Land Concentration and Long-Run Development in the Frontier United States

Cory Smith*

October 23, 2023

Abstract

I study the long-run economic effects of land concentration on the American frontier. Using quasi-random variation in initial land allocations from a checkerboard formula, I analyze a large database of property assessments and find that historical concentration reduced modern land values by 4.5% and fixed capital by 23%. Modern effect sizes are 23%—64% of their historical equivalents, indicating significant rates of both persistence and convergence over the last 150 years. Using archival data on tenant contracts, I argue that the low-powered incentives of share agreements discouraged investment by large-scale owners with long-term effects.

Keywords: land concentration, land allocation, American frontier
JEL codes: O13, N51, Q15

*University of Maryland, cbsmith@umd.edu. This research was generously funded by the George and Obie Shultz Fund. I am extremely grateful to my advisors Dave Donaldson, Daron Acemoglu, and Abhijit Banerjee as well as Treb Allen, Melissa Dell, Tavneet Suri, Esther Duflo, the participants in the MIT Development Lunch, and the participants at the NBER Summer Institutes for their feedback on innumerable versions of this project. This project could not have been completed without the help of Madyha Leghari, Raquel Meneses, Mahvish Shaukat, Ben Milner, Jetson Leder-Luis, Madeline McKelway, Sarah Abraham, Alex Olssen, Ryan Hill, Donghee Jo, Lisa Ho, Moya Chin, Pablo Azar, Colin Gray, and Nathan Zorzi, who unerringly gave their time to help develop ideas, create graphic designs, plan presentations, and proofread.

1 Introduction

In many agricultural economies, land is concentrated in the hands of large-scale owners who hire tenants to work their properties. Classic economic theorists such as Smith (1776) and Marshall (1890) regarded this arrangement as unproductive. Smith highlighted the role of investment, writing that "a great proprietor is [seldom] a great improver," while Marshall focused on tenants' supply of effort. Both critiqued crop-sharing contracts that divided output between owners and tenants and thus limited individual incentives. However, Marshall's analysis focused on static models, whereas Smith's focus on investment highlights the potential for dynamic effects.

The short- and long-run effects of land concentration could differ for several reasons. The Coase theorem suggests that under ideal conditions, initial allocations are unimportant as markets can reallocate land to the most productive owners (Coase, 1960). However, imperfect conditions can hinder or slow this process (Bleakley and Ferrie, 2014). When labor markets make effort and investment contractable, landowners can replicate first-best outcomes with their tenants (Cheung, 1969). Even when land concentration has long-term effects, their sign is theoretically unclear. In contrast to classic incentive problems, large-scale ownership could be beneficial due to economies of scale (Allen, 1988) or to owners' ability to provide public goods (Dell, 2010). Low-powered incentives could also prevent the over-exploitation of land by tenants on short-term contracts (Dubois, 2002).

It is empirically challenging to study the long-term effects of land concentration because patterns of land ownership are shaped by endogenous social processes. Government policies, geographic features, and the overall wealth distribution can all plausibly affect concentration. Previous research has focused on the productivity response to state-driven land reforms, which capture immediate adjustments to heterogeneous policies (Banerjee, Gertler and Ghatak, 2002; Ghatak and Roy, 2007). However, to study long-term effects, we would ideally consider areas where land was historically allocated arbitrarily and compare them over time.

In this paper, I estimate the long-term effects of land concentration on the American frontier, using quasi-random variation from land allocation policies. Although this region is sometimes popularized as the domain of independent pioneers, in practice many parts of it were also held by large-scale landlords who engaged in tenant farming. To capture this contrast, my work compares two major pillars of 19th-century American land policy: the 1862 Homestead Act and railroad land grants. The Homestead Act aimed to reserve the frontier for small-scale owner-cultivators, requiring that settlers live on their property for five years and initially restricting their plots to 160 acres. Paradoxically, the government also used millions of acres of farmland as in-kind payments to railroad companies. These companies auctioned off their land in much larger blocks, typically to wealthy purchasers. The two policies thus sharply contrasted in whether they discouraged or encouraged land concentration. Collectively, Homestead and railroad land grants governed the settlement of approximately 25% of the country's area in a foundational time for many of its new states.

These two policies were applied in an arbitrary manner that increased land concentration in alternating square miles. Before settlers arrived, land had previously been divided into a square grid by the Public Lands Survey System (PLSS), and each square mile "section" received an identification number from 1 to 36. In policy areas, railroad companies received the odd-numbered sections, while the even-numbered ones were primarily settled via the Homestead Act. This formula was often called a "checkerboard" as it resulted in an alternating pattern of ownership between small-scale owners (via the Homestead Act) and large-scale owners (via railroad land purchases). The formula balanced the quality of land in odd- and even-numbered sections, with geographic characteristics typically differing by less than 0.002 standard deviations.

While farm sizes varied and were determined by individual owners, railroad lands were almost always part of larger properties. By default, even-numbered squares were divided into four 160-acre Homestead plots, roughly typical for a single-family farm in the late 1800s. Railroad buyers faced no particular size limits. Many bought multiple 640-acre squares with holdings reaching into the thousands of acres. Alone, a single 640-acre square would have been among the largest 2% of contemporary farming units in both my six sample states and the US overall (Census of Agriculture, 1880).

I find that historical land concentration led to fewer investments and a less intense practice of agriculture for approximately 150 years after the initial allocations. Today, odd-numbered sections have 4.5% lower land value and 23% less fixed capital relative to their even-numbered neighbors. Using archival data, I trace these differences back to the early 1900s. Odd sections had less "improved" land cleared for crop cultivation and more "unimproved" land devoted to less intensive forms of agriculture, such as cattle grazing. Modern even/odd differences are 23%–64% the size of equivalent ones in the early 1900s, suggesting that both persistence and convergence shape the evolution of economic activity over time. Markets tend to reallocate land in a Coasian manner as the importance of the initial allocation diminishes. However, this process unfolds very gradually, as I can still discern non-trivial differences in land use and value in 2017.

These results rely on original data sources that measure farm investment and productivity at the microscale, allowing me to fully exploit the natural experiment of the checkerboard formula. My modern data come from property assessments covering 12 million properties over 380,000 square miles, with $600 billion of agricultural land in 2017. To provide evidence on mechanisms, I digitize archival records documenting farm-level ownership and operation details in the early 1900s, which contain important information unavailable in agricultural census microdata. I also assemble historical georeferenced data on land use, schools, and taxpaying behavior. These data allow me to trace the land policies' effects back in time, providing important information on the timing and sources of land concentration's effects. While these data are more limited in geographic scope, differences in modern outcomes are consistent across most subsamples of the data, suggesting the fundamental mechanisms at play were widespread.

Turning to mechanisms, I provide evidence that land concentration's long-run impacts in this setting stemmed from tenant farming contracts. Most directly, I use archival data to show that land concentration increased rates of non-owner operation by 10 percentage points (27%) in a 1940 survey and crop share agreements by 3.7 points (18%) in a 1965 one. I additionally conduct a heterogeneity analysis using county-level variation in the frequency of share agreements among tenants. Counties with high rates of share tenancy experience the largest land value loss in odd sections; counties with high rates of cash rental experience few or no losses. These differences are not driven by different state or land quality composition. The heterogeneity analysis supplements the direct evidence and is consistent with classical theories that emphasize the principal/agent issues in share agreements. Both analyses broaden the classical theories by considering how these issues dynamically affect investment and usage over periods much longer than an individual contract.

I find little evidence for other mechanisms in this context. The main results are similar across all states and railroad companies in my sample, meaning that none of their individual policies explain the effects. Quantitative and qualitative discussion of the Homestead Act rules out its peculiar features as explanations. I individually argue against a longer list of alternate mechanisms including urban growth, lack of secure property rights, fragmented property ownership, and financial speculation. The tight zeros on land quality differences in the modern data address both concerns of an imbalance across the two land allocation policies as well as post-allocation environmental degradation. Differences in public goods provision at the square mile level are small and disappear when accounting for population differences. While my analyses occur at the square mile level and are not suited to testing whether land concentration affected political outcomes at higher levels, conversely almost all political units are controlled for by my models' fixed effects.

One significant barrier to reallocation was the difficulty smallholders faced in raising capital. Under ideal conditions, large-scale owners could have simply sold their property to neighboring Homestead farmers or similar new owners who would use it more intensively, in line with the predictions of the Coase theorem. In practice, this was likely difficult. I provide empirical support for this barrier by considering a sample of property tax records from 1900. Relative to their large-scale neighbors, even-section owners took longer to raise the money to pay their taxes and were more likely to do so through an intermediary. While other frictions were likely present in historic land markets, the difficulty of raising money for small-scale owners is a direct example of an impediment to the Coase theorem's predictions. The result is also consistent with the motivation for the Homestead Act's provision of free land, aimed at attracting settlers who could not otherwise afford to purchase it.

In the final part of my analysis, I show that the impacts of the railroad land grants do not shrink at a higher level of aggregation. In theory, Homestead lands could have benefited at the expense of railroad lands, meaning that the policy simply reallocated scarce resources across farms and the even/odd differences overstated the total impact. I test for this effect by comparing non-railroad lands at the policy boundary with those just outside in a regression discontinuity (RD) framework. Surprisingly, I find that even sections within the checkerboard have lower valuations than land just outside the grant area despite being initially allocated under the same policies.

This paper contributes to several strands of literature. A very long tradition in economics has debated the relative efficiency of different modes of land ownership. Classic theories proposing that share contracts reduce (short-term) production by limiting tenant effort have received experimental support (Burchardi et al., 2018). More recent theory explains these contracts' existence as second-best responses to constraints on resources, risk sharing, or monitoring (Reid, 1977; Winters, 1974; Alston and Higgs, 1982; Eswaran and Kotwal, 1985; Naidu, 2010). In development economics, these questions are intimately related to the study of land reforms, which have had notably heterogeneous effects (Ghatak and Roy, 2007). Some research has found that stronger tenant incentives increase output (Banerjee, Gertler and Ghatak, 2002; Markevich and Zhuravskaya, 2018), but others find mixed (Montero, 2022; Besley and Burgess, 2000) or negative effects (Adamopoulos and Restuccia, 2019). Broadly, negative impacts could be due to disruptions from property confiscation or the loss of scale economies. On the question of size, the smallest farms in developing countries exhibit diseconomies of scale (Foster and Rosenzweig, 2022), though there is debate over causality (Benjamin, 1995; Desiere and Jolliffe, 2018). In the United States, size and productivity are positively correlated (Sumner, 2014). Scale economies might also foster mechanization or other sources of agricultural productivity growth (Olmstead and Rhode, 2001; Allen, 1988; Hornbeck and Naidu, 2014). I contribute to this literature by providing causally identified evidence on the long-run impact of land concentration, with its potential to foster both tenant farming and scale economies.

This paper also contributes to work on the effects of property rights on economic development, particularly within the American West (Hornbeck, 2010; Iwanowsky, 2019; Dippel, Frye and Leonard, 2020; Hagerty, 2022). Several studies have documented persistent impacts from systems of drawing parcel boundaries (Bleakley and Ferrie, 2014; Libecap and Lueck, 2011). In contrast, I explore the effects of concentrated land ownership within a fixed parcel system.

Finally, this paper also contributes to work on the economic history of the American frontier (Bazzi, Fiszbein and Gebresilasse, 2017; Raz, 2021; Allen and Leonard, 2021). An important strand of this body specifically focuses on the Homestead Act, with recent work by Mattheis and Raz (2023) finding that small-scale, private ownership was superior to Homestead Act settlement. To the extent that homesteaders compared unfavorably to those who purchased land, my results understate the inefficiency of large landlords. A small literature has also investigated the effect of legal and regulatory issues of the checkerboard pattern. In particular, Alston and Smith (2022) similarly conducts a section-level analysis focused on legal controversies surrounding Montana's Northern Pacific Railroad grant. My paper adds to this literature by broadly evaluating the railroad grant policy across multiple states and companies, focusing on its effects on land concentration.

The rest of the paper is as follows. Section 2 discusses the historical background of American land policy and the railroad land grant formula. Section 3 provides an illustrative conceptual framework. Section 4 describes my data sources, and Section 5 presents results confirming initial land concentration and land quality balance. Section 6 details my main results on land values and investment, and Sections 7–8 discuss mechanisms, frictions, and aggregate effects. Section 9 addresses alternative mechanisms, and Section 10 concludes.

2 Historical Background

2.1 American Land Policy

The rapid expansion of the United States and its dispossession of Native American peoples allowed the country to demarcate frontier areas in a highly regularized manner. Beginning in 1785, this was done via the PLSS, which divided the new areas into an essentially square grid. The grid's main units were 6-by-6 mile squares called townships, further subdivided into 36 sections of 1 square mile (640 acres). Each section was identified by a number between 1 and 36, which corresponded to its location within a township. Appendix[1] Figure A.1a shows an example of this division, depicting townships with their numbered sections. Most states created after American independence adopted the PLSS, and its grid pattern still determines many of today's parcel boundaries.

The 1862 Homestead Act markedly changed American land policy to favor small family farmers, making it a watershed moment in the country's history. Initially, the government sold land at a typical price of $1.25 per acre with few restrictions on scale. The latter

[1] Appendices A-E, including tables and figures, will be available online.

policy thus favored aspiring large-scale owners as those of modest means lacked access to the capital necessary for purchasing and developing these farms (Gates, 1936). Building off the earlier Preemption Act (1841), The Homestead Act offered farmers a maximum-sized "quarter section" (160 acres) for a small fee if they agreed to farm it for five years; more eager settlers could purchase the title after just six months. Afterward, settlers received, unrestricted ownership of their land.

The Homestead Act's property size cap, nearly free land, and residence requirement combined to encourage owner-cultivator settlement. Large-scale owners would have been uninterested in the smaller plots, and those who aimed to become landlords would have been unable to meet the residency requirements. Although some "dummy entrymen" fraudulently served as placeholders for wealthier buyers (Bradsher, 2012), neither fraud nor adjustments to the Homestead Act overturned its promotion of small-scale ownership relative to alternative policies (see Section 5). With 1.6 million parcels granted nationally (Edwards, 2008), Homesteading quickly became the dominant form of federally administered settlement in my sample states (see Table 2). I thus use "federal" and "Homestead" settlement interchangeably.[2] I discuss further details of the Homestead Act and their relevance for my results in Section 9.2.

Paradoxically, other contemporary American land policies promoted concentrated landholding, and the most important of these were arguably railroad land grants. Beginning in 1850, federal and state governments funded railroad construction by giving companies thousands of square miles of unsettled land as in-kind payments. Companies sold these lands, notably without any restriction on purchase size or expectation that owners personally work their land (Ellis, 1946). As such, "the land policies of the railroads encouraged speculative and large-scale purchases with the result that millions of acres ... were rented or leased to incoming settlers who had expected to find free land" (Gates, 1936). This dynamic held true across many different railroad companies as "the plain fact is that basically their [sales] policies were identical" (Greever, 1951).

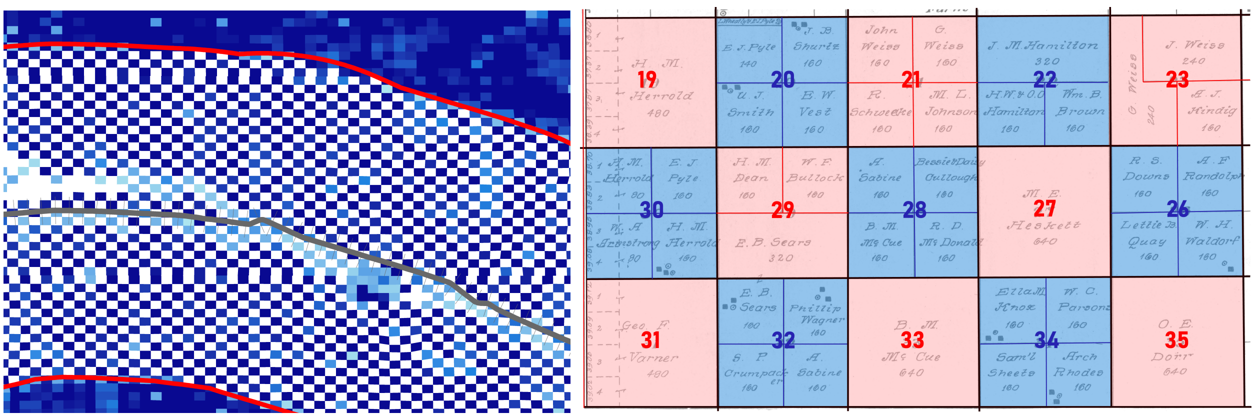
The Homestead and railroad grant policies were the two largest elements of American land policy at the time, making their potential impacts quite large. Homestead grants probably amounted to around 270 million acres of farmland (Edwards, 2008) and railroad land grants another 170 million (Decker, 1964). Together, these policies governed how roughly one-quarter of the continental US was settled and developed.

[2] In other states with more pre-Homestead settlement, this would be less applicable.

2.2 The Railroad Land Grant Formula

The allocation of lands to Homestead or railroad grants was determined formulaically, and this forms the core natural experiment of the paper. Companies were awarded land near their tracks, but the federal and state governments were reluctant to give away too much. They thus settled on a formula that gave railroads "every other" PLSS section (square mile) of land, ensuring an equal division of the policy area. This was implemented by reserving even-numbered sections (2, 4, 6...) for Homesteads and odd-numbered sections (1, 3, 5...) for railroads. In my sample states, a large majority of settled government land within the grant boundaries was given to Homesteaders (see Table 2). The primary exceptions were the "education sections" administered by local government, pre-determined to be number 16 in Florida and 16 and 36 in my other five states.[3] The grant areas were typically determined based on sharp cutoffs, with companies receiving land within a fixed distance of their track.

Figure 1: Railroad Grants and Checkerboarding



(a) Federal Land Transfers, Nebraska **(b) Farm Properties 1910, Finney County, Kansas**

Notes: Panel (a) shows in blue the percentage of land transferred by the federal government to settlers according to US Bureau of Land Management records. White sections indicate no federal transfers, typically due to railroad ownership. The area is centered around the Union Pacific line in western Nebraska. Panel (b) shows 1910 farm properties ("plat map") in Finney County, Kansas with an overlaid color scheme.

Compliance with the even/odd formula was high but not perfect. Settlers of odd-numbered sections who preceded the railroad were allowed to keep their claims. Some individuals also effectively managed to purchase even-numbered sections by having accomplices fraudulently pose as Homesteaders. Because these deviations are unlikely to have been

[3]These pre-determined lands were reserved for sale by local governments to fund education and the arbitrary locations of these sections make them an interesting topic for future research. However, they were often mismanaged by local administrators and were excluded from some key data sources; these dynamics fall outside this paper's scope. See Swift (1911); Schaede (2021).

random, I use an intent-to-treat (ITT) strategy, comparing even and odd non-education sections.

Visually, the grant formula led to what is known as the alternating checkerboard pattern. Figure 1a shows in blue the fraction of each section transferred to settlers by the federal government according to US Bureau of Land Management (BLM) records. Odd sections were held by railroads and so are typically colored white, leading to a side-by-side contrast within the grant area. Figure 1b shows how this pattern manifested in the sizes of early farm properties of one township. Sections 19, 27, 31, 33, and 35 (all odd) were held by single owners, whereas 20, 21, 26, 28, 30, 32, and 34 (mostly even) were split into four standard 160-acre Homestead farms. Although the map does not highlight it, odd-section owners could and did obtain multiple 640-acre sections.

2.3 Contemporary Farm Sizes

The Homestead Act's 160-acre size was fairly typical for late 1800s farms, though many contemporaries regarded it as insufficient in undeveloped frontier lands. Census of Agriculture (1880) reported an average farm size of 162 acres for my sample of six frontier states and 134 acres for all US states. Still, Ely and Wehrwein (1940) viewed 160 acres as "uneconomical" in some frontier areas, both for productivity and for providing a family's needs. The first point is most important for this paper as scale economies would favor land concentration. This paper does not take a stance on the historically important question of what acreage would have ensured a good standard of living for Homesteaders. Instead, it focuses on how the Act's focus on smallholder agriculture influenced investment and land productivity over long periods. Congress did subsequently adjust the Act's acreage limit with a number of amendments, though many of these came at the tail end of the settlement process and were targeted at unsettled, low-productivity areas.[4]

Purchases of railroad land had no specific size limit; buyers could typically obtain as much as they wanted. While land beyond a 640-acre section was connected only diagonally, this was likely large enough to realize most contemporary economies of scale. In 1880, only 1.6% of farms in my six sample states and 2% in the US overall exceeded this size (Census of Agriculture, 1880).[5] Buyers could and did purchase multiple squares, meaning individual holdings reached into the thousands of acres; see Section 5.3. In the long run, railroad land purchasers could also connect these already large squares by buying even-section plots.

[4]For example, the 1909 Enlarged Homestead Act increased the limit to 320 acres in areas of poorer quality.

[5]For this calculation, I assume a uniform distribution of farms in the census's 500–999 acre category.

3 Conceptual Framework

To structure my empirical work, I develop an illustrative model that describes how historical concentration affects land investment and ownership over the long run. I compare two types of landowners with different production technologies: small-scale owners who work their own land and large-scale landowners who rent their properties to tenants. When cash rent is possible, the two technologies perform equally, as per the Marshallian intuition. However, in the presence of limited resources, large-scale owners must adopt second-best share contracts, which results in lower productivity compared to small-scale producers. As a result, large owners avoid the sharecropping problem by not improving their lands.

I embed the static problem in a dynamic framework where small-scale owners face stochastic costs to raising capital. Consequently, optimal reallocations may not occur, and initial ownership has persistent effects. In any given time period, land initially allocated to larger-scale owners has less expected investment than land initially allocated to small-scale owners. However, the frictions do not always prevent reallocation, meaning that levels of investment and usage converge over time. As $t \to \infty$, Coasian convergence is achieved in the limit despite the market's imperfections.

3.1 Static Problem

Production follows a standard principal-agent framework applied to agriculture. The world consists of a parcel of land, its owner, and its operator who may or may not be distinct. Owners may engage in low-intensity agriculture on unimproved ($I = 0$) land, which always produces an output of 1. Alternatively, they may engage in high-intensity agriculture on improved ($I = 1$) land. In this case, output is stochastic at $Y = A > 1$ in the case of success and is $Y = 0$ in the case of project failure. The probability of success is equal to the effort e of the operator. Owners of $I = 0$ land may improve it to $I = 1$ for a cost $r > 0$. Effort and investment are costly to the operator and owner, respectively.

Landowners come in two types. Small-scale owners S operate their own parcel and choose their level of investment and effort. For high-intensity ($I = 1$) agriculture, large-scale owners L must contract with a tenant. Here, effort is unobservable, so owners of improved land are limited to offering payments based on project success or failure. Owners face a limited liability constraint based on tenants' ability to pay upfront and must offer them a contract giving them at least 0 expected utility. For simplicity, all agents are risk neutral. As in Banerjee, Gertler and Ghatak (2002), limited liability determines whether share, cash, or intermediate contracts are optimal for the landowner. Owner-operators face no constraints and so choose first-best effort, but tenants of large-scale owners replicate this only when cash rent is possible and otherwise receive suboptimal effort.

Note that S owners produce weakly more on improved ($I = 1$) land, with the difference holding strongly when L owners use share agreements. In those cases, L owners can avoid the contractual issues by not improving their properties and leaving them in the $I = 0$ state.

3.2 Dynamic Problem

The parcel is initially allocated to either an S- or L-type owner with $I = 0$, reflecting Homestead and railroad distributions of frontier land. At the start of a new period, the previous owner sells the parcel and exits the model along with the operator. The landowner makes a take-it-or-leave-it offer to an agent of either type. S types, however, face stochastic costs in purchasing, reflecting their lack of access to capital.[6] The costs are distributed according to F_S. Before the sale, improved land has a $0 < \delta < 1$ chance of depreciating into unimproved land, and then transaction costs are determined. Owners discount across periods at a rate of $0 < \beta < 1$.

3.3 Predictions

The model's main predictions and their empirical equivalents are summarized in Table 1. Initial owner type can play an important or trivial role depending on the model parameters. The trivial cases occur in the following cases: (1) land quality A is sufficiently low so that neither owner type will invest, (2) large-scale owners can offer cash rent and make the same decisions as small-scale owners, and (3) transaction costs are such that only one type of owner purchases land with probability 1. In these situations, we should see minimal effects from initial land concentration. I test the first two cases empirically, considering cases of low-quality land and counties where cash rent is common.

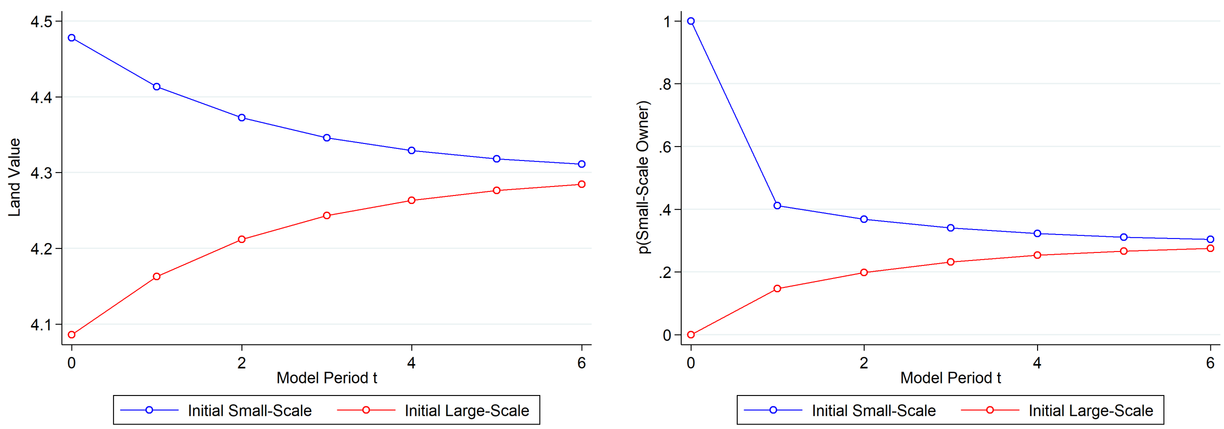
Table 1: Model Predictions and Corresponding Empirical Results

Scenario	Prediction	Empirical Results
Initial land concentration (general case)	Reduced investment	Table 4
	Persistent concentration	Figure 6a
	Attenuation over time	Figures 4, 6a
Initial land concentration (cash rent common)	Few or no effects	Figure 5
Initial land concentration (low land quality)	Few or no effects	Appendix Table A.4

[6] Other interpretations give the same or similar results, for example, the costs of reallocation across different owners.

In the non-trivial case, small-scale owners invest strictly more than large-scale owners with persistent effects over time. Here, the game follows a Markov process with $I = 0$ or $I = 1$ being the only relevant state at the point just before frictions are realized. Large-scale owners value improved land relatively less than small-scale owners due to contractual frictions and are less likely to purchase improved land. As a result, initial land concentration (L ownership) results in persistently less investment on average. These differences attenuate over time in a standard process of Markovian convergence as $t \to \infty$. Example results for specific parameter values are given in Figure 2. Proofs and further details are available in Appendix Section A.

Figure 2: Sample Model Dynamics



(a) Land Value **(b) Probability of Small-Scale Ownership**

Notes: The figure shows expected land value and the probability of smallholder ownership for initial small- versus largeholder land. Parameter values are given by $A = 2$, $\beta = 0.75$, $\delta = 0.25$, and improvement costs of $r = 1$. Effort costs are quadratic: $\frac{1}{2}e^2$. Tenants face a limited liability and an outside option of 0, leading to an even-split sharecropping contract. S-type buyers face uniform costs on [0,2].

3.4 Historical Support

The conceptual framework proposed in this text is in line with historical and contemporary evidence about the US agricultural economy circa 1900. Gates (1942) discusses how land concentration reduced investment on the American frontier, writing that "[large-scale] ownership and tenancy did not always result in the best use of the land." Kansas farms surveyed by Grimes (1919) specifically reported lower rates of investment for farms under share versus cash rent because "the landlord receives his share of the benefits without sharing the expense."

Despite potential disadvantages, the USDA (1923) notes that the "concentration of land ownership in large holdings is favorable to landlordism and tenancy." The ability of landlords to use cash rather than share rent was particularly limited by resource constraints: "When tenants are able to pay cash ... landlords are more likely to be willing to rent for cash than when the opposite conditions prevail." Finally, Rajan and Ramcharan (2015) note that constraints to credit were an important determinant in whether "tenants or farm workers could buy land off landlords, eliminating the agency costs associated with tenancy."

4 Data

4.1 Modern Outcomes

Land values are a natural outcome to study for agricultural economies and reflects the net present value of profits to current and future owners (Borchers, Ifft and Kuethe, 2014). I draw especially on land assessments by county and state governments for this project. Assessors either attempt to find comparable properties recently sold or estimate the net income of the property based on current environmental and use characteristics. Thus, a cattle ranch is evaluated based on the assessor's belief about the net income from cattle even if they believe wheat farming would be more profitable. Although each county's assessment procedures have unique elements, all comparisons in this paper are within county. This consistency alleviates concerns that individual farmers might systematically misreport characteristics of their farms in surveys (Desiere and Jolliffe, 2018).

I assemble 2017 assessment data from six US states covering 12 million parcels spread over 380,000 square miles accounting for $600 billion in agricultural land. The records detail each property's total value, the value of "improvements" (buildings and fixed capital, e.g., barns, irrigation systems), housing value, and the number of parcels. To measure the appropriateness of land use, I analyze the USDA's satellite-derived CropScape dataset that codes usage into distinct crop, grassland, and developed area categories at the 30-by-30 meter level. I then use these data to compute an estimated "use value" for land based on the expected profits from its current agricultural use. To estimate profits, I combine the CropScape data with productivity measures from the FAO GAEZ dataset and USDA and other price and profitability rates (see Appendix Section B.8). Thus, land with the same underlying geographic features can vary in its value based on its current usage. In 13 of my 322 counties, substantial areas of government land are not assessed, leading many economically active parcels to receive a $0 value. For these counties, I treat the satellite-derived use value as the total property value, though the results are essentially unchanged if these counties are instead dropped. Appendix B details the sample construction procedure and further GIS sources. Modern coverage is shown in Appendix Figure A.1b.

4.1.1 Validating Assessed Values

I validate the relevance of assessed land values by comparing them to sale prices in Florida, where the assessment data include sale amounts for 2016-2017. Appendix Table A.1 shows that assessed total values per acre is highly correlated with sale prices per acre. Aggregated to the PLSS section level as in my main analysis, the elasticity is 0.94. Assessed values remain highly predictive at the property level and after including controls. This result mirrors other literature, which finds that assessed values are highly predictive of sales values, though other factors often have residual information value (Bigelow, Ifft and Kuethe, 2020). Similarly, the usage-based values I construct are highly correlated with assessed values, even adjusting for township fixed effects. In both cases, valuations excluding improvements have either smaller or no additional predictive power.

4.2 Historical and Archival Sources

To elucidate the timing and mechanisms of effects from historical land concentration, I turn to historic records detailing land ownership, land use, and population from the late 1800s and early 1900s. My research either was conducted in person at the Nebraska State Archives or used digitized records from individual counties, the Library of Congress, and FamilySearch.org (see Appendix Table A.2 for the full list of sources). Where property or sales records give owner names, I link these owners to census microdata as described in Appendix Section B.5.

Because historical land records are typically collected and held locally, historical data for this project are often only available for individual counties or states. For this reason, such analyses are limited to subsamples where key outcomes are available, and stronger assumptions of external validity are required. However, my main results on modern property values hold across a variety of state and land quality subsamples (Section 9.3). As such, the same mechanisms likely applied broadly, meaning that the archival data are still informative about the complete sample.

4.3 Summary Statistics

Summary statistics are given for key variables in Table 2. Columns (1)–(7) show average or median statistics at the PLSS section (square mile) level, and columns (8)–(9) show averages for 1940 counties. Overall, the sample is large and features a diverse set of agricultural conditions. The main sample includes about 130,000 grant area sections, just under half of which today grow some crops. 87% of settled Homestead-eligible (even, non-education) sections used the Homestead Act or its extensions. Another 8.5% of land in these areas

Table 2: Summary Statistics

	PLSS Sections						Counties (1940)	
	(1) Number Sections	(2) Soil Quality (z-score)	(3) Homestead (%) (Even Only)	(4) Crops (%)	(5) Value $000	(6) # Parcels (Median)	(7) Tenant Farm (%)	(8) Share Farm (% Tenant)
Whole Sample	386,224	0	84.1	49.6	6,712	3	37.8	39.8
RR Grant Areas	132,463	-.0455	86.7	47.9	2,231	2	38.8	42.8
(FL) Florida RR	1,406	-.0716	93.6	37.9	30,270	70.1	13.5	21.2
(FL) Pensacola	7,857	.455	90.3	55.7	5,844	18	29.9	43.7
(KS) Atchison & Santa Fe	7,176	1.08	96.5	87.7	1,033	5	47	61
(KS) Union Pacific	12,512	1.13	92	85.6	895	4	43.6	51.8
(MT) Northern Pacific	62,253	-.777	79.6	29.6	1,361	1	26.9	41.2
(NE) Burlington	9,011	1.73	93.5	95.3	2,923	6	55.8	33.7
(NE) Union Pacific	14,868	.814	96.6	80.4	2,273	5	52.3	40.5
(OR) Oregon & California	1,422	1.44	27.1	41.2	16,134	10	18	19.4
(WY) Union Pacific	15,958	-.8	89.6	12	1,296	.505	15.6	36.9

Notes: The table presents summary statistics for different geographic units. Column (1) shows results for section sample size, column (2) for the gSSURGO crop productivity index (full-sample z-score), column (3) for percentages of non-railroad land transferred under the Homestead Act, column (4) for the percentage of sections with at least 1% in crops per the USDA CropScape data, column (5) for total property values, column (6) for the median number of parcels, column (7) for county-level average rates of non-owner-operated (tenanted) 1940 farms, and column (8) for county-level averages of the shares tenant farms as a fraction of all tenant farms in 1940.

was similarly targeted toward smallholders, and only 5% were open to concentration due to settlement via public sale.[7] Finally, despite the popular image of the West as a bastion of pioneer independence, a diversity of landholding patterns prevailed. About 40% of farms operated under tenant contracts, and 40% of those used share arrangements.

Comparable statistics are given for the archival sample counties in Appendix Table A.3. These samples cover a range of areas. For most variables, there is at least one above average and one below average area.

5 Empirical Strategy

5.1 Unit of Analysis

For all regressions, the unit of analysis is a PLSS section, and all outcomes are aggregated at this level. Importantly, PLSS sections are pre-defined, natural units of area. Alternative units such as farms or parcels are formed endogenously and are shaped by the land allocation policy itself (see Figure 3b), making them more complex to analyze. Because assessors evaluate property at the parcel level, rather than the owner level, the vast majority of

[7] Of course, even these were not necessarily concentrated. For "other small grants," I include those given for military service, to Native Americans, and through the Bankhead–Jones Farm Tenant Act of 1937.

assessments fall within a single section; 93% of grant area sections in the modern data are formed from whole parcels. In the minority of cases where a parcel is split across multiple sections, I allocate its value uniformly by area. In the case of statistics about owners, I assign parcels the value of their owner's characteristics and compute the area-weighted average. Thus, outcomes always reflect the characteristics of a typical unit of land.

5.2 Even/Odd Regression

Within the checkerboard areas, I compare even (small-scale/Homestead) and odd (large-scale/railroad) sections: since the even-odd distinction stemmed from surveying decisions made before the railroad grants, there should be no unobserved average quality differences between the two groups. I run regressions of the form

$$y_i = \alpha \mathrm{RR}_i + X_i\beta + \varepsilon_i, \tag{1}$$

where i is a non-education[8] PLSS section (roughly 1 square mile) within a grant boundary, RR_i is a dummy variable for a railroad (i.e., odd) section, y_i is some outcome, and X_i are controls that typically include township fixed effects.[9] In my baseline results here and in other sections, I include controls for (log) section area, mean elevation, average terrain slope, the miles of streams, average soil quality, an indicator for entirely missing or unproductive soil, and latitude and longitude by state. For fat-tailed outcomes like property valuations that sometimes include 0, I transform them using the inverse hyperbolic sine (asinh) function, which allows me to interpret coefficients roughly as percentage changes.

Note that this is an ITT regression and likely underestimates the true effects of land concentration. I nonetheless retain the ITT design as deviations from the intended policy are unlikely to be random and because no comprehensive database of railroad land grants exists to rescale the reduced-form results by a first stage. Attenuation might also occur due to sales across section boundaries, a topic I explore in Section 8.

Textbooks often use checkerboard patterns as an example of spatial autocorrelation (Grekousis, 2020, Chapter 4), so by default I use Conley standard errors with a 100-mile bandwidth to allow for the important possibility of long-range spatial correlation (Kelly, 2019; Conley, 2010). Section 6 shows this approach is most conservative and results in higher standard errors relative to a range of other methods. For regressions over smaller

[8]See Section 2.1. Education sections were also pre-specified by a PLSS section number, meaning they similarly should not statistically differ in quality.

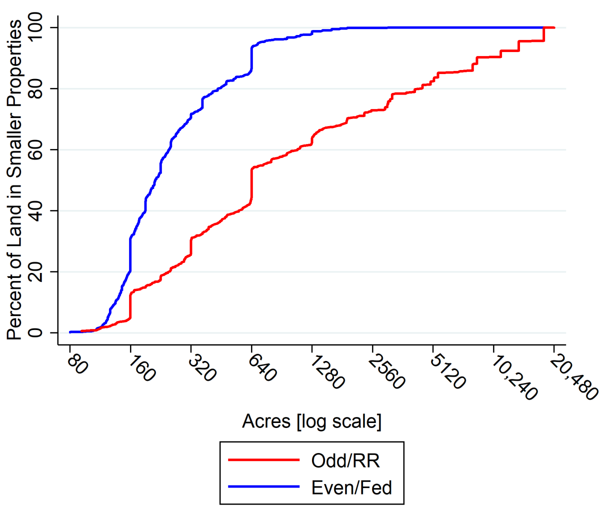
[9]County and state × railroad grant area fixed effects are also included. However, these are effectively subsumed by the township fixed effects as townships are only very rarely split across counties, states, or company grants.

areas (e.g., counties) where this approach is not possible, I instead cluster errors at the township level.

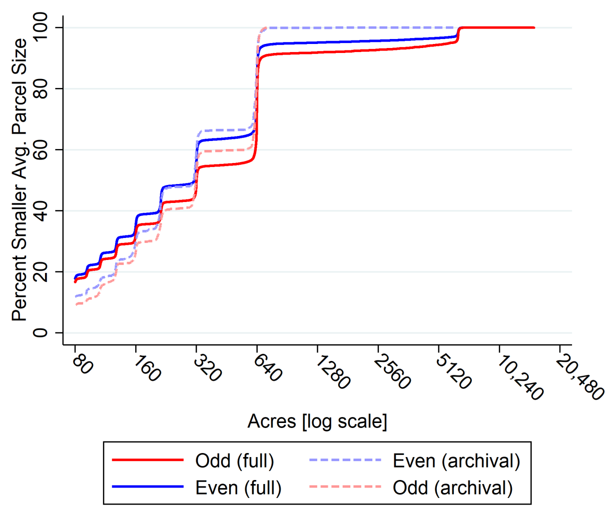
5.3 First-Stage Results

I begin my analysis by confirming that the railroad grant policy did, in fact, increase land concentration. Figure 3 presents two measures, both computed as averages at the PLSS section level. Figure 3a presents archival data from one Nebraska county about initial owners' total holdings, and Figure 3b presents data on the average size of a parcel in the full 2017 data.[10] Federal land records have been largely digitized nationwide, but comprehensive records of railroad land sales do not exist, requiring me to collect local archival data. In this case, I digitized original deeds of sale from the Union Pacific Railroad Company to individual owners and combined this information with similar federal (BLM) data.

Figure 3: Railroad Grants and Land Concentration



(a) Initial Owner Acres (Archival), 1800s



(b) Average Parcel Acres (2017)

Notes: Panel (a) shows the CDF of (log) owned property sizes of initial land allocations in Lincoln County, Nebraska, based on archival data. Panel (b) shows the CDF of acres per 2017 parcel in both the full sample and in Lincoln County.

Both measures indicate that odd (railroad) sections were more concentrated. In my preferred measure from the archival allocation data, odd-section properties are 390% larger than their even-section neighbors, driven by a shift across the distribution, including the

[10]Average parcel area is defined as the section's area divided by the number of parcels in the PLSS section. Parcels split among multiple sections are counted fractionally. Initial ownership is defined as the total holdings of the initial owner across all sections in the county regardless of allocation policy.

extreme right tail. 47% of odd sections are in properties larger than one PLSS section (640 acres), and 22% are in properties over 3200 acres. In contrast, the corresponding figures for even sections are 7% and 0%, respectively.

Modern parcel data indicate that this effect was not limited to that particular county. For administrative reasons, parcels are rarely combined, meaning that their modern boundaries partially reflect initial divisions. However, the first owners typically held many modern parcels, meaning that those differences likely attenuated over time. With this caveat, odd-section parcels are 13% larger in the county analyzed in Figure 3a and are 20% larger in the full sample. Appendix Figure A.2 shows that this result holds across a range of state and land quality subsamples, indicating that railroad land grants increased concentration broadly.

5.4 Land Quality Balance

Table 3: Balance on Geographic Characteristics

	(1) Soil (z-score)	(2) Slopes (z-score)	(3) Streams (z-score)	(4) Elevation (z-score)	(5) log(Area)	(6) log(RR Dist)
RR Effect	-0.00047	-0.00027	-0.0014	-0.00049*	0.00011	-0.0011
	(0.00097)	(0.00035)	(0.0045)	(0.00028)	(0.00049)	(0.00078)
Sample	All	All	All	All	All	All
Grant × State FEs	Y	Y	Y	Y	Y	Y
County FEs	Y	Y	Y	Y	Y	Y
Township FEs	Y	Y	Y	Y	Y	Y
SEs / Clusters	Spatial	Spatial	Spatial	Spatial	Spatial	Spatial
N	132,463	132,463	132,463	132,463	132,463	132,463
$\mathbb{E}[y]$	-.046	1.2	.55	1.7	-.017	2.5

Notes: The table estimates the direct, even/odd comparison equation (1) on gSSURGO crop productivity in column (1), terrain slopes ("ruggedness") in column (2), miles of streams in column (3), elevation in column (4), log section area in column (5), and log distance to the railroad in column (6). Columns (1)–(4) are normalized as full-sample z-scores.

Table 3 uses estimates of equation (1) to confirm that even and odd sections are balanced across a range of geographic characteristics, with the coefficients very small and precisely estimated. Z-scores of land quality characteristics are estimated in columns (1)–(4) and show differences of 0.0014 standard deviations or less, with the largest standard error of 0.0045. Similarly, differences larger than 0.0027 log points in section area or distance to a railroad can be ruled out. These differences are not statistically significant except for elevation at the 10% level. The tight null result on distance from the railroad highlights the fact that this paper has little to say about railroads per se: the checkerboard formula allocated land symmetrically across the two policies, and the typical section was more than 10 miles from the track.

Because land quality measures are based on modern data, these results point against environmental degradation or mismanagement as a major explanatory factor. This result is consistent with Burchardi et al. (2018), who find no soil quality change from experimental variation in land contracts. Hagerty (2022) similarly finds minimal soil differences across California irrigation districts with higher water allocations despite long-term changes in crop choice. While my tests cannot rule out environmental effects over spans larger than a PLSS section, by definition those cannot affect the section-level regressions of this paper.

6 Results: Land Values and Investments

In this section, I compare even and odd sections on dimensions of land values, investment, and other measures of inputs. I use both the full 2017 data and archival subsamples to compare the effects over time.

6.1 Land Values

Historic concentration lowered modern land values by about 4.5% as shown in Table 4, Panel A, and the results are statistically significant at the 1% level across a number of standard error methodologies. In each column, my preferred Conley standard errors with a 100-mile bandwidth are given in parentheses. Standard errors using county clustering, township clustering, and simple heteroskedastic-robust errors follow sequentially. In all cases, the Conley errors are the most conservative and reflect the potentially spatial nature of treatment assignment, and I therefore adopt them as the default. The columns sequentially add a rich set of controls including soil quality and township fixed effects, with column (6) including all controls as my preferred specification. The estimated differences are minor, with column (1) and column (6) differing by approximately 0.0015.

The negative effect of historic concentration is striking for two reasons. First, it demonstrates that the drawbacks larger owners faced outweighed any advantages in scale economies or access to capital. These owners also purchased their land and might have been positively selected relative to Homesteaders who obtained free land.[11] These findings suggest that the system of tenant farming employed by larger owners persistently reduced investments. Second, the fact that an initial allocation of land had persistent effects over 150 years suggests

[11] See Mattheis and Raz (2023) for a detailed study of this question.

that US land markets fell short of the ideals formulated in the Coase theorem. I explore these mechanisms further in Section 7.

The placebo regressions in Panel B of Table 4 further support a causal interpretation of my results. These regressions compare even and odd sections in areas one or more miles from the grant areas. Assessed values of even- and odd-numbered sections differ by only small, statistically insignificant, and tightly estimated results. This is sensible as there are no reasons to think that even and odd sections should have differed except for the application of the checkerboard policy.

Table 4: Effects on (asinh) Total Property Value

	Panel A: Main Sample					
	(1) Value	(2) Value	(3) Value	(4) Value	(5) Value	(6) Value
RR Effect	-0.047***	-0.047***	-0.046***	-0.044***	-0.046***	-0.045***
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.014)
	[0.0086]	[0.0086]	[0.0085]	[0.0087]	[0.0086]	[0.0084]
	{0.0049}	{0.0049}	{0.0049}	{0.0095}	{0.0048}	{0.0048}
	⟨0.010⟩	⟨0.010⟩	⟨0.0088⟩	⟨0.0073⟩	⟨0.0069⟩	⟨0.0068⟩
State FEs	Y	Y	Y	Y	Y	Y
Grant × State FEs		Y	Y	Y	Y	Y
County FEs			Y	Y	Y	Y
Township FEs				Y	Y	Y
Geo Controls					Non-soil	Y
N	132,463	132,463	132,463	132,463	132,463	132,463
$\mathbb{E}[y]$	$2,134k	$2,134k	$2,134k	$2,134k	$2,134k	$2,134k
	Panel B: Placebo Sample					
	(1) Value	(2) Value	(3) Value	(4) Value	(5) Value	(6) Value
Odd Section	-0.000014	-0.00014	0.00012	-0.00061	-0.0016	-0.0012
	(0.0052)	(0.0051)	(0.0048)	(0.0049)	(0.0053)	(0.0050)
State FEs	Y	Y	Y	Y	Y	Y
Grant × State FEs		Y	Y	Y	Y	Y
County FEs			Y	Y	Y	Y
Township FEs				Y	Y	Y
Geo Controls					Non-soil	Y
N	230,483	230,483	230,483	230,483	230,483	230,483
$\mathbb{E}[y]$	$9,562k	$9,562k	$9,562k	$9,562k	$9,562k	$9,562k

Notes: The table reports regressions of the even/odd comparison of equation (1) on (asinh) total section property value. Panel A uses the main sample (inside the grant areas), and Panel B uses a placebo sample of areas one mile or more outside the grant area. Geographic controls are listed in Section 5.2. For Panel B, Grant × State areas are defined based on the closest actual grant area. Panel A uses Conley standard errors (100-mile bandwidth), county clusters, township clusters, and heteroskedastic-robust methods.

The reduction in land values is not driven by particular states, subsamples, or functional form but instead applies across most of the grant area. Appendix Figure A.3 runs OLS regressions separately for each state split by gSSURGO soil quality. Broadly, most of the estimates are negative, individually statistically significant, and similar across states (conditional on quality). Quantile regressions of equation (1) show that the leftward shift occurs broadly across the distribution (Figure A.5), and alternatives to the inverse hyperbolic sine leave the estimates essentially unchanged (Table A.4, Panel A).

One exception is land with very low-quality soil: in approximately the bottom 20% of soil productivity across states, historic land concentration seems to have little effect on property value, as shown in Figure A.3. In these areas, the high-intensity farming necessary to feed a family on a Homestead plot was likely impractical.[12] This suggests that persistent changes in land use might drive the land value effects, as I discuss in the next section.

6.2 Land Investments

Fewer investments and less intensive land use are the proximate causes of the lower land values seen in Section 6.1. In this section, I especially examine owners' decision to "improve" their lands as recorded in assessment data. "Improved land" refers to land that has been cleared and developed for agricultural purposes, primarily crop cultivation. Improving land was a substantial investment, often comparable in cost to the value of the land itself (Coffin, 1902). However, this investment could substantially boost production relative to "unimproved" acres left in their natural state. Unimproved land still had economic uses, particularly in ranching and grazing cattle on native vegetation. Assessors historically tracked these differences as well as the value of fixed capital ("improvements") like barns and silos. Improved land and improvement value constitute my primary measures of investment, as in the conceptual framework.

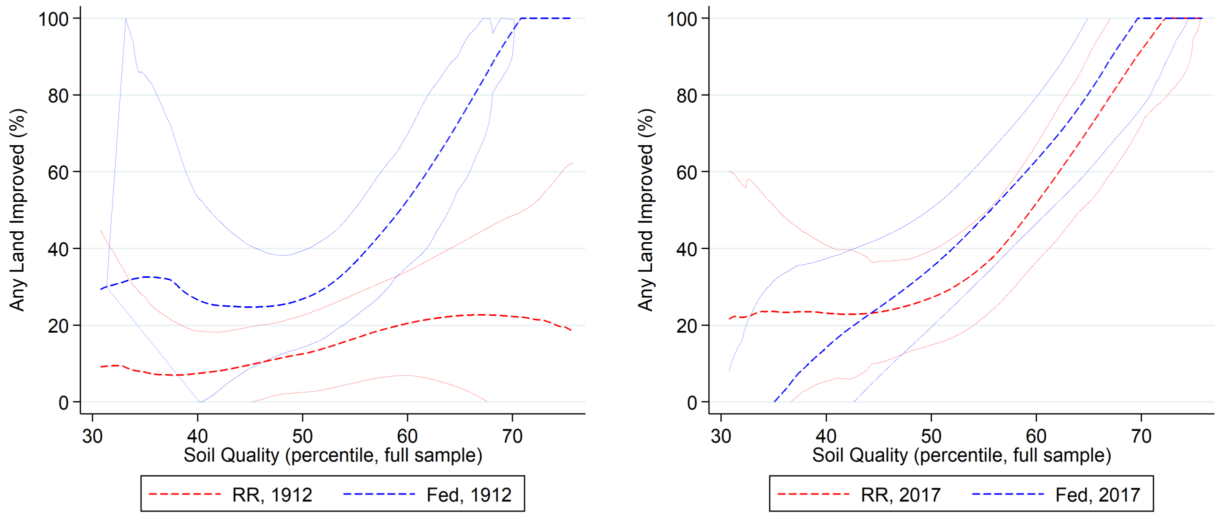
Land concentration reduced initial investments, and these effects are persistent (but smaller) today. Figure 4 illustrates the dynamic for one archival sample, comparing the percentage of sections with any improved cropland in 1912 and 2017. In both periods, historic concentration reduces improvements. Strikingly, the historical differences are largest in areas of high crop productivity as large owners' investments were largely unresponsive to land quality. In modern data, the differences are much smaller, and convergence has shifted investment toward the Homestead rather than the railroad pattern. Only the 1912 differences are statistically significant in this county, but Table 5 shows that they are significant in the full sample.

Table 5 also reinforces these findings with a broader set of investment measures. Columns (1)–(3) report on the extensive margin of land improvement (i.e., positive improved land acres) and columns (4)–(6) on the value of assessed value of fixed improvements (i.e., farm buildings).[13] Both outcomes tell the same story. Land concentration lowered investments

[12] 5% of sections with this quality grow crops, whereas 63% in the other four quintiles do. For a non-parametric version of this point, see Appendix Figures A.2-A.4.

[13] The assessed total values in 1912 were, unfortunately, only computed using a single rate each for improved

Figure 4: Land Improved by Settlement Type (archival)



(a) Improved Land and Soil Quality, 1912 **(b) Improved Land and Soil Quality, 2017**

Notes: The figure shows the section-level existence of improved land for even and odd sections as a function of gSSURGO soil quality (percentile of full sample). Improved land derives from 1912 property assessments, Morrill County, Nebraska (panel (a)) and 2017 land use (CropScape) with 1% or more land in crops (panel (b)). The samples are consistent, considering only sections assessed in 1912.

substantially in 1912, with 24 percentage point fewer sections improving any land and fixed improvement falling by 77%. In 2017, the differences are substantially smaller and only statistically significant in the full sample. While the historical case study area is much smaller than the full modern sample, the results have the same sign and are not statistically distinguishable from the full sample's.

6.2.1 Could Fewer Improvements be Better?

Three aspects of these results cut against the idea that fewer improvements simply reflect larger owners economizing on fixed costs. First, property values reflect expected profits, and these are lower in historically concentrated lands. Second, while value-based measures of improvements potentially incorporate fixed costs, the same pattern of results is also present in measures based on acreage usage that do not. Both the extensive margin of improved acreage and expected profits based on land use decrease (Tables 5 and 6). Appendix Table A.4 finds similar reductions across many other functional forms, including improvement value per owner. Third, land concentration historically reduced crop farming most in the best, rather than worst, croplands (Figure 4a), inconsistent with comparative advantage. Table

and unimproved land compared to the more detailed procedures used today. As such, improved land alone captures all differences in 1912 assessed value.

Table 5: Land Investments

	Any Improved Land (%)			(asinh) Improvement Value		
	(1)	(2)	(3)	(4)	(5)	(6)
	1912 Sample	1912 Sample (in 2017)	Full Sample (in 2017)	1912 Sample	1912 Sample (in 2017)	Full Sample (in 2017)
RR Effect	-24.3**	-5.67	-1.48***	-0.77**	-0.25	-0.23***
	(7.76)	(6.37)	(0.42)	(0.26)	(0.42)	(0.047)
Grant × State FEs	Y	Y	Y	Y	Y	Y
County FEs	Y	Y	Y	Y	Y	Y
Township FEs	Y	Y	Y	Y	Y	Y
Geo Controls	Y	Y	Y	Y	Y	Y
SEs / Clusters	Township	Township	Spatial	Township	Township	Spatial
N	101	101	132,462	101	101	132,463
$\mathbb{E}[y]$	26%	42%	48%	\$3.2k	\$15k	\$1,277k

Notes: The table shows even/odd comparisons per equation (1) covering the extensive margin of land improvement in columns (1)–(3), as in Figure 4, and the assessed value of improvement in columns (4)–(6). The samples used are Morrill County, Nebraska in 1912 for (1) and (4); the same sections in 2017 for (2) and (5); and the full 2017 data for (3) and (6).

A.4, Panel B similarly shows that the effects on land value and use intensity are small in the bottom quintile of soil productivity and are instead driven by the better quintiles. Large-scale owners thus invested less according to many measures rather than making the most out of a fixed number of improvements.

6.3 Other Inputs and Land Use

Table 6 shows that historic land concentration reduced observable labor and capital inputs, consistent with lower intensity use. Although I lack survey-based microdata on most farms, administrative data directly or indirectly capture both variables. For labor, I note that farm operators typically reside on their properties, especially in the 19th and early 20th centuries. The resident population of a section thus approximates available labor per square mile. Using 1940 census maps for Nebraska, I find that land concentration lowered the number of rural farms[14] by 25%; the satellite-measured population is still lower by 16% in 2019 (9.1% for the whole sample). For capital, archival data from property assessments in one Nebraska county in 1965 show 26% less capital (farm tools and equipment).

Modern data show that historically concentrated land is still used more homogeneously and less intensively in 2017. Columns (5)–(6) use the USDA's satellite-derived, 30m × 30m pixel data on land use. For agricultural activities, I code an expected profit for each pixel based on current usage, expected yields, prices, and costs as described in Appendix Section B.8. By this measure, historic concentration lowered use-based profits by 2.8%, similar to the

[14]This definition excludes urban households. The modern definition includes them, but only 0.5% of sections contain a town, and the (asinh) transformation makes the regression insensitive to outliers.

Table 6: Land Use and Non-Land Inputs

	(1) (asinh) Farms [1940]	(2) (asinh) 2019 Pop [1940 sample]	(3) (asinh) 2019 Pop [Full sample]	(4) (asinh) Capital [1965]	(5) (asinh) Use Value [2017]	(6) # Uses [2017]
RR Effect	-0.25***	-0.16***	-0.091***	-0.25***	-0.028**	-0.069***
	(0.012)	(0.0089)	(0.014)	(0.063)	(0.014)	(0.023)
Grant × State FEs	Y	Y	Y	Y	Y	Y
County FEs	Y	Y	Y	Y	Y	Y
Township FEs	Y	Y	Y	Y	Y	Y
Geo Controls	Y	Y	Y	Y	Y	Y
SEs / Clusters	Spatial	Spatial	Spatial	Township	Spatial	Spatial
N	18,622	18,622	132,463	2,084	132,462	132,462
$\mathbb{E}[y]$	2	18	23	$13k	$380k	4.2

Notes: The tables shows even/odd comparisons per equation (1) covering (asinh) number of farms in 1940 in Nebraska in column (1); (asinh) satellite-derived populations in 2019 for the 1940 and full sample in columns (2)–(3); assessed farm equipment in 1965 in Lincoln County, Nebraska in column (4); (asinh) values based on land use, per Section B.8 in column (5); and the number of distinct land uses per CropScape in 2017 in column (6).

loss in land values. Finally, column (6) counts the number of distinct economic uses.[15] Odd sections exhibit fewer distinct land use choices, indicating homogeneity. This homogeneity likely reflects a mix of persistent ownership concentration and persistent patterns of land use.

Together, these results suggest that historically, land concentration led to a group of owners who were unable to improve and take full advantage of their properties' potentials. While the effect of concentration has attenuated over time and modern differences amount to 23%–64% of their historical magnitudes, the differences are still large enough to lower property values by 4.5% in 2017. In the next section, I explore in further detail why larger landowners were discouraged from investing.

7 Mechanisms: Tenancy Contracts

In this section, I examine the relationship between the persistently low investment by railroad landowners and the agency problems associated with tenant contracts. Using archival survey data, I explore whether historic land concentration led to increased rates of tenant farming. Classical economists believed that share agreements were especially prone to agency problems compared to cash rent. I test this theory with a heterogeneity analysis to determine if land concentration had more negative effects in counties where share tenancy was more common.

[15]Defined as a separate CropScape coding of a particular crop, pasture, or urban/developed area.

7.1 Tenant Contracts: Direct Measures

I start by establishing that land concentration causally increased the use of tenant farming over a long time span. Geocoded information on tenant contracts is rare for the post-railroad grant era due to the loss of most of the agricultural census microdata and the latter's lack of property information. I thus present evidence from two local surveys: Kansas county census documents from 1940 and personal assessments and grain taxation from one Nebraska county in 1965. In both surveys, the property location, owner, and operator are given. When the owner and operator differ, I code the property as tenanted. In the 1965 survey, operator and landlord output shares are also reported, and I code properties with a positive landlord share of output as under a share agreement. Because of this survey's focus on crop tax liability, share contracts account for a large majority of contracts reported; see Appendix B.

Table 7 estimates the increase in tenant contracts across each survey and contract definition. Column (1) pools both surveys and shows a 7.0 percentage point increase of a section having a tenant operator, roughly a 24% increase from the Homestead (even-section) mean. Columns (2)–(3) replicate this result for each survey individually, and column (4) replicates it for share contracts in the 1965 sample. The relative increases are similar in magnitude, with column (4) representing a 18% increase in rates of share contracting from the homestead mean. Columns (3)–(4) are significant at the 10% level, while columns (1)–(2) are significant at the 1% level.

Table 7 uses additional archival evidence to show that odd-section owners were less likely to live near their properties, indicating they were absentee landowners. I link owner names in the early 1900s Nebraska assessments to census microdata as described in Appendix B.5. Overall, odd-section owners are 8 percentage points less likely to be linked to someone in their property's county, suggesting absenteeism. While not all owners match to someone in the census, my overall success rate is similar to those in other analyses[16] and is unlikely to be differential across even and odd sections. Column (6) considers the percentile rank of owners' last name in the census. While odd-section owners have slightly less common names by about 1 percentile, this difference is not statistically significant, unlike the county-specific match difference. Railroad owners' names are typically in the 94th percentile of commonness of all census names, whereas Homestead owners' are typically in the 95th.[17]

The duration of the tenancy effect, rather than its sign, is the most notable result. Because the Homestead Act deliberately selected owners who would personally work smaller

[16]I uniquely match 39% of owner names to the census; a subset of these matches are located within the property's county. For comparison, Abramitzky, Boustan and Eriksson (2014) achieve a 16% match rate of people across census years. Their process differs from mine, however, as it draws upon a rich set of census information, whereas I can only consider the names of individuals and counties where their properties are located.

[17]This calculation excludes corporate ownership, where last names are not well defined.

Table 7: Tenancy and Absentee Ownership

	Tenant Agreement				Absenteeism	
	(1) Both Surveys	(2) KS Only	(3) NE Only	(4) Share Contract	(5) Owner In County (%)	(6) Last Name (percentile)
RR Effect	6.99***	10.2***	3.74*	3.74*	-8.24***	-1.13
	(2.41)	(1.62)	(2.03)	(2.02)	(2.57)	(2.21)
Sample Years	1940 1965	1940	1965	1965	1900 1912	1900 1912
Grant × State FEs	Y	Y	Y	Y	Y	Y
County FEs	Y	Y	Y	Y	Y	Y
Township FEs	Y	Y	Y	Y	Y	Y
Geo Controls	Y	Y	Y	Y	Y	Y
SEs / Clusters	Spatial	Spatial	Township	Township	Township	Township
N	1,571	738	832	832	613	581
$\mathbb{E}[y]$	33%	43%	23%	22%	12%	94

Notes: The table shows even/odd comparisons per equation (1) covering initial owners in Lincoln County, Nebraska in columns (1)–(2); owners per archival assessments of Perkins County in 1990 and Morrill County in 1912 (Nebraska) in columns (3)–(4); and 1940 farm surveys in multiple Kansas counties in columns (5)–(6). Columns (5)–(6) report on the links to the census microdata described in Section B.5.

plots of land, it is unsurprising that railroad landowners more frequently employed tenant and share contracts while residing away from their properties. However, the agricultural surveys measuring tenancy studied in Table 7 cover a period roughly 70 to 100 years after the initial land allocation. Based on the attenuation of other effects shown in Section 6, it is likely that the even/odd difference in tenant farming also used to be much larger, though unfortunately I am unaware of any earlier surveys that directly measure tenancy.

The durable increase in tenant arrangements of all kinds, and share tenancy in particular, provides direct evidence for the classical mechanisms of contracting frictions as explanations for the long-term investment changes driven by land concentration. In the next section, I provide further evidence that share contracts in particular explain the results.

7.2 Share Tenancy: Heterogeneity Evidence

Classical economic theory argued that share contracts, as opposed to cash rents, lowered agricultural output. In this section, I investigate whether this dynamic could extend over long periods by influencing land use and investment. Two distinct processes could contribute to this outcome. First, landowners using share agreements might be less inclined to develop their land. Second, landowners anticipating the challenges associated with share arrangements might forgo crop cultivation entirely. Instead, they would favor less intensive forms of agriculture, such as ranching, on unimproved land. As discussed mathematically in Section 3, opting for unimproved land avoids the principal/agent costs of crop share contracts but also leaves the land underutilized.

I make use of the fact that different areas of the United States varied in their propensity to use share contracts rather than cash rent. In my sample counties in 1940,[18] anywhere between 10% and 70% of farms were run by tenants, with share tenants accounting for 0%–91% of this group. If the classic incentive problem accounts for the results, counties prone to share tenancy should be most affected. A long literature in economics has explored why share contracts can remain second-best solutions in equilibrium, proposing factors of resource constraints, adjustment costs, or risk aversion (Banerjee, Gertler and Ghatak, 2002). Geographic characteristics play an important role in this, for example by influencing the stakes of monitoring tenants (Alston and Higgs, 1982).

In support of the classic theories, Figure 5 presents a county-level heterogeneity analysis of equation (1) based on the 1940 fraction of tenant farms using crop share contracts. Counties with high rates of share tenancy experience the largest losses in modern property values in odd-numbered sections; counties where share tenancy is uncommon have few or no negative effects. Combined with the Section 7.1 result that land concentration results in more dependence on tenant farming, the larger effects in counties prone to share tenancy suggest that the incentive problems discussed by Smith and Marshall account for the lower land values today.

I perform several robustness checks on the heterogeneity analysis to rule out other explanations for this result. First, different levels of soil productivity do not explain the pattern. Figure 5b plots the land value effects by soil quality for areas above and below median share tenancy rates. Across all land quality values, the high share tenancy sample shows larger effects. Appendix Table A.5 presents a similar result in regression form, showing that the heterogeneity is not due to state, railroad company, or soil type.

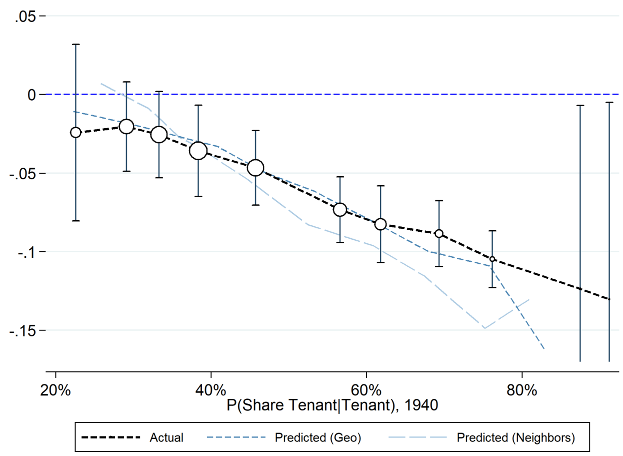
I also consider whether the analysis is complicated by the fact that 1940 is a post-treatment year.[19] Note that since share tenancy prevalence is defined at the county level, individual square sections have very limited contributions to their own prevalence value. In addition, Figure A.6 incorporates two external predictions for prevalence. Using predicted prevalence based on either county-level geographic characteristics or the average value of neighboring counties' share tenancy prevalence, the pattern remains essentially unchanged.

In summary, direct evidence suggests that land concentration led to a long-term increase in rates of tenant farming, including share tenancy. However, the prevalent form of tenancy varied substantially across counties. Consistent with the classic theories emphasizing the incentive problems of share tenancy, counties where tenants mostly used cash rent saw few

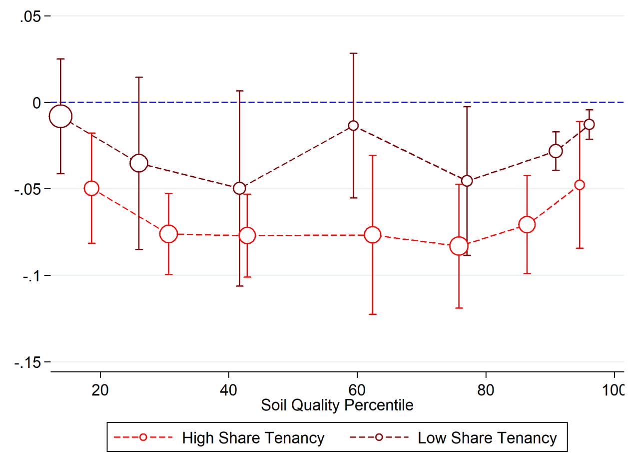
[18]The first agricultural census year in which all my counties are present.

[19]Ideally, a pre-allocation measure of share tenancy prevalence could be used. However, because railroad land was allocated essentially at the start of settlement, no direct pre-allocation data exist.

Figure 5: Effects on Property Values by Share Tenancy and Soil Quality



(a) Land Value Effects by Share Tenancy



(b) Land Value Effects by Soil Quality

Notes: The figure shows even/odd comparisons per equation (1) on (asinh) 2017 land value using samples defined by a county's fraction of 1940 tenants using share agreements and soil quality. Controls are included as in Table 4, column (6). Panel (a) uses the actual 1940 share tenancy rates, rates linearly predicted on all geographic characteristics in Table 4, and rates linearly predicted on the average rate of share tenancy in neighboring counties. Panel (b) uses samples of counties above and below the median rate of share tenancy.

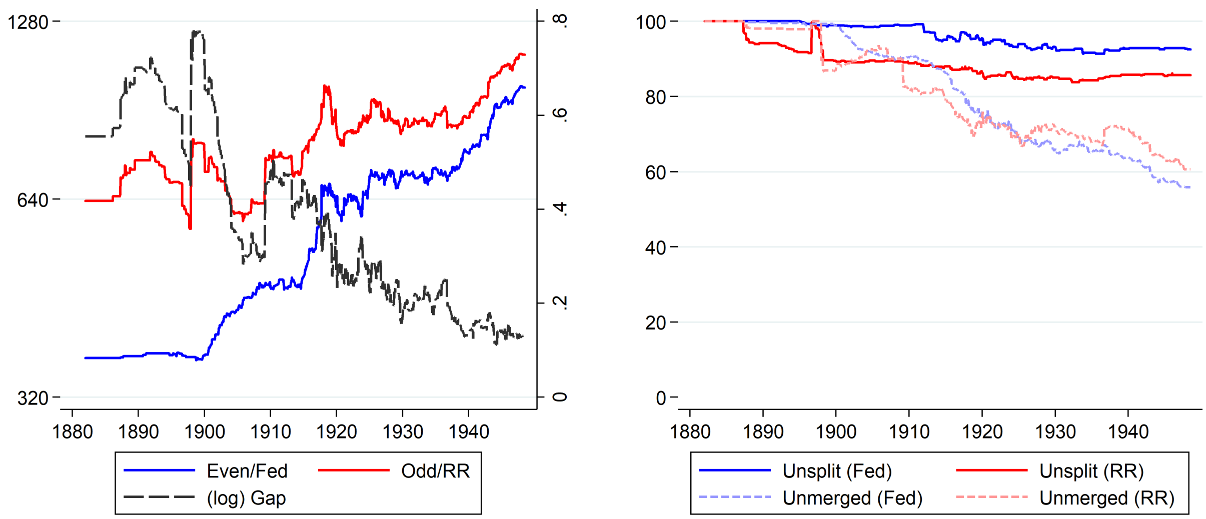
or no long-term impacts from land concentration. In counties where geographic or economic conditions predicted share tenancy would prevail, however, the negative impacts of land concentration were larger and long-lasting. The duration of this result suggests that land markets in this context fell short of the Coasian ideal. In a Coasian world, initial allocations of any sort should not have long-term impacts. In the next section I examine why Coasian reallocation has proceeded slowly.

7.3 Reallocation Speed

Although the even/odd-section difference is still present in modern data, Coasian reallocation did occur over long time spans. Section 6 showed that modern differences in land investment and use range from 23%–64% of their historic size, depending on the outcome. I use a case study of one Nebraska county's individual land transactions to provide more detail about how this process unfolded. The data cover all land transfers from the county's creation in 1882 until 1947 and are available at the parcel level, representing one-sixteenth of a section (40 acres).

The data indicate that convergence is a slow but steady process. Figure 6a plots the average owned property size difference by month for even- and odd-numbered sections. Over the course of 65 years, the odd/even size ratio falls from a high of 2.2 to a low of 1.1. No particular period accounts for this convergence; instead, it occurs as part of a regular trend. Convergence is primarily achieved as even-section properties grow in size faster than odd-numbered ones. The economic history of this period has emphasized that changes in technology such as the tractor made it more realistic for individual families to operate larger farms (Sumner, 2014; Olmstead and Rhode, 2001), and these changes likely favored size increases in even-sections over odd-numbered ones.

Figure 6: Ownership Dynamics (Archival Sample)



(a) Land Concentration (Acres, Log Scale) **(b) Property Retention (%)**

Notes: The figure shows monthly data on land transfers from Banner County, Nebraska. Panel (a) plots the average property sizes, and panel (b) considers similarity to initial 1882 properties, defined as "unsplit" (largest portion of initial property with a single current owner) or "unmerged" (largest portion of current property owned by a single 1882 owner).

Perhaps surprisingly, odd-section properties are rarely split into smaller pieces. With an ideal market, odd-section owners of full squares could have split them into four 160-acre parcels to mirror the Homestead allocation illustrated in Figure 3b. The aggregate property sizes in Figure 6a suggest this did not occur since odd-section property sizes rarely shrink. More detailed data graphed in Figure 6b confirm this intuition by comparing properties in each month to their initial boundaries in 1882. Over the 65-year span of data, about 86% of the odd-numbered section parcels remain within the same 1882 property, indicating that land is mostly transferred intact to the next owner. However, any splitting that occurs is typically complete by 1900, and reallocation instead happens due to parcels being combined over time. As a result, by the end period, the typical odd-numbered section property can only trace 61% of its area to a single 1882 property.[20] This fraction shrinks over time as

[20]For example, an 1882 property split into three pieces of 70%, 20%, and 10% of its original size would be considered 70% "unsplit," and each piece would be 100% "unmerged." If a 100-acre farm in 1900 was

more land is combined into larger holdings. Since dividing large properties would have been the most direct way to reverse land concentration, its rarity points to the constraints faced by small owners in obtaining property.

The lack of complete convergence by the end of the sample period cannot be attributed to any formal legal barrier. Both Homestead and railroad landowners received full titles, and the land market was thick enough to facilitate frequent transfers. Appendix Figure A.7 shows that, in these data, even and odd sections were settled around the same time and nearly all properties changed hands at least once, often more, by 1920. Appendix Section C.7 delves further into this topic and presents data demonstrating that Homestead owners had more limited access to capital, reflecting the Act's goal of providing land to those who could not afford it. Still, other factors may have contributed to the incomplete convergence.[21] As with the original Coase Theorem, many kinds market imperfections would yield the same key outcome of a persistent initial allocation.

8 Aggregate Effects

To determine whether the railroad grant policy had broader impacts beyond the direct comparisons outlined in equation (1), I investigate its effects at higher levels of aggregation. Due to the long time period between the initial allocation and my modern data, there are several reasons why the even/odd comparison may not fully reflect the policy's broader effects. Homestead lands could have obtained resources at the expense of railroad lands, or individual landowners in each group could have obtained land in the other type of square over time. To address this question, I compare non-railroad land in the checkerboard area to land outside the checkerboard. I exploit the fact that the grant boundaries were arbitrarily determined by formula and use a geographic RD around those borders.

8.1 RD Specification

I use an RD design to compare federal (Homestead) land within the checkerboard areas to those outside it. Because some of the borders of this grant area were determined by political features,[22] I consider only grant boundaries set formulaically by fixed distances to a railroad

formed from three complete properties of 70, 20, and 10 acres, it would be considered 70% unmerged and 100% unsplit.

[21]For important work on this topic, see Bleakley and Ferrie (2014).

[22]For example, the Osage Reservation (Kansas) and Crow (Montana) Reservation borders are excluded; see Appendix Section B.2 for a complete discussion.

tracks in this analysis. I run regressions of the form

$$y_i = \alpha[\text{NearRR}]_i + f(d_i) + X_i\beta + \varepsilon_i, \tag{2}$$

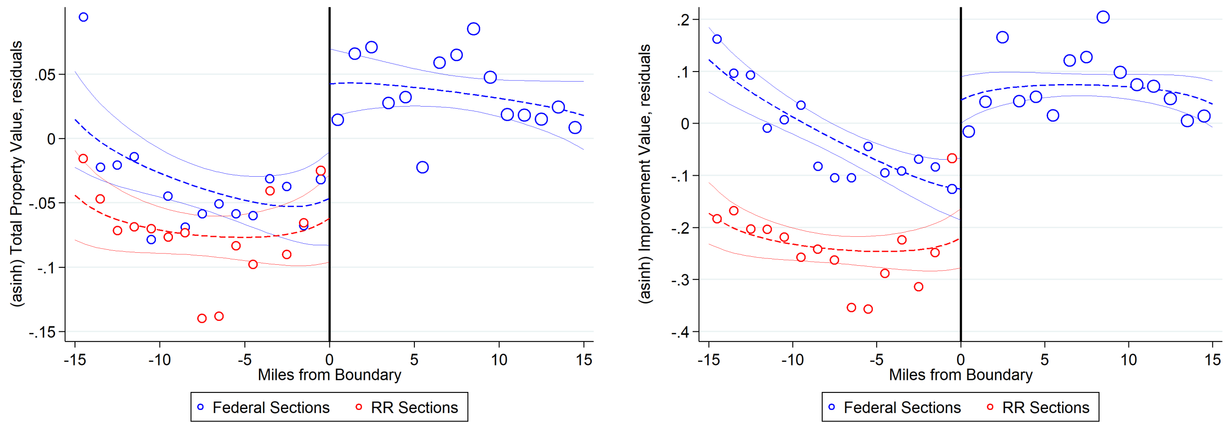
where i is a non-education, federal section within a bandwidth of the boundary; $[\text{NearRR}]_i$ is a dummy for a section being located within the railroad land grant; d_i is the distance to the boundary; f is a local linear function on either side of the cutoff, estimated separately for each state × railroad grant pair; and X_i are controls including county fixed effects, the geographic controls from the even/odd specification, and latitude and longitude by state. For my baseline results, I estimate this equation by the Calonico, Cattaneo and Titiunik (2014) methodology with a triangular kernel while clustering at the county level. For consistency, I use the optimal bandwidth for the regression in 6.47 miles across all regressions. I explore robustness to the sample and bandwidth in Appendix Figure A.8 and Table A.6.

8.2 The Checkerboard Effect

Figure 7 indicates that the aggregate effect of the checkerboard policy was, if anything, more negative than what the even/odd comparisons suggest. To compare consistently settled areas, I focus on federally administered (Homestead) lands. The total and improvement value of these lands in 2017 sharply increase at the border, indicating that Homestead sections within the checkerboard lost value despite not being allocated to railroad owners. Combining these results with the even/odd comparisons, the best-performing groups are, in order, pure Homestead lands, Homestead lands within the checkerboard, and railroad lands. These groups correspond to pure control areas, control units in treated areas, and treated units, considering railroad allocations as a "treatment."

I examine potential reasons for a negative checkerboard effect in Section 9.3 and Appendix Table A.6. The effects of Figure 7 are statistically significant and not due to differences in geographic characteristics at the border, which remain small and tightly estimated. A slightly higher fraction of federal land is settled, but settlers linked to census microdata also appear relatively similar. One possible explanation is that large-scale owners with property in the checkerboard area formed an important pool of buyers for lands adjacent to their sections, outbidding credit-constrained small-scale owners. This type of purchase across units of treatment would have reduced even/odd differences within the grant area, leading to an understatement of the policy's aggregate effects.

Figure 7: RD Graphs for (asinh) Total and Improvement Value



(a) Total Value **(b) Improvement Value**

Notes: The figure shows modern data on (asinh) total and improvement value near the checkerboard border. Railroad (odd-numbered sections within the checkerboard) land is depicted in red, and federal land (all other land) is in blue. Outcomes are adjusted for the controls listed in Section 5.2.

9 Alternative Mechanisms

In this section, I examine alternate mechanisms that might explain the long-run differences in even and odd sections. I specifically consider non-agricultural components of land value, other elements of the railroad grant policy, other elements of the Homestead Act, and land fragmentation. I also consider whether the results applied broadly or were limited to particular subsets of my data. Further channels are addressed in Appendix C, including public goods, the timing of settlement, and speculative investment.

9.1 Urban Density and Agglomeration

One major alternative story would be that differences in non-agricultural land explain my results. Since the Homestead Act was designed to increase settlement, it could have fostered the creation of towns and cities whose per-acre land values vastly exceed those of farms. I address this explanation below. Because the unit of observation is a land area, agriculture is overwhelmingly the dominant economic use, and land for towns and cities covers only 0.4-3% of the sample depending on the definition. This small part of the sample is balanced across even and odd sections and, based on a number of analyses, has a very limited impact on the final result in either direction.

Table A.7, Panel A shows that differences in the number of towns per square-mile section are small and tightly estimated according to multiple definitions.[23] Effect estimates in

[23] I consider both sections with positive area in a Census Designated Place (CDP) and towns defined as

columns (1)-(2) are one town per thousand square miles or less. Neither is significant at the 5% level though (2) is significant at the 10% level. Both coefficients are positive, which, if anything, would suggest odd sections benefited from having slightly more towns. Similarly, there are tight zeros on sections having 100+ or 1000+ people. Even sections are persistently more likely to have higher farm density, however. They have increased densities at 1+ or 10+ people per square mile, or approximately 2-3 households.

Panel B explores robustness of the main results on land value to different checks intended to limit the role of towns. Column (1) is the baseline result. Column (2) top codes values to the 5$^{\text{th}}$ percentile, primarily affecting high-value urban areas. Column (3) removes areas within 2 miles of the railroad track, reflecting the high density of towns located near railroads. Columns (4)-(6) drop sections with towns,[24] using different definitions. All of these leave the main result essentially unchanged. Overall, urban areas account for relatively few observations that do not explain the main result.

I address a related question about whether the density of farms in a square mile could have agglomeration or other spillover benefits in Section C.2.

9.2 Provisions of the Homestead Act

Another explanation would be that other elements of the Homestead Act induced settlers to intensively invest. For example, if settlers were required to grow crops to receive a title, my results could simply reflect legal compulsion. Some historical sources suggest that the Act's farming requirement was implemented this way, with settlers needing to "cultivate some portion of the land" (Bradsher, 2012).

However, official regulation interpreting the Homestead Act portrays a more flexible policy where settlers were free to choose their form of agriculture. In the 1880 Luning decision, the Department of the Interior explicitly addressed the crop requirement. It ruled that crop farming was not required as "stock-raising and dairy production are so nearly akin to agricultural pursuits as to justify the allowance of entry [i.e., title]" (Department of the Interior, 1880; General Land Office, 1884).[25] Bradsher (2012) also clarifies that "no specified amount of cultivation or improvements was required." That is, settlers were not required to do any specific activity, but rather demonstrate that they intended to settle and work their land in "good faith."

individual points in Schmidt (2018). Both definitions use the year 2000 as the date.

[24] Because towns could be a downstream outcome, these estimates should be treated with caution. However, because the dropped observations are a small and balanced part of the sample, any bias is likely small. It is also reassuring that none of the results differ much from the baseline.

[25] 1880 precedes a large majority of settlement in my sample. Figure A.9 shows similar effects occur in areas settled before and after 1880.

Two pieces of quantitative evidence also suggest a flexible Homestead policy. First, if Homesteaders were compelled to generate (assessed) improved land, Homestead grants should be rare in areas without improved acres. Instead, among eligible sections in my 1912 archival data, 91% of lands without assessed improvements were successfully Homesteaded; 93% of these were obtained for free.[26] These figures provide a lower bound for approval as the small portion of non-Homestead land in these areas could have been settled under other policies for reasons unrelated to the level of improvements. Even alone, the lower bound is quite high and is most consistent with a flexible Homestead policy. In my data, settlers who made no improvements (per the assessor) were not typically, if ever, denied a title.

Second, the overall pattern of my results is not consistent with compulsion. If settlers felt compelled to grow crops in unsuitable areas to comply with the law, the biggest even/odd differences should occur in the least suitable lands. Instead, in Figures 4 and A.3, the opposite pattern emerges. The land value and investment differences are smallest in the low quality lands and largest in lands well-suited for farming. Compulsion to grow crops should have been minimal in productive areas since economic considerations would lead settlers to do so anyway. Figure 4a shows that even sections historically did indeed grow more crops in productive areas, but odd-section investment was low even on high quality lands.

The Act's focus on farming was clearly important, especially for establishing the first stage in this study. Without it, there was no guarantee Homesteaders would be small-scale farmers rather than people simply interested in a free plot of land. However, the best evidence from my archival data and regulations clarifying the law's implementation is that settlers could pick the kind of farming they practiced. As such, even sections' increased rates of the specific (assessed) improvements I analyze were likely due to other factors.

9.3 Other Railroad Grant Area Policies

If not elements of the Homestead Act, perhaps other aspects of railroad companies' sales could be the primary mechanism. Notably, Alston and Smith (2022) details the legal troubles of the Northern Pacific Railroad (NPRR). The company's "violations, controversies, and investigations... had no peers" and this plausibly created uncertainty for those who bought NPRR land (Draffan, 1998). More broadly, it could be that the actions of any particular company explain the results rather than the common theme of land concentration.

Appendix Figures A.2-A.4 show that even/odd differences point in the same direction across a large majority of state and land quality subsamples. Odd sections broadly exhibit

[26] 79% were transferred under a typical Homestead arrangement and 5% under the 1873 Timber Act extension (granting Homesteaders free land for raising trees). 7% were sold under the Homestead Act (allowed after 6 months of residence and "improvements"). 9% were sold outside the Homestead Act. These percentages are 87%, 5%, 4%, and 5% for the whole sample of eligible sections.

increased parcel size, decreased rates of crop farming, and decreased land values. Figure A.9 also shows that the land value differences are similar across counties settled over a span of five decades, implying no particular year or event drove the result.[27] The consistent pattern in odd sections across a range of states, companies, and time periods indicates the need for a similarly broad mechanism. As noted earlier, the "basically identical" policies of railroad companies consistently led to large-scale sales (Greever, 1951). A corollary of this analysis is that the legal troubles specific to the NPRR do not explain the results. Dropping the grant, in fact, slightly strengthens the land value differences; see Appendix Section C.3 for more discussion on this issue.

A separate concern is that the Homestead Act was implemented differently with smaller, 80-acre allotments in railroad grant areas. This concern affects the RD interpretation and is discussed in Appendix Section C.4.

9.4 Land Fragmentation

One peculiar feature of the railroad checkerboard is that exceptionally large landowners initially had to be content with diagonally connected sections (640 acres). In developing countries, fragmentation is thought to negatively impact agriculture by inhibiting scale economies or increasing workers' travel times, though in practice such costs may be modest (Ali, Deininger and Ronchi, 2019). While 640 acres is far larger than farms found in most developing countries and diagonal connections are less likely to impede travel times, the costs associated with fragmentation could be responsible for underinvestment in my setting.

However, in practice, odd sections were not particularly fragmented on diagonals and most owners in my archival data were able to obtain land connecting such holdings. Appendix Figure A.10b shows that both odd- and even-section properties retained over 90% of their contiguous sizes when removing diagonal connections. Odd sections were also part of larger holdings regardless of the metric used. Appendix Figure A.10 shows measures of total holdings, comparing total acres owned, contiguous acres owned, and non-diagonally contiguous acres. While measuring size contiguously mechanically lowers property sizes, odd sections are larger across all measures, and the dynamics across time are also similar.[28] Odd-section owners thus had greater potential for scale economies regardless of how size is defined.

Finally, it is important to note that a section of 640 acres was already at the upper end of operational sizes in the late 1800s. As noted in Section 2.3, a 640-acre farm would have been

[27] None of this is intended to rule out heterogeneity. Rather, it suggests that the effect's direction is consistent, with potentially different magnitudes.

[28] Similar statistics hold in other archival data: in the early 1900s assessments, odd-section properties are 70% larger than even-section ones even when diagonal connections are ignored.

in the top 2% largest farming units in my sample states and the US overall. While many owners benefited from larger holdings, in practice these were operated separately according to the census.

10 Conclusion

Whether land is allocated to smallholders or concentrated among large landowners is a distinguishing feature of many agrarian economies. In this paper, I use a natural experiment to study the impact of that distinction on the American frontier. The effects of land concentration on economic development were primarily negative, with larger owners investing less in their properties due to the agency problems of tenant farming and absentee ownership. Although the effects of land concentration have partially faded over time, the process took many decades and differences in land use and value are still apparent 150 years after the initial allocation. This incomplete convergence indicates American land markets historically fell short of the Coasian ideal but may yet achieve it over even longer horizons.

The results in this paper speak to the concerns classic economists like Smith and Marshall expressed about land concentration. When landowners are constrained in their contract choice, they may adopt second-best share contracts that impede incentives. The evidence in this paper suggests that these agency problems can have long-term impacts via investment and land use. As such, the allocation of important assets like land is not only relevant to contemporary well-being, but can also shape economic development in subsequent generations.

References

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. 2014. "A nation of immigrants: Assimilation and economic outcomes in the age of mass migration." *Journal of Political Economy*, 122(3): 467–506. 24

Acemoglu, Daron, Suresh Naidu, Pascual Restrepo, and James A Robinson. 2019. "Democracy does cause growth." *Journal of Political Economy*, 127(1). 46

Adamopoulos, Tasso, and Diego Restuccia. 2019. "Land reform and productivity: A quantitative analysis with micro data." National Bureau of Economic Research. 4

Ahola, Jason. 2013. "Estimating carrying capacity: How many cows can I graze?" 45

Ali, Daniel Ayalew, Klaus Deininger, and Loraine Ronchi. 2019. "Costs and benefits of land fragmentation: evidence from Rwanda." *The World Bank Economic Review*, 33(3): 750–771. 34

Allen, Douglas W, and Bryan Leonard. 2021. "Property right acquisition and path dependence: Nineteenth-century land policy and modern economic outcomes." *The Economic Journal*, 131(640): 3073–3102. 5

Allen, Robert C. 1988. "The growth of labor productivity in early modern English agriculture." *Explorations in Economic History*, 25(2): 117–146. 1, 4

Alston, Eric, and Steven M Smith. 2022. "Development Derailed: Policy Uncertainty and Coordinated Investment." *The Journal of Law and Economics*, 65(1): 39–76. 5, 33, 48

Alston, Lee J, and Robert Higgs. 1982. "Contractual mix in southern agriculture since the Civil War: Facts, hypotheses, and tests." *Journal of Economic History*, 327–353. 4, 26

Banerjee, Abhijit V, Paul J Gertler, and Maitreesh Ghatak. 2002. "Empowerment and efficiency: Tenancy reform in West Bengal." *Journal of political economy*, 110(2): 239–280. 1, 4, 9, 26, 40

Bazzi, Samuel, Martin Fiszbein, and Mesay Gebresilasse. 2017. "Frontier Culture: The Roots and Persistence of "Rugged Individualism" in the United States." National Bureau of Economic Research. 5

Benjamin, Dwayne. 1995. "Can unobserved land quality explain the inverse productivity relationship?" *Journal of Development Economics*, 46(1): 51–84. 4

Besley, Timothy, and Robin Burgess. 2000. "Land reform, poverty reduction, and growth: Evidence from India." *The Quarterly Journal of Economics*, 115(2): 389–430. 4

Bigelow, Daniel P, Jennifer Ifft, and Todd Kuethe. 2020. "Following the market? Hedonic farmland valuation using sales prices versus self-reported values." *Land Economics*, 96(3): 418–440. 13

Bleakley, Hoyt, and Joseph Ferrie. 2014. "Land openings on the georgia frontier and the coase theorem in the short-and long-run." 1, 4, 29, 50

Borchers, Allison, Jennifer Ifft, and Todd Kuethe. 2014. "Linking the price of agricultural land to use values and amenities." *American journal of agricultural economics*, 96(5): 1307–1320. 12

Bradsher, Greg. 2012. "How the West Was Settled." *The US National Archives and Records Administration*. 6, 32

Burchardi, Konrad B, Selim Gulesci, Benedetta Lerva, and Munshi Sulaiman. 2018. "Moral hazard: Experimental evidence from tenancy contracts." *The Quarterly Journal of Economics*, 134(1): 281–347. 4, 18

Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik. 2014. "Robust data-driven inference in the regression-discontinuity design." *The Stata Journal*, 14(4): 909–946. 30

Census of Agriculture. 1880. "United States Census of Agriculture." 2, 8

Cheung, Steven. 1969. "The theory of share tenancy." *The theory of share tenancy.* 1

Coase, Ronald H. 1960. "The problem of social cost." *Journal of law and economics*, 1–44. 1

Coffin, LS. 1902. "Breaking prairie." *The Annals of Iowa*, 5(6): 447–458. 20

Conley, Timothy G. 2010. "Spatial econometrics." In *Microeconometrics*. 303–313. Springer. 15

Decker, Leslie Edward. 1964. *Railroads, lands, and politics: the taxation of the railroad land grants, 1864-1897.* Brown University Press. 6

Dell, Melissa. 2010. "The persistent effects of Peru's mining mita." *Econometrica*, 78(6): 1863–1903. 1, 46

Department of the Interior. 1880. "Homesteads – T.W. Luning." *Copp's Landowner*, 7(9): 135. 32

Desiere, Sam, and Dean Jolliffe. 2018. "Land productivity and plot size: Is measurement error driving the inverse relationship?" *Journal of Development Economics*, 130: 84–98. 4, 12

Dippel, Christian, Dustin Frye, and Bryan Leonard. 2020. "Property Rights without Transfer Rights: A Study of Indian Land Allotment." National Bureau of Economic Research. 4

Donaldson, Dave, and Richard Hornbeck. 2016. "Railroads and American economic growth: A "market access" approach." *The Quarterly Journal of Economics*, 131(2): 799–858. 42

Draffan, George. 1998. "Taking Back Our Land." *Transitions*, 12: 5–31. 33, 48

Dubois, Pierre. 2002. "Moral hazard, land fertility and sharecropping in a rural area of the Philippines." *Journal of Development Economics*, 68(1): 35–64. 1

Edwards, Richard. 2008. "Why the Homesteading Data Are So Poor (and What Can Be Done about It)." *Great Plains Quarterly*, 28(3): 181–190. 6

Ellis, David Maldwyn. 1946. "The Forfeiture of Railroad Land Grants, 1867-1894." *The Mississippi Valley Historical Review*, 33(1): 27–60. 6, 48

Ely, Richard, and George Wehrwein. 1940. *Land Economics.* The Macmillan Company. 8

Eswaran, Mukesh, and Ashok Kotwal. 1985. "A theory of contractual structure in agriculture." *The American Economic Review*, 75(3): 352–367. 4

Foster, Andrew D, and Mark R Rosenzweig. 2022. "Are There Too Many Farms in the World? Labor Market Transaction Costs, Machine Capacities, and Optimal Farm Size." *Journal of Political Economy*, 130(3): 636–680. 4

Gates, Paul Wallace. 1936. "The homestead law in an incongruous land system." *The American Historical Review*, 41(4): 652–681. 6, 49

Gates, Paul Wallace. 1941. "Land policy and tenancy in the prairie states." *The Journal of Economic History*, 1(1): 60–82. 50

Gates, Paul Wallace. 1942. "The role of the land speculator in western development." *The Pennsylvania Magazine of History and Biography*, 66(3): 314–333. 11, 46

Gates, Paul Wallace. 1954. "The Railroad Land-Grant Legend." *The Journal of Economic History*, 14(2): 143–146. 49

General Land Office. 1884. "Circular from the General Land Office showing The Manner of Proof to Obtain Title to Public Lands Under the Pre-Emption, Homestead, and Other Laws." 32

Ghatak, Maitreesh, and Sanchari Roy. 2007. "Land reform and agricultural productivity in India: a review of the evidence." *Oxford Review of Economic Policy*, 23(2): 251–269. 1, 4

Greever, William S. 1951. "A Comparison of Railroad Land-Grant Policies." *Agricultural History*, 25(2): 83–90. 6, 34

Grekousis, George. 2020. *Spatial analysis methods and practice: describe–explore–explain through GIS.* Cambridge University Press. 15

Grimes, Waldo Ernest. 1919. *Farm Leases in Kansas.* Agricultural Experiment Station, Kansas State Agricultural College. 11

Hagerty, Nick. 2022. "The Scope for Climate Adaptation." 4, 18

Hornbeck, Richard. 2010. "Barbed wire: Property rights and agricultural development." *The Quarterly Journal of Economics*, 125(2): 767–810. 4

Hornbeck, Richard, and Suresh Naidu. 2014. "When the levee breaks: black migration and economic development in the American South." *American Economic Review*, 104(3): 963–90. 4

Iwanowsky, Mathias. 2019. "Property Rights, Resources, and Wealth: The Public Grazing Solution in the United States." mimeo. 4

Kelly, Morgan. 2019. "The standard errors of persistence." 15

Libecap, Gary D, and Dean Lueck. 2011. "The demarcation of land and the role of coordinating property institutions." *Journal of Political Economy*, 119(3): 426–467. 4

Markevich, Andrei, and Ekaterina Zhuravskaya. 2018. "The economic effects of the abolition of serfdom: Evidence from the Russian Empire." *American Economic Review*, 108(4-5): 1074–1117. 4

Marshall, Alfred. 1890. "Principles of economics." 1

Mattheis, Ross, and Itzchak Tzachi Raz. 2023. "There's No Such Thing As Free Land: The Homestead Act and Economic Development." 5, 18

Montero, Eduardo. 2022. "Cooperative property rights and development: Evidence from land reform in El Salvador." *Journal of Political Economy*, 130(1): 48–93. 4

Naidu, Suresh. 2010. "Recruitment restrictions and labor markets: Evidence from the postbellum US South." *Journal of Labor Economics*, 28(2): 413–445. 4

Olmstead, Alan L, and Paul W Rhode. 2001. "Reshaping the landscape: the impact and diffusion of the tractor in American agriculture, 1910–1960." *The Journal of Economic History*, 61(3): 663–698. 4, 28

Rajan, Raghuram, and Rodney Ramcharan. 2015. "The anatomy of a credit crisis: The boom and bust in farm land prices in the United States in the 1920s." *American Economic Review*, 105(4): 1439–1477. 12

Raz, Itzchak Tzachi. 2021. "Learning is Caring: Soil Heterogeneity, Social Learning and the Formation of Close-knit Communities." 5

Reid, Joseph D. 1977. "The Theory of Share Tenancy Revisited–Again." *Journal of Political Economy*, 85(2): 403–407. 4

Schaede, Ursina. 2021. "The long run impact of funding for public education: Evidence from Illinois land grants." 7

Schmidt, Benjamin. 2018. "Creating Data: The Invention of Information in the nineteenth century American State." *http://creatingdata.us*. 32, 45, 69

Smith, Adam. 1776. "The wealth of nations." 1

Sumner, Daniel A. 2014. "American farms keep growing: Size, productivity, and policy." *Journal of Economic Perspectives*, 28(1): 147–66. 4, 28

Swift, Fletcher Harper. 1911. *A history of public permanent common school funds in the United States, 1795-1905.* H. Holt and Company. 7

USDA. 1923. *Farm Ownership and Tenancy.* Yearbook of the Department of Agriculture. 11

Winters, Donald L. 1974. "Tenant farming in Iowa, 1860-1900: a study of the terms of rental leases." *Agricultural History*, 48(1): 130–150. 4

Appendices For Online Publication Only

A Model Proofs

A.1 Setup and Arguments

For simplicity, I assume a monetary-equivalent quadratic effort costs of $\frac{1}{2}ce^2$ (operator) and $r > 0$ for improvements (owner).[29] Since effort is unobservable, owners of improved land are limited to offering payments of h in a high state and l in a low state. Tenants have access to resources W and consequently landlords face a limited liability constraint are $h, l \geq -W$. For simplicity, the tenant's outside option is 0. The timing of the game is as follows: the owner chooses whether to invest, non-operator owners offer a take-it-or-leave-it contract specifying h and l, the operator chooses a level of effort, agricultural output is realized and any contracts implemented, and the owner sells the land along with any investments. The only relevant prices are thus the expected values for unimproved and improved land, p_0 and p_1 respectively.

Applying the same proofs as Banerjee, Gertler and Ghatak (2002), first-best (surplus maximizing) effort is $\frac{A}{c}$. a share contract equally splitting output occurs when $W \leq \frac{1}{8}\frac{A^2}{c}$ inducing suboptimal effort $\frac{A}{2c}$, a cash rent with first-best effort occurs when $W \geq \frac{1}{2}\frac{A^2}{c}$, and an intermediate contract with effort equal to $\sqrt{\frac{2W}{c}}$ in other cases. Denote by $Y_S = \frac{A^2}{2c}$ the small-holder's payout from one period of working improved land and Y_L the large-holder's payout from one period of working improved land. By the above $\frac{A^2}{4c} \leq Y_L \leq \frac{A^2}{2c}$ where the extremes are realized by a share and cash rent scenario respectively.

A.2 Equilibrium Characterization

The equilibrium is characterized by expected land prices p_0, p_1 for unimproved and improved land respectively. In the case of improved land, each owner type values the investment and works it to the best of their ability. In the case of unimproved land, owners must decide whether to invest. In both cases, land will be valued as the benefit of one period of production, a discounted sales price, and will be purchased by the buyer who values it most adjusted for transaction costs. With $f_S \sim F_S$ denoting small owners' transaction cost,

[29] Up to shifting values of limited liability and the outside option, the costs could equally be borne by the operator.

$$p_0 = \mathbb{E}\left[\max\left(1 + \beta p_0 - f_S, Y_S + (1-\delta)\beta p_1 + \delta\beta p_0 - r - f_S,\right.\right.$$
$$\left.\left. 1 + \beta p_0, Y_L + (1-\delta)\beta p_1 + \delta\beta p_0 - r\right)\right] \quad (3)$$
$$p_1 = \mathbb{E}\left[\max\left(Y_S + (1-\delta)\beta p_1 + \delta\beta p_0 - f_S, Y_L + (1-\delta)\beta p_1 + \delta\beta p_0\right)\right] \quad (4)$$

where the four cases in equation (3) represent a smallholder's purchase (not investing/investing respectively) and a largeholder's purchase (not investing/investing) of unimproved land. The two cases in equation (4) represent a small- and large-holder's purchase of improved land, accounting for possible depreciation.

A.3 Markovian Convergence

Agents do not interact with each other over multiple periods, so the model's state reduces to whether land is improved just before transaction costs are realized. Similarly, the probabilities $q_0 = p(I_t = 1|I_{t-1} = 0)$ and $q_1 = p(I(t) = 1|I_{t-1} = 1)$ completely determine the probabilistic dynamics. For this subsection, I focus on the case where S-type owners choose to upgrade land and L-type owners do not. Since S-type owners have greater one-period profit from improved land, the reverse is not possible. If both owners make the same investment decisions, convergence trivially follows in one period. Finally, I focus on the case where both S and L types have positive probability of buying unimproved land. Existence is given in the parameter values for Figure 2.

Depreciation $\delta > 0$ combined with the case assumption that L types have a positive probability of buying $I = 0$ land and do not improve means that $q_0, q_1 < 1$. Since $\delta < 1$, $q_1 > 0$. Since S types have a positive probability of buying unimproved land and improving it and $\delta < 1$, $q_0 > 0$. The transition matrix

$$T = \begin{bmatrix} 1 - q_0 & 1 - q_1 \\ q_1 & q_0 \end{bmatrix}$$

therefore has all strictly positive entries and is irreducible and aperiodic, meaning convergence as $t = \infty$ to the same probabilistic distribution of investments regardless of initial S-type (with investment) or L-type (with no investment) states.

Under different conditions, $I = 0$ can be an absorbing state: if neither type invests or if L types do not invest and have a 100% probability of purchasing $I = 0$ land.

B Data Sources and Sample Construction

B.1 Property Tax Assessments

Florida and Montana property taxes are publicly available as GIS files. I obtained Kansas, Oregon, and Wyoming taxes through either state- or county-level tax officials. For Nebraska, I webscraped county-level data hosted by GIS Workshop, covering almost all counties. A large majority of assessments list the PLSS section (or, rarely, sections) of each property. In counties where section information was not comprehensively provided, I relied on GIS parcel maps (Florida, Wyoming) or geocoded property address (Kansas).

Some data are only reported comprehensively for specific states. Land use data including active grassland and pasturing are reported for Kansas, Montana, and Nebraska. Florida, Nebraska, Kansas, and Oregon report owner name and address. Similar data are reported partially for Montana and Wyoming, but both contain substantial unsettled lands in the public domain which are coded as owned by the federal government. These lands are typically leased to nearby farmers, meaning that ownership data has a different interpretation in these parcels compared to parcels outside the public domain. For thirteen counties in the sample, exempt government lands are absent in the dataset and for these I the CropScape-derived use value in place of total valuation.

B.2 Grant Boundaries and Sample Construction

As noted in Section 2.2, most railroad grant areas are within a pre-specified distance of the company's railroad track. For these areas, I use historical maps to find the relevant radius for the grant and draw a buffer around the railroad. In some cases, multiple effective distances applied, e.g., because companies were granted additional "miles" on some sections to substitute for excluded land elsewhere. In such cases, I choose the outermost distance as relevant. Since most railroad locations have not changed, I use modern-day GIS information from ESRI on their location as it is most precise. I confirm the grant railroad location with the 1890 railroad data from Donaldson and Hornbeck (2016).

Some grants had non-formulaic borders. For example, companies lost land that intersected with the Crow (Montana) and Osage (Kansas) reservations. In these cases, I use a mix of historical maps, court records, and Bureau of Land Management (BLM) transfer records to determine the boundaries of the grant. Maps show the rough locations of non-formulaic grants and the BLM records permit an exact mapping through the evidence of checkerboard patterns around individual PLSS Sections. In the rare cases these records are incomplete, I use the BLM Tractbooks to determine the areas railroads received grants. Using the land grant boundary lines I construct, I code any PLSS section which intersects them as being within the grant area. As noted in Section 7.3, only pure formulaic boundaries are considered as part of the RD. Borders from other political boundaries, railroad start and end points, or formula violations are not considered.

B.3 Historic Farm Microdata

For farm ownership and operational details, I draw upon 1940 “county census” documents for Kansas, preserved by Ancestry and the Kansas State Historical Society. These were used to produce (bi)annual reports on agricultural activity by the Kansas State Board of Agriculture. For each farm in 1940, the records list the operator, the PLSS section, acreage, land use, and owner information. I selected geographic coverage based both upon the presence of railroad grant land and the existence of complete records at the district level. In the long survey, some assessors chose to leave ownership blank entirely or selectively.

I also include 1965 “personal assessments” from Lincoln County, Nebraska. These are essentially tax filings based on personal property and, crucially, grain production. For the purposes of this project, they include a property’s location, the owner’s farm equipment used on it, and a breakdown of grain output by “operator” and “landlord” shares while giving the identity of both in this case. While some respondents record a contract without a share arrangement, this is uncommon meaning the survey primarily measures share tenancy.

B.4 Land Transfer Records

I measure land concentration and sale volume with two data sources. First, the Bureau of Land Management General Land Office records offers complete coverage of initial federal transfers. To the best of my knowledge, no comprehensive database of railroad transfers exists. I thus supplement these records with archival work on railroad company transfers in Lincoln County, Nebraska, which preserved its railroad sale deeds. Historical assessment and tax records were also useful for determining land concentration’s impact on investment over time. To this end, I digitized the 1900 assessment records from Perkins County, Nebraska, and the 1912 assessors’ records from Morrill County, Nebraska. The Morrill records additionally record the fraction of improved land and the value of improvements, but unfortunately total valuations are not useful as land in this era was typically fixed at a particular value per acre.

For panel data, I digitized Register of Deeds transfer records for Banner County, Nebraska available at the sixteenth section (40-acre) parcel level. I selected these counties based on data quality, availability, and their possession of substantial portions of land inside and outside railroad grant areas.

These data were collected by assessors for the Kansas State Board of Agriculture annual reports. They document contains outcomes rarely recorded at the farm level in this period, including an exact PLSS section, the operator, and the owner, revealing rates of owner-driven cultivation.

B.5 Linking to Census Microdata

I often match property owners to the most recent US Census microdata prior to the assessment/sale. Since property taxes typically only includes the owner's name, I lack key pieces of information common in other linking procedures such as an owner's age or birthplace. In all cases, I can make use of the property's county. In the case of the initial sales matching for Lincoln County, Nebraska I am also able to use a listed county of origin.

I first compute a name match score between the property owner and all Census individuals, considering only the first listed owner in the uncommon case of joint property. For both the first and the last name, I compute the Jaccard string similarity index, the fraction of unique bigrams in either name that are contained in both the owner and proposed match names. In the case of single-letter first names, I substitute a value of 90% if the two names begin with the same letter. Thus, "John Smith" would be considered a good although not perfect match for "J. Smith." I compute the overall name match as the average similarity between the first and last names.

The second element of the matching procedure is how to value location. In the case of the Lincoln County, Nebraska initial sales, I consider the owner's listed county of origin, state of origin, and finally Lincoln County itself. For historical property tax matching, I consider the property's county and state only since I lack information on the owner's origin. Taking the name match value given as above, I apply a 20-percentage point premium to the Census individual's score if they reside in the listed county of origin or property value's location; I apply a 10-percentage point premium to their score if they reside in the same state as the owner or owner's property respectively. The individual with the highest match score, including location premia, is my preferred match. To exclude false matches, by default I consider tied duplicate matches or those with string similarity below 75% as non-matches.

B.6 Population and Public Goods

I obtain 2019 population at 30 meter by 30 meter resolution from the Humanitarian Data Exchange. For historical population, I use the detailed 1940 Census "enumeration district" maps showing the location of every rural farm, school, church, and other structures. The number of farmsteads serves as a good proxy for the rural population as almost all would have resided in farm buildings. I consider schools, churches, cemeteries, and community buildings as public goods. For the modern road network, I use the Federal Highway Administration's 2015 HPMS data. For town locations, I use both the Schmidt (2018) point file and the 2000 Census TIGERLINE place polygons.

B.7 Geographic Characteristics and Land Use

Elevation data are from the SRTM 250-meter resolution database. A related database from the FAO contains the terrain slope characteristic, a key agricultural input. In the small number of areas where these data are unavailable, I impute elevation and slopes, regressing the measure on latitude and longitude in each county and using the predicted value. For river and stream length, I use data from ESRI. For soil quality characteristics, I use the USDA's gSSURGO database. For crop productivity, I draw upon their "nccpi2 (all)" aggregated measure of soil productivity for different crops.

B.8 Land Use Value Calculation

I construct a pure "use value" of land using satellite data on land use (USDA's CropScape), models of agricultural productivity (the FAO's GAEZ), and data on crop prices.

For pixels coded for crop use, I first consider the expected crop yield according to the FAO's GAEZ data. I use the GAEZ "high input" scenario as this most accurately reflects agricultural processes in developed countries like the United States. For the small number of crops are not listed in GAEZ data, I use USDA-reported average national yields. To compute revenue, multiply by crop farmgate prices. I primarily use FAO-reported prices, but where these are missing I use USDA prices or prices from other sources.

For pasture and grassland pixels, I use the GAEZ yield for "pasture grass" as the expected yield of forage. Following Ahola (2013), I assume an average cow weight of 1000 pounds, that each cow eats 2.6% of its weight per day, and that 30% of forage is accessible. This analysis assumes, somewhat generously, that each grassland pixel is actively grazed. For non-developed, non-agricultural pixels, I assume a value of $0 in production.

Using the USDA's Commodity Cost and Returns dataset, I estimate annual production profits and convert these into net present valuations for cattle and each major US crop. I compute the profit margin as the ratio of revenue minus operating costs, hired labor, and taxes/insurance divided by revenue. About 1% of land in my sample has crops not covered therein and for those I use a 10% profit margin. I convert estimated annual profits to valuations by discounting at 5% rate, typical for assessors, and sum within section.

My final measure of use value also includes the valuation from urban areas. CropScape classifies such developed areas into "open," "low," "medium," and "high." Since valuations from this use do not come from production, they must necessarily be imputed. I regress total valuations in counties with complete information on the total amount of land in each category of development according to CropScape, combining the last two being combined as few pixels are coded as either. I include Township fixed effects and the main geographic characteristics and round the results. This procedure estimates a $12.5 million / square mile value for open development, $125 million for low development, and $300 million for medium/high. For comparison, Omaha, NE has roughly 180,000 households typically worth $200,000 (Zillow estimate) and an area of roughly 130 square miles, yielding about $277 million in value per square mile.

C Further Robustness and Mechanisms

C.1 Public Goods and Politics

Land concentration's effects could have operated through political, rather than economic, channels. Previous research has documented the effects of political capture by landed elites' capture and coercion of political systems (Acemoglu et al., 2019). Gates (1942) provides qualitative support for this notion in my context, writing that the large-scale owners "were influential in local and state governments which they warped to suit their interests." Other work, however, has found that landowning elites can use their influence to solve collective action problems and increase public goods provision (Dell, 2010).

Political and governmental mechanisms are unlikely to play a role on the scale I study. While land concentration may have shaped state or local policies, these would have to be selectively applied at the square mile level to explain my results. I test for such manipulation by considering two classes of outcomes at the square mile level: public goods provision and political activities by individual landowners. I code the former using 1940 census maps showing schools, churches, and community halls in rural areas as well as the modern road network. I code the second an archival sample in the early 1900s, linking owners to the names of those running for local, subcounty offices[30] in 1912 to measure officeseeking.

Appendix Table A.8 estimates the effects of the checkerboard allocation policy on public goods provision and owners' political activities. Broadly, these analyses show either economically small or statistically undetectable results. In a large sample of 1940 areas, odd sections have slightly fewer public goods than their even neighbors, with a difference of one community hall per 1000 square miles and being 0.002 miles (11 feet) further from a road being statistically significant. While these results are statistically different, they are not

[30]For example, irrigation district commissioners

economically meaningful and unlikely to have influenced the property value results. For individual behavior, odd-section owners if anything are to be less likely to seek local political office. While the sample size is small for the latter estimate and the confidence interval includes 0, it suggestively points against frequent political activity by large-scale landowners in this context.

Importantly, even sections within the checkerboard do not appear to be statistically different than their non-checkerboard counterparts in terms of in public goods or owner political activities. In Table A.8 Panel B, estimates of equation (2) on all six outcomes are not statistically significant. On public goods, the coefficients are again small and not consistent in sign, pointing against a broad effort to reduce or increase government spending in this area. This result then suggests a more mundane explanation for the differences in Panel A: they simply reflect section-level differences in population. As shown in Table 6, there were 25% fewer farming households on odd sections in 1940. Overall, odd and even sections in the checkerboard have very similar rates of public goods per household.[31] Finally, the absenteeism of odd-section landowners detailed in Section 7.1 would have made it harder for them to seek local office, explaining why only 3% did so.

Overall, these results suggest land concentration in this context did not lead to political capture at the square mile level. However, it is important to note that land concentration might still have changed public policy and politics in larger units like the county, state, or nation. Because these units are controlled for in my main results, such differences are unlikely to explain the investment differences across even and odd sections. At higher levels, the channel remains plausible and an important avenue for future research.

C.2 Rural Density

Relatedly, perhaps the density of farms per se had positive effects on productivity. For example, farmers might have cooperated with or learned from their neighbors and odd sections do have fewer farms. However, that does not necessarily mean they had fewer neighbors. The interspersed nature of the checkerboard in fact meant that farmers on both types of section would be part of very similar communities. For example, someone living on (odd) section 23 would have had neighbors in adjacent (even) sections 22, 14, 24, and 26.

Table A.7, Panel C examines this issue empirically. For each section, it picks a point at random and measures the modern[32] population living within 1-10 miles from that point. Using this slightly broader definition, there are almost no differences in density across even

[31]Odd sections have 0.051 schools, 0.0073 churches, and 0.0011 community halls per farm in the full sample compared with 0.046 schools, 0.0055 churches, and 0.0013 community halls for even sections.

[32]While I use modern data due to it being available in a disaggregated form, Table 6 shows that these persistently reflect historical even/odd differences.

and odd sections. All estimated differences are smaller than 0.5% and only one (at 3 miles) is significant; it would indicate odd sections had a higher rather than lower density. So, although even squares had more farms operating within their boundaries, those farmers had similar numbers of total neighbors. Unless density spillovers occurred only within the artificial boundaries of PLSS squares, they cannot explain my results.

C.3 Property Rights

Railroad owners could have invested less in their land because they felt their ownership was not secure. The slow speed of some (but not all) companies to either build their tracks or sell their land sparked "forfeiture" movements to reclaim their unsold sections. Within my sample, the detailed overview in Ellis (1946) lists movements targeting the Northern Pacific Railroad (NPRR) and Oregon/California Railroad (OCRR) companies. The others in my sample, such as the Union Pacific, were more compliant and not targeted; see Appendix Section C.5 for evidence on settlement speed. Although individual settlers were never targeted, in principle they may still have felt uncertainty ex-ante. Alston and Smith (2022) argues the NPRR's grant was uniquely troubled by this and other legal ambiguities as the company's "violations, controversies, and investigations... had no peers" (Draffan, 1998). Thus, the NPRR and to a lesser extent the OCRR are the grants where insecure property rights should have been most prominent.

Two analyses indicate that these forms of insecurity do not explain my results. In Appendix Table A.9, odd sections in untargeted grants experienced slightly greater land value reductions: dropping the NPRR and OCRR from the sample modestly increases the estimates' magnitude and significance. Second, I analyze the frequency of lawsuits (*lis pendens* notices) in the archival sales data from Nebraska. 28% of land in the sample experiences a lawsuit over the period, with the rate actually lower (insignificantly) in odd sections. Overall, odd sections' reduced valuations seem to result from a consistent pattern in and out of contested grants like the NPRR. Based on available data, odd-section owners in other areas did not face greater legal issues with their titles.

C.4 Homestead Implementation in Railroad Grant Areas

Some historical sources argue that the Homestead Act was implemented differently in railroad grant areas. For example, proponents of the grant policy argued that doubling federal land prices in the grant area could compensate the government's loss of half its land. Other proposals would have set the standard settler plot size at 80 acres rather than 160. If implemented, these policies would complicate the RD analysis in Section 8 as multiple variables changed at the border. The even/odd regression's interpretation would be substantively unaffected since even squares would still be reserved for individual families. However, the exact policy details would require correction.

Based on historical and quantitative evidence, these proposals were not implemented to a significant degree in my sample. First, as noted by Gates (1954), the 80-acre rule was abandoned in 1879, preceding almost all settlement of my areas; the doubled $2.50/acre price was not meaningful given that the vast majority of settlers opted for free land under the Homestead Act. Appendix Table A.6 is consistent with this narrative. There is no detectable difference in federal land grant sizes within the border and, belying a higher price, slightly more land was transferred. The statistically insignificant 12-acre point estimate would represent a 3% decrease in contrast to the 50% implied by the 80- versus 160-acre distinction. Thus, there is little qualitative or quantitative evidence that federal settlement policy changed at the borders in my sample.

C.5 Date of Settlement

If either Homestead or railroad lands were systematically settled earlier, any differences today could simply reflect some sort of first-mover advantage or head start from the earlier group. While comprehensive data on railroad sales are unavailable across all railroad grants, the archival sales data offer one window into this question. Appendix Figure A.7b shows the fraction of land that had at least one (non-railroad, non-federal) owner by year. Railroad and Homestead land were settled around the same time in this county, with neither group consistently experiencing a faster process.

It is possible that the relative timing differed in other grants, particularly the ones subject to the forfeiture movement discussed in Section C.3. However, these archival data suggest that that was not always the case. As noted earlier, including the railroad grants subject to forfeiture, if anything, weaken the main results.

C.6 Speculation

Gates (1936) and other historians viewed some large-scale railroad land buyers to be "speculators." One interpretation of this view is that those owners held their land off the market, aiming to let it appreciate in value following population increases rather than from their own investments. My results would thus primarily represent the long-term effects of a free-riding problem rather than land concentration per se. However, Gates connected speculation with land concentration, writing that tenant farming was one "of the worst effects of the resulting large-scale ownership [from these purchases]." Many speculators had long-term ambitions of "establishing for themselves a permanent investment from which they and their descendants might draw rents as the landed aristocracy of England had done for centuries" (Gates, 1941).

Quantitative evidence also cuts against the idea of odd-sections being held idly off the market. Appendix Figure A.7a indicates that they were transferred by their owners somewhat more frequently than even sections.[33] Very simply, odd-section owners were not typically holding their properties off the market, and by 1920, 99% of odd-sections had been transferred at least once and 59% had been transferred three or more times. Delayed investments from land held off the market should also not have been more harmful in areas with high rates of share tenancy, the pattern documented in Section 7.2. Both qualitative and quantitative evidence indicates that odd-section owners were in fact working their farms, making additional use of tenant farming relative to even-section owners. This element, rather than idle land, is key to the results.

C.7 Credit Constraints

In this section, I use archival data to provide evidence that even-section owners had more limited access to capital than odd-section owners. This finding supports the hypothesis that credit constraints played an important role in the slow reallocation process discussed in Section 7.3. However, other frictions could certainly have contributed and this paper does not aim to exhaustively list them; see Bleakley and Ferrie (2014) for other important work on this topic. As with the original Coase Theorem, many possible market imperfections would lead to the same key result that the initial allocation had long-term effects.

To measure landowners' access to capital, I use an archival sample of property tax records dating from 1900. Property taxes were a substantial cash obligation in this setting, meaning that difficulties paying them reflected a general lack of access to cash. For each parcel, the records list the land's owner, the date of the tax payment, and by whom the tax was paid. Essentially, all taxes are eventually paid in this context, but on average it took 24 months, and 71% of owners used an intermediary, indicating substantial difficulties for these settlers. Figure A.11 shows that even-section owners had more difficulty accessing capital than their odd-section neighbors. On average, they paid their taxes off 5 months later, and the effect spread evenly across the distribution. Even-section owners were also 6.7 percentage points (t=2.05) more likely to use an intermediary to pay.

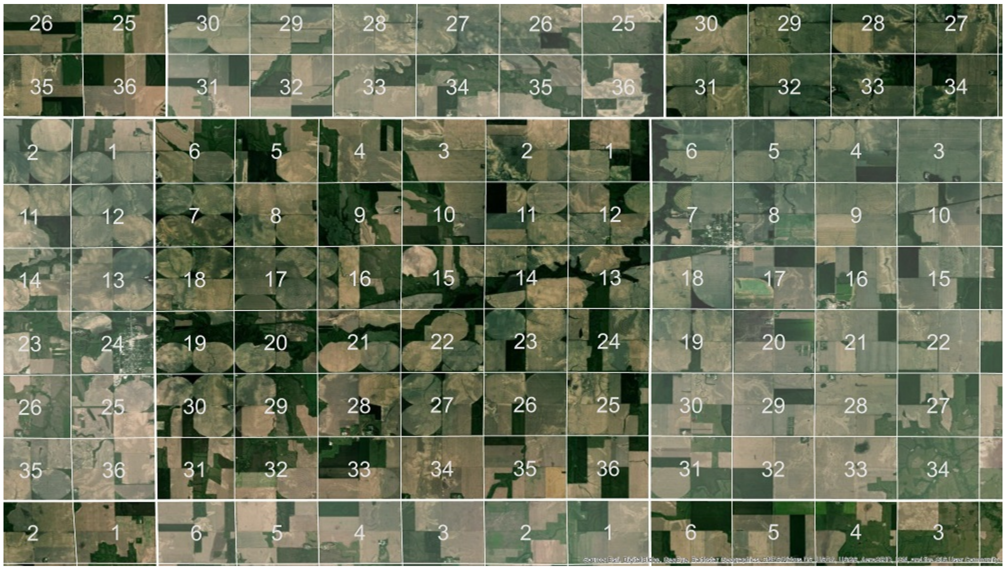
These delays and heightened reliance on intermediaries suggest that even-section owners had more limited access to capital than their odd-section counterparts. This result aligns well with the historical context. Railroad land buyers were by definition capable of purchasing

[33] The data shown do not consider the "first" transfer from either the federal government or the railroad company in this graph.

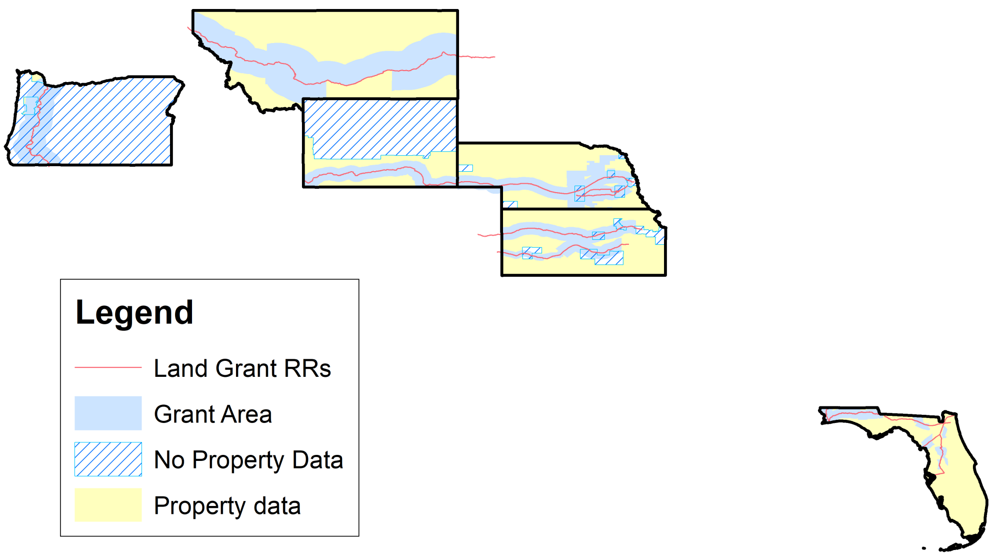
property, whereas the Homestead Act aimed to distribute land to individuals less able to do so.

D Appendix Tables and Figures

Figure A.1: The Public Lands Survey System

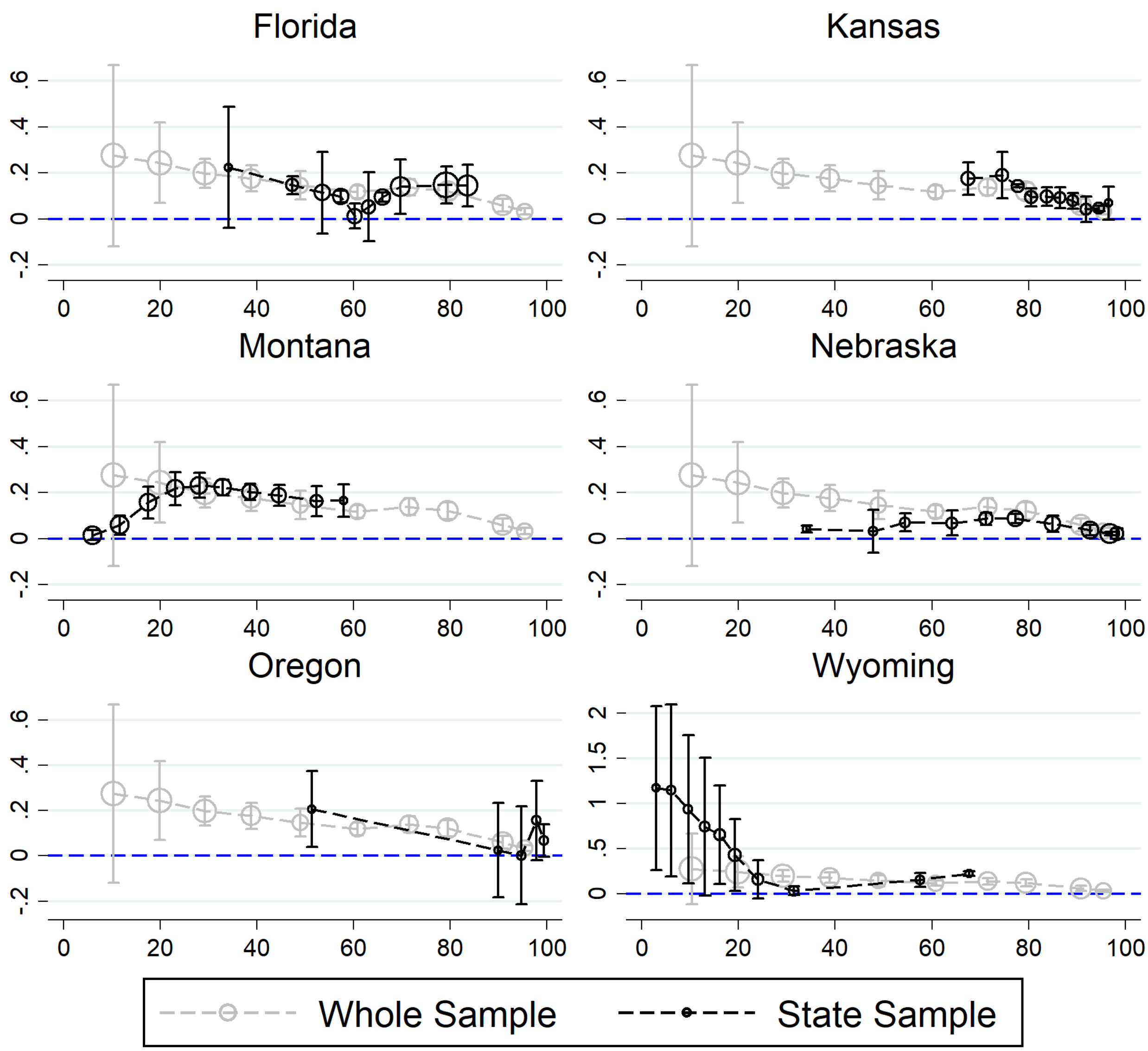


(a) Nebraska Townships and Numbered Sections



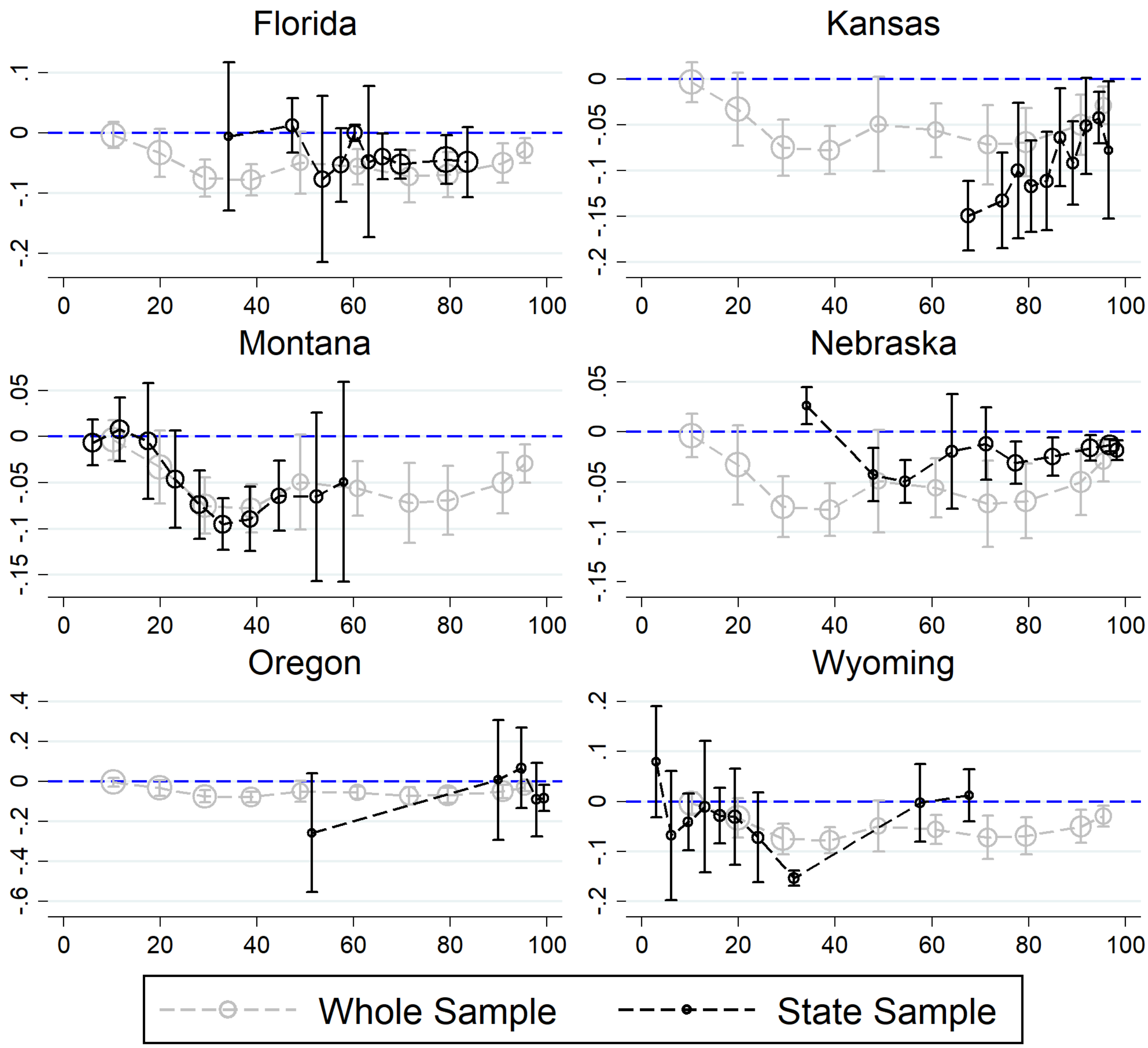
(b) Extent of Railroad Land Grants

Figure A.2: Effects on (log) Average Parcel Size



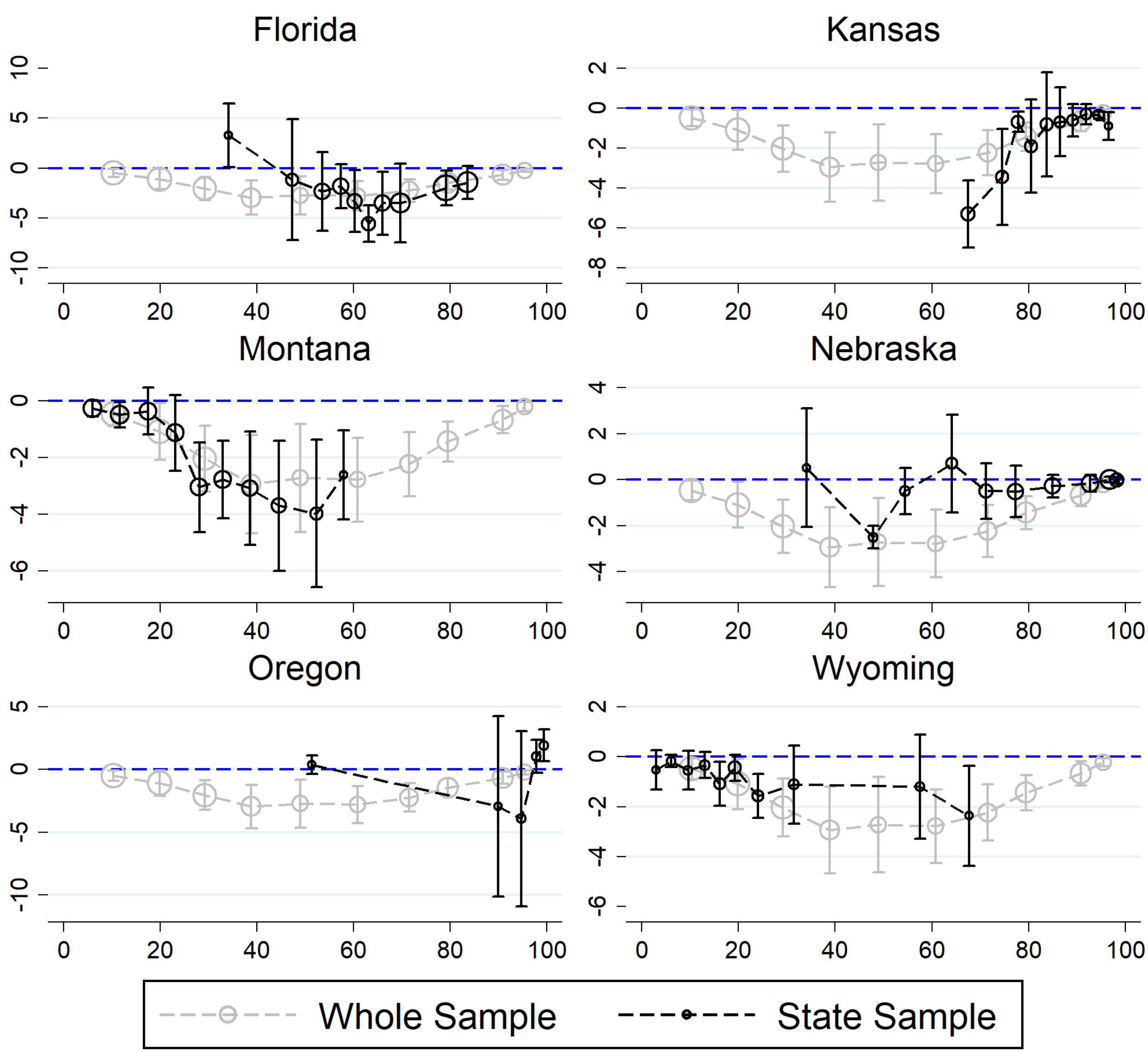
Notes: This graph shows full-sample and state-sample estimates of railroad land ownership on log(Acres / parcels) in a PLSS section as per the even/odd comparison of equation (1). Parcels that cover multiple sections are counted fractionally across each section so their total contribution sums to 1. Each dot represents a subsample of sections based on land quality according to the gSSURGO database. X-axis values reflect the average, full-sample percentile of land quality within the sample. State samples are chosen to reflect a 20-percentile range of the land quality within their state.

Figure A.3: Effects on (asinh) Land Value



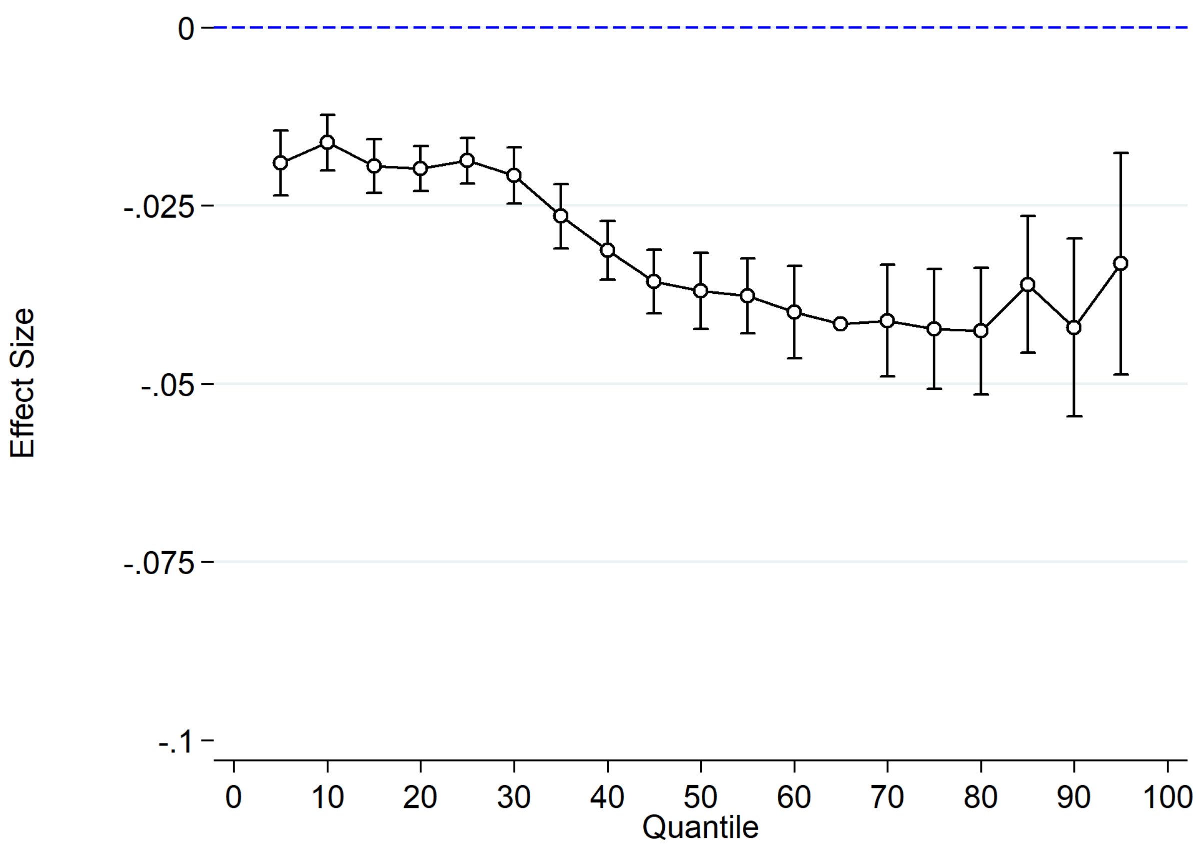
Notes: This graph shows full-sample and state-sample estimates of railroad land ownership on the (asinh) modern land value in a PLSS section as per the even/odd comparison of equation 1. Each dot represents a subsample of sections based on land quality according to the gSSURGO database. X-axis values reflect the average, full-sample percentile of land quality within the sample. State samples are chosen to reflect a 20 percentile range of the land quality within their state. Bars depict 95% confidence intervals.

Figure A.4: Effects on Crop Farms



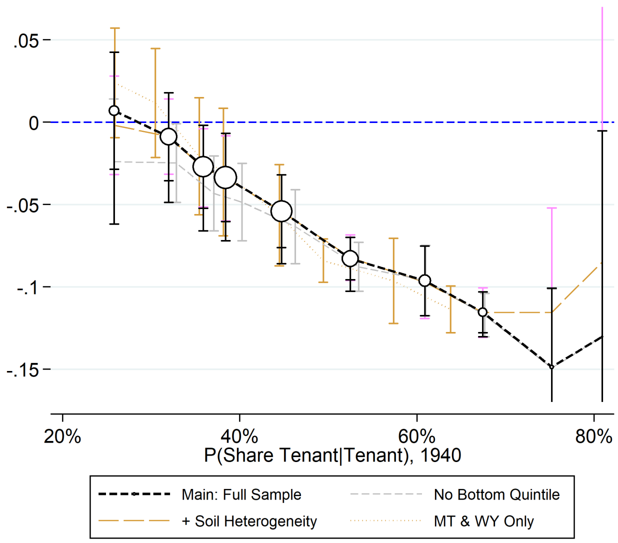
Notes: This graph shows full-sample and state-sample estimates of railroad land ownership on the percent of PLSS sections growing crops on 1% or more of their area, using the even/odd comparison of equation (1). Each dot represents a subsample of sections based on land quality according to the gSSURGO database. X-axis values reflect the average, full-sample percentile of land quality within the sample. State samples are chosen to reflect a 20 percentile range of the land quality within their state. Projected values are censored to remain within the 0-100% range. Bars depict 95% confidence intervals.

Figure A.5: Quantile Effects on (asinh) Land Value

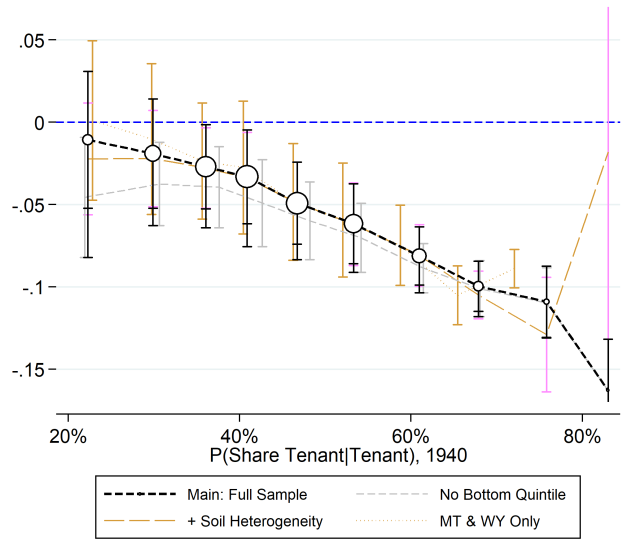


Notes: Quantile effects of equation (1), replicating the specification of Table 4, column (6).

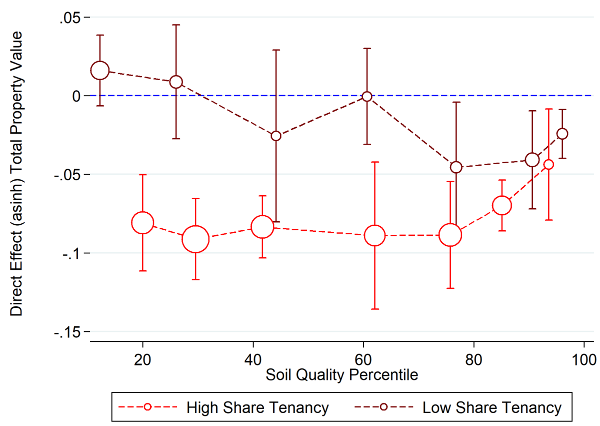
Figure A.6: Effects on Property Values by (Predicted) Share Tenancy



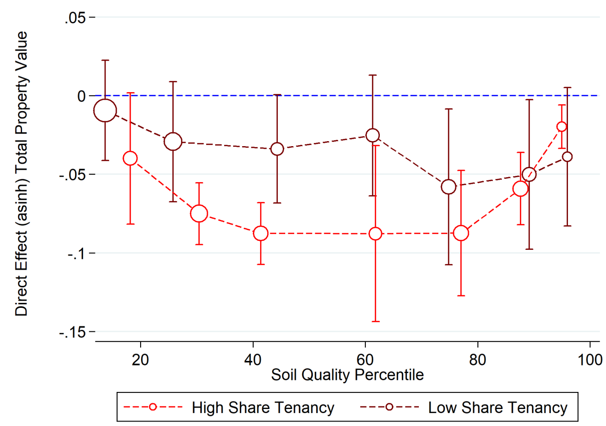
(a) Neighbors' Share Tenancy



(b) Geo-Predicted Share Tenancy



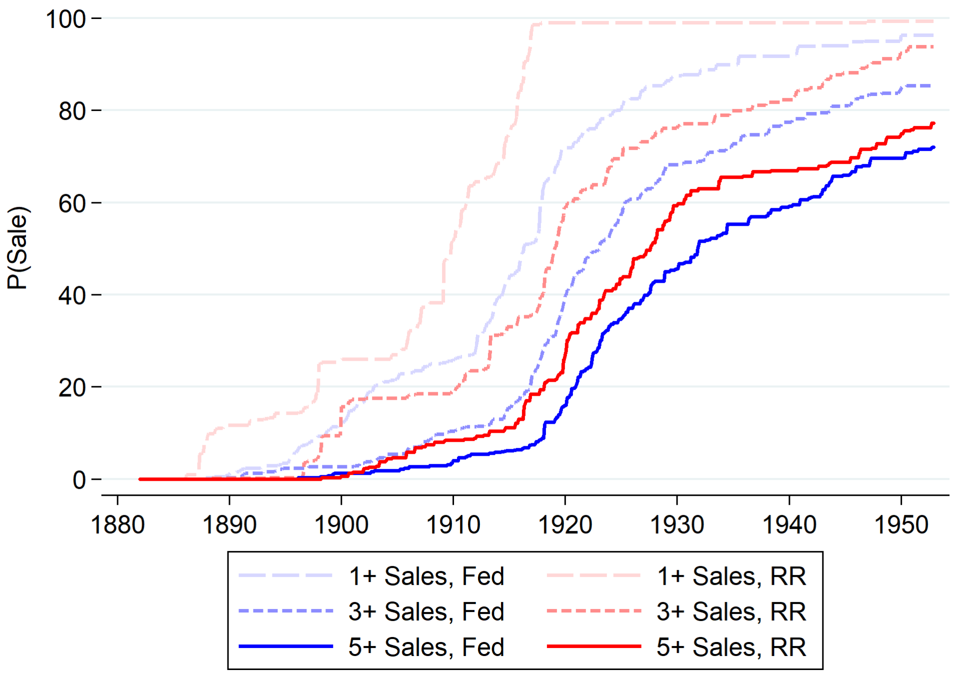
(c) Neighbors' Share Tenancy



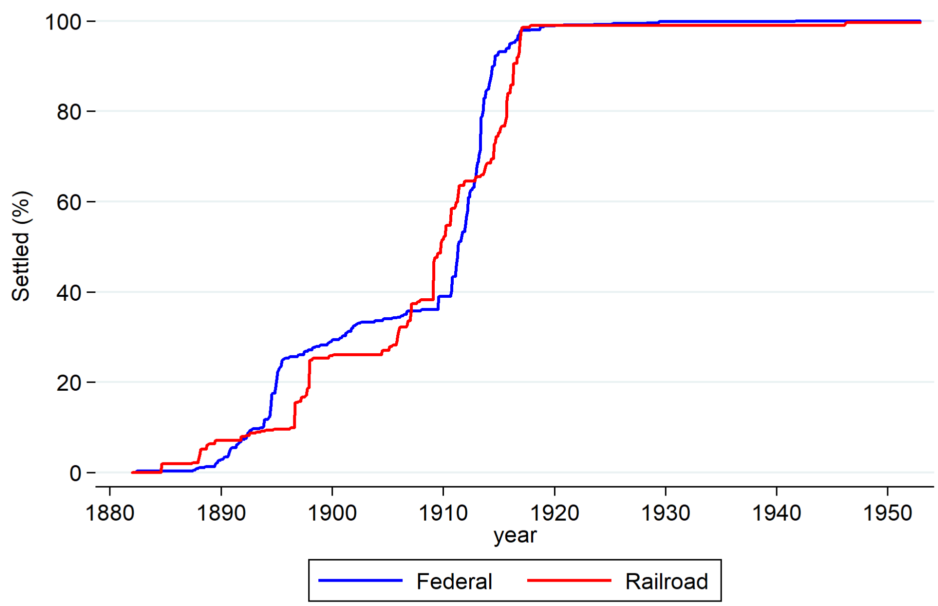
(d) Geo-Predicted Share Tenancy

Notes: This figure replicates the share tenancy heterogeneity results in Figure 5, using predicted rates of share tenancy instead of the county's own value. As such, the source of heterogeneity for any particular section is not determined by the section's own rates of share tenancy. (a)-(b) replicate heterogeneity based on predicted levels of share tenancy. (c)-(d) compare effects by soil quality percentile for areas above and below median rates of predicted share tenancy. (a), (c) use the average rate of share tenancy in neighboring counties within the same state as the prediction. (b)-(d) regress counties' share tenancy rates on the county-average values of the geographic characteristics and log(county area), using the regression-predicted values.

Figure A.7: Extent of Sales Over Time (archival sample)



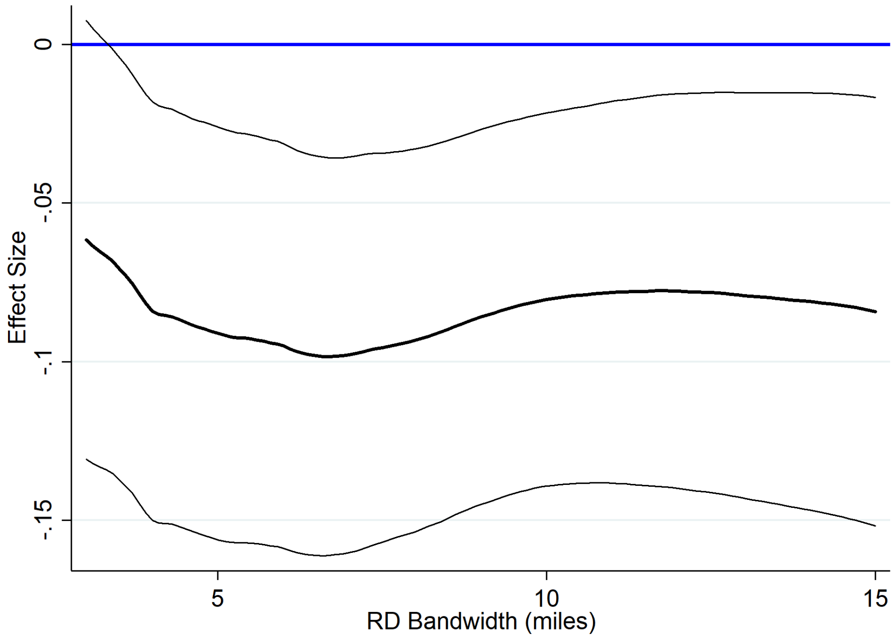
(a) Probability 1+, 3+, 5+ Cumulative Sales



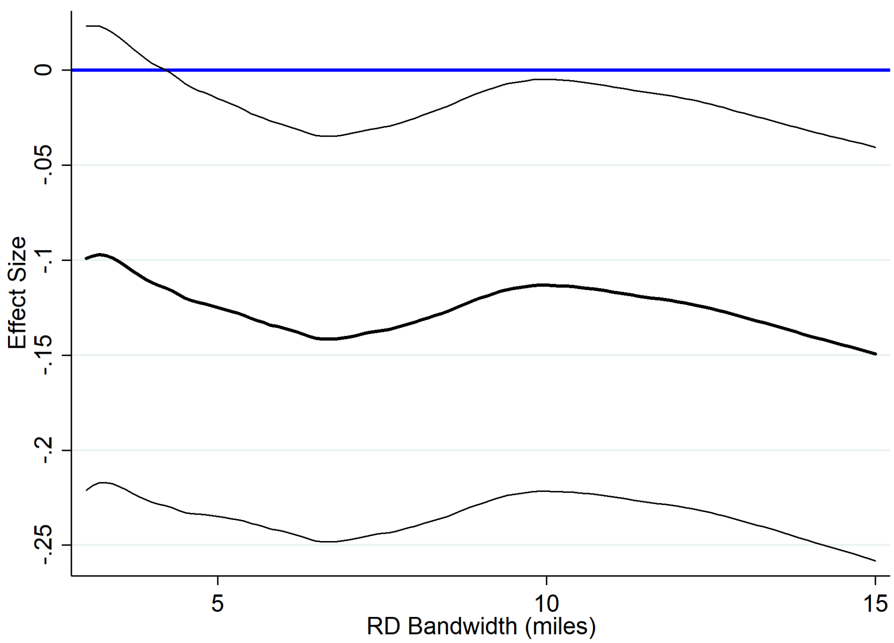
(b) Ever Settled

Notes: (a) depicts the percent of railroad versus federal parcels in Banner, NE that had been transferred 1+, 3+, or 5+ times by individual owners (i.e. excluding initial transfers from either the US government or railroad). (b) depicts the fraction of land that had been transferred to its first owners (ignoring the US government and railroad companies).

Figure A.8: RD Bandwidth Robustness



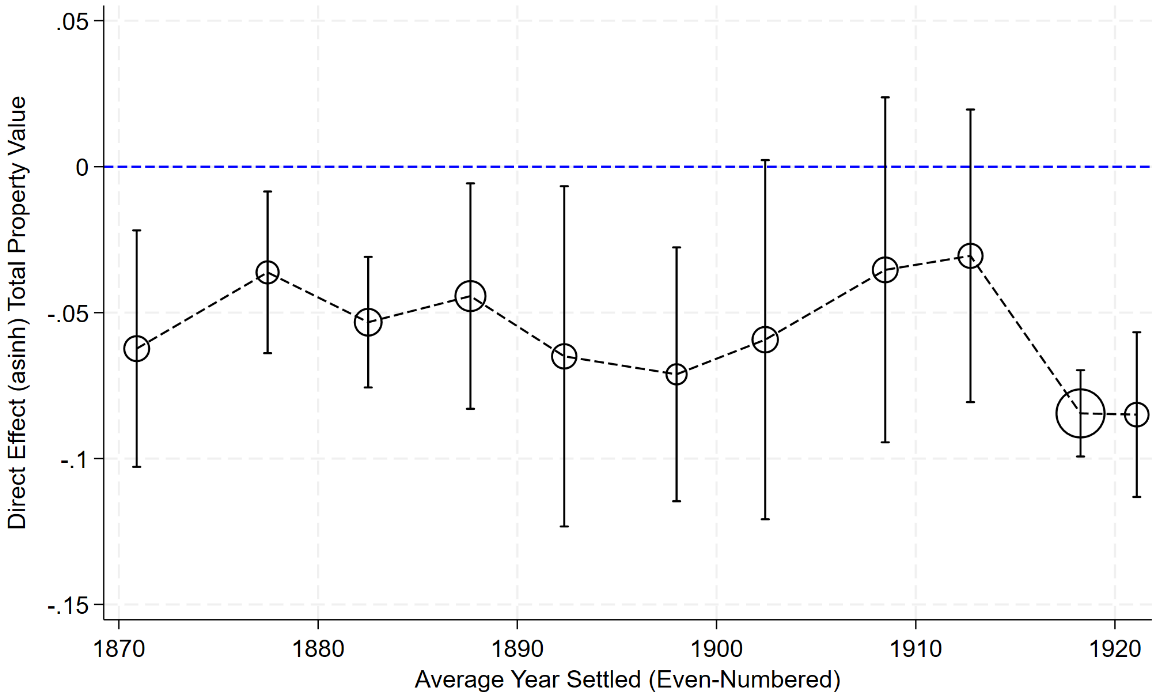
(a) Total Property Value Effect by Bandwidth



(b) Improvement Value Effect by Bandwidth

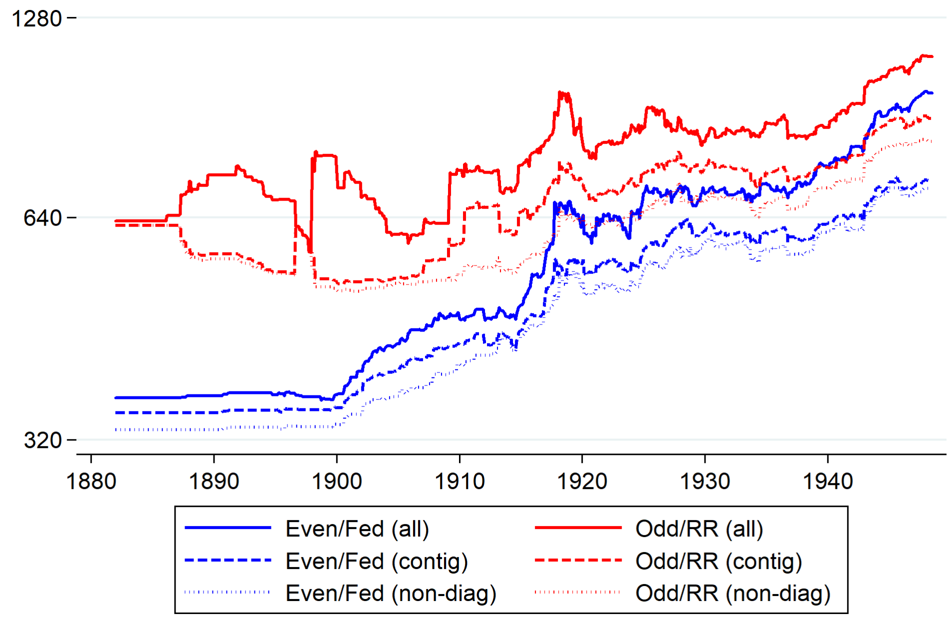
Notes: Estimates of equation (2) as a function of bandwidth, with 95% confidence intervals plotted. (a) plots (asinh) 2017 assessed total value, (b) plots (asinh) assessed improvement value.

Figure A.9: Effects by Year of Average County Settlement

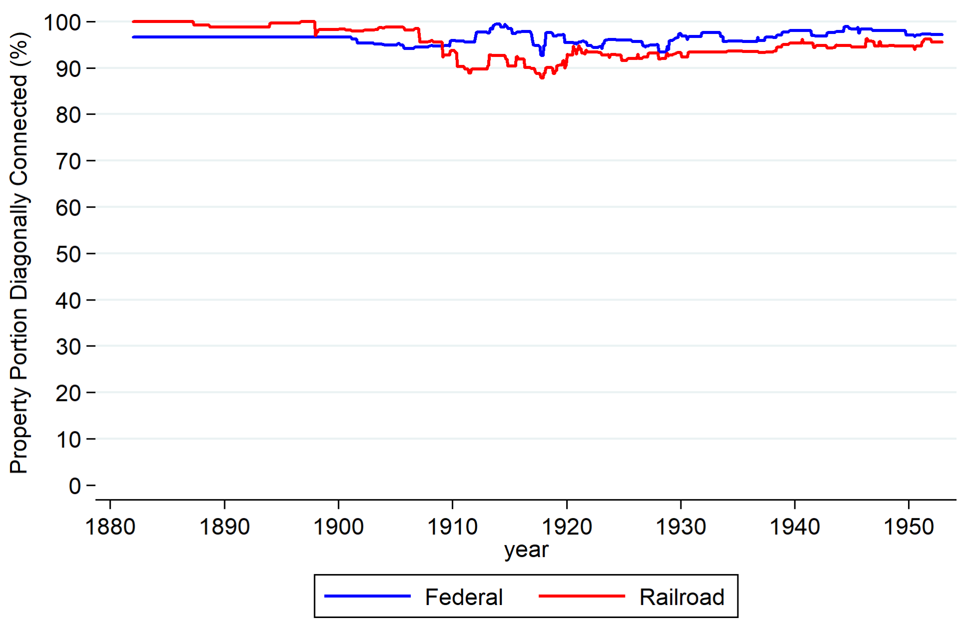


Notes: This graph shows full-sample and state-sample estimates of railroad land ownership on the (asinh) modern land value in a PLSS section as per the even/odd comparison of equation 1. Each dot represents a subsample of sections based on the average land was settled in the non-railroad lands of a county. This value is computed for each PLSS section as the average year of federal settlement for non-education, non-railroad sections within the county, excluding the section itself (i.e., "leave one out"). Bars depict 95% confidence intervals.

Figure A.10: Alternative Property Size Measures



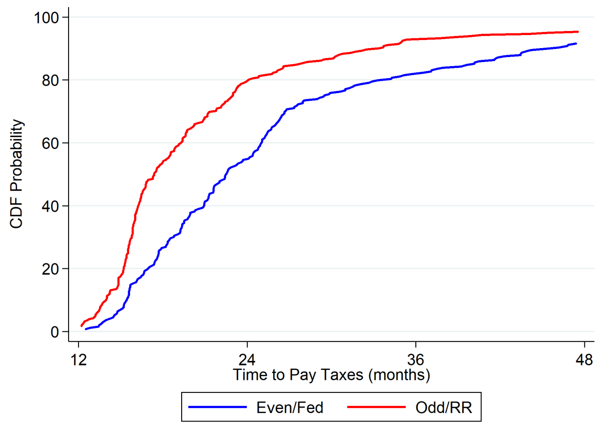
(a) Measures of Property Sizes



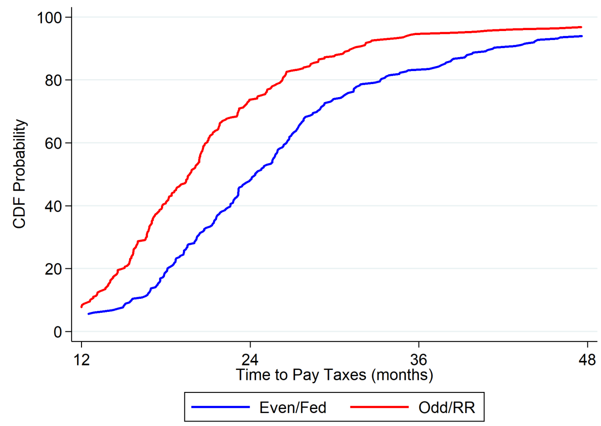
(b) Diagonal/Contiguous Property Size Ratio

Notes: property size measures over time based on archival sales data. (a) plots section-average log property sizes for even/odd sections based on an owner's entire holdings, contiguous holding, and contiguous holdings excluding diagonal connections. (b) plots the section-average ratio between the third and second of these groups.

Figure A.11: Months to Pay 1900 Taxes (Archival Sample)



(a) Unadjusted



(b) Adjusted for Controls

Notes: The figure shows cumulative distribution functions of section-average time-to-pay property taxes from Perkins County, Nebraska in 1900. Panel (b) adjusts for controls, keeping the even-section mean fixed.

Table A.1: Sales Price and Assessed Value per Acre

	Panel A: Sales vs. Assessed Value, Section Level			
	(1)	(2)	(3)	(4)
	log(Sale Price/Acre)	log(Sale Price/Acre)	log(Sale Price/Acre)	log(Sale Price/Acre)
log(Total Val/Acre)	0.94***	0.51***	0.61***	0.68***
	(0.0080)	(0.044)	(0.038)	(0.038)
log(Land Only Val/Acre)		0.14***	0.18***	0.11***
		(0.051)	(0.034)	(0.033)
Sample	All	All	Agricultural	Agricultural
Township FEs				Y
SEs / Clusters	Township	Township	Township	Township
N	22,970	6,250	6,250	6,250
$\mathbb{E}[y]$	7.9	7.9	8.5	8.5

	Panel B: Sales vs. Assessed Value, Parcel Level			
	(1)	(2)	(3)	(4)
	log(Sale Price/Acre)	log(Sale Price/Acre)	log(Sale Price/Acre)	log(Sale Price/Acre)
log(Total Val/Acre)	0.80***	0.63***	0.60***	0.70***
	(0.020)	(0.043)	(0.046)	(0.037)
log(Land Only Val/Acre)		0.23***	0.14***	0.038
		(0.035)	(0.036)	(0.028)
Sample	All	All	Agricultural	Agricultural
Township FEs				Y
SEs / Clusters	Township	Township	Township	Township
N	913,886	850,494	11,104	11,104
$\mathbb{E}[y]$	9	8.9	8.6	8.6

	Panel C: Use vs. Assessed Value, Section Level			
	(1)	(2)	(3)	(4)
	asinh(Value)	asinh(Value)	asinh(Value)	asinh(Value)
asinh(Use Value)	0.42***	0.46***	0.25***	0.23***
	(0.0057)	(0.0087)	(0.0040)	(0.0045)
asinh(Ag. Use Value)		-0.066***		0.10***
		(0.0080)		(0.0038)
Sample	All	All	All	All
Township FEs			Y	Y
SEs / Clusters	Township	Township	Township	Township
N	339,482	339,482	339,482	339,482
$\mathbb{E}[y]$	7,434	7,434	7,434	7,434

Notes: This table correlates sales, assessed, and use values per acre. Sales data come from Florida tax records from 2016-17. It considers properties sold in 2016-17 with a positive sales price and reported acreage. Both the total property valuation and the valuation excluding "improvements" (buildings) are considered. Panel C correlates author-generated values based on land use with assessed values for the sample of property tax counties (regardless of railroad grant status); see Appendix Section B.8. Panel B uses data at the property level. Panels A and C aggregate values to the PLSS section level, as in the paper's main regressions for equation (1).

Table A.2: Archival Sample Description

Area (year)	Source	Variables	Where Used
Lincoln County, NE (1800s)	First individual owners (Federal: BLM records Railroad: deeds of sale)	Owner name, county of origin, property description	Figure 3, Table 7
Morrill County, NE (1912)	Land assessment, local elections	Owner name, property description, improved land, improvement value, officeseeking	Figure 4, Table 5, Table 7, Table 8, Table A.4, Table A.8
Nebraska (1940)	Census enumeration district maps	Number of farms, schools, and other public goods by PLSS section	Table 6, Table A.8
Lincoln County, NE (1965)	Personal property assessment	Farm equipment, (share) tenancy by PLSS section	Table 6, Table 7
Perkins County, NE (1900)	Land assessment	Owner name, property description	Figure A.11, Table 7, Table 8
Kansas (1940) 30 townships Barton, Dickinson, Harvey, Pottawatomie counties	State agricultural survey	Operator name, property description, owner name	Table 7
Banner County, NE (1882-1948)	Land transfer records	Owner name, recipient name, property description, deed type	Figure 6, Figure A.7a, Table A.9

Notes: Descriptions of archival samples used in this paper. Kansas 1940 survey samples include only townships with complete lists of owners and operators. In many cases, the list of owners was left blank. Banner County, NE land records cover the 17N townships.

Table A.3: Summary Statistics (Archival Samples)

	PLSS Sections						Counties (1940)	
	(1) Number Sections	(2) Soil Quality (z-score)	(3) Homestead (%) (Even Only)	(4) Crops (%)	(5) Value $000	(6) # Parcels (Median)	(7) Tenant Farm (%)	(8) Share Farm (% Tenant)
RR Grant Areas	132,463	-.113	86.7	47.9	2,231	2	38.8	42.8
Lincoln, NE (1800s, 1965)	2,084	-.107	99.2	44.4	1,078	2	49.2	35.2
Morrill, NE (1912)	101	-.273	94	41.6	307	1	55.7	63.3
Nebraska (1940)	18,622	.92	96.1	82.8	2,296	5	51.5	39.8
Perkins, NE (1900)	537	.345	98.3	98.9	1,398	4	54.4	55.9
KS State Census (1940)	738	1.45	84.4	96.5	2,626	7	42.3	43.6
Banner, NE (1882+)	204	-.278	96.5	72.5	338	2	42.7	71.7

Notes: The table presents summary statistics for different geographic units, focused on the archival data samples of Table A.2. Column (1) shows results for section sample size, column (2) for the gSSURGO crop productivity index (full-sample z-score), column (3) for percentages of non-railroad land transferred under the Homestead Act, column (4) for the percentage of sections with at least 1% in crops per the USDA CropScape data, column (5) for total property values, column (6) for the median number of parcels, column (7) for county-level average rates of non-owner-operated (tenanted) 1940 farms, and column (8) for county-level averages of the shares tenant farms as a fraction of all tenant farms in 1940.

Table A.4: Land Values — Functional Form and Heterogeneity

	Panel A: Functional Form (Total Property Value)				
	(1) asinh(x) (baseline)	(2) ln(1+x)	(3) ln(max(1,x))	(4) $x > 0$ (%)	(5) $x >$ median (%)
RR Effect	-0.045***	-0.045***	-0.046***	-0.015	-1.49***
	(0.014)	(0.013)	(0.013)	(0.026)	(0.39)
Grant × State FEs	Y	Y	Y	Y	Y
County FEs	Y	Y	Y	Y	Y
Township FEs	Y	Y	Y	Y	Y
SEs / Clusters	Spatial	Spatial	Spatial	Spatial	Spatial
N	132,463	132,463	132,463	132,463	132,463
$\mathbb{E}[y]$	$2,134k	$2,134k	$2,134k	1.0e+02%	50%

	Panel B: Functional Form (Investments)				
	1912 Sample			Full 2017 Sample	
	(1) Imp. Value > 0 (%)	(2) (asinh) Imp. Value / owners	(3) Acres Imp. (% Section)	(4) Imp. Value > 0 (%)	(5) (asinh) Imp. Value, non-home
RR Effect	-24.1**	-1.14***	-9.93**	-3.68***	-0.16***
	(8.01)	(0.19)	(4.09)	(1.00)	(0.034)
Grant × State FEs	Y	Y	Y	Y	Y
County FEs	Y	Y	Y	Y	Y
Township FEs	Y	Y	Y	Y	Y
Geo Controls	Y	Y	Y	Y	Y
SEs / Clusters	Township	Township	Township	Spatial	Spatial
N	101	82	101	132,463	132,463
$\mathbb{E}[y]$	23%	$2.7k	13%	43%	$412k

	Panel C: Heterogeneity				
	(1) (asinh) Value Total	(2) (asinh) Value Ag.	(3) Crop Farm (%)	(4) (asinh) Value Improvements	(5) (asinh) Pop
RR Effect	-0.060***	-0.046***	-1.83***	-0.30***	-0.11***
	(0.013)	(0.014)	(0.48)	(0.039)	(0.011)
RR × Low	0.056***	0.066***	1.33***	0.24***	0.077***
	(0.011)	(0.016)	(0.44)	(0.038)	(0.012)
Grant × State FEs	Y	Y	Y	Y	Y
County FEs	Y	Y	Y	Y	Y
Township FEs	Y	Y	Y	Y	Y
Geo Controls	Y	Y	Y	Y	Y
SEs / Clusters	Spatial	Spatial	Spatial	Spatial	Spatial
N	132,463	132,462	132,462	132,463	132,463
$\mathbb{E}[y]$	$2,231k	$380k	48%	$1,277k	23

Notes: This table extends Table 6.2 with alternative functional forms for the outcomes and heterogeneity. Panel A focuses on functional form. (1) and (4) considers the extensive margin of improvements. (2) studies the (asinh) value of improvements divided by the number of individual owners who own land in the section. (3) studies the fraction of a section's land marked as improved. (5) focuses on the value of improvements excluding homes and dwellings. Panel B considers an interaction effect with low land quality, defined as a gSSURGO quality measure in the bottom 20% of the sample. All data are from the full modern sample and respectively use modern total property value, imputed use value based on satellite data, the extensive margin of crop farming, the value of improvements, and population.

Table A.5: Land Concentration Effects by Share Tenancy Prevalence

	(1)	(2)	(3)	(4)	(5)	(6)
Odd × County Share Tenancy		-0.16***	-0.14***	-0.13***	-0.16***	-0.11**
		(0.044)	(0.048)	(0.039)	(0.043)	(0.042)
Odd Section	-0.045***	0.022		-0.0025	0.020	
	(0.014)	(0.024)		(0.022)	(0.023)	
Grant × State FEs	Y	Y	Y	Y	Y	Y
County FEs	Y	Y	Y	Y	Y	Y
Township FEs	Y	Y	Y	Y	Y	Y
SEs / Clusters	Spatial	Spatial	Spatial	Spatial	Spatial	Spatial
Odd × Grant × State			Y			Y
Odd × Low Soil Quality				Y		Y
Odd × Soil (linear)					Y	Y
N	132,463	132,463	132,463	132,463	132,463	132,463
$\mathbb{E}[y]$	5.8	5.8	5.8	5.8	5.8	5.8

Notes: This table tests for heterogeneous effects of land concentration (odd sections) by rates of share tenancy. The outcome is full-sample (asinh) total property value in 2017. Share tenancy prevalence is defined as the fraction of a county's tenant farms under share agreements in 1940. The specification replicates the even/odd comparison of equation (1) with heterogeneous effects for the listed variables. Column (1) replicates Table 4 column (6). Columns (3)-(6) allow for heterogeneity by state and grant area, a section having "low" (botton-quintile) soil quality, and a linear interaction for soil quality. Soil quality refers to the gSSURGO crop productivity index. The baseline effect of "odd section" is subsumed in specifications with fixed effects and omitted from those columns of the table.

Table A.6: Checkerboard Area Effects

Panel A: Main Estimates

	Baseline			Drop 1-Mile Donut		
	(1) (asinh) Total Value	(2) (asinh) Value Improvements	(3) (%) Any Improved	(4) (asinh) Total Value	(5) (asinh) Value Improvements	(6) (%) Any Improved
In Checkerboard [Even]	-0.098**	-0.14**	-1.27	-0.18***	-0.18*	-1.46
	(0.038)	(0.065)	(0.97)	(0.066)	(0.095)	(1.73)
Grant × State FEs	Y	Y	Y	Y	Y	Y
County FEs	Y	Y	Y	Y	Y	Y
Geo Controls	Y	Y	Y	Y	Y	Y
SEs / Clusters	County	County	County	County	County	County
N	32,520	32,520	32,519	27,223	27,223	27,222
N (clusters)	162	162	161	162	162	161
$\mathbb{E}[y]$	$1,673k	$997k	51%	$1,673k	$997k	51%

Panel B: Land Quality Balance

	(1) Soil (z-score)	(2) Slopes (z-score)	(3) Streams (z-score)	(4) Elevation (z-score)	(5) log(Area)	(6) log(RR Dist)
In Checkerboard [Even]	-0.012	0.0062	0.024	0.0018	0.0032	-0.0058*
	(0.014)	(0.0063)	(0.025)	(0.0059)	(0.0034)	(0.0033)
Grant × State FEs	Y	Y	Y	Y	Y	Y
County FEs	Y	Y	Y	Y	Y	Y
Geo Controls	N	N	N	N	N	N
SEs / Clusters	County	County	County	County	County	County
N	39,838	39,838	39,838	39,838	39,838	39,838
N (clusters)	162	162	162	162	162	162
$\mathbb{E}[y]$	.059	.88	.47	1.6	-.017	3.2

Panel C: Federal Settler Characteristics

	(1) Acres Granted	(2) Ever Granted (%)	(3) Public Land 2017 (%)	(4) Occ. Income	(5) Farm Home (%)	(6) Owns Home (%)
In Checkerboard [Even]	-12.3	1.28*	-0.84	-0.11	2.19	-1.65
	(9.72)	(0.76)	(0.71)	(0.40)	(1.76)	(3.31)
Grant × State FEs	Y	Y	Y	Y	Y	Y
County FEs	Y	Y	Y	Y	Y	Y
Geo Controls	Y	N	Y	Y	Y	Y
SEs / Clusters	County	County	County	County	County	County
N	24,128	32,520	31,762	7,109	8,158	2,967
N (clusters)	157	162	161	135	137	104
$\mathbb{E}[y]$	396 ac	58%	27%	14%	56%	71%

Notes: RD comparisons of federal sections per equation (2). Panel A considers (1)-(2) 2017 total and improvement value (3) owned acreage in 1900s assessments (4) lack of census microdata link to 1900s owner (5)-(6) number of distinct CropScape land uses and extensive margin of crop farming. Panel B considers the geographic characteristics analyzed in Table 3. Panel C considers (1) average acres per grant, top-coded at the 95th percentile (2) the percentage of land ever granted (3) the percentage of public land in 2017 (4)-(6) consider the characteristics of settlers linked to census microdata in the decade before their grant.

Table A.7: Distribution of Population Effects

	Panel A: Town Outcomes, Even/Odd					
	(1)	(2)	(3)	(4)	(5)	(6)
	# Towns CDPs	# Towns Schmidt (2018)	Pop $\geq$ 1 (%)	Pop $\geq$ 10 (%)	Pop $\geq$ 100 (%)	Pop $\geq$ 1000 (%)
RR Effect	0.00029	0.0010*	-3.63***	-1.02***	-0.046	0.0085
	(0.00024)	(0.00059)	(0.66)	(0.30)	(0.054)	(0.019)
Sample	All	All	All	All	All	All
Grant × State FEs	Y	Y	Y	Y	Y	Y
County FEs	Y	Y	Y	Y	Y	Y
Township FEs	Y	Y	Y	Y	Y	Y
Geo Controls	Y	Y	Y	Y	Y	Y
N	132,463	132,463	132,463	132,463	132,463	132,463
$\mathbb{E}[y]$	.024	.0039	33%	11%	3%	.58%

	Panel B: Value Outcomes, Even/Odd					
	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline	5% Top Code	RR Dist >2	No Towns (CDPs)	No Towns Schmidt (2018)	No Towns (Pop ≥100)
RR Effect	-0.045***	-0.045***	-0.046***	-0.045***	-0.047***	-0.046***
	(0.014)	(0.014)	(0.014)	(0.015)	(0.015)	(0.015)
Grant × State FEs	Y	Y	Y	Y	Y	Y
County FEs	Y	Y	Y	Y	Y	Y
Township FEs	Y	Y	Y	Y	Y	Y
Geo Controls	Y	Y	Y	Y	Y	Y
N	132,463	132,463	122,293	126,105	131,946	128,539
$\mathbb{E}[y]$	$2,231k	$920k	$1,687k	$1,185k	$2,093k	$901k

	Panel C: (asinh) Population Within X Miles					
	(1)	(2)	(3)	(4)	(5)	(6)
	1 mile	2 miles	3 miles	4 miles	5 miles	10 miles
RR Effect	0.0012	-0.0040	0.0044***	0.0032	-0.0013	0.0012
	(0.0033)	(0.0034)	(0.0012)	(0.0025)	(0.0018)	(0.0033)
Grant × State FEs	Y	Y	Y	Y	Y	Y
County FEs	Y	Y	Y	Y	Y	Y
Township FEs	Y	Y	Y	Y	Y	Y
Geo Controls	Y	Y	Y	Y	Y	Y
N	132,463	132,463	132,463	132,463	132,463	132,463
$\mathbb{E}[y]$	74	292	648	1,140	1,774	74

Notes: This table tests for effects of railroad land grants on town formation. Panels A-B use PLSS-section level data on the fraction of land in a census place, the Schmidt (2018) number of towns, and the satellite-based population. Panel A explores the even/odd comparison from equation (1). Panel B uses asinh total value as the outcome on equation equation (1). Columns (1)-(2) are the baseline result and a top-code of the outcome at the 5th percentile. Columns (3)-(6) drop different observations, specifically sections within 2 miles of the railroad and sections with towns defined as Census Designated Places, the Schmidt (2018) points, or having a population of more than 100. Panel C studies the (asinh) population in 2019 within specified distances of a random point within each section.

Table A.8: Public Goods and Political Behavior

	Panel A: Even/Odd				
	(1)	(2)	(3)	(4)	(5)
	Schools	Churches	Community Halls	Road Distance	Owner Seeks Office (%)
RR Effect	-0.014	-0.00022	-0.0010***	0.0021***	-3.61
	(0.0100)	(0.00078)	(0.00035)	(0.00076)	(5.05)
Sample	NE & KS 1940	NE & KS 1940	NE & KS 1940	All 2015	Morrill 1912
Grant × State FEs	Y	Y	Y	Y	Y
County FEs	Y	Y	Y	Y	Y
Township FEs	Y	Y	Y	Y	Y
Geo Controls	Y	Y	Y	Y	Y
N	18,622	18,622	18,622	132,463	82
$\mathbb{E}[y]$	.096	.013	.0025	1.1 mi	5.5%
	Panel B: In Checkerboard (Federal Only)				
	(1)	(2)	(3)	(4)	(5)
	Schools	Churches	Community Halls	Road Distance	Owner Seeks Office (%)
In Checkerboard [Even]	-0.011	0.00067	-0.00060	0.051	-5.21
	(0.012)	(0.0047)	(0.0011)	(0.042)	(3.45)
Sample	NE & KS 1940	NE & KS 1940	NE & KS 1940	NE & KS 1940	Morrill 1912
Grant × State FEs	Y	Y	Y	Y	Y
County FEs	Y	Y	Y	Y	Y
Geo Controls	Y	Y	Y	Y	Y
N	18,514	18,514	18,514	296,289	525
$\mathbb{E}[y]$	.086	.01	.00054	1.3 mi	4.9%

Notes: This table studies the presence of public goods, tax records, and officeseeking on PLSS sections. Panel A compares even and odd sections using equation (1). Panel B considers even (Homestead) sections within the grant area differ from those outside using equation (2). Columns (1)-(3) count the number of schools, churches, and community halls according to 1940 census enumeration district maps. (4) measures the distance from the section's centroid to the closest road in 2015. (5) uses an archival case study from Perkins, NE in 1900 which counts the number of months owners took to pay their property tax bill. (6) uses an archival case study from Morrill, NE in 1912 and computes the fraction of owners in a section who ran for county and sub-county elected office.

Table A.9: Property Rights and Legal Disputes

	(1) asinh(Value) Baseline	(2) asinh(Value) No NPRR	(3) asinh(Value) No NPRR, OCRR	(4) Recorded Lawsuit (%)
RR Effect	-0.045***	-0.049***	-0.047***	-2.63
	(0.014)	(0.014)	(0.014)	(5.54)
Grant × State FEs	Y	Y	Y	Y
County FEs	Y	Y	Y	Y
Township FEs	Y	Y	Y	Y
Geo Controls	Y	Y	Y	Y
SEs / Clusters	Spatial	Spatial	Spatial	Township
N	132,463	70,210	68,788	204
$\mathbb{E}[y]$	\$2,231k	\$3,003k	\$2,732k	28%

Notes: The table shows even/odd comparisons per equation (1). (1)-(3) replicate Table 4, Panel A. (2) drops the Northern Pacific Railroad grant and (3) additionally drops the Oregon and California Railroad grant. (4) analyzes the Banner County sales data with the outcome being the fraction of land in a section that ever experienced a lawsuit (*lis pendens* notice) during the period.