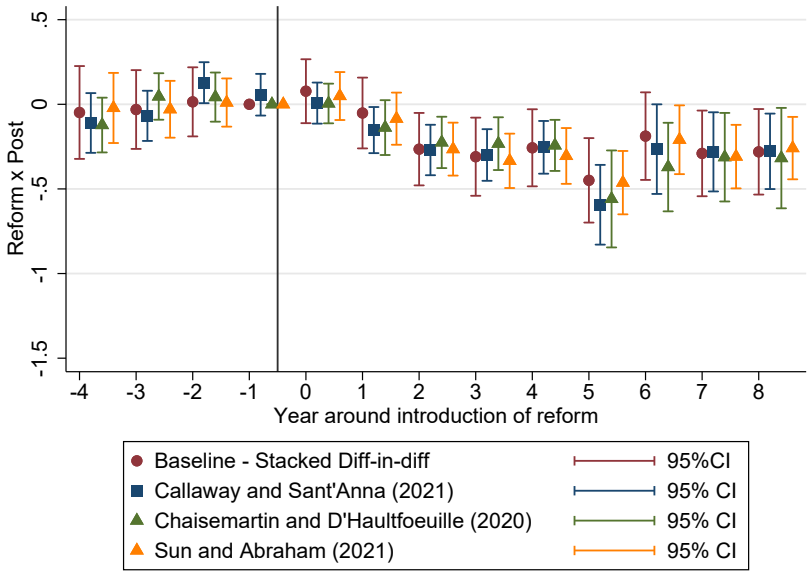
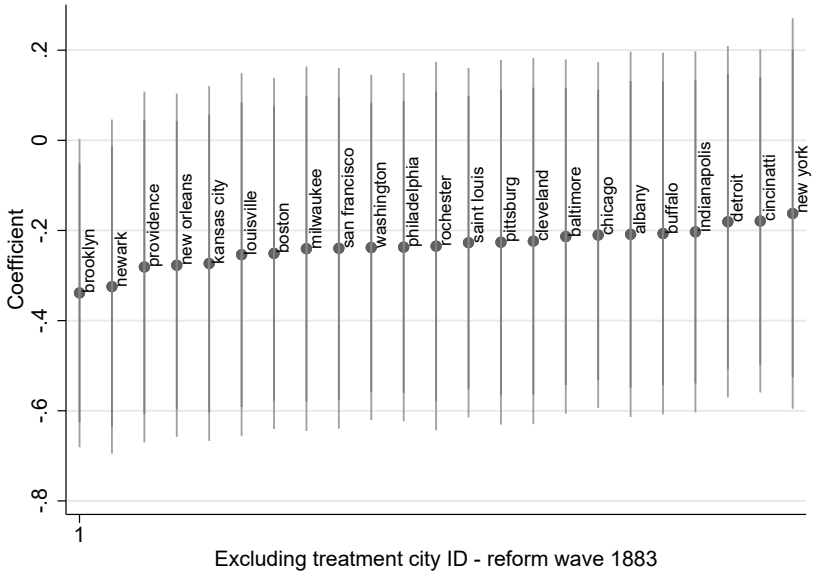
A Online Appendix for "Strengthening State Capacity: Civil Service Reform and Public Sector Performance during the Gilded Age," by Abhay Aneja and Guo Xu

Figure A1: Event study evidence – stacked DD and twoway FEs



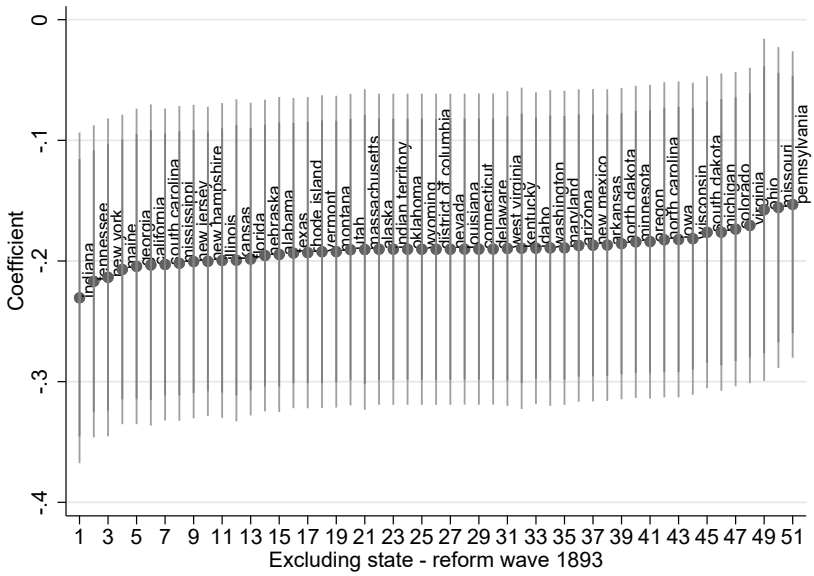
Notes: Figure compares the dynamic estimates for the stacked difference-in-differences with twoway FE panel estimates following Callaway and Sant'Anna (2021); de Chaisemartin and D'Haultfœuille (2020); Sun and Abraham (2021). The estimations of the twoway FEs are implemented using the Stata commands `csdid`, `did_multiplegt`, and `eventstudyinteract`. Reporting 95% confidence intervals.

Figure A2: Delivery errors – robustness of the 1883 reform wave results, dropping one treatment city at a time



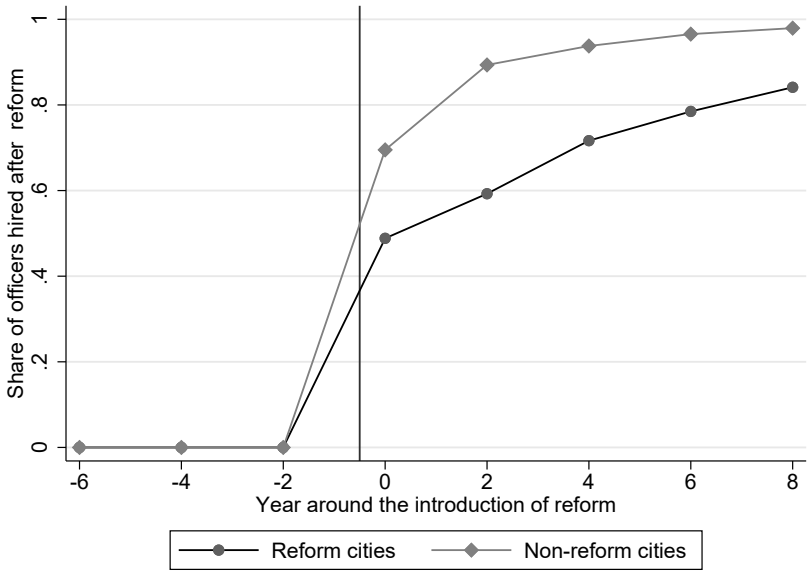
Notes: Reporting coefficients of the Reform × Post estimate, restricting the analysis to only the 1883 reform period (Table 2, column 2) and excluding each of the 23 treatment cities, one at a time. Reporting 95% confidence intervals with light gray vertical lines, and 90% confidence intervals with dark gray lines. Standard errors clustered at the city-level.

Figure A3: Delivery errors – robustness of the 1893 reform wave results, dropping one state at a time



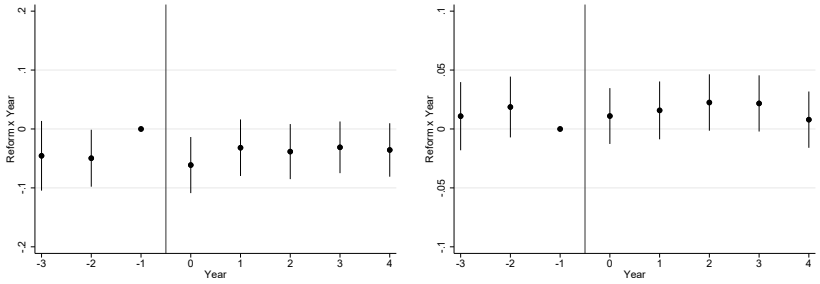
Notes: Reporting coefficients of the Reform × Post estimate, restricting the analysis to only the 1893 reform period (Table 2, column 2) and excluding each state one at a time. Reporting 95% confidence intervals with light gray vertical lines, and 90% confidence intervals with dark gray lines. Standard errors clustered at the city × reform-wave level.

Figure A4: Share of civil servants hired after the reform event, by reform status



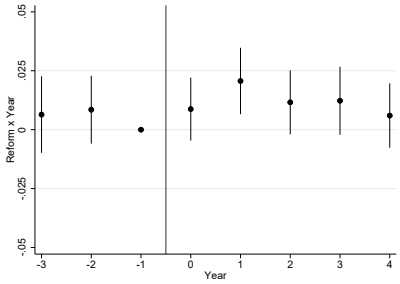
Notes: Figure shows the share of civil servants who were hired after the reform event (i.e., prior to 1883 or 1893, respectively) in reformed and unreformed cities. The year of entry is measured as the first year in which a given individual is observed in the personnel data. The outcome is shown relative to the year after the introduction of the reform.

Figure A5: Individual-level characteristics of hires and civil service reform

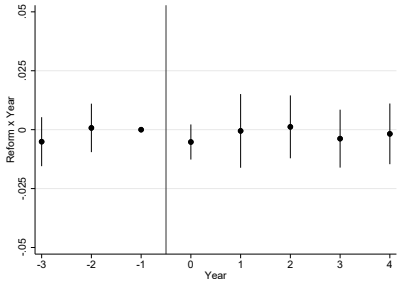


(a) Same state

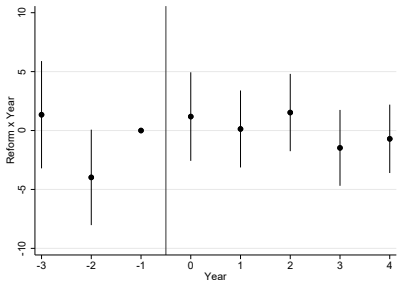
(b) Foreign-born



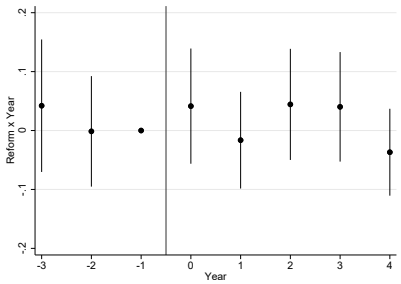
(c) German



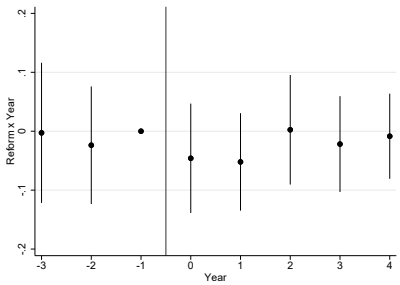
(d) Irish



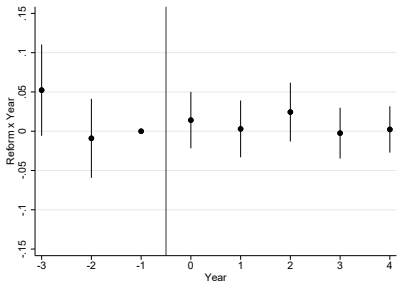
(e) Age



(f) Literacy



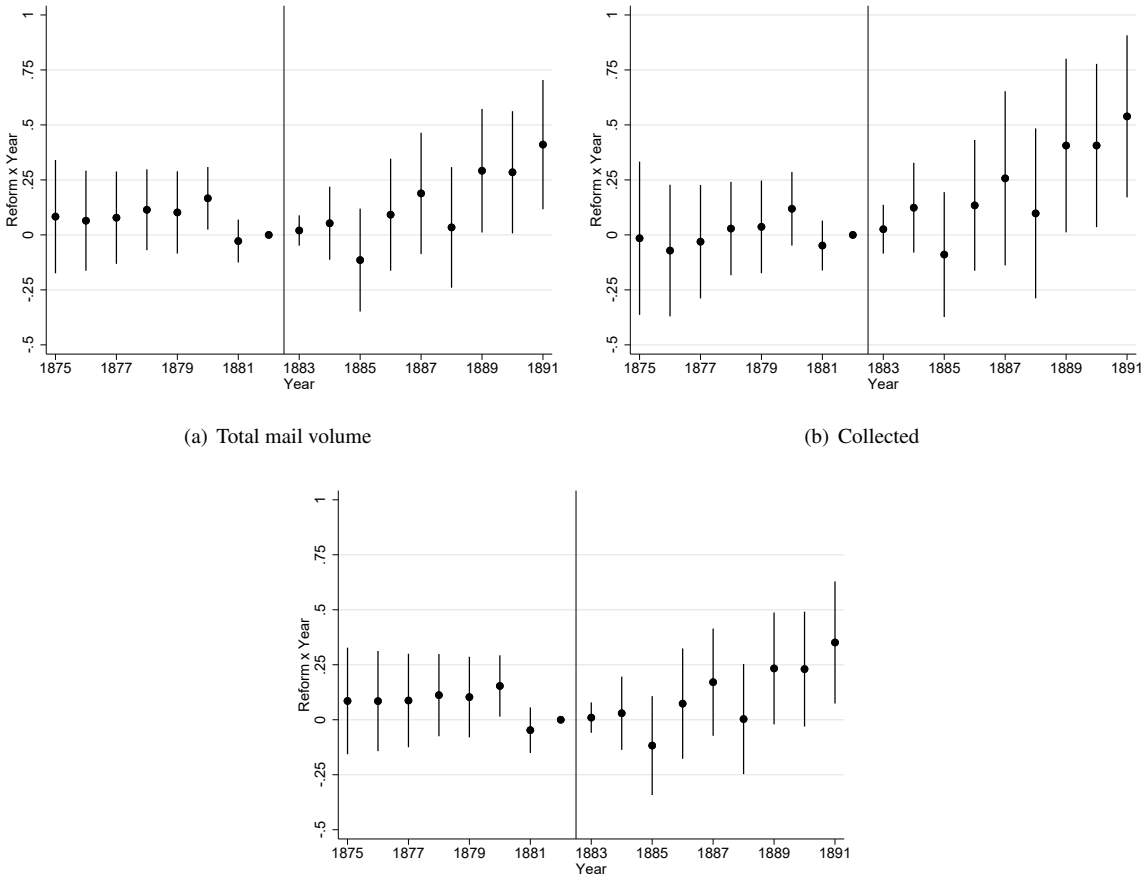
(g) Female



(h) White

Notes: Figure reports estimates from an augmented version of Equation 1 (corresponding to Table 5), where the estimated difference between the treatment and control cities is allowed to vary for each year around the introduction of the reform. Outcome for each estimation listed below each of six panels. Reporting 95% confidence intervals. Standard errors clustered at the city $\times$ reform year-level.

Figure A6: Total mail volume, collections, deliveries and civil service reform (1883 reform wave)



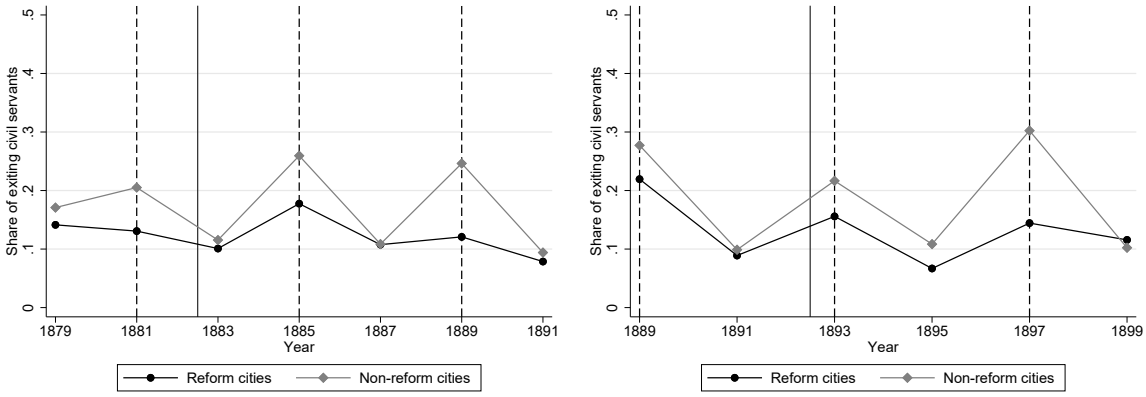
(a) Total mail volume

(b) Collected

(c) Delivered

Notes: Figure reports estimates from an augmented version of Equation 1 (corresponding to Table 6, columns 2–4), where the estimated difference between the treatment and control cities is allowed to vary for each year around the introduction of the reform. Outcome for each estimation listed below each of 3 panels. Reporting 95% confidence intervals. Standard errors clustered at the city-level.

Figure A7: Exit rate for reform vs. non-reform cities around the reform, 1883 and 1893 reform waves

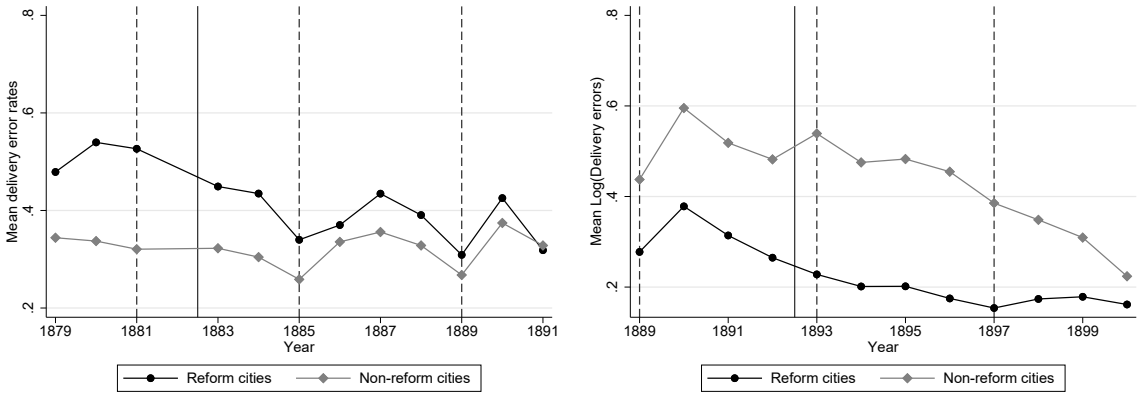


(a) 1883 reform wave

(b) 1893 reform wave

Notes: Figure shows the raw share of civil servants who exit the postal service in a given year. Exit is defined as observing an individual for the last time in the personnel data. Panel (a) focuses on the 1883 reform wave, and Panel (b) focuses on the 1893 reform wave. Dashed lines mark presidential election years and the solid line marks the reform.

Figure A8: Delivery error rates for reform vs. non-reform cities, 1883 and 1893 reform waves

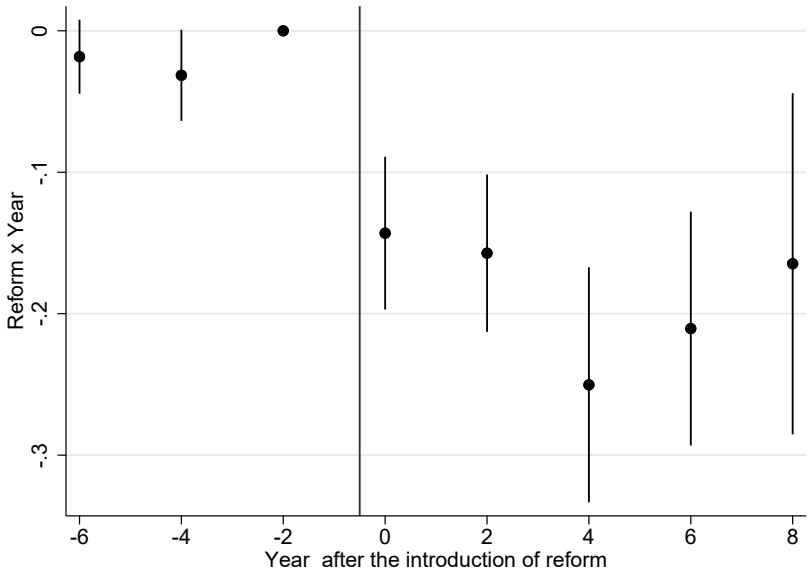


(a) 1883 reform wave

(b) 1893 reform wave

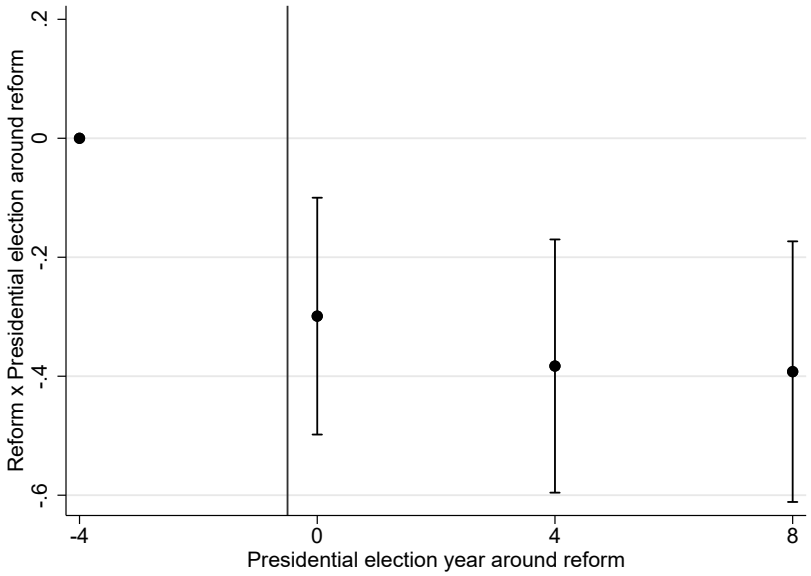
Notes: Figure shows the raw mean delivery error rates in a given year. Panel (a) focuses on the 1883 reform wave, and Panel (b) focuses on the 1893 reform wave. Dashed lines mark presidential election years and the solid line marks the reform.

Figure A9: Civil service reform and exit, pre-reform entrants



Notes: Figure reports estimates from an augmented version of Equation 1 (corresponding to Table 9, column 2), where the estimated difference between the treatment and control cities is allowed to vary for each year around the introduction of the reform. Reporting 95% confidence intervals. Standard errors clustered at the city × reform year-level.

Figure A10: Number of political newspapers in reform vs. non-reform cities around reform years



Notes: Figure reports estimates from an augmented version of Equation 1 (corresponding to Table 10, column 1), where the estimated difference between the treatment and control cities is allowed to vary for each year around the introduction of the reform. Reporting 95% confidence intervals. Standard errors clustered at the city × reform year-level.

Table A1: Descriptive statistics of reformed and unreformed post-offices – additional characteristics

	(1)	(2)	(3)	(4)	(5)
	Mean	Difference treatment-control in reform wave			
	control	1883	1894-1892	1893	Pooled
County-level (all in 100%)					
Share of federal government workers	0.186	0.202	-0.017	0.057	0.060
		(0.210)	(0.046)	(0.044)	(0.040)
Share of postal workers	0.069	0.020	0.020	0.010	0.011
		(0.016)	(0.006)	(0.002)	(0.002)
Share of state government workers	0.144	0.136	0.078	0.021	0.032
		(0.019)	(0.009)	(0.004)	(0.004)
Share of telephone workers	0.003	0.013	0.012	0.003	0.004
		(0.002)	(0.002)	(0.000)	(0.000)
Share of telegraph workers	0.067	0.042	0.038	0.032	0.033
		(0.010)	(0.010)	(0.003)	(0.002)
Share of railway workers	0.934	0.080	0.617	0.350	0.354
		(0.132)	(0.156)	(0.054)	(0.049)
Share of education workers	0.761	-0.172	0.008	0.176	0.143
		(0.028)	(0.038)	(0.017)	(0.016)
Total number of counties	1,122	477	1,050	994	1,134
- of which with treatment:	0	23	28	431	464

Notes: Table reports additional census characteristics. Column 1 shows the mean for the unreformed (control) cities. Columns 2-4 show the difference between reformed and unreformed cities. Column 5 shows the pooled difference, conditional on reform wave FEs. Observation counts report the maximum number of counties. See Appendix B for a description of the data sources. Robust standard errors are reported in parentheses.

Table A2: Robustness of inference to alternative clustering of standard errors – all outcomes

	(1)	(2)	(3)	(4)	(5)	
		Estimated standard error (SE) clustering at:				
Dependent variable	Coefficient estimate	City × Reform-wave	City	County × Reform-wave	County	
Log(Number of delivery errors)	-0.200	(0.0804)	(0.0803)	(0.0819)	(0.0815)	Table 2, Column 3
Log(Volume/carrier)	0.137	(0.0683)		(0.0689)		Table 3, Column 2
Log(Cost/volume)	-0.129	(0.0684)		(0.0690)		Column 4
Log(Postal staff)	0.017	(0.0406)	(0.0405)	(0.0402)	(0.0401)	Table 4, Column 2
Delivery errors/staff	-0.057	(0.0222)	(0.0219)	(0.0220)	(0.0217)	Column 4
Same state	-0.014	(0.0144)	(0.0130)	(0.0144)	(0.0132)	Table 5, Panel A, Col 1
Foreign-born	0.007	(0.0072)	(0.0065)	(0.0073)	(0.0067)	Column 2
German	0.007	(0.0041)	(0.0035)	(0.0042)	(0.0036)	Column 3
Irish	-0.001	(0.0035)	(0.0033)	(0.0035)	(0.0032)	Column 4
Age	0.826	(1.0379)	(0.9325)	(1.0476)	(0.9311)	Panel B, Column 1
Literacy	-0.002	(0.0267)	(0.0243)	(0.0254)	(0.0227)	Column 2
Female	-0.015	(0.0263)	(0.0241)	(0.0268)	(0.0245)	Column 3
White	-0.000	(0.0122)	(0.0121)	(0.0122)	(0.0118)	Column 4
Log(Total volume)	0.129	(0.1302)		(0.1298)		Table 6, Column 2
Log(Collected)	0.257	(0.1658)		(0.1675)		Column 3
Log(Delivered)	0.104	(0.1205)		(0.1201)		Column 4
Exit rate	-0.130	(0.0532)	(0.0533)	(0.0532)	(0.0532)	Table 7, Column 2
Mean experience	0.779	(0.1376)	(0.1376)	(0.1375)	(0.1375)	Column 5
Error rate	-0.075	(0.0414)	(0.0410)	(0.0413)	(0.0409)	Table 8, Column 1
Log(Volume/carrier)	0.194	(0.0717)		(0.0717)		Column 2
Log(Cost/volume)	-0.157	(0.0731)		(0.0720)		Column 3
Exit	-0.160	(0.0216)	(0.0243)	(0.0217)	(0.0243)	Table 9, Column 2
Newspaper number – political	-0.195	(0.0786)	(0.0788)	(0.0781)	(0.0782)	Table 10, Column 1
Newspaper number – independent	0.085	(0.0426)	(0.0431)	(0.0425)	(0.0429)	Column 2
Circulation – political	-0.266	(0.7787)	(0.7822)	(0.8119)	(0.8177)	Column 3
Circulation – independent	0.800	(0.4774)	(0.4778)	(0.4784)	(0.4790)	Column 4
Circulation share – political	-0.044	(0.0252)	(0.0260)	(0.0257)	(0.0264)	Column 5

Notes: Summary table that shows for each outcome reported in the paper the robustness of the standard errors to alternative levels of clustering. Each row summarizes the results for one outcome. Column 1 shows the point estimate of the main coefficient of interest. Columns 2–5 report the different standard errors. Column 2 is the preferred level of clustering.

Table A3: Civil service reform and delivery errors – robustness to alternative samples (I)

	(1)	(2)	(3)
	Log(Number of delivery errors)		
Mean of dep. var	4.459	4.285	5.321
Reform × Post	-0.194	-0.149	-0.234
	(0.063)	(0.058)	(0.191)
Reform wave × City FEs	✓	✓	✓
Reform wave × Year × State FEs	✓	✓	✓
Reform wave × Year FEs × Postal employment	✓	✓	✓
Sample	1883 & 1893	1883–1893	Free delivery
Observations	18,856	84,008	8,338

Notes: Relating delivery errors to the civil service reform in a stacked event-study design, centered around each reform year. Column 1 corresponds to the baseline specification in Table 2. The sample in Column 2 includes the 29 cities that become reformed between 1884–1892 due to having (endogenously) passed the 50-employees threshold. Column 3 restricts the sample to cities that have free postal delivery services during our study period. The unit of observation is the reform wave × city × year. Reform is a dummy that is 1 if the city was covered by the civil service reform in the reform wave, and 0 otherwise. Post is a dummy that is 1 if the year is after the reform year of interest. Standard errors clustered at the city × reform-wave level.

Table A4: Civil service reform and expansion of complementary infrastructure

	(1)	(2)	(3)	(4)
	Log Δ employment 1880-1900			
	Telegraph		Railway	
Mean of dep. var	0.0255	0.0255	0.400	0.400
Reform × Post	0.001	0.004	0.096	0.102
	(0.003)	(0.004)	(0.066)	(0.086)
Reform wave × State FEs	✓	✓	✓	✓
Reform wave × Postal employment	✓		✓	
Reform wave × PDS controls		✓		✓
Observations	1,669	1,669	1,669	1,669

Notes: Long regression relating changes in (log) employment among telegraph (columns 1-2) and railway (columns 3-4) workers to the expansion of the civil service reform. The employment numbers are measured for the county in which a city is located, and are computed based on the full-count Decennial Censuses for 1880 and 1900. The unit of observation is the reform wave × city. Reform is a dummy that is 1 if the city was covered by the civil service reform in the reform wave, and 0 otherwise. Post is a dummy that is 1 if the year is after the reform year of interest. Standard errors clustered at the city × reform-wave level.

Table A5: Expansion of postal inspectors and civil service reform

	(1)	(2)	(3)
	Number of inspectors		Any inspector
Mean of dep. var	0.0309	0.0309	0.00606
Reform × Post	0.012	0.006	-0.003
	(0.015)	(0.012)	(0.002)
Reform wave × City FEs	✓	✓	✓
Reform wave × State FEs	✓	✓	✓
Reform wave × Postal employment	✓		
Reform wave × PDS controls		✓	✓
Observations	17,980	17,980	17,980

Notes: Relating total postal employment of inspectors (inspectors of mail depredations, inspectors of money-order services, inspectors of free delivery service) to the civil service reform in a stacked event-study design. The unit of observation is the reform wave × city × year. Reform is a dummy that is 1 if the city was covered by the civil service reform in the reform wave, and 0 otherwise. Post is a dummy that is 1 if the year is after the reform year of interest. Column 1 includes the (time-interacted) total number of postal workers as control variable. Columns 2–3 include time-interacted controls selected via post-double-selection (PDS, see Belloni et al. (2014)). Standard errors clustered at the city × reform-wave level.

Table A6: Civil service reform and delivery errors – robustness to alternative samples (II)

	(1)	(2)	(3)	(4)
	Log(Number of delivery errors)			
Mean of dep. var	4.459	4.304	4.430	4.423
Reform 1883 × Post	-0.194	-0.191	-0.193	-0.192
	(0.063)	(0.069)	(0.064)	(0.065)
Sample	Baseline	No port cities	No reformed customs offices	No municipal reforms
Reform wave × City FEs	✓	✓	✓	✓
Reform wave × Year FEs × State FEs	✓	✓	✓	✓
Reform wave × Year × Postal employment	✓	✓	✓	✓
Observations	18,856	17,063	18,700	18,046

Notes: Relating delivery errors to the civil service reform in a stacked event-study design, centered around each reform year. Column 1 presents the results for the baseline sample. Column 2 drops all port cities (i.e., cities with a customs office). Column 3 drops all cities that also experienced civil service reform within the customs office. Column 4 drops all cities that experienced a municipal civil service reform episode during the study period (data from Rauch (1995)). The unit of observation is the reform wave × city × year. Reform is a dummy that is 1 if the city was covered by the civil service reform in the reform wave, and 0 otherwise. Post is a dummy that is 1 if the year is after the reform year of interest. Standard errors clustered at the city × reform-wave level.

Table A7: Salary and civil service reform

	(1)	(2)	(3)
	Log(Annual salary)		
Mean of dep. var	6.146	6.427	5.696
Reform × Post	0.053	0.045	0.058
	(0.041)	(0.041)	(0.064)
Reform wave × City FEs	✓	✓	✓
Reform wave × Experience FEs	✓	✓	✓
Reform wave × Year × State FEs	✓	✓	✓
Reform wave × Year × Job FEs	✓	✓	✓
Reform wave × Year × Controls	✓	✓	✓
Sample	Full sample	Pre-reform entrants	New hires
Observations	142,770	68,191	62,642

Notes: Relating individual-level (log) annual salary to the civil service reform in a stacked event-study design. The unit of observation is an individual × reform wave × city × year. Reform is a dummy that is 1 if the city was covered by the civil service reform in the reform wave, and 0 otherwise. Post is a dummy that is 1 if the year is after the reform year of interest. Column 1 reports results based on the full sample. Column 2 restricts the sample to individuals who entered before the reform, and column 3 restricts the sample to new hires. All specifications include (time-interacted) total postal employment and (log) city population as controls. Standard errors clustered at the city × reform-wave level.

Table A8: Individual-level census match rate of hires and civil service reform

	(1)	(2)	(3)
	Hire is matched to census=1		
Mean of dep. var	0.332	0.332	0.332
Reform × Post	-0.010	-0.007	-0.001
	(0.012)	(0.012)	(0.016)
Reform × City FEs	✓	✓	✓
Reform × Year × State FEs	✓	✓	✓
Reform × Year × Job FEs		✓	✓
Reform × Controls			✓
Observations	68,977	68,977	68,977

Notes: Relating a dummy for whether a recruited civil servant matches to the Decennial Census to the implementation of the civil service reform. Newly recruited civil servants are identified as workers first observed in the personnel data (source is the Official Registers series). To avoid truncation (since all workers are first observed in the earliest year of our data), we exclude the first year of our personnel records, thus covering 1879–1901. The unit of observation is an individual × reform wave × city × year. Reform is a dummy that is 1 if the city was covered by the civil service reform in the reform wave, and 0 otherwise. Post is a dummy that is 1 if the year is after the reform year of interest. All specifications include (time-interacted) total postal employment and (log) city population as controls. Standard errors clustered at the city × reform-wave level.

Table A9: Civil service reform and mail delivery/collection, disaggregated by type (1883 reform wave)

Panel A: Volume delivered (in log)	(1)	(2)	(3)
	Letters	Postal cards	Newspapers
Mean of dep. var	13.68	12.15	13.19
Reform 1883 × Post	0.114	0.178	0.146
	(0.123)	(0.143)	(0.129)
Observations	2,945	2,945	2,945
Panel B: Volume collected (in log)	(1)	(2)	(3)
	Letters	Postal cards	Newspapers
Mean of dep. var	13.17	11.75	11.04
Reform 1883 × Post	0.301	0.089	0.124
	(0.162)	(0.177)	(0.233)
City FEs	✓	✓	✓
Year FEs	✓	✓	✓
Year FEs × PDS controls	✓	✓	✓
Observations	2,945	2,945	2,945

Notes: Relating mail delivered and collected through the city free delivery service to civil service reform, focusing separately on each type of mail item for the 1883 reform wave. The unit of observation is the city × year, and the sample period is 1875–1891. Reform 1883 is a dummy that is 1 if the city was covered by civil service reforms in 1883, and 0 otherwise. Post is a dummy that is 1 if the year is after civil service reform takes effect under the Pendleton Act. Standard errors clustered at the city-level.

Table A10: Delivery errors and aggregate volume, before and after the reform (1883 reform wave)

	(1)	(2)
	Log(Delivery errors)	
Mean of dep. var	8.747	5.574
Log(Aggregate volume)	1.899	-0.170
	(0.397)	(0.217)
Log(Aggregate volume) × Post	-0.515	0.309
	(0.308)	(0.225)
Sample	Reformed	Unreformed
City FEs	✓	✓
Year FEs	✓	✓
Year FEs × Postal employment	✓	✓
Difference Log(Aggregate volume) × Post	-0.825	
	(0.368)	
Observations	275	2,210

Notes: Relationship between delivery errors and aggregate mail volume, broken down by reform vs. unreformed cities (columns 1–2), before and after federal civil service reform. The sample period is 1875–1891. The unit of observation is the city × year. Reform is a dummy that is 1 if the city was covered by the civil service reform under the Pendleton Act, and 0 otherwise. Post is a dummy that is 1 if the year is after civil service reform takes effect under the Pendleton Act. Standard errors clustered at the city-level.

B Data sources of covariates

Table B1: Description of baseline covariates

Variable	Description	Source
Post office staff	Number of post officers (clerks, carrier, postmaster) in city	Official Registers of the U.S.
City-level population	Total city/town population	US census
Land-grant universities	Number of land-grant universities in county	IPEDS
Western Union office	Dummy for whether a city has a Western Union office in 1874	W. Union telegraph directory
Year post office opened	The year the post office was established in city	Report of Postmaster General
Southern state	AL, AR, DE, FL, GA, KY, LA, MD, MS, MO, NC, OK, SC, TN, TX, VA, WV	Own calculation
Distance to D.C.	Distance (in miles) between town/city and D.C.	Own calculation
Latitude	Town/city latitude	Google Maps API
Longitude	Town/city longitude	Google Maps API
County-level population	Total county population	US census, Haines (2010)
Foreign-born share	Sbare of foreign-born in the county	US census, Bazzi et al. (2020)
Urban share	Urban population share in county	US census, Bazzi et al. (2020)
Non-white share	County-level share of non-white population	US census, Bazzi et al. (2020)
Frontier county	Dummy for whether a county is a "frontier county"	Bazzi et al. (2020)
Railway	Dummy for whether a county has railroad access	Bazzi et al. (2020)
Canal	Dummy for whether a county has canal access	Bazzi et al. (2020)
Manufacturing establishments	County-level share of manufacturing establishments	US census, Haines (2010)
Share literate	Share of literate in the county	US census
Labor force participation rate	Share of county population in labor force	US census
Occupational income score	County-level average occupational income score	US census
Share of Democrat votes	County-level Democrat congressional vote share	Clubb et al. (2006)
Share of Republican votes	County-level Republican congressional vote share	Clubb et al. (2006)
Turnout	County-level turnout	Clubb et al. (2006)
Number of party switches	Share of elections between 1872-1882 in which county's majority party changes	Clubb et al. (2006), own calculation

Variable	Description	Source
Share of workers in education	County-level share of workers in educational services (IND1950=888)	US census
Share of workers in federal government	County-level share of workers in federal public administration (IND1950=916, excluding postal service)	US census
Share of workers in state government	County-level share of workers in state public administration (IND1950=926)	US census
Share of workers in telephone	County-level share of workers in telecommunications – telephone (IND1950=578)	US census
Share of workers in telegraphy	County-level share of workers in telecommunications – telegraph (IND1950=579)	US census
Share of workers in railway	County-level share of workers in railroads and railway express service (IND1950=506)	US census
Share of workers in post office	County-level share of workers in postal service (IND1950=906)	US census

Notes: Summary description of all covariates (see also Table 1 and Appendix Table A1) and their data sources.

Table B2: Post-double-selection covariates for each regression

Table	Column	Baseline covariates selected via Post-Double-selection (Total #)
Table 2	Column 3	Post office staff, Log(City-level population), Share of workers in education, Western Union office (4)
	Column 4	Post office staff, Log(City-level population), Share of workers in state government, Manufacturing establishments (4)
	Column 5	Post office staff, Log(City-level population), Share of employees in education (3)
Table 3	Column 2	Post office staff, Log(City-level population) (2)
	Column 4	Post office staff, Log(City-level population) (2)
Table 4	Column 2	Log(City-level population), Log(County-level population), Share of employees in education (3)
	Column 4	Log(City-level population), Share of employees in education, Western Union office (3)
Table 6	Columns 2, 4	Post office staff, Log(City-level population), Share of employees in railway (3)
	Column 3	Post office staff, Log(City-level population) (2)
Table 7	Column 2	Post office staff, Manufacturing establishments (2)
	Column 3	Post office staff, Urban share, Share of employees in telephone, Log(County-level population), Manufacturing establishments (5)
	Column 6	Post office staff, Year post office opened, Log(City-level population), Manufacturing establishments (4)
Table 8	Column 1	Post office staff, Log(City-level population) (2)
	Column 2	Post office staff, Log(City-level population), Share of workers in telephone, Share of workers in railway (4)
	Column 3	Post office staff, Log(City-level population), Share of workers in telephone (3)
Table 10	Columns 1-2	Post office staff, Log(Occupational income score), Western Union office, Log(City-level population), Log(County-level population) (5)
	Column 3	Post office staff, Log(Occupational income score), Western Union office, Log(City-level population), Log(County-level population), Manufacturing establishments (6)
	Column 4	Post office staff, Log(Occupational income score), Western Union office, Canal, Log(Distance to D.C.), Log(City-level population), Log(County-level population), Manufacturing establishments (8)
	Column 5	Post office staff, Log(Occupational income score), Western Union office, Log(City-level population), Log(County-level population) (5)

Notes: Covariates selected via Post-Double-selection (Belloni et al., 2014) for each regression specification.

C Census linking

We match the personnel records from the Official Registers of the United States (the "Registers") to the U.S. full count Decennial Census to obtain additional individual background characteristics. Each postal worker is matched based on the full name (first name, middle name/initial, last name), birth state, and current state of employment. We proceed by using different combinations, successively relaxing the matching restrictions:

- Step 1: First name + middle name + last name + birth state + current state
- Step 2: First name + middle name initial + last name + birth state + current state
- Step 3: First name + last name + birth state + current state
- Step 4: First name + last name + birth state

Given the limited number of variables that are available for matching, we pursue a conservative approach to ensure we do not overmatch by linking incorrect individuals (i.e., false positives). To start with, we always match individual names exactly. Second, we discard candidate matches of Census respondents who were younger than 18 and older than 65 when they are observed in the personnel records. Third, we restrict our matches to only individuals who are uniquely matched to the Census.

Although the Census data should, in principle, allow nearly every postal worker to be matched, match rates obtained through automated linking methods during this historical period rarely exceed 30–40% (Abramitzky et al., 2021). In this setting, there are multiple reasons why a postal worker may not be matched to the Decennial Census. First, transcription errors may occur both in the personnel records and the historical census data. Second, name variations may exist in the Decennial Censuses (e.g., Rick vs. Richard). Third, postal workers with common names residing in populous states will often have multiple potential counterparts in the census (e.g., John Smith from New York), making it difficult to identify the correct individual in the absence of unique identifiers such as social security numbers. Fourth, since the Decennial Census data is only available at a decadal frequency, individuals may have passed away or migrated between the year they were recorded in the personnel records and the year the census was taken. To increase the odds of finding individuals in the census data, we thus link each individual observed in the personnel record to the U.S. Decennial Censuses of 1880 and 1900.[37] Overall, we obtain a match rate of 34%. This match rate is comparable to those obtained in related census-linking exercises (Abramitzky et al., 2021; Aneja and Xu, 2021; Moreira and Perez, 2022a,b)

[37] While aggregate data exists for the 1890 Decennial Census, the micro-level data for the 1890 U.S. Decennial Census is unavailable as the records were destroyed in a fire in 1921.

Table C1: Characteristics of Census-linked vs. non-linked workers

	(1)	(2)	(3)	(4)
	Mean characteristics		Differences	
	Matched	Unmatched	Raw	Conditional
Panel A: Individual characteristics				
Log(Salary)	5.80	5.83	-0.025	0.016
			(0.007)	(0.007)
Clerk	0.59	0.60	-0.011	-0.006
			(0.002)	(0.002)
Same state	0.52	0.66	-0.141	-0.147
			(0.002)	(0.002)
Foreign-born	0.11	0.10	0.008	0.011
			(0.001)	(0.002)
German	0.02	0.02	0.000	0.002
			(0.000)	(0.001)
Irish	0.022	0.021	0.001	0.002
			(0.001)	(0.001)
Panel B: City-level characteristics				
Post office staff	143.41	181.56	-38.150	-26.230
			(1.654)	(1.515)
Log(City population)	10.59	10.82	-0.230	-0.139
			(0.011)	(0.010)
Observations	65,496	126,307	191,803	191,803

Notes: Column 1 shows the mean for the postal workers who could be matched to the Decennial Census. Column 2 shows the mean for the postal workers who could not be matched to the Census. (control) cities. Column 3 shows the raw mean difference between matched vs. unmatched cities. Column 4 shows the mean difference between matched vs. unmatched cities, conditional on year FEs. Observation counts report the maximum number of observations. See Appendix B for a description of the data sources. Robust standard errors are reported in parentheses.

In Appendix Table C1, we compare the traits of Census-linked individuals to those who were not linked. We report differences in means both unconditionally (column 3) and conditional on year FEs (column 4). While matched and unmatched postal workers significantly differ on many observable characteristics, these differences are, in terms of magnitude, relatively small. For example, the raw mean salary difference between matched and unmatched workers is only 1.6%, and differences across other individual characteristics are likewise economically small (Panel A). In Panel B, we report the mean characteristics of the cities in which the matched and unmatched postal workers work. Match rates are significantly higher in smaller post offices and cities.

The observed differences in the characteristics of matched and unmatched officers shown in Appendix Table C1 raise the question whether selection can affect our findings on worker quality (Table 5). If the match rate is significantly associated with the reform rollout, for example, differences in match rates may partly mask any actual change in the characteristics of post-reform hires. Reassuringly, however, we do not find that the match rate is significantly correlated with the rollout of the civil service reform (Appendix Table A8). Finally, we can use inverse probability weights (IPW). IPW is a non-parametric procedure by which individual observations are re-weighted according to the estimated probability that they are part of the matched sample. IPW purges

estimates of selection bias provided that selection is well captured by observable characteristics. Appendix Table C2 shows the reweighted results based on the observed characteristics of Appendix Table C1. As the table shows, the results remain comparable and we do not observe significant changes in the characteristics of hired officers post-reform. In terms of point estimates, the magnitudes remain economically small.

Table C2: Individual-level characteristics of hires and civil service reform – IPW reweighting

	(1)	(2)	(3)	(4)
	Age	Literacy	Female	White
Mean of dep. var	28.40	0.864	0.114	0.966
Reform × Post	0.906	0.001	-0.016	-0.002
	(1.058)	(0.027)	(0.027)	(0.013)
Reform wave × Year × State FEs	✓	✓	✓	✓
Reform wave × Year × Job FEs	✓	✓	✓	✓
Reform wave × City FEs	✓	✓	✓	✓
Reform wave × Controls	✓	✓	✓	✓
Observations	22,465	22,465	22,465	22,465

Notes: Relating individual-level characteristics of recruited civil servants to the implementation of the civil service reform, using inverse probability weights (IPW). The Census-linked sample is reweighted to be representative of the population in terms of salary, occupation (clerk/carrier), being born in the same state as the state of service, and being foreign-born. Newly recruited civil servants are identified as workers first observed in the personnel data. To avoid truncation (since all workers are first observed in the earliest year of our data), we exclude the first year of our personnel records, thus covering 1879–1901. The unit of observation is an individual × reform wave × city × year. Reform is a dummy that is 1 if the city was covered by the civil service reform in the reform wave, and 0 otherwise. Post is a dummy that is 1 if the year is after the reform year of interest. All specifications include (time-interacted) total postal employment and (log) city population as controls. Standard errors clustered at the city × reform-wave level.