Effects on Legislative Reentry of the Introduction of Merit Civil Service Appointments in US States*

Miriam Golden[†] Eugenia Nazrullaeva[‡] Dylan Potts[§]

This version 16 January 2026

*The authors thank Stephane Wolton for discussions and comments at an early stage of this project as well as participants at the Political Economy Workshop at the Norwegian Business School, May 22–24, 2022; the 2022 Applied Economics Workshop at Petralia Sottana, 20–27 July; the 13^{th} Annual Conference of the European Political Science Association, Glasgow, 22–24 June 2023; the Institutions and Political Economy Research Group at the University of Barcelona; the IE Business School, Madrid; the Cener for the Study of Democracy, Development and the Rule of Law, Stanford University; and a Workshop on the Political Economy of Legislatures and Legislators, European University Institute, October 3–4, 2024. We are especially grateful for suggestions from Michael Becher, Gary Cox, Carl Dahlström, Seth Hill, and Hye Young You. For research assistance, we thank Sohaila Abdelaal, Ivan Formichev, Brian Hamel, Diala Husseini, Mamoor Ali Khan, Siyu Liang, Hailee Nguyen, Ali Abdul Rehman Unar, and Nikita Savin. Funding in support of data collection was provided by the Academic Senate of the University of California at Los Angeles, the Department of Economics Warwick Internship Scheme for Economics (WISE) at the University of Warwick, and the European University Institute. The research reported here was pre-registered under the same title with the EGAP Registry, timestamp 02/01/2020 – 12:33, EGAP Registration ID 20200201AA.

[†]Corresponding author. European University Institute and the University of California at Los Angeles (currently Stanford University and the University of California at Los Angeles), golden@ucla.edu

[‡]University of Konstanz, evgeniya.nazrullaeva@uni-konstanz.de (currently University of Liverpool)

[§]London School of Economics and Political Science (currently Independent Scholar) potts@lse.ac.uk

Abstract

Prior research shows that control over patronage appointments confers an electoral advantage. Civil service reform requiring appointments by examination should therefore weaken incumbents. We study the staggered introduction of civil service legislation to the 50 US states between 1900 and 2016 and show that states that adopted reform experienced substantial increases in the proportion of legislators who served consecutive terms, with rates rising by approximately five percentage points in the post-reform period. Although reform weakened party control of legislatures, it paradoxically strengthened individual legislative careers. We discuss possible selection and performance mechanisms consistent with this unexpected result. Our analysis of the limited data available is consistent with selection effects, with more professional new entrants into legislative office entering post-reform. These results contribute a new understanding of the factors that made 20th century US legislative politics a candidate-centered era. [137 words]

Introduction

Control over patronage appointments is believed to advantage those already holding elected office.[1] Analyzing data from US states between 1885 and 1995, Folke et al. [2011] shows that political parties, especially entrenched parties, are less likely to retain majority control of a state legislature after civil service reform adoption. The reason is that "Patronage jobs constitute a valuable resource for the party in power" [Folke et al., 2011, 567], a resource rendered unavailable once a requirement that appointments occur via competitive examinations is enacted. Parties can no longer pressure appointees to the public bureaucracy to contribute campaign funds, nor can they appoint their loyalists to public office by fiat.

In this short note, we extend this research to ask how the adoption of merit civil service regulations over more than a hundred years across the 50 US states affected the political careers of individuals — rather than political parties — elected into state legislatures. Although the abolition of patronage reduced the abilities of incumbent parties to retain legislative majorities, expectations about how individual legislators fare under civil service reform are ambiguous. Perhaps they are replaced wholesale by a more professional post-reform cohort specializing in reelection. Perhaps, instead, adaptable legislators reoptimize, enacting policies that allow them to retain elected office even after they are prohibited from cycling in and out of the bureaucracy.

Using a complete newly-compiled dataset covering more than a century of state legislative elections, we document the surprising result that states that adopted civil service reform experienced substantial *increases* in rates at which individual legislators reentered the subsequent session — on average, increases of more than five percentage points, a sizeable share of the century's total increase. Along with Progressive Era reforms such as direct primaries and changes in ballot type, state adoption of civil service reform later in the 20th century represents a previously-unexplored institutional factor with substantial political consequences for legislative careers.

We probe two different channels that might account for the pattern that we uncover. The first

[1] The research reported in this note was pre-registered; the pre-analysis plan (PAP) is available at [redacted]. Page length limitations prevent us from including it as an appendix but deviations from the PAP are reported in Appendix H.

is post-reform entry into office by more politically ambitious and professional individuals, those more interested in and capable of building legislative careers (a selection effect). The second is the delivery of more visible and consequential legislation, thanks to the greater legislative experience of those elected working in partnership with the more professional post-reform bureaucracy (a performance effect). Data limitations prevent more than suggestive evidence regarding either mechanism, but the data that is available corroborates the importance of selection rather than performance. These provisional results suggest the value of future data collection and analysis.

Theory, Hypotheses, and Measures

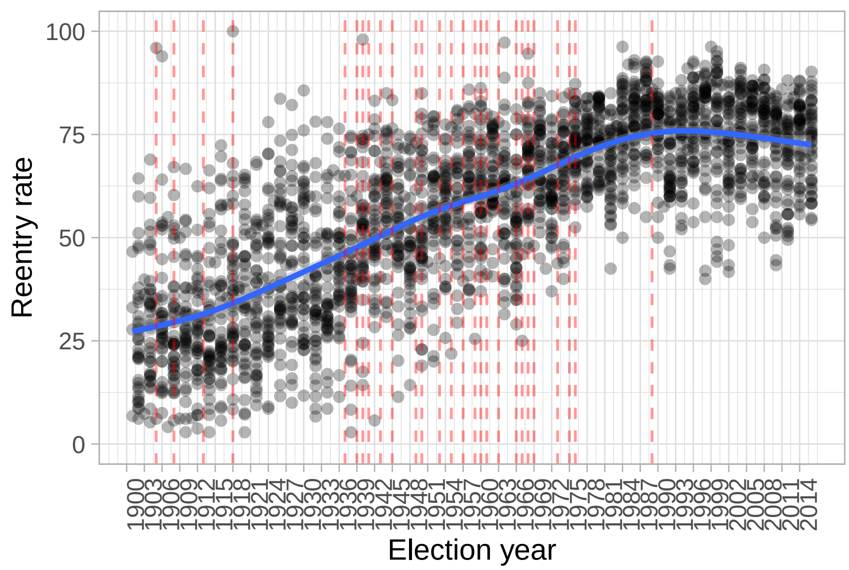
Why would civil service reform alter legislative career patterns? By eliminating the bureaucracy as a rotation option for state legislators, reform increased the opportunity cost of losing an election. Under patronage, a legislative seat functioned as a waystation in a party-managed career: parties rotated loyalists through elected offices, appointed positions, and party roles. Reform transformed the legislature into a destination — a standalone profession rather than one stop in a party-controlled trajectory. This structural shift should attract different types of candidates. The old guard, selected for party loyalty and willingness to rotate, possessed skills — patronage distribution, party service — largely obsolete post-reform. New entrants, by contrast, were selected for traits suited to the emerging candidate-centered environment: constituency service, policy expertise, and independent electoral appeal [Cain et al., 1987, Cox and Katz, 1996, Engstrom and Kernell, 2014]. This institutional reset of career expectations should attract more legislatively-oriented individuals (selection) while simultaneously enabling them to deliver more effectively thanks to a professional bureaucracy (performance). We uncover evidence only for the former.

While the federal government adopted civil service reform with the 1883 Pendleton Act, it took state legislatures another 106 years to pass parallel legislation. New York and Massachusetts legislated reform in the late 19th century; the remaining 48 states adopted it over the course of the 20th century, with the exception of never-reformed Texas. A flurry of adoptions took place after 1939, when Congress amended the Social Security Act to require state level departments administering federal social security funds be staffed with merit appointees [Ash et al., 2022].

Nonetheless, diffusion is generally depicted as endogenous to each state [Ruhil and Camões, 2003]. The heart of reform legislation consisted of the requirement that aspiring entrants sit competitive examinations [Ujhelyi, 2014b], designed to undercut the spoils system [Hoogenboom, 1961] and to improve bureaucratic effectiveness [Lewis, 2007].

Figure 1 depicts state-level average rates of incumbent reentry from 1900 to 2016 for 45 US states, with reform years indicated by vertical red lines. (Discussion of omitted states appears in Appendix A.) Reform is concentrated in the period from 1936 to 1976. Pre-World War II reentry rates remained below 50 percent, and the average US statehouse experienced reentry rates over 50 percent only in the 1950s.

Figure 1: Reentry rates of incumbents in US state legislatures, 1900–2016



Note: Each dot represents a state legislature.
Line with confidence intervals produced by LOESS smoothing.
Red vertical lines indicate civil service reform adoption in at least one state.

We take dates of civil service reform from Ash et al. [2022, table A1, col 4, p. 33], which reviews and adjudicates discrepancies between Ting et al. [2013] and Ujhelyi [2014a]. See Appendix B for a reanalysis using the latter two sets of adoption dates. No state abolished civil service reform once enacted. Our analysis covers as many states as possible for the post-1900 period, when all reforms other than those in New York and Massachusetts take place.

We merge data on civil service reform adoption dates with candidate-level information about state legislative election results. For the latter, we digitize archival data to assemble a complete dataset on all elected officials serving in all state legislatures since 1900 [redacted]. The dataset we create complements [Ansolabehere et al., 2017, Klarner, 2018] which, however, were missing most observations for the first half of the 20th century. For details, see Appendix C.

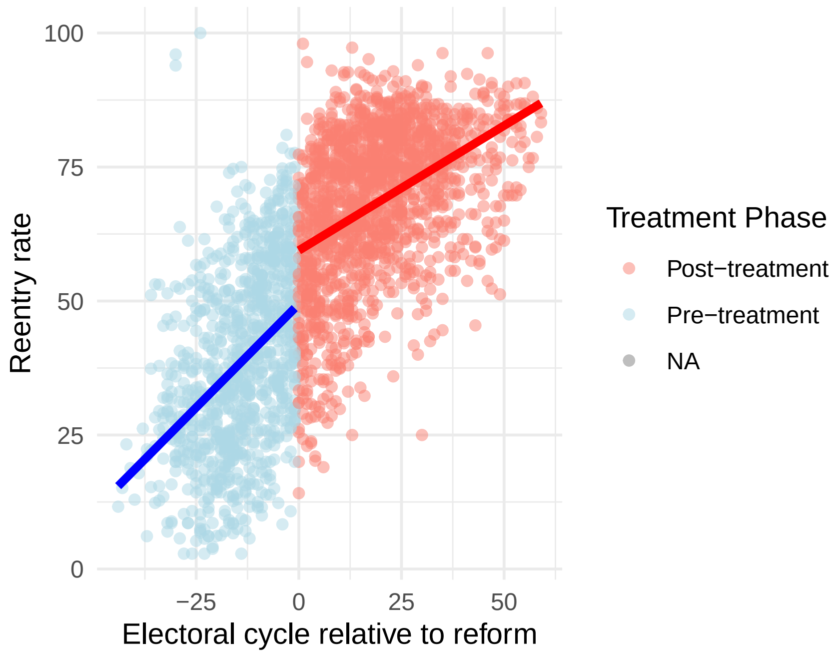
The data we assemble includes legislator name, election year, and state but no other substantive information; we do not have consistent and complete information on partisan affiliation, occupation, age, gender, or any other observable characteristic of legislators. Nor do we have information on which incumbents reran except for incomplete data for nine states (analyzed in Appendix F.1), so while we study *reentry* into the next legislative period, we cannot directly assess *reelection* rates — although to the extent that the decision to rerun is endogenous to the likelihood of winning, our measure of reentry provides a reasonable proxy. The total number of elections analyzed is 2,535, and our dataset includes 127,313 individual legislators.

We measure the variable of theoretical interest, the reentry rate at time t, as the proportion of legislators serving in the $t-1$ legislature who also serve in the legislature at time t.

As documented in Figure 1, state legislatures in the Unites States see increasing reentry rates of individual representatives during the 20th century, with a clear upward trend until the late 1980s. In Figure 2, we show the relationship between civil service reform and reentry rates after pooling the data from all elections and all states. There is a visible and large discontinuity between elections held before and after reform. We observe a roughly 10 percentage point difference in reentry rates at the discontinuity. The figure also shows that rates increase both pre- and post-treatment. The goal of our statistical analysis is to partial out the contribution that civil service reform makes to this underlying increase.

Reform was staggered across 45 states over more than a century, making the unconditional parallel trends assumption required for modern difference-in-differences (DiD) methods [Roth et al., 2023, Xu, 2023] implausible a priori. States that reformed earlier are likely to differ systematically from later adopters in ways that affect baseline reentry trajectories. We therefore employ PanelMatch [Imai et al., 2023], which conditions on pre-treatment outcome histories rather than assuming unconditional parallel trends. This approach asks whether states with sim-

Figure 2: Pre- and post-treatment reentry rates



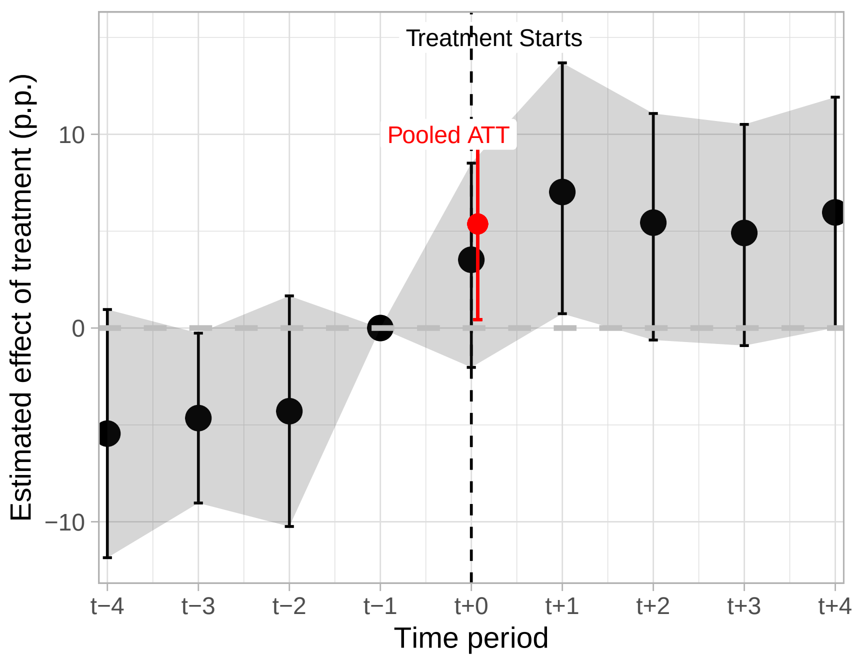
ilar recent reentry patterns diverge after one adopts reform. While no observational method can definitively rule out confounding, this design addresses the most direct threat to inference: that reforming states were already on different reentry trajectories before reform. PanelMatch estimates condition on pre-treatment reentry trends, addressing the concern that reform-adopting states may differ from non-adopters in their baseline trajectories. The post-reform divergence that we estimate therefore represents the additional effect of reform beyond the secular increase in reentry rates documented in Figure 1, exactly the quantity of theoretical interest. In Appendix E, we also report results of estimates using DiD methods developed by Gardner [2022] and by Borusyak et al. [2024]; results are consistent with those generated by PanelMatch.

Results

PanelMatch [Imai et al., 2023, Rauh et al., 2025] compares units with the same treatment history; that is, it compares a state that implements reform to unreformed states. Comparisons are made only between contemporaneous states, guarding against generic time-trends contributing to the results. For instance, a state which reforms in 1955 is compared to a set of unreformed states in 1955. To reduce pre-tend differences in reentry rates, we additionally refine comparison states by matching them on pre-treatment outcomes — that is, on their reentry rates before reform. Throughout, we implement PanelMatch using lags of four periods with homogeneous treatment status to generate matched sets, each of which includes five other states. Intuitively, this process can be considered similar to a series of synthetic control estimations; for each treated unit, a sample of five not-yet-treated units are used as counterfactuals.

In Figure 3, we show the estimated PanelMatch effect of civil service reform on reentry rates using Malhanobis matching. The pooled ATT results show an increase in reentry rates in reformed states of approximately 4.7 percentage points. The effect appears cumulative and grows during electoral cycles after reform. We present placebo treatment effects with 95% confidence intervals from pre-reform periods. The PanelMatch estimates show no apparent violation of the sequential ignorability assumption — placebo coefficients are statistically indistinguishable from zero — conditional on the matching strategy. This evidence supports the sequential ignorability assumption that is required by PanelMatch.

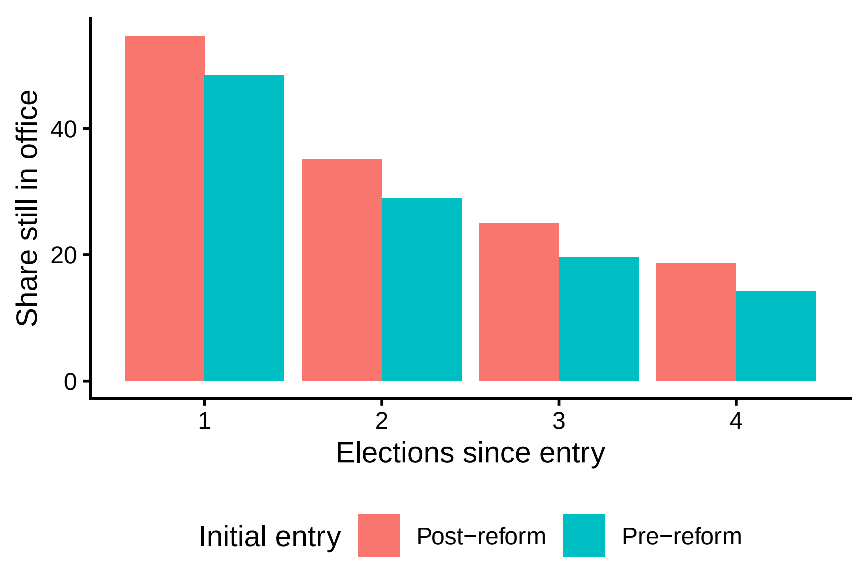
Figure 3: Mahalanobis PanelMatch estimates



Two possible mechanisms We now investigate two possible mechanisms to explain why reentry rates rose in states adopting reform at rates well above the underlying upward trend. To study political selection, we examine the length of political careers before and after reform. We map the rate of decay in retention of politicians across electoral cycles and compare politicians elected in the three elections before reform (the pre-reform generation) to those elected in the three cycles after reform (the post-reform generation). If a new type of electorally more ambitious and professional post-reform politician enters, we anticipate retention rates increase specifically among post-reform entrants. This identification strategy exploits the temporal discontinuity created by reform to separate cohort effects from period effects. If, on the other hand, extant politicians adapt to changing circumstances, we expect little difference in retention rates between pre- and post-reform politicians.

Figure 4 plots reentry rates of politicians in the four legislative sessions after entry for different cohorts. There is a consistent increase in the longevity of careers for post-reform entrants compared to pre-reform entrants, as longer legislative careers stretching across many elections become more common after adoption of civil service reform. We see an approximately six percentage point difference in shares of legislators serving in the next term between cohorts who enter pre-reform and those entering post-reform, and a continued difference between the two groups in the rates at which they achieve third, fourth, and fifth terms. Legislators who first enter after reform is adopted are more likely to have long careers and to serve as many as five cycles (ten years), for instance. Legislators who enter under a patronage regime have shorter careers overall.

Figure 4: Longevity in office of pre- and post-reform politicians



Bars represent shares of legislators first elected before civil service reform (blue)
and those first elected after reform was adopted (red).
The sets show shares reentering for a first, second, third, and fourth term.

This descriptive analysis is consistent with the interpretation that reform alters the pool of politicians: the new-guard experiences higher reentry rates because of unobserved underlying characteristics. We do not see evidence that supports the interpretation that the old guard changed its behavior to improve career longevity. Our estimates are that 12 percent of politicians who enter just before reform serve 10 years whereas 17 percent of politicians who enter just after reform serve equally long. This is a large increase.

To investigate performance effects, we collect, harmonize, and analyze annual data on per capita state-level expenditure disbursements. In the absence of more exact measures, such as the number or types of bills passed, infrastructure improvements, or sector-specific expenditure data, total state expenditures (including federal transfers) broadly proxy the delivery of public goods by states. We use PanelMatch to study whether civil service reform increases state-level per capita spending, and estimate the effect of reform on log per capita expenditures for four subsequent electoral cycles. We find precisely estimated null results; there is no evidence that reform ushers in more spending for service delivery to voters. Details are in Appendix G.

Interpretations and conclusions

Our results show that civil service reform during the 20th century across the US states contributed to the creation of a more professional class of state legislators, one that more often achieved legislative reentry and remained in office for more terms. Literature suggests that this group became more independent of political party organizations, in contrast to the patronage era, when party machines controlled candidate selection [Carson and Roberts, 2005]. Its members created long-lasting legislative careers instead of rotating between elected and appointed offices. Examining per capita state expenditures, we see no increase in this measure of post-reform legislative performance. There is no evidence that legislatures delivered more, only that post-reform legislators themselves more frequently reentered the next session.

The findings reported in Folke et al. [2011] shed light on how individuals interested in serving in elected office reoptimize strategically following reform adoption. Civil service reform makes it more difficult for pre-reform *legislators* to retain office, which is consistent with the increased likelihood of the *party* in power losing its majority. The increasing partisan competitiveness that civil service reform introduces weakens the ability of party organizations to control candidate selection and election campaigns, pushing the work of creating a personal following onto individual politicians. New post-reform politicians enter office, and their subsequent electoral successes demonstrate better capacities at individual vote-getting. Our data corroborates that post-reform entrants become more successful at reentry than their pre-reform counterparts; the theory of the personal vote displacing political parties is consistent with our finding of increased post-reform individual reentry rates. Civil service reform weakens political parties, and individuals interested in political careers respond by building longer careers in the legislature.

References

Ansolabehere, S., Ban, P., and Snyder, Jr., J. M. (2017). State legislative historical elections. Available at Harvard Dataverse: https://doi.org/10.7910/DVN/LEMNXZ.

Ash, E., Morelli, M., and Vannoni, M. (2022). Divided government, delegation, and civil service reform. *Political Science Research and Methods*, 10(1):82–96.

Borusyak, K., Jaravel, X., and Spiess, J. (2024). Revisiting Event-Study Designs: Robust and Efficient Estimation. *The Review of Economic Studies*, 91(6):3253–85.

Cain, B., Ferejohn, J., and Fiorina, M. (1987). *The Personal Vote: Constituency Service and Electoral Independence*. Harvard University Press, Cambridge, Massachusetts.

Carson, J. L. and Roberts, J. M. (2005). Strategic politicians and U.S. House elections, 1874–1914. *Journal of Politics*, 67(2):474–96.

Cox, G. W. and Katz, J. N. (1996). Why did the incumbency advantage in u.s. house elections grow? *American Journal of Political Science*, 40(2):478–97.

Engstrom, E. J. and Kernell, S. (2014). *Party Ballots, Reform, and the Transformation of America's Electoral System*. Cambridge University Press, Cambridge.

Folke, O., Hirano, S., and Snyder, J. M. (2011). Patronage and elections in the U.S. states. *American Political Science Review*, 15(3):567–85.

Gardner, J. (2022). Two-stage differences in differences. Available at arXiv: https://arxiv.org/abs/2207.05943.

Hoogenboom, A. (1961). *Outlawing the Spoils: A History of the Civil Service Reform Movement, 1865–1883*. Greenwood Press, Westport, CT.

Imai, K., Kim, I. S., and Wang, E. H. (2023). Matching methods for causal inference with time-series cross-sectional data. *American Journal of Political Science*, 67(3):587–605.

Klarner, C. (2018). State legislative election returns, 1967–2016. Available at Harvard Dataverse: https://doi.org/10.7910/DVN/3WZFK9.

Lewis, D. E. (2007). Testing Pendleton's premise: Do political appointees make worse bureacrats? *Journal of Politics*, 69(4):1073–88.

Rauh, A., Kim, I. S., and Imai, K. (2025). Panelmatch: Matching methods for causal inference with time-series cross-section data. arXiv preprint arXiv:2503.02073.

Roth, J., Sant'Anna, P. H., Bilinski, A., and Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics*, 235(2):2218–44.

Ruhil, A. V. and Camões, P. J. (2003). What lies beneath? The political roots of state merit systems. *Journal of Public Administration Research and Theory*, 13(1):27–42.

Ting, M. M., Snyder, James M., J., Hirano, S., and Folke, O. (2013). Elections and reform: The adoption of civil service systems in the U.S. states. *Journal of Theoretical Politics*, 25(3):363–87.

Ujhelyi, G. (2014a). Civil service reform. *Journal of Public Economics*, 118:15–25.

Ujhelyi, G. (2014b). Civil service rules and policy choices: Evidence from US state governments. *American Economic Journal: Economic Policy*, 6(2):338–80.

Xu, Y. (2023). Causal inference with time-series cross-sectional data: A reflection. Available at SSRN: https://ssrn.com/abstract=3979613.

Appendices to "Effects on Legislative Reentry of the Introduction of Merit Civil Service Appointments in US States"

Contents

A States included in the analysis

Four states (Mississippi, Maryland, Louisiana, and Alabama) hold elections every four years instead of biannually. We permanently exclude them from analysis rather than impute data every other year. We exclude Nebraska because it is unicameral. Alaska and Hawaii only joined the Union in 1959. Alaska adopted civil service reform in 1960 and Hawaii in 1955; reentry data for both states begins in 1958, when the first elections were held for statehouses into which incumbents could reenter. In effect, therefore, both states are always coded as reformed for the entire period for which we have reentry data. Our dataset thus covers 45 out of 50 states after 1959, when Alaska and Hawaii joined the Union, and 43 out of 48 before 1959. Although states use a variety of electoral systems [Carey et al., 2000, p. 675n5], we do not take these into account.

B Robustness to alternative coding of dates of civil service reform

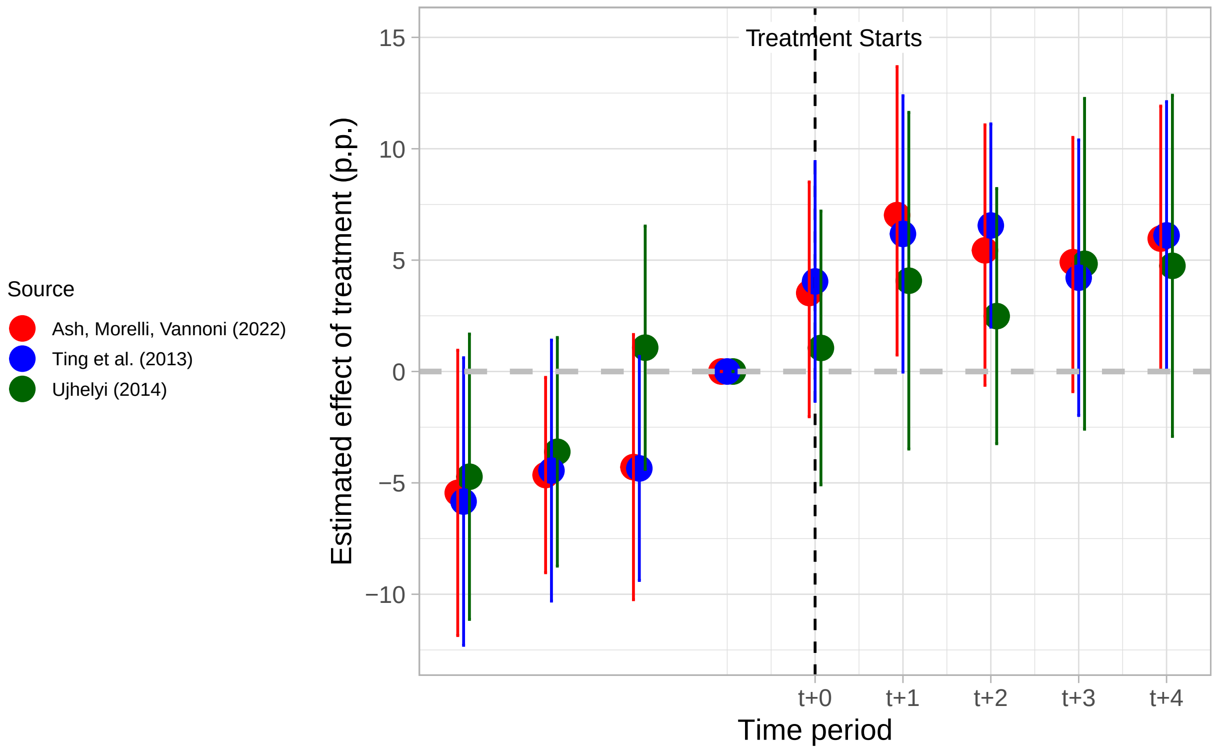
To check that results are not sensitive to the reform dates provided by Ash et al. [2022], we reestimate our main specification using dates from Ting et al. [2013] and from Ujhelyi [2014]. Results are depicted in Figure B-1. Effect sizes are largely consistent, although results using Ujhelyi [2014] are not statistically significant.

C Reentry of individual legislators

The main variable of theoretical interest in the reentry of legislators into the next legislative session. A legislator reenters if she was elected at time t in constituency i, conditional on her election in the same or a different constituency in the same state at time $t-1$. If a legislator skipped a legislative period, we do not code her as reentering.

We match individual legislators across sessions using their full last name and first name initials. Because we have only the first name initial for 9.9 percent of individuals, we match on first name initials for all legislators. This provides a consistent matching procedure but probably produces some false positives.

Figure B-1: Mahalanobis PanelMatch estimates: comparison of three sources for reform dates



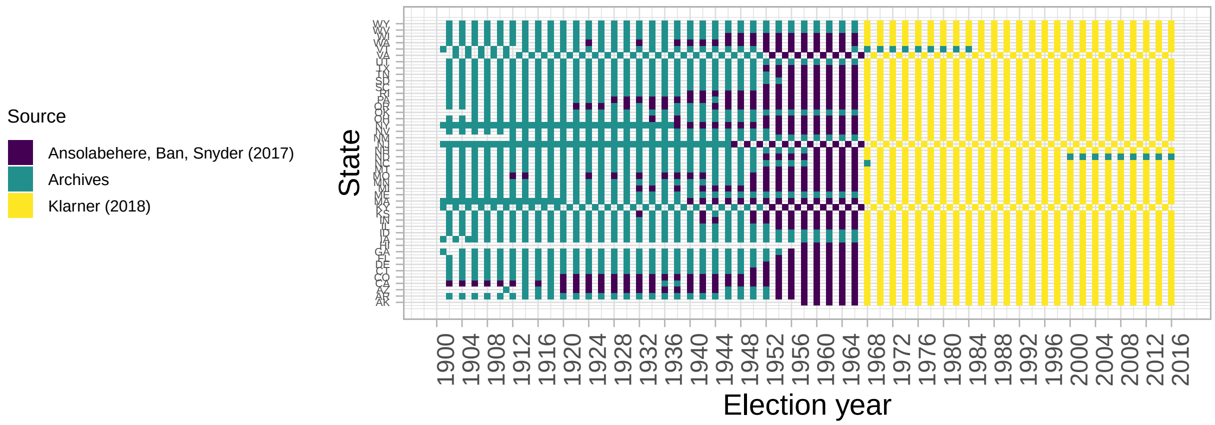
C.1 Sources of reentry data

We combine data sources to create a complete dataset containing names of legislators. We begin with a dataset that provides candidate-level state legislative election information for many states between 1890 and 1978 [Ansolabehere et al., 2017]. However, it has limited data prior to 1952 and after 1966. We combine the Ansolabehere et al. [2017] dataset with Klarner [2018], which provides candidate-level state legislative returns from 1968 to 2016. In the dataset used in the analysis reported in this paper, data taken from Ansolabehere et al. [2017] provide 15.6 percent of individual-level observations and those from Klarner [2018] 40.2 percent.[2]

After combining data from Ansolabehere et al. [2017] and Klarner [2018], we still have 46 percent missingness for the years 1900 to 2016. To complete the dataset, we collect missing information directly from state legislative offices and archives.

[2]Carsey et al. [2008] provides an alternate dataset that covers the period from 1967 to 2003 and contains a variable for incumbent status. We use Klarner [2018] instead of Carsey et al. [2008] because the former has longer coverage.

Figure C-2: Data source for each state-election year



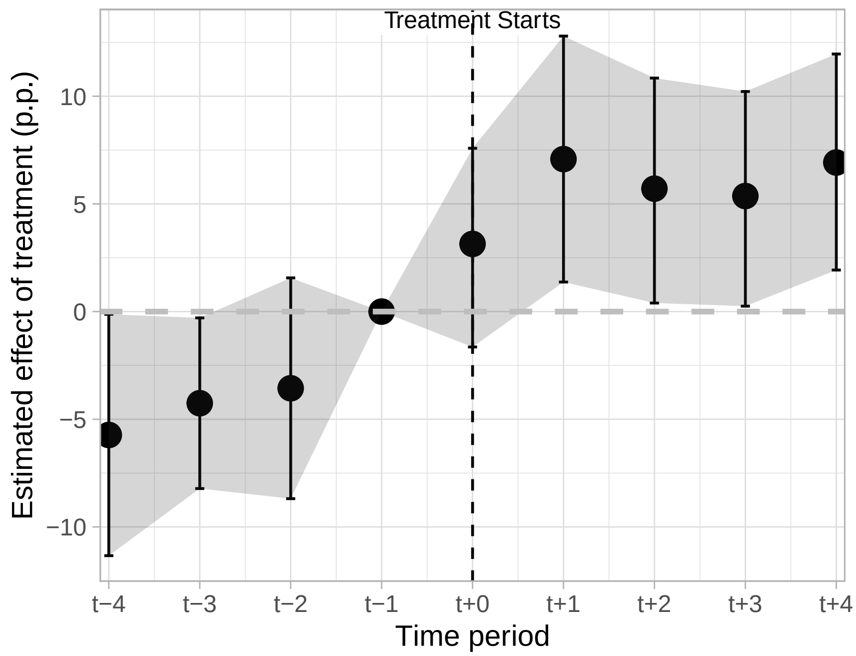
D Term limits

Term limits directly alter the pool of politicians and the maximum possible reentry rate. Term limits emerged in the 1990s and do not coincide temporally with civil service reform. However, reentry rates in the post-reform period could be influenced non-randomly by term limits. Of the 50 states, 17 adopted term limits at some point during the years we study. We assemble data on term limits from the National Conference of State Legislatures [National Conference of State Legislatures, 2024] and reestimate our main results excluding term-limited states. Results are reported in Figure D-3 and look similar in magnitude to the main results.

E Robustness checks of main results

We now detail our methodological approach and present robustness checks. We begin by explaining in more depth our choice of PanelMatch over alternative difference-in-differences estimators. The core challenge in our setting is that staggered adoption across 45 states over more than a century makes the required unconditional parallel trends implausible. States that reformed in the Progressive Era (e.g., Ohio in 1912) operated in fundamentally different political environments than states reforming during the Great Society period (e.g., New Mexico in 1961). Expecting these states to follow identical counterfactual reentry trajectories absent reform is not credible. We therefore require an estimator that conditions on pre-treatment dynamics rather than assuming unconditional parallel trends. PanelMatch accomplishes this by matching treated states to control

Figure D-3: PanelMatch estimates excluding states with term limits



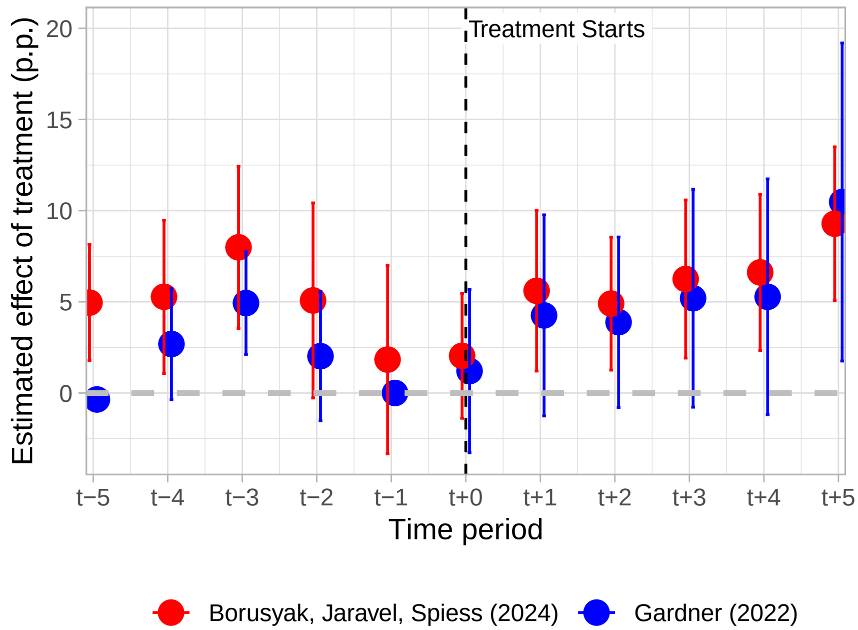
states with similar recent reentry histories, asking: among states on similar trajectories, what happens when one reforms? Crucially, comparisons are made only between contemporaneous states — a state reforming in 1955 is matched to unreformed states in 1955 — guarding against generic time trends contributing to the results. This is a more defensible causal identification strategy than one that asks whether all treated and control states would have evolved identically.

E.1 Alternative difference-in-differences estimators: imputation methods

We present results of both Gardner [2022] and Borusyak et al. [2024] imputations in Figure E-4. We see a perceptible increase in reentry rates after reforms are passed. Beginning in the second post-reform electoral cycle, we find an estimated increase of approximately five percentage points, which then stabilizes with a point estimate of around six percentage points across the following four cycles. Throughout the post-reform period, we see statistically significant effects.

In the pre-treatment period, we observe statistically significant estimates comparable to treatment effects when using these imputation methods, confirming our expectation that unconditional parallel trends does not hold in this setting. This is unsurprising: states that reformed earlier were likely on different baseline trajectories than later reformers due to differences in polit-

Figure E-4: Treatment effects over time using imputation



ical development, urbanization, and party system characteristics. Rather than viewing this as a threat to our analysis, we interpret it as confirmation that methods requiring unconditional parallel trends are inappropriate for our setting. This motivates our use of PanelMatch in the main text, which conditions on pre-treatment outcome histories rather than assuming parallel trends unconditionally.

The parallel trends assumption in Imai et al. [2023] is assumed to hold conditional on "the treatment, outcome, and covariate histories" (sequential ignorability), although it does not entirely rule out the possibility of unobserved confounders. The placebo treatment effects from pre-reform periods shown in Figure 3 show no apparent violations of the sequential ignorability assumption — the placebo coefficients are statistically indistinguishable from zero — conditional on the matching strategy. We implement the sensitivity analysis proposed by Rambachan and Roth [2023] using their HonestDiD package to quantify how much confounding would be required to explain away our results. Estimates remain statistically significant when allowing parallel trends violations of up to 20 percent of the maximum observed pre-treatment violation; beyond this threshold, significance is lost. This provides a concrete benchmark: unobserved confounding must generate differential trends at least 20 percent as large as the maximum pre-treatment difference to attribute our findings entirely to bias. We view it as plausible that our conditional matching strategy — which balances states on recent reentry histories — reduces confounding below this threshold, but we cannot rule out violations. This uncertainty is inherent to observational studies of historical institutional change.

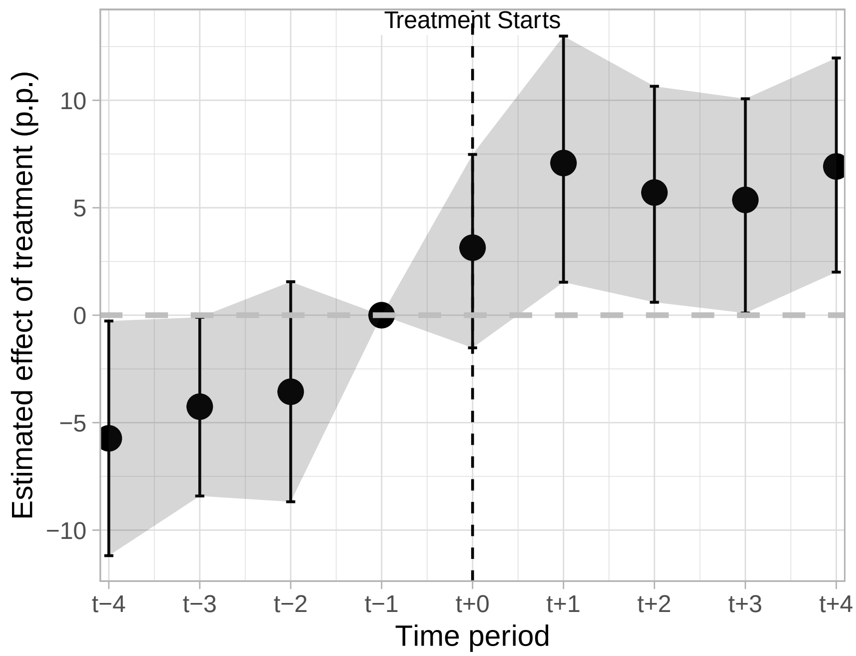
Point estimates are remarkably consistent across all three estimation strategies that we use. PanelMatch yields an estimated effect of approximately 4.7 percentage points; Gardner [2022] and Borusyak et al. [2024] likewise yield estimates in the 5–6 percentage point range. This convergence across methods with different identifying assumptions provides a form of robustness: if results were driven by a particular violation of a particular assumption, we would expect estimates to diverge across methods. Instead, we find that the magnitude of the treatment effect is stable regardless of the specific estimation approach used, with differences arising primarily in statistical precision and pre-trend behavior. This pattern is consistent with a genuine treatment effect of approximately five percentage points, contaminated by modest pre-existing differences between treated and control states that different methods handle differently.

E.2 PanelMatch robustness checks

In the body of the paper, we present matched estimates. In Figure E-5, we show results from unmatched estimates with 95% confidence bands around the coefficient point estimates. Results show that in the immediate aftermath of reform, reentry rates in reformed states increase relative to those in unreformed states. In the period immediately following adoption (*t+1*), the estimated effect is around six percentage points, although not statistically significant. In the second election after reform, the effect increases in size to around 10 percentage points and is statistically significant at the conventional 5% level. The effect stabilizes in the following electoral cycle and then slightly increases once again in the region of 12 percentage points four election cycles after reform.

The effects of reform are substantively significant over the elections following its adoption. The smallest effect of five percentage points in reform's immediate aftermath is larger than the effect we attain from the TWFE estimates (reported in Column B in Table F-2, with state-specific

Figure E-5: Unrefined PanelMatch coefficient estimates

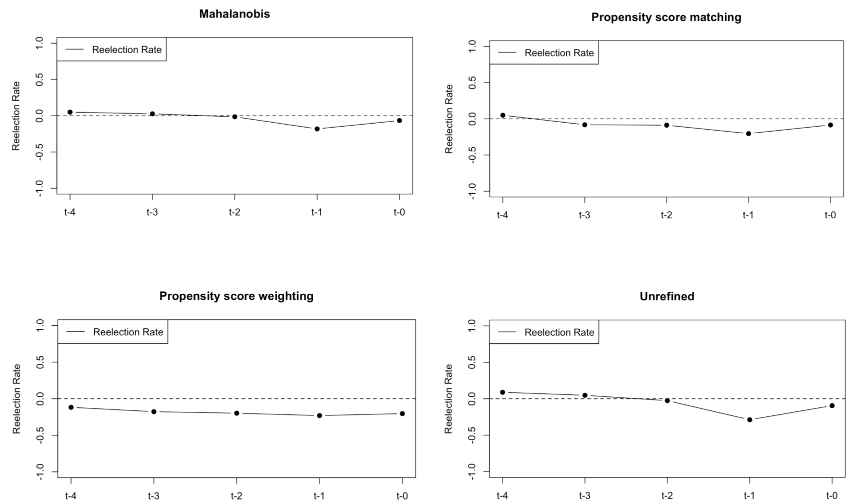


trends included) . In addition, the effect persists. We see a permanent shift compared to the pre-reform period, a shift that grows steadily as politicians in the legislature more often achieve reentry.

We study three different matching procedures used to make treated and control states more similar. Throughout, we utilize pre-reform reentry rates to create appropriate counterfactuals. In Figure E-6, we display how the three different matching algorithms impact differences between treated and control in the pre-treatment period. The plots show the standard deviations from the treated units; a standard deviation of zero indicates a perfect pre-reform fit. We aim to construct a matched set with low pre-treatment differences to create a comparison where pre-reform trends in reentry rates are less pronounced.

The final panel in Figure E-6 shows pre-trends in reentry rates between treated and control states using unrefined data. We see a good fit in the unrefined estimates, particularly in periods $t-4$ through $t-2$, and a small difference in reentry rates in the period $t-1$. The first panel in E-6 shows that the Mahalanobis algorithm slightly improves the fit in $t-1$. Conversely, results of two propensity-score matching methods that are displayed in the other two panels widen the differences compared to unrefined estimates. In the analyses reported in the body of the paper,

Figure E-6: Covariate balance



Note: Plots generated using the 'get_covariate_balance' function with main = 'Mahalanobis'. The latest version of PanelMatch (v3) no longer supports Mahalanobis for covariate balance and the package is not backward compatible. To reproduce the figures, use PanelMatch version 2.2.2.

we use the Mahalanobis-derived set of states since this provides more homogeneous reentry rate trajectories for treated and control units.

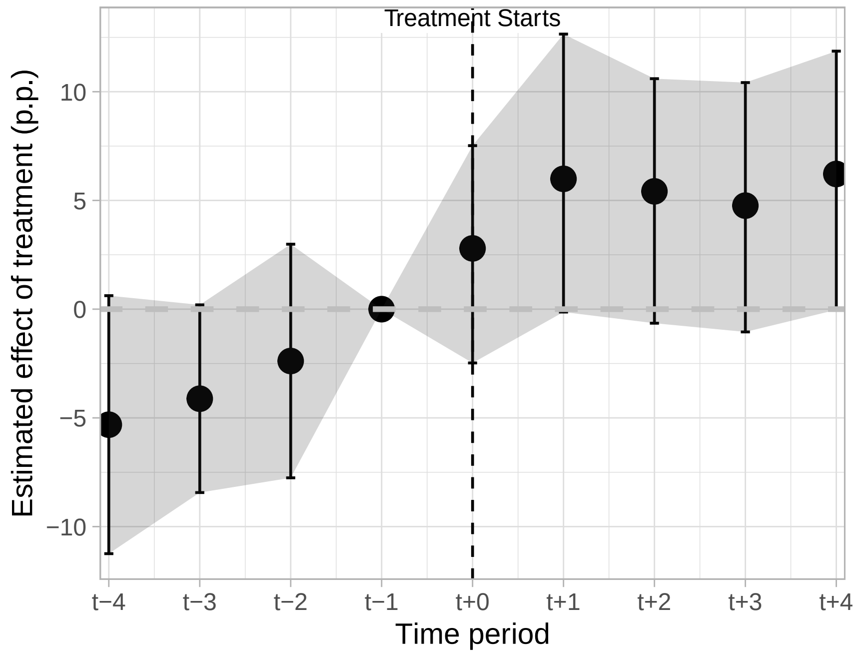
To probe the robustness of our results, we rerun PanelMatch estimation and exclude Texas, since it never adopted reform. (We have already reported results when we exclude states with term-limited politicians in Section D.) Removing the never-taker state (Texas) is a useful exercise since there could be systematically different dynamics operating in a state that never reformed.

In Figure E-7, we show the main results excluding the never-treated observation. Estimates are similar to our main results. In the election cycles following reform, we see increased reentry rates even when Texas is not included in the comparison group.

F Rerunning and reelection rates over time in selected states

We successfully assembled temporally incomplete rerunning data for nine states, two of which (New York and Pennsylvania) adopt reform legislation in unobserved years. The data is shown in Figure F-8. Coding rerunning is challenging because it requires full lists of candidates in order to know whether an incumbent appears on the ballot for the next legislative session. Rerunning could be consequential if decisions by incumbents not to run again depress reentry rates; this is

Figure E-7: PanelMatch estimates excluding Texas



a mechanical effect, since reentry is not possible without rerunning. Theoretically, we expect rerunning and reentry to track each other, since the decision to rerun should be endogenous to the likelihood of reelection. For the seven states that adopt civil service reform in years that coincide with available rerunning and reentry data, we observe gradual increases in both rerunning and reentry over time.

F.1 Comparing the rerunning-reelection gap before and after reform

A possible secondary test of selection theory is to examine whether the gap between rerunning and reelection rates is affected by civil service reform or whether rerunning and reelection display stable relationships. Evidence against a selection effect could consist of finding a significant difference in the gap before and after reform, where the difference between rates of rerunning and reelection was greater before than after reform. This would show that many incumbents tried to gain reelection but failed before reform, suggesting frustrated political ambition. Evidence consistent with a selection effect would find no significant difference between rerunning and reelection rates before and after reform. We undertake regression analyses for the states on which we have rerunning data. In the analysis, we omit Illinois because the pre-reform period is too short to meaningfully interpret, we omit New York because reform took place before 1900, when our data begin, and we omit Pennsylvania, which adopted reform after our rerunning data end. This leaves only six states. Table F-1 displays the estimates of how reform changes the gap between rerunning and reelection rates for the available states.

Table F-1: Effects of civil service reform on the gap between rerunning and reelection

	Dependent variable: Rerunning-reelection gap					
	CT	IA	IN	MI	OH	WA
	(1)	(2)	(3)	(4)	(5)	(6)
After reform	0.011	−0.009	0.013	−0.038	0.052	−0.037
	(0.013)	(0.029)	(0.033)	(0.028)	(0.068)	(0.031)
Constant	0.101^{***}	0.130^{***}	0.175^{***}	0.119^{***}	0.169^{**}	0.125^{***}
	(0.009)	(0.012)	(0.021)	(0.019)	(0.064)	(0.017)
Observations	32	37	30	37	18	33
R^2	0.021	0.003	0.006	0.048	0.036	0.043
Adjusted R^2	−0.012	−0.026	−0.030	0.021	−0.025	0.013
Note:				*p<0.1; **p<0.05; ***p<0.01		

We find null results for all states. The rates at which rerunners are elected are not affected by reform. Instead, the two rise together, with greater rerunning rising in parallel with reelection rates. This evidence is consistent with a selection theory, although the data obviously has limited numbers of observations.

We also examine whether the passage of civil service reform increases rerunning rates in the six states for which we have data and report the results in Table F-2. We observe no effect, in line with our result on the gap between rerunning and reelection remaining consistent pre- and post-reform. After reform is passed, rerunning and reelection rates both increase and a consistent share of rerunners are returned to office. Civil service reform does not make politicians more successful at reelection; instead, it appears that politicians hone their skills and as a by-product, reelection rates increase in the period we study.

Table F-2: Effects of civil service reform on rerunning and reelection in six states

Note: The six states are CT, IA, IN, MI, OH, and WA.

Using more extensive data, we thus confirm earlier research reporting a close correspondence between rerunning and reelection rates for a sample of states during a ten year period (1925–1935), which was interpreted as demonstrating that most incumbents did not wish to retain office [Hyneman, 1938].

G Government spending

To assess if reform affects legislative performance, we study whether reform appears causally linked to increases in state-level public spending. We collect and harmonize annual data on per capita expenditures of states across various years [United States Government, nd]. As Figure G-9 shows, there are null effects of reform on spending.

We also present auxiliary analyses regarding the relationship between state spending and reentry. We examine the correlation between per capita state expenditures and reentry rates to examine whether politicians are rewarded at the ballot box for prior government spending. Results are depicted in Figure G-10.

Figure G-10: Per capita state expenditures and reentry rates

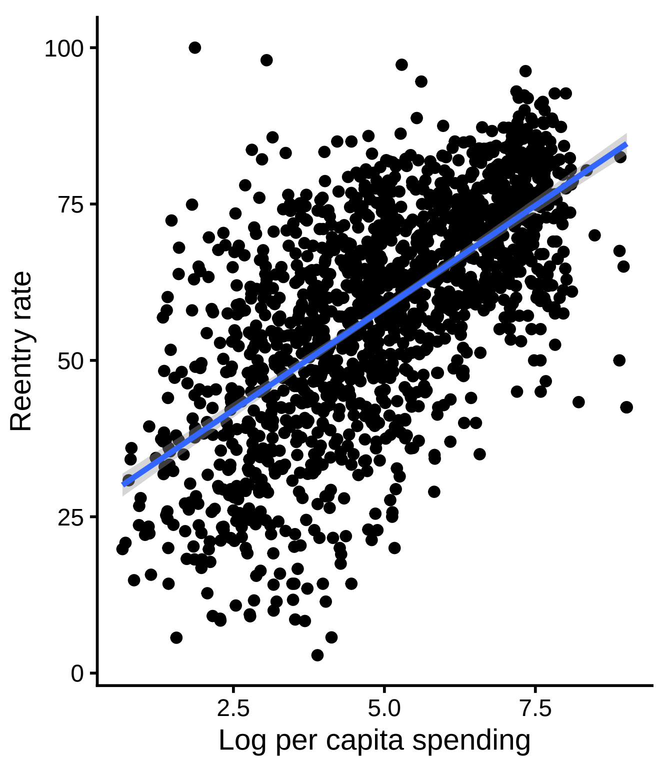


Figure G-10 shows that reentry rates and state government spending are closely correlated. We suspect both to be features of the larger transformation of state politics that took place during the 20th century, as states gained more fiscal resources. Years when per capita spending is greater also have higher reentry rates to their state legislatures.

We statistically estimate the relationship between lagged spending and reentry rates, testing the hypothesis that greater spending per capita heightens reentry rates in subsequent electoral cycles. Results appear in Table G-3 and show that lagged spending has a significant and substantively large positive effect on reentry rates. The effect is largest further back in time, around five electoral cycles prior (10 years). Nevertheless, there is also a slightly smaller effect of spending four years before the current election (see column 2). In general, per capita spending appears to impact the ability of politicians to gain reentry, with the stock of historic spending working in favor of politicians gaining office again. Thus, legislative political careers become more stable with the growth in government.

Table G-3: Effect of historic (lagged) spending on reentry rates

	Dependent variable: Reentry rate				
	1 cycle lag	2 cycle lag	3 cycle lag	4 cycle lag	5 cycle lag
	(1)	(2)	(3)	(4)	(5)
Log expenditures per capita	0.470	3.277***	4.311***	4.986***	6.160***
	(1.270)	(1.230)	(1.234)	(1.257)	(1.261)
Cycle FEs	Y	Y	Y	Y	Y
State FEs	Y	Y	Y	Y	Y
Observations	1,537	1,537	1,537	1,537	1,537
Note:					*p<0.1; **p<0.05; ***p<0.01

H Deviations from the pre-analysis plan (PAP)

- **Deviation 1**: In our pre-analysis plan, we specified that "our analysis covers the period from 1946 to 2016," based on the justification that "much greater election data availability exists for the entire matrix of reformed and unreformed states in the postwar period." We also noted that "over 60 percent of reform legislation was adopted in the years following the end of World War II," which implies that focusing on the post-WWII period would exclude roughly 40 percent of the already-treated states. During data collection, we obtained additional funding to digitize pre-1945 House journals, enabling us to expand sample back to 1900. Assuming nearly half the sample (out of 50 states) as already treated by 1946 weakens the power of our main analysis to detect the effects of civil service reforms. In addition, by starting in 1946, we miss the implications of the 1939 Congressional amendment of the Social Security Act to encourage merit civil service appointments at the state level. In Figure H-11, we provide a side-by-side comparison of the original 1946–2016 sample with the extended 1900–2016 sample. Results are largely consistent across both the full (post-1900) and truncated (post-1946) samples.

- **Deviation 2**: The PAP proposed that we "employ both a staggered difference-in-differences model and an event study approach," estimating the main equation of interest "using OLS," specifically referring to the standard two-way fixed effects approach. While we proposed using the Goodman-Bacon decomposition approach for sensitivity analysis, we did not pre-register alternative estimation methods at that time. Since June 2020, when we filed the PAP, the literature on staggered difference-in-differences estimation, has evolved rapidly. At the time, we were uncertain about which method would be most suitable for our analysis. Our main findings rely on PanelMatch estimators developed by Imai et al. [2023], as well methods proposed by Gardner [2022] and Borusyak et al. [2024], all of which became publicly available only after we wrote the PAP. Nonetheless, we beleive it makes most sense to use the best most recent methods rather than adhere mechanically to TWFE.

- **Deviation 3**: Our pre-analysis plan specified that the event study would examine 10 election cycles before and after reform (including a sensitivity check where our period varies from three cycles before the reform to ten cycles after). However, in the data we collected, there are not enough observations to identify the coefficients for 10 lags and leads when estimating the main regression specification (Equation 1). Therefore, our main event-study model specifications test five periods before and after reform. In addition, our pre-analysis plan specified a prediction (consistent with a selection mechanism) that effects on reentry rates would weakly increase over time. Our results indicate that the estimated coefficients remain stable at around five percent and show no increase over time following treatment, findings that are only partially consistent with our pre-registered prediction.

- **Deviation 4**: In our PAP, we wrote: "To check that our results are not sensitive to the inclusion of any particular state, we will reestimate our specification dropping each state one at a time. This will also allow us to verify that results hold even without the inclusion of open-primary Louisiana and unicameral Nebraska."

  In the paper, we do not drop states one at a time but instead permanently exclude four states (Mississippi, Maryland, Louisiana, and Alabama) that hold elections every four years. We also exclude Nebraska because it is unicameral.

  In addition to dropping these states, we recode off-year two-year election cycles to align with presidential election years. For instance, states holding elections in odd-numbered years (e.g., Virginia in 1901, 1903, etc.) are recoded so that the 1901 election is matched with the 1900 cycle and the 1903 election with the 1902 cycle. This ensures all state elections align temporally for the panel analysis.

- **Deviation 5**: The timing of reform may not be as-if random. To check that our results are not driven by legislators trying to pass civil service reform to win the upcoming election, we initially thought (following Folke et al. [2011]) that we might reestimate our preferred specification dropping the electoral cycle just before the reform and the electoral cycle immediately after reform (conditional on data availability). We decided against this approach since we now use the PanelMatch estimator, which corrects for pre-trends. We believe this is a superior to dropping data and relying on the problematic TWFE estimation strategy because we retain all available data while correcting for anticipation effects via matching.

- **Deviation 6**: We initially aimed to provide qualitative evidence on the determinants of civil service reforms and remove states where adoption of the reform is most likely to be correlated with our dependent variable. After consideration of the data collection process, we judged this unfeasible. Conducting a qualitative evaluation of primary sources, even from a limited number of states, would be tantamount to a separate research project, one which evaluates the processes and dynamics behind the passage of reforms. In this paper, we instead focus on the effects of the meritocratic reforms (consequences) and leave the analysis of their introduction (causes) to future research.

- The PAP stated: "To check that the South does not exhibit different trends, we code the states from the deep South with a dummy δ_S and rerun our analysis excluding Southern states from our sample." We show that the result holds without the southern states in Table H-4.

Table H-4: Two-way fixed-effects estimates of reentry rates excluding southern states

	Dependent variable: Reentry rate	
	(1)	(2)
Post-reform rate	7.399***	4.134**
	(2.400)	(1.818)
Election cycle FE	yes	yes
State FE	yes	yes
State-specific trends	no	yes
N	1,957	1,957
R^2	0.768	0.822

Notes: ***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.
Standard errors, clustered by state, in parentheses.
Excludes AL, AR, FL, GA, KY, LA, MS, NC, SC, TN, TX, VA, WV.

- In the PAP, we stated: "To check whether our results remain insensitive to possible differences in reelection rate data assembled by Ansolabehere et al. [2017], by Klarner [2018], and by ourselves, we will add fixed effects for each data source." In Table H-5, we add data-source fixed effects (using a TWFE estimation). The main results remain unchanged.

Table H-5: Two-way fixed-effects estimates

	Dependent variable: Reentry rate		
	(1)	(2)	(3)
post_reform	7.751***	3.650**	7.783***
	(2.260)	(1.542)	(2.319)
Election cycle FE	yes	yes	yes
State FE	yes	yes	yes
State-specific trends	no	yes	no
Data source FE	no	no	yes
N	2,535	2,535	2,535
R^2	0.775	0.826	0.775

Notes: ***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.
Standard errors, clustered by state, in parentheses.

I Data availability and codebook

The complete dataset used in this analysis, along with a detailed codebook, is publicly available on the Harvard Dataverse [redacted]. The codebook provides variable definitions, coding procedures, and data sources for all variables used in the analysis.

References

Ansolabehere, S., Ban, P., and Snyder, Jr., J. M. (2017). State legislative historical elections. Available at Harvard Dataverse: https://doi.org/10.7910/DVN/LEMNXZ.

Ash, E., Morelli, M., and Vannoni, M. (2022). Divided government, delegation, and civil service reform. *Political Science Research and Methods*, 10(1):82–96.

Borusyak, K., Jaravel, X., and Spiess, J. (2024). Revisiting Event-Study Designs: Robust and Efficient Estimation. *The Review of Economic Studies*, 91(6):3253–85.

Carey, J. M., Niemi, R. G., and Powell, L. W. (2000). Incumbency and the probability of reelection in state legislative elections. *Journal of Politics*, 62(3):671–700.

Carsey, T. M., Niemi, R. G., Berry, W. D., Powell, L. W., and Snyder, Jr., J. M. (2008). State legislative elections, 1967–2003: Announcing the completion of a cleaned and updated dataset.

State Politics & Policy Quarterly, 8(4):430–43.

Folke, O., Hirano, S., and Snyder, J. M. (2011). Patronage and elections in the U.S. states. *American Political Science Review*, 15(3):567–85.

Gardner, J. (2022). Two-stage differences in differences. Available at arXiv: https://arxiv.org/abs/2207.05943.

Hyneman, C. S. (1938). Tenure and turnover of legislative personnel. *Annals of the American Academy of Political and Social Science*, 195(2):21–31.

Imai, K., Kim, I. S., and Wang, E. H. (2023). Matching methods for causal inference with time-series cross-sectional data. *American Journal of Political Science*, 67(3):587–605.

Klarner, C. (2018). State legislative election returns, 1967–2016. Available at Harvard Dataverse: https://doi.org/10.7910/DVN/3WZFK9.

National Conference of State Legislatures (2024). The term-limited states. Available at https://www.ncsl.org/about-state-legislatures/the-term-limited-states. Accessed: 2024-12-03.

Rambachan, A. and Roth, J. (2023). A more credible approach to parallel trends. *Review of Economic Studies*, 90(5):2555–91.

Ting, M. M., Snyder, James M., J., Hirano, S., and Folke, O. (2013). Elections and reform: The adoption of civil service systems in the U.S. states. *Journal of Theoretical Politics*, 25(3):363–87.

Ujhelyi, G. (2014). Civil service reform. *Journal of Public Economics*, 118:15–25.

United States Government (n.d.). United States Census Bureau: Historical data. Various years. Available at: https://www.census.gov/programs-surveys/gov-finances/data/historical-data.html. Accessed: 2024-09-12.

Figure F-8: Proportions of incumbents rerunning and reelected in selected states, various years, 20th century

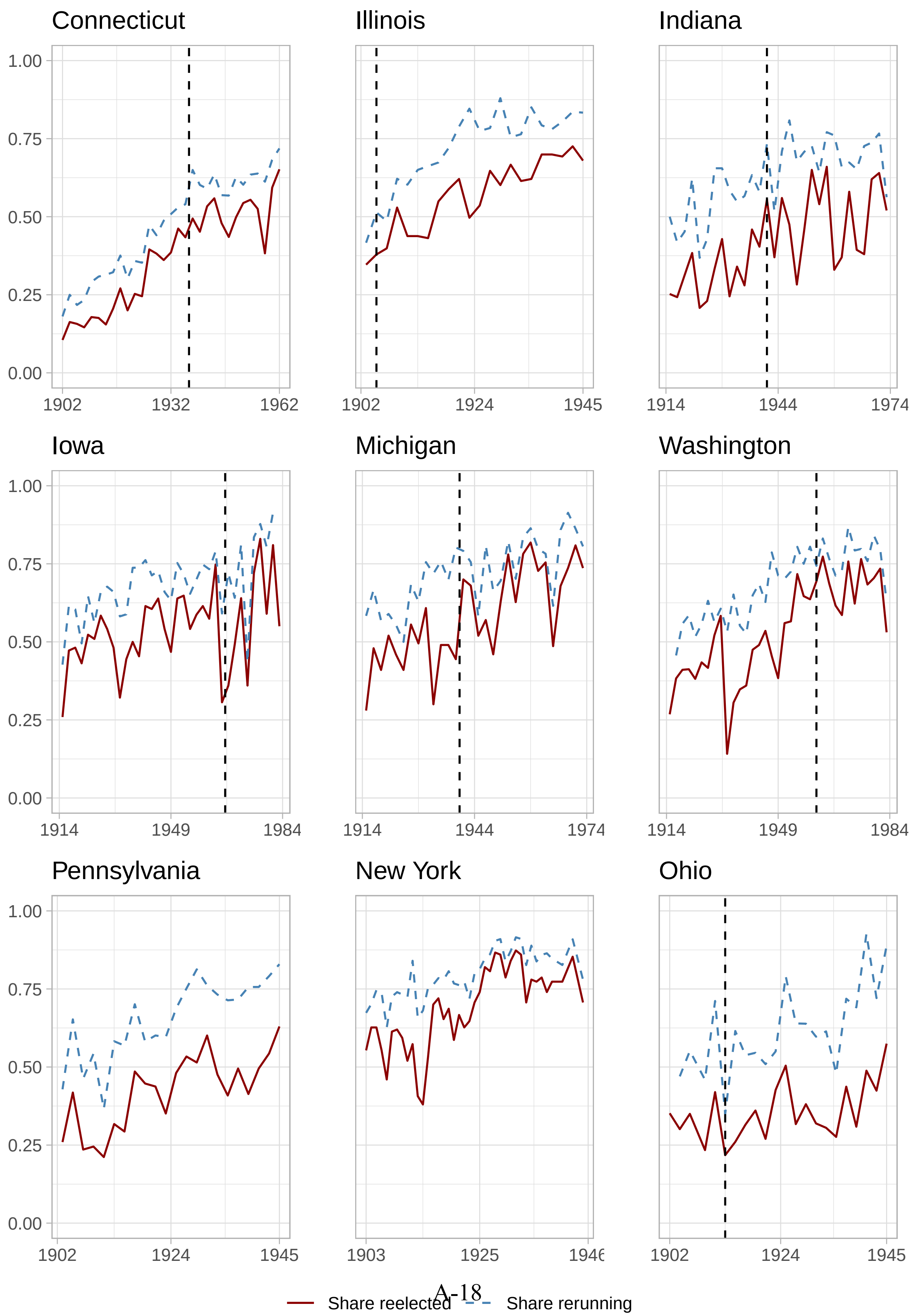


Figure G-9: Effects of civil service reform on per capita state expenditures

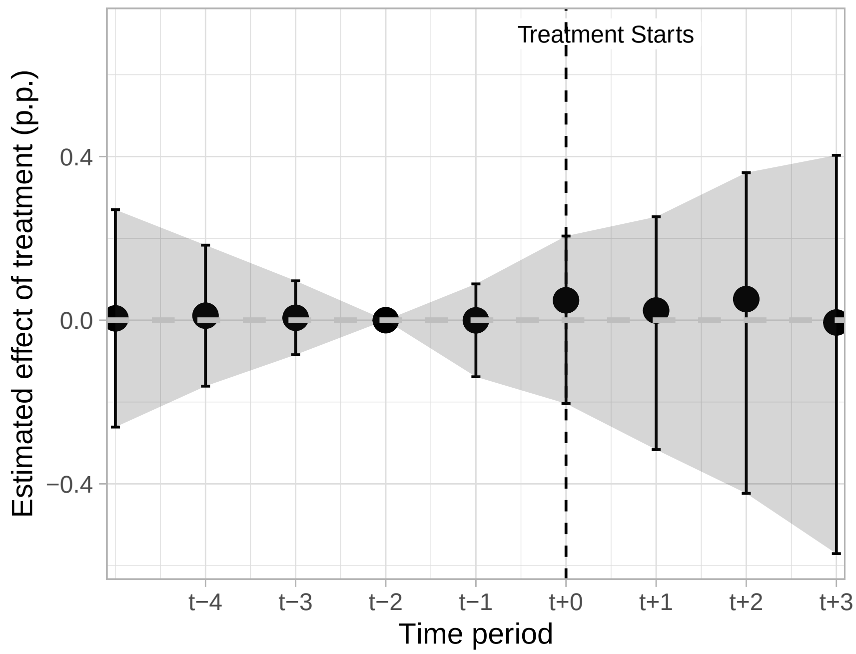


Figure H-11: Mahalanobis PanelMatch estimates for the full dataset compared with post-1946 data