed in a managerial role in the county, controlling also for fixed effects for the number of workers employed. As shown in column 1 of Panel B, holding fixed the size of the workforce in the county, counties with a lower travel time to D.C. have a significantly higher probability of being delegated managerial power. Relative to a 32 percent mean probability, decreasing $Log\ Time\ to\ DC_{ct}$ by one standard deviation increases the probability of delegation by 39 percent. Importantly, this result runs counter to the predictions of purely cognitive models of organizations without agency problems, which predict *less* decentralization of decision power as communication costs decrease. It is instead consistent with models predicting an increase in decentralization as incentive conflicts between principals and local offices decrease (Dessein, 2002; Bloom et al., 2012). This suggests that, in this specific context and historical period, faster travel to D.C. was relatively more valuable for addressing monitoring problems than for enabling central managers to assist local workers with complex decision-making.

Taxation vs public goods provision. Finally, we expect agency problems to be especially

[69] In our data, these occupations were often compensated "per day worked", "per hour worked", or on the basis of observed production (e.g. "per inspection" or "per drill hole").

severe for tasks related to taxation. The revenue collected by the federal government during the nineteenth century mainly came from taxation of custom and from excise taxes on goods such as tobacco and liquor. Since the output produced by these activities (i.e. the amount of taxes collected) depended on tax officers' measurement of the underlying values of the goods being taxed, and since these values were hard to observe from D.C., these government functions were especially prone to corruption and passive waste. On the other side, we expect agency problems to be less significant for activities related to the provision of public goods, whose output can be more easily observed. Consistent with this idea, column 2 of Table 3, panel B, shows that travel time to D.C. is significantly associated with the presence of bureuas of the federal government related to taxation. Columns 3 and 4 of the same panel show that state presence in bureaus in the War and Navy department, or in the Interior department, is not significantly affected by a lower travel time to D.C., consistent with these bureaus being tasked with public goods provision.[70]

6.2 From patrimonial to bureaucratic organization

The theoretical framework in section 2.1 highlights how enhanced monitoring capacity should be associated with a less patrimonial organization of the state. As monitoring costs decline, the relative value of employing workers embedded in personalistic, trust-based relationships with their superiors diminishes. In what follows, we present empirical evidence consistent with this hypothesis. All the results in this section are estimated on the (selected) subsample of county-year observations with state presence.

Workers' turnover. A patrimonial organization is associated with high turnover of the workforce: since trust-based relationships between workers and their superiors substitute for formal monitoring, the career of a worker is tightly linked to that of the superior. A lower travel time between a county and D.C., by increasing D.C.'s monitoring capacity, should reduce the extent of workers' turnover. We test this hypothesis in Table 4, where we estimate equation 2 using as dependent variable the share of workers in county c and year t who leave within the next two years.[71] In light of our discussion in the previous section, we expect

[70] Online Appendix Table A7 and Online Appendix Figures A13-A18 show robustness checks for the additional measures of state presence discussed in this section for which we find a significant effect of travel time to D.C. Results are highly robust to our extensive battery of robustness tests, with two exceptions when we use the number of clerks as dependent variable. First, while the effect of travel time to New York City on the number of clerks is insignificant, the confidence intervals are large in panels (b) and (d) of Online Appendix Figure A14, so that we cannot reject that the effect of travel time to New York City is different than travel time to D.C. Second, travel time to D.C. is significantly associated with the number of clerks employed in the previous decade, although the estimated effect is substantially smaller than the contemporaneous effect.

[71] Specifically, a worker leaves if, in the subsequent volume of the Official Register, they are not employed

monitoring to reduce turnover for occupations characterized by more agency problems. Thus, we separately analyze turnover for white collars (managers and clerical workers) and non-clerical workers. Column 1 shows that turnover among white collars is significantly higher in counties with a longer travel time to D.C.: two counties that are one standard deviation apart in their (log) travel time to D.C. have a turnover rate that differs by 10 percentage points (or 16 percent of the mean turnover rate among white collars during the nineteenth century). In contrast, travel time to D.C. is not associated with turnover among non-clerical workers (column 2).

Personal connections among workers. Since, when monitoring costs are high, personal networks are an important source through which the principal selects trustworthy workers, personal connections among co-workers are more common in patrimonial organizations. Using shared last names among workers in the same local office as a proxy for personal connections, we can show suggestive evidence that personal connections among co-workers are less likely when monitoring capacity increases. The dependent variable in column 3 of Table 4 is the share of workers in county c and year t who share their last name with a co-worker. In order to account for the mechanical correlation between the number of workers and the probability of having workers sharing a last name, we include fixed effects for the number of workers employed. The estimated coefficient on $Log\ Time\ to\ DC_{ct}$ implies that a one standard deviation higher (log) travel time to D.C. is associated with a doubling of the share of workers sharing the last name with a co-worker.

Relevance of trust. As an additional test of the monitoring mechanism, we now show that a lower travel time from a county to D.C. significantly increased the employment of a group that, absent formal monitoring, was arguably not trusted by D.C., namely Southern-born individuals after the Civil War.

Figure 5 motivates our empirical test. It plots the evolution over time in the number of federal workers, differentiating between those who were born in a Confederate state and those who were born in any other state. We normalize the two series by the population of these two regions. Individuals born in a confederate state were less represented in the federal bureaucracy even before the Civil War, with about 0.1 employees per 1,000 inhabitants, compared to about 0.2 employees per 1,000 inhabitants for the other states. However, the representation of the two groups starts to diverge significantly after the war. At the onset of the conflict, there is a sizeable decline in the number of Southern-born federal bureaucrats. More surprisingly, the North-South employment gap is persistent (see gray series): while the employment of both groups constantly increases after 1861 as the federal government

in the same local office. We exclude observations in 1905, since we do not have information on which employees leave by 1907.

expands its scope, the difference in workers per capita between Southern and Northern states increases from about 0.1 in 1859 to about 0.4 in 1869, and remains constant over the next decades. We interpret this as evidence of the federal government's lower trust towards workers from former confederate states after the end of the conflict.

However, as we show in column 4 of Table 4, this gap was on average less pronounced in counties that could be monitored more easily from D.C. Focusing on the post-war period, we estimate a version of equation 2 using as dependent variable the share of workers in county c and year t who are born in the South. The estimated coefficient implies that a one standard deviation decrease in $Log\ Time\ to\ DC_{ct}$ is associated with a 14 percentage points increase in the share of workers born in the South (this is a 52 percent increase relative to an average share of 26.4 percent in the post-was period). As we show in column 5, this relationship was absent before 1861.

This result provide suggestive evidence in favor of monitoring capacity being an important mechanism: the government's ability to monitor the behavior of its agents throughout the territory reduces reliance on trust between the central managers in D.C. and workers employed away from D.C.

7 Conclusion

The historical development of states progressed from small, patrimonial organizations to large, territorially expansive states governed in a bureaucratic way. There were undoubtedly many important drivers of the development of modern states. In this paper, we focus on one — technologies that reduce agency costs between rulers and officials, — and we show that it was important in explaining the growth and organizational change of the U.S. federal government. We leverage a unique dataset on the presence and internal organization of the U.S federal government during the nineteenth century and develop an identification strategy that exploits the expansion of the railroad network to isolate changes in travel costs between D.C. and individual locations across the U.S. territory. Our findings show that reductions in travel time between a county and D.C. account for a significant share of the geographic expansion of the federal government during the nineteenth century. Several empirical tests suggest that increased monitoring capacity was an important mechanism driving this effect.

The findings of this paper underscore the importance of agency problems and the role of monitoring-enhancing technologies in shaping the historical growth and organizational evolution of centralized states. While this idea has a long standing tradition in the literature in sociology on the historical development of state capacity (Weber, 1978; Coleman, 1990; Kiser and Schneider, 1994), we provide the first empirical test of this mechanism. Our results sug-

gest that agency problems between top mangers and lower-level officials are crucial not only to understand the functioning of bureaucracies at a given point in time, but also to explain their growth and organizational evolution in response to technological progress in communication and transportation. This insight supports arguments, such as in Heldring (2025), suggesting that different systems of bureaucratic organization—characterized by varying degrees of discretion over personnel decisions—may have been optimal at different stages of historical development.

Although our analysis centers on innovations in monitoring capacity, several alternative mechanisms contributed to the expansion and organizational change of centralized states during the nineteenth century. For instance, independent of agency problems, an improved ability to communicate across space facilitated collaboration among federal bureaucrats in the provision of public goods. Moreover, the same technologies that enabled better monitoring also spurred economic growth. Our finding that a county's market access is significantly associated with the geographic expansion of the federal government suggests that both the increased taxable surplus and the rising demand for public services that resulted from increased growth played a relevant role. Systematically exploring these additional mechanisms, in this or other historical contexts, offers a promising direction for future research.

In the specific context of the United States, the Civil War clearly marked a turning point in the evolution of the federal government. We hope that our novel dataset will serve as a foundation for future research examining the impact of this and other historical events on the development of the federal government. While our analysis focuses on the federal bureaucracy, it is important to recognize that state and local governments were primarily responsible for public goods provision during the nineteenth century. A systematic investigation into the determinants of local state capacity represents an important direction for future research.

Finally, although the provision of those public goods—such as education, healthcare, and local infrastructure— that are typically considered drivers of economic growth was primarily the responsibility of lower levels of government, increases in the federal government's monitoring capacity may have influenced local economic growth through lower corruption at the local level. While a comprehensive analysis of this question lies beyond the scope of this paper, Online Appendix Table A8 shows that counties that experience a decrease in travel time to D.C. see an increase in the size of manufacturing establishments—measured as average revenue or number of workers—and higher levels of urbanization.[72] We view a

[72]Data on average establishment size is available at the county-level from the 1860, 1870, 1880, 1890, and 1900 censuses. We re-estimate our main specification from column 6 of Table 2 (excluding controls for local economic growth).

systematic investigation of how federal monitoring capacity affects the performance of local offices—and, by extension, local economic activity—as an additional promising direction for future research.

Data availability statement

The data and code underlying this research is available on Zenodo at https://doi.org/10.5281/zenodo.17573595.

Bibliography

Abbott, C. (1989): "Perspectives on Urban Economic Planning: The Case of Washington, D.C., Since 1880," *The Public Historian*, 11, 5–21.

Akhtari, M., D. Moreira, and L. Trucco (2022): "Political turnover, bureaucratic turnover, and the quality of public services," *American Economic Review*, 112, 442–493.

Allen, R., M. Bertazzini, and L. Heldring (2023): "The economic origins of government," *American Economic Review*, 113, 2507–45.

Allen, R. C. (1997): "Agriculture and the Origins of the State in Ancient Egypt," *Explorations in Economic History*, 34, 135–154.

Aneja, A. and G. Xu (2023): "Strengthening State Capacity: Civil Service Reform and Public Sector Performance during the Gilded Age," .

Angelucci, C., S. Meraglia, and N. Voigtländer (2024): "Organizing a Kingdom," Tech. rep., National Bureau of Economic Research.

Armstrong, J. A. (2015): *The European administrative elite*, Princeton University Press.

Ashraf, N., O. Bandiera, E. Davenport, and S. S. Lee (2020): "Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the Delivery of Public Services," *American Economic Review*, 110, 1355–94.

Ashraf, N., O. Bandiera, and K. B. Jack (2014): "No margin, no mission? A field experiment on incentives for public service delivery," *Journal of Public Economics*, 120, 1–17.

Atack, J. (2013): "On the use of geographic information systems in economic history: The american transportation revolution revisited," *The Journal of Economic History*, 73, 313–338.

Atack, J., F. Bateman, M. Haines, and R. A. Margo (2010): "Did railroads induce or follow economic growth?: Urbanization and population growth in the American Midwest, 1850–1860," *Social Science History*, 34, 171–197.

Atack, J. and R. A. Margo (2011): "The impact of access to rail transportation on agricultural improvement: The American Midwest as a test case, 1850–1860," *Journal of Transport and Land Use*, 4, 5–18.

Bandiera, O., M. C. Best, A. Khan, and A. Prat (2021): "The Allocation of Authority in Organizations: A Field Experiment with Bureaucrats," *Quarterly Journal of Economics*, 136, 2195–2242.

Bates, R. H. (1987): *Essays on the Political Economy of Rural Africa*, University of California Press.

Bazzi, S., M. Fiszbein, and M. Gebresilasse (2020): "Frontier Culture: The Roots and Persistence of "Rugged Individualism" in the United States," *Econometrica*, 88, 2329–2368.

Becker, S., A. Ferrara, E. Melander, and L. Pascali (2022): "Wars, Taxation and Representation: Evidence from Five Centuries of German History," *CEPR Press Discussion Paper No. 15601*.

Besley, T., D. Bogart, J. Chapman, and N. Palma (2025): "Justices of the Peace: Legal Foundations of the Industrial Revolution," Tech. rep., Arthur Lewis Lab, The University of Manchester.

Besley, T., R. Burgess, A. Khan, and G. Xu (2022): "Bureaucracy and development," *Annual Review of Economics*, 14, 397–424.

Besley, T. and T. Persson (2008): "Wars and state capacity," *Journal of the European Economic Association*, 6, 522–530.
——— (2009): "The Origins of State Capacity: Property Rights, Taxation, and Politics," *American Economic Review*, 99, 1218–44.
——— (2011): "Pillars of prosperity," in *Pillars of Prosperity*, Princeton University Press.
——— (2013): "Taxation and development," in *Handbook of public economics*, Elsevier, vol. 5, 51–110.
Bloom, N., L. Garicano, R. Sadun, and J. Van Reenen (2014): "The Distinct Effects of Information Technology and Communication Technology on Firm Organization," *Management Science*, 2859–2885.
Bloom, N., R. Sadun, and J. Van Reenen (2012): "The Organization of Firms Across Countries," *The Quarterly Journal of Economics*, 127, 1663–1705.
Cantoni, D., C. Mohr, and M. Weigand (2024): "The Rise of Fiscal Capacity: Administration and State Consolidation in the Holy Roman Empire," Tech. Rep. 5.
Carneiro, R. (1970): "A Theory of the Origin of the State," *Science*, 169, 733–738.
Cassese, S. (1982): *Il sistema amministrativo italiano*, Roma - Istituto Poligrafico e Zecca dello Stato.
Chambru, C., E. Henry, and B. Marx (2024): "The Dynamic Consequences of State-Building: Evidence from the French Revolution," *American Economic Review*, 114, 3578–3622.
Chandler, A. D. (1962): *Strategy and Structure: Chapters in the History of the Industrial Enterprise*, Cambridge, MA: MIT Press.
Chen, C. (2017): "Management Quality and Firm Hierarchy in Industry Equilibrium," *American Economic Journal: Microeconomics*, 203–244.
Chiovelli, G., L. Fergusson, L. Martinez, J. D. Torres, and F. Valencia Caicedo (2023): "Bourbon Reforms and State Capacity in the Spanish Empire," Tech. rep., Working Paper.
Chiovelli, G., M. S., and E. Papaioannou (2024): "Landmines and Spatial Development," *Working paper*.
Coleman, J. S. (1990): "Rational Organization," *Rationality and Society*, 94–105.
Colonnelli, E., M. Prem, and E. Teso (2020): "Patronage and selection in public sector organizations," *American Economic Review*, 110, 3071–3099.
Crenson, M. A. (1975): *The Federal Machine: Beginnings of Bureaucracy in Jacksonian America.*, Johns Hopkins University Press.
Dal Bo, E., F. Finan, and M. A. Rossi (2013): "Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service," *The Quarterly Journal of Economics*, 128, 1169–1218.
Deserranno, E. (2019): "Financial incentives as signals: experimental evidence from the recruitment of village promoters in Uganda," *American Economic Journal: Applied Economics*, 11, 277–317.
Dessein, W. (2002): "Authority and Communication in Organizations," *The Review of Economic Studies*, 69, 811–838.
Dincecco, M. and G. Katz (2016): "State capacity and long-run economic performance," *The Economic Journal*, 126, 189–218.
Donaldson, D. and R. Hornbeck (2016): "Railroads and American economic growth: A "market access" approach," *The Quarterly Journal of Economics*, 131, 799–858.
Duflo, E., R. Hanna, and S. P. Ryan (2012): "Teacher Performance Pay: Experimental Evidence from India," *American Economic Review*, 102, 1241–1278.
Dzansi, J., A. Jensen, D. Lagakos, and H. Telli (2025): "Technology and Tax Capacity: Evidence from Local Governments in Ghana," Tech. rep., Working Paper.
Elmes, W. (1879): *The Executive Departments of the United States at Washington*, McGill and Co.
Evans, P. and J. E. Rauch (1999): "Bureaucracy and growth: A cross-national analysis of the effects of" Weberian" state structures on economic growth," *American sociological review*, 748–765.
——— (2000): "Bureaucratic Structure and Bureaucratic Performance in Less Developed Countries," *Journal of Public Economics*, 75, 49–71.
Feigenbaum, J. J., J. Lee, and F. Mezzanotti (2022): "Capital Destruction and Economic Growth: The Effects of Sherman's March, 1850-1920," *American Economic Journal: Applied*, 14, 301–342.
Fenske, J. (2014): "Ecology, Trade, and States in Pre-Colonial Africa," *Journal of the European Economic Association*, 612–640.
Finan, F., B. A. Olken, and R. Pande (2017): "The personnel economics of the developing state," *Handbook of economic field experiments*, 2, 467–514.

Finer, S. E. (1997): *The History of Government from the Earliest Times: Ancient monarchies and empires (Vol. 1)*, Oxford University Press.
Fish, C. R. (1905): *The civil service and the patronage*, Longmans, Green, and Company.
Fishlow, A. (1965): *American Railroads and the Transformation of the Ante-bellum Economy*, Harvard University Press.
Fogel, R. W. (1965): "Railroads and American economic growth: Essays in econometric history," *Business History Review (pre-1986)*, 39, 130.
Folke, O., S. Hirano, J. M. Snyder, and M. Ting (2012): "Elections and reform: The adoption of civil service systems in the U.S. states," *Journal of Theoretical Politics*, 25.
Garfias, F. and E. A. Sellars (2021): "Fiscal Legibility and State Development: Evidence from Colonial Mexico," *Unpublished manuscript, Yale University. url: https://www. dropbox. com/s/g06yaf7ib7m6u2t/FiscalLegibilityState Development. pdf*.
Garicano, L. (2000): "Hierarchies and the Organization of Knowledge in Production," *Journal of Political Economy*, 108, 874–904.
Gennaioli, N. and H.-J. Voth (2015): "State capacity and military conflict," *The Review of Economic Studies*, 82, 1409–1448.
Giroud, X. (2013): "Proximity and Investment: Evidence from Plant-Level Data," *Quarterly Journal of Economics*, 128, 861–915.
Gumpert, A., S. Henrike, and M. Antoni (2022): "Firm Organization with Multiple Establishments," *Quarterly Journal of Economics*, 1091–1138.
Hacker, J. D. (2010): "Decennial Life Tables for the White Population of the United States, 1790-1900." *Historical Methods*, 43, 45–79.
Hamilton, A. (1795): *The Works of Alexander Hamilton, (Federal Edition)*, vol. 6, G. P. Putnam's Sons.
Heldring, L. (2025): "Historical Government: Origins, Evolution, and Varieties," Tech. rep., Working Paper.
Hoogenboom, A. A. (1968): *Outlawing the spoils: a history of the civil service reform movement, 1865-1883*, vol. 50, University of Illinois Press.
Hornbeck, R. (2010): "Barbed Wire: Property Rights and Agricultural Development," *The Quarterly Journal of Economics*, 125, 767–810.
Hornbeck, R., S. Hsuan-Ming Hsu, A. Humlum, and M. Rotemberg (2024): "Gaining Steam: Incumbent Lock-in and Entrant Leapfrogging," *Working paper*.
Hornbeck, R. and M. Rotemberg (2024): "Growth Off the Rails: Aggregate Productivity Growth in Distorted Economies," *Journal of Political Economy*, 132, 3547–3602.
Iyer, L. and A. Mani (2012): "Traveling agents: political change and bureaucratic turnover in India," *Review of Economics and Statistics*, 94, 723–739.
Jayachandran, C., M. Thorpe, R. Subramanian, and V. Nagadevara (2020): *Business clusters: partnering for strategic advantage*, Taylor & Francis.
Jefferson, T. (1854): *The Writings of Thomas Jefferson*, vol. 4, Taylor Maury.
Jensen, A. (2022): "Employment Structure and the Rise of the Modern Tax System," *American Economic Review*, 112, 213–34.
Kelley, E. M., G. Lane, and D. Schonholzer (2021): "Monitoring in Small Firms: Experimental Evidence from Kenyan Public Transit," Tech. rep., Working Paper.
Khan, M. Y. (2023): "Mission Motivation and Public Sector Performance: Experimental Evidence from Pakistan," Tech. rep., Working Paper.
Kiser, E. (1994): "Markets and hierarchies in early modern tax systems: a principal-agent analysis," *Politics & Society*, 22, 284–315.
Kiser, E. and S. M. Karceski (2017): "Political Economy of Taxation," *Annual Review of Political Science*, 75–92.
Kiser, E. and J. Schneider (1994): "Bureaucracy and efficiency: an analysis of taxation in early modern Prussia," *American Sociological Review*, 187–204.
Lamphere, G. N. (1881): *The United States Government: Its Organization and Practical Workings*, Lippincott and co.
Lee, M. M. and N. Zhang (2017): "Legibility and the informational foundations of state capacity," *The Journal of Politics*, 79, 118–132.

Lucas Jr, R. E. (1978): "On the size distribution of business firms," *The Bell Journal of Economics*, 508–523.

Mann, M. (1986): *The Sources of Social Power. Volumen 1: A History of Power from the Beginning to AD 1760.*, Cambridge University Press.

Mayoral, L. and O. Olsson (2019): "Pharaoh's Cage: Environmental Circumscription and Appropriability in Early State Development," *Available at SSRN 3496225*.

Mayshar, J., O. Moav, and Z. Neeman (2017): "Geography, Transparency, and Institutions," *American Political Science Review*, 111, 622–636.

Mayshar, J., O. Moav, and L. Pascali (2022): "The origin of the state: Land productivity or appropriability?" *Journal of Political Economy*, 130, 1091–1144.

Moreira, D. and S. Pérez (2020): "Civil Service Reform and Organizational Practices: Evidence from the 1883 Pendleton Act," Tech. rep., mimeo.

——— (2022): "Who Benefits from Meritocracy?" Tech. rep., National Bureau of Economic Research.

Muralidharan, K. and V. Sundararaman (2011): "Teacher Performance Pay: Experimental Evidence from India," *Journal of Political Economy*, 119, 39–77.

Nerlove, M. (1966): "Railroads and American economic growth," *The Journal of Economic History*, 26, 107–115.

Olson, M. (1993): "Dictatorship, Democracy, and Development," *American Political Science Review*, 87, 567–576.

Ornaghi, A. (2016): "Civil service reforms: Evidence from US police departments," *Job Market Paper*.

Penrose, E. T. (1959): *The Theory of the Growth of the Firm*, New York: Wiley.

Prince, C. E. and M. Keller (1989): *The U.S. Customs Service: a bicentennial history*, Department of the Treasury U.S. Customs Service, Washington D.C.

Redding, S. J. and M. A. Turner (2015): "Transportation Costs and the Spatial Organization of Economic Activity," *Handbook of Urban and Regional Economics*, 1339–1398.

Riaño, J. F. (2021): "Bureaucratic nepotism," *Available at SSRN 3995589*.

Ruiz, É. (2013): "Trop de fonctionnaires? Contribution à une histoire de l'État par ses effectifs (France, 1850-1950)," Ph.D. thesis, Ecole des Hautes Etudes en Sciences Sociales (EHESS).

Sanchez de la Sierra, R. (2020): "On the Origins of the State: Stationary Bandits and Taxation in Eastern Congo," *Journal of Political Economy*, 128, 32–74.

Schonholzer, D. and P. Francois (2023): "Environmental Circumscription and the Origins of the State," *Working paper*.

Scott, J. C. (1998): *Seeing like a state: How certain schemes to improve the human condition have failed*, Yale University Press.

——— (2017): *Against the grain: A deep history of the earliest states*, Yale University Press.

Sequeira, S., N. Nunn, and N. Qian (2020): "Immigrants and the Making of America," *The Review of Economic Studies*, 87, 382–419.

Shahe Emran, M., M. Morshed, and J. E. Stiglitz (2021): "Microfinance and Missing Markets," *Canadian Journal of Economics*, 54, 34–67.

Short, L. M. (1923): *The Development of National Administrative Organization in the United States (A Study by the Institute for Government Research)*, Johns Hopkins Press.

Snowberg, E. and M. Ting (2019): "An Organizational Theory of State Capacity," Tech. rep., Working Paper.

Spenkuch, J. L., E. Teso, and G. Xu (2021): "Ideology and performance in public organizations," Tech. rep., National Bureau of Economic Research.

Tilly, C. (1975): *The Formation of National States in Western Europe*, Princeton University Press.

——— (1990): *Coercion, capital, and European states*, Oxford: Blackwell.

Turquan, V. (1899): *Essai de recensement des employés et fonctionnaires de l'État, suivi d'une statistique des pensionnaires de l'État, par Victor Turquan...*, Au secrétariat de la Société d'économie sociale.

Ujhelyi, G. (2014): "Civil service rules and policy choices: evidence from US state governments," *American Economic Journal: Economic Policy*, 6, 338–380.

U.S. Customs Service, . (1988): *A history of enforcement in the United States Customs Service, 1789-1875.*, San Francisco, California.

Voth, J. and G. Xu (2020): "Discretion and destruction: promotions, performance, and patronage in the

Royal Navy," Tech. rep., CEPR Discussion Paper No. DP13963.
Wallis, J. J. (2000): "American Government Finance in the Long Run: 1790 to 1990," *The Journal of Economic Perspectives,*, 14, 61–82.
Washington, G. (1855): *The writings of George Washington*, vol. 10, Little, Brown and company.
Weaver, J. (2021): "Jobs for Sale: Corruption and Misallocation in Hiring," *American Economic Review*, 110, 3093–3122.
Weber, M. (1978): *Economy and society: An outline of interpretive sociology*, University of California press.
Webster, D. (1904): *The Writings and Speeches of Daniel Webster: Writings and speeches hitherto uncollected*, vol. 4, Little, Brown Company.
White, E. N. (2004): "From privatized to government-administered tax collection: tax farming in eighteenth-century France 1," *The Economic History Review*, 57, 636–663.
White, L. D. (1954): *The Jacksonians; a study in administrative history, 1829-1861*, Macmillan.
Wittfogel, K. A. (1976): *Oriental despotism*, New Haven: Yale University Press.
Woodson, T., L. N. Smith, and H. H. Pearson (1985): *The centenary edition of the works of Nathaniel Hawthorne. Vol. 16.*, Ohio State University.
Xu, G. (2018): "The costs of patronage: Evidence from the british empire," *American Economic Review*.

Table 1: Local economic development and state presence

	(1)	(2)	(3)	(4)	(5)
	Dependent variable: *State Presence*				
Log Population	0.015***				0.010***
	(0.002)				(0.003)
Urbanization		0.247***			0.206***
		(0.029)			(0.029)
Share Manu. Emp.			0.088		0.076
			(0.056)		(0.055)
Railroad				0.068***	0.054***
				(0.008)	(0.008)
Observations	96,875	96,427	94,897	96,875	94,897
Mean dep. var.	0.307	0.309	0.310	0.307	0.310

Notes: The unit of observation is a county-year. *State presence* is an indicator taking value one if the federal government is present in the county. *Log Population* is the logarithm of county population, *Urbanization* is the share of the county population living in places with more than 2,500 inhabitants, *Share Manu. Emp.* is the share of the county population employed in manufacturing, and *Railroad* is an indicator taking value one if the county contains any railroad track. All specifications control for county fixed effects, year fixed effects, and a second-order polynomial in the straight line distance between the county and D.C. interacted with year fixed effects. Standard errors are clustered at the county level. D.C. and the Alaska districts are excluded. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 2: Travel time to D.C. and state presence

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable: *State Presence*					
Log Time to DC	-0.219***	-0.122***	-0.108***	-0.111***	-0.109***	-0.120***
	(0.030)	(0.033)	(0.033)	(0.033)	(0.034)	(0.035)
Observations	96,875	96,759	94,791	94,071	94,071	94,071
Mean dep. var.	0.307	0.307	0.310	0.312	0.312	0.312
Std. dev. Log Time to DC	0.814	0.814	0.817	0.816	0.816	0.816
Railroads controls	No	Yes	Yes	Yes	Yes	Yes
Econ. returns controls	No	No	Yes	Yes	Yes	Yes
Immigration	No	No	No	Yes	Yes	Yes
Civil war	No	No	No	No	Yes	Yes
Endowments	No	No	No	No	No	Yes

Notes: The unit of observation is a county-year. *State presence* takes value one if the federal government is present in the county. *Log Time to DC* is the log of travel time (in minutes) between D.C. and the county's centroid. All specifications control for county fixed effects, year fixed effects, and a second-order polynomial in the distance between the county and D.C. interacted with year fixed effects. Railroad controls include an indicator taking value one if the county contains any railroad track, a cubic polynomial in the length of railroad track in the county, and the county's log market access. Controls for economic returns include the log of the county's total population, the share of the county's population that is employed in manufacturing, urbanization, and a second-order polynomial in distance from the state capital. The control for immigration is the share of foreign born in the county. Civil war controls include the number of casualties, the number of battles, an indicator equal to one if there was any major battle in the county, and an indicator equal to one if the county is within 5 miles of the Sherman's march, all interacted with year fixed effects. Endowments control include average agricultural suitability, average rainfall, waterpower potential, and the number of mineral deposits, all interacted with year fixed effects. Standard errors in parentheses, clustered at the county-level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3: Travel time to D.C. and employment in specific tasks

Panel A:	(1)	(2)	(3)	(4)
	Log (1 plus) number of workers employed as:			
	Managers	Clerks	Non-clerical High skills	Non-clerical Low skills
Log Time to DC	-0.085***	-0.251***	-0.068	0.036
	(0.029)	(0.052)	(0.043)	(0.068)
Observations	94,071	94,071	94,071	94,071
Std. dev. dep. var.	0.320	0.649	0.626	0.713
Std. dev. Log Time to DC	0.816	0.816	0.816	0.816
Panel B:		State presence in specific bureaus:		
	Manager present	Taxation	War and Navy	Interior
Log Time to DC	-0.138**	-0.265***	0.024	0.021
	(0.069)	(0.038)	(0.023)	(0.023)
Observations	28,954	94,071	94,071	71,735
Mean dep. var.	0.318	0.196	0.062	0.072
Std. dev. Log Time to DC	0.893	0.816	0.816	0.828

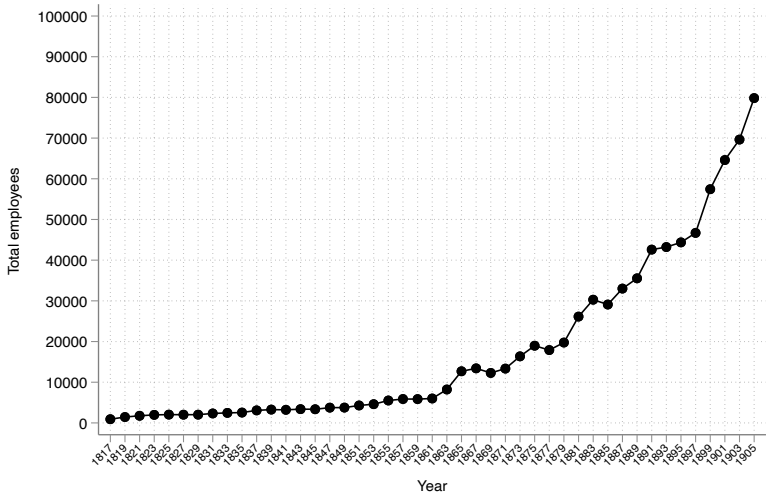
Notes: The unit of observation is a county-year. *Log Time to DC* is the log of travel time (in minutes) between D.C. and the county's centroid. In Panel A, the dependent variable in columns 1-4 is the logarithm of (one plus) the total number of workers employed in managerial, clerical, non-clerical high skills, and non-clerical low skills occupations, respectively. In Panel B, the dependent variable in column 1 is an indicator equal to one if there is at least one manager in the county; the dependent variable in columns 2-4 is an indicator equal to one if a bureau of the federal government related to taxation is present in the county, if the War or Navy department is present in the county, and if the Interior department is present in the county, respectively. All specifications include the full set of controls from the specification in column 6 of Table 2. In column 1 of Panel B, we restrict the sample only to county-year observations with state presence, and we additionally control for fixed effects for the number of workers in the county. Standard errors in parentheses, clustered at the county-level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 4: Travel time to D.C. and patrimonial features of the organization

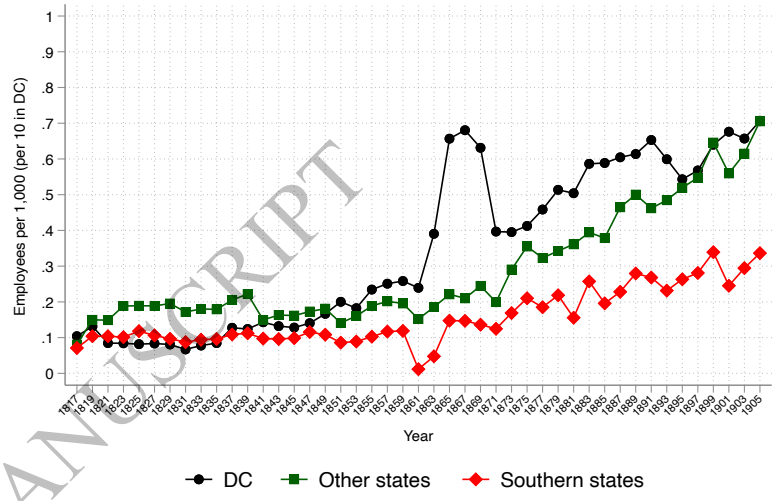
	(1) Turnover White Collars	(2) Turnover Non-clerical	(3) Shared Last name	(4) Born in South post 1861	(5) Born in South pre 1861
Log Time to DC	0.108**	-0.083	0.044***	-0.159***	-0.000
	(0.049)	(0.055)	(0.016)	(0.045)	(0.088)
Observations	20,305	18,746	28,945	22,097	4,827
mean DV	0.611	0.585	0.040	0.264	0.305
sd Log Time to DC	0.911	0.927	0.893	0.863	0.919

Notes: The unit of observation is a county-year. *Log Time to DC* is the log of travel time (in minutes) between D.C. and the county's centroid. *Turnover White Collars* (respectively, *Turnover Non-clerical*) is the share of white collars (respectively, non-clerical workers) who leave their local office between year t and year t-2. *Shared Last Name* is the share of employees who share a last name with at least one other worker in the same local office. *Born in South* is the share of employees who were born in a state that was part of the Confederacy. All specifications include the full set of controls from the specification in column 6 of Table 2. In column 3 we additionally control for fixed effects for the number of workers in the county. In all columns, we restrict the sample only to county-year observations with state presence. In columns 1 and 2 we are also excluding the year 1905, as we cannot measure turnover for employees in 1905. In column 4 we include only years after 1859, and in column 5 we include only years before 1861. Standard errors in parentheses, clustered at the county-level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

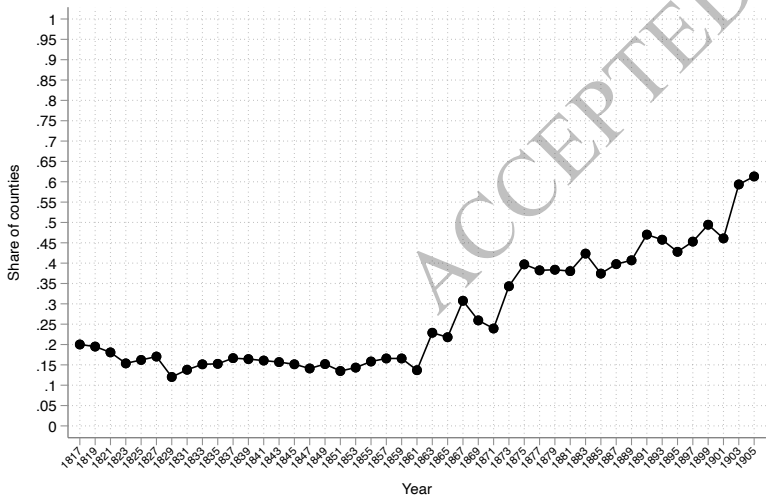
Figure 1: Growth of the U.S. Federal Government, 1817-1905



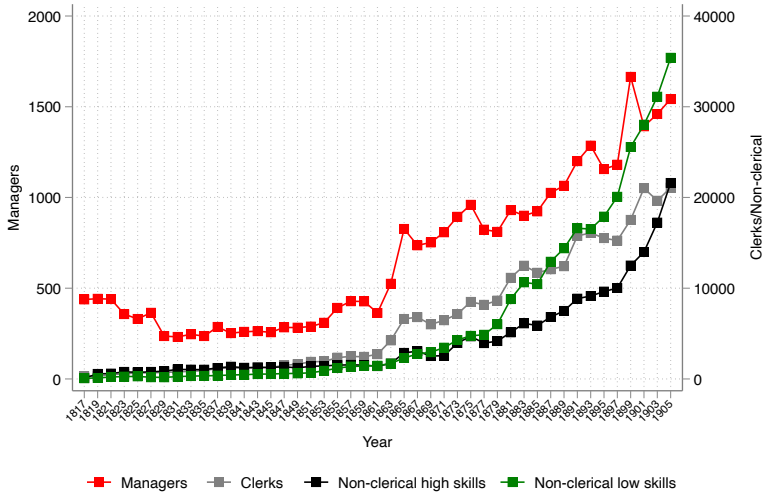
(a) Total number of employees



(b) Employees per capita in D.C. vs South vs Outside South



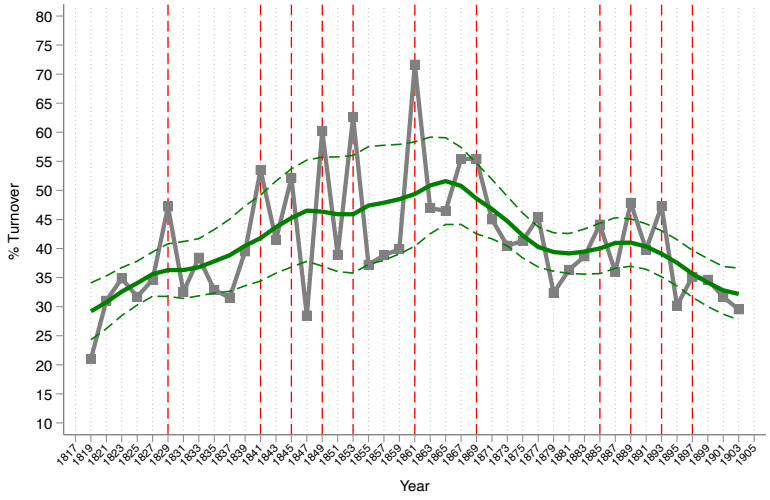
(c) Share of counties with state presence



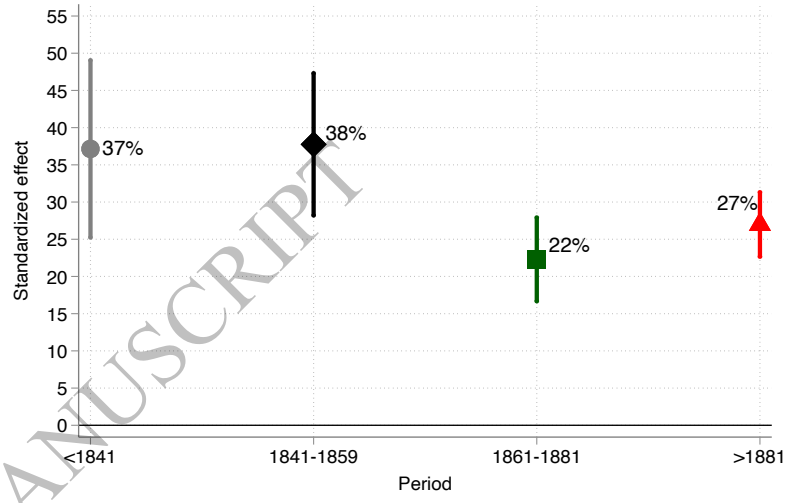
(d) Number of employees by occupational category

Notes: The figure shows the evolution over time of: total number of employees (Panel A); number of employees in D.C. per 10 inhabitants (in black) in Southern states per 1,000 inhabitants (in red) and in other states per 1,000 inhabitants (in green) (Panel B); share of counties with federal government presence (Panel C); number of employees by occupational category (Panel D).

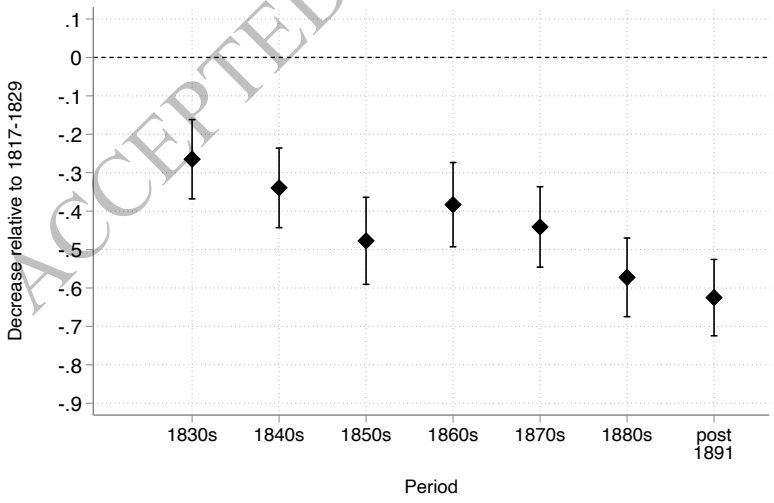
Figure 2: Organizational Features of the U.S. Federal Government



(a) Turnover



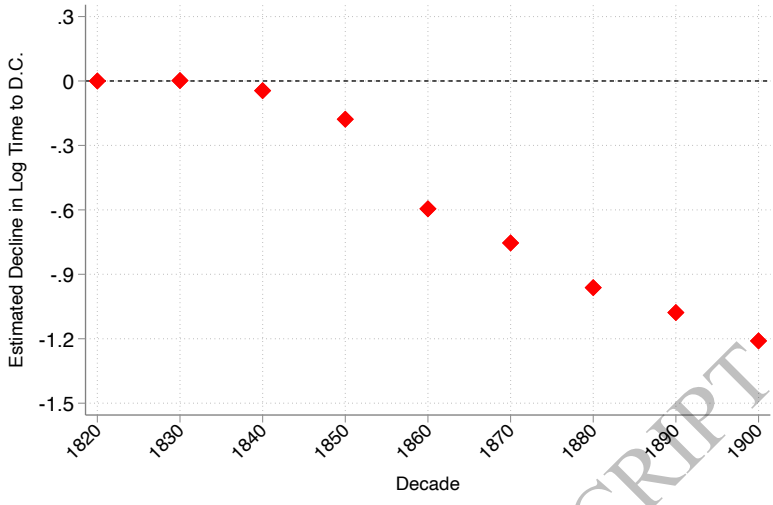
(b) Correlation between a worker's and her supervisor's turnover



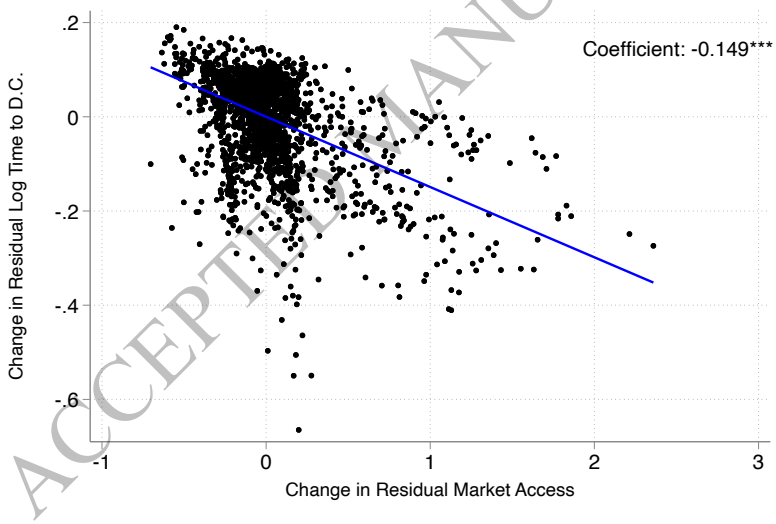
(c) Decrease of probability that co-workers share last name

Notes: The figure shows aggregate turnover, i.e. the share of employees leaving the bureaucracy, over 1817-1905 (Panel A), the standardized coefficients on β_τ from equation 1, with 95 percent confidence intervals based on standard errors clustered at the organizational unit times year level (Panel B), the decrease in the probability that an employee shares the same last name with a co-worker, for each decade, relative to the period 1817-1829 (Panel C). The red vertical lines in Panel A indicate years in which the party of the President changes. See section 4.3 for full details.

Figure 3: Decline in travel time to D.C. and relationship with changes in market access



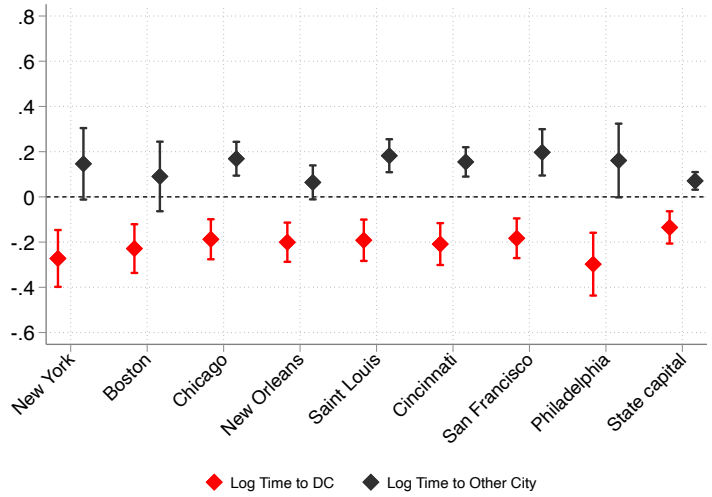
(a) Average decline in travel time to D.C.



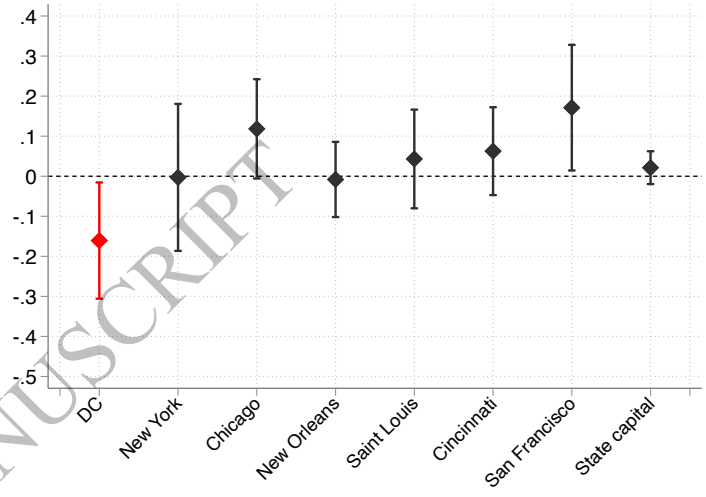
(b) Residual change in travel time to D.C. and residual change in market access, between 1870 and 1880.

Notes: The top panel shows the estimated average decline in travel time between D.C. and counties throughout the U.S. between 1830 and 1900, relative to 1820. It is computed by regressing $Log\ Time\ to\ DC_{ct}$ on a set of indicators turning to one for specific decades (and excluding 1820) and county fixed effects. The bottom panel shows the relationship between changes in log travel time to D.C. between 1870 and 1880 and changes in log market access over the same period, after partialing out changes in local railroad construction between 1870 and 1880 and a second order polynomial in the straight line distance between the county and D.C. Each point is a county that appears both in 1880 and in 1890. In the top right of the panel, we report the estimated coefficient and significance level from a linear regression.

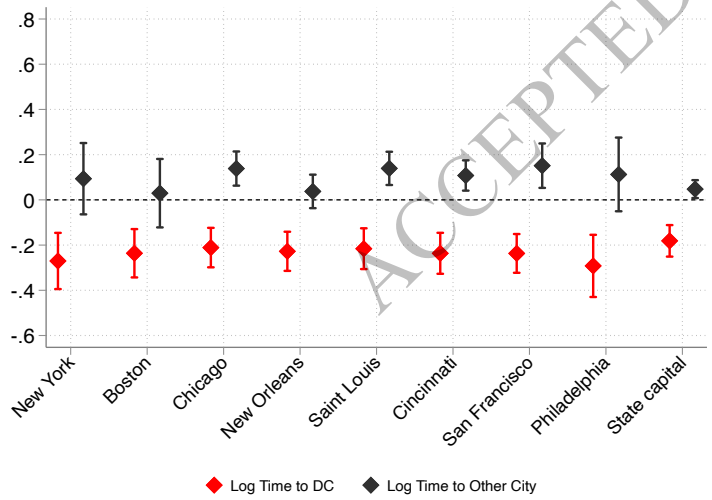
Figure 4: Lower travel time to other cities is not associated with more state presence



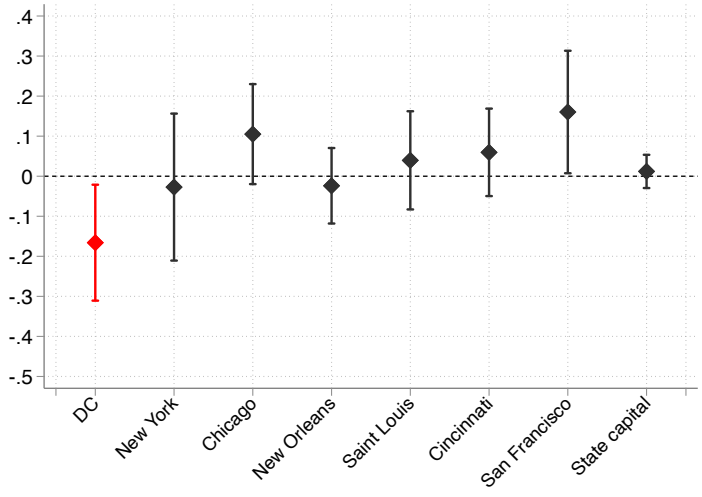
(a)



(b)



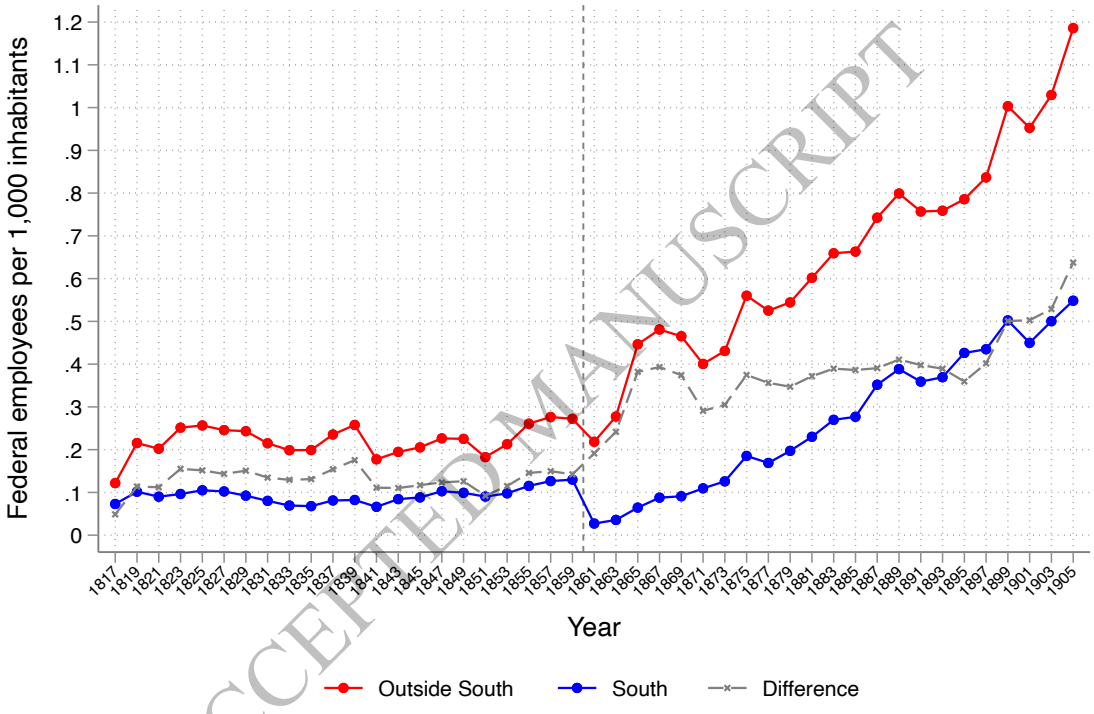
(c)



(d)

Notes: Results from estimation of the specification in column 6 of Table 2, adding the log of travel time between city P and the county's centroid. Panel (a) reports effects for D.C. (in red) and for each other city (in black) from a series of regressions for each city P, while Panel (b) reports results from a single regression with all cities included. Panels (c) and (d) report results from similar regressions, but with $log(MA)$ calculated using travel time instead of transportation costs. 95 percent confidence intervals are based on standard errors clustered at the county-level.

Figure 5: Civil War and decline in Southerners' employment



Notes: The figure plots the evolution over time in the number of federal employees who were born in a confederate state (i.e., Alabama, Arkansas, Florida, Georgia, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Texas, Virginia) (in blue) and those who were born in any other state (in red). Both series are divided by the population of these two regions. The gray line plots the difference in employees per capita between the two regions.