NBER WORKING PAPER SERIES

# STATE CAPACITY AS AN ORGANIZATIONAL PROBLEM. EVIDENCE FROM THE GROWTH OF THE U.S. STATE OVER 100 YEARS

Nicola Mastrorocco
Edoardo Teso

Working Paper 31591
http://www.nber.org/papers/w31591

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2023, Revised January 2024

We are grateful to Tianyi Wang for sharing data on the electric telegraph. For comments we are indebted to Charles Angelucci, Oriana Bandiera, Robert Gibbons, Simon Hix, Massimo Morelli, Elias Papaioannou, Torsten Persson, Martin Rotemberg, Shanker Satyanath, Jim Snyder, Guido Tabellini, Michael Ting, our discussants Erika Deserranno, 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. Siddhant Agarwal, Michael Giordano, Jimin Han, Benjamin Lualdi, and Navika Mehta provided outstanding research assistance. All errors are solely our responsibility. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

State Capacity as an Organizational Problem. Evidence from the Growth of the U.S. State Over 100 Years

Nicola Mastrorocco and Edoardo Teso

NBER Working Paper No. 31591
August 2023
JEL No. D73,M51,N41,P0

## ABSTRACT

We study how the organization of the state evolves over the process of development of a nation, using a new dataset on the internal organization of the U.S. federal bureaucracy over 1817- 1905. First, we show a series of novel facts, describing how the size of the state, its presence across the territory, and its key organizational features evolved over the nineteenth century. Second, exploiting the staggered expansion of the railroad and telegraph networks across space, we show that the ability of politicians to monitor state agents throughout the territory is an important driver of these facts: locations with lower transportation and communication costs with Washington DC have more state presence, are delegated more decision power, and have lower employee turnover. The results suggest that high monitoring costs are associated with small, personalistic state organizations based on networks of trust; technological innovations lowering monitoring costs facilitate the emergence of modern bureaucratic states.

Nicola Mastrorocco
University of Bologna
Department of Economics
Piazza Scaravilli 2, 40126
Bologna Italy
nicola.mastrorocco@unibo.it

Edoardo Teso
Kellogg School of Management
Northwestern University
2211 Campus Drive
Evanston, IL 60208
and NBER
edoardo.teso@kellogg.northwestern.edu

A data appendix is available at
http://www.nber.org/data-appendix/w31591
A Supplementary Data Appendix is available at
https://static1.squarespace.com/static/58210c07197aea631e39422b/t/64d5e9b4b1a949729471c37b/1691740640960/Supplementary_Appendix.pdf

# 1 Introduction

Economists consider state capacity an engine of economic growth. Effective states – centralized organizations with the ability to raise revenue and provide public goods through a vast territory – are only a recent historical phenomenon, and are still lacking in several developing countries (Dincecco and Katz (2016)). Most of this literature has focused on the analysis of rulers' incentives to set up a state apparatus.[1] Less attention has been devoted to the natural next step in the process of establishment of state capacity: once rulers have an *incentive* to establish a state apparatus, how do they concretely *organize* it to effectively perform its functions?

Building on a long tradition in sociology (Weber (1978); Kiser and Schneider (1994)), in this paper we stress the importance of agency problems in influencing the growth and evolution of modern state organizations. At the hearth of this theory is the observation that the principal-agent problems that characterize all organizations – how to monitor the behavior of agents whose incentives are not perfectly aligned with those of the principal – are particularly severe for states, because of the need to monitor officials over vast territories. How does a government solve this organizational problem?

In this paper, we turn to the development of the U.S. federal state apparatus over the nineteenth century to explore this question. We show that the ability of a state apparatus to grow, and the way in which it is organized at different stages of development, are significantly affected by the development of technological innovations that ameliorate the monitoring problems faced by the government when managing the state apparatus.

Our study relies on a large data collection effort that allows us to study the evolution of the U.S. federal bureaucracy over most of the nineteenth century. We construct a new micro-database combining newly digitized federal employees' personnel records and hand-collected information on the internal organization of the bureaucratic state for the period 1817-1905. Leveraging this unique data, the first part of the paper shows a series of novel facts on the growth and organizational development of the U.S. federal state over the nineteenth century. In the second part of the paper, we hypothesize that innovations in the government's monitoring capacity, which decreased agency problems between the government and its agents throughout the territory, are an important driver of these empirical facts. We provide causal evidence in support of our hypothesis by exploiting the expansion of the railroad and the telegraph networks – which increased the government's monitoring capacity by decreasing communication and transportation costs.[2]

[1] Building on Tilly (1975)'s argument that "war made states," the threat of external conflict has been considered a relevant driver of investment in state capacity (Besley and Persson (2008, 2010)).

[2] Importantly, as we discuss at length, many alternative mechanisms are likely to be important drivers

In order to build our dataset, we start by digitizing personnel records of the U.S. federal bureaucracy between 1817 and 1905. We digitize every volume of the Official Register of the United States, a biennial government publication that listed the names of all the employees of the federal government, together with their occupation, salary, department and office of employment, location of employment, and place of birth.[3] We link employees over time, in order to track their careers in the federal bureaucracy. We further complement this data in several ways. First, we re-construct the hierarchy of this organization throughout our sample period. Second, we categorize each job in the data into homogeneous occupational categories, dividing employees into homogeneous layers based on the type of work that they perform. Third, we geo-code each place of employment. This unique dataset allows us to observe the internal organization of a state over an unusually long time-span, and during a period of intense technological and economic development of a nation.

The first contribution of our paper is to show a series of novel descriptive facts on the growth and organizational development of the U.S. federal state over the nineteenth century. First, we show that the federal state apparatus grew very modestly throughout the first part of the nineteenth century, and experienced a very rapid growth thereafter. We document that the ability of the state to reach more locations throughout the territory was an important driver of this growth, and that this ability was limited in the first part of the nineteenth century. Interestingly, we show that the presence of federal bureaucrats does not substitute for state capacity by other levels of government (Wallis, 2000).

Second, we show that growth in manufacturing activity was positively correlated with the presence of the state in a location, consistent with greater incentives to increase state capacity in response to economic growth. However, in the first half of the nineteenth century, this link between economic growth and state presence is significantly weaker for locations that are more distant from the headquarter of the organization (Washington DC); distance does not play this moderating role in the second half of the nineteenth century.

Third, we show that the organization of the federal state changed significantly over the nineteenth century. Until the 1850s, the organization was characterized by high employee turnover (especially in years of turnover in the party of the President), by a tight link between a worker's career and her supervisor's career, and by very limited delegation of managerial power outside of the headquarter. These patterns were increasingly less significant in the second half of the century: turnover rates decreased, the career of a worker was less correlated

of the facts that we document. Our goal is to show that changes in the government's monitoring capacity are a significant driver of these facts, without denying the importance of alternative drivers of the evolution of the U.S. federal state over the nineteenth century.

[3]For reasons explained in Section 2, we digitize information on employees throughout all departments with the exception of the Postal Office.

with that of her supervisor, and locations outside of DC were delegated significantly more managerial power.

While the descriptive facts that we present are obviously driven by many historical developments, we argue that a decrease in the monitoring costs faced by the principal over the nineteenth century significantly affected the evolution of this organization. In the early years of the American federal state, Presidents and their cabinet (the principals) had low monitoring capacity in their relationship with federal employees (the agents), given the presence of high communication and transportation costs. In light of these constraints, a second best solution was to rely on the employment of trusted individuals: the optimal organizational form was of a *personal* nature, with relationships of trust between principals and employees, or between supervisors and their immediate subordinates, replacing effective monitoring of performance (Crenson, 1975). Given the limited supply of individuals who could be trusted, this organization faced difficulties in growing and in delegating managerial power away from the headquarter, especially in places located further away from DC. In addition, it was characterized by high turnover rates, as new principals or new supervisors needed to replace old employees with new, trusted ones. As technological innovations lowered communication and transportation costs, principals' greater ability to control agents allowed the transition to a larger *Weberian* organization. The lower need to rely on trust to ensure performance allowed faster organizational growth, made it easier to delegate decision power away from the headquarter, and decreased reliance on employee turnover.

The second contribution of our paper is to provide empirical evidence in support of this hypothesis. We show that changes in monitoring capacity are an important driver of the evolution of the U.S. federal state that we document.[4] To do this, we exploit the introduction and expansion across space of the two most prominent technological innovations that increased a principal's ability to control agents across space in the nineteenth century's U.S. – the railroad and the telegraph networks.

Our main identification strategy leverages the expansion of the railroads network, which dramatically decreased communication and transportation costs over the nineteenth century. We use data from Donaldson and Hornbeck (2016) to calculate the travel time between Washington DC and each county for each decade between 1820 and 1900. We then investigate whether counties that experience larger decreases in travel time to DC experience greater changes in the presence and organizational features of the federal bureaucracy. Importantly, we take several steps to ensure that our estimates capture the effect of decreased monitoring costs faced by DC, rather than other factors that are correlated with both railroads construction and the development of the federal state in a location. First, since railroads

[4]In section 4.4 we discuss additional interpretations for the facts that we document.

construction in a county is likely correlated with local economic growth, we directly control for local railroad construction in a county. By doing so, our estimates of the impact of travel distance to DC are identified from more-distant changes in the railroad network, which are thus arguably uncorrelated with local economic shocks in the county. Second, since more distant changes in the railroad network are associated with changes in market access (the ability of a county to trade with other counties), which in turn affected county's agricultural land values (Donaldson and Hornbeck, 2016) and manufacturing activity (Hornbeck and Rotemberg, 2021), we directly 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, (i) have similar local railroad construction in their territory, (ii) have similar changes in their market access, but (iii) experienced different changes in their travel time to DC *specifically*.

Consistent with our interpretation, we show that decreasing transportation and communication costs between D.C. and a county significantly increased the presence of the federal state in the county. We also show that, conditional on our set of controls, travel times to a set of other important economic centers are not associated with a larger presence of the federal state: these placebo tests lend further support to our interpretation that travel time to DC is associated with state presence through a monitoring mechanism. We also find that, consistent with our hypothesis, the development of the railroad network also affected the way in which the federal state was organized in a location: conditional on the federal state being present in the county, a lower travel time to DC increased the probability of observing delegation of managerial power to the county, and decreased turnover among the county's workforce. Importantly, we do not see differential pre-trends in the development of the federal state associated with changes in travel time between a county and DC. We also show that 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.

We obtain similar results when we exploit the expansion of the telegraph network between 1845 and 1852 (Wang, 2020): counties that have more telegraph connections to DC have a larger presence of the state, more delegation of managerial power, and lower employee turnover.[5] Differently from our empirical strategy focusing on the railroads, we have limited

[5]Given the fast expansion of the telegraph network, after 1852 all major U.S. towns and cities had a telegraph connection, limiting the time period for which we can rely on meaningful variation in connection to DC.

ability to address the possible endogeneity of the timing of the expansion of the telegraph network. However, we view the fact that we obtain similar results when focusing on a county's communication costs to DC via the telegraph as further corroborating our interpretation linking monitoring capacity and state development.

Finally, we provide indirect evidence pointing to the role of increased monitoring capacity substituting for trust relationships between the headquarter and the federal agents. We show that, after the Civil War, there was a sizable decline in the share of Southern-born federal bureaucrats, consistent with a lower level of trust towards individual from former confederate states. However, a better connection between a location and DC reduces the North-South employment gap: counties that become better connected to DC thanks to the expansion of the railroad network experience an increase in the share of Southern-born federal employees. In other words, when information and monitoring costs were lower, less trustworthy individuals were relatively more likely to be employed.

***Related Literature.***

Our findings contribute to a number of literatures. First, we speak to a growing literature on state formation and the development of state capacity, dating back to Zophy (1975), Tilly (1990), and Bonney (1999). Recent contributions document the relationship between state capacity and economic development (Besley and Persson, 2011, 2013; Dincecco and Katz, 2016), as well as the impact of administrative reforms expanding the reach of the state (Chambru et al., 2022; Chiovelli et al., 2023). An influential strand of this literature has emphasized the role of the threat of war in providing incentives to set up a centralized state apparatus (Besley and Persson (2008, 2010); Gennaioli and Voth (2015); Cantoni et al. (2019); Becker et al. (2022)). Other papers study state formation as a result of citizens' need to solve collective actions problems (Allen et al. (2020)) or of rulers' desire for extraction (Scott (2017); Mayshar et al. (2022); Allen (1997); Schönholzer (2017); Mayoral and Olsson (2019)). While a common denominator among these previous studies is their focus on rulers' incentive to set up a state apparatus, our paper studies the natural next step in this process: once these incentives are in place, how does a government organize its state apparatus to concretely perform its functions?

In emphasizing how the evolution of the state organization depends on developments in technologies of control, our paper resonates with theories linking state centralization to fiscal legibility, namely the ability of rulers to obtain information about the population and the state of the economy (Scott, 1998; Lee and Zhang, 2017; Mayshar et al., 2017; Garfias and Sellars, 2021).[6]

[6]In addition, a number of studies have linked the development of modern tax systems to the availability of hard (Dzansi et al., 2022) or soft (Jensen, 2022) information about taxpayers.

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 solve principal-agent problems within these organizations. Our paper underlines how these principal-agent relationships are crucial not only to our understanding of the functioning of bureaucracies at a given point in time, but also to explain their growth and organizational evolution over the process of development. In a recent theoretical contribution, Snowberg and Ting (2019) model a bureaucracy as a knowledge hierarchy, abstracting from agency problems within the organization. Our paper offers a complementary view by stressing the importance of agency problems in shaping the evolution of state organizations.

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, and the effects of transitioning to a merit-based Weberian 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 at different stages of development, characterized by different levels of government's monitoring capacity.[7]

Finally, 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).[8] Our results are in line with Chen (2017)'s insight that better communication technology increases delegation and organizational growth in a monitoring-based hierarchy model. They are also consistent with versions of cognitive models of the the firm (Garicano, 2000) with agency problems between the principal and lower level managers (Bloom et al., 2012).[9] Kelley et al. (2021) shows that technologies that improve owners' monitoring ability lead them to expand the size of their firm, and Giroud (2013) shows that

[7] As we discuss in section 4.4, this is not inconsistent with an increase in efficiency after the introduction of objective and meritocratic selection procedures, which might accelerate the transition to a less personalistic organization.

[8] 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.

[9] Bloom et al. (2012) show evidence consistent with a positive relationship between trust and the willingness to delegate decision power from the firm's headquarter to its subsidiaries, which in turn leads to larger firm size.

decreasing travel time between headquarters and plants increases plant-level investment, by facilitating monitoring and access to information. Our work highlights that similar mechanisms are also relevant to understand the process of development of state capacity.[10]

# 2 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 state between 1817 and 1905. In this section, we describe our data collection effort and the sources of the data.[11]

## 2.1 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.[12] We digitized all issues of the Registers between 1817 and 1905. 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. We have digitised a total of 15,801 pages. 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.

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.[13] Finally, we drop employees in navy yards and in the engineer department

[10] In exploiting the introduction of the railroads and of the electric telegraph as shocks to the government's monitoring capacity, we also contribute to the rich literature on the expansion of the railroads (Fogel (1965); Nerlove (1966); Atack et al. (2010); Atack and Margo (2011); Donaldson and Hornbeck (2016); Hornbeck and Rotemberg (2021)) and of the telegraph (Field (1992); Wang (2020); García-Jimeno et al. (2022)).

[11] We describe the data on the railroads and on the telegraph network in Section 5. The interested reader can find full details on the data construction in the Data Construction Appendix.

[12] The Registers were initially compiled and published by the Department of State, and since 1861 by the Department of Interior.

[13] 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 often report only the initials of the first names.

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.[14] Our final panel includes a total of 304,410 unique employees, and 810,942 employee-year observations.

This data source allows us to observe a rich set of characteristics of all the individuals employed by the Federal government.[15] 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.[16] 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 (see section 2.3).

We link employees over time, in order to track their careers in the federal bureaucracy.[17]

## 2.2 Geo-location of places of employment

The Register contains information on each worker's location of employment. Online Appendix Figure A2 shows an extract from the 1875 Register, highlighting the locations under the "where employed" column.

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.[18]

Of the 810,942 observations in our dataset, 800,538 (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 768,041 observations that are located within the United States, we can recover information on the county of employment for 95% of ob-

[14] None of the central results of the paper are affected by this choice.

[15] From 1817 until 1877, the Registers included all individuals employed as of September 30, while since 1879 they included all individuals employed as of June 30.

[16] Employees could be paid either a fixed annual amount, or a variable amount depending on the days, weeks, or months of employment throughout the year. We calculate each employee's total annual compensation by multiplying daily, weekly, or monthly pay rates and assuming that the individual was employed for the entire year. In relatively rare cases, the compensation is expressed as a variable amount depending on a number of tasks performed (e.g., "per inspection" or "per drill hole").

[17] 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.

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

servations. 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. In total, the data include 9,651 unique geo-located places of employment.

## 2.3 Construction of the hierarchical structure

To construct a consistent hierarchy of the US Federal bureaucracy across time we exploit the fact that, from 1817 to 1905, the Official Register was arranged in a tabular format. This layout provided 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 state, we construct a consistent hierarchy of the organization by following the evolution of its units over time.[19] 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 offices (or bureaus) within each department. Some examples of offices 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. The third layer is composed of the divisions within each office. We use the generic term division to refer to the different sub-units in which offices can be divided. For example, the Customs Office is composed of several customs districts; the General Land Office is composed of several surveyor districts. The fourth layer is composed of the different local offices within each division. 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.[20]

The reconstruction of this hierarchy allows us to recover the chain of command in the organization, assigning all workers to their direct supervisor. The direct supervisor can either be present in the specific location, in case we observe a worker employed in a supervisory or managerial capacity in the location, or can be someone at a higher organizational layer (at

[19]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)

[20]The hierarchy is not complete. That is, in any given year, we can find departments that are not organized in offices, offices that are not organized in divisions, or divisions that are not organized in local offices.

the division, or at the office/bureau level).[21]

### 2.4 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 11,930 unique occupation codes.

We group occupations into five categories based on the type of task performed.[22] The first category includes the top managers of the organization: the heads of department, deputy heads of department, and heads of offices. The second category includes workers employed in a supervisory or managerial capacity (for example: chief of divisions, chief clerks, chief of regional offices). The third category includes clerical occupations (for example: clerks, copyists). The fourth category includes professional occupations (for example: engineers, doctors). Finally, the fifth category includes jobs requiring a relatively low level of skills (for example: laborer, messengers). Of the 11,930 occupation codes in the data, 2.1% are categorized as top managers, 11.8% as supervisors/managers, 26.6% as clerical workers, 34.5% as professionals, and 25% as low skills workers.

These five occupational categories can be arranged in a hierarchy, with top managers at its top, followed by managers, by clerical and professional occupations, and finally by low skills workers. Importantly, this hierarchy of jobs maps into the average annual pay that we observe in the data for each of these categories: on average, top executives earn $3,709, managers earn $2,230, workers in clerical positions earn $1,179, those in professional occupations earn $974, and those in lower skills positions earn $524.

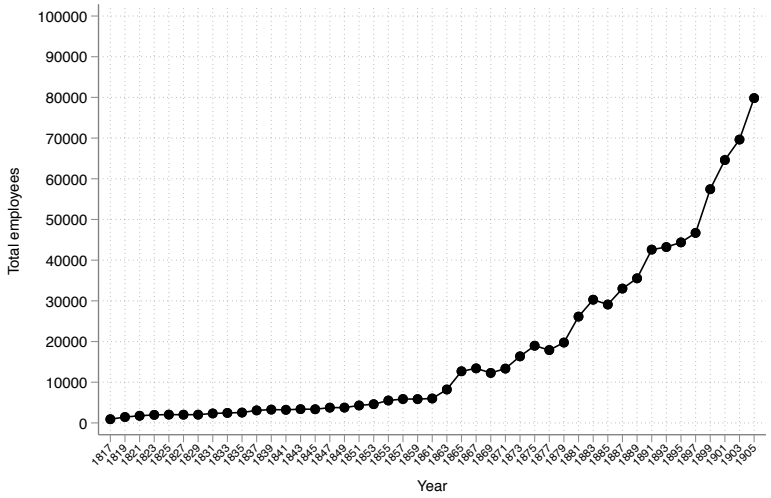
## 3 Descriptive Facts on the Development of the U.S. State

In this section, we show a series of novel descriptive facts on the growth and organizational development of the U.S. federal state. We divide these facts into three broad groups. First, we document the extent, timing, and sources of the growth in the presence of the federal state over the nineteenth century. Second, we document where the federal state was more likely to be present. Third, we document the evolution in the way in which the federal state was organized.

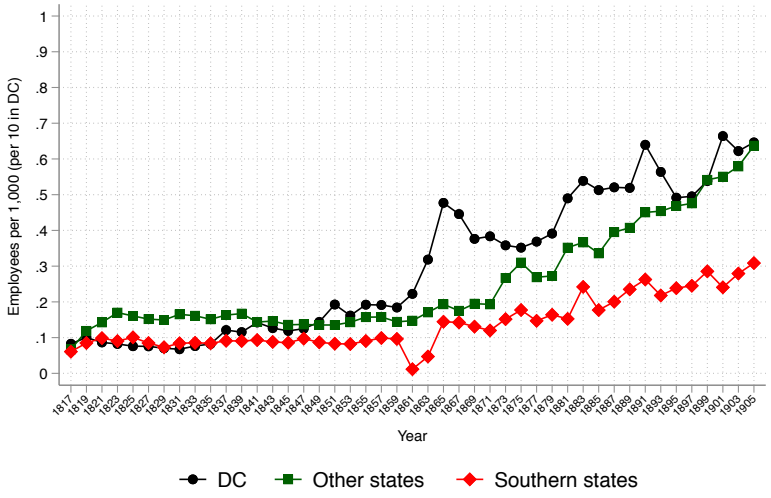
[21] See the next subsection for a description of our grouping of jobs into occupational categories.

[22] A similar occupational classification is employed in *The Executive Civil Service of the United States of Commerce and Labor* (Census, 1904). We heavily rely on this publication in our manual coding of occupations.

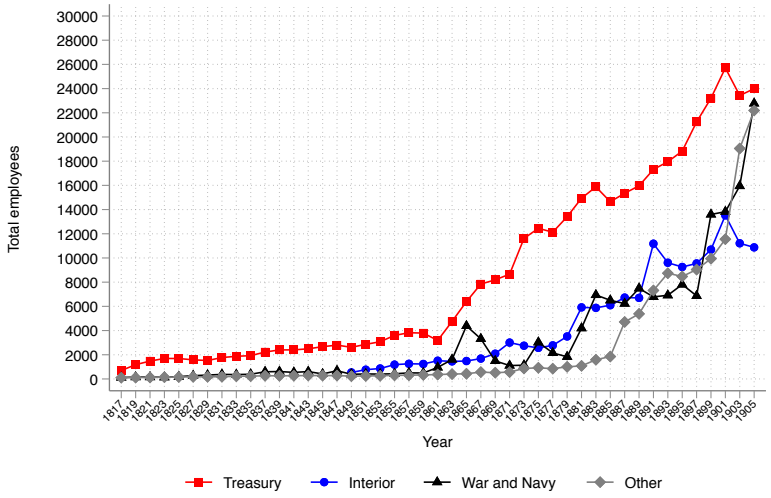
## Figure 1: Growth of U.S. Federal Bureaucracy, 1817-1905



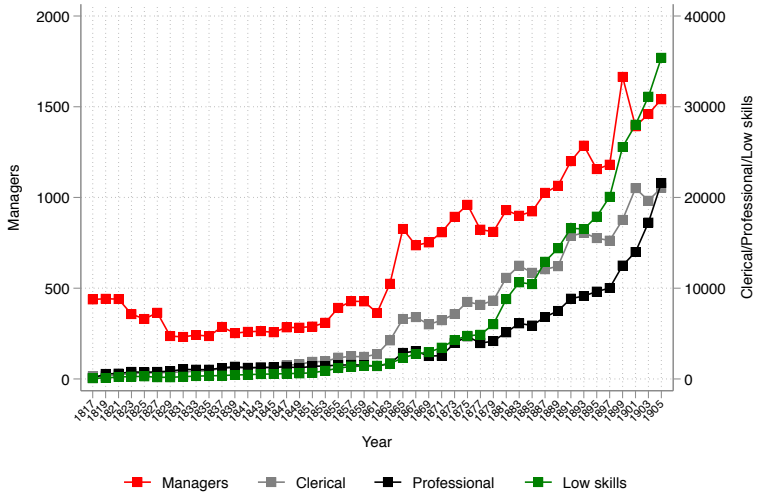

(a) Total number of employees




(b) Employees in DC vs South vs Outside South




(c) Number of employees by department




(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 DC per 10 inhabitants (in black) in Southern states per 1,000 inhabitants (in red) and on other states per 1,000 inhabitants (in green) (Panel B); number of employees by main department (Panel C); number of employees by occupational category (Panel D).

## 3.1 Timing and sources of the growth in state presence

### 3.1.1 Timing of growth

Figure 1, Panel A, plots the total number of federal employees in each year between 1817 and 1905. The federal state grew very slowly in size in the first part of the nineteenth century. Starting from the early 1860s, the size of the state started to increase at a rapid pace.[23] These patterns are even clearer when we normalize the size of the federal workforce by the U.S. population (Online Appendix Figure A4) Panel B of Figure 1 shows that most of the growth in the size of the state in the 1860s is driven by a sudden and significant increase in the number of employees in DC. The number of employees outside of DC (henceforth, "in the field") grows more moderately in the 1860s and starts to grow faster in the 1870s.

What were the main goals of the federal state over the nineteenth century? Figure 1, Panel C, 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.[24] By the 1880s, a large number of additional, smaller departments started to employ a large number of employees. Online Appendix Table A1 shows the main tasks of the federal state outside DC, 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.

Figure 1, Panel D, provides a breakdown of employees by occupational category. The number of individuals in managerial positions (left axis) did not significantly increase until the 1850s – something that we will further explore in section 3.3.3. In contrast, the number of employees employed in clerical, professional, and low skills positions (right axis) slowly but steadily increased between 1817 and the end of the 1850s. By the end of the sample period, the U.S. federal bureaucracy exhibits a pyramidal structure, with the bottom of the hierarchy (low skills employees) comprising the largest group, followed by an intermediate

[23] The Official Register of 1817 lists the names of 917 employees; by 1859, this number increased to 5,856, with an average of 235 added jobs per year. The federal bureaucracy added an average of 1,286 jobs per year from 1861 to 1869, 1,493 jobs per year from 1871 to 1879, 3,157 jobs per year between 1881 and 1889, and 5,537 jobs per year between 1891 and 1905. In the last year covered by our data, the federal state employs 79,835 individuals.

[24] In this figure we combine the War and Navy departments, but they were distinct departments throughout the entire period. Besides being responsible for the defense of the country, the War and Navy departments were also tasked with the building of critical infrastructure. The Interior department, established in 1849, was responsible for a variety of functions broadly related to domestic affairs, including the disposition of public lands, pensions, Indian affairs, and the granting of patents.

layer of clerical and professional employees, and by a smaller layer of managers.

Interestingly, as we show in Online Appendix Figure A5, the presence of federal bureaucrats is, if anything, positively correlated with the presence of employees of local and state governments in a county, suggesting that state capacity at the federal level does not substitute for state capacity by more local level of governments.[25]

### 3.1.2 Drivers of growth

There are three possible sources of growth in a state organization. First, a state can grow because it starts to perform a higher number of functions (the "functions" component of state growth). Second, a state can grow because it increases the number of locations across the territory in which it is present (the "geographic expansion" component). Third, a state can grow by increasing the intensity of its presence, i.e. by increasing the number of employees performing a given function in a given location (the "intensity" component).

Figure 2, Panel A, shows the growth in the number of offices (or bureaus) of the U.S. federal state over the nineteenth century, which we consider as a proxy for a specific function performed by the state. Their number steadily increased in the first half of the century, from 25 in 1817 to 46 in 1859. The rate of growth was higher in the second half of the century, when the organization added an average of 3.7 new functions every two years, reaching a total of 132 separate offices in 1905.

In contrast, as shown in Figure 2, Panel B, the state did not start to expand its geographical presence until the 1860s. We plot the share of U.S. counties where we observe a presence of the federal state (i.e. with at least one individual employed within the county borders).[26] This share hovered around 15 percent between 1817 and 1859, and does not display any increasing trend over this period. In the second half of the nineteenth century, the state begins to increase its presence across the territory: it is present in 24% of counties by 1871, in 38% of counties by 1881, and in 61% of counties by 1905.[27] Online Appendix Figure A8 shows the presence of the state across space at four points in time. While by 1859 the frontier had moved West, the portion of the territory with state presence had remained constant, while by 1881 we observe a marked increase in state presence across the territory.

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

[26] For each year, the number of counties with potential state presence (i.e., the denominator of this share) is the number of counties in States and Territories that were included in the most recent census. In this way, we account for the enormous territorial expansion of the U.S. over the nineteenth century.

[27] In Online Appendix Figures A6 and A7 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 (which shows that these patterns do not depend by the westward expansion of the country over the nineteenth century), or if we weight each county by the fraction of the U.S. population living in the county in a specific year.

In Figure 2, Panel C, we show how the average number of employees for each county-office pair, i.e. our measure of the intensity of state presence, changed over time. We observe a steady growth in this measure during the sample period, from 1.9 average employees in 1817 to 6.7 in 1859 and to 14.5 in 1905.

In Figure 2, Panel D, we provide a formal decomposition of state growth between these three sources. We compute counterfactual growths between 1817 and 1859, and between 1859 to 1905, had each of the three components remained constant at its level at the beginning of the period.[28] The growth of the U.S. federal state between 1817 and 1859 was entirely driven by the functions component and by the intensity component, which were responsible for about 40 percent and 60 percent of the growth, respectively. Consistent with the trends in Panel B of the figure, the geographic expansion component did not lead to any state growth in the 1817-1859 period. In contrast, after 1859, the geographic expansion component accounted for about 29 percent of the growth of the state, with the intensity component accounting for 32 percent and the functions component for the remaining 39 percent.

We can summarize this first set of descriptive facts with the following:

**Descriptive fact. *1:*** *The U.S. federal state grew mainly since the 1860s, and started to expand to new locations:*

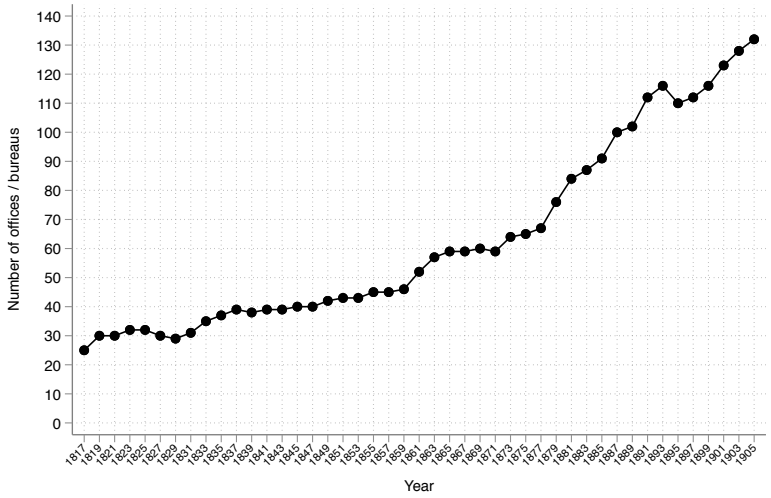
> *(1a) There was a slow growth in the size of the state before the 1860s, and significantly higher growth since the 1860s.*
>
> *(1b) An important driver of growth since the 1860s was the increased presence of the state in more locations across the territory. This driver of growth was not present before the 1860s.*

[28] Specifically, we define the total number of workers employed by the state in year $t$ as:

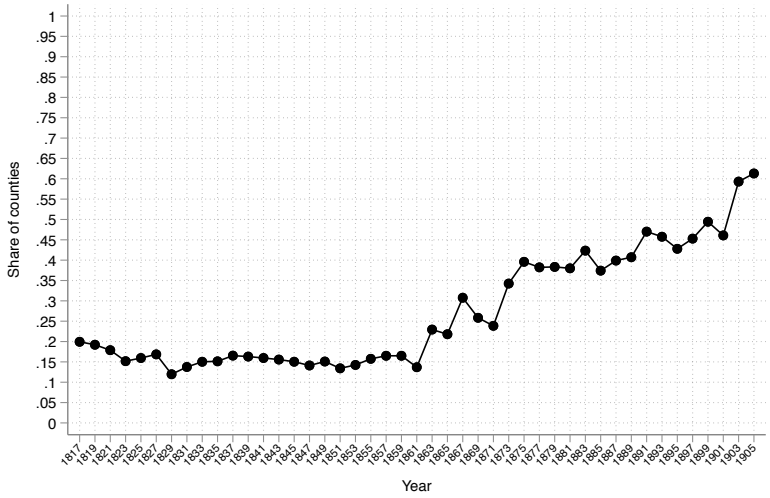
$$\text{Workers}_t = B_t \times \frac{1}{B_t} \sum_b L_{bt} \times \frac{1}{\sum_b L_{bt}} \sum_{blt} W_{blt} \tag{1}$$

where $B_t$ is the number of offices in year $t$, $L_{bt}$ is the number of counties where office $b$ is present in $t$, and $W_{blt}$ is the number of workers employed in office-county $bl$ in $t$. The three terms captures the function, geographic expansion, and intensity components, respectively. We compute each of the three terms for 1817, 1859, and 1905, their change from 1817 to 1859 and from 1859 to 1905, and counterfactual growths in $\text{Workers}_t$ had each of the three components remained constant at its level at the beginning of the period.

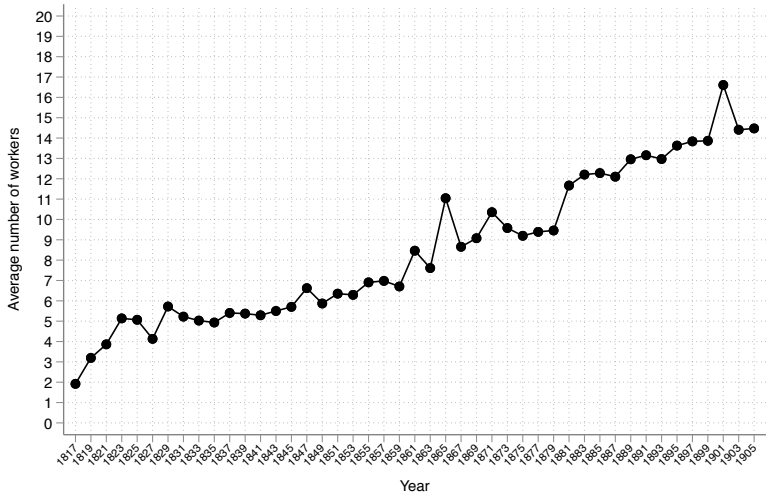
## Figure 2: Decomposing the Sources of Growth



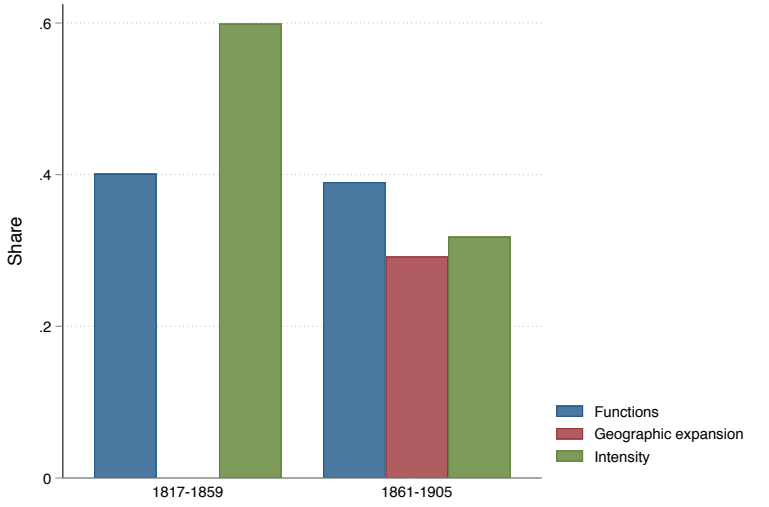

(a) Number of offices/bureaus




(b) Share of counties with state presence




(c) Average workers in county-office




(d) Decomposition

*Notes:* The figure shows the number of offices/bureaus over time (Panel A), the share of counties with state presence over time (Panel B), the average number of workers in county-offices over time (Panel C), the share of each component's contribution to state growth between 1817-1859 and 1861-1905, following equation 1 (Panel D).

## 3.2 Where was the state more likely to be present?

We first investigate whether economic growth is associated with greater state presence. To this end, we construct a panel at the county-year level. Our measure of economic growth is the logarithm of the share of a county's population that is employed in manufacturing. We rely on this measure since it is available in all decades throughout the entire 1820-1900 period, with the exception of 1830.[29]

The first column of Table 1, and Panel A of Online Appendix Figure A9, present results of a regression of an indicator for the presence of a federal employee in the county on the share of manufacturing employment, controlling for county and state-year fixed effects. A one standard deviation increase in the share of manufacturing employment is associated with a 1.3 percentage points increase in the probability of state presence (a 4 percent increase relative to the mean probability). The relationship between state presence and manufacturing growth is strong also when using an extensive margin measure, namely the logarithm of one plus the total number of federal employees in the county (column 2 of Table 1 and Panel B of Online Appendix Figure A9).[30]

These results are consistent with theories on the determinants of state creation that emphasize the link between state presence and incentives for extraction by the state, as counties with greater presence of manufacturing have greater potential to generate revenue. Additionally, to the extent that counties with greater manufacturing intensity have higher returns from public goods, our results are also consistent with theories of state formation emphasizing citizens' demand for government.[31]

Next, we investigate how a location's distance from the headquarter of the organization limits the ability of the state to establish its presence in response to growth in manufacturing. Specifically, we interact the share of manufacturing employment with a variable measuring the distance (in thousands miles) between a county's centroid and DC. We separately estimate this specification for the 1817-1859 and the 1861-1905 periods, since physical

[29]Since the variable takes value zero for about 8 percent of the observations, we use the logarithm of one plus the manufacturing employment share. The variable is available at the decade-county level, thus each county-year $ct$ is assigned the value of county $c$'s manufacturing employment share at the beginning of $t$'s decade. In essence, we ask whether a county's level of manufacturing development at the beginning of a decade is associated with a greater presence of the state in the following ten years.

[30]Throughout this section, we do not consider DC in our analysis, 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.

[31]The positive association between state presence and manufacturing employment exists both when we focus only on "extractive" bureaus, namely customs and internal revenue and when we focus only on the other, non-extractive bureaus (see Online Appendix Table A2).

**Table 1: Manufacturing Growth, Distance from DC, and State Presence**

| | (1) State presence | (2) Log tot. employees | (3) State presence | (4) State presence | (5) Log tot. employees | (6) Log tot. employees |
|---|---|---|---|---|---|---|
| Log Share Manu. Emp. | 0.340*** | 2.386*** | 1.265*** | -0.260 | 1.592*** | 2.530*** |
| | (0.109) | (0.349) | (0.432) | (0.227) | (0.560) | (0.812) |
| Log Share Manu. Emp. X Distance | | | -1.352*** | 0.140 | -1.758*** | -0.117 |
| | | | (0.455) | (0.117) | (0.626) | (0.404) |
| Observations | 89,870 | 89,870 | 28,985 | 60,885 | 28,985 | 60,885 |
| Sample | All | All | 1817-1859 | 1861-1905 | 1817-1859 | 1861-1905 |
| Std dev Dep. Var. | 0.465 | 0.988 | 0.359 | 0.488 | 0.629 | 1.092 |
| Std dev Log Share Manu. Emp. | 0.037 | 0.037 | 0.038 | 0.036 | 0.038 | 0.036 |
| Std dev Distance | - | - | 0.664 | 0.864 | 0.664 | 0.864 |

*Notes:* The unit of observation is a county-year. State presence takes value one if the federal state is present in the county. Log tot. employees is the logarithm of one plus the total number of employees employed in the county. Log(Share Manu. Emp.) is the logarithm of the share of a county's population that is employed in manufacturing. Distance is the distance (in thousands miles) between the county's centroid and DC. All specifications control for county fixed effects and state-year fixed effects. The sample in columns 1-2 includes all odd years between 1821-1905, with the exception of 1831-1839, while it includes all odd years between 1821-1859, with the exception of 1831-1839 in columns 3, 5, and all odd years between 1861-1905 in columns 4, 6. Standard errors in parentheses, clustered at the county-level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

distance represented a more significant impediment in the first part of the nineteenth century. Columns 3-6 of Table 1 present the results. In the 1817-1859 period, the association between a county's manufacturing employment share and state presence (column 3) or total employees (column 5) is significantly weakened by an increase in distance between the county and DC. This is not true for the period 1861-1905 (columns 4 and 6).

We can summarize this second set of descriptive facts with the following:

**Descriptive fact.** ***2:*** *The state grew more in more prosperous locations. Distance from the headquarter (DC) offsets this relationship, but only in the first half of the nineteenth century.*

## 3.3 How was the state organized at different stages of development?

### 3.3.1 Employee Turnover

Our dataset can be used to document how the organization of the state changed over the nineteenth century.

The first dimension that we analyze is the degree of employee turnover in the organization. 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 $t$.[32]

Figure 3, 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. 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.[33]

In Online Appendix Figure A10 we separately plot turnover rates in DC and outside of DC (i.e."in the field"). Turnover rates are consistently lower in DC than in the field. This is not due to the different nature of jobs and bureaus between DC and the field: when we regress an indicator equal to one if the employee leaves the organization on an indicator for DC, including a set of year-bureau-position type fixed effects, being employed in DC is associated with a 40 percent reduction in turnover probability (Online Appendix Table A3).

### 3.3.2 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 ask whether the turnover of a supervisor leads also her direct subordinates to leave. We assign employees in each year and organizational unit (i.e., a specific local office of a division within a bureau) to their direct supervisor (or supervisors), as described in section 2.3.

We employ our panel at the employee-year level, and we estimate the following model:

$$\text{Turnover}_{it} = \alpha_t + \gamma_{b(it)} + \delta_{l(it)} + \sum_{\tau} \beta_\tau \text{Share Supervisor Turnover}_{it} + \epsilon_{it} \qquad (2)$$

[32]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 increased over the nineteenth century (Hacker, 2010). Thus, the rate of employees' exit because of death can be assumed roughly constant over our sample period.

[33]Online Appendix Figure A11 plots turnover by occupational category. We observe similar temporal trends for all the categories, with a steady increase in turnover until the end of the 1850s and a declining trend thereafter. In the first half of the nineteenth century, spikes in turnover were significantly higher for managerial positions and professional positions, followed by low skills positions, and by clerical occupations.

The variable Turnover$_{it}$ is an indicator equal to one if employee $i$ leaves her organizational unit in year $t$. We are interested in whether an employee's turnover is related to the turnover of her most immediate supervisors, Share Supervisor Turnover$_{it}$, namely the share of $i$'s supervisors who leave the organizational unit in year $t$.[34] We include year fixed effects, $\alpha_t$, which absorb any time-level shock affecting organizational turnover (e.g. presidential transitions). We further include bureau fixed effects, $\gamma_{b(it)}$, and location fixed effects, $\delta_{l(it)}$, in order to account for the tendency of some bureaus and some locations, respectively, to exhibit high personnel turnover. We allow the relationship between 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.

Figure 3, panel B, presents the standardized effects, namely the coefficient $\beta$ 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 more pronounced before 1861.

### 3.3.3 Delegation of managerial power

The third organizational dimension that we explore is the extent to which managerial power was delegated outside of DC. In Figure 3, Panel C, we plot the number of employees in managerial positions located away from DC, for each year in the 1817-1859 period. There is no growth in the number of field managers between 1817 and 1859, with the number of field managers actually decreasing during the 1820s, and staying constant until the mid-1850s. Their number started growing in the 1860s, and experienced a sustained growth over the second half of the nineteenth century: by 1905, the number of field managers has approximately tripled relative to the 1850s.

We also show that this increase went hand in hand with the likelihood that a local office outside of DC had an additional managerial layer between them and the top managers in DC (i.e. we observe a worker employed in a managerial occupation either in the local office, or at the division level in the hierarchy). We estimate a regression at the local office - year

[34] An employee has a median of 3 supervisors.

level:

$$\text{Additional layer}_{ot} = \alpha_{l(o)} + \gamma_{b(o)} + \sum_{\tau} \beta_{\tau} \mathbb{1}[\text{year}_t \in \tau] + \delta_1 W_{ot} + \delta_2 W_{ot}^2 + \epsilon_{ot} \tag{3}$$

where Additional layer$_{ot}$ is an indicator equal to one if local office $o$ has an additional managerial layer in year $t$. We estimate how the probability of having an additional layer varied over the nineteenth century, by including three indicators for the same periods of equation 2 (between 1841 and 1859, between 1861 and 1881, and after 1881, with the years before 1841 as excluded category). We include location fixed effects ($\alpha_{l(o)}$) and bureau fixed effects ($\gamma_{b(o)}$) to account for specific characteristics of a location or of a bureau that might affect their organization. We additionally include a second order polynomial in the size of the workforce in the local office, in order to control for the fact that an average increase in office size might mechanically increase the probability of observing an additional managerial layer.

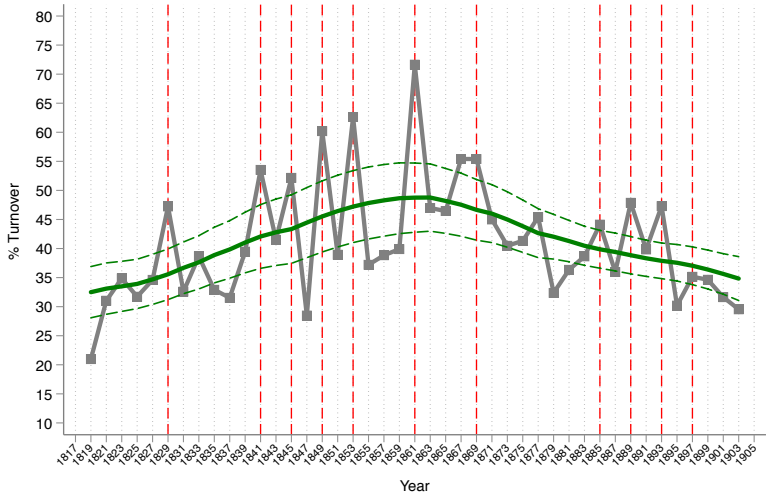
Figure 3, Panel D, presents the coefficients $\beta_{\tau}$ normalized by the mean of the dependent variable in the years before 1841. The likelihood that workers in a local office have an additional managerial layer between them and DC is similar between 1841 and 1859 relative to the pre-1841 period. This likelihood increases by 5 percent between 1861 and 1881 and by 6 percent post 1881.

We can summarize this third set of descriptive facts with the following:

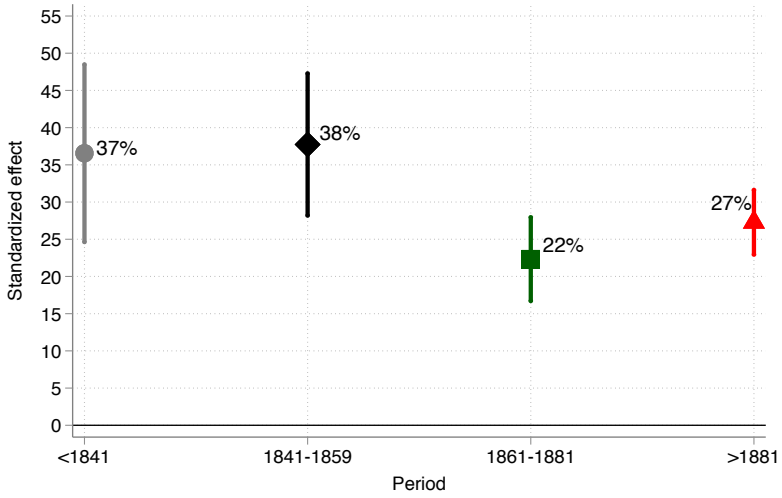
**Descriptive fact. *3:*** *The organization of the state apparatus started to change since the 1860s:*

> *(3a) In the period 1817-1850s, there was an increasing presence of employee turnover when the party of the President changed.*
>
> *(3b) In the period 1817-1850s, there was a tight link between workers' and their supervisors' careers.*
>
> *(3c) In the period 1817-1850s, there was no growth in delegation of power outside DC.*
>
> *(3d) Since the 1860s, we see a progressive change in these organizational features, with lower turnover, a less tight link between workers' and supervisors' careers, and more delegation of power outside DC.*

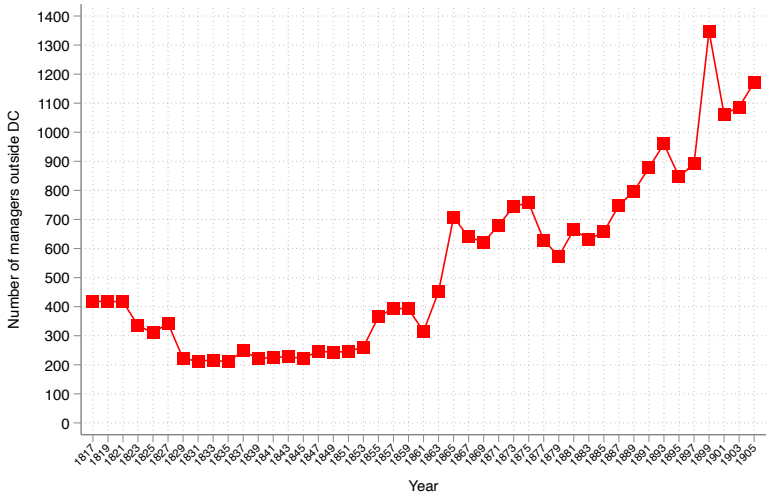
## Figure 3: Organizational Features of the U.S. Federal Bureaucracy



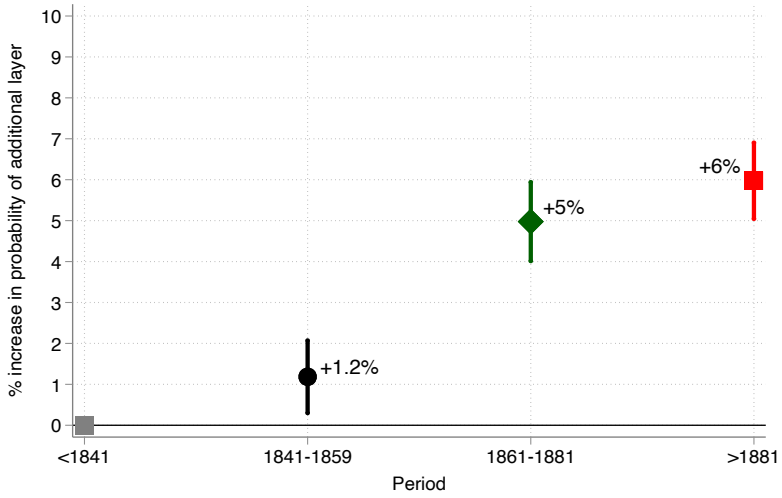

(a) Turnover




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




(c) Number of field managers




(d) Additional managerial layer in local offices

*Notes:* The figure shows aggregate turnover, i.e. the share of employees leaving the bureaucracy, over 1817-1905 (Panel A), the standardized coefficients on $\beta_\tau$ from equation 2, with 95 percent confidence intervals based on standard errors clustered at the organizational unit times year level (Panel B), the number of employees in managerial positions located away from DC, over 1817-1905 (Panel C), and the coefficients $\beta_\tau$ from equation 3, normalized by the mean of the dependent variable in the years before 1841 (with 95 percent confidence intervals). The red vertical lines in Panel A indicate years in which the party of the President changes. See section 3.3 for additional details.

# 4 Interpreting the Descriptive Facts

A long tradition in sociology has linked the emergence of modern bureaucracies to the development of technologies decreasing communication and transportation costs. This argument dates back to Max Weber, who underlined how *"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). Scholars have argued that a gradual decrease in communication and transportation costs, by increasing rulers' ability to monitor throughout the country, facilitated the transition from tax farming regimes, in which agents paid a fixed rent to the ruler for the right to levy taxes in a specific territory, to centralized bureaucracies employing tax collectors for a fixed wage (Kiser and Schneider, 1994; Kiser, 1994; White, 2004).

We hypothesize that a similar mechanism might have been important to facilitate the transition of the U.S. federal bureaucracy from a small *personal* organization to a larger *Weberian* organization.

Delegation from principals to lower-level agents is crucial to ensure the growth of an organization (Lucas Jr, 1978). However, delegating tasks goes hand in hand with agency problems: how can the principal (in this case, the politicians in power in DC and their cabinets) ensure that the agents (in this case, the individuals employed in the federal bureaucracy) will not follow their own personal interests at the expense of the interest of the principal? While these agency problems are present in any organization, they are particularly challenging for states, since the principal employs agents throughout a vast territory.

We argue that the development in technologies of control throughout the nineteenth century led to changes in the way through which the principals attempted to ameliorate these agency problems. When communication and transportation costs were high, and thus monitoring capacity was low, the presence of trust between principals and agents was used to substitute for effective monitoring. The development of technologies that increased monitoring capacity made it optimal to progressively adopt a modern bureaucratic organization of the state apparatus.

## 4.1 Delegation and monitoring problems in the U.S. federal bureaucracy

The increasing need to delegate tasks to subordinates as an organization seeks to expand in size was emphasized by Treasury Secretary Alexander Hamilton, who in 1778 wrote to Secretary of War McHenry: *"I observe you plug in a vast mass of details. I know from experience that it is impossible for any man, whatever his talents or diligence, to wade through*

*such a mass. 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)

However, politicians in DC encountered frequent challenges in supervising the behavior of the field employees. This was true both for workers involved in simple tasks, and for agents that were delegated significant decision powers. For instance, Land Office administrators *"strung out along the frontier [...] were relatively secure from the prying eyes of Washington bureaucrats,"* leading to frequent cases of fraud and corruption, or of *"plain indifference to public duties"* (Crenson, 1975, pp. 86-87).

Cases of corruption and lack of effort in the performance of duties were common among employees in the custom houses (Prince and Keller, 1989). Custom collectors, whose pay was partially a function of the amount of trade at their port, often undervalued imports, in order to attract ships to their port and secure higher fees for themselves. (White, 1954, p. 179).

## 4.2 Low monitoring capacity and personal state organization

In the early decades of the nineteenth century, high communication and transportation costs throughout the U.S. made monitoring of field employees difficult. Systems of supervision were sometimes used, but the large distances between DC and the various field offices made these tools insufficient to ensure adequate monitoring. For example, while the Commissioner of the Land Office had established a system of inspections of local offices, the inspectors visited each office only once a year, making it easy for local officials to conceal any wrongdoing in the performance of their duties (Crenson, 1975, pp. 92). Some officials, like the collectors of the customs, were incentivized to exert effort by having their compensation partially dependent on the value of the goods ascertained at their port. However, as discussed above, when coupled with a lack of adequate monitoring, this system was likely to introduce distorsions.

In this context, the presence of trust between principals and agents could ameliorate agency problems. The U.S. federal government focused mainly on the selection margin to ensure an adequate performance by federal bureaucrats. Political leaders frequently underlined the individual's fitness for office, moral character, and political opinions friendly to the administration as important requirements for selection (Fish, 1905).[35] The First Comptroller believed that *"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 indi-*

[35]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).

*vidual intrusted with the receipt and disbursement of the public funds"* (Senate Doc. 1 25th Congress, 1837).[36]

Personal networks were an essential tool in order to identify individuals who could be trusted to adequately and dutifully perform their tasks.[37] Defending his own choice for the position of collector of the customs of New Haven, Thomas Jefferson writes that *"From private sources it was learned that his understanding was sound, his integrity pure, his character unstained."* (Jefferson, 1854). Similarly, department leaders often relied on personal connections to identify possible candidates for appointment. For instance, Secretary of State Daniel Webster in 1851 asked a correspondent for *"the name of a man, the fittest, within your knowledge, to be Naval Officer. He must be a firm an energetic friend to the present Administration; not too old, all together trustworthy and enjoying public confidence"* (Webster, 1904). Members of Congress were often asked to identify trustworthy individuals from their districts (White, 1954, p. 116). In turn, personal relationships between office chiefs and their subordinates were also common and considered essential to ensure trust within each organizational unit: subordinates were tied to their chief *"by personal loyalty, friendship, and, not infrequently, kinship."* (Crenson, 1975, p. 72).

In sum, the U.S. federal bureaucracy during the first decades of the nineteenth century had the characteristics of a *personal* organization: it was based on the personal character of the individuals employed, and on relationships of trust between leaders and subordinates, while bureaucratic procedures for monitoring behavior were scarce and often ineffective (Crenson, 1975). The government's discretionary power over appointments and removals allowed political leaders to assign federal jobs to individuals who could be sufficiently trusted.

While a personal organization might be an efficient response to structural conditions that make monitoring difficult, is has two important drawbacks. First, frequent turnover of officials led to loss of experience in the bureaucracy.[38] Second, since the supply of trustworthy individuals that can be found through personal networks is limited, this placed constraints

[36] This echoes the theoretical framework in Bloom et al. (2012). They present an extension of the cognitive model of the firm in Garicano (2000), in which employees can take "wrong" actions in the performance of the tasks that are delegated to them. The presence of trust between principals and employees make the former more willing to delegate to the latter, allowing the firm to grow.

[37] 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).

[38] As William Coleman, the editor of the New York Evening Post, remarked in 1801: *"If every change of a chief magistrate is to produce a similar change of subordinate officers [...] their places are to be supplied by a new set of men who have every thing to learn [...] Government will be entirely deprived of all the benefits of experience, and the management of public offices, perpetually shifting from one tyro in office to another, will forever be kept in infancy and weakness"* (Coleman, 1801).

on the organization's growth potential. As the required level of trust is increasing in an agent's decision powers, delegation of managerial tasks away from the headquarter of the organization is particularly challenging.

This is consistent with the descriptive facts that characterized the federal bureaucracy in the first half of the nineteenth century. The importance of trust in filling bureaucratic positions led to high employee turnover (Fact 3a), as a new administration needed to fill positions with trusted bureaucrats. In addition, the need to maintain relationships of trust between supervisors and subordinates led to a tight link between their careers (Fact 3b). Since reliance on personal networks for staffing an organization naturally leads to a limited supply of trusted individuals, this limited the ability of the state to (i) grow in size (Fact 1a), (ii) expand its presence across the territory (Fact 1b), especially in more remote locations (Fact 2), and (iii) delegate managerial power to the periphery (Fact 3c).

## 4.3 Increase in monitoring capacity and transition to a Weberian bureaucracy

Over the course of the nineteenth century – and especially in the second half of the century, – the expansion of the railroad networks and of the electric telegraph increased political leaders' monitoring ability by decreasing the costs of communicating with, and, in the case of the railroads, traveling to, locations away from DC. In turn, this allowed the transition to a *Weberian* bureaucratic organization of the state apparatus, characterized by a fixed hierarchy of officials – rather than one with frequent turnover, – where each agent could be more effectively monitored to ensure adequate performance.

Importantly, as this organization is no longer limited to employing trusted individuals, it faces less challenges in expanding its size.

This is once again consistent with the descriptive facts that we showed. Over the second half of the nineteenth century, employee turnover and the link between workers' and their supervisors' careers decreased in importance, and delegation of managerial power outside of DC became more common (Fact 3d); the substitution of reliance on trust with effective monitoring as a way to ensure performance allowed the organization to grow (Fact 1a) and expand to new locations (Fact 1b).

## 4.4 Additional mechanisms

While we argue that innovations in the government's monitoring capacity, and the corresponding transition from trust to monitoring as a tool to ensure performance, are relevant drivers of the descriptive facts that we showed in section 3, this is obviously not the only

mechanism that is consistent with these facts. In this section, we discuss three additional, arguably important mechanisms.

### 4.4.1 Demand shocks

The large increase in the size of the federal bureaucracy in the second half of the nineteenth century is also consistent with higher incentives to invest in state capacity because of demand-side shocks. Two relevant shocks during this period were the American Civil War and the progressive industrialization of the country. As the principals responded to these shocks by increasing the size of the workforce, and its presence across the territory, the need to manage a larger state apparatus might have also facilitated the shift to a more modern, Weberian form of organization.

In particular, the American Civil War represents a potentially relevant driver of the development of the federal bureaucracy, as the federal government needed to invest in state capacity in order to repay the debt accumulated during the war years.[39] In Figure 1, we show that the beginning of the civil war in 1861 coincides with an important inflection point in the growth path of the U.S. federal bureaucracy, consistent with the incentive to invest in state capacity to repay the federal debt playing an important role. Importantly, all our results in the next section control for year fixed effects, in order to account for aggregate time-varying shocks in the federal government's incentive to increase the size of the federal bureaucracy.

### 4.4.2 Patronage as electoral tool

The monitoring mechanism on which we focus takes a benevolent view of the the principals in DC, who are interested in organizing the federal state in order to maximize its productivity. However, staffing of the federal bureaucracy also responded to electoral considerations, and federal employment was an important tool to build support for the party. The political discretion over appointments and removals over most of the nineteenth century not only allowed politicians to employ individuals that could be trusted, but also opened the door to a "spoils system" where political support could substitute for qualifications (Fish, 1905; Hoogenboom, 1968).

[39] A common argument among both historians and economists is that the prospects of external war may lead to the development of more effective states. The relationship between the American Civil war and the development of the federal state is debated among scholars of American history. On the one hand, some consider the war a major turning point in the development of the American State (Beard, 1927; Hacker, 1940), labeling it the "Second American Revolution" (Ransom, 1998), as the war concentrated power away from states and in the hands of the federal government. On the other hand, other scholars have argued that the civil war might have retarded industrialization and, in turn, the development of state capacity (Cochran, 1961).

The large spikes in turnover in the years of presidential transitions, shown in Figure 3, are also consistent with this mechanism. Interestingly, we see some evidence that turnover starts declining before the meritocratic reforms which decreased the President's control over bureaucratic hiring.[40] Our results in the next section show that increases in monitoring capacity were associated with a reduction in turnover, as trust relationships between principals and agents were less necessary in determining performance. However, higher monitoring capacity does not decrease electoral incentives to hire copartisans. Thus, our results are not inconsistent with an increase in efficiency after the passage of reforms that curbed the electorally motivated turnover that still in part characterized the federal bureaucracy at the end of the nineteenth century.

### 4.4.3 Decrease in communication costs in absence of agency problems

Cognitive models of organizations predict a relationship between communication costs, size, and delegation of decision power, even in absence of agency problems (Bloom et al., 2014; Gumpert et al., 2022). Workers tasked with production in a location face problems for which they might need the help of the principal in the headquarter. Each location might employ a manager, who solve some of the problems that would otherwise flow to the principal. Lower communication costs with the headquarter decrease the amount of time that the principal needs in order to solve problems arising in the location. This, in turn, has two effects. First, it increases the probability of observing employment, and the size of the workforce, in the location. Second, it decreases the amount of managerial delegation to the location, holding the size of employment in the location fixed.

Thus, similar to a mechanism of higher monitoring capacity, also these models predict that lower communication costs between DC and a location should increase the presence of the federal government in the location. However, while looking at a decrease in communication costs through the prism of a higher principal's monitoring ability predicts *more* decentralization of managerial power, abstracting from agency problems predicts *less* decentralization as communication becomes cheaper.

Both forces are likely at play in our context: the development of the railroad and the telegraph networks allowed principals in DC both to monitor field offices better and also, abstracting from agency problems, to communicate information more cheaply. However, the results in the next section – where we find increases in managerial delegation as a location becomes better connected to DC, holding fixed the size of its workforce – are consistent with

[40]The 1883 Pendleton Act introduced meritocratic hiring in the federal bureaucracy, but initially only a small share of positions were affected by the reform. Most of the positions transitioned to meritocracy only in subsequent decades.

the historiography underlining the presence of severe agency problems in the relationships between the principals in DC and federal agents in the field in the nineteenth century U.S.

# 5 Innovations in Monitoring Capacity as Drivers of State Development

In this section, we provide an empirical test of our hypothesis. First, we exploit the expansion of the railroad network, whose features allow us to develop an identification approach to control for a host of factors correlated with both increases in DC's monitoring ability and with the development of the federal state in a location. Second, we show that we obtain similar results when exploiting the expansion of the telegraph network. Third, we provide evidence that suggests that lower monitoring costs are associated with a reduction in reliance on trust as a way to staff the organization.

## 5.1 The expansion of the railroad network

### 5.1.1 Data and estimating equation

Our goal is to measure how the expansion of the railroad network decreased the travel time between DC and different counties and to study whether this had an impact on the presence and the organization of the federal bureaucracy. To do so, 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.

Following Donaldson and Hornbeck (2016), we calculate the shortest path between DC and the centroid of each county. These shortest paths are calculated as the shortest travel times (measured in minutes), using a combination of travel by wagon, navigation, and railroad. Relative to Donaldson and Hornbeck (2016), who are interested in the lowest-cost freight routes and thus need to specify transportation cost parameters, we specify travel time parameters. The resulting measure, $Log\ Time\ to\ DC_{ct}$, is a continuous treatment variable that provides the log travel time (in minutes) in year $t$, between DC and the centroid of county $c$.[41] Online Appendix A12 shows the expansion of the railroad network over time.

[41] In other words, we exploit variations in travel time to DC that are driven by the expansion of the network across time and space, and not the staggered arrival of railroads in a specific county which, as discussed in the next section, is arguably endogenous.

Online Appendix Figure A13 shows how the average travel time between DC and other counties decreased over time between 1830 and 1900, from more than 100 hours in 1830 to less than 40 hours in 1900.

With this measure at hand, we estimate the following regression model on a county-year panel between 1821 and 1905:[42]

$$y_{ct} = \alpha_c + \gamma_t + \beta^R Log\ Time\ to\ DC_{ct} + \delta_t Distance_c + X_{ct}\theta + \epsilon_{it} \quad (4)$$

where $y_{ct}$ is one of our outcomes of interest measured in county $c$ and year $t$. We include a set of county fixed effects ($\alpha_c$), which capture time-invariant county-specific unobservables which affect the development of the state, and of year fixed effects ($\gamma_t$ ), which account for aggregate time-varying shocks in federal state development (e.g., for the Civil War). The matrix $X_{ct}$ includes a set of controls which we discuss in the next section. The coefficient $\beta^R$ measures whether outcome $y_{ct}$ changes differentially in counties that become better connected to DC (i.e. which experience a decrease in travel time to DC).[43] We also control for the straight line distance between county $c$ and DC, interacted with year fixed effects, allowing for differential changes over time in the outcome variables in counties with different geographic distances from DC.

### 5.1.2 Threats to identification

The key threat to identification is that the expansion of the railroads network, and in particular the way in which this affects travel time between a county and DC at a given point in time, might be endogenous: counties that experience a reduction in travel time to DC could have experienced a change in the presence and organization of the federal state even absent a decrease in the monitoring costs faced by DC. In particular, two are the main concerns for identification.

First, as discussed by Atack et al. (2010), railroad promoters and investors sought locations with high profitability, and were more likely to target counties with higher growth in population density and agricultural productivity. In addition, new railroad construction might increase local manufacturing activity through higher demand for construction materials (Fishlow, 1965). A crucial concern is then that changes in the presence and organization of the state associated with a reduction in $Log\ Time\ to\ DC_{ct}$ are the result of time-varying

[42]Given the near absence of any railroad in 1830, the travel times between DC and each county is the same before 1830, which allows us to extend the sample used for estimation back to 1821.

[43]Since the railroads network database is available at 10-years interval, 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 the first years of each decade give qualitatively similar results (see Online Appendix Table A6).

shocks in a county's economic growth rather than a decrease in travel time between county $c$ and DC.

In order to address these concerns, as in Donaldson and Hornbeck (2016), we can exploit the fact that variation in travel time between county $c$ and DC is driven by both (1) railroad construction in county $c$, and (2) changes in other, more distant portions of the railroad network. This allows us to shut down variation driven by (1), by controlling for railroads construction in county $c$, and to only rely on variation driven by (2), i.e. exploiting only variation in $Log\ Time\ to\ DC_{ct}$ driven by railroads expansion in other parts of the network (and thus arguably uncorrelated with local economic shocks in county $c$ and year $t$). Specifically, $X_{ct}$ includes an indicator taking value one if county $c$ contains any railroad track in year $t$, and a variable measuring the length of railroad track in county $c$ and year $t$. After the inclusion of these controls, $\beta$ is identified from more-distant changes in the railroads network that lead to a decreased travel time between county $c$ and DC.

A second, and related, concern is that more distant changes in the railroad network which reduced $Log\ Time\ to\ DC_{ct}$, are also associated with an increase in county $c$'s market access.[44] Since this in turn led to an increase in county agricultural land values (Donaldson and Hornbeck, 2016) and manufacturing activity (Hornbeck and Rotemberg, 2021), this might create a spurious correlation between $Log\ Time\ to\ DC_{ct}$ and our outcomes of interest. In order to account for this, we directly control for a measure of market access as in Hornbeck and Rotemberg (2021).[45] In doing so, we exploit the fact that expansions of the network which create similar changes in a county's market access do not necessarily result in equal changes in a county's travel time to DC. To illustrate this point, Online Appendix Figure A14 shows the relationship between a county's change in $Log\ Time\ to\ DC_{ct}$ between 1880 and 1890 and the county's change in log market access over the same period. While there is a significant negative relationship between the two changes, this correlation is not perfect, and two counties with similar increases in market access might experience different decreases in travel time to DC.[46]

Thus, our identifying assumption is that, conditional on our set of controls, state devel-

[44]Market access captures how easily county $c$ can trade with all other U.S. counties, assigning higher weights to counties with greater population.

[45]Formally, we control for log market access, where market access of county $c$ at time $t$ is defined as $MA_{ct} = \sum_{d \neq c}(1 + t_{cdt}/P)^{-\theta} L_{dt}$, where $t_{cdt}$ is the per ton county-to-county transportation costs (as in Donaldson and Hornbeck (2016)), $P$ is the average price per ton of transported goods between counties $c$ and $d$ at time $t$, $\theta$ is a measure of trade elasticity, and $L_{dt}$ is the population of county $d$ in year $t$. We follow Hornbeck and Rotemberg (2021) and use a value for $\theta$ of 3.05 and a value for $P$ of 38.7.

[46]In some of our specifications, we also control for log population and for the share of manufacturing employment in the county. Given the possibility that these constitute "bad controls," as state presence might itself affect population and manufacturing growth, we also show that their exclusion does not affect the estimates.

opment in a county with a decrease in travel time to DC would have been similar to other counties, if not for the increase in DC's ability to monitor the workforce in that county. While it is impossible to directly test this assumption, we believe that our identification strategy makes the exclusion restriction (that a lower travel time to DC affects state development only through enhanced monitoring ability) plausible, as it relies on the comparison of counties that, as the railroad network expands, (i) have similar local railroad construction in their territory, (ii) experience similar changes in their ability to trade, but (iii) experience different changes in their travel time to DC *specifically*.

Moreover, we also present two indirect tests that lend additional support to this assumption. First, we show that a shorter travel time between a county and other important cities did not increase the presence of the state in the county: it is time distance to DC specifically that predicts the presence of the federal state. Second, we show that changes in $Log\ Time\ to\ DC_{ct}$ are not associated with differential pre-trends in the development of the federal state.[47]

### 5.1.3 DC's monitoring capacity and state presence

Table 2 presents results from estimating equation 4. Column 1 reports estimates from the simple specification including only year fixed effects, county fixed effects, and the straight line distance between the county and DC interacted with year fixed effects. A faster connection between a county and DC thanks to the expansion of the railroad network increases the probability of observing a presence of the federal state in the county. Specifically, a one standard deviation decrease in $Log\ Time\ to\ DC_{ct}$ is associated with an increase in the probability of state presence of 0.34 standard deviations.

Column 2 reports estimates from a specification that also controls for local railroad construction in the county. While the estimated impact of travel time to DC decreases once we exploit only variation stemming from more distant changes in the network, it remains significant and substantial in magnitude. This estimate is not significantly affected when we additionally control for a county's market access (column 3), or for a county's population and share of manufacturing employment (column 4). The estimate in column 4 shows that a one standard deviation decrease in travel time to DC led to an increase in the probability of observing the federal state in the county by 0.26 standard deviations.

[47] 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.

**Table 2: Railroad expansion, DC's monitoring capacity, and state presence**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| | State presence | | | | Log Clerks | Log Profess. | Log Low skills |
| Log Time to DC | -0.193*** | -0.136*** | -0.146*** | -0.149*** | -0.268** | -0.057 | 0.113 |
| | (0.030) | (0.031) | (0.032) | (0.032) | (0.107) | (0.107) | (0.135) |
| Local Railroads | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Log Market Access | | | ✓ | ✓ | ✓ | ✓ | ✓ |
| Additional controls | | | | ✓ | ✓ | ✓ | ✓ |
| Observations | 99,673 | 99,673 | 99,673 | 97,618 | 29,418 | 29,418 | 29,418 |
| Sd dep. var. | 0.4583 | 0.4583 | 0.4583 | 0.4595 | 0.9261 | 0.9484 | 1.1160 |
| Sd Log Time to DC | 0.8019 | 0.8019 | 0.8019 | 0.8048 | 0.8815 | 0.8815 | 0.8815 |

*Notes:* The unit of observation is a county-year. *State presence* takes value one if the federal state is present in the county. *Log Clerks*, *Log Profess.*, *Log Low skills* are the logarithm of the total number of employees employed in clerical, professional, and low skills positions, respectively. *Log Time to DC* is the log of total time (in minutes) between DC and the county's centroid. All specifications control for county fixed effects, year fixed effects, and the straight line distance between the county and DC interacted with year fixed effects. In columns 2 to 7 we additionally control for an indicator taking value one if the county contains any railroad track, and the length of railroad track in the county. In columns 3 to 7 we additionally control for the county's log market access. In columns 4 to 7 we additionally control for the log of the county's total population and the log of the share of the county's population that is employed in manufacturing. Standard errors in parentheses, clustered at the county-level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

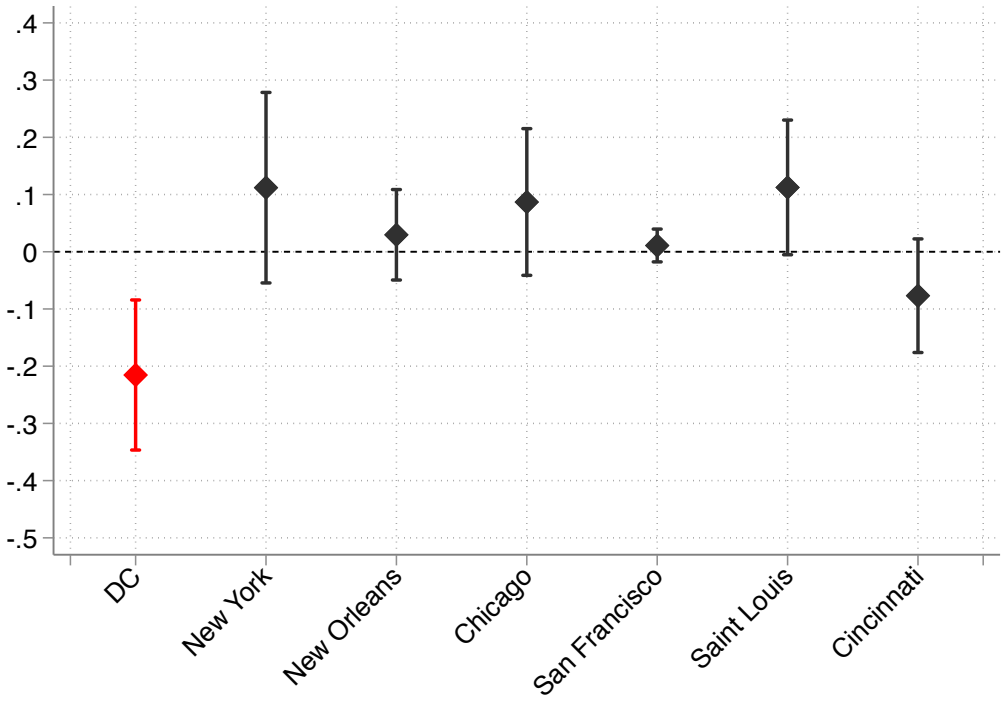
Columns 5-7 investigate which specific occupations were mostly affected by the increased monitoring ability of DC. In counties with state presence, reducing the traveling time to DC by one standard deviation increases the size of the clerical force by about 0.26 standard deviations. In contrast, we find no intensive margin effects for professional and relatively low skills positions.

Online Appendix Table A4 shows that we obtain very similar estimates if we limit the sample only to counties in states that were already part of the U.S. in 1821 (at the beginning of the sample period). This suggests that the results are not merely driven by the ability of the railroads to extend westward the American frontier (Bazzi et al., 2020).

A possible concern is that, even after accounting for a county's overall market access, a better connection to DC might per se matter for state development because of the specific economic significance of this city, and not because it is the capital of the federal government. To check whether this is the case, we re-estimate equation 4 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 city $P$. Each "placebo city" $P$ was an important economic center but, differently from DC, was not the center of power of the federal government.[48] We report the results in

[48] Additionally, as we do for DC, we control for the straight line distance between the county and each city $P$ interacted with year fixed effects.

**Figure 4: Travel time to other cities is not associated with state presence**



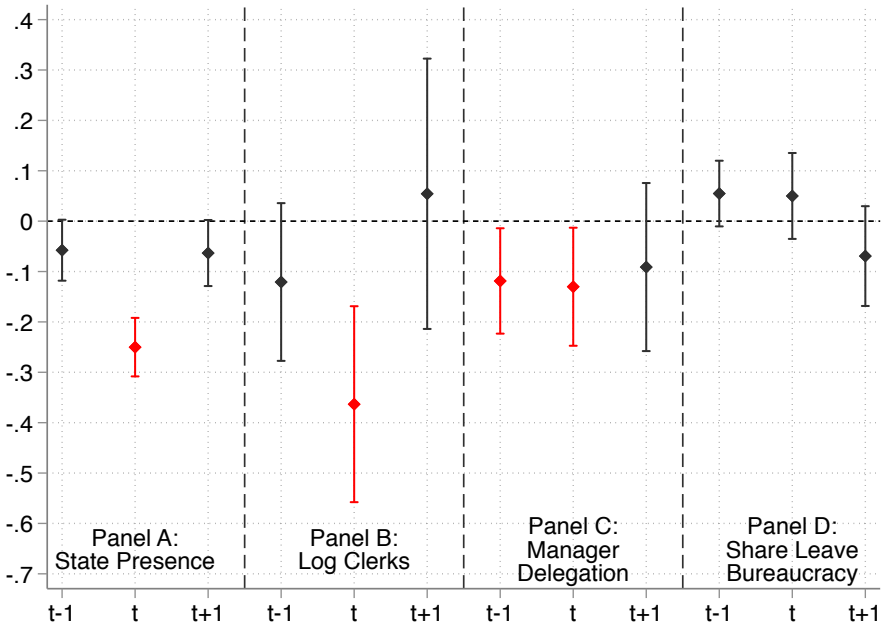

*Notes:* The figure shows results of estimating the specification in column 4 of Table 2 with the additional inclusion of 6 variables (*Log Time to P*) which are the log of total time (in minutes) between city $P$ and the county's centroid, as well as the straight line distance between the county and each city $P$ interacted with year fixed effects. We report the estimated effect for DC (in red) as well as for the 6 additional cities (in black), with 95 percent confidence intervals based on standard errors clustered at the county-level.

Figure 4 where we plot the estimated effect for DC as well as for New York, New Orleans, Chicago, San Francisco, Saint Louis and Cincinnati.[49] Consistent with our interpretation, after accounting for local railroad construction and for overall market access, only travel time to DC explains the presence of the federal state in a county, while travel time to other important cities is not associated with differences in state presence.

Finally, Figure 5 shows that changes in travel time to DC are not associated with differential pre-trends in the development of the federal state. We re-estimate equation 4 additionally including the travel time to DC in the previous and in the next decade (as well as controls for previous and future local railroad construction and market access).

[49]New York was the largest city in the North-East (was well as the largest city in the country), 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 city in the Midwest at some point during our sample period. Substituting travel time to New York City with travel time to Boston (which are highly correlated variables) produce similar results, with a coefficient on travel time to Boston of -0.047 (standard error =0.090).

**Figure 5: Past and future travel time to DC and the development of the state**




*Notes:* The figure shows the impact of the contemporaneous value of *Log Time to DC* (time $t$), as well as of past (time $t-1$) and future (time $t+1$) values of *Log Time to DC*, on the dependent variable shown at the bottom of each panel. The figure plots coefficients and 95 percent confidence intervals based on standard errors clustered at the county level. Estimates in red indicate a p-value$<0.05$. Each regression includes as right-hand side variables all the variables included in the specification in column 3 of Table 2, and additionally includes the following variables: *Log Time to DC* ten years in the future, *Log Time to DC* ten years in the past; an indicator taking value one if the county contains any railroad track ten years in the future, an indicator taking value one if the county contained any railroad track ten years in the past; the length of railroad track in the county ten years in the future; the length of railroad track in the county ten years in the past.

Panels A and B of the figure show that only current travel time to DC is associated with a larger presence of the state. In contrast, future travel time to DC (the estimates corresponding to $t+1$ in the figure) does not predict the development of the state in a location. Thus, conditional on our controls, we see similar pre-trends in state development in counties that will experience different changes in travel time to DC in the future. This lends further support to our assumption that counties that experience similar local railroad construction and similar changes in market access, but lower decreases in travel time to DC, represent a good control group for counties experiencing a faster decrease in travel time to DC in the same decade.[50]

[50]Interestingly, we do not see an effect of travel time to DC in the previous decade (the estimates corresponding to $t-1$ in the figure), consistent with an immediate response of DC to improved monitoring capacity.

**Table 3: Railroad expansion, DC's monitoring capacity, and state organization**

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| | Manager Delegation | | | | Share Leave Bureaucracy | | | |
| Log Time to DC | -0.277*** | -0.186*** | -0.180*** | -0.184*** | 0.055* | 0.051 | 0.057* | 0.062* |
| | (0.053) | (0.059) | (0.059) | (0.060) | (0.032) | (0.034) | (0.034) | (0.035) |
| Local Railroads | | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ |
| Log Market Access | | | ✓ | ✓ | | | ✓ | ✓ |
| Additional controls | | | | ✓ | | | | ✓ |
| Observations | 30,239 | 30,239 | 30,239 | 29,366 | 28,068 | 28,068 | 28,068 | 27,193 |
| Sd dep. var. | 0.4708 | 0.4708 | 0.4708 | 0.4673 | 0.3781 | 0.3781 | 0.3781 | 0.3771 |
| Sd Log Time to DC | 0.8805 | 0.8805 | 0.8805 | 0.8812 | 0.8928 | 0.8928 | 0.8928 | 0.8940 |

*Notes:* The unit of observation is a county-year. *Manager Delegation* is an indicator equal to one if there is at least one manager in the county. *Share Leave Bureaucracy* is the share of employees who left the federal bureaucracy between year $t$ and year $t-2$. *Log Time to DC* is the log of total time (in minutes) between DC and the county's centroid. All specifications control for county fixed effects, year fixed effects, and the straight line distance between the county and DC interacted with year fixed effects. In columns 2-4 and 6-8 we additionally control for an indicator taking value one if the county contains any railroad track, and the length of railroad track in the county. In columns 3-4 and 7-8 we additionally control for the county's log market access. In columns 4 and 8 we additionally control for the log of the county's total population and the log of the share of the county's population that is employed in manufacturing. In columns 1-4 we additionally control for a set of fixed effects for the total number of federal employees in the county. Standard errors in parentheses, clustered at the county-level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

### 5.1.4 DC's monitoring capacity and the organization of the state

Table 3 shows that a lower time travel between a county and DC not only influenced whether the federal state was present in a county, but also the way in which the state was organized in that county. Importantly, the estimates are once again robust to the inclusion of the most stringent set of controls.

We find a significant effect of decreased monitoring costs on the degree of delegation of managerial power outside of DC. The estimate in column 4 shows that a one standard deviation decrease in $Log\ Time\ to\ DC_{ct}$ is associated with an increase of about 0.35 standard deviations in the probability of observing managerial delegation to the county. Importantly, since the presence of employees with managerial responsibilities might simply be a by-product of having a larger workforce, these specifications additionally control for a full set of fixed effects for the total number of federal employees in the county.

In addition, increased monitoring capacity is also associated with less employee turnover. In counties with a longer travel time to DC, the share of employees who leave the bureaucracy within the next two years (i.e. they are not present in the subsequent volume of the Official Register) is significantly higher. Two counties that are one standard deviation apart in their

travel time to DC have a turnover rate that differs by 0.15 standard deviations.[51]

Panels C and D of Figure 5 show that we do not see differential pre-trends in these outcomes in counties that will experience different changes in travel time to DC in the future. We do find that past travel time to DC is associated to changes in managerial delegation and personnel turnover, suggesting that changes in monitoring capacity lead to more gradual adjustments along these margins than along the margin of state presence. Finally, Online Appendix Table A5 shows that we obtain very similar estimates if we limit the sample only to counties in states that were already part of the U.S. in 1821.

## 5.2 The expansion of the telegraph network

In this section, we leverage the expansion of the telegraph network across the U.S., in order to measure the ease of communication between DC and different locations at different points in time. The first telegraph line, connecting DC with Baltimore, opened in 1844. Private investors soon expanded the telegraph network, which by the early 1850s had connected all major urban centers (Highton, 1852).

We rely on data from Wang (2020), who collected information on the year in which different locations were connected to the telegraph network between 1844 and 1852.[52] The data collection effort by Wang (2020) ends in 1852 since comprehensive information on the telegraph network after 1852 is unavailable. In addition, the rapid expansion of the network limits the extent of variation after the mid-1850s, as by then most major centers had a telegraph connection. For each year between 1844 and 1852, we compute the variable $Telegraph\ Connections_{ct}$, namely the number of telegraph stations in each county $c$ and year $t$. Online Appendix Figure A15 shows the geographical distribution of the variable from 1845 to 1853.

With this measure at hand, we estimate the following regression model on a county-year panel between 1839 and 1953:[53]

$$y_{ct} = \alpha_c + \gamma_t + \beta^T Telegraph\ Connections_{ct} + \delta_t Distance_c + X_{ct}\theta + \epsilon_{it} \tag{5}$$

where all variables are defined as in equation 4, and $X_{ct}$ includes log population and the share of manufacturing employment. The coefficient $\beta^T$ measures whether outcome $y_{ct}$ changes

[51] We exclude from this specification observations in 1905, since we do not have information on which employees leave by 1907.

[52] We are very grateful to Tianyi Wang for providing access to the data.

[53] We choose the year 1939 as the first year in this estimating sample in order to include three years (1939, 1941, 1943) in the "pre-telegraph" era. Results in which we restrict the sample to the 1841-1853 period or to the 1843-1853 period are qualitatively similar.

**Table 4: Telegraph connections, DC's monitoring capacity, and the development of the state**

| | (1) State Presence | (2) Log Clerks | (3) Log Profess. | (4) Log Low skills | (5) Manager Delegation | (6) Share Leave Bureaucracy |
|---|---|---|---|---|---|---|
| Telegraph Connections | -0.001 | 0.081*** | -0.005 | 0.034* | 0.027** | -0.046*** |
| | (0.008) | (0.030) | (0.020) | (0.018) | (0.011) | (0.015) |
| Observations | 15,583 | 2,212 | 2,212 | 2,212 | 2,167 | 2,212 |
| Std. dev. dep. var. | 0.3555 | 0.7776 | 0.9406 | 0.7459 | 0.4945 | 0.3805 |
| Mean dep. var. | 0.1484 | 0.6946 | 0.7545 | 0.4963 | 0.4255 | 0.5248 |
| Std. dev. Telegraph | 0.3259 | 0.5563 | 0.5563 | 0.5563 | 0.5541 | 0.5563 |

*Notes:* The unit of observation is a county-year. *Telegraph Connections* is the number of locations connected to the telegraph in the county. *State presence* takes value one if the federal state is present in the county. *Log Clerks*, *Log Profess.*, *Log Low skills* are the logarithm of the total number of employees employed in clerical, professional, and low skills positions, respectively. *Manager Delegation* is an indicator equal to one if there is at least one manager in the county. *Share Leave Bureaucracy* is the share of employees who left the federal bureaucracy between year $t$ and year $t-2$. All specifications control for county fixed effects, year fixed effects, the straight line distance between the county and DC interacted with year fixed effects, the log of the county's total population, and the log of the share of the county's population that is employed in manufacturing. In column 5 we additionally control for a set of fixed effects for the total number of federal employees in the county. Standard errors in parentheses, clustered at the county-level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

differentially in counties that become better connected to DC thanks to a higher number of telegraph stations.

The identifying assumption is that, absent the telegraph, the presence and organizational features of the federal state would have evolved similarly in counties connected to DC via the telegraph and in counties without this connection. In this case, our ability to assuage concerns regarding the identifying assumption is more limited relative to our analysis exploiting the development of the railroads. This is due to the technological features of the telegraph network: if a location has a telegraph, it is connected with the same speed to DC and to all other nodes of the network, irrespective of the specific structure of the network.

Importantly, we find that a county's change in federal state presence between 1833 and 1843 (i.e. in the decade immediately before the beginning of the telegraph era) is uncorrelated with its future number of telegraph connections.[54] Notwithstanding, we acknowledge that the identification assumption in this section is less likely to be perfectly satisfied, even conditional on controls, and thus these results should be seen as more suggestive.

Table 4 shows the results of estimating equation 5. Increasing telegraph connections

[54] A regression of the change in *State Presence* between 1833 and 1843 on the number of telegraph connections between 1843 and 1852, controlling for the change in county's population and in the county's share of manufacturing employment, give a coefficient of -0.005 (standard error 0.011).

does not increase the probability that a county switches to having a state presence, but, conditional on state presence, we observe a significant increase in both the number of clerks and of blue collar workers: a one standard deviation increase in the number of locations with telegraph connections in a county is associated with increases in clerical and blue collar workers of 0.06 and 0.03 standard deviations, respectively. We also observe significant effects of telegraph connections on the organizational features of the federal bureaucracy: counties with access to more telegraph connections to DC are more likely to be delegated managerial power and have a lower turnover of their workforce.

Despite the more suggestive nature of this empirical exercise, we view the fact that we obtain results that are similar to those obtained exploiting the expansion of the railroads network as further corroborating our interpretation linking monitoring capacity to state development.

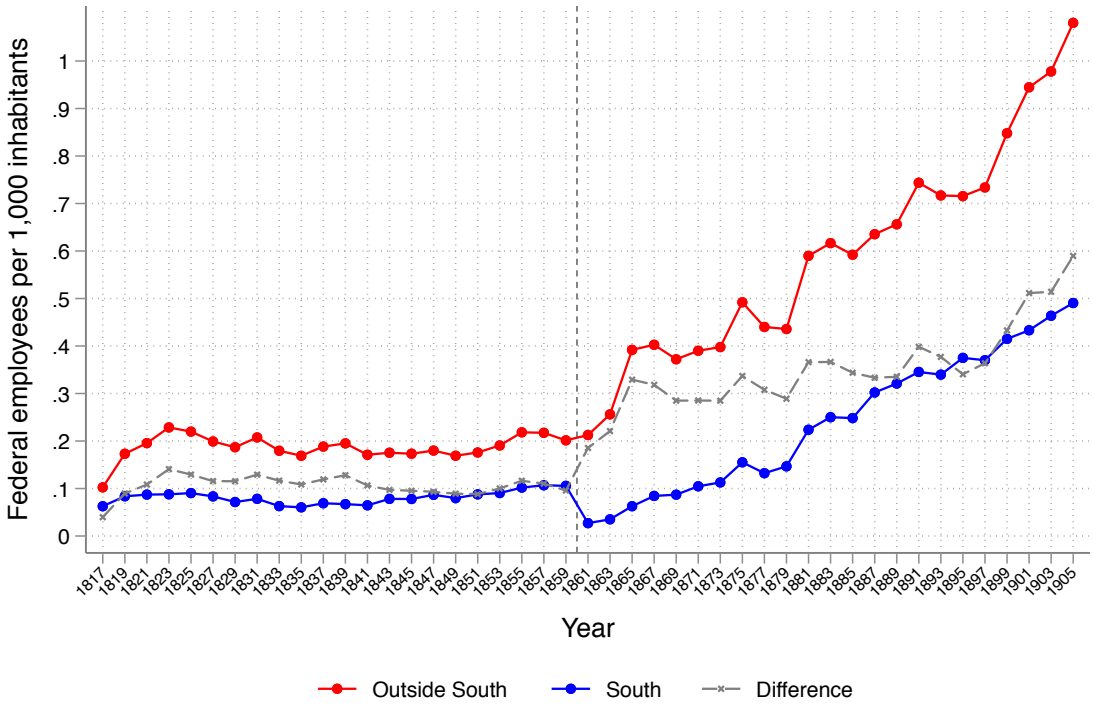
## 5.3 Monitoring capacity reduces reliance on trust

The results in the previous sections show that lower communication and transportation costs between DC and a county are associated with an increased likelihood of state presence, a larger presence of the state, more delegation of managerial power, and lower employee turnover in the county. Our interpretation for these results is that innovations in technologies of control, by increasing the government's monitoring capacity, created the conditions for a shift from a personal organization to a more modern bureaucratic organization, with lower reliance on networks of trust as a way to select bureaucrats.

In order to further corroborate this interpretation, we now provide suggestive evidence that a lower time distance between a county and DC decreased reliance on trust as a way to staff the bureaucracy in that county. We show that counties that become "better connected" to DC thanks to the railroads network expansion see an increased presence of workers who are relatively less trusted by the government after the civil war, namely those born in former confederate states.

Figure 6 motivates our empirical test. It plots the evolution over time in the number of federal employees, 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. Employees from confederate states 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 civil war. At the onset of the conflict, there is a sizeable decline in the number of Southern-born federal bureaucrats.

**Figure 6: 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.

More surprisingly, the North-South employment gap is persistent (see gray series): while the numbers of Southern-born and Northern-born bureaucrats constantly increase after 1861 as the federal state expands its scope, the difference in employees per capita between Southern and Northern states increases from about 0.1 in 1859 to about 0.3 in 1865, 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. We exploit this fact to investigate whether an increase in the federal government's monitoring capacity is associated with an attenuation of this North-South employment gap.

Table 5 reports the results of estimating Equation 4, using as dependent variable the share of employees in county $c$ and year $t$ who were born in a confederate state.[55] The estimate in column 1 implies that a one standard deviation decrease in travel time to DC increases the share of workers born in a confederate state by 0.3 standard deviations. Consistent

[55]Since our data on the telegraph network ends before the civil war (in 1852), we cannot exploit the expansion of the telegraph for the analysis in this section.

**Table 5: Increased monitoring capacity increases the share of Southern employees**

| Dep. var. is Share of workers born in a Confederate state | | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| | *Full sample* | *Post civil war* | *Pre civil war* |
| Log Time to DC | -0.137*** | -0.357*** | 0.031 |
| | (0.036) | (0.051) | (0.054) |
| Observations | 27,153 | 21,945 | 5,058 |
| Std. dev. dep. var. | 0.3996 | 0.3963 | 0.4132 |
| Std. dev. Log Time to DC | 0.8740 | 0.8410 | 0.9285 |

*Notes:* The unit of observation is a county-year. The dependent variable in all columns is the share of a county's employees who were born in Alabama, Arkansas, Florida, Georgia, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Texas, Virginia. Log Time to DC is the log of total time (in minutes) between DC and the county's centroid. Controls are the same as in column 4 of Table 2. The sample in column 1 includes all counties with state presence in all odd years between 1821-1905. The sample in column 2 is limited to the 1861-1905 period, and the sample in column 3 is limited to the 1821-1859 period. Standard errors in parentheses, clustered at the county-level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

with increased monitoring capacity substituting for reliance on trust as a way to staff the bureaucracy, the entire effect is concentrated in the post-civil war period. In the 1861-1905 period, a one standard deviation decrease in $LogTimetoDC$ leads to an increase in the share of southern-born employees of 0.76 standard deviations. In contrast, in the 1821-1859 period, there is no significant relationship between DC's monitoring capacity and the share of southern-born employees in a county.

We interpret this result as suggestive of the theoretical mechanism behind our results. Lower transportation and communication costs, by enhancing the government's ability to monitor the behavior of its agents throughout the territory, decrease the need for employing trusted individuals. In addition, the results in Table 5 points towards an important role of higher monitoring capacity in attenuating persistent employment discrimination against groups who are relatively less trusted by the government.

# 6 Conclusion

Mann (1984) defines *infrastructural power* as *"the capacity to implement logistically political decisions throughout the realm"*. A large literature has investigated the incentives to set up a state apparatus with the capacity to implement these decisions. However, once these incentives are in place, how is a state concretely organized? In this paper, we study this broad question leveraging a unique dataset that allows us to investigate the internal organization

of a state over an unusually long time-span. We assembled a new micro-database which combine personnel records of the U.S. federal bureaucracy over the period 1817-1905, and hand-collected information on the internal organization of the bureaucracy.

Our novel data allow us to document a number of novel descriptive facts on the development of the U.S. federal bureaucracy. First, we show that the state expanded in size mainly since the 1860s, and that an important driver of this growth was its ability to reach new locations. Second, the presence of the federal state was higher in more prosperous locations, but, in the first part of the nineteenth century, distance from DC limited the association between state presence and growth. Third, the organization of the state started to change since the 1860s, with a lower reliance on employee turnover, a less tight link between workers' and their supervisors' careers, and an increasing delegation of managerial power away from DC.

We interpret these facts through the lenses of principal-agent theory. In presence of low monitoring capacity, the state had low growth potential, and the optimal way to manage the state apparatus resembled a personal organization, with relationships of trust replacing effective monitoring. Technological innovations that lowered monitoring costs were conducive to organizational change, making it optimal to adopt a modern, Weberian organizational form, and allowing faster organizational growth. Exploiting the staggered introduction of the railroads and telegraph network across different locations over the nineteenth century, we provide evidence in support of our interpretation.

Our results underline how principal-agent relationships 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 over the process of development: changes in a ruler's ability to monitor state agents affect both the growth potential of a state apparatus and its organizational form. This highlights how different systems for organizing a state, characterized by different degrees of principals' discretion over personnel choices, might be optimal at different stages of development.

While our study focuses on innovations in monitoring ability as a driver of change in state organizations, we have discussed alternative mechanisms that are arguably also important to explain the facts that we document. We believe that an investigation of these mechanisms, in this or other contexts, is an important avenue for future research.

# Bibliography

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 (2020): "The economic origins of government," .

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," .

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.

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.

Beard, C., B. M. (1927): *The Rise of American Civilization*, Macmillan.

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., 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.

——— (2010): "State capacity, conflict, and development," *Econometrica*, 78, 1–34.

——— (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.

Bonney, R. (1999): *The rise of the fiscal state in Europe c. 1200-1815*, Clarendon Press.

Cantoni, D., C. Mohr, and M. Weigand (2019): "The rise of fiscal capacity," Tech. rep., Discussion Paper.

Census, B. o. (1904): *The Executive Civil Service of the United States*, Government Printing Office.

Chambru, C., E. Henry, and B. Marx (2022): "The Dynamic Consequences of State-Building: Evidence from the French Revolution," Tech. rep., Working Paper.

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.

Cochran, T. (1961): "Did the Civil War Retard Industrialization?" *The Mississippi Valley Historical Review*, 48, 197–210.

Coleman, W. (1801): *An Examination of the President's Reply to the New-Haven Remonstrance*, George Hopkins.

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.

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 (2022): "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.

Field, A. J. (1992): "The magnetic telegraph, price and quantity data, and the new management of capital," *The Journal of Economic History*, 52, 401–413.

Finan, F., B. A. Olken, and R. Pande (2017): "The personnel economics of the developing state," *Handbook of economic field experiments*, 2, 467–514.

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.

García-Jimeno, C., A. Iglesias, and P. Yildirim (2022): "Information networks and collective action: evidence from the women's temperance crusade," *American economic review*, 112, 41–80.

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.

Hacker, L. (1940): *The Triumph of American Capitalism: The Development of Forces in American History to the End of the Nineteenth Century*, Columbia University Press.

Hamilton, A. (1795): *The Works of Alexander Hamilton, (Federal Edition)*, vol. 6, G. P. Putnam's Sons.

Highton, E. (1852): *The Electric Telegraph Its History and Progress by Edward Highton*, J. Weale.

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. and M. Rotemberg (2021): "Growth Off the Rails: Aggregate Productivity Growth in Distorted Economies," .

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 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. (1984): "The autonomous power of the state: its origins, mechanisms and results," *European Journal of Sociology/Archives européennes de sociologie*, 25, 185–213.

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.

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.

Ransom, R. L. (1998): *The Economic Consequences of the American Civil War*, Kluwer Academic Publishers.

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*.

Schönholzer, D. (2017): "The origin of the state: Incentive compatible extraction under environmental circumscription," *Available at SSRN 2944106*.

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.

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): "Reflections on the History of European State-Making," Princeton University Press.

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

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

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.

Wang, T. (2020): "The Electric Telegraph, News Coverage and Political Participation," *Working paper*.

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.

Xu, G. (2018): "The costs of patronage: Evidence from the british empire," *American Economic Review*.

Zophy, J. W. (1975): "The Formation of National States in Western Europe: Tilly, Charles, ed.:(Studies in Political Development) Princeton: Princeton University Press 711 pp., Publication Date: July 25, 1975," *History: Reviews of New Books*, 4, 5–5.