AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

No endorsement of AgEcon Search or its fundraising activities by the author(s) of the following work or their employer(s) is intended or implied.

“When is Long-run Agglomeration Possible? Evidence from County Seat Wars”

Cory Smith, University of Maryland, cbsmith@umd.edu
Amrita Kulka, University of Warwick, Amrita.Kulka@warwick.ac.uk

Selected Paper prepared for presentation at the 2024 Agricultural & Applied Economics Association Annual Meeting, New Orleans, LA; July 28-30, 2024

Copyright 2024 by Cory Smith and Amrita Kulka. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.

When Is Long-Run Agglomeration Possible? Evidence from County-Seat Wars*

Cory Smith[†] ⓡ[‡] Amrita Kulka[§]

April 15, 2024

Abstract

We study the factors that foster long-run urban growth using a unique setting of close elections that determined “county seats” (capitals) in the frontier United States. Using a regression discontinuity design, we show that winning towns experienced rapid population and income growth as migrants coordinated on them as their counties’ economic centers. We show that this coordination was largest in smaller towns and in the years close to county creation. Using generalized random forests, we also show that the economic benefits were not zero sum locally: specific choices of county seat could increase long-run county population and income. County administration was limited in this era and did not play a substantial role in this process. Instead, these results are consistent with theories of economic geography in which there are strong local returns to scale and equilibria can be selected through historical and political events, even when geographic fundamentals are similar.

*We thank Ryan Schwartz, Fatima Najeeb, Ben Kremer, Roque Bescos, Jake Burgess, and Risa Wagner for their excellent research assistance. We also thank Treb Allen, Dave Donaldson, James Fenske, Hoyt Bleakley, Peter Nencka, Ezra Karger, Enrico Berkes, Randall Walsh, and other seminar participants at Dartmouth College and the University of Warwick for their advice and comments. This research was generously supported by the USDA through an AFRI grant.

[†]University of Maryland (AREC), corybsmith@gmail.com

[‡]Author order was randomly determined via coin toss.

[§]University of Warwick, amrita.kulka@warwick.ac.uk

1 Introduction

When and where do population shocks permanently shape economic geography? When individuals and firms prefer locating in larger towns and cities, agglomeration externalities can lead to multiple population equilibria in which individuals coordinate on a central location. In these cases, small shocks can have large and persistent effects on economic geography if they change which location is chosen. However, this dynamic is not universal: even some large-scale shocks only temporarily affect population distributions (for example, Davis and Weinstein (2002)). Overall, little is known about which historical conditions and shocks lead to coordination and which do not.

Several factors could influence how effectively people coordinate in their decisions of where to live and work. If switching from established equilibria is difficult, coordination during the initial periods of history will be most significant for determining modern outcomes. Alternatively, technological advances might make later events more relevant. Community characteristics could also shape this process. For example, if agglomeration externalities are only relevant for communities above a minimum size, smaller towns will rarely benefit from coordination. On the other hand, the preexisting advantages of large cities could render them insensitive to future shocks. There are two key challenges in empirically testing the importance of these kinds of factors. First, any examined variation must differentiate the impact of specific shocks from the fixed locational advantages that contribute to density. Second, the variation must occur across a range of time periods and community types.

We study long-run coordination using a natural experiment set in towns on the American frontier. Historically, the selection of the "county seat" (capital) served to centralize the populations of new counties formed in the nineteenth and early twentieth centuries. While county seats directly provided only a small number of administrative jobs, they nonetheless typically became the area's largest communities. Historical accounts of this period indicate that growth occurred because the status of the county seat enabled a town and its civic leaders to become the primary destination for new migrants and thereby coordinate the resulting population growth on their location. To establish the causality of this process, we collect an original data set on elections that determined the seats and analyze it with a regression discontinuity (RD) design. We show three key results that highlight the impact of coordination on towns' economies in both the short and long run, the importance of the timing of the shock, and the county-level impact of town-level coordination.

First, we show that county-seat elections led to rapid coordination in favor of the winner. On average, winning towns become 1.17 log points (223%) denser than their rivals by 2010, representing about an 11% shift of each county's population. Historical census data show that the increased population dates almost entirely to the first years after the election, with few long-term adjustments afterward.

Despite the relatively small size of towns in this setting, winners experience agglomeration benefits typically associated with larger cities: income measures rise by 8%–18%, white-collar employment by 10 percentage points, and per capita patents by 43%. Using consistent measures of income and employment from individual census responses, we show that towns' economic adjustments track their population growth, with most changes occurring soon after the elections. Overall, the results illustrate that coordination played a substantial, causal role in the long-run economic geography of these areas. Small numbers of votes in the early years of settlement selected a population and economic center from a set of alternatives.

Within the close elections in our sample, our results imply that county-seat elections determined the residence of about 5 million people today, or 7% of our sample counties' populations. Administrative and legal jobs connected to the county seat account for only a very small portion of this migration. At around 1% of employment even in winning towns, there are too few such jobs to account for the growth effects we see without multipliers vastly larger than estimated in previous studies. Therefore—in contrast to much of the literature on capital cities—the main mechanism behind winners' economic transformations appears to not be a direct capital-city effect but the agglomeration benefits from their long-term increase in the population. We provide evidence that winning towns become attractive to new residents by providing the sort of amenities that larger cities generally provide.

Our analysis requires the extensive collection of original data. We compiled results for 1117 county-seat election results in 811 counties, individually drawn from newspaper archives, county histories, and administrative records. In 2010, these counties had 71 million residents, or 23% of the nation's population. Building on recent advances in standardizing locations in public microdata (Berkes et al. 2023), we link competing locations to subcounty data from census records, post office locations, geographic characteristics, and tax filings. The result is a detailed political and economic history of our sample counties. The sample spans the population distribution of American communities, with a typical place home to several thousand people in 2010. The very upper tail consists of large cities including Phoenix, Dallas, and Miami. While most elections occurred close to the time of county formation, an important minority occurred in subsequent decades.

For our second set of findings, we identify the conditions that make it easiest to achieve coordination. We explore the heterogeneity of the above effects by considering both the timing of elections and the size of the community (Lin and Rauch 2022). Initial conditions play an important role in shaping the size of the county-seat shock. Exploiting variation in the timing of elections, we show that the magnitudes of population and income effects steadily diminish as an election is held later. This suggests that it was easier to influence the initial distribution of people than it was to change an existing equilibrium, in line with the idea that there are persistent durable investments that fix the location of places after a point. The result also provides further evidence that it is not the county government per se that is driving the results. If this was the primary factor, changing the county seat should have led to economic growth in both earlier and later periods.

In a similar analysis, smaller communities experience the largest population and income benefits from the county-seat shock. As population is an endogenous outcome, we focus on heterogeneity by a town's predicted size based on characteristics fixed at the time of the election. Towns in the top third of predicted density experience positive but smaller income and population gains relative to the full sample. Since the population and income effects attenuate in tandem, this implies that a fixed density gain has similar effects across a wide range of community sizes.

Third, we show that the choice of town as county seat can have long-term consequences for the county as a whole. Using the machine learning method of generalized random forests (Athey et al. 2019), we explore whether coordinating on specific locations affects outcomes at the county level. This algorithm allows us to make out-of-sample predictions of the ideal seat for maximizing long-run county density. Using a county-level RD design, we show that, on average, picking the selected location increased county density by approximately 51%. These same choices also appear to increase modern county income by 5%, which is significant at the 10% level. Although the location of the county-seat itself was a zero-sum choice, this shows that its effects on economic geography were not. However, the county-level changes we estimate are smaller than the town-level ones, indicating significant attenuation at this expanded geographic scale.

Our paper supports theories in which historical and political shocks can shift population equilibria under the right conditions. Our context is suited to answering this question as the election-based RD design balances geographic and other locational fundamentals across competing towns. We show that this coordination is most easily achieved in early periods, in smaller towns, and on smaller geographic scales. These results are relevant for designing place-based policies, but they represent mixed news: Rural areas facing population loss can clearly benefit from increased density, as there is no minimum threshold for agglomeration to be beneficial. But it is harder to shift populations today than in the past, and investments in these communities might not precipitate coordination as they once could.

We contribute to several strands of literature in urban economics and economic history. First, we add to our understanding of the effects of agglomeration forces, particularly over long periods. Why economic activity is geographically concentrated has been studied both across cities (Rosenthal and Strange 2004; Duranton and Puga 2004; Ellison and Glaeser 1999; Glaeser et al. 1992; Krugman 1991; Kleinman et al. 2023; Desmet and Rossi-Hansberg 2013; Redding et al. 2011; Combes et al. 2012; Michaels and Rauch 2018) and within cities (Ahlfeldt et al. 2015; Heblich et al. 2020). Other work (Greenstone et al. 2010; Qian and Tan 2021) examines the spillovers of economic activity within local regions, especially American counties. Our paper contributes to knowledge about the less understood effects of the population and occupational structure of smaller centers of local economic activity (towns) on their immediately surrounding areas. Our results also agree with the finding that the gains from agglomeration are unevenly distributed and can increase economic inequality (Baum-Snow et al. 2018; Baum-Snow and Pavan 2013; Ahlfeldt and Pietrostefani 2019). Commonly studied reasons for agglomeration include locations' fundamentals and natural advantages, productivity gains from proximity to dense workers and firms (Ahlfeldt et al. 2015; Heblich et al. 2020; Qian and Tan 2021), and knowledge spillovers (Davis and Dingel 2019; Combes et al. 2008; De La Roca and Puga 2017). We consider a novel initial source of agglomeration stemming from the population increase following a county-seat election victory. This allows us to study the effects of initial population distributions on longer-term outcomes.

Second, by exploiting historical variation from county-seat elections, we contribute to the literature on the factors that shape the growth and location of towns and cities (Bleakley and Lin 2012; Davis and Weinstein 2002; Banzhaf and Walsh 2013; Shertzer et al. 2018; Ager et al. 2020; Harari 2020; Brown and Cuberes 2020; Dell and Olken 2020; Howard and Ornaghi 2021). County-seat elections provide a clean source of identification of the effects of agglomeration separately from the fundamentals of a given location. Using this strategy, we are able to study the long-run effects of agglomeration, thereby shedding light on the persistence of historical shocks to productivity and infrastructure (Allen and Donaldson 2020; Hanlon and Miscio 2017). We exploit widespread, cross-town variation in large parts of the United States, which broadens our understanding of agglomeration in geographically or industrially concentrated areas (see Hanlon and Heblich (2020) for an overview). This connects to a broader literature on economic development on the American frontier in general (Donaldson and Hornbeck 2016; Hornbeck and Naidu 2014; Acemoglu et al. 2016; Nagy 2020; Bleakley and Rhode 2023). County-seat elections were a ubiquitous and salient feature of American life outside the original colonies, and the question of their impact on location choices and local economies is important in its own right.

Our paper also adds to the subset of this literature on the role of capital cities in economic development (Bluhm et al. 2021; Campante and Do 2014; Ades and Glaeser 1995; Bai and Jia 2021; Chambru et al. 2021), which finds that an expanded public sector presence can bring complementary investments in regional and national capital. However, the small scale of local government in our context means our primary mechanisms differ from those researched in this body of work. The rapid growth induced by county-seat selections in the frontier period greatly exceeds the population effects in other contexts, and we find null or negative effects on per-person public-goods spending. Overall, we argue that the long-term county-capital effect is driven by the historical migration effects rather than the political economy aspects discussed in this literature. If our effects were driven only by public employment, it would imply a multiplier of around 75, about two orders of magnitude higher than that found by other papers (Faggio and Overman 2014; Becker et al. 2021).

Finally, our paper speaks to the literature on the effectiveness and persistence of place-based policies (Ehrlich and Seidel 2018; Glaeser and Gottlieb 2008; Kline and Moretti 2014; Bartik 2020). Since county-seat contests occurred in largely rural counties, our findings on the effects of population growth help in understanding the benefits of population growth outside the superstar cities often studied. Studies of smaller communities are especially limited in many settings in which data on smaller towns are aggregated into larger administrative areas, limiting the ability of researchers to accurately observe them in any detail (Puga 2010). In contrast, our data collection efforts allow us to track small towns over a long historical span. Our results showcase an example of large population growth that was not catalyzed through high rates of government spending. Small locations benefited from the resulting agglomeration economies, perhaps even more so than larger ones. However, our analysis also suggest that the place-based shocks have the largest effects during the initial periods of a place's formation. As such, achieving growth now is likely more difficult than it was historically.

Section 2 discusses the historical significance of county-seat elections, Section 3 describes our data collection and assembly, and Section 4 introduces our research design and confirms the relevance (first stage) and exogeneity (balance) of the close-election RD design. Section 5 shows the long-run population and economic impacts of county-seat selection, and Sections 6–7 discuss variation of the results based on time and community or geographic scale. Section 8 tests alternative mechanisms, and Section 9 concludes.

2 Historical Background

2.1 County-Seat Functions

New counties on the American frontier faced the decision of where to locate their seat to conduct official business. Tangibly, a county seat required a courthouse, storage of official records, and related employment. County political and administrative officials such as commissioners, judges, or the county clerk would typically live or work in the county seat. The exact functions of county government and court systems depend on which functions states have chosen to devolve to localities, but some commonalities exist. County governments are usually responsible for property tax assessment, expenditure on schools, and maintenance of police forces. Some also provide other local public goods such as health facilities, utilities, or parks. County courts process many nonjury crimes (for example, traffic violations) and misdemeanor cases. Intangibly, being named the county seat accorded a large amount of prestige, suggesting a town's premier position within a county: "Towns desired county seats . . . because the designation brought increased status for the town" (Paher 1969).

Jobs in legal and administrative functions in county government represent only 0.24% of jobs nationally. Total county employment accounts for 1.7%, but the majority of these jobs are tied to specific communities rather than the seat (Census 2012); common examples include teachers and police officers. The proportions are larger but still tiny among winning towns; see Appendix Figure C.2. Despite this, a county's seat is its largest town in 78% of cases.[1]

2.2 County-Seat Elections

Many counties chose seats by holding county-wide elections. Most elections involved two locations competing for a majority, but automatic runoff and first past the post were also common. Less commonly, pre-election policies at the state level were designed to favor incumbents or other locations and required supermajorities to defeat these locations. In approximately 80% of counties in our sample, we observe only a single election, usually around the time of a county's creation (or incorporation). Consequently, election results were usually permanent, though not always.[2]

2.3 Historical Impact of the County Seat

> "Lamesa won the county seat election by five votes. . . . A town meeting the next day invited the citizens and merchants of Stemmons to move to the new county seat. . . . The offer was accepted and effected within several days."
> —Texas State Historical Association (*Handbook of Texas*, "Dawson County")

Despite the small size of county governments, county-seat status was a significant draw for

[1] Authors' calculation for all counties nationwide.

[2] Exceptions could occur based on subsequent elections or later legislative action to designate a new seat. To avoid issues related to the endogenous calls for elections, our main results focus on only the first election for each county; see Section 4.2.

migrants in the frontier period: "Status in the governmental hierarchy can be recognized as an important economic centralizing influence ... the important growth of any frontier village to larger status dates from its becoming a county seat" (Knight 1973). Securing the county seat was a clear signal that a town would become the premier community in its area, meaning the choice of seat served as a coordinating mechanism for early migrants. Guidebooks written for migrants to western states would often highlight county-seat status, indicating its salience; see Bleakley et al. (2023).

Individual county histories emphasize that these population influxes were large and directly followed the election. In Dawson County, Texas, the population movement happened within days, as noted in the epigraph. In Yolo County, California, Woodland displaced the incumbent county seat Washington in 1862. Understandably, "the people of Washington were loath to relinquish the prestige and advantages derived from having the seat of government in their town." In contrast, Woodland "entered upon an era of business and social activity. Buildings were erected rapidly, business developed and new people sought a home in the thriving new town" (Gregory 1913). Advocates for the competing communities were cognizant of the high stakes involved, leading to bitter disputes playing out in speeches and newspaper editorials. Some observers thus termed these contests "county seat wars" even though actual violence was quite rare (Schellenberg 1987).

Several further examples illustrate that broad and rapid movement of population was a general response to a county-seat election. In Wichita County, Kansas, Leoti triumphed over Coronado in 1888 and an "exodus began immediately ... during the fall and winter of 1888-1889 all of the hundred or more buildings in Coronado were moved to Leoti" (DeArment 2006). In Plainview, Texas, when "the first courthouse was ... completed ... [w]ithin a year the town grew to a population of seventy-five and had a hotel" (Davis 1952).

Histories also describe organized coordination. As mentioned in the epigraph, in Dawson County, Texas, this took the form of a town meeting enticing the citizens and merchants to migrate to the winning town. Another example is Phoenix, Arizona. After an 1871 county-seat election victory, its "commissioners quickly followed up their electoral success by offering more lots [parcels] for sale" (Mawn 1977). Here, the commissioners served as prominent organizers who correctly anticipated and capitalized on increased demand for residence in their town.

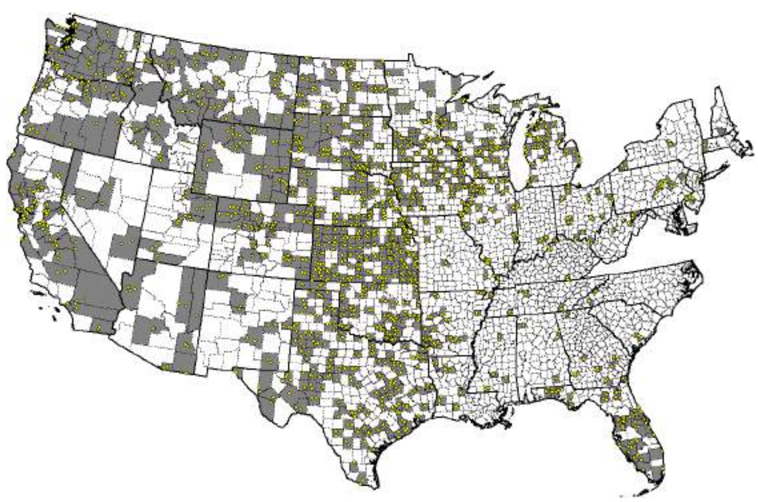
3 Data

We provide here a brief overview of our data sources and variable construction. More details are included in Appendix Section A.

3.1 County-Seat Elections

We assemble an original data set of county-seat elections. No comprehensive data set on such elections exists with national or even state-level coverage, meaning an extensive county-by-county search was required. We draw most heavily from historical newspapers, which frequently detailed elections' exact vote totals. We also consult county clerks, county histories, historical societies, and other administrative sources. Our data set consists of 1117 county-seat elections across 811 counties in 42 states. These data are mapped in Figure 1.

Figure 1: Sample Counties and Towns



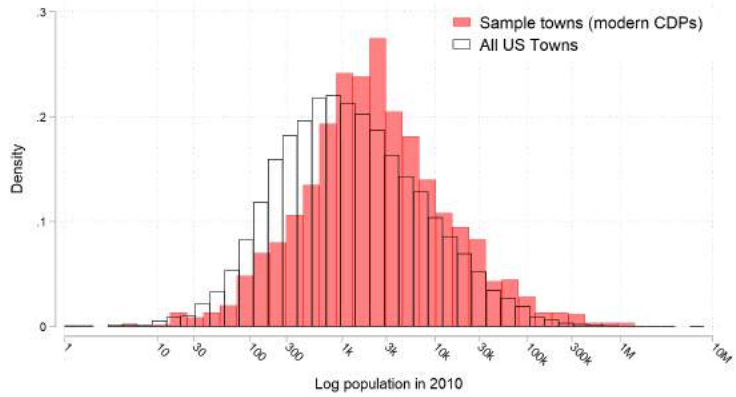
Collectively, the counties in our sample were home to 71 million people in 2010, or 23% of the nation's population. Figure 2 plots the 2010 populations of the census places in our sample, compared to all other census places. While our sample locations are modestly larger on average, the two distributions are fairly similar. Appendix Figure C.1 replicates this graph for towns with close (within the main RD bandwidth) elections, and it leaves the pattern substantively unchanged.

3.2 Other Data

We combine the county-seat election data on historical populations with modern population and economic data from the US Census. We base 2010 population figures on census-block data and historical town-level population data from Schmidt (2018). For modern income, we use Internal Revenue Service (IRS) statistics for 2010, reported at the zip code level. We draw census-block-level data on population and job characteristics from the 2010 LODES data and census figures.

For other historical data, we turn to IPUMS public census microdata from 1940 and

Figure 2: 2010 Populations of Sample Census Places



earlier and a geocoded panel data set on US post office locations throughout the nation from Blevins (2021). Elevation data are constructed from the SRTM data set, and stream shapefiles are from ESRI. Railroad data are from Atack (2016) and "U.S. National Transportation Atlas Railroads" (2015, distributed by ESRI).

3.3 Linking Data Sources

Some care is required to consistently link the economic data to the towns. In the large majority of cases (82% within close elections), the competing towns are US census-designated places in 2010. In earlier census years, these communities typically corresponded exactly to one or more enumeration districts. In these cases, we link to the corresponding census blocks or enumeration districts, making use of the Census Place Project for historical microdata (Berkes et al. 2023). For the remainder, we match to small subcounty units of the rural areas close to the original site: enumeration districts (census microdata), zip codes (IRS tax data), or 2010 census blocks. See Appendix Section A.14 for further details.

This linking procedure requires us to accurately code the locations of all abandoned locations. To do so, we extensively review historical county maps and histories to locate these towns and locate the original site in all cases within the RD bandwidth. The final baseline sample consists of 2130 unique locations, with 1935 geolocated.

3.4 Summary Statistics

Appendix Table C.1 presents summary statistics for the locations[3] within our unrestricted sample. The median election occurs in 1884, and 95% of elections fall between 1840 and

[3]Reflecting their weight in the regression, towns that appear in multiple elections are counted separately each time in the summary statistics.

1922. Because we geolocate all towns within our RD bandwidth, it is rare for relevant data to be missing, though it does occur when other data sources (for example, IRS filing data) omit values.

The modern statistics reflect a process of structural transformation as depicted in Appendix Figure C.2. On average, towns begin with a primarily agricultural economy that transitions into a range of blue- and white-collar occupations in subsequent decades. In contrast to the economies of most national capitals, local government and lawyer/judge positions do not account for a significant share of jobs. Even among winners in post-election years, these positions account for less than 1% of employment.[4] Along with other evidence, these statistics inform our choice to focus on the effects of density and agglomeration rather than multiplier effects from local government.

4 Econometric Approach

4.1 Regression Discontinuity

Our main identification strategy consists of RD analyses. In keeping with the best practices recommended in the literature (Calonico et al. 2014; Gelman and Imbens 2019), we use the following specification:

$$y_i = \alpha \mathbb{1}\left(\text{Win Election}\right)_i + f(\text{Victory Margin})_i + X_i\beta + \varepsilon_i \tag{1}$$

Here, f is a local linear function in vote percentages separately estimated for each side, X_i are controls, and ε_i are error terms clustered by election. By comparing towns that narrowly won to those that narrowly lost, α captures the causal effect of a county-seat election victory on agglomeration measures. Controls typically include county fixed effects and geographic controls (listed in Appendix Section A.12) that cover land quality, pre-election demographics, railroads, and other infrastructure. The pre-election restriction is necessary to avoid the use of endogenous controls.

For comparability, we use a single default bandwidth of 14.5 percentage points (pp) selected via the Calonico et al. (2014) procedure for the first-stage regression without controls. Appendix Section B.1 shows robustness to bandwidth choice, with bandwidths ranging from 5 to 20pp, none of which substantially change the estimates.

[4] See Appendix Section A.15 for definitions. In the census microdata, these figures are also increased because the data merge county- and municipal-government employment into the category of local government. Mechanically, the statistics on just county-government employment would be lower.

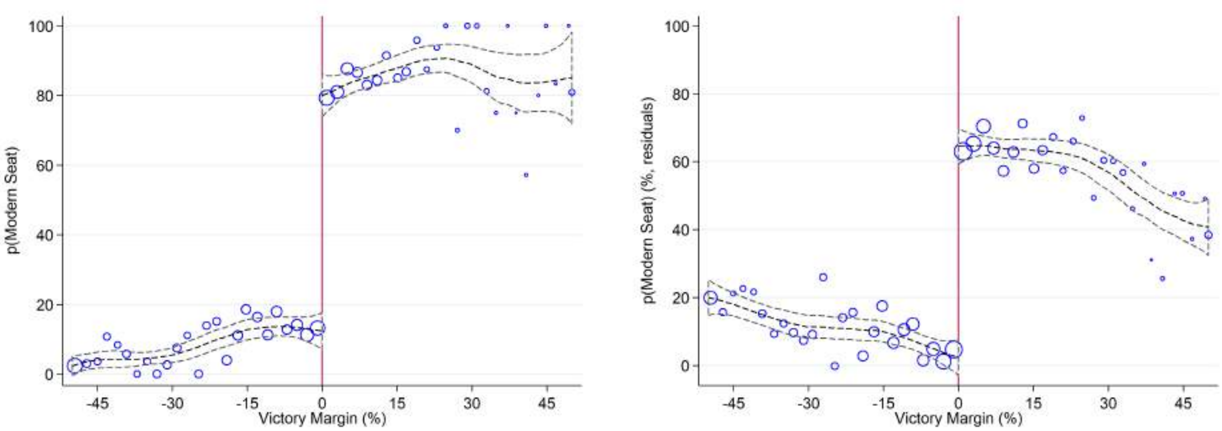
4.2 Sample Selection

Our baseline sample includes only the first election for each county. Although 80% of counties in our sample have just one recorded election, this selection procedure avoids potential endogenous selection for the remainder.[5] When we consider effects based on election timing in Section 6, we include all elections to maximize our sample for later periods.

4.3 First Stage and Balance

We first confirm that these elections actually determine county-seat choices and that winning is empirically balanced across predetermined characteristics. Figure 3 shows the RD plot and indicates that victory leads to a 67 percentage-point increase in probability of being the modern county seat. We should not expect a 100 percentage-point effect here, as subsequent elections or legislative action could ultimately overturn victories. Still, the estimate indicates that the majority of elections were final; consequently we present subsequent results as reduced-form impacts of victory.

Figure 3: Probability of Being Modern County Seat



(a) No Controls **(b)** All Controls

Notes: Binned scatter plot and local linear fit on the probability of being a county seat in the present day. Size of circles indicates number of observations in a bin. Panel (a) ignores controls, and Panel (b) residualizes the y-axis on county fixed effects and geographic controls. Geographic controls are listed in Section A.12. [no controls] RD Estimate: +67.56pp, (z = 16.32)*** [controls] RD Estimate: +66.30pp, (z = 20.08)***

Balance tests for a set of pre-election and geographic characteristics are shown in Figure

[5]Hypothetically, a motivated election loser could call for another election, leading to a sample imbalance in a set of all elections; focusing on a single election removes this possibility. The first election includes runoff rounds where applicable.

C.3. Hypothetically, imbalances across the victory threshold could have occurred if elections were manipulated or if our data-collection process was biased. However, in our tests the point estimates are typically small and none approach statistical significance; the highest magnitude among nine z-scores is 1.17, about what would be expected from chance alone.

4.3.1 Electoral Fraud and Manipulation

The statistical balance shown above makes it unlikely that fraud or electoral manipulation substantially affects our analyses. Violating the RD design's continuity assumption requires that one party has the ability to "precisely control" the vote share around the win/loss cutoff (Jeong and Shenoy 2022; Das 2023), either through fraud or exact knowledge of vote shares. For example, a ruling authority could add just enough votes to ensure a victory for their preferred choice. However, our balance checks show that politically protected incumbent towns are slightly less likely to win (not statistically significant).

Qualitative evidence is consistent with the idea that electoral manipulation was unlikely to affect the RD results. Precise knowledge of voting intentions would have been difficult given the period's technology and changing populations. While losing towns sometimes alleged fraud, these complaints typically referred to "unfair" electoral tactics or disputes over voter registration, neither of which is a method of precise control (Schellenberg 1987).[6]

5 Population, Amenities, and Agglomeration

In this section, we use estimates of equation (1) to analyze the effects of election victory on population and town amenities. We then analyze how changes in population density changed winning locations' economies.

5.1 Population Growth

The historical accident of a close county-seat election victory had large effects on modern populations. Figure 4 shows the impact of winning on a town's 2010 population using four metrics. Our preferred measure is 2010 population density within one mile of the town center

[6] For example, this cited work relates a story in which one town's voters were unable to go to the polling place "because they had found a barrel of whiskey that day . . . and many were too drunk" to make the journey. This sort of tactic or allowing unregistered voters to submit ballots could be undemocratic and shift votes, but it could not do so in a discontinuous manner. By analogy, modern candidates may increase their vote shares through campaigns and advertisements without affecting the validity of an RD. Even if candidates violate election regulations in doing so, vote differences will not show up exactly at the threshold of victory.

(Panel [a]). By this measure, winning towns increase in density by 1.17 log points[7] (223%) over their rivals—that is, slightly more than tripling their density. Since many of the towns in our sample are quite small, a one-mile radius is typical for their extent. Increasing the radius to two miles delivers a slightly smaller but still substantial increase of 0.92 log points (Panel [c]). The effect on towns' official-census populations is even larger at about 2.61 log points (1263%; Panel [b]). We prefer the radius-based estimate, as the census-based data may be influenced by the administrative decision to declare a location a town. Measuring populations as a percentage of the county total, a county-seat victory shifts about 11% of the population, a substantial figure (Panel [d]). While the previous measures could be influenced by large relative effects on small populations, this last measure shows that the population effect was important in absolute terms as well.[8] All metrics point to large and sustained population increases in the winning locations relative to the losing locations.

The results suggest that the county seats proved attractive to many early migrants and had large impacts on the early distribution of the population. Extrapolating from the result on county-population fraction, 5 million people were shifted simply by the chance occurrence of close elections within our sample and bandwidth. This figure and our previous estimates highlight that we are able to study the impact of relatively large shifts in the population distribution. The approximate tripling of population density also contrasts with the small number of jobs associated with county seats. As we noted in Section 2, only 0.24% of modern employment is connected to administrative or legal functions of county government. It is hard to imagine that such a small sector directly drove such a large shift in population. Instead, county-seat status likely served as a signal of the durability and prominence of a location, which attracted and thereby coordinated migrants; we return to this point in Section 8.

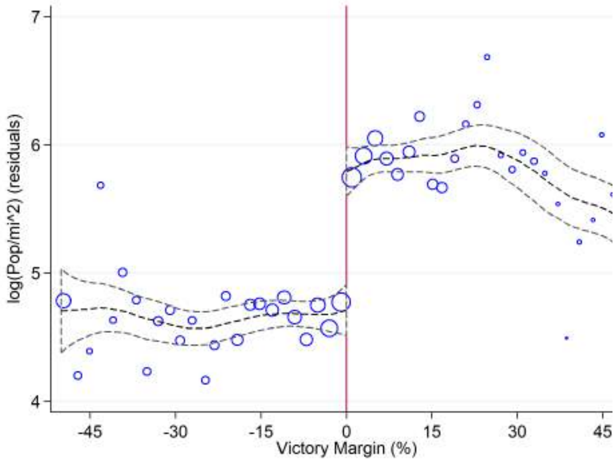
We can interpret the population result through the lens of models of city size and path dependence. Bleakley and Lin (2015) suggest that persistent long-run differences in city size can arise as a consequence of strong local agglomeration forces even if places have similar locational fundamentals. This notion is consistent with our results: winning and losing locations are balanced on geographic fundamentals and other pre-election characteristics but experience different population growth paths as the result of a close election victory

[7] For this and other outcomes that include zero in their support, *log* denotes $\log(\max(1, x))$—that is, we bottom-code pre-log values at 1. Among the solutions proposed for log-like transformations by Chen and Roth (2023), this is mathematically equivalent to their suggestion of manually calibrating the effect of a 0/1 change in which we place no weight on this transition. We view this calibration as appropriate given our focus on urban areas. A density of one person per square mile, for example, is extremely sparse and would not represent meaningful progress toward urbanization relative to even lower densities.

[8] For example, a town that grew from 5 to 50 people would show very large proportional growth, though the growth might not be important in absolute terms.

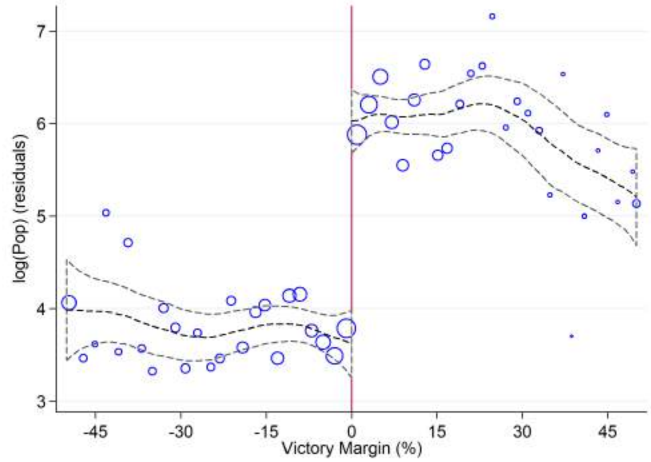
Figure 4: Impacts on Town Population

(a) log(Density) < 1 mile, 2010 census blocks



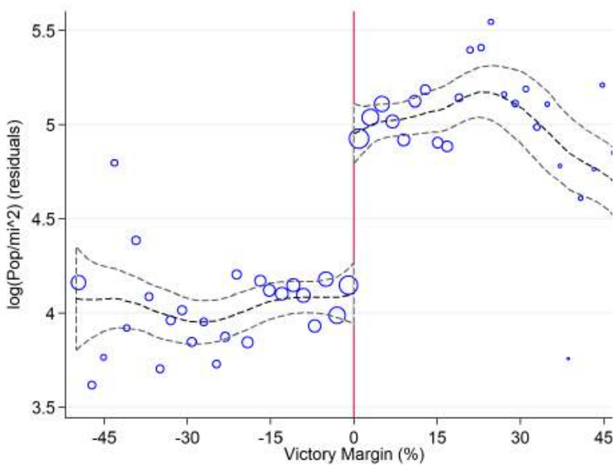
RD Estimate: +1.17, (z = 9.67)***

(b) log(Town-Census Population), 2010



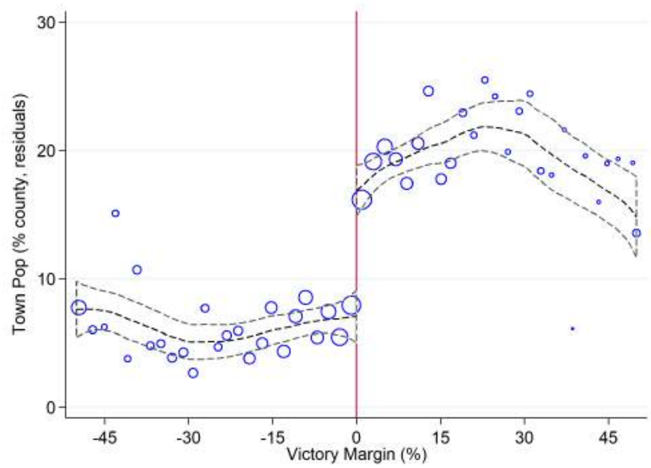
RD Estimate: +2.61, (z = 11.23)***

(c) log(Density) < 2 miles, 2010 census blocks



RD Estimate: +0.92, (z = 9.46)***

(d) Town-Census Population (% of County's)



RD Estimate: +10.86pp, (z = 8.14)***

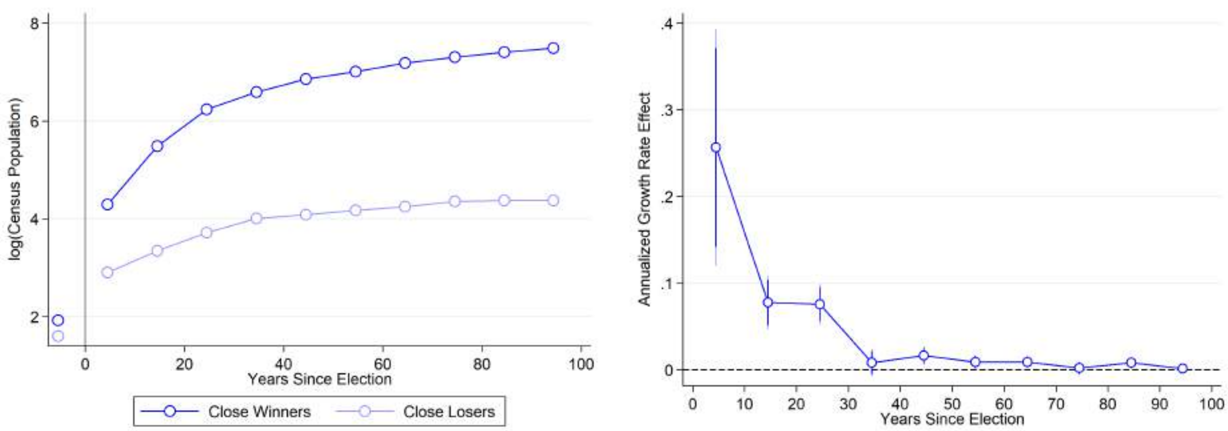
Notes: Binned scatter plot and local polynomial fit for multiple preperiod or geographic characteristics. RD estimates and z-scores are presented below as text. "Density" refers to population per square mile for 2010 census blocks within the specified radius. "Census population" refers to a town's population using the census-defined boundaries of the town. "Log" refers to $\log(\max(1, x))$. Geographic controls are listed in Section A.12.

that helps centralize a county's population. In Sections 5.2—5.3, we provide evidence that becoming the county seat led to immediate, considerable increases in population followed by the provision of amenities typically found in larger towns per se rather than amenities specific to capitals. We thus interpret county-seat elections as a case in which a shock selected one of many population equilibria and thereby shaped the long-run distribution of city size.

5.2 Timing of Population Effects

The population effects estimated above primarily reflect high rates of historical population growth in the winning locations. Figure 5 charts town-census population growth in competing towns in the years following elections. Panel (a) plots the average log population[9] of close winners and losers (within the RD bandwidth) over time. Both types of location start off with similarly low populations and both grow on average. However, their paths diverge after the election with close winners growing much more rapidly. Panel (b) uses RD estimates of equation (1) to measure the causal effects of post-election annual growth rates. Notably, the largest growth rates occur in the first decade, with an estimated 25% annual growth. The subsequent two decades increase this gap, albeit at a slower rate. Finally, there are either small or null effects in most of the subsequent decades.

Figure 5: Town-Census Population over Time



(a) log(Pop) Close Winners/Losers

(b) Annualized Growth Rates (RD)

The high growth rate in winning locations occurs because they attract a disproportionate share of counties' early migrants. As most elections were held close to the time of counties' creation, when population sizes were low, there was large potential for growth. Notably, there are few adjustments to the winner/loser population gap after the early decades of high growth. Accordingly, we primarily interpret the establishment of a county seat as a shock to early population size through migrants' historical choice of destination. Strong local agglomeration effects led to a rapid transition to the new and persistent spatial equilibrium.

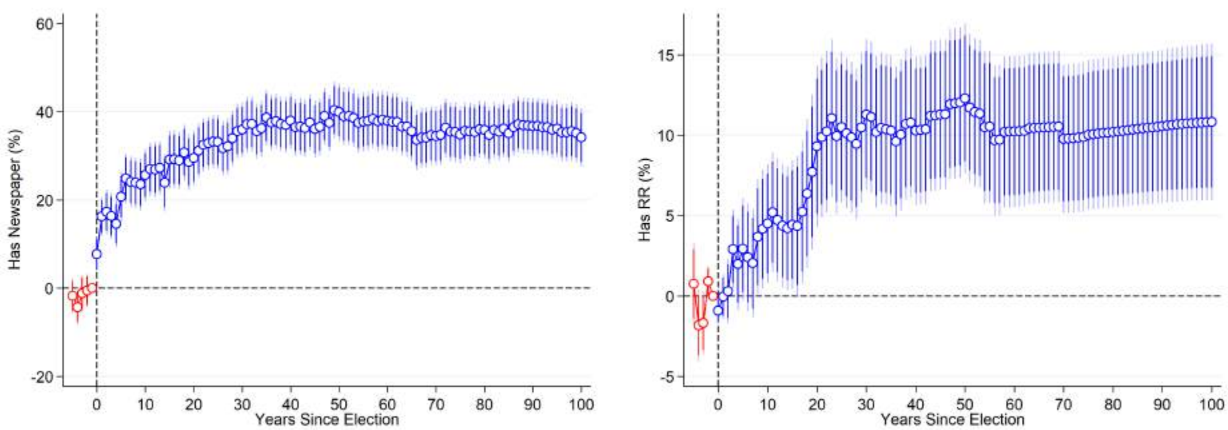
[9]Unfortunately, the density data we use for the modern period are unavailable historically. As noted earlier, we bottom-code populations at 1, so levels and possibly differences may reflect census recognition of their existence. However, the annual measures in Section 5.3 reflect town size but are not subject to the formalization difference.

The statistics here also counter a story of direct harm to the typical losing location. On average, losers gained rather than lost population, albeit at a lower rate than winners.[10]

5.3 Amenities

Amenity provision was likely a key reason early migrants coordinated on county seats as central cities. Other benefits highlighted by previous work on capital cities, such as government employment (Bai and Jia 2021; Chambru et al. 2021), are much more limited in the context of the small county governments we study.

Figure 6: Amenity Provision (annual data)



(a) Has Newspaper (%)

(b) Has Railroad (%)

Notes: Effects over time on the presence of a newspaper or railroad per Atack (2016). Each point summarizes an RD analysis of equation (1) a specified number of years after the election. Regressions include county fixed effects and controls listed in Section A.12.

Figure 6 shows that county seats rapidly gained amenities as a result of election victories. We focus on two we can observe on an annual basis: town newspapers and railroad access. Each point in the figure represents the result of an RD analysis of equation (1) and an outcome a specified number of years after the election. In both cases, the gains are fairly rapid, though railroads' arrival slightly lags the population increase and newspaper establishment; its most rapid rise begins around 20 years after elections. As with the town population effects in Section 5.2, the differences stabilize after 30 or so years. The rapid increase in amenities in the post-election years underlines the causal importance of county seats as influences on early-stage economic geography.

[10] Of course, this represents an average change, and some individual losing locations may have lost population as a result of the election rather than only having forgone an increase.

This section provides intuition for the coordination effect and population growth in winning locations. In the rural counties that represent the bulk of our sample, a number of important amenities were located in a handful of locations. Living in a county's largest town gave migrants the best chance of accessing these amenities. However, as discussed in Section 8.2, election winners do not appear to have more amenities conditional on their population, relative to losers. That is, county seats do not appear to provide more public goods than towns of similar size. The amenities in a county seat thus appear to be a result of its higher population rather than the location of the government per se, highlighting the importance of the agglomeration channel.

5.4 Agglomeration in the Long Run and Transition Dynamics

In the long run, the substantial growth in early population altered the economies of winning locations. Similar to what we found for amenities, despite their small size, county-seat election winners transition in ways similar to larger cities: toward higher-income, service-oriented, and innovation-driven economies. Using both census microdata from 1850 to 1940 and 2010 subcounty data, we show that these transformations occurred early in history and persisted.[11]

Across a range of measures, historical county-seat victories increase towns' income. Table 1 shows this increase using two sources: 1940 census wage and salary income, and aggregated 2010 IRS tax returns. Workers in winning towns report 18% higher wages in 1940; 2010 income in tax data increase by 8.0%. Because reporting areas and procedures differ between the 1940 and 2010 sources, we do not compare them to arrive at a measure of convergence; we assess convergence using comparable data in Figure 7. However, both indicate a sustained, economically meaningful increase in income in response to a county-seat victory. Appendix Table C.2 shows the robustness of this result to alternative measures of income and wealth in 1940 and 2010. Within-occupation wages, occupational income, home values, and white-collar employment all rise. Table 1 further shows that these income gains occur as part of a transition to an education-based, white-collar economy. Years of education (in 1940) and the fraction of workers with a bachelor's degree (in 2010) both rise. Per capita patents, as measured in Berkes et al. (2023) and linked to census microdata, rise by approximately 43% of the close-loser sample mean.[12]

The speed of the economic transition closely mirrors those of the population and amenities

[11] See Section 3.3 for details on linking and Section 3 for sources.

[12] This effect is not driven mechanically by winning locations' higher populations or transport connections. We link inventors to the census using only name and county information and ignore the filing location. These rates are in per capita terms and so mechanically adjust for population size in that respect. See Appendix Section A.14.5.

Table 1: Effects on Income and Proxies

	1940				2010	
	(1)	(2)	(3)	(4)	(5)	(6)
	log(Wages)	log(Wages) (non-gov)	Ed. Years	Patents / 1000	log(Income)	Bachelor's Worker (%)
Win	0.18***	0.16***	0.45***	0.022***	0.080***	1.52***
	(0.020)	(0.019)	(0.044)	(0.0041)	(0.019)	(0.39)
County FEs	Y	Y	Y	Y	Y	Y
Geo	Y	Y	Y	Y	Y	Y
SEs / Clusters	Election	Election	Election	Election	Election	Election
BW (pp)	14.5	14.5	14.5	14.5	14.5	14.5
N	1044	1044	1044	1044	1020	990
N (clusters)	538	538	538	538	537	541
$E[y]$	\$730	\$721	8.6	.06	\$48,375	16%

Notes: RD estimates of impacts of county-seat victory on income and proxies. Columns (1)–(4) use census microdata from 1940 and respectively examine the log of wages and salary income; log of wages and salary income excluding government and legal workers; years of education; annual per capita patents (per Berkes (2018)) through 1950. Column (5) examines zip-code-level income in IRS tax data. Column (6) examines the fraction of jobs with workers with bachelor's degrees in 2010 LODES. Geographic controls are listed in Section A.12. See Section A.14.5 for patent linking and Section A.14 for linking generally.

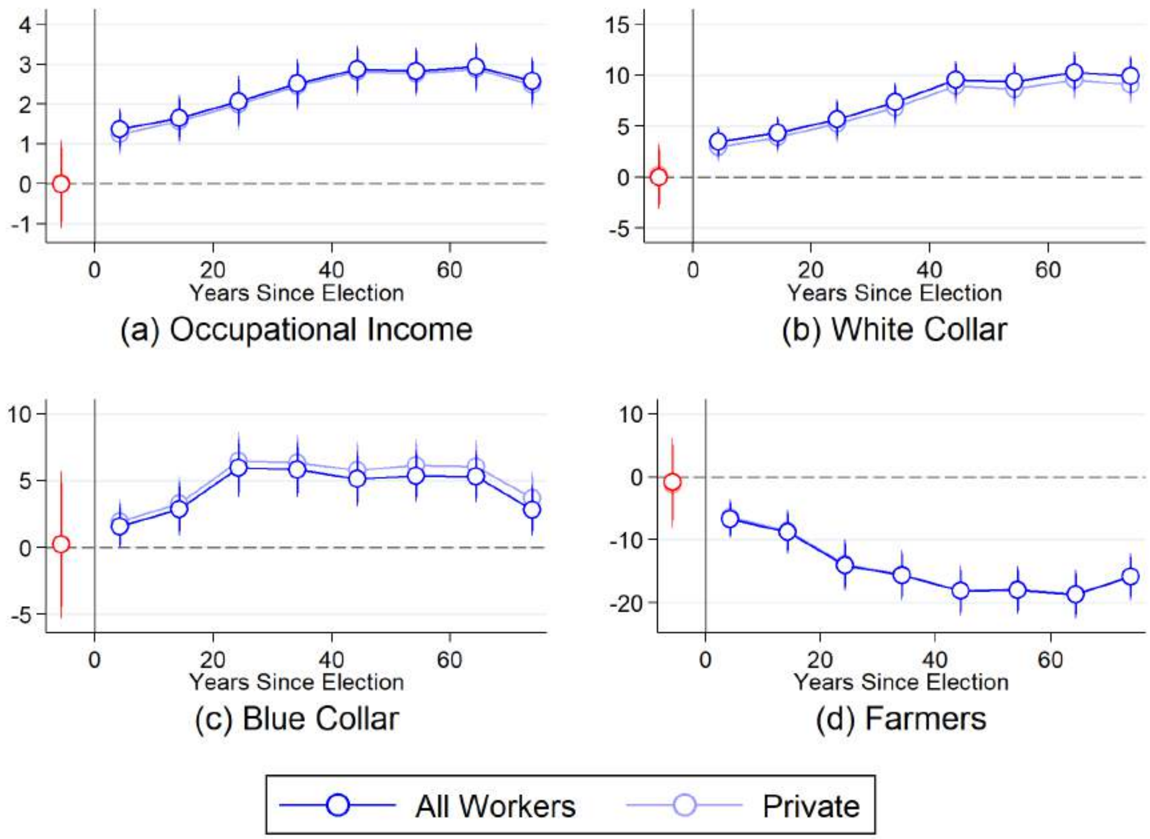
estimated in Section 5.3. Figure 7 depicts the economic transformations over time according to decadal census data. Each data point represents an RD estimate of equation (1) in years relative to the election, averaged in decadal bins.[13] Winning towns experience a structural transformation, shifting jobs out of agriculture and into a broader range of blue- and white-collar occupations. Occupational-income scores, the best proxy for income in these data, similarly increase by a maximum of 2.5 points. This represents about 23% of the standard deviation in 1850, or an increase of about 13% relative to the towns' pre-election averages, similar in size to the direct income measures presented in Table 1.

Very little of the economic adjustment shown above is tied to the public sector. If we exclude public workers from the occupational changes in Figure 7 and the wage estimates in Table 1, the estimates remain essentially unchanged.[14] Appendix Figure C.4 shows a similar result by breaking down the 1940 occupational changes by private versus public sector. The largest shifts are toward white-collar workers in the private sector, mainly at the expense of agriculture. Administrative employment in local government increases by a smaller amount

[13]For example, the first post-election data point averages results 0–9 years after the election, the second 10–19 years after, and so on.

[14]For Table 1, compare regressions (1) and (2). (2) drops lawyers and judges in the private sector whose jobs might be tied to the courthouse.

Figure 7: Occupation and Income Effects (census)



Notes: This figure plots RD estimates for town-level statistics over time. Each data point represents an RD estimate of equation (1) in years relative to the election, averaged in decadal bins. Occupation data including income scores are computed based on census microdata. Each panel contains two sets of estimates: one for all workers and one for all workers except those who work in public administration.

from a base of less than 1% of jobs. As we discuss in Section 8.1, the private sector growth cannot be directly explained by the public employment without multipliers more than an order of magnitude higher than those found in the literature.

In short, historical population-density growth in these primarily small locations precipitated the same changes we associate with density in larger urban areas. These changes stem from adjustments in the first decades after election victory and become permanent, indicating path dependence. Given the importance of this coordination to towns' long-run development, we next turn to the question of under what conditions coordination could occur.

6 Shock Timing

We turn to the question of when the county-seat shocks had the greatest effects on long-run economic geography. Such critical junctures may have occurred early on if initial conditions were most important in shaping long-run outcomes before population equilibria become fixed. Alternatively, they may have occurred contemporaneously with later events such as the arrival of railroads or the adoption of certain technologies. To test these theories, we take advantage of the critical fact that we observe county-seat elections at many points in time. Empirically, we rerun our RD estimates from Section 5.4 and compare effects across elections that occur increasingly later. To keep time on a consistent scale across areas that were settled in different periods, we categorize elections based on the years elapsed since county creation. An election in the year a county was created occurs in year 0, which can be interpreted as an initial condition.

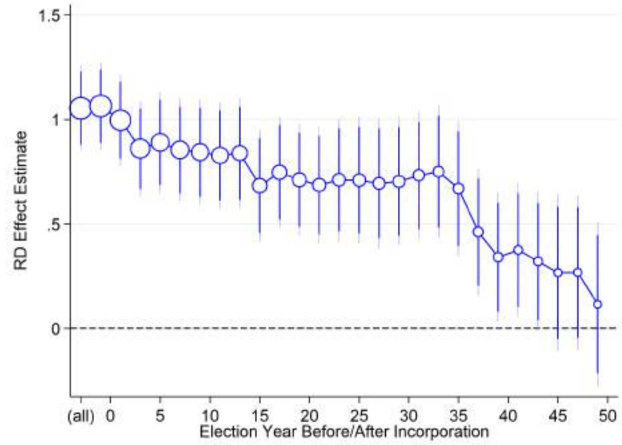
Figure 8 indicates the importance of initial conditions for the county-seat shock. Each point represents an RD estimate of equation (1), where the sample consists of all elections that occur at least a certain number of years after county creation;[15] moving from left to right, we can compare effects across elections in earlier versus later years. Effects on population (Panel [a]) begin close to our preferred estimates in Section 5.4 and attenuate to small, statistically insignificant values after about four decades. Estimated effects on income (Panels [c] and [d]) move in a similar pattern, beginning with large increases and attenuating to zero after several decades.[16] Appendix Figure C.5 produces a similar analysis for incumbent county seats. Compared to non-incumbents that get elected, incumbents that get reelected experience much smaller effects that become statistically insignificant after only about five years. This small effect compared to incumbents that lost elections narrowly is consistent with the notion that the benefits of county-seat status accrue rapidly and are locked in.

Notably, this attenuation is not caused by a lower chance of becoming the seat of county government: if anything, later elections are slightly more likely to determine the modern capital (Panel [b]). This observation also cuts against the notion that our effects are due to the county government per se, rather than the large migration observed only in early periods. If county governments themselves drove economic growth, we would expect to see

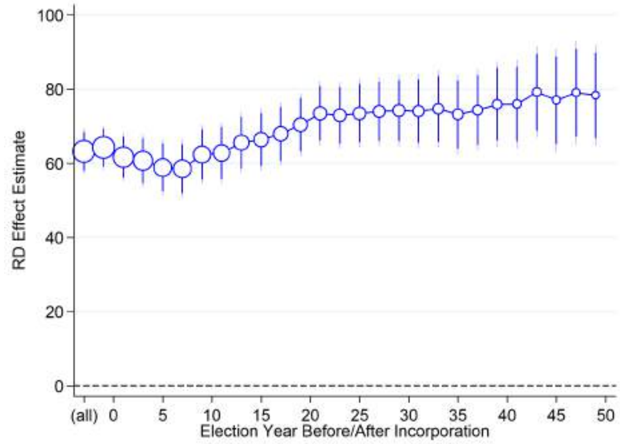
[15] As noted in Section 3, to maximize our sample of non-initial elections, we include all elections in our database for this analysis regardless of whether we observe a previous county-seat election.

[16] The pattern for later years in Panel (d) is not monotonic, with a seeming fall and rise to borderline significance in later years. However, it is hard to interpret this as a consistent pattern given the number of outcomes and sample years we consider. Instead, in the economic measures behind Panels (a), (c), and (d), only the early years consistently give positive effects, and Panel (d) has a less clear attenuation than Panels (a) and (c).

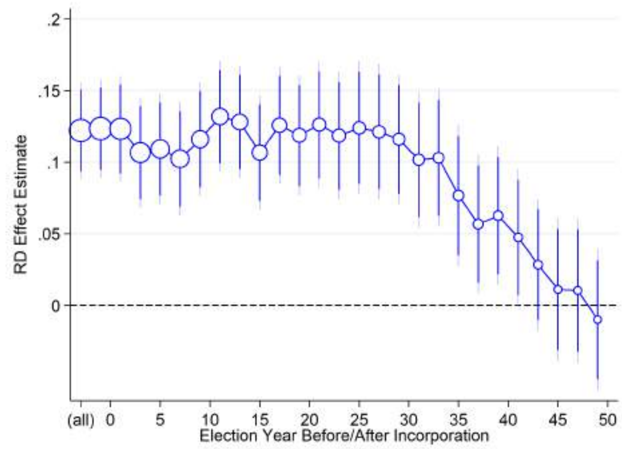
Figure 8: Effect Size Relative to County Creation



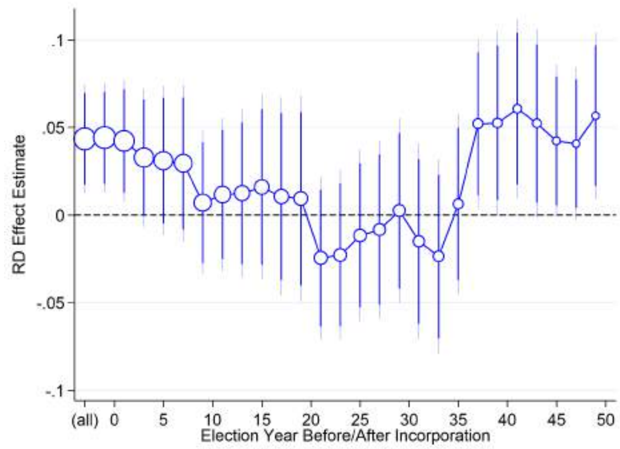
(a) (log) Population Density within 1 Mile



(b) Is Modern Seat (%)



(c) (log) 1940 Wages and Salaries, Private



(d) (log) Adjusted Gross Income (IRS zip code level)

Notes: RD estimates of log population density of census blocks within one mile of town centers (Panel [a]), the identity of the modern seat (Panel [b]), (log) 1940 reported wages and salaries for nongovernmental employees (Panel [c]), and log adjusted gross income at the zip code level (Panel [d]). Each point represents a regression on a subset of elections based on their year relative to county incorporation. The first bin analyzes the full sample, and subsequent ones restrict the sample to elections at least a specified number of years after county creation. Confidence intervals of 90% (thick line) and 95% (thin line) are shown over each estimate. Controls include the geographic characteristics listed in Section 4 and county fixed effects.

the strongest results in later periods, as they are more definitive for determining the modern seat.

These results suggest a theory for why previous research finds heterogeneous long-term effects from historical shocks. Early in an area's history, multiple spatial equilibria are possible, and shocks may have large impacts because they affect equilibrium selection. In later years, population distributions become increasingly fixed through local returns to scale that lead to agglomeration effects and durable capital investments (see Section 5.3), so these shocks matter less. Other research has explored shocks in a variety of settings, some of which are in initial time periods and others much later; the effect sizes could depend critically on this timing.[17] Our context allows us to explore the role of timing and this theory fits both the quantitative and qualitative evidence. In the early years of a county, populations were more fluid: towns had only short histories, and new migrants were constantly arriving. The value of county-seat status was highest in this period. In later years, towns were sufficiently established that relocating the seat did little more than move a small number of jobs in county administration.

This pattern also further indicates that county government per se is relatively unimportant as an explanation for the economic adjustments we observe. Relatively late elections consistently changed the county seat without many observable economic or population changes. If county-government spending or employment drove the results, the effects should have been strongest here.

7 Effect Size and Scale

We now consider two possible ways in which the long-term impacts of a county-seat shock scale up: the size of the affected community and the geographic scale on which we measure outcomes. Regarding the first way, the benefits of agglomeration might require a certain minimum scale, meaning we would see small effects for very small communities. Conversely, agglomeration could apply broadly. Regarding the second way, we might similarly expect to find smaller or larger effects at extended geographic scales. We specifically consider whether the choice of seat could influence county-level outcomes.

7.1 Machine Learning Approach

To address questions of heterogeneity, we use a recently developed machine learning package known as generalized random forests (GRF) (Athey et al. 2019). This methodology extends

[17] As examples, Bleakley and Lin (2012); Ager et al. (2020) find large roles for geography-based initial conditions in shaping economic geography, but Becker et al. (2021) finds only "modest" private sector growth from a "substantial" shift in public employment due to the West German capital shifting to Bonn. Davis and Weinstein (2002) finds only temporary impacts from large-scale bombing of Japanese cities in WWII, but suggests that initial conditions may be quite important given the range of densities dating from the Stone Age.

the classic random forest approach (which predicts fixed variables) to the prediction of heterogeneous treatment effects. Given a set of inputs and the RD equation (1), the algorithm attempts to predict which potential county seat would deliver the greatest impact. We use both this functionality and the standard (fixed variable) prediction.

Throughout, we use the predetermined controls listed in Section A.12 as predictors. We only estimate out-of-sample predictions for each location, clustered at the (pre-election) county level. That is, only data from outside a particular county are used to estimate predictions for towns within that county.

7.2 Agglomeration across the Distribution of Community Size

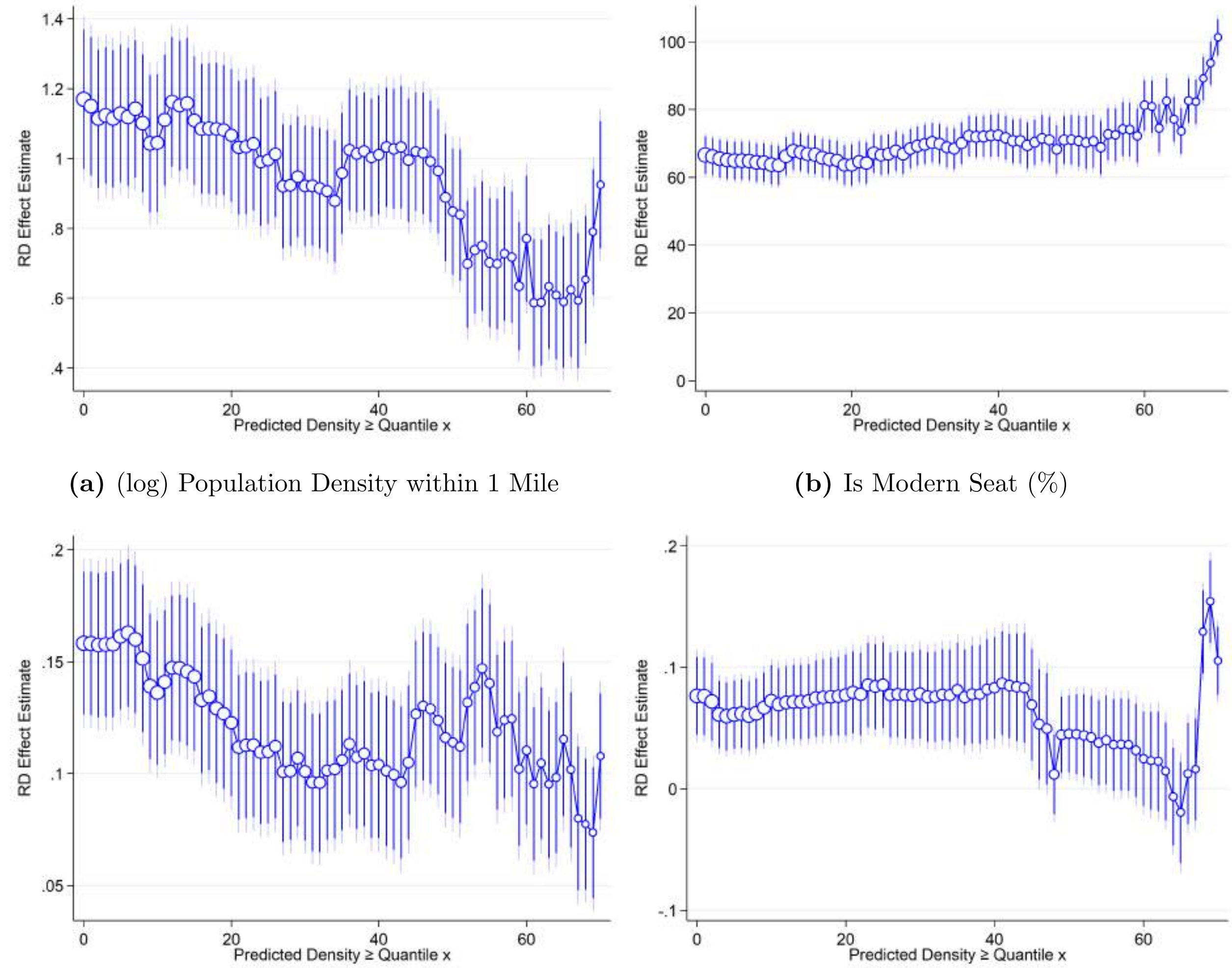
We begin by testing whether larger or smaller communities benefited most from the population influx following a county-seat election victory. In theory, either small or large communities could have been more affected. If agglomeration economies are limited for communities below a specific size, we would expect small benefits for the smallest communities. On the other hand, towns of all sizes might benefit similarly from increases in density.

Since density is endogenous to our treatment, we test these theories using predictions of density from pretreatment characteristics. Appendix Figure C.7 depicts the fit between actual density and the out-of-sample predictions from a random forest model. Below, we use these predictions to select subsamples of locations, focusing on those with higher quantiles of (predicted) density. We then rerun equation (1) for towns of increasingly larger (predicted) size.

Figure 9 shows the results of this heterogeneity analysis and illustrates that, if anything, smaller towns receive greater benefits from county-seat shocks. In Panel (a), the predicted 2010 density change drops from 1.17 to around 0.5 for towns at and above the 60th percentile of predicted density. Similarly, the effect on (log) private wages in 1940 drops from 0.16 to around 0.08 and 2010 zip-code-level income from 0.08 to 0.05. These declines are not due to weakening in the first stage of remaining the county seat, which shows a slight positive trend.

These results indicate there is not a minimum scale for the benefits of density to matter. We are wary of computing an elasticity, but the effects on income appear proportional to those for density across the (predicted) size distribution. Smaller communities also appear to have the largest relative benefits from the shock, perhaps reflecting the more limited set of attractions to migrants in those towns. At first glance, both points are positive news for place-based policies targeted toward smaller communities. Smaller communities can certainly benefit from density, and shocks of a fixed size may affect them relatively more. Our results

Figure 9: Effect Size by Predicted Density



(a) (log) Population Density within 1 Mile

(b) Is Modern Seat (%)

(c) (log) 1940 Wages and Salaries, Private

(d) (log) Adjusted Gross Income (IRS zip code level)

Notes: RD estimates of log population density of census blocks within one mile of town centers (Panel [a]), the identity of the modern seat (Panel [b]), (log) 1940 reported wages and salaries for nongovernmental employees (Panel [c]), and log adjusted gross income at the zip code level (Panel [d]). Each point represents a regression on a subset of towns based on the quantile of their predicted density, as described in Section 7.2. The first bin analyzes the full sample, and subsequent ones restrict the sample to towns in higher quantiles. Confidence intervals of 90% (thick line) and 95% (thin line) are shown over each estimate. Controls include the geographic characteristics listed in Section 4 and county fixed effects.

from Section 6, however, are less positive in this respect. While small communities can benefit from returns to scale, the benefits crucially depend on their ability to cause a shift in the spatial equilibrium, and this becomes much harder over time.

7.3 Geographic Scale and Aggregate Effects

We now explore whether the county-seat decision shaped outcomes at the county level. The impacts on the entire county of choosing one town instead of another are less clear since naming a seat was almost always a zero-sum game.[18] However, some locations may have been better suited for the population influx caused by county-seat status, and the county as a whole would have benefited when they were selected.

Methodologically, we use the GRF method to estimate treatment effects on county-level outcomes. Considering only the top two towns within an election by vote margin, we define the ideal choice as the one predicted at the time of the election to maximize the modern (2010) density within the original county borders.[19] We then test the machine learning model's predictions by rerunning equation (1) at the county level. With one observation per election, we define the running variable as the margin of this ideal location. We necessarily use county-level equivalents of the town-level controls[20] and choose a new optimal bandwidth; see Appendix B.2. Because of the two-step nature of the procedure, we present bootstrapped estimate distributions in Appendix Section B.3. Overall, this procedure translates place-level shocks into county-level ones.

Table 2 shows that the model's density-maximizing choice increased both county density and income. Relative to the town-level predictions, these regressions have much less power, with about 400 counties inside the bandwidth. With this caveat, the RD analysis indicates that the predicted ideal seat would have increased county density by 51% and income by 5.1% , using our preferred estimates with all controls given in columns (3) and (6). Bootstrapping these estimates with 500 trials using county-level clustering as described in Appendix Section B.3 respectively gives p=0.012 for density and p=0.076 for income; we interpret the latter result as marginally significant.

Two things are notable about these results. First, the aggregate effects of the county-seat choice are not zero, particularly for density and plausibly for county income. The density result confirms the relevance of the GRF model's choices and our method of translating a place-level shock into a county-level one. The income result suggests that county-level density remains relevant for economic outcomes, as with the town-level one. An important caveat is that with this methodology, we cannot disentangle whether these impacts are driven by larger spatial spillovers that certain towns have from whether some towns themselves have features that enable them to grow more.

Second, the coefficients estimated here are notably smaller than the town-level results

[18] A very small number of counties divide their functions across multiple seats, but this is exceptional.

[19] We construct this using GIS files of historical county boundaries and of 2010 census blocks.

[20] For example, county fixed effects cannot be included here; town-level variables have no sensible meaning.

Table 2: Effects on County-Level Outcomes, 2010

	log(Pop/mi^2)			log(Income), IRS		
	(1)	(2)	(3)	(4)	(5)	(6)
Best Wins	0.84***	0.63***	0.51***	0.083**	0.073***	0.051**
	(0.30)	(0.15)	(0.14)	(0.042)	(0.023)	(0.022)
State FEs		Y	Y		Y	Y
State × x,y		Y	Y		Y	Y
County Geo			Y			Y
SEs / Clusters	Robust	Robust	Robust	Robust	Robust	Robust
BW (pp)	9.25	9.25	9.25	9.25	9.25	9.25
N	387	387	387	387	387	387
$E[y]$	159/mi^2	159/mi^2	159/mi^2	\$48,776	\$48,776	\$48,776

Notes: RD estimates of impacts of the ideal county-seat victory on (pre-election) county outcomes in 2010, as described in Section 7.1. Values are constructed based on 2010 census blocks and zip code polygons, assuming a constant population density within each boundary. Columns (1)–(3) consider (log) population density per square mile. Columns (4)–(6) consider (log) average income, defined as average filer adjusted gross income. County geographic controls are listed in Appendix B.2. Bootstrapped estimates and significance are shown in Appendix Section B.3.

shown in Table 1 and Figure 4. The effects of a shock to one particular place are attenuated at a broader scale, even if they do not shrink to zero. That there is variation at all indicates that voters' goals or knowledge diverged from the optimizations performed by our machine learning algorithm. In some respects, this result should not be surprising. Aggregate outcomes over a century after the election were largely irrelevant to most voters. Most tended to vote for their own community (if possible) or the one with which they had the most affinity (if not). We discuss the divergence between the algorithm's and voters' preferences more in Appendix Section B.4.

8 Alternate Mechanisms

The most obvious alternative to our story of long-run agglomeration centers on the direct effects of county government. In this section, we first consider effects of growth directly due to government administrative employment (Bai and Jia 2021) and then consider the effects of a county capital's ability to create higher-quality public goods (Bluhm et al. 2021; Campante and Do 2014).

8.1 Government Sector

A priori, the small size of the county-government sector means its growth is unlikely to drive the population and economic changes we document. Figure 4 shows that population density triples as a result of an election victory. Yet only about one in 400 jobs nationally is connected to county administrative and legal functions. Similarly, Figure 7 shows that occupational trends are unchanged by excluding public-sector workers.

If public employment accounted for the job growth we see, the implied multiplier on public hiring would be exceptionally large. In 1940, about 1% of employment in close winners consisted of the sorts of local-government legal or administrative jobs that are potentially connected to the county seat; for close losers, the figure is around 0.4%. Given the 223% increase in density, a multiplier of approximately 75 would be required to explain this change.[21] But Faggio and Overman (2014) find no public multiplier and Becker et al. (2021) find one of about 0.7. Focusing on the private sector, Moretti (2010) finds a multiplier of 1.6 for the private, nontradable sector. It is unlikely county government has a multiplier of 50 times the largest of these. Instead, the population growth seems to reflect the special role of the county seat in agglomerating early migrants, discussed in Sections 2 and 5.1.

As a final check, we compare our main effects in states where county government plays a larger role to those where it has a smaller role. We measure this by counting the number of major county officers (for example, sheriff or clerk) as listed in Murphy (2009) and dividing states into above and below median. Appendix Table C.3 shows that states with larger roles for county government have similar effect sizes to those with smaller roles, with some point estimates being larger and others smaller. If county-government functions were an important mechanism, we should see distinctly larger effects in states where county government is more important.

8.2 Public Goods and Service Provision

Another potential explanation for our results is that being the county seat enabled a town to obtain a disproportionate share of public goods, similar to the role of national capitals

[21] Density increases by roughly 223% and the fraction of local-government administrative/legal jobs by 0.62 percentage points from a base of 0.4%. This gives an estimate of $2.23/\left((1+2.23)\times(0.4+0.62)\%-0.4\%\right)-1 \approx 75$. This is an underestimate, as we cannot distinguish county from municipal employment in these data; this fact increases the denominator. Here, we approximate the total increase in jobs with the 2010 increase in density shown in Figure 4 and occupational change as shown in Figure C.4. While these estimates come from different years, Figure 5 shows that population differences stabilize rapidly, indicating it is a good approximation. Since census-enumeration-district areas for the 1940 population are not standardized or recorded, we prefer the standardized density comparison available in 2010.

discussed in Ades and Glaeser (1995). In two analyses, we show that this is unlikely, examining both continuously measurable and lumpy public goods. Our results suggest that winning locations are essentially normal for their size and look like other large towns, similar to what we found in Section 5. We interpret this as additional support for the coordinating role of county seats.

8.2.1 Continuous Measures of Public Goods

Table 3: School-District Characteristics

	(1) log Students	(2) log Exp / Student	(3) log Rev / Student	(4) Teachers / Student
Win	0.17***	0.024	-0.0049	-0.0026***
	(0.047)	(0.015)	(0.011)	(0.00080)
County FEs	Y	Y	Y	Y
Geo	Y	Y	Y	Y
SEs / Clusters	Election	Election	Election	Election
BW (pp)	14.5	14.5	14.5	14.5
N	1049	1049	1049	1048
N (clusters)	541	541	541	541
$E[y]$	4,698	$14,812	$14,872	.075

Notes: RD estimates of impacts of county-seat victory on locations' school districts. Column (1) reports on log students. Column (2) reports on log expenditure per student. Column (3) reports on log revenue per student from the following sources: total, federal, state, local, and property taxes. Column (4) reports on teachers per student. Columns (2)–(3) are top-coded at the 1% level to remove outliers with few students. Column (4) is bottom-coded at the 1% level to account for districts with zero listed teachers. Geographic controls as listed in Section A.12.

We begin by examining schools—a major focus of most county governments and a case in which a public good scales naturally with population. We turn to Common Core school-district data for the 2017–18 school year for outcomes.[22] Table 3 shows the results. Unsurprisingly, an election victory increases the number of students in a district, though by less than the population of the winning town (column [1]).[23] Regarding the money available for students, the results suggest there is little difference. The point estimates show that winning results in a 2.4% expenditure increase per student but also 0.5% lower revenue, neither of which is statistically significant. Finally, column (4) suggests a small but statistically significant decrease in teachers per student, with about one fewer teacher per 400 students, or

[22] The earliest for which all key data are available.

[23] This gap likely reflects the fact that low-enrollment school districts were merged to achieve administrative scale.

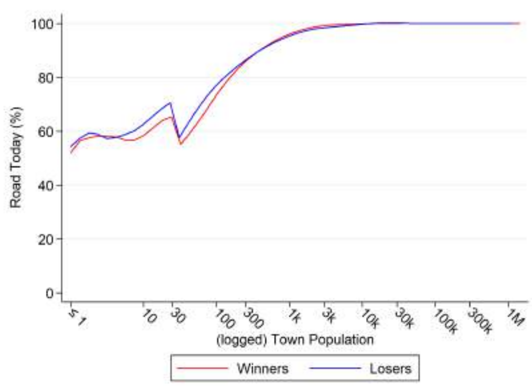
3% of the sample mean. This last fact may represent a real loss to schools in a district or simply reflect the minimum scale of needing one teacher for very small classrooms.

Taken together, the results cut against the notion that migrants were attracted to the county seat to access better public goods. If anything, a small drop in the teacher-to-student ratio suggests a slight decrease in school quality.

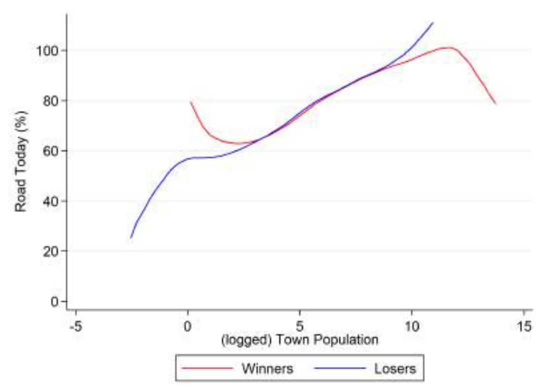
8.2.2 Lumpy Public Goods

Figure 10: Public Goods versus Population, Election Winners and Losers

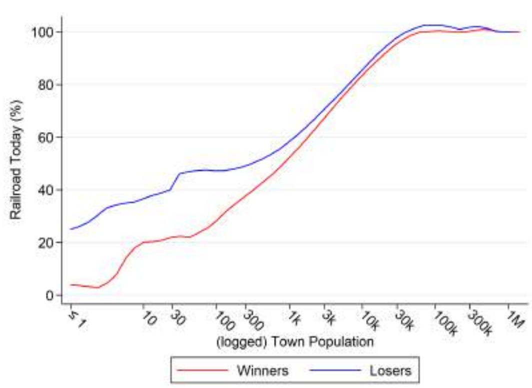
(a) Roads (unconditional)



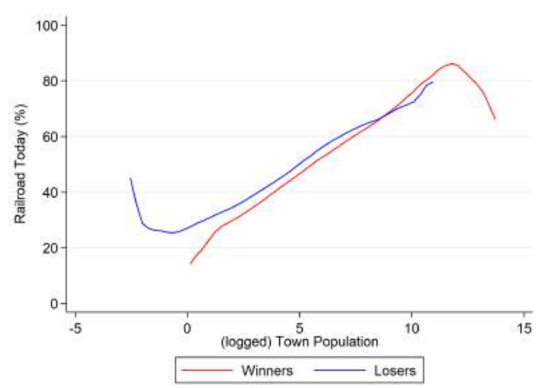
(b) Roads (residualized)



(c) Railroads (unconditional)



(d) Railroads (residualized)



Notes: Relationship between various measures of public goods and population for close (within the default RD bandwidth) election winners and losers. Each line represents a local linear regression of the presence of the feature on log formal town population. The outcomes are respectively a road or railroad in the modern census-designated place or within one mile of the location. Panels (a) and (c) present the unconditional relationship, while Panels (b) and (d) show the relationship on residuals after regressing on the geographic characteristics listed in Section A.12 and county fixed effects.

We also argue there are no notable differences in the presence of discrete public goods, adjusted for population. We study the presence of modern roads, railroads, and post offices.

Figure 10 plots the relationship for the presence of the public good in 2010 with respect to log town population for close[24] election winners and losers in our main sample. Similar relationships hold both unconditionally (Panels [a] and [c]) and when the full sample is residualized on county fixed effects and geographic controls (Panels [b] and [d]).[25] Conditional on town population, close winners are slightly less likely to have a railroad, about as likely to have a road connection, and slightly more likely to have a post office (see Appendix Figure C.9). The relationship thus shows little evidence that county seats received a disproportionate share of public goods.

Because we are studying a single historical shock, there are important limits to the conditional analyses we can conduct. Small towns are naturally much less likely than large towns to have public goods such as railroad connections, meaning that the large population increase from an election victory mechanically produces higher provision in an absolute sense, even absent favoritism. However, Figure 10 conditions on a downstream control, which can bias estimates. But given that the relationship of public goods and population is so similar among close winners and losers, such biases would have to roughly cancel those from favoritism to explain our results. We thus consider these analyses a supporting evidence against a direct role for county government, in concert with the other evidence presented in this section.

9 Conclusion

This paper uses the historical shock of close county-seat elections to study when such elections have long-term effects on towns and larger areas. County-seat elections determined the location of county government and gave winning towns a reputational boost that served to coordinate the population when new cities and towns were forming. Using an RD design, we showed that county-seat selection catalyzed a structural transformation in winning locations that durably increased their population and income. Election victories precipitated rapid population and economic changes that were mostly completed within a few decades and remained persistent thereafter.

We establish several important and novel facts about the temporal and spatial dynamics of this shock. First, the shock matters most in the early periods of frontier settlement, before a population equilibrium is fully settled. Shocks that occur several decades after a county has existed have little effect. This implies that population distributions become increasingly

[24] Within the default RD bandwidth.

[25] Both the outcome and (logged) town population are residualized. For ease of visualization, we shift the residuals by a constant such that the mean of each variable among election losers is preserved.

fixed over time and our shock has diminishing effects on long-run outcomes. Second, smaller places experience relatively larger effects from the county-seat shock. This result suggests that economic change is easier in the smaller, rural locations that place-based policies often target as long as a change in spatial equilibrium can be induced. Further, even small towns benefit from density in similar ways to large cities. Using out-of-sample machine learning predictions, we also demonstrated that the choice of county seat was not a zero-sum game at the county level: some towns' status as county seat increases county-level income and density more in the long run than others' status does.

Our results are consistent with theories in which human behavior plays an important role in selecting among multiple steady states in economic geography. In our setting, the political act of selecting a county seat had a large centralizing effect on the population despite the limited changes it effected in employment and spending. Geographic and other fixed characteristics are balanced in our RD design, showing that locational fundamentals do not wholly determine the distribution of town sizes, though they clearly remain important. Our results also show that the extent to which such equilibria can be shifted varies greatly across contexts. Most notably, population distributions become less sensitive to the same shock as initial conditions become more fixed over time.

While caution is necessary in considering the external validity of results in any particular setting, the analysis has several implications for economic theories of agglomeration. As noted in Section 1, previous literature found mixed results on the long-run effects of historical shocks on economic geography. One plausible explanation is that shocks matter more when they come at an early time or another critical juncture. Once populations are in an equilibrium, it becomes much harder to transition. Second, larger communities and geographic areas may be less sensitive to shocks of all sorts. Relative to the smallest towns in our sample, bigger cities and whole counties have a broader set of influences and advantages that make them less sensitive to any particular event. Thus, the study of large-scale economic shocks may require settings like ours in which researchers individually track small-scale communities over time.

References

Daron Acemoglu, Jacob Moscona, and James A Robinson. State capacity and American technology: evidence from the nineteenth century. *American Economic Review*, 106(5): 61–67, 2016.

Alberto F Ades and Edward L Glaeser. Trade and circuses: explaining urban giants. *The*

Quarterly Journal of Economics, 110(1):195–227, 1995.

Philipp Ager, Katherine Eriksson, Casper Worm Hansen, and Lars Lønstrup. How the 1906 san francisco earthquake shaped economic activity in the american west. *Explorations in Economic History*, 77:101342, 2020.

Gabriel M Ahlfeldt and Elisabetta Pietrostefani. The economic effects of density: A synthesis. *Journal of Urban Economics*, 111:93–107, 2019.

Gabriel M. Ahlfeldt, Stephen R. Redding, Daniel M. Sturm, and Nikolaus Wolf. The economics of density: Evidence from the Berlin Wall. *Econometrica*, 83(6):2127–2189, 2015.

Keith A Albright. Using subcounty population estimates as controls in weighting for the american community survey, 2011.

Treb Allen and Dave Donaldson. Persistence and path dependence in the spatial economy. Technical report, National Bureau of Economic Research, 2020.

Jeremy Atack. Historical geographic information systems (gis) database of u.s. railroads. Technical report, 2016.

Susan Athey, Julie Tibshirani, and Stefan Wager. Generalized random forests. *Forthcoming in the Annals of Statistics*, 47(2), 2019.

Ying Bai and Ruixue Jia. The economic consequences of political hierarchy: evidence from regime changes in china, 1000-2000 ce. *The Review of Economics and Statistics*, pages 1–45, 2021.

H Spencer Banzhaf and Randall P Walsh. Segregation and tiebout sorting: The link between place-based investments and neighborhood tipping. *Journal of Urban Economics*, 74:83–98, 2013.

Timothy J. Bartik. Using place-based jobs policies to help distressed communities. *Journal of Economic Perspectives*, 34(3):99–127, 2020.

Nathaniel Baum-Snow and Ronni Pavan. Inequality and city size. *Review of Economics and Statistics*, 95(5):1535–1548, 2013.

Nathaniel Baum-Snow, Matthew Freedman, and Ronni Pavan. Why has urban inequality increased? *American Economic Journal: Applied Economics*, 10(4):1–42, 2018.

Sascha O Becker, Stephan Heblich, and Daniel M Sturm. The impact of public employment: evidence from bonn. *Journal of Urban Economics*, 122:103291, 2021.

Enrico Berkes. Comprehensive universe of us patents (cusp): data and facts. *UMBC Economics Department Collection*, 2018.

Enrico Berkes, Ezra Karger, and Peter Nencka. The census place project: A method for geolocating unstructured place names. *Explorations in Economic History*, 87:101477, 2023.

Hoyt Bleakley and Jeffrey Lin. Portage and path dependence. *The Quarterly Journal of Economics*, 127(2):587–644, 2012.

Hoyt Bleakley and Jeffrey Lin. History and the sizes of cities. *American Economic Review*, 105(5):558–563, 2015.

Hoyt Bleakley and Paul Rhode. Was free soil magic dirt? endowments versus institutions in the antebellum united states. 2023.

Hoyt Bleakley, Paul W Rhode, et al. De tocqueville, population movements, and revealed institutional preferences. *Journal of Historical Political Economy*, 3(2):179–210, 2023.

Cameron Blevins. *The Postal West*. Dissertation, 2015.

Cameron Blevins. *Paper Trails: The US Post and the Making of the American West*. Oxford University Press, 2021.

Richard Bluhm, Christian Lessmann, and Paul Schaudt. The political geography of cities. 2021.

J Brown and David Cuberes. The birth and persistence of cities: Evidence from oklahoma's first fifty years of urban growth, 2020.

Sebastian Calonico, Matias D Cattaneo, and Rocio Titiunik. Robust data-driven inference in the regression-discontinuity design. *The Stata Journal*, 14(4):909–946, 2014.

Filipe R Campante and Quoc-Anh Do. Isolated capital cities, accountability, and corruption: Evidence from us states. *American Economic Review*, 104(8):2456–81, 2014.

Census. 2012 census of governments: Finance—state and local government. 2012.

Cédric Chambru, Emeric Henry, and Benjamin Marx. The dynamic consequences of state-building: evidence from the french revolution. 2021.

Jiafeng Chen and Jonathan Roth. Logs with zeros? some problems and solutions. *arXiv preprint arXiv:2212.06080*, 2023.

Pierre-Philippe Combes, Gilles Duranton, and Laurent Gobillon. Spatial wage disparities: Sorting matters! *Journal of urban economics*, 63(2):723–742, 2008.

Pierre-Philippe Combes, Gilles Duranton, Laurent Gobillon, Diego Puga, and Sébastien Roux. The productivity advantages of large cities: Distinguishing agglomeration from firm selection. *Econometrica*, 80(6):2543—-2594, 2012.

Sabyasachi Das. Democratic backsliding in the world's largest democracy. *Available at SSRN 4512936*, 2023.

Charles G. Davis. Plainview, tx (hale county). *Handbook of Texas Online*, 1952. URL https://www.tshaonline.org/handbook/entries/plainview-tx-hale-county.

Donald R Davis and Jonathan I Dingel. A spatial knowledge economy. *American Economic Review*, 109(1):153–70, 2019.

Donald R. Davis and David E. Weinstein. Bones, bombs, and break points: The geography of economic activity. *American Economic Review*, 92(5):1269–1289, 2002.

Jorge De La Roca and Diego Puga. Learning by working in big cities. *The Review of Economic Studies*, 84(1):106–142, 2017.

Robert K DeArment. *Ballots and bullets: The bloody county seat wars of Kansas*. University of Oklahoma Press, 2006.

Melissa Dell and Benjamin A Olken. The development effects of the extractive colonial economy: The dutch cultivation system in java. *The Review of Economic Studies*, 87(1): 164–203, 2020.

Klaus Desmet and Esteban Rossi-Hansberg. Urban accounting and welfare. *American Economic Review*, 103(6):2296–2327, 2013.

Dave Donaldson and Richard Hornbeck. Railroads and American economic growth: A "market access" approach. *The Quarterly Journal of Economics*, 131(2):799–858, 2016.

Gilles Duranton and Diego Puga. Micro-foundations of urban agglomeration economies. In *Handbook of regional and urban economics*, volume 4, pages 2063–2117. Elsevier, 2004.

Maximilian v Ehrlich and Tobias Seidel. The persistent effects of place-based policy: Evidence from the west-german zonenrandgebiet. *American Economic Journal: Economic Policy*, 10(4):344–74, 2018.

Glenn Ellison and Edward L Glaeser. The geographic concentration of industry: does natural advantage explain agglomeration? *American Economic Review*, 89(2):311–316, 1999.

Giulia Faggio and Henry Overman. The effect of public sector employment on local labour markets. *Journal of urban economics*, 79:91–107, 2014.

Andrew Gelman and Guido Imbens. Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, 37(3):447–456, 2019.

Edward L Glaeser and Joshua D Gottlieb. The economics of place-making policies. Technical report, National Bureau of Economic Research, 2008.

Edward L Glaeser, Hedi D Kallal, Jose A Scheinkman, and Andrei Shleifer. Growth in cities. *Journal of political economy*, 100(6):1126–1152, 1992.

Michael Greenstone, Richard Hornbeck, and Enrico Moretti. Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings. *Journal of Political Economy*, 118(3):536–598, 2010.

Thomas Jefferson Gregory. *History of Yolo County, California*. Historic Record Company, 1913.

W Walker Hanlon and Stephan Heblich. History and urban economics. *National Bureau of Economic Research Working Paper Series*, (w27850), 2020.

W Walker Hanlon and Antonio Miscio. Agglomeration: A long-run panel data approach. *Journal of Urban Economics*, 99:1–14, 2017.

Mariaflavia Harari. Cities in bad shape: Urban geometry in india. *American Economic Review*, 110(8):2377–2421, 2020.

Stephan Heblich, Stephen J Redding, and Daniel M Sturm. The making of the modern metropolis: evidence from london. *The Quarterly Journal of Economics*, 135(4):2059–2133, 2020.

Richard Hornbeck and Suresh Naidu. When the levee breaks: Black migration and economic development in the American South. *American Economic Review*, 104(3):963–90, 2014.

Greg Howard and Arianna Ornaghi. Closing time: The local equilibrium effects of prohibition. *The Journal of Economic History*, 81(3):792–830, 2021.

Dahyeon Jeong and Ajay Shenoy. Can the party in power systematically win a majority in close legislative elections? evidence from us state assemblies. *The Journal of Politics*, 84 (2):1149–1164, 2022.

Benny Kleinman, Ernest Liu, and Stephen J Redding. Dynamic spatial general equilibrium. *Econometrica*, 91(2):385–424, 2023.

Patrick Kline and Enrico Moretti. Local economic development, agglomeration economies, and the big push: 100 years of evidence from the Tennessee Valley authority. *The Quarterly Journal of Economics*, 129(1):275—-331, 2014.

Oliver Knight. Toward an understanding of the western town. *The Western Historical Quarterly*, 4(1):27–42, 1973.

Paul Krugman. History versus expectations. *The Quarterly Journal of Economics*, 106(2): 651–667, 1991.

Jeffrey Lin and Ferdinand Rauch. What future for history dependence in spatial economics? *Regional Science and Urban Economics*, 94:103628, 2022.

Geoffrey P Mawn. Promoters, speculators, and the selection of the phoenix townsite. *Arizona and the West*, 19(3):207–224, 1977.

Guy Michaels and Ferdinand Rauch. Resetting the urban network: 117–2012. *The Economic Journal*, 128(608):378–412, 2018.

Enrico Moretti. Local multipliers. *American Economic Review*, 100(2):373–377, 2010.

Kathryn Murphy. County government structure: A state by state report. *A Publication of the Reseach Division of the National Association of County's County Services Department. Washington, DC March*, 2009.

Dávid Krisztián Nagy. Hinterlands, city formation and growth: Evidence from the u.s. westward expansion. *Working Paper*, 2020.

Stanley W Paher. *Significant County Seat Controversies in the State of Nevada*. PhD thesis, University of Nevada, Reno, 1969.

Diego Puga. The magnitude and causes of agglomeration economies. *Journal of Regional Science*, 50(1):203–2019, 2010.

Franklin Qian and Rose Tan. The effects of high-skilled firm entry on incumbent residents. 2021.

Stephen J Redding, Daniel M Sturm, and Nikolaus Wolf. History and industry location: evidence from german airports. *Review of Economics and Statistics*, 93(3):814–831, 2011.

Stuart S. Rosenthal and William C. Strange. Evidence on the nature and sources of agglomeration economies. *Handbook of urban and regional economics*, 4:2119—-2172, 2004.

James A Schellenberg. *Conflict between communities: American county seat wars*. Pwpa Books, 1987.

Benjamin Schmidt. Creating data: The invention of information in the nineteenth century american state. *http://creatingdata.us*, 2018.

Allison Shertzer, Tate Twinam, and Randall P Walsh. Zoning and the economic geography of cities. *Journal of Urban Economics*, 105:20–39, 2018.

Cory Smith. *Land Concentration and Long-Run Development in the Frontier United States*. 2023.

When is Long-Run Agglomeration Possible?
Evidence from County-Seat Wars

by Cory Smith ⓡ Amrita Kulka

ONLINE APPENDIX

A Data Sources and Variable Construction

A.1 Election Margin of Victory

We compile an original database of county-seat elections from a range of sources including historical newspaper articles, individual county histories, and administrative data. In each case, we record the number of votes received by each location and the required vote percentage for victory, if applicable. We code elections into three categories: automatic runoff/required-majority elections, first past the post elections, and elections with preset super-majority requirements (usually required of a challenger to an incumbent). We code these electorally favored towns as "protected incumbents."[26]

In some cases, our sources only report the absolute vote margin of the winning town, for example "Location A wins over Location B by 112 votes." In cases where exactly two locations compete, we impute the number of total votes based on county population. We log-linearly interpolate county populations based on census decades and, for our sample of elections where total votes are reported, we calculate the median ratio of votes for the top two locations to population, approximately 23%. For elections where only the margin is recorded, we use this ratio to impute the total number of votes. These two pieces of information allow us to construct vote percentages. In the most common case of majority/first past the post[27] elections, we use the margin to mean the gap between winner and loser. In the above example, if we imputed 300 total votes, Town A would have 206 votes and Town B would have 94. In the case of elections with fixed vote percentages, we interpret the margin as the number of votes the largest vote-getter surpassed (or failed to surpass) the required margin. We base these interpretations on examples where we see different sources report both margins and exact vote tallies. In the case where the election In cases where three or

[26]In other parts of the paper, we use incumbent to mean the preexisting county seat. These two can differ: not all incumbents received different treatment in the election. A small number of non-incumbents were also favored when laws required supermajorities to choose a location, for example, farther from the county center. Because these were ex-ante restrictions imposed before votes were cast, they do not bear on the RD's continuity assumption.

[27]For the two-town case, these are mathematically equivalent

more locations competed, we consider the information incomplete and do not enter it into the sample.

In the case where 50% or another fixed percentage of the vote is required, we compute each town's vote percentage minus the applicable threshold as the vote percentage margin. For example, a town receiving 56% of the vote in a majority-required election would have a value of 6%. If it required a 60% supermajority to win, it would have a value of -4%. In a small number of cases, multiple locations ran against an incumbent and a specific vote percentage/supermajority was required; in the case that none of the challengers achieved this threshold, the incumbent would remain the seat. In this instance, we code the challengers' margins as before and the incumbent's margin as the negative of the top challenger's.[28] In the case of automatic runoff elections, we code each round as a separate election where applicable.

In first past the post elections, we initially compute the margin as the location's vote percentage of the location minus the runnerup's vote percentage (for the winner) or the location's vote percentage minus the winner's (for the losers). To achieve consistency with the automatic runoff systems, we then divide these margins by two.[29]

We finally top- and bottom-code margins at +50pp and -50pp respectively, largely for visual clarity of the RD graphs. These reflect the maximum and minimum possible margins for simple majority votes and only affects a small number of observations with other election types. None of these elections would be included within the RD bandwidth under either their new or restricted value.

A.2 Town Population and Location

Our primary source for town location and population is Schmidt (2018) which aggregates data on town population and location deriving from the US census. However, some of our locations do not appear in this data set, either because of an oversight in including census data or because the census never recognized the location as a community. In the former case, we supplement the main Schmidt (2018) data by merging it with data directly from the census including location, modern, and historical populations.

For locations never listed in the census as towns, we attempt to locate them based on historical maps and descriptions of the county. Because linkage failures could be differential

[28] For example, imagine A, B, and C receive 50%, 40%, and 10% of the vote. C is the incumbent and challengers require 60% of the vote. Then the vote margins would be -10%, -20%, and 10% respectively.

[29] Consider the case of a two-location race where majority-required and first past the post elections are equivalent. Say the locations receive 65% and 35% of the vote. Comparing the winner's vote percentage to the required 50% would produce a margin of 15%. Comparing the winner's percentage to the loser's would produce a margin of 30%. Dividing the latter value by two restores the equivalence.

based on election victory, we ensure that we locate every competing location within 15 points of election victory (in either direction).

For locations not recorded in a census decade, we code the town population as 0[30] and interpret this value as the "formal" town population. Because informal and unincorporated communities still exist, we thus supplement this measure with boundary-neutral definitions of population, for example census block population density within 1 mile of the town center.

A.3 Zip-level Income Data

For location income data, we turn to the IRS Zip-level aggregates of income tax returns. Because these derive from administrative data, they have several advantages over their survey-based alternatives like the American Community Survey (ACS). Many of the communities we study are comprised of fewer than 1000 people, meaning that the sample of respondents even in the 5-year ACS would be small. Because some locations we study also no longer exist as census places (CDPs), we would also require comparability between subcounty/minor civil divisions (MCDs) and CDP within the ACS. In general, research on the ACS shows that its accuracy falls for smaller areas, posing a significant issue for its use in our study (Albright 2011). In contrast, administrative sources have a claim to either minimal sampling variation (even a small area with 1000 people is large statistically) or none (because it reflects the universe of an appropriately defined set of people).

We link each location to its Zip code via a polygon sourced from ESRI. By default, we consider the IRS definition of "adjusted gross income" (AGI) as our measure of income.

A.3.1 Zip-level Income Inequality

IRS data also have the advantage of providing AGI in predefined brackets: under \$25k, \$25-50k, \$50-75k, \$75-100k, \$100-200k, and over \$200k. This differentiation allows us to approximately compute top income shares to measure inequality.

We estimate top income shares via the following procedure by approximating income as constant within each of the six brackets (and equal to the reported average). We then report income shares among specified top percentages of filers. This method offers simplicity, though it likely understates inequality via the approximating assumption.

[30] Note that, when taking logarithms, this values are <u>not</u> coded to missing as we typically use a $\log(\max(1, x))$ functional form; see Appendix Section A.13.

A.4 LODES

For data on job characteristics, we turn to the 2010 Census LODES data set. These data report information on 2-digit industry codes, approximate wage terciles, and education requirements of jobs at the census block level (with added noise).

For this project, we select the version of LODES based on worker residence as opposed to the job location. There are two primary reasons for this. First, city and town sizes are typically computed based on residence and not the workforce; for consistency, it makes sense to discuss residents' characteristics. Second, smaller communities may have workforces that work in other communities, meaning that their values could be missing — confusing the analysis. Appendix Section A.14.3 discusses how locations are linked to census blocks.

A.5 School District Spending

Data on school district boundaries and revenue/spending come from the National Center for Education Statistics (NCES) for the year 2017, the earliest publicly available for all our relevant variables.

A.6 Railroads and Other Transportation

Historical railroad data are from Atack (2016) and the "National Transportation Atlas Railroads" (modern, distributed by ESRI). Because the former database only covers years up until 1911 and the latter database only covers the modern period, we adopt the following approach when analyzing railroad construction in a panel setting. In years prior to the election or prior to 1911, we exclusively use the Atack (2016) database and code railroads as existing based on the "in operation by/InOpBy" field and the year of data. For years after 1911 but still prior to the relevant election, we consider only railroads existing in 1911. Although this definition might attenuate results by ignoring track built in the intervening years, we are reluctant to add 21^{st}-century information; this would invalidate the use of this outcome as a pre-election placebo variable. Finally, for years following both 1911 and the election year, we linearly interpolate the existence of a railroad from the 1911 Atack (2016) file and the modern data, using 2010 as the endpoint for the latter.

For modern road data, we use the US Geological Survey (USGS) National Transportation Dataset (NTD) shapefile of major roads.

In the panel setting, we code the existence of railroads based on the distance between the location's coordinates and the nearest active section of track. We primarily consider the existence of railroads within one or two miles of the location's coordinates. For anaysis of modern characteristics, we code a location as having a road or railroad if either (a) it has a segment within 1 mile of its coordinates or (b) the census place (CDP) of which it is part intersects with any segment.

A.7 Post Offices

Geocoded data on US post offices, along with their dates of opening and closing, are available from Blevins (2015, 2021). In our analysis, we typically only count "active" post offices — ignoring any that closed prior to the specified year.

A.8 County Government

Aggregate employment in county government comes from Census (2012). We use these to produce aggregate statistics of county government jobs in particular roles (for example administrative/judicial) for 2012. When comparing to 2010 total jobs, we deflate the figures by the average national employment change from 2010 to 2012.[31]

A.9 Geographic Controls

Raw sources for geographic data are listed here: elevation data are from the SRTM data set, stream shapefiles are via ESRI, terrain slopes from the 2016 LANDFIRE data, and soil quality from the USDA gSSURGO data set. For soil quality, we use versions of the National Commodity Crop Productivity Index (NCCPI) computation to measure soil productivity and the frequency of flooding variables. The latter are broken into discrete categories based on flooding frequency.

For elevation, terrain slopes, flooding frequency variables (by category), and soil quality we compute area-weighted averages within 1 mile of the town center. For the NCCPI soil productivity measures, we code missing values as 0, reflecting the fact that these are usually water bodies or areas unsuitable for farming.[32] For streams, we compute the (log) distance to the nearest type of stream. For this purpose, we consider only natural bodies, ignoring human-made channels such as canals.

[31] Aggregated changes in county government employment would have been more ideal but are not available.

[32] Because we consider both the soil properties and the presence of soil types unsuitable for their calculations to be natural features unrelated to election characteristics, the imputation here should not create bias. As with any geographic characteristic in our sample, if the economic development shock caused by our election caused environmental changes or affected the observability of the characteristic, including it as a control would lead to the bias via inclusion of a "downstream" variable. However, Section 4.3 finds no empirical evidence of imbalance on this and other characteristics.

A.10 PLSS and Settlement Data

For other preperiod election controls we turn to locations' section (approximate square mile) grid square on the Public Lands Survey System (PLSS) and consider pre-election settlement characteristics as reported by the Bureau of Land Management (BLM)'s General Land Office (GLO) records.

From the PLSS itself, we consider the area of the section, whether it functioned as an education section, and whether the section number was even or odd. The latter two pieces of information reflect government land settlement policies of education and railroad land grants; see (Smith 2023) for historical details and relevance. From the GLO, we compute several variables reflecting government-administered settlement up to the year before the election (i.e. preperiod): the log total number of government land distributions ("patents"), the fraction of the section settled under the Homestead Act, and the fraction of the section settled according to "lots" rather than direction-based subsections (for example "quarter sections" of approximately 160 acres). This last variable is useful because agricultural land was typically divided on the basis of gridded subdivisions of sections (for example half, quarter, sixteenth... sections), land for urban areas was typically allotted based on more idiosyncratic "lots."

For areas outside the PLSS system (primarily Texas and New England), we code all these variables to 0.[33]

A.11 Newspaper Presence

The Library of Congress's *Chronicling America* directory includes approximately 157,000 newspapers along with their state and town and years active. We link this list of newspapers to our town sample and define a town as having a newspaper in a year if it has any linked paper active. Note that the directory is a larger sample than the searchable list of digitized newspapers.

A.12 List of Geographic Controls (Town-Level RD)

For the town-level implementation of RD equation (1), we include the following geographic controls: mean elevation, log distance to the (pre-election) county center, average terrain slopes ("ruggedness"), the presence of a (pre-election) post office, log (pre-election) town

[33] Because both pre-election settlement characteristics and the location with respect to the PLSS should be irrelevant for winning close elections, this imputation should not create any bias. If locating within or outside the PLSS was connected with election manipulation, then in theory this control could exacerbate bias. However, this is extremely unlikely as almost all counties would have been entirely within or outside of the PLSS; Section 4.3 also finds no empirical evidence of imbalance on this and other characteristics.

census population, "protected incumbent" status (a town competitors require a supermajority to beat, usually due to incumbency), miles to the nearest stream, average soil quality ("nccpi3all") and land flooding ("flodfq" categories) in the USDA gSSURGO database, and the presence of a (pre-election) railroad within 1 or 2 miles, and the (pre-election) presence of a newspaper. For locations surveyed in the public lands survey system (PLSS), we include the PLSS section area, even/odd status, status as an education section; see Smith (2023) for a discussion of their relevance. We include the log number of (pre-election) GLO settlement deeds ("patents") and the fraction of these given out under the Homestead Act (or 0% in the case of no patents). We include a binary for being within the PLSS grid, to obviate the importance of the value chosen when the settlement variables cannot be defined. Finally, we include variables defined in the pre-election census data: fraction of workers in agriculture, fraction racially white, fraction female, fraction of people living on farms, fraction of workers in blue collar occupations, average occupational income score. Where unavailable, we use the previous sets of control variables to predict them in an OLS regression.

For terrain and land quality data, we define the value as the average within one mile of the town center. Pre-election is defined as the earliest year available before the election year; in most cases, this is simply the previous year. For census data, it refers to the previous census.

A.13 Logarithms

When presented with fat-tailed variables that have zero in their support, researchers often search for a transformation that roughly matches the properties of the logarithm but gives nonmissing values for zero. By default, we adopt the functional form $\ln(\max(1, x))$ in these cases. This matches the Chen and Roth (2023) suggestion of manually specification and it places as little weight as possible on the 0/1 transition. We view this as appropriate for our variables. For example, a town with exactly 1 resident likely differs little from an area without a town. In most cases, the latter scenario corresponds to an area of sparse but nonzero population.

A.14 Linking

We herein discuss our procedures for linking our locations to different data sources. As detailed in Section 3.3, we have both the name and location centroid of all towns within the bandwidth to use in this process.

A.14.1 Census Microdata 1850-1940

Our linkage here builds off the Census Place Project (CPP) (Berkes et al. 2023) which we supplement with manual adjustments in cases of unclear links. For a large majority of locations, we find a "Census Place" in the above data that has a similar name (per a fuzzy string match) or is very close to the location. The town's characteristics then become those of the people recorded as living in the single or multiple enumeration districts.

In cases where no Census Place matches on either criterion, we adopt a procedure of manual correction. Using the original enumeration district maps in the National Archives,[34] we manually locate the coordinates of the town and link it to the listed district. In cases where these maps are unavailable, as a last resort we average across all enumeration districts that are not part of the CPP. This reflects a decision that in these cases the location would not have formed a sufficiently large town to be classified as a "place" in the CPP.

Because it is uncommon for our locations to exist more than one census year prior to the election, we restrict the final sample to one preperiod of data and subsequently all census years up to 1940.

A.14.2 Census Microdata for 1890

The absence of census microdata for 1890 impose several challenges for our linking exercise. We herein detail our procedures around this year, though we note these choices only affect a small part of the overall sample. Throughout, we maintain the validity of the RD design in pre-election census years. That is, we do not average or otherwise mix pretreatment and post-treatment years.

First, for election years 1891-1900, 1890 is a pretreatment year. We here use 1880 values in place of 1890 values — an earlier preperiod for all variables. For elections 1881-1890, 1890 is the first post-treatment year and we use 1900 values instead — a later post-period. When the election year is 1880 or earlier, all of 1880, 1890, and 1900 are post-treatment years. Here, we linearly interpolate all variables from 1880 to 1900 (i.e., average 1880 and 1900) for the 1890 value. Finally, there are no choices to make for election years after 1900: 1900 serves as the preperiod census and nowhere do we use even earlier preperiod data.

A.14.3 2010 Census Blocks

For the 2010 LODES data, we primarily link locations to census places (CDPs) on a block level using our linkage between our locations and census blocks. We use the residential

[34] We extend our thanks to Risa Wagner and Jake Burgess for excellent research assistance in this endeavor.

version of the LODES data set. For the minority of cases where no CDP exists, we use the workplace locations within 1 mile of the central location.

When linking locations to census blocks based on distance (for LODES, population, or other data), we assume a constant distribution of people (or jobs) within the area of each block. We further assume other characteristics (for example ethnicity, job wages...) are constant. Under these assumptions, we intersect a circle of the specified radius with all census blocks and compute the relevant totals (for example total population) and weighted averages from all intersected census blocks, based on the area of intersection.

A.14.4 2010 Zip Codes

We link each location to the 2010 Zip code in which it lies.

A.14.5 CUSP Patent Data to Census Microdata

In this section, we describe our procedure for linking the CUSP patent data to the census Berkes (2018). Our broad approach is to use a fuzzy string match on the first and last name of each inventor to census microdata within each county. We record the year of a patent as the year of issuance and link only within the county recorded by CUSP to the previous census within 10 years. The result is we code, in each year, the fraction of the population that is an "inventor": that has had a patent issued in a particular year. We link inventors to the most recent available census with microdata (i.e., 1850-1940 with 1890 as unavailable).

Several choices are notable here. First, we do not use the "city" information present in the CUSP data. Based on discussions with the authors and our own read of the data, "city" location may in some cases record a mailing address rather than a residential address. Given that not all small locales had post offices, treating the postal address as the true address would result in a mechanically positive effect in favor of the winners – we aimed to avoid inflating our results in this manner. Our read of the data is also that the issuance data of the patent is most complete relative to other years in the data, particularly for the earliest years. Empirically, it produces a higher match rate within the described counties relative to other listed years in the CUSP data.

A.15 Industry Definitions and Classifications

This section describes how we classify industries into "blue collar," "white collar," and farming.

A.15.1 Census Microdata 1850-1940

These definitions are based on the 1950 occupation codes. Farmers constitute codes 100-199 (farm owners, tenants, and managers) and 800-899 (farm laborers and foremen). "Blue collar" constitutes codes 500-699 (craftsmen and operatives) and 900-971 (nonfarm laborers). "White collar" constitutes codes 0-99 (professional and technical services) and 200-499 (managers, clerical, and sales).

We also consider several other specific industries and occupations in these data. Lawyers and judges are defined by occupation code 55. Local government covers industry code 936. Note that that county government specifically is a subset of this code which also would include municipal government positions. Non-local public administration covers codes 906, 916, 926, 946.

A.15.2 LODES Data 2010

These definitions are based on 2-digit NAICS industry codes that are further aggregated by the LODES data set; we do not observe occupation in these data. We define "blue collar" as codes 21-33 (mining, utilities, construction, manufacturing). "White collar" constitutes codes 42-92 (largely administrative, clerical, and service sectors). Informational/tech jobs refers to NAICS code 51 specifically.

B Robustness and Other Results

In this section, we demonstrate robustness to a range of sample definitions and coding decisions. Although Section 4.3 showed no detectable imbalances in geographic or pre-election characteristics, other arbitrary choices in specification could theoretically drive our results, diminishing their meaning.

B.1 Bandwidth Selection

Appendix Figure C.6 shows that bandwidth choice does not fundamentally change the results of two of our main variables: population density and (log) income. For the first outcome, decreasing the bandwidth from 20[35] to 5 modestly raises the effect size while leaving all results highly significant. For the second, decreasing the bandwidth again largely leaves the

[35] Note that a small number of communities are not geolocated outside of the 15-point bandwidth, meaning that their outcomes are generally missing. Since this missingness may be endogenous to the election result, we prefer bandwidths within 15 points but include up to 20 points for completeness.

point estimate unchanged, although with the smaller sample size at lower bandwidths, the estimates lose power and do not achieve statistical significance.

B.2 Sample, Optimal Bandwidth, and Controls

Our sample consists of the baseline ("first") elections with at least two competitors.[36] Within these, we focus on the top two competing locations by vote margin.

As with the town-level results in Section 4, we calculate the county-level optimal bandwidth using the no-controls version on county density.

For county-level controls, we adopt equivalents of the town-level controls listed in Appendix Section A.12. Most importantly, we can no longer include county fixed effects as they would essentially subsume the whole sample. In their place, we substitute a linear function of county latitude and longitude by state, i.e., state fixed effects and county latitude/longitude interacted with state. For the other controls, we include county-level versions of the town-level variables.

B.3 Bootstrapping and GRF Estimation

To determine the statistical significance of the GRF verification in Section 7.3, we bootstrap the primary estimates presented in Table 2, columns (3), (6). While we are unaware of theoretical work characterizing the bootstrap's properties for GRF models, to the best of our knowledge this represents the best method for calculating statistical uncertainty given the 2-step procedure in the aforementioned table.

We bootstrap the estimation in the following manner. For each trial, we construct a new sample by bootstrapping in blocks at the county level. We re-run the GRF prediction algorithm, considering each selected county as unique[37] and re-run the RD estimates of equation (1) based on those predictions. In both cases, the bandwidth does not change from the main estimation.[38] As a placebo test, we additionally calculate the RD estimates for randomly selected "ideal" options.

The results from 500 rounds of this procedure are graphed as histograms in Figure C.8. In recording statistical significance, we consider the fraction of estimates that fall below 0. Because we select the seat that the GRF model predicts will maximize density (and

[36] A very small number of elections in our sample were uncontested, meaning that the winning town has a victory margin of 50 percentage points (100% - 50%). While we technically include these in our sample, the margin prohibits them from entering RD estimates of equation (1).

[37] That is, if a single actual county is drawn twice, each selection is treated as though it were a separate county, albeit with the same characteristics.

[38] In a similar way, the standard errors of an RD typically do not change for a given bandwidth regardless of whether that bandwidth is manually or calculated as "optimal."

presumably income), we treat positive county-level RD estimates as verifications and negative ones as failures. By this metric, we obtain p=.012 for density and p=.076 for income. Unsurprisingly, about half of the placebo estimates are below 0: p=.47 and p=.52 for density and income respectively.

B.4 The Election Mechanism

Voters in the historical county-seat elections ultimately made their decisions based on different characteristics than those favored by our machine learning models. Figure B.1 panel (a) presents a graphs of towns' vote margins and the ML-predicted treatment effects on income. These correlations are fairly weak and usually negative, meaning voters tended to choose "worse" towns from the ML perspective. Similarly, panel (b) shows the probability that the better option is chosen in each election as a function of how close the election was. While the balance of predetermined variables at the vote threshold ensures that the town chosen by the ML model can only win close elections 50% of the time, in theory voters could have selected that town more frequently in less close elections. Instead, the better option is only chosen around 40% of the time for noncompetitive elections.

To illustrate why voters and the machine-learning models diverged, Table B.1 presents determinants of both the models' predictions and actual vote margins for selected pre-election controls. Although the coefficients can differ depending on the model and whether election-level fixed effects are included, several broad patterns stand out. First, the machine learning models de-prioritize productive agricultural land. Seats with dry, nonswampy land and good soil are predicted to decrease county-level outcomes. More centrally-located towns also have lower predicted effects, especially for density. However, these were not necessarily remote locations. Higher rates of "patents" (settlers), the presence of a post office, and the presence of a railroad all positively correlate with the ML predictions. Finally, towns that have high ex-ante predictions of income in 2010 are predicted to perform better for the county.

Voters mostly did not prioritize the same characteristics and instead tend to vote for the largest, most prominent towns. In large part, these correlations are simply the mechanical results of how voting operated in these contexts. According to historical sources, most voters support their own towns in elections, so bigger towns received more votes. State laws often protected incumbents, requiring a supermajority of votes to unseat them. Towns with newspapers, railroads, and central locations could also have been more prominent in voters' minds.

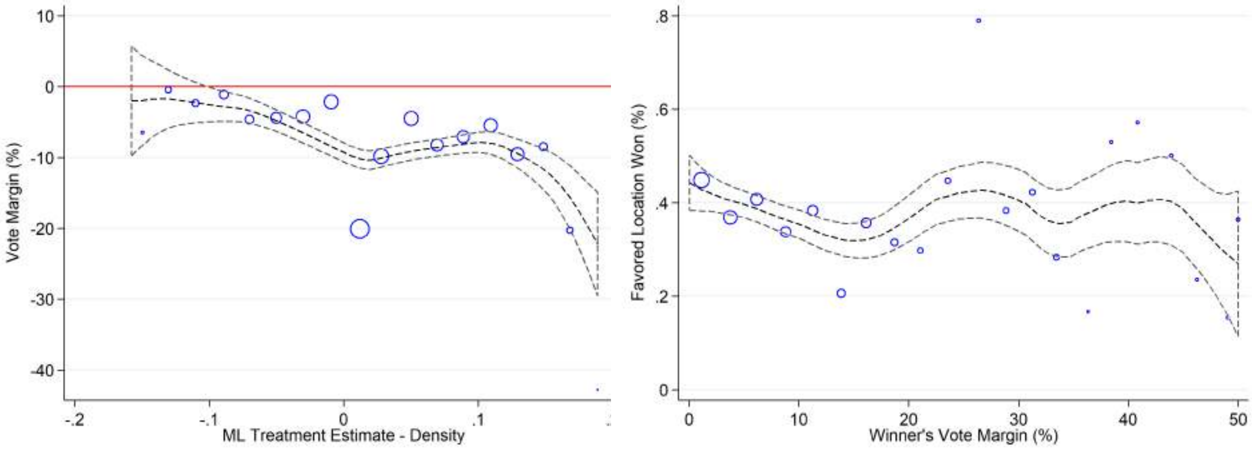
Overall, these results suggest that the machine learning models could prioritize change when voters selected the status quo. ML-favored locations were not ideal agricultural loca-

Table B.1: Determinants of Town Vote Margins

	ML: Density		ML: Income		Vote Margin	
	(1)	(2)	(3)	(4)	(5)	(6)
(log) Distance Center	0.031***	0.059***	-0.00021	0.000075	-0.0038*	-0.062***
	(0.0013)	(0.0024)	(0.00016)	(0.00034)	(0.0021)	(0.0096)
Non-swamp	-0.074***	-0.073***	-0.011***	-0.010***	-0.0057	-0.024
	(0.0061)	(0.0093)	(0.00096)	(0.0015)	(0.014)	(0.039)
Soil Quality	-0.061***	-0.050***	-0.0023**	-0.0069***	-0.0061	0.023
	(0.0064)	(0.014)	(0.00094)	(0.0022)	(0.011)	(0.058)
log(Land Patents)	0.0059***	0.0058***	0.00038**	0.00073**	-0.0070***	-0.0031
	(0.0011)	(0.0022)	(0.00016)	(0.00030)	(0.0023)	(0.0081)
Has PO	0.0097***	0.010**	0.00099**	0.0012*	-0.016*	-0.013
	(0.0033)	(0.0042)	(0.00048)	(0.00067)	(0.0088)	(0.018)
RR	0.0068**	0.0085*	-0.0020***	-0.00038	-0.0080	0.059**
	(0.0031)	(0.0051)	(0.00043)	(0.00079)	(0.0075)	(0.024)
Predicted $$, 2010						
Protected Incumbent	0.00037	0.0025	0.00083	0.00073	0.075***	0.16***
	(0.0041)	(0.0045)	(0.00066)	(0.00069)	(0.013)	(0.021)
log(Town Pop)	0.0013***	-0.0000076	0.0000086	-0.000042	0.0019*	0.016***
	(0.00048)	(0.00084)	(0.000066)	(0.00011)	(0.0012)	(0.0032)
Has Newspaper	-0.0011	-0.00061	-0.00051	0.000017	0.041***	0.11***
	(0.0028)	(0.0039)	(0.00040)	(0.00055)	(0.0075)	(0.016)
Election FEs		Y		Y		Y

Notes: OLS determinants of machine learning predictions for county-level density, county-level income, and town vote margins respectively; see Section 7.1. Predicted income in 2010 refers to a random forest prediction of zip-level income in 2010 IRS data, based on the controls listed in Section A.12. Protected Incumbent refers to cases where election regulations favored a particular location (for example, an incumbent) by requiring a supermajority to unseat it.

Figure B.1: Vote Margins and ML-Selected "Ideal" Seats



(a) Vote Margin and Density Effect

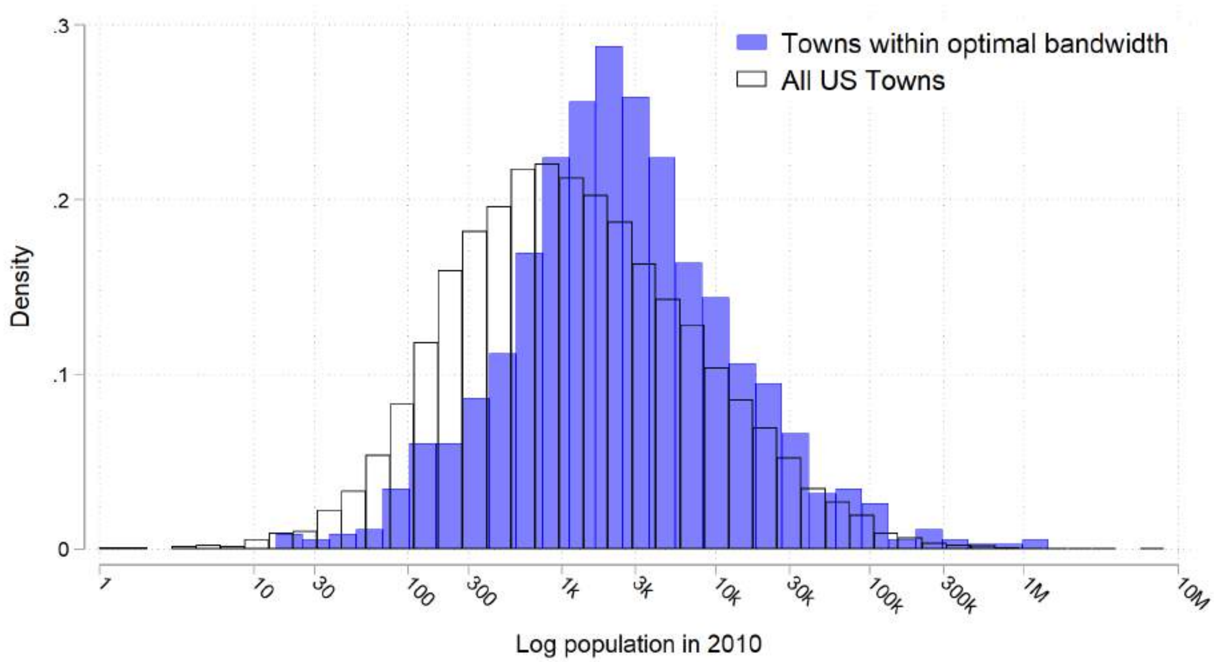
(b) p(Ideal Seat [Density]) and Margin

Notes: Correlations of vote margins and the "ideal" county seat, as chosen by the machine learning process described in Section 7.1. (a) shows the correlation of the predicted treatment effect and the vote margin. Vote margin is defined in the following way: For example, say Towns A and B have respectively a vote margin of +5pp and -5pp. Town B is considered to increase the county's population by more than Town A according to the algorithm. Then, the county's vote margin for picking the "best" seat is -5pp as the best seat was not chosen. (b) depicts the probability of the best option being chosen within a county as a function of how close the election is, defined as the absolute value of the vote margin for the town with the highest vote margin (equal to the winner's vote margin if a winner exists). In (a) the sample is the baseline set of elections. In (b), we further restrict the sample to elections with winners (no runoffs) and at least two contestants.

tions likely favored by the first settlers in a county, but they were places that were nonetheless growing and whose characteristics predicted higher incomes in the modern economy of 2010; these locations may not have been large at the time of the election, but they had the potential for growth. However, the politics of these elections meant larger, more established towns at the time of election were favored. Voters may well have been quite rational in making these choices: there is little reason they should have cared about the average characteristics of their county a century or more after their vote. As shown in Section 6, the benefits to a community from an election victory accrued relatively quickly and these were likely far more important for any individual voter. However, these results do illustrate that initial choices for counties' population centers could be locked in. Once a town was selected as a county seat and its population expanded, it would be heavily advantaged in future elections even if an alternative choice could have improved long-run economic outcomes.

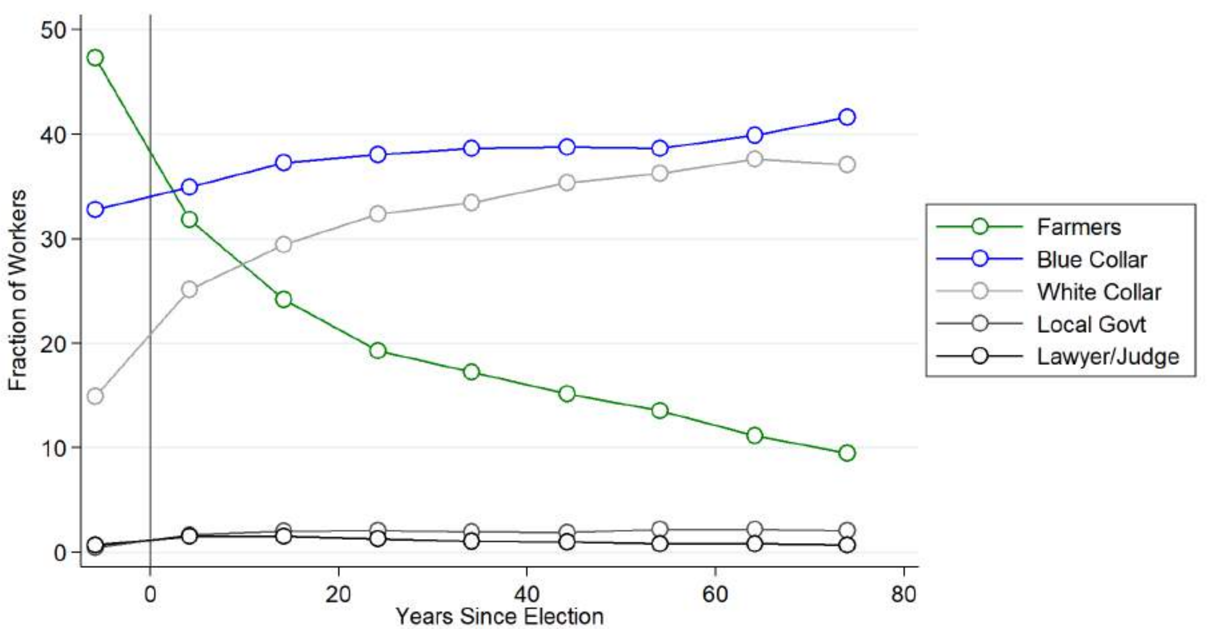
C Appendix Tables and Figures

Figure C.1: Modern Town Populations, Close Elections



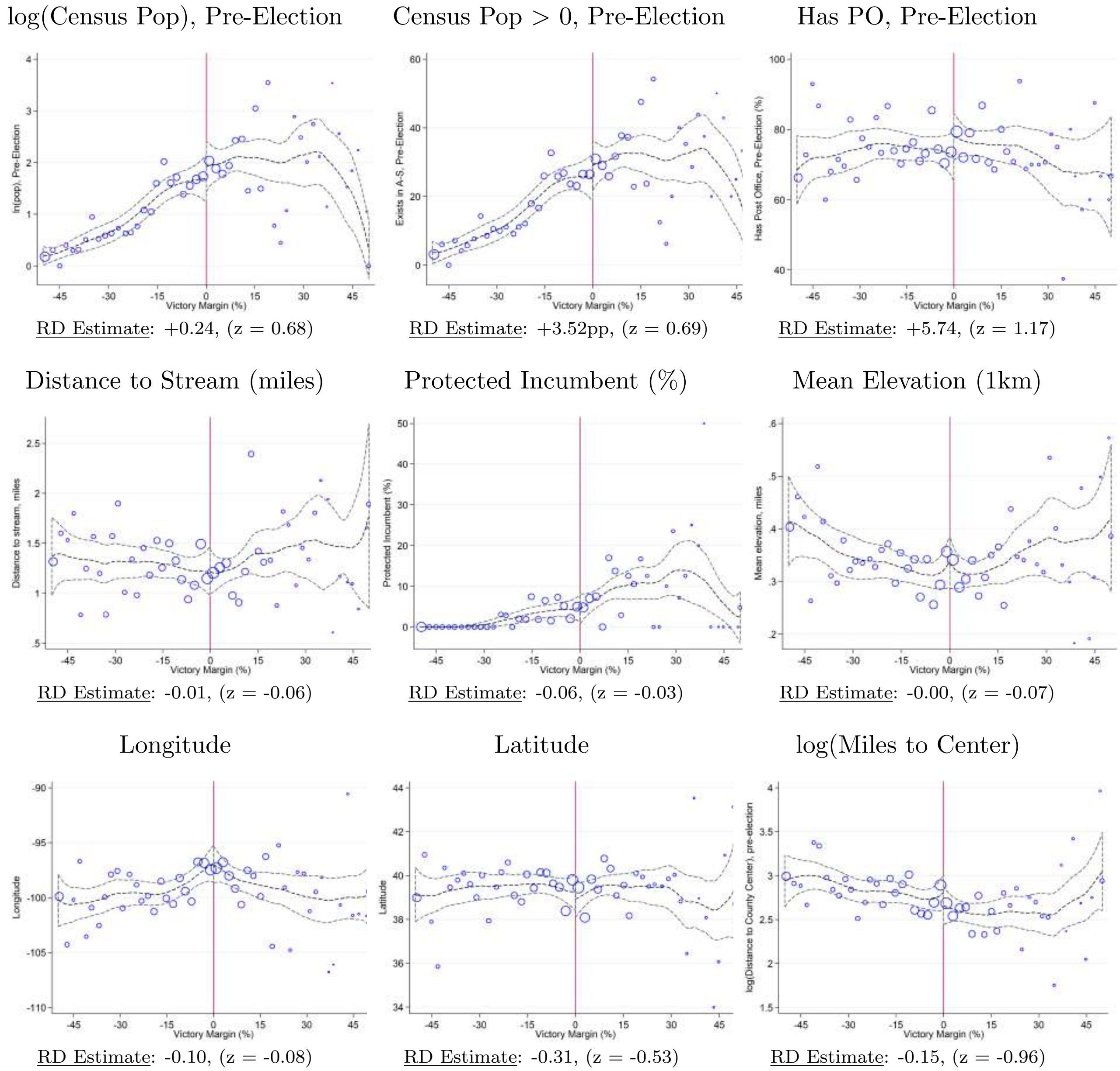
Notes: Distribution of population of census CDPs (a) for all US towns (b) our sample towns, conditional on being within the bandwidth for victory in a county-seat election.

Figure C.2: Industry Mix in Census Microdata: Close Winners



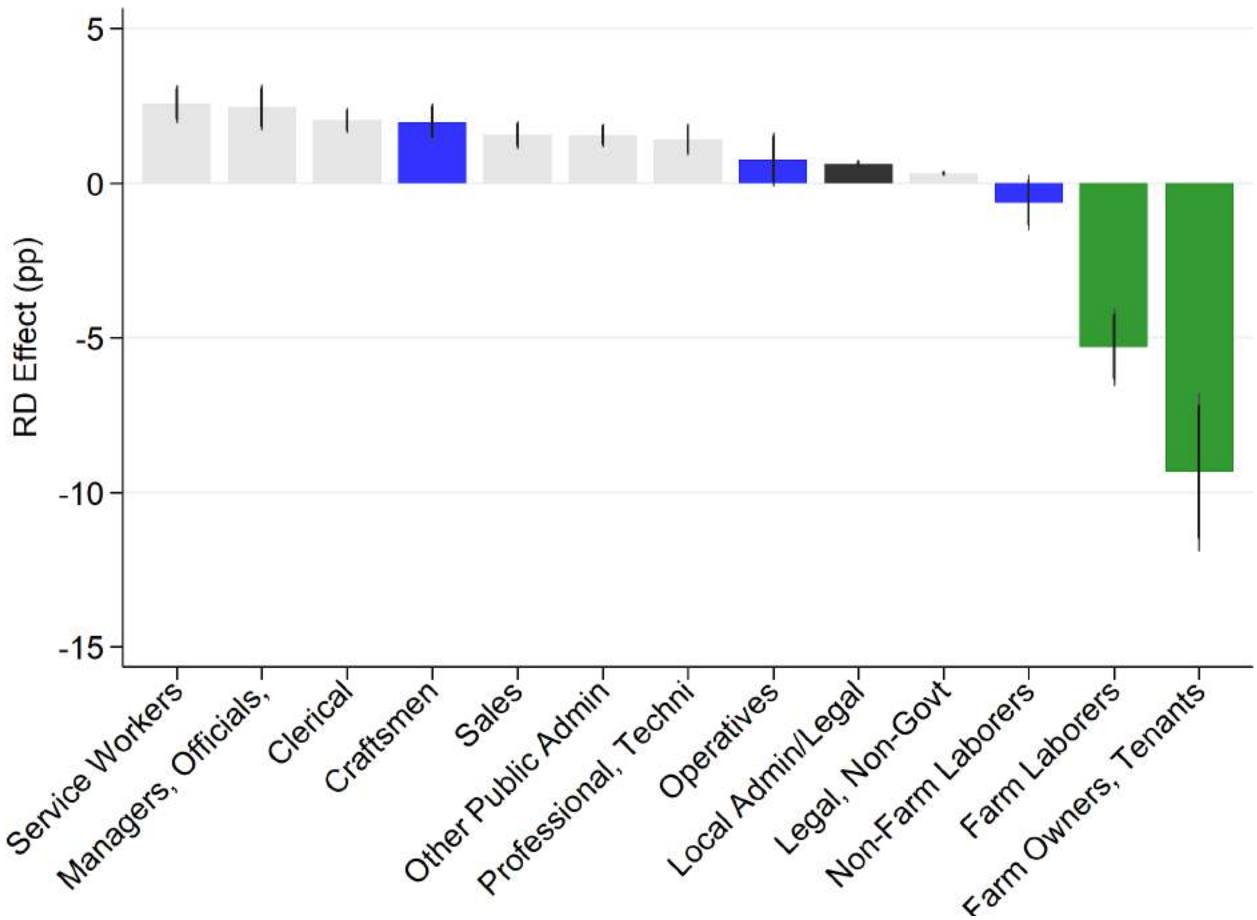
Notes: Average industry statistics in census microdata in years relative to the county-seat election. The sample consists of "close winners": winning locations within the main RD bandwidth.

Figure C.3: RD Balance Tests — Multiple Outcomes



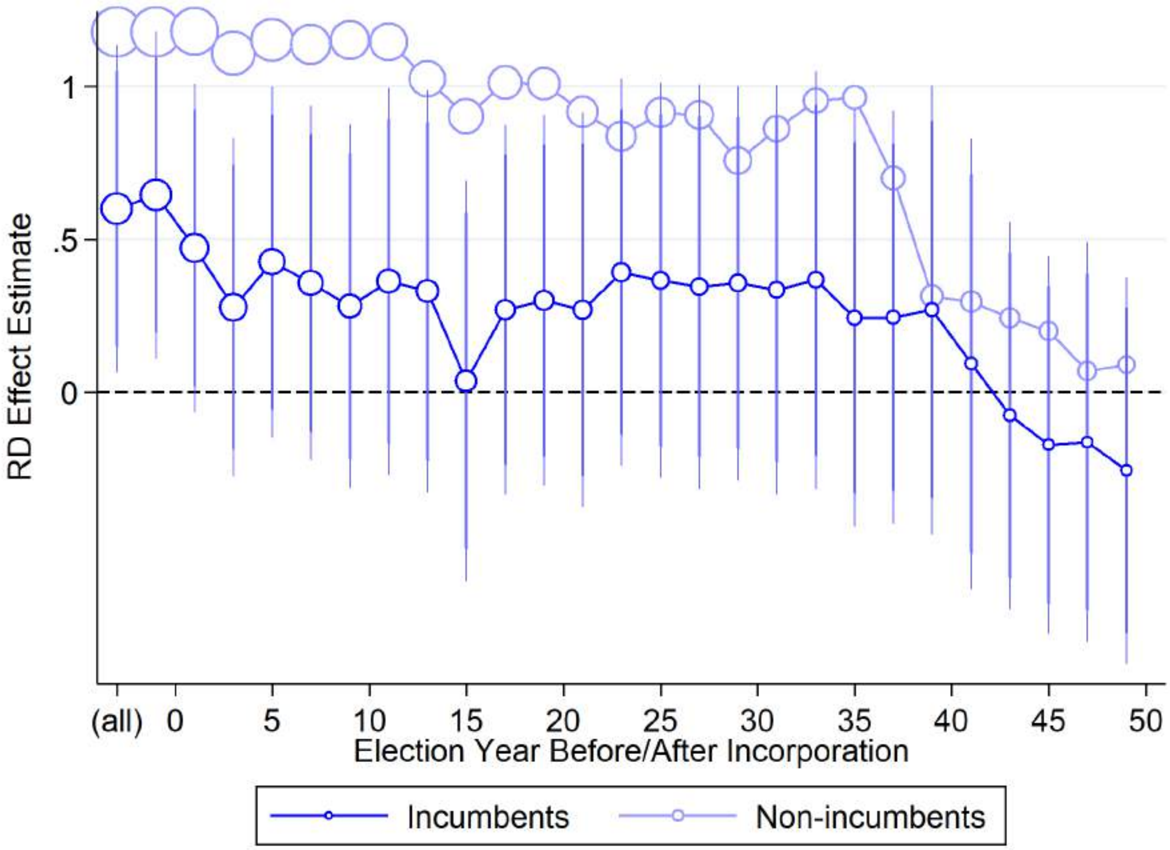
Notes: Binned scatter plot and local polynomial fit for multiple preperiod or geographic characteristics. RD estimates and z-scores are presented below as text.

Figure C.4: Occupational Change in 1940



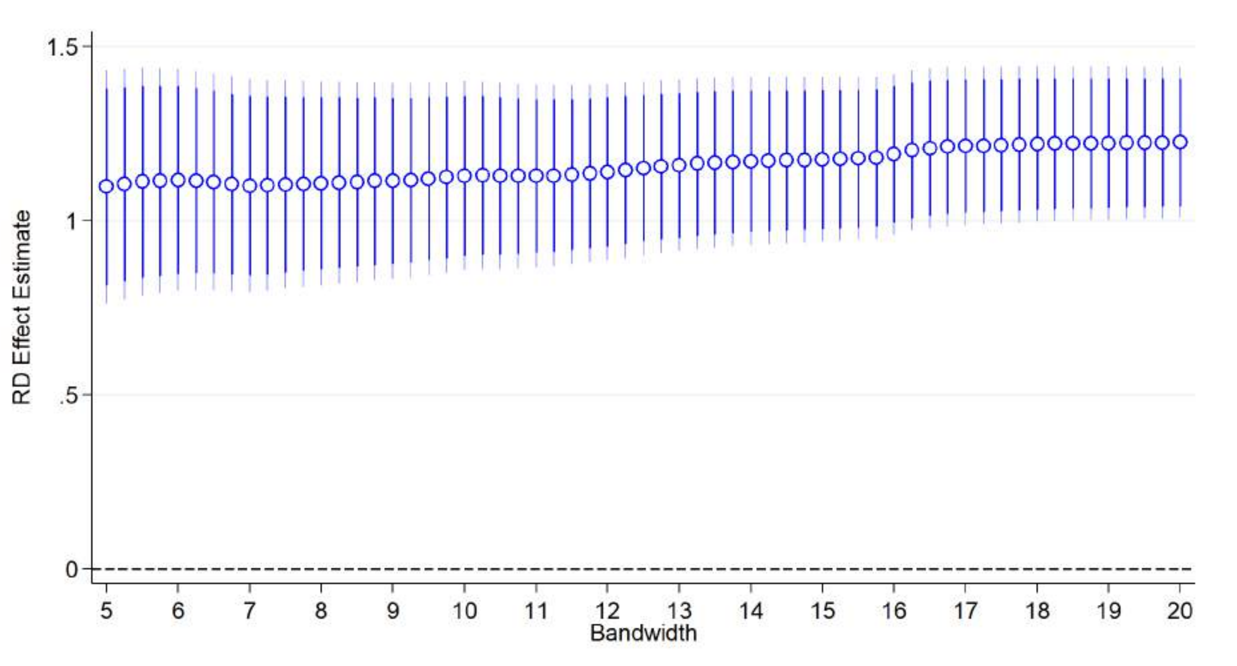
Notes: RD estimates of equation (1) for 1940 job categories for winning towns' enumeration districts compared to losing towns' within optimal bandwidth. Each bar represents a regression for a particular category. Local government and other public administration jobs are defined based on industry. The remaining categories exclude government workers and are coded based on occupation. Colors indicate "white collar" (gray), "blue collar" (blue), agricultural (green), or local government (black) occupations. For occupation definitions see Section A.15. Geographic controls as listed in Section A.12.

Figure C.5: Density Effect Size and Timing: Incumbent Towns

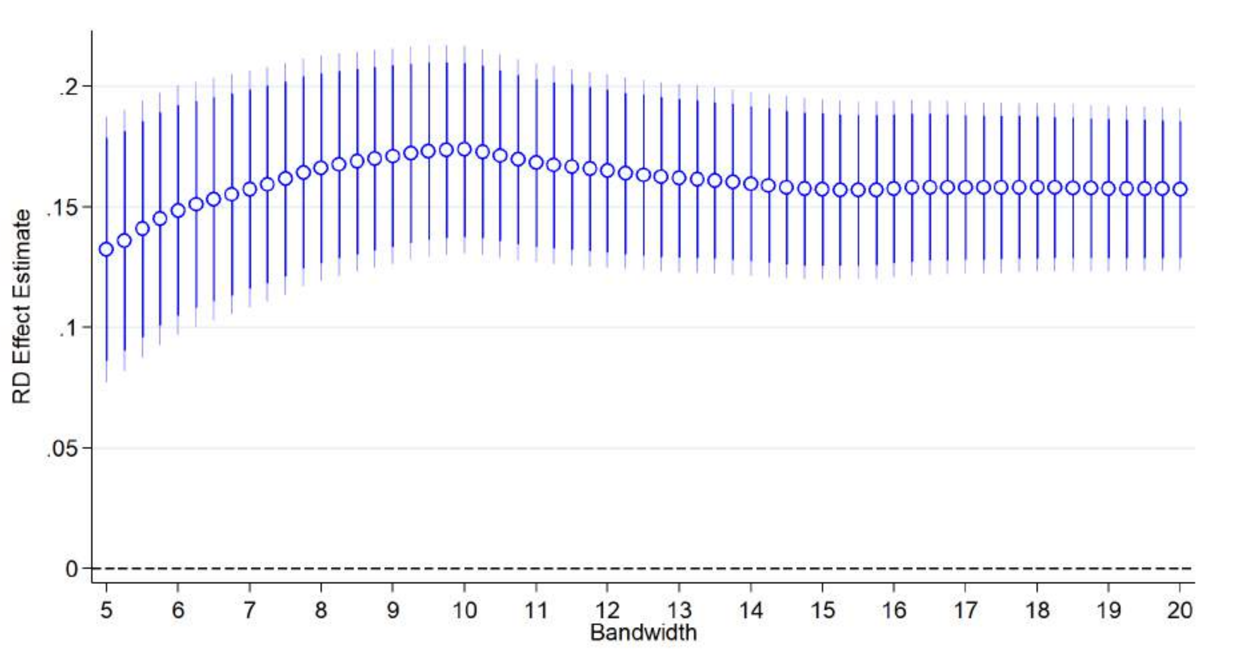


Notes: This figure replicates the analyses of Figure 8a. It splits the sample by towns listed in our archival sources as incumbents and the remaining sample ("non-incumbents"). Standard errors on the latter are removed for visual clarity. Because the incumbents group has, at most, one observation per election, we necessarily remove county fixed effects from the control set as their inclusion would subsume almost the full sample. As a historical note, we code all seats as incumbents regardless of origin. For example, county commissioners in some cases named a "temporary" seat prior to the election. We would code this location as an "incumbent" even if it never won an election in our data.

Figure C.6: Bandwidth Selection



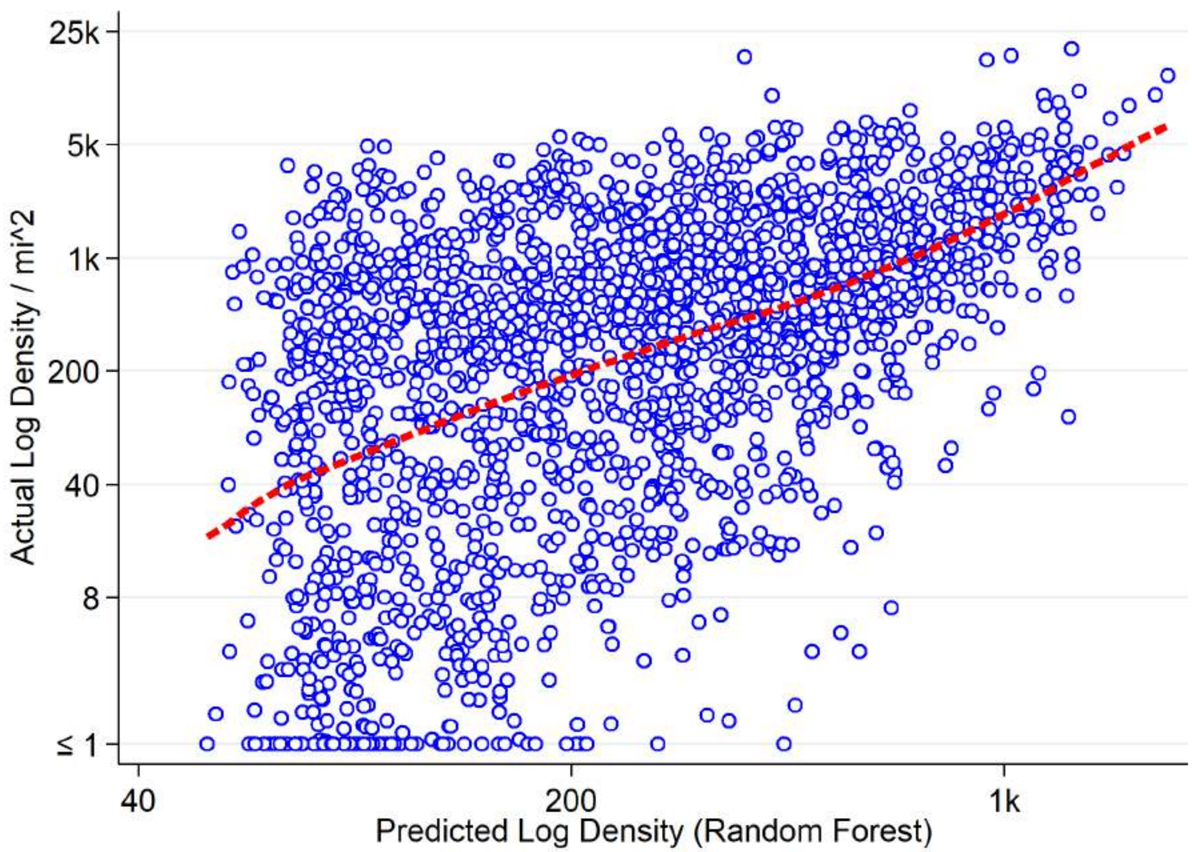
(a) (log) Population Density within 1 mile



(b) (log) 1940 wages, private, nonlegal

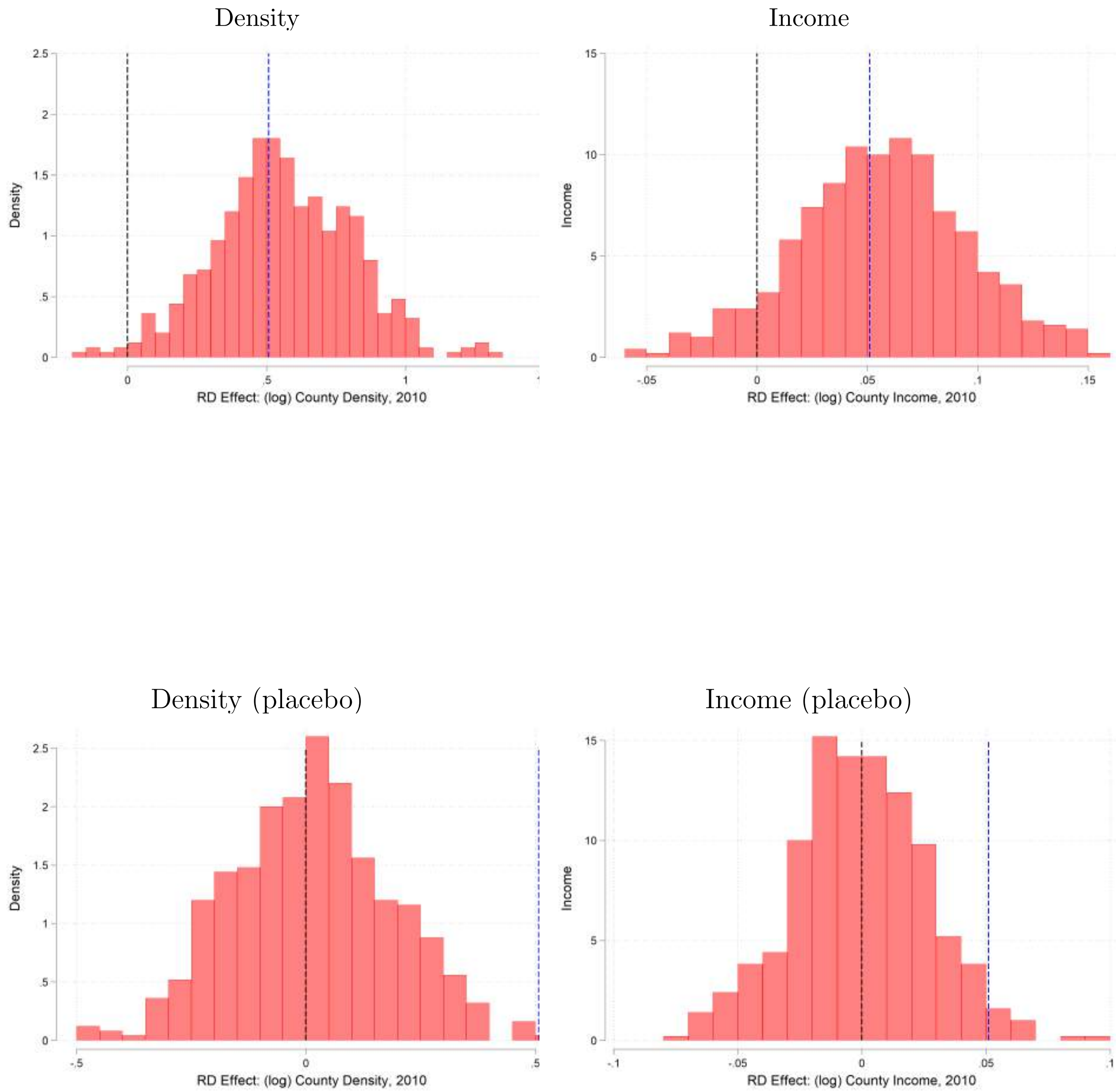
Notes: RD estimates on log population density of census blocks within 1 mile of town centers (panel (a)) and 1940 wages excluding government and legal workers (panel (b)). Each point represents a regression at the specified bandwidth. Confidence intervals of 90% (thick line) and 95% (thin line) shown over each estimate. Controls include the geographic characteristics listed in Section 4 and county fixed effects.

Figure C.7: Predicted vs. Actual (log) Density



Notes: the y-axis shows actual (log) density in 2010 within 1 mile of the county center according to census blocks. The x-axis depicts out-of-sample random forest predictions of that outcome using the controls listed in Section A.12 as the predictors. A local linear regression fits a dashed line between the two measures.

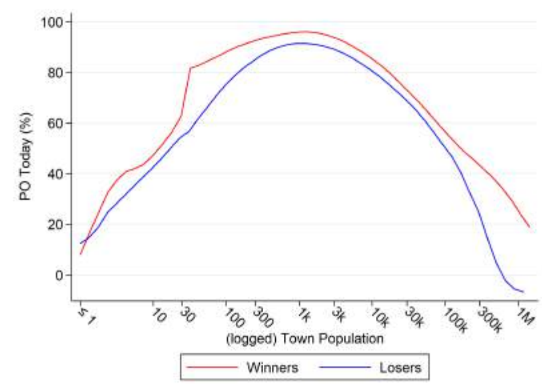
Figure C.8: Bootstrapped County-level RD Estimates



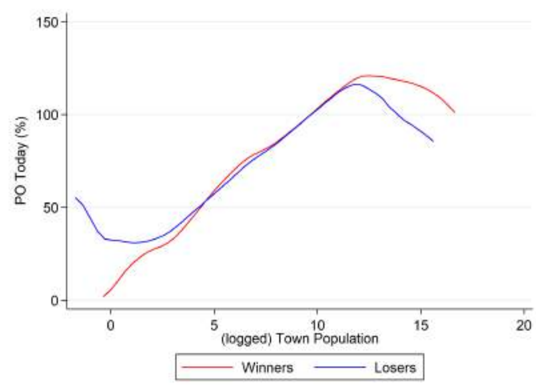
Notes: histograms of bootstrapped GRF estimations as described in Appendix Section B.3 and Section 7.3. Blue vertical lines show the main, full-sample estimate and black vertical lines are marked at 0.

Figure C.9: Public Goods versus Population, Election Winners and Losers: Post Offices

(a) Post Offices (unconditional)



(b) Post Offices (residualized)



Notes: Relationship between various measures of public goods and population for close (within the default RD bandwidth) election winners and losers. Each line represents a local linear regression of the presence of the feature on log formal town population. The outcomes is a post office within one mile of the center. The left graphs present the unconditional relationship, the right graphs show the relationship on residuals after regressing on the geographic characteristics listed in Section A.12 and county fixed effects.

Table C.1: Location-Level Summary Statistics

	N	Mean	SD	p(10)	p(25)	p(50)	p(75)	p(90)
Modern Seat (%)	2,642	38.6	48.7	0	0	0	100	100
Town Pop 2010	2,642	9679.9	57340.7	0	0	927	3560	13542
Pop/mi^2 $<$ 1 mile	2,443	924.3	1475.2	7.51	70.7	416.1	1120.0	2462.3
Election Year	2,642	1884.6	23.5	1856	1870	1884	1903	1914
Election: Years Post Creation	2,642	19.0	22.1	0	2	11	30	49
Election Number	2,642	1.34	0.75	1	1	1	1	2
Vote Margin (%)	2,642	-7.88	21.9	-42.2	-21.8	-4.08	4.94	16.5
Longitude	2,443	-98.7	10.5	-117.0	-102.6	-97.2	-92.7	-85.5
Latitude	2,443	39.6	5.15	31.7	36.4	40.1	43.7	45.9
Zip Income ($)	2,334	48196.5	19879.6	36322.0	39993.6	45216.3	51209.5	59481.6
Top 5% Share	2,334	21.1	6.39	14.4	17.5	20.5	24.3	27.8
White Collar, 1940 (%)	2,448	32.4	14.7	8.39	20.9	36.6	43.8	48.7
White Collar, 2010 (%)	2,259	76.6	18.5	64.6	74.1	80.1	85.6	90.8
Miles to County Center	2,443	35.4	84.9	2.74	5.61	12.0	29.3	64.8

Notes: Summary statistics on location-level data. Shown respectively are the nonmissing count, the arithmetic mean, standard deviation, and 10$^{\text{th}}$, 25$^{\text{th}}$, 50$^{\text{th}}$, 75$^{\text{th}}$, and 90$^{\text{th}}$ percentile values.

Table C.2: Effects on Income and Proxies

	1940				2010		
	(1) log(Wages) Above Avg.	(2) log(Home Value)	(3) Occscore	(4) White Collar (%)	(5) Top 5% Share	(6) Top $\frac{1}{3}$ Wage	(7) White Collar (%)
Win	0.077***	0.24***	2.48***	9.71***	2.88***	4.14***	4.73***
	(0.013)	(0.036)	(0.29)	(1.00)	(0.43)	(0.61)	(1.14)
County FEs	Y	Y	Y	Y	Y	Y	Y
Geo	Y	Y	Y	Y	Y	Y	Y
SEs / Clusters	Election	Election	Election	Election	Election	Election	Election
BW (pp)	14.5	14.5	14.5	14.5	14.5	14.5	14.5
N	1044	1041	1044	1044	1020	990	990
N (clusters)	538	538	538	538	537	541	541
$E[y]$	$95	$1,920	23	34%	21%	27%	77%

Notes: RD estimates of impacts of county seat victory on various long-run outcomes. (1) shows the impact on log(wages) minus the log of the national average of wages for all workers in the same occupation. (2) shows the impact on the log of home value in the 1940 micro census. (3) shows the impact on the 1940 census occupational income score ("occscore"). (4) shows the impact on white-collar jobs. (5) shows the impact on the share of income going to the top 5% in zip-level IRS data. (6) shows the impact on the fraction of LODES jobs in the highest wage tercile. (7) shows the impact on white-collar jobs in the LODES data. For sector definitions see Section A.15. Geographic controls as listed in Section A.12.

Table C.3: Main Effects by Scope of County Government

	Officers Below Median			Officers Above Median		
	(1) ln(Pop) [< 1 mile]	(2) log(Wages) 1940	(3) log(Income) 2010	(4) ln(Pop) [< 1 mile]	(5) log(Wages) 1940	(6) log(Income) 2010
Win	1.08***	0.16***	0.066***	1.16***	0.12***	0.12***
	(0.15)	(0.023)	(0.023)	(0.17)	(0.029)	(0.030)
County FEs	Y	Y	Y	Y	Y	Y
Geo	Y	Y	Y	Y	Y	Y
SEs / Clusters	Election	Election	Election	Election	Election	Election
BW (pp)	14.5	14.5	14.5	14.5	14.5	14.5
N	667	661	653	385	383	367
N (clusters)	339	337	336	202	201	201
$E[y]$	1,108	$712	$48,642	900	$735	$47,901

Notes: RD estimates of impacts of county seat victory on various long-run outcomes separated by high/low scope of county government. For each state, we count the number of major county offices as listed in Murphy (2009) and split the sample into above and below median states. The outcomes are respectively population density within 1 mile; log(wages) in 1940 for nongovernmental, non-lawyer/judge workers; and 2010 Zip-level income per IRS data. Full-sample estimates are given in Figure 4 and Table 1. Geographic controls as listed in Section A.12.