

**The Control of Politicians in Normal Times
and Times of Crisis:
Wealth Accumulation by U.S. Congressmen, 1850-1880**

Supplemental Materials

(Online Appendix)

Appendix A:

Robustness Checks and Alternative Mechanisms

Robustness Checks

Tables A.1 to A.4 present results for additional robustness checks and placebo tests.

Alternative Mechanisms

Tables A.5 to A.6 present results regarding other possible interpretations of the results reported in Tables 4-6.

Majority Party Status. It seems plausible that members of the majority party, who control the main levers of power, could enrich themselves more easily than those in the minority. We therefore investigated the interaction between winning and majority party status in the RDD samples. We report these results in Appendix Table A.5. For the 1850s, 1870s, and 1860s overall, there are no statistically significant differences between majority party winners and minority party winners. Regarding the Civil War period the evidence is mixed. For the dependent variable *Ending Wealth*, the estimates suggest that majority party winners accumulated substantially more wealth than minority party winners, and the difference is statistically significant. For *Ending Servants*, however, the difference between majority and minority party winners is small, negative, and statistically insignificant. Issues of partisanship and majority party control clearly deserve further investigation, but these are beyond the scope of our paper.¹

Wealth Accumulation *After* Serving in Congress. The results in Table 4 of the paper provide no systematic evidence of abnormal wealth accumulation *while serving in office*, for those who served during the 1850s or 1870s. One possibility is that congressmen are able to benefit from the connections and networks established while in office in order to accumulate wealth *after* they leave Congress. Congressmen meet many rich and powerful people while serving in Washington, and some of these people may become clients after the congressmen leave office (this seems especially likely for lawyers). Alternatively, congressmen may exploit their connections with those who continue in office – e.g., by receiving preferential treatment for contracts or “insider” information – after they leave office themselves and become less scrutinized by voters and the media. For the British case, Eggers and Hainmueller (2008) find that Conservative MPs profited from office largely through lucrative outside employment they often later acquired as a result of their political positions.

To explore this possibility, we examined whether those who served in Congress accumulated more wealth than those who ran but lost in the decade *after* they were out of office. For example, we consider wealth accumulation between 1860 and 1870 for those who served

¹Simpler OLS regressions also reveal no significant partisan differences. We regressed each of the three dependent variables on a dummy variable for majority party status, restricting the sample to those who won during the Civil War (we also include controls for lagged wealth, age, occupation and state fixed-effects). The point estimates are positive, but statistically insignificant.

during the 1850s (31st-36th Congresses) but did not serve during the 1860s. We report this in Appendix Table A.6. The estimates provide no robust evidence of abnormal returns in the decade after serving in Congress. For all three dependent variables, the points estimates are generally small, sometimes negative, and never statistically significant.

Attraction of Venal Candidates. Another possibility is that the political environment during this period attracted more venal candidates, who anticipated that federal politicians would have greater opportunities for war profiteering than others. Greater wealth accumulation by congressmen during this period may partly reflect a change in the *type* of individuals who ran for Congress. However, this hypothesis receives only limited support in the data. In particular, there is little change in the occupational background of the overall pool of candidates – winners and losers – who ran for Congress during the Civil War years. For the Civil War congresses, 62% of the candidates were lawyers, 15% were farmers, and 20% were merchants, manufacturers or bankers. These figures are not very different for candidates who ran during the 1850s or during the non-war years in the 1860s – 54% of candidates who ran during the 1850s were lawyers, 16% were farmers, and 19% were merchants, manufacturers or bankers. The corresponding figures for the non-war years during the 1860s were 62%, 13%, and 20% respectively. Of course, it is possible that candidates differed along dimensions that we cannot observe, so this hypothesis should be explored in more detail by future research.

Rents during Normal Times. A final potential interpretation of the results in Table 4 is that the rate of rent extraction during 1850s, 1870s, and the second half of the 1860s was similar to that during the Civil War, but the amounts were difficult to detect because the federal budget was so much smaller. In other words, political institutions may have been equally effective during the Civil War in monitoring and keeping the levels of rent extraction at the same proportion as during the non-war years, but we can only detect systematic rents during the Civil War years due to the much larger size of federal spending. This interpretation, while plausible, is unlikely to explain our results. First, the estimated coefficients in Table 4 for the 1850s, 1870s, and non-war years of the 1860s are all *negative*. Second, our analysis of the LaCrosse and Milwaukee railroad scandal – reported in Appendix Table B.4 – reveals that we can detect bribes as small as \$5,000 using census wealth data. As noted above, the point estimates in Table 4 imply that total rents during the Civil War years represented about .2% of federal spending. If congressmen had collected rents at the same rate during the non-war years of the 1860s, then rents would have been almost \$6,000 per congressman. In any case, there is no special reason why we may care more about rents as a fraction of total federal spending as opposed to their absolute value. Even if the fraction of federal spending appropriated by congressmen remained constant during the Civil War years, the dollar amount of rents was substantially larger during this period.

Table A.1: Robustness Checks for RDD Estimates in Table 4									
	2% Bandwidth			3% Bandwidth			5% Bandwidth		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Ending Wealth	Ending Log Wealth	Ending Servants	Ending Wealth	Ending Log Wealth	Ending Servants	Ending Wealth	Ending Log Wealth	Ending Servants
<i>Panel A: 1850-1860</i>									
Winner 1850s	-1804 (5907)	-0.065 (0.160)	-0.212 (0.177)	-1153 (4039)	-0.014 (0.136)	-0.141 (0.140)	-10636 (3916)	-0.332 (0.207)	-0.201 (0.197)
Obs.	176	155	181	248	220	249	372	328	373
<i>Panel B: 1860-1870</i>									
Winner 1860s	2618 (6208)	0.189 (0.182)	0.354 (0.174)	3026 (3166)	0.206 (0.128)	0.333 (0.130)	849 (5880)	0.266 (0.201)	0.488 (0.210)
Obs.	164	160	186	242	233	286	365	353	420
<i>Panel C: 1860-1870, Civil War vs. Non-War</i>									
Winner Civil War	16515 (4250)	0.375 (0.199)	0.597 (0.189)	18387 (3474)	0.385 (0.141)	0.582 (0.143)	12462 (4689)	0.267 (0.147)	0.549 (0.151)
Winner Non-War	-3344 (3815)	-0.028 (0.193)	-0.239 (0.186)	-1140 (3242)	0.018 (0.136)	-0.169 (0.139)	2314 (4605)	0.247 (0.145)	-0.116 (0.152)
p-value of F-test	0.002	0.188	0.005	0.000	0.085	0.001	0.076	0.910	0.001
Obs.	164	160	186	242	233	286	365	353	420
<i>Panel D: 1870-1880</i>									
Winner 1870s	–	–	0.117 (0.178)			0.043 (0.138)			-0.021 (0.215)
Obs.			148			236			343

Table A.2: Robustness Checks on RDD Estimates in Table 5									
Contracting State vs. Other States									
	2% Bandwidth			3% Bandwidth			5% Bandwidth		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Ending Wealth	Ending Log Wealth	Ending Servants	Ending Wealth	Ending Log Wealth	Ending Servants	Ending Wealth	Ending Log Wealth	Ending Servants
Civil War, Contract State	34152 (5285)	0.455 (0.232)	0.756 (0.213)	29975 (4404)	0.429 (0.167)	0.636 (0.168)	25206 (5253)	0.316 (0.168)	0.614 (0.172)
Civil War, Other State	-1924 (7172)	0.192 (0.363)	0.100 (0.367)	6644 (6416)	0.281 (0.252)	0.443 (0.264)	4308 (6688)	0.164 (0.224)	0.394 (0.238)
Non-War, Contract State	-197 (4721)	0.017 (0.214)	-0.200 (0.201)	-1386 (4127)	0.017 (0.158)	-0.196 (0.161)	2401 (5035)	0.225 (0.163)	-0.173 (0.169)
Non-War, Other State	-7115 (8019)	-0.229 (0.476)	-0.470 (0.462)	-4606 (6292)	0.021 (0.275)	-0.089 (0.271)	222 (6758)	0.298 (0.233)	0.025 (0.241)
p-value of F-test 1	0.000	0.533	0.116	0.002	0.617	0.532	0.005	0.543	0.404
p-value of F-test 2	0.000	0.207	0.004	0.000	0.098	0.001	0.001	0.679	0.001
p-value of F-test 3	0.000	0.195	0.017	0.000	0.207	0.025	0.001	0.944	0.029
Obs.	164	160	186	242	233	286	365	353	420

Median regression estimates for *Ending Wealth* dependent variable (columns 1,4,7). OLS estimates for *Ending Log Wealth* and *Ending Servants* dependent variables (columns 2,3,5,6,8,9). The p-values are for F-tests of the hypothesis that the β for Civil War, Contract State Winners is equal to: (1) β for Civil War, Other State Winners, (2) β for Non-War, Contract State Winners, and (3) β for Non-War, Other State Winners. Local-linear control function included in regressions using the 5% Bandwidth.

Table A.3: Military Committees vs. Other Committees Full Sample (Not Restricted to 3% Window)			
	(1)	(2)	(3)
	Ending Wealth	Ending Log Wealth	Ending Servants
<i>Panel A: Difference-in-Difference Estimates, 1860-1870</i>			
Civil War, Military Comm	21311 (4535)	0.359 (0.153)	0.554 (0.153)
Civil War, Other Comm	7131 (2592)	0.136 (0.089)	0.428 (0.092)
Non-War, Military Comm	3603 (3446)	0.183 (0.119)	-0.272 (0.121)
Non-War, Other Comm	1082 (2662)	0.076 (0.092)	-0.020 (0.092)
Diff-in-Diff Estimate	11660	0.115	0.379
p-value	0.071	0.599	0.087
Obs.	775	741	905
<i>Panel B: Placebo Difference-in-Differences, 1850-1860</i>			
Civil War, Military Comm	-9062 (4836)	-0.044 (0.227)	0.164 (0.172)
Civil War, Other Comm	-483 (2889)	0.065 (0.136)	0.038 (0.104)
Non-War, Military Comm	8851 (4111)	0.126 (0.191)	0.148 (0.148)
Non-War, Other Comm	151 (3247)	0.035 (0.155)	0.085 (0.117)
Diff-in-Diff Estimate	-17279	-0.200	0.062
p-value	0.020	0.562	0.813
Obs.	554	419	658

Independent variables defined as follows: Civil War, Military Comm = 1 for those who served during the Civil War, and served on committees with military-related jurisdictions. Difference-in-Difference = (Civil War, Military Comm – Civil War, Other Comm) - (Non-War, Military Comm – Non-War, Other Comm). Median regression estimates for *Ending Wealth* dependent variable (column 1). OLS estimates for *Ending Log Wealth* and *Ending Servants* dependent variables (columns 2-3). The p-values are for F-tests of the hypothesis that the difference-in-difference estimates are equal to zero.

Table A.4: Contracting States and Military Committees in Other Decades			
	(1)	(2)	(3)
	Ending Wealth	Ending Log Wealth	Ending Servants
<i>Panel A: 1850-1860, Contracting States vs. Other States</i>			
Contract State	-9327 (5752)	-0.428 (0.271)	-0.401 (0.268)
Other State	3606 (7649)	-0.311 (0.357)	-0.339 (0.362)
p-value of F-test	0.051	0.704	0.840
Obs.	248	220	249
<i>Panel B: 1870-1880, Contracting States vs. Other States</i>			
Contract State	—	—	-0.038 (0.297)
Other State	—	—	0.013 (0.326)
p-value of F-test			0.856
Obs.			236
<i>Panel C: 1850-1860, Military Committees vs. Other Committees</i>			
Military Comm	6490 (2252)	0.219 (0.113)	0.009 (0.117)
Other Comm	3578 (1619)	0.173 (0.081)	0.020 (0.086)
p-value of F-test	0.215	0.696	0.925
Obs.	756	652	794
<i>Panel D: 1870-1880, Military Committees vs. Other Committees</i>			
Military Comm	—	—	0.068 (0.124)
Other Comm	—	—	0.024 (0.091)
p-value of F-test			0.726
Obs.			652

See Table 5 and 6 for details regarding variables and estimation methods. The p-values are for F-tests of the hypothesis that the estimated coefficients on the two reported variables are equal.

Table A.5: RDD Estimates of Wealth vs. Serving in Congress, by Majority Party Status			
	(1)	(2)	(3)
	Ending Wealth	Ending Log Wealth	Ending Servants
<i>Panel A: 1850-1860</i>			
Winner 1850s, Majority	-9116 (10618)	-0.417 (0.273)	-0.402 (0.275)
Winner 1850s, Minority	-10143 (11412)	-0.378 (0.291)	-0.372 (0.287)
p-value of F-test	0.896	0.849	0.884
Obs.	248	220	249
<i>Panel B: 1860-1870</i>			
Winner 1860s, Majority	3528 (7967)	0.124 (0.252)	0.504 (0.263)
Winner 1860s, Minority	-2778 (8154)	-0.096 (0.265)	0.560 (0.272)
p-value of F-test	0.270	0.232	0.769
Obs.	242	233	286
<i>Panel C: 1870-1880</i>			
Winner 1870s, Majority	—	—	0.071 (0.289)
Winner 1870s, Minority	—	—	-0.129 (0.297)
p-value of F-test			0.319
Obs.			236

Table A.5: RDD Estimates of Wealth vs. Serving in Congress, by Party Status (continued)			
	(1)	(2)	(3)
	Ending Wealth	Ending Log Wealth	Ending Servants
<i>Panel D: 1860-1870, Civil War vs. Non-War</i>			
Winner Civil War, Majority	29882 (6886)	0.392 (0.202)	0.562 (0.200)
Winner Civil War, Minority	11004 (8348)	0.197 (0.247)	0.610 (0.250)
Winner Non-War, Majority	-2757 (6850)	-0.003 (0.202)	-0.179 (0.207)
Winner Non-War, Minority	2924 (8179)	-0.209 (0.246)	-0.142 (0.249)
p-value of F-test 1	0.035	0.466	0.858
p-value of F-test 2	0.507	0.433	0.891
Obs.	242	233	286

Median regression estimates for *Ending Wealth* dependent variable (column 1). OLS estimates for *Ending Log Wealth* and *Ending Servants* dependent variables (columns 2-3). In Panels A, B and C, the p-values are for F-tests of $H_0: \beta \text{ for Winner, Majority} = \beta \text{ for Winner, Minority}$. In Panel D, the p-values for F-test 1 are for $H_0: \beta \text{ for Winner Civil War, Majority} = \beta \text{ for Winner Civil War, Minority}$, and the p-values for F-test 2 are for $H_0: \beta \text{ for Winner Non-War, Majority} = \beta \text{ for Winner Non-War, Minority}$.

Table A.6: RDD Estimates of Effect of Serving in Congress on Wealth After Leaving Congress			
	(1)	(2)	(3)
	Ending Wealth	Ending Log Wealth	Ending Servants
<i>Panel A: 1860-1870</i>			
Winner 1850s	-5259 (11481)	0.025 (0.330)	0.354 (0.323)
Obs.	159	155	167
<i>Panel B: 1870-1880</i>			
Winner 1860s	—	—	0.004 (0.382)
Obs.			144
<i>Panel C: 1870-1880</i>			
Winner Civwar	—	—	0.143 (0.286)
Winner Non-War	—	—	-0.293 (0.260)
p-value of F-test	—	—	0.263
Obs.			144

Median regression estimates for *Ending Wealth* dependent variable (column 1). OLS estimates for *Ending Log Wealth* and *Ending Servants* dependent variables (columns 2-3).

Appendix B: Data Issues

In this section we discuss various issues associated with our data sources and data collection process. First, we describe in more detail some characteristics of the political environment during this period. Then, we provide an in-depth discussion of the reliability of the census wealth, as well as a detailed description of the process for matching politicians to their census records, and for the coding of servants.

Electoral Data

The electoral data consists of election results for each election to the U.S. House of Representatives between 1845 and 1875. These data were collected by the Inter-University Consortium for Political and Social Research (ICPSR), and we revised and updated the ICPSR dataset using Dubin (1998).²

Additional information on the winners of each election is available from a biographical dataset compiled by the ICPSR, as well as the *Biographical Directory of the U.S. Congress*.³ These provide information on the year and place of birth, profession and career, and the county of residence at different points in time. We use Martis (1982) to match counties and cities to congressional districts. These sources were useful for cleaning the electoral database (e.g., finding cases in which the election winner did not serve in congress), and also for matching candidates to census records. Finally, we use Canon, Nelson and Stewart (1998) to construct measures of party leadership and committee positions.

It is important to describe several features of the electoral and political environment of the mid-19th century, because they are relevant for our analysis and quite different from the environment today. As shown in Table B.1, there were over 2,300 races to the House of Representatives between 1845 and 1875, involving about 3,000 distinct candidates. Congressional elections were quite competitive compared to today. Approximately 50% of all winners received less than 55% of the vote (as a percentage of the top two candidates' votes), and about 33% received less than 53% of the vote

Also, there were very few “career congressmen” compared to today. Fewer than 24% of those who won their first race ran for Congress in more than two elections, and only 16% served three or more terms. The numbers are similar for those who won their first race by a narrow margin – only 21% ran more than twice for Congress and only about 12% served three or more terms. In sum, few congressmen seem to have been interested in long careers in Congress, irrespective of whether they first entered this office by a small or large margin. However, many of those who served in Congress served in other offices, both before and after their congressional service, so a larger number of men were “career politicians.”

Those who lost their first race almost never served in Congress – 80% of those who lost their first race never even ran again, and fewer than 9% served in Congress during our period. As one would expect, those who lost their first race by a small margin were more likely to

²ICPSR Number 1. Dubin (1998) is essential not only for providing more complete and accurate election returns, but also for providing the first and middle names or initials of many candidates for which the ICPSR data provides only the last name, or the last name plus the first initial. Dubin (1998) also provides information on many special elections that are not included in the ICPSR data set.

³ICPSR Number 7428, and <http://bioguide.congress.gov>.

try again and succeed in the future: about 28% ran again and about 16% ended up serving in Congress.

Finally, a note on how we define a candidate’s first race. Some candidates – about 5% of our sample – ran more than once, for non-consecutive congresses. In these cases we define a “spell” as a set of consecutive election attempts separated by at least one Congress in which they did not run. We treat the spells as separate “quasi-experiments” and consider the vote share in the first election of each spell. Thus, we count as “close winners” or “close losers” those who won or lost the first race of any of their election spells by a small margin.

Close Elections and Sorting at the Threshold

Recent papers by Snyder (2005), Carpenter et al. (2011) and Caughey and Sekhon (2012) criticize RDD studies that rely on close elections, arguing that there are anomalies even very near the 50% threshold. They show that in U.S. House elections, incumbents win noticeably more than 50% of the very close races – especially those where the winning margin was less than 1% – and that candidates from the party in control of state offices, such as the governorship, secretary of state and state house and senate, hold a systematic advantage in extremely close elections.

These papers do not analyze the time period we study, so we provide some evidence here. In races where the winner’s margin was 1% or less and an incumbent was running, exactly 50% of the incumbents won and 50% lost. In the 2% window, 56% of the incumbents won, but this figure is not statistically different from 50% at the .05 level. In the wider 3% window 60% of incumbents won, and the percentage is statistically significant. We check the hypothesis explored in Carpenter et al. (2011), by studying outcomes from the point of view of the party controlling the governor’s office at the time of the election. In races where the winner’s margin was 1% or less, the candidate whose party controlled the governor’s office won 53% of the time, but this figure is not statistically different from 50% at the .05 level. In the 2% window, the candidate from the governor’s party won 54%, and again the percentage is not statistically different from 50% at the .05 level. In the wider 3% window, 56% of the candidates from the governor’s party won, and the percentage is statistically different from 50%. In fact, these outcomes are not too surprising, since as Snyder et al. (2013) show, in districts with a “normal vote” different from 50% we actually *expect* candidates from the favored party to win more than 50% of the time except in extremely small windows around the threshold. And incumbents, as well as candidates whose party won the governorship, tend to be from the favored party rather than the disadvantaged party in a district.

Interestingly, the “sorting” at the threshold is concentrated in the post-Civil War congresses. In the congresses of the 1850s (31st-36th Congresses) there is no significant evidence of sorting. The figures for the percentage of close races won by the incumbent for the various windows are as follows: 46% for the 1% window, 57% for the 2% window, and 57% for the 3% window. Similarly, the figures for the percentage of close races won by the governor’s party are as follows: 47% for the 1% window, 52% for the 2% window, and 53% for the 3% window. None of these are statistically different from 50% at the .05 level. We find the same patterns – no significant evidence of sorting – for our main sub-sample of interest, the Civil War years (37th-39th Congresses). The figures for the percentage of close races won by the incumbent for the various windows are as follows: 50% for the 1% window, 51% for the 2% window, and 55% for the 3% window. Similarly, the figures for the percentage of close races

won by the governor’s party are as follows: 51% for the 1% window, 53% for the 2% window, and 56% for the 3% window. Again, none of these are statistically different from 50% at the .05 level.

Census Wealth Data

In this section we present our census records matching procedure and address any potential selection issues that may arise from our inability to find the census record of every single candidate in every census year. Next, we present evidence from previous studies, and new evidence that we have compiled ourselves, that suggest that census wealth data can be trusted and does not introduce any major biases for the purposes of our analysis.

Matching of Candidates to their Census Records

We attempted to find the census record in each census year of every candidate for the U.S. House during the period 1845-1875. We restrict our analysis to candidates who obtained at least 25% of the vote in at least one election. All census records before 1940, including slave schedules, are available in ancestry.com. This is a genealogical website that provides images of the original census records and a search engine to locate records by first, middle and last name, as well as year and place of birth, and place of residence. We initially used PERL scripts to automatically match candidates to census records using the first and last name, as well as geographic information based on the county or counties located in the congressional seat sought. In the case of winners we used information provided in the biographical databases on the year of birth, county, and town of residence to further narrow the search. In the case of losers, for which more precise biographical information was not always available, we matched candidates by first and last name and verified that they were living in a county contained in the congressional district in which they were running. Despite the automated matching done by the scripts, the data collection process was still very labor intensive since we had to manually enter wealth figures and occupations. Also, the scripts found fewer than 50% of the cases, due to typographical errors in the information provided by ancestry.com or to candidates who moved. We had to locate other cases by searching manually, checking alternative spellings and misspellings of names, checking miss-coded birth years, and searching in other counties and states for candidates who moved. Figure B.1 shows one sample census record, the page on which Abraham Lincoln was listed in 1860 (noted by the arrow). This illustrates the various types of data that had to be coded manually for each record – in particular, occupation, wealth, and the number of servants.

We successfully located and entered data from about 8,000 census records, out of a universe of 10,250 cases spanning the period 1850-1880.⁴ This corresponds to an overall success rate of nearly 80%. For the 1850-1870 censuses the success rate is even higher, 84%. The success rates in these censuses is 92% for candidates who served in Congress during our period, and 77% for those who did not. In the RDD sample using a 3% bandwidth, the corresponding figures are 92% and 82%, respectively. The lower success rate for losers is not surprising, since we did not always have biographical information that allowed us to perform a more detailed

⁴The biographical information allowed us to know the year of death of those who served. Naturally, we did not attempt to find the census record of those who were dead in a given census year.

search. Our overall success rate compares favorably to that reported by previous studies. Steckel (1988) reports a 59% success rate when trying to match over 1,800 household heads from 300 different counties in the 1850 and 1860 censuses. Ferrie (1996) reports a success rate of only 19% when trying to match a sample of over 25,000 males included in the IPUMS sample for 1850 to the 1860 census.

Naturally, we were not able to find the census record of every single candidate in every single year. This could lead to concerns of selection bias in our sample. One encouraging fact is that our overall success rates were similar across the different census years – we found 81% of the census records in 1850, 86% of the records in 1860 and 83% of the records in 1870. Failure to match a congressional candidate to its census record in a given census year could happen for a variety of reasons. First, there is the possibility of underenumeration. Evidence reported by Steckel (1988) suggests enumeration rates were around 85%.⁵ In addition, there were frequent typos in the transcription of the original census records which made it harder to find some of the candidates. Steckel (1988) examined a sample from the 1860 census, and found that 8.8% of the transcriptions were searchable errors (minor mistakes or typos), while 15.8% constituted non-searchable errors (that is, errors that would have made it impossible to find an individual). Migration and death were additional factors which complicated the matching of individuals, though this was less of a problem for winners for whom we had not only their exact year of death but also some information on migration reported in the congressional biographies. For the case of the losers however, it is likely that some of the candidates we failed to match in the later census years had already passed away.

A large fraction of the candidates we failed to match were those individuals with very common names. Also, in some cases we found two or more census records with the same first name, last name, and middle initial in the same congressional district and of the relevant age. In order to minimize our type I error, we decided to exclude these candidates with very common names that we could not match. However, we are confident that this should not introduce any systematic bias in our sample. In fact, for their 1850 and 1860 samples Steckel (1988) and Ferrie (1996) ran logit regressions of a “common name” dummy against characteristics such as location of residence (region and city size) and other personal characteristics such as real and personal wealth, ethnicity, illiteracy and occupation. Their results show that while common names occur less often in southern states and in cities with less than 75,000 inhabitants, having a common name is not correlated with real or personal wealth. In order to explore in greater detail all these matching issues Steckel (1988) ran a logit regression of a “failure to match” dummy against different personal and geographic characteristics and found that for people in the North Central and Mountain and Pacific regions, those who were illiterate, foreign born, or living in cities larger than 75,000 inhabitants were less likely to be matched, while those with large real estate wealth or living in smaller cities were easier to match (though the coefficient on wealth is quite small). Similarly, Ferrie (1996) found that the probability of a successful match was higher for households in the northeastern states, for married individuals, and for household heads involved in farming activities, and it was

⁵However, as mentioned by Wright (1970b, 149), there were large efforts toward avoiding underenumeration and the instructions on coverage were explicit and italicized in the government circulars: “The assistant marshall shall make the enumeration by actual inquiry at every dwelling house, or by personal inquiry of the head of every family, and not otherwise.” Also, the assistant marshals were sworn to carry out their instructions and violations were subject to penalties. Another incentive for avoiding the undercounting of individuals was provided by the fact that compensation for assistant marshals was on a per-entry basis.

lower for foreign-born and older individuals.

For our purposes, the only concern would be if we were differentially likely to find the census records of close winners (relative to losers) or of individuals with different wealth levels. Thus, in columns 1 and 2 of Appendix Table B.2 we report the estimates of linear probability models for the close election sample where the dependent variable is a “failure to find” dummy in 1860 and 1870. The independent variables of interest are a dummy variable indicating whether an individual served in Congress in the prior decade, the log of wealth reported in the previous census year, and the interaction of these two. The point estimates reveal that we are neither more nor less likely to find the census records of those who won their first race by a narrow margin for those who were originally richer. This suggests that the failure to find some census records does not introduce a systematic bias in our analyses.

Reliability of the Census Data

The wealth data provided in census records was self-reported by the respondents, and was not checked for accuracy in other ways by government officials. Given this, it is important to discuss the possible problems associated with these data, as well as work done by previous authors and ourselves to establish its reliability.

All censuses were administered in person, by U.S. assistant marshals. The exact instructions given to the enumerators for the 1860 census were as follows. For real estate wealth:

“Under heading 8, insert the value of the real estate owned by each individual enumerated. You are to obtain this information by personal inquiry of each head of a family, and are to insert the amount in dollars, be the estate located where it may. You are not to consider any question of lien or encumbrance; it is simply your duty to enter the value as given by the respondent.”

Similarly, for personal wealth:

“Under heading 9, insert (in dollars) the value of personal property or estate. Here you are to include the value of all the property, possessions, or wealth of each individual which is not embraced in the column previous consist of what it may; the value of bonds, mortgages, notes, slaves, live stock, plate, jewels or furniture; in fine, the value of whatever constitutes the personal wealth of individuals. Exact accuracy may not be arrived at, but all persons should be encouraged to give a near and prompt estimate for your information. Should any respondent manifest hesitation or unwillingness to make a free reply on this or any other subject, you will direct attention to Nos. 6 and 13 of your general instructions and the 15th section of the law.”

According to Wright (1970b, 38), there was some concern at the time regarding the “suspicions of the interviewees about the intentions of the enumerators and about the uses to which the information divulged would be put.” Also, some observers expressed fears that the information might be used for tax purposes, although much of this was probably political posturing. For instance, a southern journal inquired whether “this Federal prying into the domestic economy of the people was not a precursor to direct taxes” (Wright, 1970b, 38).

Williamson and Lindert (1991) discuss the reliability of the 1860 census returns, in particular the possibility that respondents gave casual, and therefore inaccurate, answers to the census takers. They note that “a large number of [households] may have reported zero wealth in order to avoid the bother of estimating asset value (in fact 38% of free adult males reported property less than \$100 in the 1860 census sample) but it is hard to tell what share of these actually reported zero wealth” (Williamson and Lindert, 1991, 41). They add that “at the other end of the wealth spectrum, one might speculate that the very rich overstated their wealth in the 1860 and 1870 censuses, but this is a hard conjecture to sustain [and] we know of no clear bias in the estimates, either for the 1774 or for 1860 [censuses].”

There are, however, several reasons to believe that these issues are not a major concern for our purposes. First, the information collected by census officials was, as a matter of policy, strictly confidential. The U.S. Secretary of the Interior stated the policy as follows:

“... all marshals and assistants are expected to consider the facts intrusted to them as if obtained exclusively for the use of the Government, and not to be used in any way to the gratification of curiosity, the exposure of any man’s business or pursuits, or for the private emolument of the marshals or assistants, who, while employed in this service, act as agents of the Government in the most confidential capacity.”

This policy was reinforced for the 1870 census, with the following:

“No graver offense can be committed by assistant marshals than to divulge information acquired in the discharge of their duty. All disclosures should be treated as strictly confidential, with the exception hereafter to be noted in the case of the mortality schedule [where professional review by a local physician was authorized]. Information will be solicited of any breach of confidence on the part of assistant marshals. The [Department of Interior] is determined to protect the citizen in all his rights in the present census.”

Moreover, Wright (1970b) adds that “enumerators were instructed to approach every family ‘with civil and conciliatory manners’ and to ‘secure confidence and good will.’” They were to warn recalcitrant respondents of the penalties for refusal to answer or for giving false information. Above all, the information received was to be kept strictly confidential, and interviewees were to be assured that this was true. By 1860 one would suppose that much of the apprehension of the earlier years would have been dissipated by virtue of experience. Steckel (1990) notes that although the census did not verify self-reported wealth figures, it would have been difficult to conceal real estate holdings and thus these figures were probably reported reliably.

Second, even if some respondents were worried that the information provided would not in fact be kept confidential, there was no clear incentive for under-reporting or over-reporting wealth. There was no federal tax on wealth at the time, and no estate tax. Personal vanity, however, might have lead to some over-reporting.

Census Wealth Data in Previous Studies

Several previous studies have assessed the reliability of the census data in different ways. Soltow (1975) used random samples from the 1850-1870 censuses to analyze the evolution of wealth inequality in the U.S. He found that “wealth averages for the samples in the years 1850-1870 are generally in line with estimates made by various authorities on wealth distribution. Growth rates are similar to those found for GNP per worker by Kusnetz and commodity output per worker by Gallman” (Soltow, 1975, 6). He also found evidence that the census wealth figures were consistent with aggregates obtained by county assessors. The fact that patterns of wealth and wealth inequality suggested by the census data were in line with those described by other sources provides evidence in favor of the reliability of census wealth.

Another group of studies compared wealth reported in the census sheets with taxable wealth. In an early paper, Gallman (1969) used a random sample from a 1 percent sample of census sheets from Baltimore, New Orleans, St. Louis, Maryland (excluding Baltimore) and Louisiana (excluding New Orleans), and found that “the aggregate value of property reported on the population schedules exceeded the value of property assessed for tax purposes by more than 50% and the estimated true value of taxable property by almost 20%” (Gallman, 1969, 17). The discrepancies could be explained by the fact that not all property was subject to tax, and also that some of the property listed in the tax records belonged to corporations and other institutions (not enumerated in the population census). Moreover, individuals owning personal property worth less than \$100 were apparently not obliged to list their property in the census but presumably were obligated to list for tax purposes. Gallman (1969, 17) concludes that “the large value of property reported on the population schedule, relative to the estimated true value of taxed property, is good evidence that the enumerators and respondents met their obligations.”

One of the most relevant studies for our purpose is Steckel (1994), who matched 20,000 households from the federal census of Massachusetts and Ohio with real and personal property tax records from 1820 to 1910. Simple scatter plots of taxable wealth against census wealth reveal that for Massachusetts most observations line up around the 45 degree line, indicating an average coincidence of census and taxable wealth. The data from Ohio shows that census wealth tends to exceed taxable wealth, in line with the findings of Gallman (1969). Steckel (1994, 79) explains the discrepancies by pointing to the “old-fashioned practice of valuing property at about one-half or two-thirds of what it was worth” differences in the dates of the wealth valuations relative to the census enumerations, and the fact that some property exempt from taxation, particularly personal property, may have been included in wealth totals reported by the census. In addition, some individuals may have owned wealth in taxing jurisdictions outside their place of residence. Finally, one cannot ignore the fact that some individuals evaded taxes. In addition, the census may have reported family or household property, including that owned by children or by a spouse, with the head, whereas taxable property included only that owned personally by the head. In order to establish any systematic discrepancies between census and taxable wealth, Steckel (1994) ran regressions of taxable wealth on census wealth and characteristics of the household head, for every census year. The results suggest no systematic associations between the discrepancies and any of the variables with the exception of gender status (taxable wealth is well below census wealth for women). This, however, is easily explained by the fact that widows received

favorable tax treatment. Moreover, and despite the discrepancies between the sources of data pointed out above, inequality measures calculated with both census and taxable wealth are remarkably similar. Steckel (1994, 84) concludes by stating that “these data [wealth from census schedules] are particularly valuable for analyzing patterns of wealth holding.”

Comparing Census Wealth Data with Other Wealth Sources

A potential concern in our context, is whether politicians are more likely to misreport the true value of their wealth. In order to explore this issue, we searched for the 1850 and 1860 census records of all of the individuals in Forbes and Greene’s (1851) *The Rich Men of Massachusetts*, a book that purports to give the wealth of (most of) the richest 1,500 men in Massachusetts as of about 1851 as reported by independent parties.⁶ We matched the individuals in this book to lists of mayors, state legislators and congressmen who served during the period in order to explore any systematic discrepancies between both sources by politicians, relative to non-politicians. As shown in Appendix Table B.3, the correlation between wealth reported in this book and wealth recorded in the censuses of 1850 and 1860 is relatively high. More importantly, there is no evidence of significant under-reporting or over-reporting of politicians compared to non-politicians. This provides further confidence in the reliability of the census data.

Detecting Corruption using Census Data: The Lacrosse & Milwaukee Railroad Scandal

Finally, and perhaps most importantly for our purposes, we ask whether census wealth data can be used to detect wealth accumulation of individuals known to have been corrupt and to have received bribes. We explore this in the context of a prominent scandal: the Lacrosse & Milwaukee Railroad scandal.

In 1856 the Federal government ceded land for a major railroad project to the state of Wisconsin, but left it to the state to decide which railroad(s) would receive the grant. Several railroads competed for the land grant in 1856 and 1857, among them the Lacrosse & Milwaukee Railroad Company. In 1858, amidst fears that the railroad was in danger of going bankrupt, creditors demanded an audit of its accounts. A major scandal broke as the accounting revealed many troubling items, most importantly that the railroad had spent over \$800,000 to bribe various federal and Wisconsin officials.⁷

⁶The book provides information on total wealth while the 1850 census, as note above, reported only real estate wealth. Thus, we matched individuals in the book with the 1860 census as well as the 1850 census, in order to have a measure of total wealth despite the fact that the 1860 census measure is 9 years later.

⁷At the Federal level, the railroad paid \$105,000 to congressmen for federal approval of the land grant. However, the bulk of the bribes were paid to various Wisconsin officials, including 59 assemblymen and 19 state senators, \$10,000 to state Supreme Court Justice Abram Smith, and \$17,000 to state house clerks to expedite business. Republican Governor Coles Bashford received over \$50,000 in stocks and \$15,000 in cash. The railroad also appeared to be concerned about the scandal becoming public, since it paid \$25,000 to silence a key member of the Wisconsin state legislative committee investigating the scandal (Horace A. Tenney), and also paid bribes to journalists, including \$10,000 to the editor of the *Milwaukee Sentinel*. Government officials however, appeared to be less worried about their involvement in the scandal and taking bribes from the railroad. State assemblyman George W. Parker explained why the railroad’s payments were not bribes: They were made *after* voting for the railroad’s bill, not before. Moreover, having adjourned, they “were no longer a Legislature... [so] our acceptance could in no way be considered or regarded as a bribe... it could

Especially useful for our purposes, the committee appointed to inquire into the alleged bribery of the railroad reported the exact value of the bribes received by all Wisconsin officials involved in the scandal. The report revealed, for example, that 49 state representatives each received \$5,000 in bribes, 7 more received \$10,000 each, 1 received \$20,000, and 1 received \$25,000. State senators generally received larger bribes – 10 received \$10,000, 4 received \$20,000 and 1 received \$5,000. We attempted to find the census records in 1850 and 1860 of all those Wisconsin officials who served in the state government in 1856.⁸ This allows us to test whether those who received larger bribes, accumulated, on average, more wealth between 1850 and 1860. To do this analysis we can estimate a regression of the form:

$$Wealth_i^{1860} = \gamma_0 + \gamma_1 Wealth_i^{1850} + \gamma_2 BribeAmount_i + \gamma'_3 \mathbf{X}_i + \epsilon_i \quad (1)$$

where \mathbf{X}_i includes age and age² as controls. The coefficient of interest is γ_2 , which measures the extent to which larger bribes in 1856 are reflected in higher census wealth in 1860.

In Appendix Table B.4 we report estimates of γ_2 in equation (??). Columns 1 and 2 report the estimated coefficients when all wealth variables – wealth in 1850 and 1860 and bribe amount – are expressed in levels. Column 1 shows median regression estimates and column 2 shows OLS estimates. Column 3 reports the OLS estimate when the logs of these variables are used.⁹ The estimates reveal that the bribe amount is strongly and positively correlated with reported wealth in 1860, and the coefficient is statistically significant at conventional levels. The point estimate in column 1 implies that a bribe of \$10,000 translates into about \$10,000 in additional census wealth in 1860. These results provide further evidence of reliability of the census wealth data. They suggest that we can detect changes in wealth that occur over a decade resulting, partly, as a consequence of rents and bribes from holding office.

Wealth Misreporting and Failure to Report

Another measurement issue concerns the fact that it is sometimes difficult to distinguish between respondents with zero wealth and respondents who refused to provide any information to the census marshall, or instances where the marshal did not request the information.¹⁰ In both situations census marshals left the census record fields blank, which makes it difficult to distinguish “zero” wealth from “wealth figure not available.” It is clear that in most cases

in no way affect our honor or integrity as men... and further, that coming at the time it did, and especially when we remembered that the Company had just received at our hands 3 to \$10,000,000 worth of lands as a gratuity, we could not find it in our hearts to refuse.”

⁸We found both the 1850 and 1860 records for 94 out of 139 government officials who were in power during the 1856 legislature, when the scandal took place.

⁹We only estimate the regression for the government officials who, according to the committee’s report, received a positive bribe by the railroad. It is not clear what to assume regarding officials not listed in the report. In particular, 21 state legislators supposedly did not receive any bribe. Of these, 8 voted in favor of the Lacrosse & Milwaukee Railroad’s bill, and 13 voted against. Those who voted “yes” may have benefited in other ways – e.g., the committee report noted that John Fitzgerald had a direct interest in the railroad – and those who voted “no” might have received payments from other railroads. If we assign a value of zero bribes to those not mentioned in the report, and who voted *against* the railroad, then the estimate for the specification in levels reported in column 1 remains positive and statistically significant. However, the point estimate for the log specification falls substantially and is no longer statistically significant.

¹⁰Steckel (1994) notes that the incidence of “zero” wealth responses suggests that “some census enumerators failed to acquire accurate information on the value of wealth holdings through lack of diligence, non-compliance of the household, or ignorance of the respondent” (p. 80).

an empty wealth field corresponds to zero or very low wealth, since they are in the census records of very young individuals, and individuals with low-paying occupations such as laborers and domestic servants. However, one also finds census records of individuals known to be wealthy at the time, such as Fernando Wood, who despite reporting being rich in 1850 and 1860, did not report any wealth figure in 1870. For these individuals it is clear that the missing wealth figure did not correspond to zero wealth as can be inferred from the fact that there were various servants working for them.

The potential measurement error introduced by this issue should only be a concern for our purposes if there is a differential likelihood of not reporting any wealth by close winners and close losers. To explore this, in Appendix Table B.2 we focus on the close election sample – i.e. candidates who won or lost their first election by a margin smaller than 3%. Columns 3 and 4 of this table report the estimates from linear probability models in which the dependent variable is a dummy variable that indicates whether the candidate failed to report any wealth in 1860 and 1870 as a function of whether the individual served in Congress in the decade prior to that census year, reported wealth in the previous census, and the interaction of these two terms. All regressions include state fixed effects, age and age², and the same occupational dummies included in our main analysis. The results show that election winners, or those originally richer, are not more likely to fail to report their wealth. This gives us further confidence that this phenomenon does not introduce any systematic bias in our results.

Coding of Domestic Servants

As an alternative measure of wealth, we also collected information on the number of servants living with each individual in every census year. This is our only proxy for wealth for 1880. Servants living in every dwelling had to be reported to the enumerator and were, naturally, harder to hide and misreport than real or personal wealth figures. Servants were typically reported at the bottom of each household’s record, following the enumeration of the relatives of the household head. We classified as servants all individuals who had a variation of one of the following occupations: servant, domestic servant, cook, coachman, nurse, gardener, laundress, seamstress, washwoman, waiter, hostler, or butler. In 1850 most servants’ occupations were not listed. Inspection of the 1860 and 1870 records revealed that the majority of servants were young, foreign-born women, or were non-white. Therefore, for 1850 we classified as servants all individuals who did not share the same family name of the household head and had no occupation listed, but were Irish, German, Scandinavian, or non-white women aged 30 or younger.

Moreover, the number of servants is strongly correlated with reported wealth in the census: a regression of log total wealth against the number of servants reveals that an additional servant was associated with an increase in wealth of about 58% in 1860 and 57% in 1870. The correlations are highly statistically significant, with *t*-statistics of about 30. Evidently, the relationship between servants and wealth is highly non-linear – the fit in terms of R^2 is much higher using log of wealth than a linear regression of servants on wealth in levels. We also used information on servants to detect cases in which reported wealth figures appear to be unreliable. Consider all candidates with 1 servant. We compute the 10th percentile of the distribution of wealth for these individuals, and recode the wealth as missing for candidates whose reported wealth is below this threshold. We repeat this for all other values of the number of servants. None of our results change substantially as a result of these

transformations.

References

- Canon, David, Garrison Nelson, and Charles Stewart. 1998. "Historical Congressional Standing Committees, 1st to 79th Congresses, 1789-1947" (Computer file).
- Carpenter, Daniel, Brian Feinstein, Justin Grimmer, and Eitan Hersh. 2011. "Are Close Elections Random?" Unpublished manuscript.
- Caughey, Devin M. and Jasjeet S. Sekhon. 2012. "Regression-Discontinuity Designs and Popular Elections: Implications of Pro-Incumbent Bias in Close U.S. House Races." *Political Analysis* 19: 385-408.
- Dubin, Michael J.. 1998. *United States Congressional Elections, 1788-1997: The Official Results of the Elections of the 1st through 105th Congresses*. Jefferson, NC: McFarland and Company, Inc.
- Ferrie, Joseph. 1996. "A New Sample of Males Linked from the Public Use Micro Sample of the 1850 U.S. Federal Census of Population to the 1860 U.S. Federal Census Manuscript Schedules." *Historical Methods* 29: 141-156.
- Forbes, Abner and J.W. Greene. 1851). *The Rich Men of Massachusetts*. Boston: W.V. Spencer.
- Gallman, Robert E. 1969. "Trends in the Size Distribution of Wealth and Income". In *Six Papers on the Size Distribution of Wealth and Income. Studies in Income and Wealth*, volume 33, ed. Lee Soltow. New York: Columbia University Press.
- Inter-university Consortium for Political and Social Research and Carroll McKibbin. 1999. "United States Historical Election Returns, 1824-1968" (Computer File). ICPSR00001-v3. Ann Arbor, MI: Inter-university Consortium for Political and Social Research (distributor).
- Inter-university Consortium for Political and Social Research and Carroll McKibbin (1997. "Roster of United States Congressional Officeholders and Biographical Characteristics of Members of the United States Congress, 1789-1996: Merged Data" (Computer file). 10th ICPSR ed. Ann Arbor, MI: Inter-university Consortium for Political and Social Research (producer and distributor).
- Martis, Kenneth C.. 1982. *The Historical Atlas of United States Congressional Districts: 1789-1983*. New York: The Free Press.
- Snyder, Jason. 2005. "Detecting Manipulation in U.S. House Elections." Unpublished manuscript.
- Snyder, James M., Jr., Shigeo Hirano, and Olle Folke. 2013. "A Note on Sorting at the 50-50 Threshold in RDD Studies Using Electoral Data." Unpublished manuscript.

- Soltow, Lee. 1975. *Men and Wealth in the United States, 1850-1870*. New Haven: Yale University Press.
- Steckel, Richard H. 1988. "Census Matching and Migration: A Research Strategy." *Historical Methods* 21: 52-60.
- Steckel, Richard H. 1990. "Poverty and Prosperity: A Longitudinal Study of Wealth Accumulation, 1850-1860." *The Review of Economics and Statistics* 72: 275-285.
- Steckel, Richard H. 1994. "Census Manuscript Schedules Matched with Property Tax Lists." *Historical Methods* 27: 71-85.
- Williamson, Jeffrey G. and Peter H. Lindert. 1980. *American Inequality: A Macroeconomic History*. New York: Academic Press.

Table B.1: Summary Statistics on Congressional Races

# of Races	= 2340			
# of Races w/Margin < 55%	= 1101			
# of Races w/Margin < 53%	= 740			
# of Candidates	= 2950			
# of Democrats	= 1540			
# of Whigs	= 491			
# of Republicans	= 736			
All Candidates				
	Won 1st Race		Lost 1st Race	
# who run 1 time	600	37.9%	1925	79.7%
# who run 2 times	614	38.7%	288	11.9%
# who run 3+ times	371	23.4%	203	8.4%
# who win 0 times	0	0.0%	2217	91.8%
# who win 1 time	851	53.7%	98	4.1%
# who win 2 times	475	30.0%	62	2.6%
# who win 3+ times	259	16.3%	39	1.6%
Candidates with Close First Race				
	Won 1st Race		Lost 1st Race	
# who run 1 time	233	38.0%	474	71.6%
# who run 2 times	252	41.1%	91	13.7%
# who run 3+ times	128	20.9%	97	14.7%
# who win 0 times	0	0.0%	554	83.7%
# who win 1 time	368	60.0%	55	8.3%
# who win 2 times	173	28.2%	32	4.8%
# who win 3+ times	72	11.7%	21	3.2%

Table B.2: Assessing the Reliability of the Census Data				
	(1)	(2)	(3)	(4)
	Not Found 1860	Not Found 1870	No Report 1860	No Report 1870
Winner	-0.024 (0.034)	-0.028 (0.024)	-0.007 (0.021)	0.004 (0.033)
$\text{Log}(\text{Wealth}^{t-10})$	-0.002 (0.023)	0.007 (0.014)	0.018 (0.018)	0.004 (0.018)
$\text{Winner} \times \text{Log}(\text{Wealth}^{t-10})$	-0.011 (0.025)	0.002 (0.015)	0.009 (0.024)	0.007 (0.026)
Observations	277	303	277	303
R-square	0.079	0.107	0.081	0.069

Table B.3: Census Wealth vs. Wealth in <i>Rich Men of Massachusetts</i>			
	(1)	(2)	(3)
	Log Real 1850	Log Real 1860	Log Total 1860
RMM Wealth	0.781 (.054)	0.811 (.070)	1.010 (.059)
Politician	0.086 (.119)	0.122 (.143)	0.115 (.119)
Constant	0.770 (.620)	0.840 (.806)	-0.592 (.674)
R-square	.294	.267	.455
N	509	372	360
Correlation with RMM Wealth	.54	.52	.67

Table B.4: LaCrosse & Milwaukee Railroad Scandal			
	(1)	(2)	(3)
	Wealth 1860	Wealth 1860	Log Wealth 1860
Bribe Amount	1.000 (0.158)	1.923 (0.713)	
Log Bribe Amount			0.853 (0.249)
Observations	66	66	66
R-square	0.200	0.487	0.215

Median regression estimates in column 1 and OLS estimates in columns 2 and 3.

Figure B.1: Sample Census Page (with Abraham Lincoln)

Page No. 140

SCHEDULE 1.—Free Inhabitants in Dist. W. City of Springfield **in the County of** Springfield **State of** Ill. **enumerated by me, on the** 14 **day of** July **1860.** J. H. Christie Ass't Marshal
Post Office Springfield

1	2	3	4			7	8		10	11	12	13	14
			Age	Sex	Color		Value of Real Estate	Value of Personal Estate					
		John B. R. Worthen	5	m					Ills				
999	986	Lotus Miles	40	m		Secretary	7.000	2.500	N. Y.				
		Adella D. "	30	f					"				
		George W. Tyler	12	m					"				
		Sula M. "	3 1/2	fp					Ills				
		Reuben Randall	57	m		Servant			Conn				
		Mary D. Miles	1	f					Ills				
		Rebecca Redburn	24	m		Servant			Baden				
1000	987	Edward Bugg	48	m		Steamster	4.000	300	England				
		Nancy "	48	f					"				
		Hampton K.	19	m		Apprentice Carpenter			Ills				
1001	988	Henry Carrigan	50	m			30.000	300	Ireland				
		Bushon "	50	f					"				
		Hough "	26	m		Serv. Stable			"				
		Henry "	12	m					Ills		1		
1002	989	Abraham Lincoln	51	m		Lawyer	5.000	12.000	Ky				
		Mary "	35	f					"				
		Robert D. "	16	m					Ills		1		
		Willie W. "	9	m					"		1		
		Thomas "	7	m					"				
		M. Johnson	18	f		Servant			"				
		Phillip Dinkell	14	m					"				
1003	990	A. J. Snow	38	m				350	Ind				
		Margaret L. "	33	f					Ills				
		W. G. "	4	m					"				
		Frank "	2	m					"				
1004	991	Wm. S. Burch	46	m		Clerk	2.000	200	Ky				
		Mary E. "	15	f					Ills		1		
		R. S. "	12	m					"		1		
1005	992	Richard Doss	42	m		Bricklayer	4.000	4.500	N. Y.				
		Mahilda "	36	f					Ind				
		Aliza "	9	f					"				
1006	993	Wesley Lyon	69	m		Farmer	12.000	3.000	Pa				
		Thomas L. "	35	m					Ky				
		Houldah Bruze	42	f					"				
		George W. "	21	m		P. M.			Ills		1		
		Sophonia E. "	11	f					"		1		
		Clifton L. "	7	m					"				
		W. M. Satchee	20	f					"				
1007	994	Wm. L. Biddle	22	m					Ohio				

No. white males, 916 No. colored males, No. foreign born, No. blind,
No. white females, 119 No. colored females, No. deaf and dumb, No. insane,
No. paupers, No. service,

This corresponds to the page where Abraham Lincoln was listed in the 1860 Census. He reports real estate wealth of \$5,000, personal wealth of \$12,000, and one servant living in his dwelling. His occupation is Lawyer.

Appendix C: Before-and-After Design

A common concern with RDD estimates is that they provide *local* average treatment effects (LATE) for observations or individuals around the given threshold and thus may be uninformative regarding the effect of a given treatment on observations further away from the discontinuity (see Angrist and Pischke, 2009). This is often referred to as the lack of external validity of regression discontinuity estimates.

In the context of our analysis, one could think of several reasons why estimates for the returns to Congress based on individuals whose first election was decided by a small margin may lack external validity. One possible argument is that those who win by a narrow margin will tend to be more disciplined and less likely to engage in rent extraction since they are less safe and any minor wrongdoing may lead them to lose their seat in an upcoming election.¹¹ In this case, our regression discontinuity estimates may *underestimate* the actual rents from office enjoyed by congressmen who are safe in their seats and won their first election by a very large vote margin. However, it is easy to think of reasons why the regression discontinuity estimates *overestimate* the rents from Congress enjoyed by those individuals who did not enter Congress following a close election. If the electoral outcome of their first race was due to chance – say, an unusual partisan tide – then “close winners” may decide to make the most of their time in Congress and engage in rent-seeking, anticipating that they are unlikely to win reelection in any case.

We cannot directly assess how informative are the local treatment RDD estimates for the broader set of candidates. As an alternative, we report evidence based on a different empirical strategy – a simple “before-and-after” design – first introduced in Querubin and Snyder (2009). This approach does not rely on the comparison of winners and losers in close elections, but relies solely on data for individuals who actually won and served.

Figure C.1 below illustrates the approach.¹² Suppose we can observe the wealth of members of Congress at two different years $t-10$ and t . In Figure 4 we show this for $t = 1860$ (panel A), $t = 1870$ (panel B) and $t = 1880$ (panel C). We can then create indicator functions to classify all members of Congress who served in the years around this period. Let N_{EARLY} be an indicator function that takes a value of 1 for all members of Congress that served only during the 5 years preceding $t-10$ and zero otherwise. Similarly, T_{EARLY} takes a value of 1 for members of Congress that served only during the 5 years following $t-10$ and zero otherwise. We can also define similar indicator functions for congressmen who served around t . That is, T_{LATE} takes a value of 1 for all those who served only in the 5 years preceding t and zero otherwise while N_{LATE} takes a value of 1 for congressmen who served only during the 5 years after t and zero otherwise. We can use these indicator functions to get a rough estimate of the returns to serving in Congress in the early and late part of the decade under consideration. For example, to get an estimate of the returns to Congress in the post-war years in the second half of the 1860s we can compare the accumulation of wealth between 1860 and 1870 for representatives that only served during the five years just before 1870 (i.e. all congressmen for which $T_{LATE} = 1$) with those that only served during the five years just after 1870 (i.e. all congressmen for which $N_{LATE} = 1$). The first group was “treated” by

¹¹However, the fact that there were few career congressmen during this period makes this case unlikely.

¹²See Querubin and Snyder (2009) for a more detailed discussion.

politics – had access to congressional rents that would appear in their 1870 wealth – while the latter group was not. Similarly, we can get an estimate of the returns from a seat in Congress during the Civil War years (early 1860s) by comparing the accumulation of wealth between 1860 and 1870 for those individuals that only served during the five years just after 1860 (i.e. those for which $T_{EARLY} = 1$) with those that only served during the 5 years just before 1860 (those for which $N_{EARLY} = 1$). In this case, only the latter group was treated by politics between 1860 and 1870. We can compare the different treatment and control groups around the different census years through a simple regression of the form:

$$Wealth_i^t = \beta_0 + \beta_1 Wealth_i^{t-1} + \beta_2 T_i + \beta_3' \mathbf{X}_i + \epsilon_i^t \quad (2)$$

where $Wealth_i^t$ is the wealth of congressman i in year t , $Wealth_i^{t-1}$ is the wealth of congressman i in the previous census year, T_i corresponds to one of the “treatment” dummies defined above, and \mathbf{X}_i corresponds to a set of control variables, including age and age², occupation dummies, and state fixed-effects.

The specific sample on which the above regression should be estimated depends on whether we are estimating the returns to a seat in Congress in the early or late half of the decade under consideration. In order to estimate the returns for the late part of the decade, we should estimate the regression on the sample of individuals that served only in the five years preceding or following year t (i.e. those for which either T_{LATE} or N_{LATE} equals 1). In this case, T_i will just correspond to the indicator function T_{LATE} . If we want to estimate the returns in the early half of the decade, the estimation sample should consist of all those who only served in the 5 years preceding and following year $t - 10$ (i.e. all those for which either T_{EARLY} or N_{EARLY} equals 1). Notice that for the 1870-1880 period we can only estimate the effect on the early half of the decade as we have only collected data for those who served between 1870 and 1875.

Table C.1 presents the estimates of the main coefficient of interest – i.e., β_2 in equation (??), the coefficient on T_i . Each panel reports the results for the different decades under consideration. The results are straightforward. First, and consistent with the RDD estimates, we find no evidence of a large positive return to serving in Congress during the 1850s. Point estimates for this period are small and statistically insignificant with the exception of the estimate in column 6 for the second half of the decade that is in fact negative and statistically significant. Similarly, and consistent with the results of Table 5, we find no evidence of abnormal wealth accumulation for those who served in the first half of the early 1870s (panel C). The point estimate is small and statistically insignificant. Finally, results for the 1860s – reported in panel B – confirm the same patterns of Table 5. There is no evidence of abnormal wealth accumulation for those who served in the post-war years during the second half of the decade; estimates in columns 2, 4 and 6 are small and are not statistically significant. However, we do find evidence of a relatively large return to serving in Congress during the Civil War years in the early part of the 1860s. The point estimate for *Ending Log Wealth* reported in column 3 is 0.4, and the coefficient is highly statistically significant. The point estimates for *Ending Wealth* and *Ending Servants* in columns 1 and 5, respectively, are also large and the latter is also statistically significant at the 5% level.¹³

¹³In order to assess the validity of our approach, we also test for pre-existing differences in congressmen

The point estimates from this analysis reveal a similar pattern as that for the RDD estimates reported in Table 5. This analysis, which relies on a completely different source of variation, confirms our main findings and gives us further confidence on the relevance of our RDD estimates.

who served before and after the different census years. We do not show these results to save space but they are available upon request. Not surprisingly, congressmen who serve prior to a given census year are, on average, older than those who serve after the census year. To control for this difference, in our regressions we always include the age and squared age of the congressman to capture the (possibly nonlinear) effect that age may have on wealth accumulation. Most importantly, treated congressmen do not differ by their initial wealth, a variable that plausibly captures other relevant characteristics such as ability, education, or occupation. In addition – just as one example – we find that treated congressmen are no more or less likely to be lawyers. These similarities give us some confidence that the main difference between politicians at either side of the census year is their exposure to politics.

Table C.1: Effect of Serving in Congress on Wealth Before and After Analysis						
	Ending Wealth		Ending Log Wealth		Ending Servants	
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A: 1850-1860</i>	Served Early	Served Late	Served Early	Served Late	Served Early	Served Late
Served in Period	5665 (3453)	-1070 (3663)	0.095 (0.128)	0.035 (0.146)	-0.063 (0.188)	-0.347 (0.119)
Obs.	212	264	187	221	213	292
<i>Panel B: 1860-1870</i>	Served Early	Served Late	Served Early	Served Late	Served Early	Served Late
Served in Period	14443 (6705)	1085 (3496)	0.400 (0.147)	-0.030 (0.125)	0.351 (0.159)	0.205 (0.134)
Obs.	253	284	247	271	289	311
<i>Panel C: 1870-1880</i>	Served Early	Served Late	Served Early	Served Late	Served Early	Served Late
Served in Period	—	—	—	—	-0.019 (0.180)	—
Obs.					270	

Median regression estimates for *Ending Wealth* dependent variable (columns 1-2). OLS estimates for *Ending Log Wealth* and *Ending Servants* dependent variables (columns 3-6).

Figure C.1

Before and After Design

