NBER WORKING PAPER SERIES

WHO BENEFITS FROM MERITOCRACY?

Diana Moreira
Santiago Pérez

Working Paper 30113
http://www.nber.org/papers/w30113

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2022

We thank Luiza Aires, Mario Remigio and Lisa Pacheco for outstanding research assistance, and Enrique Pérez for help with data collection. We have benefited from the comments of Assaf Bernstein, Sandra Black, Leah Boustan, Shari Eli, James Feigenbaum, James Fenske, Ed Glaeser, Sun Go, Claudia Goldin, Walker Hanlon, Leander Heldring, Rick Hornbeck, Sarah Quincy, Chris Meissner, Marco Tabellini, Angela Vossmeyer, Tianyi Wang, Zach Ward, Guo Xu, and Noam Yutchmann, as well as by seminar participants at Corporacion Andina de Fomento, NBER Postdocs Meeting, NBER Summer Institute DAE, the Annual Cliometrics Conference, the Economic History Association Meetings, Universidade Federal de Pernambuco (PIMES) Political Economy Workshop, University of British Columbia, Melbourne University, University of Ottawa, University of Southern Denmark, Brown University, Yale University, Northwestern University, University of Toronto, Rutgers University, Pittsburgh University, Harvard University, Wilfrid Laurier University, Warwick University, UCSD, and the University of Houston. We benefited from funding from the UC Davis Small Grant in Aid of Research, the Michael Dearing Fellowship in support of Economic History Research, and the Hellman Fellowship program. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Who Benefits from Meritocracy?
Diana Moreira and Santiago Pérez
NBER Working Paper No. 30113
June 2022
JEL No. J15,J62,M5,N21

ABSTRACT

Does screening applicants using exams help or hurt the chances of lower-SES candidates? Because individuals from lower socioeconomic backgrounds fare, on average, worse than those from richer backgrounds in standardized tests, a common concern with this "meritocratic" approach is that it might have a negative impact on the opportunities of lower-SES individuals. However, an alternative view is that, even if such applicants underperformed on exams, other (potentially more discretionary and less impersonal) selection criteria might put them at an even worse disadvantage. We investigate this question using evidence from the 1883 Pendleton Act, a landmark reform in American history which introduced competitive exams to select certain federal employees. Using newly assembled data on the socioeconomic backgrounds of government employees and a difference-in-differences strategy, we find that, although the reform increased the representation of "educated outsiders" (individuals with high education but limited connections), it reduced the share of lower-SES individuals. This decline was driven by a higher representation of the middle class, with little change in the representation of upper-class applicants. The drop in the representation of lower-SES workers was stronger among applicants from states with more unequal access to schooling as well as in offices that relied more heavily on connections prior to the reform. These findings suggest that, although using exams could help select more qualified candidates, these improvements can come with the cost of increased elitism.

Diana Moreira
Department of Economics
University of California, Davis
One Shields Avenue
Davis, CA 95616
and NBER
dsmoreira@ucdavis.edu

Santiago Pérez
Department of Economics
University of California at Davis
One Shields Avenue
Davis, CA 95616
and NBER
seperez@ucdavis.edu

An online appendix is available at http://www.nber.org/data-appendix/w30113

1 Introduction

Screening applicants based on their performance in an exam is common in many contexts, ranging from the education system to the recruitment of workers in the public and private sectors. Because individuals from lower socioeconomic backgrounds fare, on average, worse than those from richer backgrounds in standardized tests, a common concern with this "meritocratic" approach is that it might have a negative impact on the opportunities of lower-SES individuals.[1] Indeed, there has been a recent push to limit the influence of exams in different selection processes, often fueled by the concern that using exams could help perpetuate (or even aggravate) existing social and economic inequalities.[2]

However, an alternative view is that exams might actually *help* the chances of applicants from disadvantaged backgrounds. For instance, in describing the benefits of using exams in hiring, a Forbes editorial argues that such an approach can provide "an unbiased, fair and consistent basis for employee selection," and thus "open the door for a more diverse set of employees" (Forbes, 2020). This view is grounded in the idea that, even if lower-SES applicants underperformed on exams, other (potentially more discretionary and less impersonal) selection criteria might put such applicants at an even worse disadvantage.[3]

Despite this ambiguity and the substantial controversy surrounding the use of exams, there is limited empirical evidence investigating whether (and when) exams actually hurt the chances of lower-SES candidates. The key challenge in providing such evidence is that it requires comparing the representation of lower-SES candidates when selection is through exams to their representation under *alternative* selection criteria. However, doing so is challenging because, for a given recruiting organization and position, applicants are usually all screened through the same procedure.

We study whether exams decrease or increase the representation of lower-SES individuals using evidence from the 1883 Pendleton Act, a landmark reform in American history which introduced competitive exams for the selection of certain federal employees. Similar to many countries today, government jobs in late 19th century US were highly coveted and paid higher salaries than comparable jobs in the private sector (Aron, 1987; Finan *et al.*, 2017). Before the passage of the Pendleton act, these jobs were allocated at the discretion of government officials and often based on political and personal connections (Aron, 1987). After its passage, in contrast, some federal positions had to be allocated to the applicants with the top scores in a high-stakes open exam. We find

[1] For evidence on the correlation between socioeconomic status and exam performance see, for instance, Council *et al.* (1989) and Camara & Schmidt (1999).

[2] In the public sector context, exams have been described as preventing the recruitment of minority police officers (see CBS, 2020). An earlier example of the pushback against exams in the public sector is the discontinuation (after a lawsuit arguing that the exam discriminated against minority applicants) of the *Professional and Administrative Career Examination* during Carter's administration. There are also several recent examples of criticisms of exams in the educational context (see, for instance, New York Post, 2018 and New York Times, 2021).

[3] Cronbach (1975) writes that "proponents of testing, from Thomas Jefferson onward, have wanted to open doors for the talented poor, in a system in which doors are often opened by parental wealth and status." Zhao (2014) describes the Chinese civil service exams as "a tool to identify and recruit the most capable and virtuous individuals into government instead of relying on members of the hereditary noble class." In the educational context, the MIT justified its recent decision to reinstate the SAT on the grounds that the exam helped lower-SES candidates (The Atlantic, 2022).

that this change *reduced* the representation of lower-SES applicants in federal jobs, while increasing the representation of the middle class. We argue that middle-class applicants benefited from the reform because they were overrepresented among "educated outsiders": individuals with high levels of education but limited connections.

The public sector is a particularly important setting for studying whether exams hurt or help the representation of lower-SES candidates. First, it is the largest employer in many countries, making its use of exams potentially consequential for a large group of individuals. Moreover, most countries use similar examinations to the ones we study to select their public employees (Teorell *et al.*, 2011). Second, it is a sector in which balancing the representation of different groups in society might be especially relevant.[4] Indeed, the issue of "representation of the less priviledged" has been at the center of recent discussions about the selection of government officials.[5]

In addition to being an important setting for studying the equity implications of exams, our context is also attractive from an empirical standpoint:

First, the reform enables us to observe the background characteristics of individuals hired to do the *same job* in the *same office*, some of whom were selected through exams and some of whom were selected through more informal criteria. Moreover, we can leverage for identification the fact that not all federal positions were initially subject to exams. Specifically, among positions in the Executive Departments in DC (our main focus in this paper), the reform exempted those at the bottom (such as laborers) and those at the top (such as bureau chiefs) of the state hierarchy. We use this feature of the reform to estimate difference-in-differences models, comparing the characteristics of employees hired before and after the reform, in exempted and non-exempted positions.

Second, we observe unusually rich information on the socioeconomic backgrounds of government employees. To assemble these data, we first digitized federal personnel records spanning 1871 to 1893, roughly a decade before and after the passage of the reform. These records include employees' names, birthplaces, salaries, and job title. We then used name-based matching techniques (Abramitzky *et al.*, 2019) to link these records to US population censuses, enabling us to observe workers' background characteristics such as parental wealth, parental occupations, country of origin, and race.

Third, unlike with more recent policy changes, our setting enables us to assess both the short- and the long-run effects of exams.[6] Doing so is important as exams' longer-term equity impli-

[4]Several studies document a link between the personal characteristics of government officials and policy outcomes, see, for instance, Keiser *et al.* (2002); Pande (2003); Chattopadhyay & Duflo (2004); Beaman *et al.* (2012); Riccucci *et al.* (2014); Xu (2020).

[5]There are several examples of efforts to increase the representation of disadvantaged groups in the public sector. For instance, in 2015 the US Department of Justice launched the "Advancing Diversity in Law Enforcement," an agency designed to help "recruit, hire, retain, and promote officers that reflect the diversity of the communities they serve" (Equal Employment Opportunity Commission, 2016). There are also initiatives to diversify the teaching workforce, see, for instance, Department of Education (2020). This emphasis is also present at higher levels of the state hierarchy; President Biden promised to nominate a Black woman to the Supreme Court, arguing that: "And, look, in terms of making everything, the corridors of powers, reflect what America looks like - that includes the White House, that includes the staff there. That includes the Cabinet and that includes the Supreme Court and the Congress."

[6]For example, some US colleges have recently dropped the SAT requirement. While an evaluation of the short-run impacts of this change has not (to the best of our knowledge) yet been conducted, it would in principle be possible.

cations might differ from their immediate ones. For instance, exams might initially benefit the chances of lower-SES candidates but then lose their equalizing force as test-preparation tools (to which the rich might have better access) emerge.[7] Indeed, consistent with the high desirability of government jobs in our setting, we document a rapid emergence of such tools–including tutoring services and test-preparation books.[8]

Our main finding is that the reform led to an immediate and persistent (for at least 10 years) decline in the representation of lower-SES individuals. First, employees hired through exams came from families that were 6 percentile ranks higher in the national wealth distribution. This increase was driven by a reduced representation of workers with parents at the bottom of the wealth distribution, together with an increase in the share of workers from upper-middle class families. Second, the reform increased the share of employees with higher-status parental occupations: we find a 5 percentage points increase in the proportion of children of professional fathers (nearly a 50% increase), together with a similar decrease in the proportion of children of blue-collar fathers. Finally, the reform also reduced the share of first- and second-generation immigrants (by 4 and 7 percentage points, respectively).[9] Interestingly, this increased elitism occurred despite the exam being based on content that should have, in principle, been accessible for applicants with only a modest educational background.[10]

Our interpretation of the findings is that the reform increased the representation of the middle-class (at the expense of lower-SES applicants) because middle-class applicants were overrepresented among "educated outsiders": individuals with high levels of education but low levels of connections.

We first introduce a simple conceptual framework that formalizes this interpretation. In this framework, access to jobs depends on two attributes, "education" and "connections", both of which are potentially correlated with applicants' social class. We conceptualize the reform as an increase in the rewards to formal education in the hiring process. Therefore, the reform helps applicants with high levels of education but low levels of connections. Whether the poor, the middle-class, or the rich increase their representation depends on the relationship between social class, education, and connections. For instance, when there is a stronger association between social class and education than between social class and connections, increasing the relative importance of education hurts the representation of lower-SES individuals.

We then present several pieces of evidence that support this interpretation:

First, we show that the reform indeed brought in "educated outsiders". Specifically, employees hired through exams were likely better educated than those appointed through more informal

However, such an evaluation would not answer how dropping the SAT would affect admitted students' backgrounds in the longer run (as students and their families adjust to the new system).

[7] For instance, a common critique of the SAT is that applicants from wealthy backgrounds have learnt how to "game" the exam through the use of tutoring, medical exemptions, etc. See for example CNBC (2019)).

[8] We provide more details on this issue in Section 2.

[9] In contrast, the reform did not significantly decrease the share of Black employees: whites remained the vast majority of employees before and after the reform.

[10] Applicants with only a "common school" (the name often used to refer to public schools in historical US) education regularly took and passed the exam. We provide further details about the exams in Section 2.

criteria: they were more likely to have held a professional occupation–such as lawyer or accountant –prior to joining government, and were also more likely to have spent their childhoods in counties with higher per capita schools and teachers.[11] Moreover, exam-based hires were also more likely to lack the type of personal and political connections that facilitated access to patronage jobs: they were less likely to have a father who was himself a bureaucrat, less likely to have grown up in DC, and less likely to hail from a county in which a majority of voters had supported the incumbent party (suggesting a decline in political favoritism).

Next, we show that middle-class individuals were likely overrepresented among the "educated outsiders". First, although such individuals were, on average, more educated than those from poorer backgrounds, in the pre-reform period they represented a similar fraction of workers in the positions that eventually became subject to exams. This similar representation is consistent with the idea that, prior to the reform, lower-SES applicants may have compensated for their lower formal education by being more likely to engage in patronage politics–presumably because their worse outside options made them an easier target for political machines. Indeed, we show that groups historically tied to political machines (immigrants and urban residents) decreased their representation after the reform. Second, we show that, after the reform, the representation of middle-class individuals in government positions became closer to their representation in comparable *private sector* white-collar jobs. This convergence suggests that the low pre-reform representation of the middle class in government jobs was unusual relative to its education.

Finally, we show that the effects of exams are heterogeneous in ways that are in line with this interpretation. First, exams had the most negative effects on the chances of lower-SES candidates when such applicants hailed from states with high inequality in access to schooling–namely, the places in which the children of the poor were the least likely to be represented among the "educated outsiders". Second, the increased elitism is concentrated in offices in which, prior to the reform, a high share of workers received a higher compensation than that which their pre-civil service occupation (our main proxy for education) would predict–namely, the offices which likely assigned a greater weight to connections in the pre-reform period. These findings suggest that, consistent with our conceptual framework, the extent to which exams decrease the representation of lower-SES candidates depends both on the distribution of education and connections across social groups and on how these two attributes are rewarded.

Related Literature. Although there is a well-documented correlation between socioeconomic status and performance in standardized exams, Autor & Scarborough (2008) is, to the best of our knowledge, the only other study that investigates the equity implications of exams by comparing this approach to an *alternative* selection criteria. Autor & Scarborough (2008) find that, in a retail firm, the introduction of job testing *did not* reduce the representation of minority applicants. Hence, our paper is *the first* that compares exams to more informal criteria and shows that exams can be detrimental for the chances of lower SES-candidates. One likely explanation for our different findings is that, due to the "democratizing" force of political patronage, the distribution

[11] As population censuses prior to 1940 do not include information on years of schooling, we cannot directly investigate if employees hired through exams had completed more years of education.

of "connections" might be more egalitarian in the public than in the private sector. Indeed, we find that the reduction in the representation of lower-SES individuals was weaker in the government offices that relied less on connections in the pre-reform period (which might have been more similar to firms in the private sector). Beyond our different focus (the public sector) and results (that exams reduce the representation of lower-SES individuals), we deviate from Autor & Scarborough (2008) in two other main ways. First, we characterize changes in employees' backgrounds beyond their minority status. This is important because exams could leave the racial composition of selected employees unchanged and yet result in a more "elitist" workforce (as we find in our context). Second, we measure both the immediate and the longer-term consequences of exams. Doing so is particularly relevant in our context as, unlike Autor & Scarborough (2008), we study a change implemented by a large and prominent employer (which could have presumably spurred the longer-term responses to exams discussed above).[12]

More broadly, we also contribute to the literature on personnel policies and workplace inequality.[13] Similar to ours, a number of studies in this literature focus on the effects of adopting more impersonal, less discretionary hiring criteria.[14] We contribute to this literature by providing some of the first evidence on the equity implications of exams, a common (and controversial) recruitment tool.[15] Our findings show that reducing hiring discretion might not necessarily translate into a higher representation of lower-SES individuals.

Finally, we contribute to the literature on the effects of civil service exams. Inspired by a tradition that emphasizes the importance of selecting the most competent workers (Weber, 2019), a number of papers in this literature have studied the extent to which exams improve bureaucrats' qualifications and organizational performance (Ornaghi, 2016; Xu, 2018; Estrada, 2019; Moreira & Pérez, 2021).[16] By contrast, and motivated by the evidence on the importance of representation in public organizations (Kingsley, 1944; Neggers, 2018; Alsan *et al.*, 2019; Xu, 2020), we study their effects on the social composition of the bureaucracy.[17] Our findings suggest that exams are not

[12] Hoffman *et al.* (2018) study the *productivity* effects of exam-based hiring but do not focus on its effects on workers' socioeconomic backgrounds.

[13] This literature has studied the equity implications of policies such as relying on employees' referrals for recruitment (Beaman *et al.*, 2018), performance pay (Castilla, 2008), enabling workers to negotiate their salaries (Biasi & Sarsons, 2022), or basing promotions on subjective evaluations of workers' potential (Benson *et al.*, 2021).

[14] For instance, Goldin & Rouse (2000) evaluate the impacts of screening applicants using blind auditions, and Li *et al.* (2020) study the implications of hiring workers using machine learning algorithms.

[15] According to the Harvard Business Review (2015), "about 76% of organizations with more than 100 employees rely on assessment tools such as aptitude and personality tests for external hiring". There is a long-standing debate on the equity implications of job testing (and of standardized tests more generally), see, for instance, Cronbach (1975) and the discussion in Autor & Scarborough (2008).

[16] In a recent paper, we investigate the consequences of the Pendleton act for the functioning of the US Customs Service (Moreira & Pérez, 2021). We deviate from Moreira & Pérez (2021) with respect to research question, data, and empirical strategy. First, while Moreira & Pérez (2021) studies the consequences of the reform for the *efficiency* of the US Customs Service, we focus on how the reform affected the social origins of civil servants across the Federal administration. To do so, we digitize personnel records spanning every executive Department in DC (rather than just the Customs Service), and collect information on employees' parental wealth and occupations by linking these records to population censuses. Finally, our current analysis exploits variation in exam requirements across *positions*, whereas Moreira & Pérez (2021) exploits variation in requirements across different customs-collection districts.

[17] There are also papers studying the consequences of civil service exams for downstream outcomes beyond the bu-

simply a tool for legitimizing the status quo, elitist social origins of civil servants (as argued by scholars such as Bourdieu (1998)), but rather that they could be consequential.[18]

2 Historical Background

2.1 Spoils System and the Civil Service Reform Movement

Prior to the Pendleton Act, hiring decisions in the federal civil service were ruled by the "spoils system". Under this system, appointment to office was based primarily on political and personal connections rather than on "fitness for office" (Ziparo, 2017). As described by Aron (1987), "who an applicant knew counted at least as much as the skills he or she could demonstrate."

Patronage positions were used to reward political supporters as well as to fuel political machines, often requiring employees to contribute a fraction of their salaries (Hoogenboom, 1968). Although allocating federal jobs was, in principle, a prerogative of the Executive, members of Congress played a crucial role as "brokers" of these jobs within their states (Fish, 1905). Indeed, clerks in DC routinely returned to their home states to vote and work for the local politicians who had facilitated their appointment (Aron, 1987).[19]

While pressure for the adoption of a merit reform had been mounting since the 1860s, the exact timing of the passing of the Pendleton act is related to two political events. First, in July of 1881, newly elected president James A. Garfield was shot by a disappointed office seeker (Garfield would die by September). This assassination put civil service reform at the center of the political stage and provided reformists with a powerful example of the negative consequences of the spoils system. Soon after the assassination, in December of 1881, Democratic senator George H. Pendleton introduced a civil service reform bill. Second, Democrats took control of the House in March of 1882. Fearing that they would lose the 1884 presidential election, Republicans supported the bill, hoping that it would help protect Republican office-holders from politically motivated dismissals (Hoogenboom, 1959). In January of 1883, President Chester A. Arthur signed the Pendleton Civil Service Reform Act into law.

2.2 The Pendleton Act

Positions Subject to Exam. The act's main provision was to establish that certain "classified" positions within the executive branch of government would need to be filled through open, com-

reaucracy, for example long-term development (Evans & Rauch, 1999; Rauch & Evans, 2000; Chen *et al.*, 2020) and political outcomes (Theriault, 2003; Folke *et al.*, 2011; Bai & Jia, 2016; Bostashvili & Ujhelyi, 2019).

[18] Our work also complements a literature on electoral rules and political representation (Myerson, n.d.; Powell, n.d.; Dal Bó *et al.*, 2017; Thompson *et al.*, 2019).

[19] Hoogenboom (1959) describes how "if a civil servant came from a state where local and federal elections were held on different dates, he was often granted two paid leaves." Similarly, a recommendation letter for an applicant stated that: "[W]e need him [Smithers] in the present close contest, for a vote may secure a congressman and put our state fully in the line of progress. Our election is only a week off today, and Mr. S. is generally on the election board in my district. By his correctness of tallying last election he detected a fraud that would have lost us a county officer." (Aron, 1987)

petitive, and anonymously graded exams (Civil Service Commission, 1883).[20] The act divided the classified (that is, subject to exams) civil service into three branches: the "classified departmental service" for employees in the executive departments in DC, the "classified Customs Service" for Customs Service employees, and the "classified Postal Service" for postal workers.

The classified departmental service in DC–our main focus–was initially restricted to employees: (1) in clerical or technical positions, and (2) with annual salaries between $900 and $1800. In addition to exempting clerical workers with very low or very high salaries, there were two other exempted groups. First, the law exempted workers in hierarchical positions (bureau chiefs, elected officers, employees requiring Senate's confirmation, etc.). Second, it exempted workers employed "merely as laborers or workmen." Hence, in essence, the law targeted the "middle" of the state hierarchy while exempting the bottom and the top.[21]

Although, initially, the act affected only 10% of the federal bureaucracy, it authorized the President to include additional positions via executive order (Civil Service Commission, 1883). In our period of analysis (up to 1893), there were three changes affecting the classified departmental service in DC: the lower salary limit for clerical workers was decreased from $900 to $720 (in 1885), the $1800 upper limit was removed (also in 1885), and the lower salary limit was finally eliminated (in 1888) (Civil Service Commission, 1885). These changes, however, affected a very limited number of positions.[22] Hence, in practice, our difference-in-differences analysis compares workers in the positions that were subject to exams at some point from 1883 to 1893 (that is, workers in clerical and technical positions) to workers in the positions that were not (that is, workers in hierarchical positions and laborers/workmen), before and after 1883.

Figure 1 shows the total number of workers in the Executive Departments in DC, as well as the share who worked in positions that became subject to exams after 1883.[23] The total number of employees grew in the decade prior to the reform (reflecting the expansion of government functions in the post-Civil war era) but stabilized in the 1880s (Libecap & Johnson, 2007).[24] The share of workers in positions that became subject to exam remained stable throughout the period, fluctuating around 60%. Note that, while this share was relatively stable from 1883 to 1893 in the *Executive*

[20] Employees in the legislative and judicial branches of government were exempted from exams.

[21] The customs and postal classified services were initially restricted to customs-collection districts and post-offices with at least 50 employees, and to employees making no less than $900 within these offices.

[22] By 1883, 90% of the workers in clerical and technical positions (that is, in the positions that would eventually all be subject to exams) were employed in jobs paying between $900 and $1800. Of the remaining 10%, about half were employed in positions paying less than $720, about 30% were in positions paying between $720 and $900, and 20% were in positions paying more than $1,800. Hence, 50% (those paying between $720 and $900 plus those paying more than $1800) of the remaining technical and clerical positions had already been added to the classified service by 1885. Note that, as the data we use were published biennially, 1885 is in practice our first post-reform year. The other remaining 50% (those in positions earning less than $720) was nearly exclusively comprised of "assistant printers" in the *Bureau of Engraving and Printing*. This position only employed women and was not apportioned among states. As this position was not counted towards apportionment, we do not have information on the names of the employees appointed to this job. We exclude workers in this position from our baseline sample.

[23] We define a position as "subject to exams" if it became subject to exams at any point between 1883 and 1893.

[24] The jump from 1881 to 1883 corresponds to expansions of the Pension Office in the Interior Department, which added nearly 800 employees, and the Medical Department in the War Department, which added nearly 300. The 1891 increase corresponds to the addition of 2,500 workers hired temporarily to tabulate the 1890 census.

Departments in DC, it *did increase* over this period in the federal administration as a whole.[25]

Additional Provisions of the Law. In addition to introducing exams, the law established that positions in the classified departmental service had to be "apportioned" among states according to their population. Consequently, applicants to these positions were in practice competing only against other applicants from their own home state. In the analysis, we sometimes include workers' home-state fixed effects so as to shut down the effects of the reform that stem from apportionment-induced changes in employees' regional origins.

Although it changed the method used to fill certain federal positions, it is important to note that the act *did not* grant tenure to employees: "classified" workers remained open to the possibility of removal as administrations changed (Johnson & Libecap, 1994).[26] Later reforms introduced the notion that employees could only be removed due to "just causes" (Johnson & Libecap, 1994).

Exam Characteristics. The law established that exams had to focus on practical knowledge relevant to an applicant's future position rather than on formal academic training.[27] Applicants to the positions of copyist or clerk (the most common occupations in the classified service) were required to complete exams in four subjects: orthography, copying, penmanship, and arithmetic.[28] These subjects corresponded to the typical curricula taught in "common schools", namely the "three Rs" of **r**eading, w**r**iting and a**r**ithmetic.[29] Applicants to positions requiring technical or scientific knowledge were further required to take "supplementary" or "special" exams. Examples of such exams include the "meteorological clerk" exam in the Department of Agriculture and the "medical examiner" exam in the Pension Office. Panels (a) to (d) in Figure A4 show one example question for each of the four required subjects for applicants to the positions of clerk or copyist. Panel (e) shows an example question for applicants to the position of "meteorological clerk".

The emphasis on practical skills differs from civil service exams adopted in other countries (Hoogenboom, 1959). For instance, Grindle (2012) argues that a reform mandating exams in 19th-century England did not affect bureaucrats' social origins as exams were designed such that they would only be accessible to those with "elite educations at Oxford and Cambridge". In contrast, the US Civil Service Commission maintained that "a common school education was sufficient to pass examination" (Hoogenboom, 1959). Indeed, applicants with only a "common school" education regularly took (and passed) the exams. Figure A5 shows the total number of applicants and the share obtaining a passing grade, by educational background: Those with a common school education were the largest group of applicants, and 55% of them actually passed the exam.

[25]For instance, the Railway Mail Service was added to the classified service in 1889.

[26]"The power to remove for even the most partisan and selfish reasons remains unchanged" (Civil Service Commission, 1883). The only exception is that workers (in *all* positions, not just those in "classified" jobs) could no longer be removed for refusing to perform a political service or paying an assessment.

[27]The typical duties of a clerk entailed "routine, repetitive tasks", often involving recording and copying (Aron, 1987). Examples of such tasks include "note signing" (for clerks in the Treasury Department) and "writing and recording patents" (for clerks in the General Land Office within the Interior Department).

[28]The exam for clerks was referred to as the *general exam*, whereas the exam for copyists was referred to as the *limited exam*. The general exam could additionally include subjects such as bookkeeping and US history.

[29]The term "common schools" has often been used in reference to public elementary schools in historical US, although the term has had different meanings at different points in history (Goldin & Katz, 2003).

How did Applicants Learn about and Prepare for the Exams? The law required that exams be held throughout the country: Figure A7 shows the location of all exams from 1886 to 1893, with each circle drawn in proportion to the number of exams per location.[30] At the beginning of the year, the Civil Service Commission issued a pamphlet with exams' dates and locations (Civil Service Commission, 1886).[31] Moreover, this information was also regularly reported in local newspapers, as illustrated by the examples in Figure A9.[32]

Exam sample questions were available from the reports of the Civil Service Commission (Civil Service Commission, 1883-1893). Over time, these sample questions also became available from non-governmental, test-preparation books.[33] Moreover, applicants could also resort to receiving help from exam tutors. These tutors were available as early as 1883 (the year of the reform), as evidenced by the presence of newspaper ads offering their services (see Figure A10). This rapid market response to the reform is consistent with the high desirability of federal jobs in this period (Aron, 1987).

Appointing Procedure. Applicants who passed the exam were added to a register of eligible candidates.[34] On the opening of a vacancy, the Civil Service Commission produced a list of the *top four* candidates for the position, with the ranking determined solely on the basis of exam scores. For positions in the classified departmental service (which were subject to apportionment rules), these four names had to belong to applicants from states with the "strongest claim" to an appointment.[35] Appointing officers were then required to choose from these four candidates, drastically reducing hiring discretion.[36] An important deviation from meritocratic principles is that recruiters could ask for an employee of a *specific gender* (for instance, a "male clerk"). Indeed, 85% of the employees hired through exams in 1883–1893 were male.

How Attractive were these Positions? Clerkships in DC were "highly coveted and difficult to secure", and the number of applicants regularly exceeded the number of available positions by a large margin (Aron, 1987). Panel (a) in Figure A6 shows the yearly number of applicants to the classified departmental service. From 1883 to 1893, nearly 150,000 individuals completed an exam to join the classified civil service, of which 30,000 applied to the departmental service in DC. Panel (b) shows that the fraction of applicants to the departmental service who obtained a passing grade was fairly stable over our period, hovering around 65%. Finally, Panel (c) shows, out of all applicants with a passing grade, the proportion who were appointed to a position. By 1893, 23%

[30] For instance, there were a total of 286 exams between June 1892 and June 1893, with at least one exam per US state (Civil Service Commission, 1893, p.141, Table 1).

[31] Figure A8 shows a calendar of examinations for the period spanning January 1886 to June of 1887.

[32] For instance, searching for the expression "civil service examination" in `newspapers.com` and restricting the search to US newspapers yielded 700 results for 1883, 1,300 for 1884, and 2,600 for 1885.

[33] For instance, in 1897 the publisher *Hinds and Noble* published the book "How to Prepare for a Civil-service Examination With Recent Questions and Answers" (Leupp, 1898).

[34] Applicants who failed the exam were prevented from retaking it for six months (Civil Service Commission, 1885).

[35] If there were no candidates from the top priority state, the Commission produced a list of candidates from the state next in the list.

[36] This number was further reduced to three in 1888 (Civil Service Commission, 1886, p.128).

of those who had obtained a passing grade in the previous decade had received an appointment.[37]

Expected Effects of the Reform. It is unclear whether such a reform would improve or worsen the representation of individuals from disadvantaged backgrounds. On the one hand, the historical literature emphasizes how applicants connected to influential individuals were more likely to secure positions under the patronage system. For instance, Ziparo (2017)'s analysis of application files finds that, among women appointed to federal jobs in the 1860s, 71% had been recommended by a member of Congress.[38] To the extent that individuals from disadvantaged backgrounds were less likely to have these social connections, the reform could have improved their representation.

Moreover, access to a "common school" education (which would have, in principle, qualified applicants for passing the exam) was relatively widespread in 19th-century US, at least for white children. Panel (a) in Figure B8 shows school attendance rates, by age and parental wealth quintile, for white children in 1870. By 1870, nearly 80% of white children age 10 attended school. Although children with wealthier parents had higher attendance rates, even among children with parents at the bottom of the wealth distribution such rates were above 60%.[39] Indeed, among adults aged 20 to 50 in 1880 (the last pre-reform census), the literacy rate was above 90% (Ruggles *et al.*, 2021).

On the other hand, Libecap & Johnson (2007) emphasize how patronage was "viewed as a means of democratizing the government" as "anyone with the right political connections could obtain a government job, at least for a short while."[40] According to this view, the rise of "mass-based political parties" in 19th-century US required "great efforts and contributions by a host of devoted party workers," thus creating patronage opportunities for the "common person" (Greene, 1984). In contrast, a system based on competitive exams faced the risk of creating a "monopoly of office holding on the part of a particular class" (Civil Service Commission, 1884, p. 49).

3 Data

3.1 Federal Personnel Records

Our main source of data are the "Official Registers of the United States" (Department of the Interior, 1871-1893) (henceforth, the Registers). The Registers contain detailed information on the Federal workforce, including employees' names, birthplaces, state of residence at the time of appointment, position, unit, and compensation. We digitized the 12 registers published between 1871

[37] These figures imply an average of 14,000 yearly applicants in 1883–1893 (about 28 out of 100,000 people based on the US 1880 population). Of those who applied to the classified departmental service, 15% ended up receiving an appointment by 1893 (23% of 65%). As a comparison, the Indian Civil Service exam is completed by about one million applicants per year (75 applicants every 100,000 people). Of these, about 1,000 are appointed yearly (0.1%).

[38] Moreover, "of the successful applicants without congressional support, two had the support of president Andrew Johnson. Generals, police commissioners, governors, bankers, mayors, and clergymen all wrote women letters of recommendation for places in Washington, D.C." (Ziparo, 2017). Similarly, Aron (1987) describes a number of cases where employees secured their position through a family connection with a member of Congress.

[39] Panel (b) in Figure B8 shows school attendance rates by parental *occupation*. Attendance rates for children aged 10 were close to 70% for children of unskilled workers (the group with the lowest average attendance).

[40] A similar quote can be found in Johnson & Libecap (1994): "if anything, patronage was seen as promoting the ideals of equality and social mobility because it allowed the common person to fill public offices."

and 1893 (the Registers were published biennially), roughly ten years before and after the passage of the Pendleton act. Although the Registers include information on members of the Army and the Navy, we focus our analysis on *civil* servants.[41] Our data include information on approximately 450,000 employee-years. Of these, about 100,000 correspond to employee-years in the executive departments in DC, and about 40,000 correspond to *new hires* in these departments (our main focus).[42] Figure A3 shows an example page corresponding to the 1881 Register, listing employees in the *Internal Revenue Service* within the Treasury Department.

3.2 Measuring Employees' Social and Professional Backgrounds

Linking the Personnel Records to Population Censuses. We collected information on employees' socioeconomic backgrounds by linking the Registers to US population censuses, using name-based matching techniques (Abramitzky *et al.*, 2019). Specifically, we used workers' names, birthplaces, and approximate ages to link each of the 1871–1893 Registers to the 1850, 1860, 1870, and 1880 censuses.[43] Through this procedure, we obtained information about: (1) employees' socioeconomic backgrounds, including race, parental wealth, parental literacy, and parental occupations; and (2) employees' own occupations prior to joining the federal government. We provide further details and sensitivity checks in Online Appendix Section A. However, we note here that: (1) employees hired through exams are not more (or less) likely to be matched to the census (Table A1), (2) the results that *do not* require the linked data (that is, those in which we focus on the country-of-origin mix of employees) are very similar when estimated in this linked sample (Table A2), (3) the results are similar when we reweight the sample to account for differences in the matching probability across individuals (Table A3 and Figure A1), and (4) the results are similar when use alternative cutoffs to determine whether we deem an observation as a match (Figure A2).

Parental Wealth. The 1860 and 1870 censuses asked all household heads to report the total dollar value of their real estate and personal property.[44] We use the combined value of real estate and personal property to rank households in the *national* (although our results are similar if we use state-specific ranks) wealth distribution, separately by census year and age of the household head.[45] For those employees for whom we observe parental wealth both in 1860 and 1870, we use

[41]"Postmasters" are a large group of civil servants whose data we have not digitized. We chose not to digitize their data as the Registers include limited identifying information about these workers. For instance, they do not include their birthplaces, and in most cases they only include first name initials rather than a complete first name.

[42]Although our data enable us to observe the same employee over multiple years, throughout the analysis we only include them the first time they show up in the data (that is, we focus on the flow of *new hires*).

[43]Specifically, we assumed workers would have been between the ages of 18 and 60 at the time of their employment in the civil service. We chose these years since 1850 is the first US population census to list persons individually, and there are no surviving records for the 1890 census.

[44]The 1850 census asked about real estate property but not about personal property. Enumerators in 1860 and 1870 were instructed to collect personal property information "inclusive of all bonds, stocks, mortgages, notes, livestock, plate, jewels, or furniture; but exclusive of wearing apparel." The 1880 census did not include either of these questions.

[45]A complication with computing such rank is that the 1860 census did not list the Black enslaved population but the 1870 census (which took place after emancipation) did. Because the formerly enslaved population (about 12% of the 1870 US population) owned little wealth, white household heads observed in 1870 would mechanically tend to have higher ranks than those observed in 1860. To avoid this issue, we construct ranks that are based just on the *white population*. In

the average rank across census years as our baseline measure.

Note that, to observe parental wealth (or, more generally, any parental characteristic), the nature of our data requires that we observe employees coresiding with their parents in the census. Hence, to minimize biases due to selective coresidence at later ages, whenever we focus on parental characteristics we restrict the sample to employees whom we observe with their parents in the census at the age of 17 or less (and *prior* to them joining the civil service).[46]

Father's Occupation and Parental Literacy. We split father's occupations into five categories: professional, non-professional white-collar, farmer, skilled blue collar, and unskilled.[47] For those employees for whom we observe their father's occupation in more than one census, we calculate the fraction of census years that their father spent in a given occupational category.[48] Finally, we also use the literacy information to construct indicators of whether employees' parents were literate.

Nativity Status and Race. We observe workers' birthplace and race, as well as the corresponding information for their parents. We use this information to construct indicators of whether workers are foreign born, whether both their parents are foreign born, and whether they are white. Note that, while we are able to observe race and birthplace for all employees that we match to at least one census, we only observe *parental* birthplace for those that we match to a census in which they were coresiding with their parents. Similarly, as the Registers include information on workers' birthplaces, we can also use these data directly (without linking to the census) when we investigate if the reform changed the likelihood that workers would be foreign born.

Summary Measures of Employees' Social Background. We compute two summary measures of employees' socioeconomic backgrounds. These measures are constructed such that a lower value corresponds to individuals from more disadvantaged backgrounds. First, we follow Kling *et al.* (2007) and compute a "summary index" equal to the unweighted average of the following standardized variables: parental wealth rank, an indicator of whether a worker's father was literate, an indicator of whether a worker's father was a professional, an indicator of whether a worker's father *was not* an unskilled worker, an indicator of whether a worker's parents were US born, and an indicator of whether the worker was white.[49] Second, we use factor analysis to com-

addition, we base the rank on households with at least one child–as this is the relevant group for our intergenerational analysis. By 1860, 87% of white household heads with at least one child had positive wealth, whereas that proportion was 80% in 1870. A related issue is that slave-owning families saw a decline in their wealth after emancipation (Ager *et al.*, 2019). Hence, families observed in 1870 would tend to be poorer than those observed in 1860. However, our results are similar if we exclude Southern employees from the sample (where most of these families resided).

[46] Among employees whom we observe at the age of 17 or less, 80% have a father present in the census.

[47] This classification corresponds to the classification using five occupational categories in Long & Ferrie (2013). We focus on father's occupations as few mothers worked outside of their households in this period. Professional occupations are those with a value below 100 in the 1950 Census Bureau occupational classification system. Examples of such occupations include accountants, lawyers, and teachers. Non-professional white-collar occupations are those with a value between 200 and 500 (for example stenographers, typists, and secretaries). Farmers are those with a value of 100. Skilled blue-collar are those with values between 500 and 700 (examples include carpenters and electricians). Finally, unskilled workers are those with a code above 700 (examples include laborers and housekeepers).

[48] For instance, when we focus on whether an individual's father had a professional occupation, we assign a value of 0.5 to those cases in which the father is listed as having a professional occupation in one census but not in the other.

[49] The variables are standardized by subtracting the control group mean and dividing by the control group standard

pute the first principal component of the same set of variables, which we then normalize to have a mean of zero and a standard deviation of one.[50]

Focusing on these summary measures offers two main advantages relative to focusing on individual characteristics. First, as we observe a large number of workers' characteristics, using a summary measure minimizes the risk of overrejecting the hypothesis that the reform did not affect workers' socioeconomic backgrounds. Second, using an index "improves statistical power to detect effects that go in the same direction" (Kling *et al.*, 2007).

Employees' Professional Backgrounds. We observe workers' occupations *prior* to joining the civil service.[51] Whenever we link an employee to multiple censuses, we focus on their most recent pre-civil service occupation. When we focus on workers' own prior occupation, we restrict the sample to workers who were at least 25 years old at the time we observe them in the census (so as to enable occupations to better reflect workers' educational and professional attainment).

3.3 Identifying Employees Appointed Through Exam

We combine the linked personnel records with data from the Civil Service Commission reports (Civil Service Commission, 1883-1893). These reports include a list of all employees hired through exams in the classified *departmental service* in DC. These lists were collected by the Civil Service Commission with the goal of keeping track of the apportionment of positions across states, and include employees' names, home state, initial department and compensation, examination taken, and appointment date.[52] Using this list, we can precisely identify which employees were hired through exams, as well as the exact exam that they took. Although these lists cover all hires to the classified departmental service, they *do not* cover employees in the classified customs and postal services (as these positions were not apportioned). Figure A11 shows an example page which lists employees appointed to the classified service in 1883.

In addition to including a list of the employees hired through exams, the reports include a detailed list of the *positions* that were subject to examinations in each of the executive departments. Figure A12 shows an example page listing the positions subject to exam in the Treasury Department. These data enable us to precisely identify the set of "treated" positions.

Summary Statistics and Sample Size. Table B2 shows summary statistics for employees in our baseline sample, separately based on whether or not they were appointed through an exam. Employees appointed through exams came from wealthier families, were more likely to have a father with a professional occupation, less likely to have an unskilled father, less likely to be foreign born or have foreign-born parents, and more likely to be white.

Finally, note that the number of observations varies depending on the specific characteristic we consider. For instance, while we only observe parental wealth for those employees that we

deviation, so that each component of the index has a mean of zero and a standard deviation of one for the control group.

[50] There is a 0.9 correlation between both measures, so, for brevity, we mostly focus on the Kling *et al.* (2007) index.

[51] Unfortunately, censuses prior to 1940 do not include information on earnings or years of schooling.

[52] The one exception is that, as described above, these lists do not include employees hired for the position of printing assistant in the Bureau of Engraving and Printing (as these positions were not apportioned).

find as children in the 1860 or 1870 censuses (as these were the only two censuses that included this variable), we observe parental occupations and parental birthplace for employees we find as children in any of the censuses of the period. Hence, we have more observations for the latter two characteristics than for parental wealth. Similarly, we observe race for any employee that we match to at least one census (regardless of the age at which we find them as observing this variable does not require observing individuals coresiding with their parents). Our findings on parental occupations and parental birthplaces are nevertheless similar if we restrict the sample to employees for whom we also observe parental wealth (see Tables B4 and B5).

4 Empirical Strategy

Our main goal is to assess the extent to which selecting employees through competitive exams changed bureaucrats' socioeconomic backgrounds. To do so, our empirical strategy compares the characteristics of employees hired before and after the reform (first difference), in positions exempted and non-exempted from exams (second difference). We estimate:

$$y_{ipt} = \alpha_p + \alpha_t + \beta Exam_p \times After_t + \gamma X_{ipt} + \epsilon_{ipt} \quad (1)$$

where y_{ipt} corresponds to a characteristic of employee i in position p in year t, α_p are position fixed effects, and α_t are hiring-year fixed effects. A position is defined as the combination of an occupation, a compensation, a bureau and a Department–for instance, *clerk, $1200, Pension Office, Interior Department*. By including position fixed effects, our analysis compares workers hired to perform the *same* job in the *same* unit, some of whom were recruited through patronage and some of whom were recruited through exams. Moreover, the inclusion of hiring-year fixed effects enables us to net out changes in worker characteristics stemming from aggregate changes in the economy (for instance, changes in the relative attractiveness of the public sector). Our interaction of interest is $Exam_p \times After_t$: $Exam_p$ is one for employees in one of the "treated" positions (that is, those that became subject to exams), and $After_t$ is one for employees hired after the reform. Finally, in some specifications we include additional control variables (described by below) as captured by X_{ipt}. Throughout the analysis, we cluster standard errors at the level of the position.

In addition, we also estimate event-study specifications of the form:

$$y_{ipt} = \alpha_p + \alpha_t + \sum_{t=1875}^{1893} \beta_t Exam_p \times \alpha_t + \gamma X_{ipt} + \epsilon_{ipt} \quad (2)$$

where the β_t coefficients describe the evolution in the characteristics of employees hired in positions subject and non-subject to exams during our sample period. The omitted category is workers hired in 1873, the first year in the data for which we can identify newly hired employees.[53] This specification enables us to investigate the extent to which the reform had different effects in

[53] While we have collected personnel records starting in 1871, 1873 is the first register year for which we know whether employees are *new hires* (based on comparing the list of employees in 1873 to the list in 1871).

the shorter (immediately after its passage) and longer (10 years after its passage) term.

As described above, the reform established that positions in the Departmental Service in DC had to be apportioned across states. Because this change could have by itself affected workers' characteristics (to the extent that it affected their regional origins), in our preferred specification X_{ipt} includes workers' home-state fixed effects. By including these fixed effects, we shut down the effects of the reform that stem from compositional changes in bureaucrats' regional origins. In practice, the inclusion of such fixed effects has only modest effects on our estimates.[54]

Challenges to Identification and Tests of the Identification Strategy. Our control group is comprised of workers both in low- (such as laborers) and high- (such as unit chiefs) pay positions. One concern is that the characteristics of such workers would have been on a different trend relative to those of workers in the positions subject to exams. This might have been the case, for instance, if the relative attractiveness of the public sector was differentially changing for workers in different parts of the skill distribution.

To address this concern, Table B3 presents, for each of our main variables of interest, F-test statistics corresponding to the hypothesis that all pre-reform event-study coefficients are equal to zero. The estimates correspond to our preferred specification, which includes home-state fixed effects in addition to the baseline variables. The table shows that, regardless of the outcome we consider, we do not reject such a null hypothesis. In Section 5, we also present graphic evidence consistent with the common trends assumption. Finally, subsection 5.1 shows the robustness of our results to using alternative definitions of the control group (including an exercise in which we simply compare workers in "treated" positions before and after the reform).

A second concern is that, to the extent that appointing officers wished to retain hiring discretion, the reform might have incentivized hiring in the exam-exempted segments of the bureaucracy. In this case, our effects could stem from changes in the control group rather than by changes in the characteristics of exam-appointed employees. Indeed, in Moreira & Pérez (2021) we document such a response in the context of the classified Customs Service: requiring that employees making $900 or more a year were hired through exams led to a near *doubling* in the share of workers making less than this cutoff.

There are three reasons why this concern is less likely to be relevant in our context (i.e. the classified departmental service). First, the historical literature suggests that such manipulation was unlikely to occur for positions in the executive departments in DC as these positions were under tighter control from the Civil Service Commission.[55] Indeed, Figure 1 shows that, in the classified departmental service, the share of positions that would have been subject to an exam remained

[54] Although employees had to provide proof of residence, a concern is that they had incentives to claim that they resided in a state with fewer appointed employees so as to increase their appointment chances. However, our results are similar if we use workers' *birthplace* fixed effects (see Table B12).

[55] For instance, Civil Service Commission (1890) writes that "Turning to the custom-houses, the Commission is able to present much less satisfactory tables. The classification of the Customs Service has always been very imperfect. It has been classified by salary rather than by employment, and has been possible to take the employees out of the classified grades by lowering their salaries or by changing their designations."

relatively flat (at about 60%) over our period.[56] Second, when we plot the data separately for the control and treatment groups, there is little indication of a sharp post-reform change in the characteristics of the control group (Figure B1). Indeed, consistent with this stability, our results are similar regardless of whether we implement our baseline difference-in-differences strategy or we perform a simple before and after comparison of the backgrounds of employees in "treated"' positions (Table B13). Third, our results are similar when we use alternative control groups comprised of workers in units in which no employee was subject to the reform (and where these spillovers were hence less likely to occur) –for instance, workers outside of DC or workers in the legislative and judicial branches of government (see Figure B2).

Sample Restrictions. In our main analysis, we restrict the sample to workers in the Executive Departments in DC. We do so because, for these workers, we have exact information on which of them were appointed through an exam (rather than inferring this information based on their position and estimated hiring date).[57] We note, however, that our results are similar if we include workers outside of DC or outside of the Executive Departments (for instance, in the Legislative or Judicial Departments) in our control group, or workers in the classified Customs Service to our treatment group (see Figures B2 and B3). In addition to focusing on the Executive Departments, we also restrict our baseline sample to *male* employees. We do so for two reasons. First, as most women changed their last name upon marriage, it is challenging to track women across sources using their names.[58] Indeed, our matching rates are lower for females (Figure A14). Second, nearly 85% of the employees appointed through exams were male, so restricting the sample to male employees further improves the comparability of the treatment and control groups. Our main results are nevertheless similar when we add females to the sample (see Figures B2 and B3). Table B1 illustrates the construction of our baseline sample.

5 Main Results: Exams and Bureaucrats' Socioeconomic Backgrounds

In this section, we ask if the reform facilitated or impeded the access of individuals from disadvantaged backgrounds to government jobs. We focus on parental wealth, parental occupations, parental literacy, worker's countries of origin, and race.

Summary Index of Social Background. We first investigate the effects of the reform on the Kling *et al.* (2007) summary index of employees' socioeconomic background. This index aggregates information on parental wealth, parental occupations, parental literacy, nativity status, and race, and is constructed such that a lower value corresponds to individuals from lower-SES backgrounds.

Panel (a) in Figure B1 shows the average of this index for newly hired workers, separately

[56] As described above, the 1891 decrease in the share of covered positions is driven by the addition of 2,500 workers in the Census office. These workers were hired temporarily to tabulate the 1890 census and were exempted from exams.

[57] This restriction excludes workers in the Executive departments outside of DC (such as those in the Postal and Customs services), workers in the Judicial and Legislative departments, and workers in miscellaneous government agencies not affected by the reform.

[58] For instance, 40% of women aged 18 to 50 with an occupation in the 1880 census were either married or widowed.

based on whether workers were employed in positions subject or not subject to exams. The figure shows that, throughout the period, workers in positions subject to exam had higher values of the index than those in exempted positions. However, this gap appears to increase after the reform.

Table 1 estimates the specification in equation 1 and confirms that the reform was associated with an increase in workers' summary index of socioeconomic background. Specifically, Column 1 shows a 0.18 standard deviation increase in the value of such index.[59] Column 3 shows a similar increase (of 0.29 standard deviations) if we instead use the first principal component of the same set of characteristics included in the Kling *et al.* (2007) index. The estimates are similar regardless of whether or not we include fixed effects for workers' state of residence (odd versus even columns), suggesting that the effects are not driven by apportionment-induced changes in workers' regional origins.

Figure 2 shows the corresponding event-study estimates, again focusing on the Kling *et al.* (2007) index. The pre-reform event-study coefficients are sometimes positive and sometimes negative, and we do not reject the hypothesis that they are all jointly equal to zero (p-value: 0.33, see Table B3). In contrast, all of the post-reform coefficients are positive and they are jointly statistically significant (p-value<0.01, see Table B3).

The figure suggests a rapid increase in the index after the reform, with the estimates then declining in size to a value of around 0.12 standard deviations. One likely explanation for the initial jump and subsequent leveling of the effects is that applicants from disadvantaged backgrounds might have required more time to "catch-up" with the contents of the exam. Such a catching-up mechanism can rationalize why the increase in applicants' socioeconomic status was stronger immediately after the reform. We emphasize, however, that our difference-in-differences estimates are not driven by this initial jump–in fact, they are not driven by any particular post-reform year, see Figure B6 which shows the robustness of our difference-in-differences estimates to excluding one post-reform year at a time.

To benchmark these magnitudes, note that the pre-reform median value of the index was 0.25 among workers employed as "clerks" and -0.12 among those employed as "laborers". Hence, the increase in the summary index corresponds to about half of the pre-reform gap between clerks, a white-collar occupation with a pre-reform median annual compensation of $1400, and laborers, a blue-collar occupation with a pre-reform median annual compensation of $660.

Parental Wealth. We next investigate the consequences of the reform for the different components of the index, starting from parental wealth ranks. Panel (b) in Figure B1 shows average parental wealth ranks for workers in positions subject and non-subject to exams, from 1873 to 1893. Similar to what we observe for the summary index, the figure suggests a differential post-reform increase in the parental wealth ranks of workers in positions subject to exams.

Columns 1 and 2 in Table 2 confirm that the reform led to an increase in employees' family wealth ranks. Specifically, employees hired through exams came from families that were 6.2 percentile ranks higher in the national wealth distribution, slightly above a 10% increase.

[59] As described in Kling *et al.* (2007), the point estimates show "where the mean of the treatment group is in the distribution of the control group in terms of standard deviation units."

In columns 3 to 6, we compute separate ranks for personal property and real estate wealth – rather than a single rank based on their combined value. Differences in real estate wealth may simply reflect regional differences in home-ownership rates (rather than true differences in parental resources). It is reassuring that the average rank increases for both measures and particularly so for personal wealth: Workers appointed through exams came from families that were 7 percentile ranks higher in the distribution of personal property wealth and 4 percentile ranks higher in the distribution of real estate property–although the increase in real estate wealth rank is not significant once we add home-state fixed effects.

Figure 3 shows event-study estimates of the effects of the reform on parental wealth ranks. The pre-reform event-study coefficients are relatively small and we do not reject the hypothesis that they are jointly equal to zero (p-value: 0.39, see Table B3). In contrast, the post-reform event-study coefficients are all positive and are jointly statistically significant (p-value<0.01, see Table B3). The estimates suggest a rapid increase in parental wealth following the reform. However, unlike when we focus on the summary index, the year-by-year estimates are less stable: they are the largest in 1885 and the smallest in 1887 and 1893.

Who Gained and Who Lost Access? The reform increased employees' average parental wealth ranks. Such increases could be compatible with increases in the representation of the middle class at the expense of the children of the poor, or with increases in the representation of the upper class at the expense of the middle (or by some combination of the two).

To investigate which groups increased and which groups decreased their representation, we split individuals based on the wealth quintile of their parents. Panels (a) and (b) in Figure 4 show, for newly hired employees in positions that became subject to exams after 1883, their distribution across family wealth quintiles in the pre- and post-reform periods. Panels (c) and (d) show the same distribution but for employees in positions that did not become subject to exams.

Among workers in positions subject to exams, those who grew up with families in the top quintile were overrepresented prior to the reform (they accounted for about 35% of workers). However, there were small differences in the relative representation of individuals from the bottom four quintiles: each of these groups accounted for about 15% of workers prior to the reform. After the reform, in contrast, we observe a sharp increase (from 15 to 25%) in the proportion of workers from the 60–80 quintile. This increase seems to come mostly at the expense of the bottom quintile.

Among those in positions exempted from exams, both the top and the bottom family wealth quintiles were overrepresented in the pre-reform period (Panel (c)). This bimodal distribution likely reflects the fact that exempted positions included both leadership (such as bureau chiefs) and low-pay positions (such as laborers). The overrepresentation of both the bottom and the top of the parental wealth distribution remained similar in the post-reform period (Panel (d)).

Panel (e) in Figure 4 confirms this pattern when we estimate equation 1 using as dependent variables indicators for belonging to each quintile of the parental wealth distribution. First, we find no change in the likelihood that an employee would belong to the top 20%. Second, we find an increased representation of families between the 60 and 80 percentiles of the wealth distribution, which comes at the expense of families in the bottom two quintiles (particularly the bottom 20%).

Father's Occupations and Parental Literacy. Table 3 shows that the reform increased the share of employees whose father had a higher-status occupation. First, employees hired through exams were 2.4 percentage points less likely to have a father with an unskilled occupation (nearly a 30% decline). Indeed, combining all blue-collar occupations (skilled blue collar plus unskilled) into a single group, we observe a 6 percentage points decline in the likelihood of having a father in this category (see Table B6). Second, exam-based hires were 5 percentage points more likely (relative to a baseline of 11%) to have a father with a professional occupation. Finally, there is also an increase in the share of employees with a farmer father, although this effect is smaller (and loses statistical significance) once we include fixed effects corresponding to workers' home states.

Similarly, Table B7 shows that individuals appointed through exams were 2.6 percentage points more likely to have a literate father (relative to a sample mean of nearly 93%), although there is no such a gap when we focus on employees' mothers.

Country of Origin and Race. Table 4 shows that the reform reduced the representation of immigrants (and their children) in government jobs. Columns 1 and 2 show that employees appointed through exams were 4 percentage points less likely to be foreign born, nearly a 40% reduction. This result could reflect the fact that immigrants, who had less exposure to US education, might have been at a disadvantage when completing the exams. The decline in the share of immigrants, however, does not seem to be simply driven by a lack of familiarity with English: Table B8 shows a large decline in the share of immigrants from *English-speaking* countries.

An advantage of using immigrant status as an outcome is that it does not require linking observations to the census (as birthplace was directly reported in the Registers). Hence, we can assess the sensitivity of the results to using either the full or the linked sample. Table A2 shows a similar decline in the likelihood that an employee would be foreign born regardless of which sample we use; if anything, the decline is larger when estimated in the full (non-linked) sample. Moreover, the difference between the estimates becomes even smaller (Table A2) as we reweight the linked sample to account for differences in the likelihood of matching an observation to the census.

In columns 3 and 4 of Table 4, we instead focus on the likelihood that an employee would have been the *child* of an immigrant. Unlike migrants themselves, their children were likely exposed to education in the US, perhaps limiting the disadvantages observed in the first generation. However, we find that exams also sharply reduced their representation: there is a 7 percentage points decline in the proportion of children of immigrants (relative to a control group mean of about 20%).

Finally, in columns 5 and 6 we investigate if the reform changed the racial mix of government employees. The dependent variable in these columns is an indicator that is one if an employee reported being white in the census. Although the reports of the Civil Service Commission argue that the reform increased the representation of African Americans, we find limited evidence that this was the case: the point estimates are very close to zero and enable us to rule out small changes in employees' racial mix.[60] This finding is perhaps not surprising in light of the fact that African

[60] The reports claim that "It is noticeable that a much larger proportion of colored people receive appointments under

Americans had limited access to educational resources but also represented a very small proportion of federal workers prior to the reform.

Heterogeneity by Type of Exam and Position. Although most employees were hired in relatively non-technical mid-tier positions such as copyists or clerks (whose exams only required a "common school" education), some workers were hired for more "elitist" positions. For example, for technical positions whose exams would have required more specialized knowledge, or for positions paying particularly high salaries. Indeed, one possibility is that the increased elitism we document was specific to these more "elitist" positions. To assess this possibility, we next investigate if the effects of the reform varied depending on the position to which a bureaucrat was appointed. To do so, we first estimate:

$$y_{ipt} = \alpha_p + \alpha_t + \beta Clerk_p \times After_t + \beta Technical_p \times After_t + \gamma X_{ipt} + \epsilon_{ipt} \quad (3)$$

where $Clerk_p$ is one if employee i is listed as having taken either the clerk or the copyist exam, and $Technical_p$ is one if the employee is listed as having taken one of the various technical exams (for instance, the exam for meteorological clerks in the Department of Agriculture).

Table B9 shows a very similar increase in the summary index of employees' socioeconomic backgrounds, regardless of the type of positions they were appointed to: employees hired through exams, both in more and less technical positions, were of higher social class than those hired through patronage. These findings suggest that the increase in applicants socioeconomic status was not driven by those positions which required advanced education.

In Table B10, we perform a similar analysis but splitting workers between those appointed into below and above median paying positions. The table shows an increase in the socioeconomic status of individuals appointed to both types of positions, although the point estimates suggest a larger increase among those appointed into the higher paying jobs.

Alternative Explanations. We next consider a number of alternative explanations, other than *the use of exams per se*, for the observed increase in applicants' socioeconomic status. First, since exams were held across the country, the reform might have facilitated the access of workers from a broader set of locations to government jobs. Second, unlike in the case of recruitment through patronage, exam dates and locations were widely and publicly advertised. Hence, an applicant from a rural area or who lived far from DC might have been more likely to be aware of government jobs than in the pre-reform era.

An implication of both of these channels is that our results should be driven by changes in employees' geographic origins. To assess this possibility, we investigate how our results change as we include: (1) birthplace, (2) childhood state, or (3) childhood state by urban/rural fixed effects. Intuitively, if the effects of the reform stemmed from changes in employees' geographic origins, we should observe muted effects once we compare individuals who grew up in similar locations.

Table B12 shows that changing employees' geographic origins does not appear to be a quantitatively important channel for explaining our results: we observe similar increases in the summary

the civil-service law than under the old patronage system." (Civil Service Commission, 1891) index of social status when looking *within* childhood locations of residence (or within birthplaces).

The third alternative explanation is that applicants from less privileged backgrounds might have fewer resources to adapt to *any new* recruitment system, irrespective of whether this new system involves an exam or not. If this "disruption" channel explained our results, then the effect of the reform should have been short-lived. However, Figure 2 shows that the reform's effect persisted: By 1893 (10 years after the reform), workers appointed through exams were still of higher social status than those hired through patronage.[61]

The final alternative explanation is that the effects that we capture are not driven by the reform itself but rather by the transition to a Democratic administration in 1884. After a long period of Republican dominance (starting with Grant's presidency in 1869), Democrats recovered the presidency in 1884, one year after the reform. Although for such a transition to explain our results it would need to be the case that it differentially affected the backgrounds of workers in positions subject to exam, we can directly show that our results are not driven by it. Specifically, as the presidency went back to a Republican in 1888 and then back again to a Democrat in 1892, we can investigate whether the effects of exams depended on whether the President was a Democrat or a Republican. Table B11 shows that the increase in workers' socioeconomic status occurred *both* under Democratic and Republican presidencies, although the point estimates are larger while Democrats were in power. Moreover, the effects are not driven by any particular presidential transition: they are of similar magnitude when we exclude one post-reform year at a time (Figure B6).

Summary of Results. The reform increased the socioeconomic status of government workers: they had higher levels of parental wealth, were less likely to have a literate father, more likely to be the children of professionals, and less likely to be the children of immigrants (or immigrants themselves). This increase in socioeconomic status occurred immediately after the passing of the reform and persisted for at least 10 years. Moreover, the increased elitism is observed even when restricting the comparison to workers with similar geographic origins, and both during Democrat and Republican presidencies. These additional findings suggest that the increased elitism was driven by the use of exams *per se* rather than by other features of the reform.

5.1 Robustness Checks

We next show that our results are robust to: (1) including additional control variables to account for potential time-varying shocks, (2) using alternative definitions of the control and treatment groups, (3) using alternative definitions of which workers are considered new hires, (4) implementing a randomization inference approach, (5) features of the linking strategy, and (6) adjusting for the fact that we cannot observe parental characteristics for individuals who moved to the US as adults.

Time-Varying Shocks and Additional Control Variables. By establishing that classified positions had to be apportioned across states, the reform increased the representation of workers from certain states and decreased the representation of workers from others. Although we include

[61] Note that introducing exams has *per se* potentially important dynamic implications, which we discuss in Section 6.1

home-state fixed effects to account for this channel, a concern is that the labor market in different states might have been on different trends, leading to differential changes in the selection of workers interested in government jobs. In this case, the effects we capture would not be those of transitioning from patronage to exams but rather the effects of increasing the representation of certain states. To address this concern, in Figures B2 and B3 we show that our results are similar when we include home-state times hiring-year fixed effects.

In Figure B5, we show that our results are similar when we exclude employees from one Executive Department at a time from the sample. The y-axis in this figure shows our estimated effects, whereas the x-axis shows the excluded department. This finding rules out the possibility that our results were driven by a change concurrent to the reform and taking place only in a specific department. Moreover, our results are similar when we include department times hiring-year fixed effects (Figures B2 and B3).

Alternative Samples and Definitions of the Control Group. In our baseline analysis, the control group is comprised of employees in the Executive Departments in DC who worked in positions exempted from exams (that is, positions either at the bottom or the top of the state hierarchy). Figures B2 and B3 show that our results are robust to using alternative control groups. First, they are similar when we use a control group constituted by either: (1) only bureaucrats at the bottom of the state hierarchy, or (2) only bureaucrats at the top. Second, they are similar when we drop workers making more than \$3000 or less than \$600 from the control group (so as to increase the treatment-control group comparability). Third, they are similar when we add workers outside of DC to the control group.[62] Fourth, they are similar when we add workers who were employed in DC but worked outside of the Executive departments (in units in which no employee was affected by the reform). Fifth, the results are similar when we add female employees to the sample (both to the treatment and the control groups).

Finally, Table B13 shows that our results are also similar when we implement a simple before and after comparison of the characteristics of employees in positions subject to exam. In this table, we restrict the sample to employees in the "treated" positions and compare their average characteristics before and after the reform (net of position fixed effects).[63]

Adding Customs Service Employees to the Treatment Group. In our baseline analysis, the treatment group is comprised of workers in the Executive Departments in DC who worked in positions subject to exams. However, our data include another group of individuals appointed through an exam: namely, employees in the classified Customs Service. Specifically, the classified Customs Service was initially restricted to customs-collection districts with at least 50 employees, and to employees making no less than \$900 within these districts.[64]

[62] When we do so, we add place of employment to our definition of a position.

[63] Specifically, we report the estimated value of β from the following equation:

$$y_{ipt} = \alpha_p + \beta After_t + \gamma X_{ipt} + \epsilon_{ipt}$$

using the sample of employees in "treated" positions.

[64] As discussed above, we do not have data on employees in the classified *Postal* Service.

We chose to focus on employees in the Executive departments in DC for two reasons. First, we have exact information on which of them were appointed through an exam. Second, Moreira & Pérez (2021) show that, in the Customs Service, the reform induced a distortion in districts' personnel structure. In particular, it caused a sharp increase in the share of workers making less than $900 (who were exempted from exams). Such a response complicates the interpretation of our difference-in-differences design, as it implies that the reform potentially affected *both* the treatment and the control positions.

Nevertheless, Figures B2 and B3 show that our results are robust to adding Customs Service employees to the treatment group. In this figure, we classify these employees as having been appointed through an exam based on their estimated hiring date, position, and collection district.

Alternative Definitions of a New Hire. In our baseline, we classify a worker as a new hire if there is no worker listed in the previous register with the same name, birthplace, appointment state, and Department.[65] A concern with this approach is that errors in the registers (or in our digitization) might lead us to deem a worker as a new hire even if that worker was already employed by the government. Although it is unclear why such errors would rationalize the effects that we observe, in Figures B2 and B3 we adopt a more stringent definition of a new hire. Specifically, we define a worker as a new hire if there is no employee in the previous register with a name within a 0.1 Jaro-Winkler string distance of their name (rather than using exact names) and regardless of birthplace.[66] The figures show that using this alternative definition, which classifies fewer employees as "new hires", yields results that are similar to the ones that we obtain with our baseline definition.

Inference. Figure B4 shows that our results are robust to implementing a randomization inference approach. To do so, we randomly classify a group of workers (of equal size of our actual treatment group) as having been hired through exams. We then estimate the "effects" of the reform using these placebo treatment groups, repeating the exercise 1,000 times. Reassuringly, these placebo estimates are centered around zero and significantly smaller than the actual estimates.

Linking Strategy. First, Figure A2 shows that our results are similar when we use more or less conservative cutoffs for deeming an observation as a match. Second, Table A3 and Figure A1 show that our results are similar if we reweight the data to account for differences in the observable characteristics of matched and non-matched employees.[67] Finally, as discussed above, our results focusing on the likelihood that employees would be foreign born (which do not require linked data) are similar when estimated using either the linked or the non-linked samples (Table A2).

Missing Data on Migrants' Parental Characteristics. Our information on parental character-

[65] We implement this approach because we do not have direct information on worker's hiring dates and hence need to infer them from comparing adjacent Registers.

[66] The Jaro-Winkler string distance is based on the number of edits that would be required to turn one string into another. The measure takes a value of zero for two identical strings and a value of one for two strings with no common characters.

[67] To do so, we estimate a probit model of the likelihood of matching to an observation in the census, including as independent variables indicators for workers' birthplace, state whence appointed, department, position, compensation, and register year. We then reweight the data based on the inverse of the estimated matching probability.

istics is based on observing children living with their parents in the US census. Hence, we do not have this information for those foreign-born employees who moved to the US as adults. This omission could be problematic because the reform reduced immigrants' representation (Table 4).

To deal with this issue, we implement an exercise, in the spirit of Lee (2009), in which we bound the bias that might result from this omission. Specifically, we reestimate our main specification in an expanded sample in which we impute foreign-born employees three alternative values of the parental summary index: (1) the 10th percentile of the value observed among all employees of the same occupation in the pre-reform period, (2) the 50th percentile, and (3) the 90th percentile. This expanded sample is constructed such that the fraction of immigrants in a given year is the same as in the actual non-linked sample.

Table B14 shows that, even under the extreme assumptions that immigrants whose parental information we do not observe were at the 10th or the 90th percentile of the summary index, there would have still been an increase in this index due to the reform.[68]

6 Why did Exams Decrease the Representation of Individuals from Disadvantaged Backgrounds?

Our interpretation of the findings is that, by increasing the relative importance of education in the hiring process, the reform improved the chances of "educated outsiders" (individuals with high education but limited connections). Because middle class applicants were overrepresented in this group, the reform increased their representation. We first provide a conceptual framework that illustrates this interpretation. We then show evidence consistent with it and discuss additional hard-to-measure channels through which the exam might have affected workers' characteristics.

6.1 Conceptual Framework

Assume that obtaining a government job depends on applicants' education ("e") and connections ("c"). We consider education broadly, including applicants' stock of knowledge as well as their ability to study for the exam. Connections could also be of various types, including family (for instance, being related to a member of Congress) and political (for instance, having worked for the incumbent party) connections. Further, assume that e and c are potentially correlated with applicants' social class (s).

Applicants are hired if they are in the top $l\%$ of candidates in terms of their combined values of e and c; that is, if:

[68] This possibility is unlikely for two reasons. First, the US was wealthier than the European countries that accounted for most immigration in this period. Hence, in the absence of migrant selectivity, migrants would have come from less privileged backgrounds than the US born. Second, the evidence on immigrant selectivity during this period shows that migrants tended to be drawn from the lower (or at most intermediate) social classes in their origin countries (see Abramitzky *et al.* (2012) and Abramitzky & Boustan (2017)). These two factors suggest that, if we could observe the parental characteristics of foreign-born employees, we would find an *even larger* increase in elitism.

$$\alpha e + (1-\alpha)c > F^{-1}(1-l) \tag{4}$$

where F^{-1} is the inverse cdf function corresponding to the distribution of $\alpha e + (1-\alpha)c$.

We interpret the reform as an increase in the value of α (the relative weight of education). Hence, a direct effect of the reform is to favor the "educated outsiders": individuals with high values of "education" (e) but low values of "connections" (c).[69]

Whether the shift towards "merit" helps the poor or the rich depends on the relationship between e, c, and s. Figure B7 illustrates three possible cases. In Panel (a), social class has a stronger correlation with education than it has with connections. In this case, introducing exams disproportionately helps the chances of the children of the rich. In Panel (b), in contrast, social class is correlated with connections but has no relationship with education. Under these conditions, introducing exams increases the representation of the children of the poor. Finally, in Panel (c) the "middle class" increases its representation after the reform: it has similar levels of connections than the "poor" but higher levels of education. Note that, even if there is a positive relationship between education and social class, increasing the weight of education *does not* necessarily favor the children of the rich.

A simplification of this framework is that it abstracts from dynamic considerations. However, applicants of different social backgrounds might differ in their ability to adapt to exams, making the effects of the reform potentially different in the short and the long run. For instance, applicants from wealthier backgrounds might have more resources to prepare for the exams (for instance, by using tutors), thus increasing such applicants' relative advantage over time.[70] Alternatively, applicants from poorer backgrounds might need more time to "catch up" with the exams' content, thus being at a relative disadvantage early on. Our findings are more consistent with this latter possibility: the increase in employees' social status was the strongest immediately after the reform.[71]

6.2 Empirical Evidence

The Reform Increased the Representation of "Educated Outsiders". Our conceptual framework predicts that the reform should have increased the representation of "educated" individuals. Although censuses prior to 1940 do not include direct information on years of schooling, they do include information on occupations. Hence, we can assess if the reform brought workers whose pre-civil service occupation likely required higher educational attainment.[72]

Table 5 shows that employees hired through exams were 8 percentage points more likely to

[69] This framework abstracts from applicants' outside options (that is, we assume that anyone above the cutoff is hired). We do so to keep the framework parsimonious as the reform did not change workers' outside options.

[70] For example, Sundell (2014) describes how "after the introduction of competitive exams for the British Indian Civil Service in the 19th century, private tutors that provided instruction specifically for the tests, "crammers", sprang up."

[71] Our framework can incorporate these dynamics by allowing for a time-varying relationship between e, c, and s.

[72] Literacy (which is included in the census) is a very coarse measure of human capital in this context as more than 90% of the adult white population was literate by 1880.

have held a *professional* occupation prior to joining the civil service (a 90% increase). These occupations are precisely the ones that would have required formal education, suggesting that the reform was successful in recruiting more educated workers. Interestingly, however, we also find an increase in the proportion of workers who were previously employed as farmers. This increase likely reflects the fact that the reform changed employees' social mix, increasing the proportion of those hailing from rural areas (see Table B15 and the discussion below).[73] These increases were mostly driven by a decrease in the likelihood that employees would have held a white-collar non-professional job prior to joining the civil service.[74]

Figure 5 shows the corresponding event-study estimates. The figure shows a rapid increase in the share of workers who had a professional occupation prior to joining the civil service, together with a decrease in the share of those with white-collar non-professional jobs. The increase in the share of workers with a professional background persists 10 years after the reform, suggesting that the reform kept attracting workers with stronger professional backgrounds in the longer term.

An additional implication of our proposed interpretation is that exams should have disproportionately benefited those individuals who grew up in areas with better access to educational resources. To test this hypothesis, we combine information on employees' childhood counties (from our linked sample) with county-level data on per capita schools and teachers in 1850 (from Haines *et al.* (2010)).[75] Panel (a) in Figure B9 shows the number of schools per children aged 5–14 in 1850, by county. The figure shows substantial heterogeneity in the local availability of schools. Counties in the top quintile of the distribution had an average of 7.3 schools every 1000 children, whereas counties in the bottom quintile had only 0.24. Panel (b) shows a similar heterogeneity when focusing on per capita *teachers*.

Panels (a) and (b) in Table 6 show that individuals appointed through exams came from counties with higher per capita schools and teachers. The outcome variable in panel (a) is the log of per capita schools in employees' childhood county, and in panel (b) is the log of per capita teachers. The results are similar when we exploit variation within states (Column 3 in each of the panels) and within urban/rural areas (Column 4), suggesting that our results do not simply capture differences across broad regions of the country. Moreover, the results are also similar when we control for parental characteristics such as occupation (Column 5), birthplace (Column 6), and wealth (Column 7). This similarity suggests that the association we document is not mechanically explained by counties' educational resources being correlated with parental resources.

We next investigate if the reform brought "outsiders", that is, individuals who lacked connections and hence were unlikely to obtain a job through patronage. A challenge in testing this hypothesis is that informal connections are–by their own nature–difficult to observe. Hence, we

[73] Also, note that farmers were not a particularly uneducated group in this period: Among white adult males aged 18 or more in 1880, those employed as farmers had a 91% literacy rate (compared to 93% among non farmers).

[74] The two largest occupations among white-collar non-professionals are "managers, officials and proprietors" and "salesmen and sales clerks". White-collar non-professionals were on average less educated than professionals: In 1870, the average "occupational education score" among white-collar non-professionals was 23.4, whereas it was 82.2 among professionals. This score measures the share of individuals in an occupation who had a college degree in 1950.

[75] We use 1850 because this is the last pre-reform census for which Haines *et al.* (2010) report these data.

proxy for them using four measures (each capturing different types of connections that applicants could have benefited from).[76] First, we use an indicator that is one if the bureaucrat's father had worked for the federal government. Second, we construct an indicator that is one for bureaucrats who had the *same surname* as a member of Congress from their home state.[77] Third, we use an indicator that is one for employees who spent part of their childhood in DC, the city that likely provided the best opportunities to develop informal political connections.[78] Finally, we construct a measure aimed at capturing workers' likely party affiliation. To do so, we combine the data on workers' county of residence prior to joining the civil service (from our linked sample) with historical county-level data on party vote shares (from ICPSR (1999)). We use this information to construct an indicator that is one if a majority of voters in the bureaucrats' county had voted for the incumbent party in the most recent presidential election.[79]

Table 7 shows that employees appointed through exams were less likely to have a father who himself worked for the federal government (although the effect is not statistically significant at the conventional levels), less likely to have spent their childhoods in DC, and less likely to hail from a county which voted for the incumbent party in the previous presidential election.

Although in principle all employees might have benefited from being "connected", it is likely that connections would have been more relevant for accessing jobs that required less technical skills. This would have been the case, for instance, if being hired required at least a minimum level of competency–thus making it harder to privilege connections when hiring for more technical positions (Brierley, 2019).

With this in mind, we assess if the decline in the importance of "connections" depended on the type of position to which the worker was appointed to. To do so, we estimate the specification in equation 3, which distinguishes between employees appointed to the less technical clerical positions and those appointed to positions requiring more specialized knowledge.

Table 7 shows that the decline in the likelihood of being connected comes exclusively from those individuals who were appointed to the relatively non-technical clerical positions: Employees in such positions were 3 percentage points less likely to have a father who worked in the Federal Government (nearly a 50% decline), 8 percentage points less likely to have spent time in DC (a 25% decline), and 5 percentage points less likely to hail from a county that voted for the incumbent party (a 10% decline). By contrast, there is a much more limited decline in these probabilities among those appointed into the more technical positions.

The Middle Class was Overrepresented Among the "Educated Outsiders". Employees hired through exams were more likely to belong to the upper-middle class. Our interpretation of this finding is that the reform increased the share of such workers because they were overrepresented among the "educated outsiders". We offer two pieces of evidence that support this interpretation.

[76] Importantly, all of these measures are pre-determined with respect to bureaucrats' employment in the civil service.

[77] The data on Congressmen names are from the *Biographical Directory of the US Congress* (Dodge & Koed, 2005).

[78] Ziparo (2017) writes that: "Living in the epicenter of national political life, applicants from Washington, D.C., had an advantage in obtaining political influence. [...] In 1861, Abraham and Mary Lincoln wrote letters of recommendation for Ann Sprigg, their landlady during Lincoln's single term in Congress in the late 1840s."

[79] That is, Republicans 1873–1885 and 1889–1893, and Democrats 1885–1889 and from 1893 until the end of our sample.

First, Panel (a) in Figure 4 shows that, prior to the reform, workers whose families belonged to the 60–80 quintile of the wealth distribution were similarly represented in the positions that eventually became subject to exams as those whose families belonged to the bottom three quintiles. Moreover, Figure B8 shows that this similar representation occurred despite the higher educational attainment of workers from the 60-80 quintile. This figure shows school attendance rates by age and parental wealth quintile, based on census data covering the entire 1870 US population.

Second, Figure B10 shows that the low representation of the 60–80 quintile in government jobs prior to the reform was unusual relative to its representation in comparable *private sector* jobs. This figure shows the distribution across parental wealth quintiles of private sector white-collar workers, based on a sample linking adults in the 1880 census to their childhood households in 1860.[80] The figure shows that, unlike in the case of civil servants prior to the reform, the likelihood of holding a white-collar job in the private sector grew monotonically with parental wealth ranks.

What Explains the Presence of Workers from Disadvantaged Backgrounds in Government Jobs Prior to the Reform? A surprising implication of our findings and conceptual framework is that, prior to the reform, applicants from poorer backgrounds must have been better "connected" than middle-class applicants. This implication is derived from the fact that these applicants had worse education than middle-class applicants (Figure B8) but nevertheless managed to obtain a similar share of government jobs (Figure 4).

A likely explanation for this pattern is that workers from disadvantaged backgrounds (who typically face worse outside options) might be more likely to be targeted for patronage jobs than those from the middle class.[81] The historical literature suggests that this was indeed the case in our context. In particular, our period of analysis coincides with the emergence of the "urban political patronage machine" (Brown & Halaby, 1987).[82] These machines emerged from the interaction between two major developments of the period: the rise of "mass-based political parties" (whose campaigns required large mobilizations of workers and resources) and the expansion of the urban working-class. These developments created opportunities for mutually beneficial exchanges between political machines and impoverished city dwellers: as the "urban immigrant and lower classes needed help", the machine provided "assistance and jobs in return for loyalty, labor, and votes" (Mashaw, 2010). A consequence of these exchanges was to draw an "unprecedented numbers of ordinary citizens into the channels of political life" (James, 2006).

Our empirical findings are consistent with this explanation. First, we observe declines in the share of workers from counties in which the incumbent party had received a majority vote, suggesting that the reform indeed reduced political favoritism in the allocation of jobs (Table 7). Sec-

[80] This sample was constructed using the exact same algorithm that we use to link the personnel records to the census.

[81] For instance, Sorauf (1960) argues that political machines "flourished especially in those urban centers inhabited by large groups of immigrants and minorities-groups not yet integrated into American life, often poor and insecure and bewildered by the traditions of American politics." Indeed, politicians' ability to use patronage jobs has diminished as "private employment has become progressively more attractive with rising wage levels, union protections and securities, unemployment compensation, pension plans, and fringe benefits."

[82] Brown & Halaby (1987) define a "political machine" as a "political party that joins a particularistic style of mobilization-one rooted in favoritism and the use of material inducements".

ond, we observe declines in the share of immigrants (Table 4), a group typically described as the primary target of urban political machines (see, for instance, Cornwell Jr (1964)). Finally, we find a sharp decline in the share of workers from urban areas (Table B15); namely the locations where machines were the most active (Brown & Halaby, 1987).

The Reform Hurt the Chances of the Poor when Inequality in Access to Education was High. One implication of our conceptual framework is that, the higher the inequality in educational resources, the more negative the impact of a shift towards "merit" on the chances of children from poorer backgrounds. To test this implication, we exploit state-level variation in differences in access to schooling by parental wealth. Note that, due to the state apportionment rules, applicants to jobs in the classified departmental service were in practice only competing against individuals from their own home state (thus making within-state inequality a relevant consideration). Specifically, we use the 1870 census to compute, for each state s:

$$Inequality_s = \frac{\%\, Children\, in\, school\, if\, family\, in\, top\, 20\%\, in\, state\, s}{\%\, Children\, in\, school\, if\, family\, in\, bottom\, 20\% in\, state\, s} \quad (5)$$

This measure corresponds to the ratio between: (1) the likelihood that a child aged 8 to 12 from a family in the top 20% of the wealth distribution would be in school, and (2) the likelihood that a child aged 8 to 12 from the bottom 20% would be in school. This ratio would be one in a state with the same access to schooling regardless of parental wealth but above one when such access favors the children of the rich. Figure B11 shows substantial regional heterogeneity in this measure, with the highest values in the South and the lowest in parts of the Northeast.

Table 8 reports results in which we split the sample based on whether states had below or above median levels of educational inequality. Panel (a) shows that the increase in the summary index of socioeconomic background is about twice as large in the high-inequality states. Note, however, that the index increases *both* in the below- and in the above-median inequality states.

In contrast, Table B16 shows that there is much more limited heterogeneity when we focus on workers' *own* occupation prior to joining the civil service: there is a similar increase in the likelihood that an employee would have held a professional occupation regardless of which group of states we consider. This finding suggests that the reform was successful in bringing "educated outsiders" from both low- and high-inequality states but what varied across states with different levels of inequality is *who* these educated outsiders were in terms of social class.

The Reform Hurt the Chances of the Poor in the Offices that Relied the Most on Connections. If lower-SES individuals were displaced by the reform because they received a disproportionate share of patronage appointments, their representation should have decreased the most in those government offices in which such patronage was more prevalent prior to the reform (that is, in those offices that relied more heavily on connections).

To test this implication, we exploit heterogeneity across offices in the extent to which, prior to the reform, employees received a higher compensation than that which their pre-civil service occupation would predict. Intuitively, offices that assigned a higher weight to connections relative to education should have paid salaries that are too high (compared to other offices) for a given pre-civil service occupation (our proxy for education). To quantify such "excess pay", we use the sample of individuals for whom we observe pre-civil service occupation to estimate (restricting the sample to the pre-reform period):

$$\log(compensation) = \alpha + \beta \log(pre\, bureaucracy\, occupational\, score) + \epsilon \tag{6}$$

where $compensation$ is the first annual compensation that we observe for a worker in the civil service, and $pre\, bureaucracy\ occupational\ score$ is the occupational score corresponding to the last occupation that we observe for that same worker prior to joining the civil service.[83] We then use the results of this regression to compute, for each office in the pre-reform period, the fraction of employees earning more than their predicted compensation.

Table 9 shows a higher increase in the summary index of employees' socioeconomic backgrounds in those offices in which, prior to the reform, a larger number of workers were paid more than their predicted compensation. This finding is consistent with the idea that, in our context, "connections" were distributed in a more egalitarian fashion than "education": the sharpest decline in the representation of lower-SES candidates happened precisely in the offices that were forced to reduce the weight of connections the most.

6.3 Additional Exam-Related Channels

Exam-Induced Change in the Applicant Pool. Our data and empirical design do not allow us to distinguish if the reform reduced the representation of lower-SES individuals because such individuals performed worse in the exams, or simply because they were discouraged by the exam and hence did not apply to begin with. We note, however, that this combined effect is what ultimately matters for representation.

Similarly, the reform may have also affected bureaucrats' characteristics indirectly, by changing the group of individuals interested in such career. For instance, the reform might have increased the prestige of holding a government job, thus increasing the appeal of such jobs for individuals of higher social status. Although this effect is inherent to any change in recruitment practices (and the combined effect of changes in screening and changes in the applicant pool is what ultimately matters for representation), two pieces of evidence suggest that the effects we document are not solely driven by changes in the prestige of holding a government job. First, if the effects were only driven by such changes, it is unclear why the effects would be stronger among applicants from states with high inequality in access to schooling. Second, the rapid change in bureaucrats' backgrounds that we observe seems inconsistent with the effects being driven by plausibly slower-to-change perceptions about the prestige of public employment.

Biases in the Exam. A final possibility is that the exam reduced the chances of applicants from disadvantaged backgrounds because its content was biased against such applicants. For instance, the exam might have included questions that were irrelevant for job performance but

[83] Specifically, we use IPUMS' *occscore* variable. This variable corresponds to median earnings by occupation based on the 1950 census.

whose answers were only accessible for upper-class individuals. Two pieces of evidence suggest that such biases are not the main channel that explains our results. First, as described in Section 2, the exam focused on practical skills relevant for the job. Indeed, applicants to different positions completed different exams.[84] Second, if such biases were the only explanation for our findings, it is unclear why, conditional on parental characteristics, individuals from counties with better educational resources would have increased their representation after the reform.[85]

7 Conclusions

By limiting discretion and making "merit" the main selection criteria, competitive entry exams are sometimes praised as providing an equal chance to all candidates regardless of background. However, critics of exams argue that they simply reproduce existing inequities rather than "leveling the playing field". Do exams help or hurt the chances of lower-SES candidates? This paper studied this question using evidence from the Pendleton Act, a landmark reform in American history which introduced competitive exams for the selection of certain federal employees. Using newly assembled data on the socioeconomic backgrounds of government employees and a difference-in-differences strategy, we find that, while the reform increased the representation of "educated outsiders" (individuals with high education but limited connections), it led to a persistent reduction in the representation of lower-SES individuals: workers hired through exams came from wealthier families, were more likely to be the children of professionals, and were less likely to be the children of immigrants (or immigrants themselves). This increased elitism was stronger among employees from states with more unequal access to schooling as well as in offices that relied more heavily on connections prior to the reform.

Our findings have implications for the broader debate on exams and meritocracy. Allocating opportunities based on exams is sometimes described as an equity-efficiency panacea, helping select the most qualified candidates while simultaneously increasing the representation of lower-SES individuals. Our results challenge this view: although using exams could, in principle, help select more qualified candidates, we show that these improvements *can* also come with the cost of increased elitism. More generally, our findings show that adopting less discretionary selection criteria might not necessarily help the chances of lower-SES individuals.

Our results and interpretation also highlight that, to assess the potential equity implications of exams, it is crucial to understand the forces shaping the distribution of the attributes (education, connections, etc.) rewarded by exams and by the alternative selection criteria. A likely driver of the increased elitism that we document is that, prior to the introduction of exams, patronage politics might have served as a "great equalizer" in the distribution of connections. If exams are replaced by other selection criteria, it is important to ask what would be the corresponding eco-

[84] Despite the general practical nature of the test, more subtle biases might have negatively affected the chances of applicants from disadvantaged backgrounds. For example, including questions related to financial issues has been shown to harm the performance of disadvantaged students (Duquennois, 2019).

[85] The poor may also suffer more frequent shocks which affect exam performance (Cotti *et al.*, 2018).

nomic and social forces equalizing connections (more so than the attributes rewarded by exams) in the respective setting.

Importantly, while we investigate how exams shaped the social origins of government officials, an important question that remains unanswered is whether the poor themselves were *on net* made worse off by the reform. The answer to this question is not an obvious one for a variety of reasons. For instance, individuals from disadvantaged backgrounds might benefit the most from having a well-functioning state, even if achieving this efficiency implies that they might lose direct access to government jobs. We hope future work can shed light on the overall distributional implications of selecting government officials using exams.

References

ABRAMITZKY, RAN, & BOUSTAN, LEAH. 2017. Immigration in American economic history. *Journal of economic literature*, **55**(4), 1311–45.

ABRAMITZKY, RAN, BOUSTAN, LEAH PLATT, & ERIKSSON, KATHERINE. 2012. Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration. *American Economic Review*, **102**(5), 1832–56.

ABRAMITZKY, RAN, BOUSTAN, LEAH PLATT, ERIKSSON, KATHERINE, FEIGENBAUM, JAMES J., & PÉREZ, SANTIAGO. 2019. *Automated Linking of Historical Data*. Tech. rept. NBER WP 25825.

AGER, PHILIPP, BOUSTAN, LEAH PLATT, & ERIKSSON, KATHERINE. 2019. *The intergenerational effects of a large wealth shock: White southerners after the Civil War*. Tech. rept. National Bureau of Economic Research.

ALSAN, MARCELLA, GARRICK, OWEN, & GRAZIANI, GRANT. 2019. Does diversity matter for health? Experimental evidence from Oakland. *American Economic Review*, **109**(12), 4071–4111.

ARON, CINDY SONDIK. 1987. *Ladies and gentlemen of the civil service: Middle-class workers in Victorian America*. Oxford University Press.

AUTOR, DAVID H, & SCARBOROUGH, DAVID. 2008. Does job testing harm minority workers? Evidence from retail establishments. *The Quarterly Journal of Economics*, **123**(1), 219–277.

BAI, YING, & JIA, RUIXUE. 2016. Elite recruitment and political stability: the impact of the abolition of China's civil service exam. *Econometrica*, **84**(2), 677–733.

BEAMAN, LORI, DUFLO, ESTHER, PANDE, ROHINI, & TOPALOVA, PETIA. 2012. Female leadership raises aspirations and educational attainment for girls: A policy experiment in India. *science*, 1212382.

BEAMAN, LORI, KELEHER, NIALL, & MAGRUDER, JEREMY. 2018. Do job networks disadvantage women? Evidence from a recruitment experiment in Malawi. *Journal of Labor Economics*, **36**(1), 121–157.

BENSON, ALAN, LI, DANIELLE, & SHUE, KELLY. 2021. *"Potential" and the Gender Promotion Gap*. Tech. rept. working paper.

BIASI, BARBARA, & SARSONS, HEATHER. 2022. Flexible wages, bargaining, and the gender gap. *The Quarterly Journal of Economics*, **137**(1), 215–266.

BOSTASHVILI, D., & UJHELYI, GERGELY. 2019. Political Budget Cycles and the Civil Service: Evidence from Highway Spending in US States. *Journal of Public Economics*, **175**(C), 17–28.

BOURDIEU, PIERRE. 1998. *The state nobility: Elite schools in the field of power*. Stanford University Press.

BRIERLEY, SARAH. 2019. Combining patronage and merit in public sector recruitment. *Journal of Politics*.

BROWN, M CRAIG, & HALABY, CHARLES N. 1987. Machine politics in America, 1870-1945. *The Journal of Interdisciplinary History*, **17**(3), 587–612.

CAMARA, WAYNE J, & SCHMIDT, AMY ELIZABETH. 1999. Group Differences in Standardized Testing and Social Stratification. Report No. 99-5. *College Entrance Examination Board*.

CASTILLA, EMILIO J. 2008. Gender, race, and meritocracy in organizational careers. *American journal of sociology*, **113**(6), 1479–1526.

CHATTOPADHYAY, RAGHABENDRA, & DUFLO, ESTHER. 2004. Women as policy makers: Evidence from a randomized policy experiment in India. *Econometrica*, **72**(5), 1409–1443.

CHEN, TING, KUNG, JAMES KAI-SING, & MA, CHICHENG. 2020. Long live Keju! The persistent effects of China's civil examination system. *The economic journal*, **130**(631), 2030–2064.

CIVIL SERVICE COMMISSION. 1883. *Report of the United States Civil-Service Commission*. US Government Printing Office.

CIVIL SERVICE COMMISSION. 1883-1893. *Report of the United States Civil-Service Commission*. US Government Printing Office.

CIVIL SERVICE COMMISSION. 1884. *Report of the United States Civil-Service Commission*. US Government Printing Office.

CIVIL SERVICE COMMISSION. 1885. *Report of the United States Civil-Service Commission*. US Government Printing Office.

CIVIL SERVICE COMMISSION. 1886. *Report of the United States Civil-Service Commission*. US Government Printing Office.

CIVIL SERVICE COMMISSION. 1890. *Report of the United States Civil-Service Commission*. US Government Printing Office.

CIVIL SERVICE COMMISSION. 1891. *Report of the United States Civil-Service Commission*. US Government Printing Office.

CIVIL SERVICE COMMISSION. 1892. *Report of the United States Civil-Service Commission*. US Government Printing Office.

CIVIL SERVICE COMMISSION. 1893. *Report of the United States Civil-Service Commission*. US Government Printing Office.

CORNWELL JR, ELMER E. 1964. Bosses, machines, and ethnic groups. *The Annals of the American Academy of Political and Social Science*, **353**(1), 27–39.

COTTI, CHAD, GORDANIER, JOHN, & OZTURK, ORGUL. 2018. When does it count? The timing of food stamp receipt and educational performance. *Economics of Education Review*, **66**, 40–50.

COUNCIL, NATIONAL RESEARCH, *et al.* 1989. *Fairness in employment testing: Validity generalization, minority issues, and the General Aptitude Test Battery*. National Academies Press.

CRONBACH, LEE J. 1975. Five decades of public controversy over mental testing. *American Psychologist*, **30**(1), 1.

DAL BÓ, ERNESTO, FINAN, FREDERICO, FOLKE, OLLE, PERSSON, TORSTEN, & RICKNE, JOHANNA. 2017. Who becomes a politician? *The Quarterly Journal of Economics*, **132**(4), 1877–1914.

DEPARTMENT OF THE INTERIOR. 1871-1893. *Official Register of the United States*. US Government Printing Office.

DODGE, ANDREW R, & KOED, BETTY K. 2005. *Biographical Directory of the United States Congress, 1774-2005: The Continental Congress, September 5, 1774, to October 21, 1788, and the Congress of the United States, from the First Through the One Hundred Eighth Congresses, March 4, 1789, to January 3, 2005, Inclusive*. Vol. 108. US Government Printing Office.

DUQUENNOIS, CLAIRE. 2019. Fictional money, real costs: Impacts of financial salience on disadvantaged students. *Unpublished manuscript*.

ESTRADA, RICARDO. 2019. Rules versus discretion in public service: Teacher hiring in Mexico. *Journal of Labor Economics*, **37**(2), 545–579.

EVANS, PETER, & RAUCH, JAMES E. 1999. Bureaucracy and growth: A cross-national analysis of the effects of" Weberian" state structures on economic growth. *American sociological review*, 748–765.

FINAN, FREDERICO, OLKEN, BENJAMIN A, & PANDE, ROHINI. 2017. The personnel economics of the developing state. *Pages 467–514 of: Handbook of Economic Field Experiments*, vol. 2. Elsevier.

FISH, CARL RUSSELL. 1905. *The civil service and the patronage*. Vol. 11. Longmans, Green, and Company.

FOLKE, OLLE, HIRANO, SHIGEO, & SNYDER, JAMES M. 2011. Patronage and elections in US states. *American Political Science Review*, **105**(3), 567–585.

GOLDIN, CLAUDIA, & KATZ, LAWRENCE F. 2003. *The" virtues" of the past: Education in the first hundred Years of the new republic*.

GOLDIN, CLAUDIA, & ROUSE, CECILIA. 2000. Orchestrating impartiality: The impact of" blind" auditions on female musicians. *American economic review*, **90**(4), 715–741.

GREENE, JACK P. 1984. *Encyclopedia of American political history: studies of the principal movements and ideas*. New York: Scribner.

GRINDLE, MERILEE S. 2012. *Jobs for the Boys*. Harvard University Press.

HAINES, MICHAEL R, *et al.* 2010. Historical, demographic, economic, and social data: the United States, 1790–2002. *Ann Arbor, MI: Inter-university Consortium for Political and Social Research*.

HOFFMAN, MITCHELL, KAHN, LISA B, & LI, DANIELLE. 2018. Discretion in hiring. *The Quarterly Journal of Economics*, **133**(2), 765–800.

HOOGENBOOM, ARI. 1959. The Pendleton Act and the civil service. *The American Historical Review*, **64**(2), 301–318.

HOOGENBOOM, ARI ARTHUR. 1968. *Outlawing the spoils: a history of the civil service reform movement, 1865-1883*. Vol. 50. University of Illinois Press.

ICPSR. 1999. United States Historical Election Returns, 1824-1968.

JAMES, SCOTT C. 2006. Patronage Regimes and American Party Development from 'The Age of Jackson'to the Progressive Era. *British Journal of Political Science*, **36**(1), 39–60.

JOHNSON, RONALD N, & LIBECAP, GARY D. 1994. Patronage to merit and control of the federal government labor force. *Explorations in Economic History*, **31**(1), 91–119.

KEISER, LAEL R, WILKINS, VICKY M, MEIER, KENNETH J, & HOLLAND, CATHERINE A. 2002. Lipstick and logarithms: Gender, institutional context, and representative bureaucracy. *American political science review*, **96**(3), 553–564.

KINGSLEY, J DONALD. 1944. Representative bureaucracy. *Representative Bureaucracy*, 12.

KLING, JEFFREY R, LIEBMAN, JEFFREY B, & KATZ, LAWRENCE F. 2007. Experimental analysis of neighborhood effects. *Econometrica*, **75**(1), 83–119.

LEE, DAVID S. 2009. Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *The Review of Economic Studies*, **76**(3), 1071–1102.

LEUPP, F.E. 1898. *How to Prepare for a Civil-service Examination: With Recent Questions and Answers*. Hinds & Noble.

LI, DANIELLE, RAYMOND, LINDSEY R, & BERGMAN, PETER. 2020. *Hiring as exploration*. Tech. rept. National Bureau of Economic Research.

LIBECAP, GARY D, & JOHNSON, RONALD N. 2007. *The Federal Civil Service System and the Problem of Bureaucracy: The Economics and Politics of Institutional Change*. University of Chicago Press.

LONG, JASON, & FERRIE, JOSEPH. 2013. Intergenerational occupational mobility in Great Britain and the United States since 1850. *American Economic Review*, **103**(4), 1109–37.

MASHAW, JERRY L. 2010. Federal Administration and Administrative Law in the Gilded Age. *The Yale Law Journal*, 1362–1472.

MOREIRA, DIANA, & PÉREZ, SANTIAGO. 2021. *Civil Service Exams and Organizational Performance: Evidence from the Pendleton Act*. Tech. rept. National Bureau of Economic Research.

MYERSON, ROGER. *American Political Science Review*, 856.

NEGGERS, YUSUF. 2018. Enfranchising your own? Experimental evidence on bureaucrat diversity and election bias in India. *American Economic Review*, **108**(6), 1288–1321.

ORNAGHI, ARIANNA. 2016. Civil service reforms: Evidence from US police departments. *Job Market Paper*.

PANDE, ROHINI. 2003. Can mandated political representation increase policy influence for disadvantaged minorities? Theory and evidence from India. *American Economic Review*, **93**(4), 1132–1151.

POWELL, G. BINGHAM.

RAUCH, JAMES E, & EVANS, PETER B. 2000. Bureaucratic structure and bureaucratic performance in less developed countries. *Journal of public economics*, **75**(1), 49–71.

RICCUCCI, NORMA M, VAN RYZIN, GREGG G, & LAVENA, CECILIA F. 2014. Representative bureaucracy in policing: Does it increase perceived legitimacy? *Journal of public administration research and theory*, **24**(3), 537–551.

RUGGLES, STEVEN, FLOOD, SARAH, GOEKEN, RONALD, GROVER, JOSIAH, MEYER, ERIN, PACAS, JOSE, & SOBEK, MATTHEW. 2021. *IPUMS USA: Version 10.0 [dataset]. Minneapolis, MN: IPUMS; 2020*.

SORAUF, FRANK J. 1960. The Silent Revolution in Patronage. *Public Administration Review*, **20**(1), 28–34.

SUNDELL, ANDERS. 2014. Are formal civil service examinations the most meritocratic way to recruit civil servants? Not in all countries. *Public Administration*, **92**(2), 440–457.

Teorell, Jan, Dahlström, Carl, & Dahlberg, Stefan. 2011. The QoG expert survey dataset. *Available at SSRN 3569575*.

Theriault, Sean M. 2003. Patronage, the Pendleton Act, and the Power of the People. *The Journal of Politics*, **65**(1), 50–68.

Thompson, Daniel M, Feigenbaum, James J, Hall, Andrew B, & Yoder, Jesse. 2019. *Who Becomes a Member of Congress? Evidence From De-Anonymized Census Data*. Tech. rept. National Bureau of Economic Research.

Weber, Max. 2019. *Economy and society*. Harvard University Press.

Xu, Guo. 2018. The costs of patronage: Evidence from the british empire. *American Economic Review*, **108**(11), 3170–98.

Xu, Guo. 2020. Bureaucratic Representation and State Responsiveness: The 1918 Pandemic in India.

Zhao, Yong. 2014. *Who's afraid of the big bad dragon?: Why China has the best (and worst) education system in the world*. John Wiley & Sons.

Ziparo, Jessica. 2017. *This Grand Experiment: When Women Entered the Federal Workforce in Civil War–Era Washington*. UNC Press Books.

TABLE 1: EFFECTS OF THE REFORM ON SUMMARY INDICES OF EMPLOYEES' SOCIOECONOMIC BACKGROUND, DIFFERENCE-IN-DIFFERENCES ESTIMATES

	Summary Index		First Principal Component	
	(1)	(2)	(3)	(4)
Exam X After	0.180***	0.169***	0.290***	0.270***
	(0.0451)	(0.0468)	(0.0811)	(0.0856)
Year FE	Yes	Yes	Yes	Yes
Position FE	Yes	Yes	Yes	Yes
App. State FE	No	Yes	No	Yes
Observations	2944	2944	2944	2944

Notes: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$. The dependent variable in columns 1 and 2 is a summary index of employees' socioeconomic background computed using the approach in Kling *et al.* (2007). The index combines information on parental wealth rank, parental occupations, parental literacy, own and parental birthplace and race. The variables composing the index are normalized such a higher value corresponds to a higher socioeconomic status. In columns 3 and 4, the dependent variable is the first principal component of the same set of characteristics as in columns 1 and 2. All columns include hiring year and position fixed effects. The odd columns further include fixed effects based on employees' state "whence appointed". The sample is restricted to newly hired employees in the Executive Departments in DC. Standard errors clustered at the position level.

TABLE 2: EFFECTS OF THE REFORM ON EMPLOYEES' PARENTAL WEALTH RANKS, DIFFERENCE-IN-DIFFERENCES ESTIMATES

	Total		Personal		Real Estate	
	(1)	(2)	(3)	(4)	(5)	(6)
Exam X After	0.0647**	0.0624**	0.0708***	0.0694***	0.0489*	0.0427
	(0.0258)	(0.0273)	(0.0240)	(0.0259)	(0.0260)	(0.0263)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Position FE	Yes	Yes	Yes	Yes	Yes	Yes
App. State FE	No	Yes	No	Yes	No	Yes
Observations	3034	3034	3034	3034	3034	3034

Notes: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$. The dependent variable in columns 1 and 2 is the rank of a bureaucrat father in the US national wealth distribution. Wealth is computed based on the combined values of real estate and personal property. These ranks are computed separately by census year (1860 and 1870) and by age (that is, relative to all fathers of the same age). When a bureaucrat is linked to more than one census with information on parental wealth, we use the average rank across both census years as our outcome variable. In columns 3 and 4, this rank is computed based solely on personal property, whereas in columns 5 and 6 it is based solely on real estate property. All columns include hiring year and position fixed effects. The odd columns further include fixed effects based on employees' state "whence appointed". The sample is restricted to newly hired employees in the Executive Departments in DC. Standard errors clustered at the position level.

TABLE 3: EFFECTS OF THE REFORM ON EMPLOYEES' PARENTAL OCCUPATIONS, DIFFERENCE-IN-DIFFERENCES ESTIMATES

	Professional		White-Collar Non-Prof		Farmer		Skilled Blue Collar		Unskilled	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Exam X After	0.0530**	0.0488**	-0.0119	-0.00378	0.0606***	0.0258	-0.0668**	-0.0455	-0.0375*	-0.0242
	(0.0239)	(0.0241)	(0.0347)	(0.0325)	(0.0234)	(0.0244)	(0.0327)	(0.0311)	(0.0216)	(0.0214)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Position FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
App. State FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	4993	4993	4993	4993	4993	4993	4993	4993	4993	4993

Notes: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$. The dependent variable in each of the columns is an indicator that takes a value of one if the father of a bureaucrat worked in a certain occupational category (as indicated by the column). When a bureaucrat is linked to more than one census with information on father's occupations, we use the fraction of census years that their father spent in a given occupational category as our outcome variable. Professional occupations are those with a value of less than 100 in the 1950 Census occupational classification system. Examples of these occupations include lawyers and accountants. Non-professional white-collar are those with a value between 200 and 500 (for example, clerks). Farmers are those with a value of 100. Skilled blue-collar are those with a value between 500 and 700 (for example, carpenters). Unskilled are those with a value of 700 or more (for example, farm laborers). All columns include hiring year and position fixed effects. The odd columns further include fixed effects based on employees' state "whence appointed". The sample is restricted to newly hired employees in the Executive Departments in DC. Standard errors clustered at the position level.

TABLE 4: EFFECTS OF THE REFORM ON EMPLOYEES' COUNTRY OF ORIGIN AND RACE, DIFFERENCE-IN-DIFFERENCES ESTIMATES

	Immigrant		Immigrant Parents		White	
	(1)	(2)	(3)	(4)	(5)	(6)
Exam X After	-0.0473**	-0.0419**	-0.0732***	-0.0693***	0.00660	0.00327
	(0.0190)	(0.0176)	(0.0237)	(0.0240)	(0.00891)	(0.00956)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Position FE	Yes	Yes	Yes	Yes	Yes	Yes
App. State FE	No	Yes	No	Yes	No	Yes
Observations	9238	9238	4822	4822	9238	9238

Notes: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$. The dependent variable in columns 1 and 2 is an indicator that takes a value of one if the worker is foreign born. The dependent variable in columns 3 and 4 is an indicator that takes a value of one if both workers' parents are foreign born. The dependent variable in columns 5 and 6 is an indicator that takes a value of one if the workers is listed as being white in the census. All columns include hiring year and position fixed effects. The odd columns further include fixed effects based on employees' state "whence appointed". The sample is restricted to newly hired employees in the Executive Departments in DC. Standard errors clustered at the position level.

TABLE 5: EFFECTS OF THE REFORM ON EMPLOYEES' OCCUPATIONAL BACKGROUND, DIFFERENCE-IN-DIFFERENCES ESTIMATES

	Professional		White-Collar Non-Prof		Farmer		Skilled Blue Collar		Unskilled	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Exam X After	0.0778*	0.0818*	-0.163***	-0.163***	0.0894**	0.0852**	-0.00165	0.00108	0.00826	0.00300
	(0.0428)	(0.0426)	(0.0564)	(0.0588)	(0.0347)	(0.0353)	(0.0270)	(0.0292)	(0.0333)	(0.0338)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Position FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
App. State FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	3582	3582	3582	3582	3582	3582	3582	3582	3582	3582

Notes: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$. The dependent variable in each of the columns is an indicator that takes a value of one if a bureaucrat worked in a certain occupational category (as indicated by the column) prior to joining the civil service. When a bureaucrat is linked to more than one census with information on adult occupations, we use the most recent occupation as long as it corresponds to a census conducted prior to the corresponding register. The sample is restricted to workers who were at least 25 year old at the time we observe them in the census. See notes to Table 3 for a definition of occupational categories. All columns include hiring year and position fixed effects. The odd columns further include fixed effects based on employees' state "whence appointed". The sample is restricted to newly hired employees in the Executive Departments in DC. Standard errors clustered at the position level.

TABLE 6: EFFECTS OF THE REFORM ON EMPLOYEES' EDUCATIONAL INPUTS DURING CHILDHOOD, DIFFERENCE-IN-DIFFERENCES ESTIMATES

(A) PER CAPITA SCHOOLS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Exam X After	0.177***	0.154***	0.130***	0.121***	0.147***	0.146***	0.123***
	(0.0375)	(0.0370)	(0.0364)	(0.0372)	(0.0388)	(0.0388)	(0.0438)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Position FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
App. State FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Childhood State FE	No	No	Yes	Yes	Yes	Yes	Yes
Urban FE	No	No	No	Yes	Yes	Yes	Yes
Parental Occupations FE	No	No	No	No	Yes	Yes	Yes
Parental Birthplace FE	No	No	No	No	No	Yes	Yes
Parental Wealth Rank	No	No	No	No	No	No	Yes
Observations	5498	5498	5498	5498	4691	4691	2866

(B) PER CAPITA TEACHERS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Exam X After	0.223***	0.189***	0.147***	0.139**	0.191***	0.190***	0.161**
	(0.0550)	(0.0540)	(0.0531)	(0.0539)	(0.0571)	(0.0571)	(0.0683)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Position FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
App. State FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Childhood State FE	No	No	Yes	Yes	Yes	Yes	Yes
Urban FE	No	No	No	Yes	Yes	Yes	Yes
Parental Occupations FE	No	No	No	No	Yes	Yes	Yes
Parental Birthplace FE	No	No	No	No	No	Yes	Yes
Parental Wealth Rank	No	No	No	No	No	No	Yes
Observations	5498	5498	5498	5498	4691	4691	2866

Notes: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$. The dependent variable in each of the columns of panel (a) is the (log) number of per capita schools in bureaucrats' childhood county of residence. The dependent variables in each of the columns of panel (b) is the (log) number of per capita teachers. When bureaucrats are linked to more than one census in which they are below the age of 18, we use the average of log per capita schools (or teachers) as the dependent variable. The data on per capita schools and teachers are from Haines *et al.* (2010). All columns include hiring year and position fixed effects. Columns 2 to 7 include additional control variable as indicated by the table. The sample is restricted to newly hired employees in the Executive Departments in DC. Standard errors clustered at the position level.

Table 7: Effects of the Reform on the Likelihood that Employees would be "Connected", Difference-in-Differences Estimates

	Father Gov. Emp.		Lived in DC		Same Surname		Incumbent Party		Vote Share	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Exam X After	-0.00863		-0.0564***		0.00346		-0.0484*		-0.0132	
	(0.0164)		(0.0190)		(0.00444)		(0.0261)		(0.00844)	
Clerk X After		-0.0326*		-0.0810***		-0.00155		-0.0531*		-0.0218**
		(0.0175)		(0.0196)		(0.00573)		(0.0275)		(0.00946)
Tech. X After		0.0266		-0.0172		0.0119		-0.0397		0.00259
		(0.0303)		(0.0253)		(0.00899)		(0.0400)		(0.0112)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Position FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
App. State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4993	4993	5860	5860	9238	9238	6416	6416	6416	6416

Notes: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$ The outcome in columns 1 and 2 is an indicator that takes a value of one if a bureaucrat's father is ever recorded in the census as working in industry 916 ("Federal public administration") based on the 1950 census industry classification. The outcome in columns 3 and 4 is an indicator that takes a value of one is a bureaucrat is ever observed living in Washington DC before the age of 18 (and prior to being employed in the federal administration). The outcome in columns 5 and 6 is an indicator that takes a value of one if a bureaucrat shared a surname with a current member of Congress from his own state of birth or appointment. The outcome in columns 7 and 8 is an indicator that takes a value of one if the incumbent party had obtained a majority vote in bureaucrat's last county of residence in the most recent presidential elections. The outcome in columns 9 and 10 is instead the vote share of the incumbent party. $Exam \times After$ is the coefficient corresponding to our baseline difference-in-differences specification. $Clerk \times After$ is an indicator that takes a value of one for employees hired through exams as clerks or copyists in the post-reform period. $Tech \times After$ is an indicator that takes a value of one for employees hired through exams in technical position in the post-reform period. All columns include hiring year, position and home-state fixed effects. The sample is restricted to newly hired employees in the Executive Departments in DC. Standard errors clustered at the position level.

TABLE 8: EFFECTS OF THE REFORM ON SUMMARY INDEX OF EMPLOYEES' SOCIOECONOMIC BACKGROUND, DIFFERENCE-IN-DIFFERENCES ESTIMATES: HETEROGENEITY BY CHILDHOOD'S STATE INEQUALITY IN ACCESS TO SCHOOLING

	Below Median Ineq.		Above Median Ineq.	
	(1)	(2)	(3)	(4)
Exam X After	0.143***	0.134**	0.228***	0.221***
	(0.0528)	(0.0551)	(0.0661)	(0.0701)
Year FE	Yes	Yes	Yes	Yes
Position FE	Yes	Yes	Yes	Yes
App. State FE	No	Yes	No	Yes
Observations	2204	2204	740	740

Notes: $***p < 0.01$, $**p < 0.05$, $*p < 0.1$. The dependent variable is a summary index of workers' socioeconomic backgrounds computed using the approach in Kling *et al.* (2007). The index combines information on parental wealth rank, parental occupations, parental literacy, own and parental birthplace and race. The variables composing the index are normalized such a higher value corresponds to a higher socioeconomic status. The sample in columns 1 and 2 in each panel is restricted to employees from states with below median inequality in access to schooling, as described in the main body of the paper. The sample in columns 3 and 4 is restricted to employees from states with above median inequality. All columns include hiring year and position fixed effects. The odd columns further include fixed effects based on employees' state "whence appointed". The sample is restricted to newly hired employees in the Executive Departments in DC. Standard errors clustered at the position level.

TABLE 9: EFFECTS OF THE REFORM ON THE SUMMARY INDICES OF EMPLOYEES' SOCIOECONOMIC BACKGROUND, DIFFERENCE-IN-DIFFERENCES ESTIMATES: HETEROGENEITY BY OFFICES' ESTIMATED "EXCESS PAY"

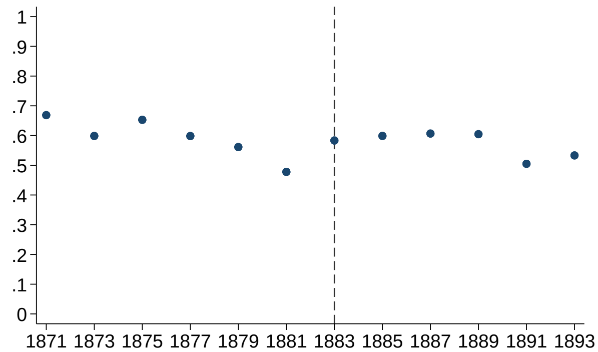
	Below Median Excess Pay		Above Median Excess Pay	
	(1)	(2)	(3)	(4)
Exam X After	0.157	0.0979	0.185***	0.178***
	(0.122)	(0.128)	(0.0509)	(0.0482)
Year FE	Yes	Yes	Yes	Yes
Position FE	Yes	Yes	Yes	Yes
App. State FE	No	Yes	No	Yes
Observations	1467	1467	1477	1477

Notes: ***p<0.01, ** p<0.05, *p < 0.1. The dependent variable is a summary index of employees' socioeconomic background computed using the approach in Kling *et al.* (2007). The index combines information on parental wealth rank, parental occupations, parental literacy, own and parental birthplace and race. The variables composing the index are normalized such a higher value corresponds to a higher socioeconomic status. All columns include hiring year and position fixed effects. The odd columns further include fixed effects based on employeesâ state "whence appointed". The sample is restricted to newly hired employees in the Executive Departments in DC. In columns 1 and 2, the sample is restricted to workers in offices where a below-median proportion of employees received a compensation higher than the predicted compensation based on their pre-bureaucracy occupation. In columns 3 and 4, it is restricted to offices where an above-median proportion of employees received a higher than predicted compensation. See the main text for our definition of "excess pay". Standard errors clustered at the position level.

FIGURE 1: COVERAGE OF EXAMS FOR WORKERS IN THE DEPARTMENTAL SERVICE IN DC

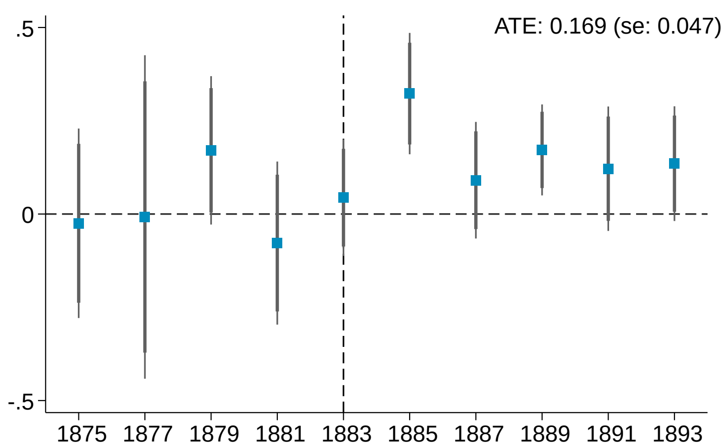
(A) TOTAL NUMBER OF WORKERS (000S)

(B) % IN POSITIONS THAT BECAME SUBJECT TO EXAMS AFTER THE REFORM



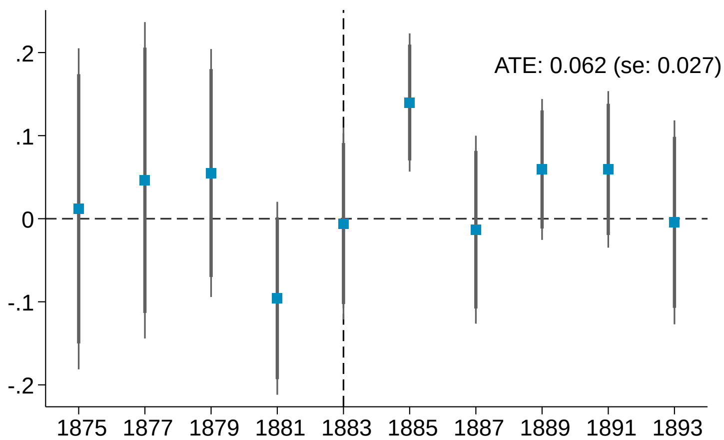
Notes: Panel (a) shows the total number of workers in the Executive Departments in Washington, DC. Panel (b) shows the proportion of such employees who worked in positions that became subject to exams after 1883 (that is, in "treated" positions). A position is coded as being subject to exams if it required an exam at any point from 1883 to 1893.

FIGURE 2: EFFECTS OF THE REFORM ON THE SUMMARY INDEX OF EMPLOYEES' SOCIOECONOMIC BACKGROUND, EVENT-STUDY



Notes: The dependent variable is a summary index of employees' socioeconomic background. The index is computed using the approach in Kling *et al.* (2007), and it combines information on parental wealth rank, parental occupations, parental literacy, own and parental birthplace and race. The variables composing the index are normalized such a higher value corresponds to a higher socioeconomic status. The figure reports estimates of event-study specifications as described in equation 2. The figure shows the estimated coefficients around 90 and 95% confidence intervals (based on standard errors clustered at the position level). All specifications include hiring year, position and state "whence appointed" fixed effects. The sample is restricted to newly hired employees in the Executive Departments in DC.

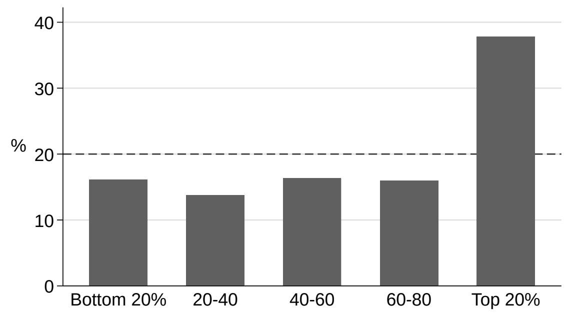
FIGURE 3: EFFECTS OF THE REFORM ON EMPLOYEES' PARENTAL WEALTH RANKS, EVENT-STUDY



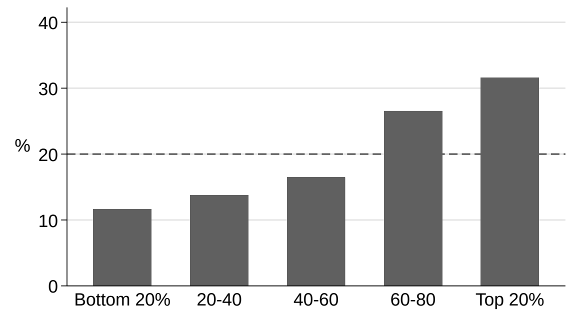
Notes: The dependent variable is the rank of a bureaucrat father in the US national wealth distribution. These ranks are computed separately by census year (1860 and 1870) and by age (i.e. relative to all fathers of the same age). The figure reports estimates of event-study specifications as described in equation 2. The figure shows the estimated coefficients around 90 and 95% confidence intervals (based on standard errors clustered at the position level). All specifications include hiring year, position and state "whence appointed" fixed effects. The sample is restricted to newly hired employees in the Executive Departments in DC.

FIGURE 4: PARENTAL WEALTH QUINTILES OF GOVERNMENT EMPLOYEES, BEFORE AND AFTER THE REFORM

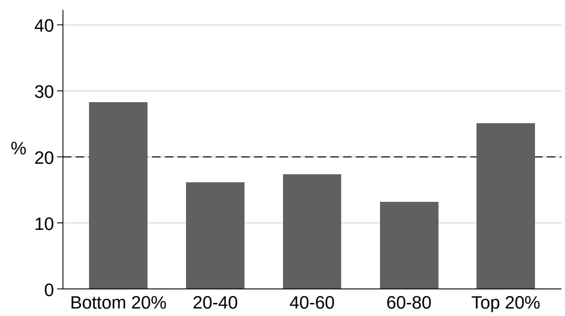
(A) SUBJECT TO EXAMS, BEFORE



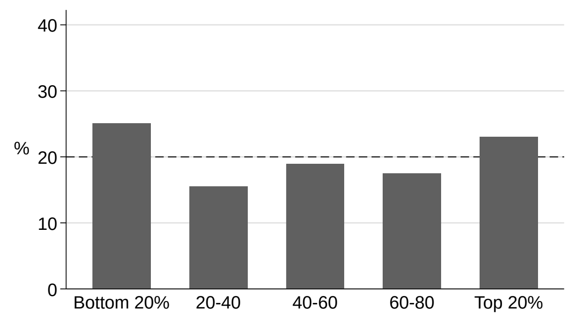
(B) SUBJECT TO EXAMS, AFTER



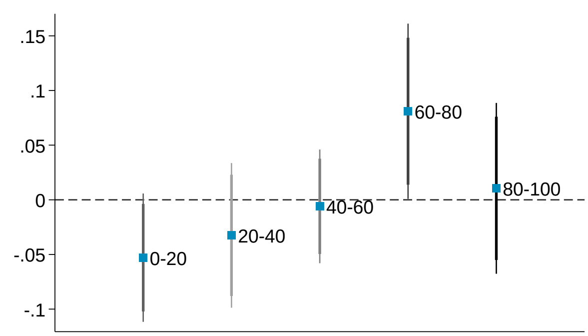
(C) NOT SUBJECT TO EXAMS, BEFORE



(D) NOT SUBJECT TO EXAMS, AFTER



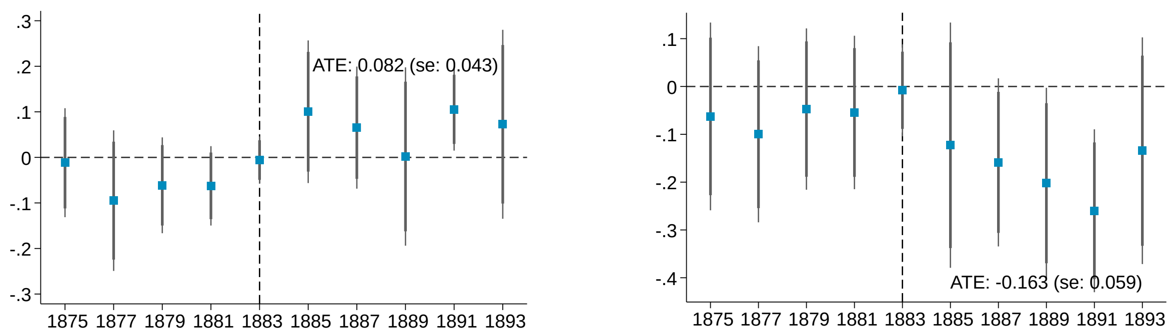
(E) DIFFERENCE-IN-DIFFERENCES ESTIMATES



Notes: Panels (a) to (d) show the distribution of workers across parental wealth quintiles for workers in positions subject and non-subject to exams, before and after the reform. Workers are classified as being in a position subject to exams if they worked in one of the position that became subject to exams in the 1883-1893 period (that is, the "treated" positions). Panel (e) shows difference-in-differences estimates in which the outcome variables are indicators for belonging to different quintiles of the parental wealth distribution. Each coefficient corresponds to a separate regression. The figure shows the estimated coefficients around 90 and 95% confidence intervals (based on standard errors clustered at the position level). All the regressions include hiring year, position and state "whence appointed" fixed effects. The sample is restricted to newly hired employees in the Executive Departments in DC.

FIGURE 5: EFFECTS OF THE REFORM ON EMPLOYEES' OCCUPATIONAL BACKGROUND, EVENT-STUDY

(A) PROFESSIONAL OCCUPATION **(B) WHITE-COLLAR NON-PROFESSIONAL**



Notes: The dependent variable in panel (a) is an indicator that takes a value of one if a worker was employed in a professional occupation prior to joining the civil service. The dependent variable in panel (b) is an indicator that takes a value of one if a worker was employed in a white-collar non-professional occupation. The sample is restricted to individuals who were at least 25 year old at the time we observe them in the census. The figures show the event-study estimates based on equation 2 in the main text. The estimated coefficients are shown around 90 and 95% confidence intervals (based on standard errors clustered at the position level). All the regressions include hiring year, position and state "whence appointed" fixed effects. The sample is restricted to newly hired employees in the Executive Departments in DC.