Railroads, Market Access, and Indigenous Land Dispossession*

Jeff Chan†, Azim Essaji‡, and Rob Gillezeau§

December 20, 2024

Abstract

The expansion of the United States across the North American continent displaced and dispossessed hundreds of Indigenous nations. Historians have long argued that this process was hastened by improved transportation links resulting from the construction of the railway. In this paper, we explore empirically the question of whether increased market access, as facilitated by railway building, accelerated the displacement and dispossession of Indigenous peoples and nations. Using spatial data containing all Indigenous land cessions up to 1894 in what is today the United States, we find that increases in market access increased the probability of Indigenous land cessions within that decade. Reductions in transport costs and growing populations in US counties both play important roles in explaining market access's overall effect. Finally, we find evidence that market access-induced land dispossession led to greater assimilation with settler populations, shorter heights for those born after dispossession, and modestly higher long-run incomes. Our findings indicate that increasing market access hastened the dispossession of Indigenous peoples and influenced on their socio-economic outcomes post-dispossession.

*We would like to thank the many scholars who have made their historical and spatial data publicly available that have helped us to engage in this research. All errors are our own.

†Department of Economics at Wilfrid Laurier University. Corresponding author: jchan@wlu.ca.

‡Department of Economics at Wilfrid Laurier University.

§Rotman School of Management at the University of Toronto.

1 Introduction

During the 19th century, the US experienced massive growth and change. In 1800, the entire population was approximately 5 million; by 1900, its population had increased to almost 76 million. This population boom was fuelled by the absorption of huge expanses of land. A country which was initially primarily focused along the east coast eventually stretched westward across the entire continent. The population's diffusion to these newly absorbed lands was facilitated by the expansion of the railroad network, which allowed Americans to traverse the breadth of their vast country in a matter of days by the end of the century. While the traditional narrative of this expansion focused on the pioneer experience and American exceptionalism, increasing attention has appropriately been placed on the Indigenous peoples who originally inhabited and continue to reside in these lands.

Indigenous peoples in the United States have clearly been the main victims of the United States' westward expansion. The increase in land belonging to the US was largely a consequence of the ceding of Indigenous peoples' lands and their subsequent relocation to reservations, often not altogether willingly. As white settlement expanded beyond the East Coast, Indigenous lands were increasingly coveted. In the South, for example, Indigenous peoples were displaced in large numbers to make way for more cotton plantations worked by slaves (Saunt, 2020). Historical narratives such as these support the notion that as Indigenous peoples' lands become more vulnerable to being taken as the frontier of settlement moved towards them. Despite this compelling narrative, there is little empirical work that has sought to provide evidence for or against this type of mechanism.

Our paper aims to fill this important gap in the literature by evaluating whether Indigenous peoples' land cessions were induced by proximity to US markets. We leverage changes in transport costs over time driven by the buildup and expansion of the railroad network during this period to obtain exogenous variation in market access for different land cessions. We then link these historical land cessions to Indigenous peoples' modern-day lands to investigate whether peoples who were displaced due to increasing access to US markets had different land characteristics in 2020.

We find that increased market access to US counties had a positive effect on a land cession being signed, in-

dicating that Indigenous peoples that were closer to US markets had a higher likelihood of being displaced. We also provide evidence that market access's effect on cession signings operated through both reductions in county-to-county transport costs brought about by the railroad expansion as well as growth in of markets size in US counties (as proxied by population). The latter result, in particular, suggests that land dispossession is increasingly likely when settlers are in closer proximity to Indigenous peoples. This finding is consistent with the historical narrative of lands being targeted due to their attractiveness to American colonizers. In contrast, we do not find any evidence that market access increased the occurrence of US-Indigenous conflicts.

We then employ a range of datasets to ascertain whether market access-induced land dispossession had any impact on Indigenous peoples' outcomes in the short and long run. We find some evidence that reservation-level incomes from 1945 to 2019 were higher in communities that had greater market access at the time of dispossession. Using US Census data, we show that Indigenous groups with higher market access at the time of their land dispossession had more children; these children were also more likely to be part white, suggesting that market access may have induced more assimilation and integration with settlers. We also find, using historical heights data for Indigenous individuals, that cohorts born after their homeland was affected by land cessions were relatively shorter if their homelands had higher market access at the time of cession signing; this suggests that market-access induced land dispossession had a particularly negative impact on Indigenous health relative to other land loss. Finally, we analyze the effect of historical market access at time of cession signing on modern-day outcomes, finding some modest positive effects on outcomes like the poverty rate and labor force participation; these results, however, could in part be driven by compositional shifts in populations for affected Indigenous peoples. Our results, taken together, provide clear evidence that railroad-driven market access increases played an important role to play in the dispossession of Indigenous lands, providing a sharp contrast to older romantic narratives of railroads paving the way for American spatial and population growth.

Our work expands the empirical knowledge of the causes and consequences of Indigenous peoples' displacement. Existing work in this area includes Carlos et al. (2022), who use descriptive data analysis to highlight the importance of Indigenous land dispossession in the American economic narrative. Feir et al. (2023a), similarly, use Canadian census data to explore the variables correlated with treaty signing and

show that Indigenous peoples whose lands were taken via treaties have worse modern-day outcomes. We contribute to this important discourse by exploring how market access could have impacted Indigenous land dispossession, which is a mechanism that has been unexplored in past work. In a companion paper (Chan et al. (2024)), we study the effects of the railroad expansion on Indigenous land dispossession via treaty-making in Canada; in contrast to the results in this paper, we find that the expansion of railways does not contemporaneously lead to the dispossession of Indigenous peoples. Rather, in Canada, dispossession of Indigenous' land via treaty occurred 10 to 20 years prior to the advent of the railway. This suggests that the while the railway was just as integral to Indigenous displacement in Canada, the process was more arguably more methodical and forward-looking. The contrast in our results also shows that the context is key to understanding whether railroads play a role in affecting the displacement of Indigenous peoples.

This paper also relates closely to a body of work on the effects of market access on various outcomes, particularly in a historical context. Donaldson and Hornbeck (2016) use railroad expansions in nineteenth century America and show that market access increased agricultural land values. Subsequent work from Chan (2022) shows that market access also increases farming output in affected US counties, with this effect operating chiefly through market access's effect on expansion of farmland and increasing population. Like this prior work, our paper adds to the understanding of how market access can lead to regional differences in outcomes and how it helped shape the United States in the nineteenth century.[1] Unlike these papers, however, our work focuses on the consequences of market access for Indigenous peoples' land, which has not been investigated empirically to our prior knowledge.

Finally, we also contribute to prior work on the persistence of historical circumstances on outcomes. Bleakley and Lin (2012), in one example of such work, show that historical portage sites, where rivers were non-navigable and required land navigation, continued to be sites where economic activity took place even after transport technology overcame the need for their original existence. Bazzi et al. (2020) show that the amount of time a US county spent on the frontier continues to affect the culture and attitudes of those living there. Feir et al. (2023b) show that Indigenous peoples who relied on the bison continued to have worse outcomes into the present, due to the virtual extinction of the species in the nineteenth century. Perhaps most

[1] Work that has studied the effects of market access in other countries include You et al. (2023), who look at market access in England and Wales between 1830 and 1911.

similarly, Chan (2024) finds that market access shocks faced by children in the early twentieth century US continued to affect their outcomes in later life. We add to this literature by providing evidence on whether the market access of Indigenous lands at time of dispossession affects the relative heights of Indigenous peoples' post-dispossession, their degree of assimilation into settler communities, their 20th century incomes, and the characteristics of their 2020 lands.

2 Historical Background

2.1 Dispossession of Indigenous Lands

This subsection provides a brief background to Indigenous land dispossession in the United States, focusing primarily on the 19th century. For a more comprehensive overview of the topic, readers can refer to Carlos et al. (2022) or to Saunt (2020), to whom we owe much for the subsequent discussion of the history of Indigenous land dispossession.

In 1789, the newly formed United States occupied a small fraction of the land mass it encompasses today. Much of what we recognize as the United States today was gradually acquired via a series of land acquisitions from Indigenous peoples. In some cases, these acquisitions were disorderly and occurred at the individual level while in others, these land acquisitions were handled with the view that Indigenous nations were sovereign nations and to be treated as such in negotiations. Treaties would transfer lands to the United States, which would then begin the often-lengthy process of surveying and selling to private buyers. The approach of negotiating with sovereign nations was stipulated in the 1790 *Non-Intercourse Act*.

This view of Indigenous sovereignty was slowly eroded over the 19th century, with the US government slowly recognizing Indigenous nations less as an equal in negotiations and more as an obstacle to overcome in expropriating lands for American settlers. Carlos et al. (2022) and Banner (2005), among others, argue that this shift was slowly achieved through a series of judicial decisions and legislation that gave the government increasingly far-reaching and unilateral authority in deciding the extent of Indigenous sovereignty, and by extension, the terms and conditions of land acquisition. The 1830 *Indian Removal Act*, for example, on paper was a simple authorization that allowed the US president to negotiate with tribes within current state and territory boundaries for their land in exchange for other lands west of the Mississippi. In practice,

however, the *Indian Removal Act* provided the legal pretext for the forced displacement of tens of thousands of Indigenous peoples in what would become known as the Trail of Tears (Saunt, 2020).

In addition to legislation such as the Indian Removal Act, a trio of Supreme Court decisions known as the Marshall Trilogy[2] also significantly redefined the relationship between the US and Indigenous peoples. First, in 1823, the Supreme Court ruled in a land dispute case between two white men that US citizens could not purchase land from Indigenous nations; in doing so, the Supreme Court effectively stated that only the US government would be able to acquire land from Indigenous nations in the future. This decision greatly reduced Indigenous nations' bargaining power. Next, in 1831 the Supreme Court heard a case brought by the Cherokee Nation against the state of Georgia. In this case, the Supreme Court ruled that it could not provide a decision as Indigenous peoples were not sovereign nations but a domestic dependent nation under the guardianship of the United States. This second ruling stripped a great deal of sovereignty away from Indigenous nations in their dealings with the US. In the final case of the trilogy, *Worcester* v. *Georgia*, the Supereme Court attempted to reverse the worst effects of its prior decisions (Breyer, 2000). It affirmed that individual states had no jurisdiction over Indigenous nations and territories, and that only the federal government could negotiate the extinguishment of Indigenous rights. The ruling in the case was blatantly ignored by the State of Georgia, and President Andrew Jackson avoided enforcement of the judgment long enough that Georgia was able to strip the Indigenous nation at the heart of the dispute, the Cherokee, of most of its land (Breyer, 2000).

Alongside this changing legal and legislative landscape, Indigenous peoples' sovereignty and property rights were increasingly infringed upon by the arrival and squatting of white settlers on their lands (Allen, 1991, 2019; Carlos et al., 2022; Allen and Leonard, 2021). This encroachment of settlers accelerated with the rapid expansion of the railroad network. While squatting was technically illegal and the US government had legal tools as its disposal to mitigate its extent, in practice settlers went largely undisturbed. In contrast to their nebulous legal status, settler-squatters became increasingly powerful in their ability to seize Indigenous lands. By the 1860s and 1870s, Indigenous nations were generally unable to stop white settlers and were increasingly pressured to give up their lands. This pressure was exacerbated by legislation like the 1862 *Homestead Act*, which guaranteed settlers ownership of their own plot of land provided that they settled it

[2]These were named for the Chief Justice at the time, John Marshall, the author of the majority opinions in the cases.

and improved it for five years. With much of these lands being Indigenous territory, the spurring of more settlers into these lands put additional pressure on Indigenous peoples to relocate to avoid political pressure and conflicts.

While the history of Indigenous land dispossession is far more complicated than what was just articulated, it is apparent that the US government became increasingly less willing to treat Indigenous peoples as having their own nations that should be treated accordingly. Instead, Indigenous peoples were treated more poorly and with less respect over time, which was enabled by the pressure placed upon them by white settlers, legislative and legal justifications, and the increasing connectedness of the country due to the railroad.

2.2 The US Railroad Network

One of the most enduring narratives of 19th century American history has been the pivotal role that the railroad played in extending the US government's influence and boundaries. In terms of timing, the correlation between the growth of the railroad network and the growth of the United States itself provides some compelling evidence for this story. In 1830, the railroad network was extremely limited in scope, with under 40 miles of track all located in the Northeast (see Figure 2). Construction of new track initially focused on expanding the nascent network in the heavily populated northeast, but by 1890 this network had become a huge, dense series of lines, with over 160,000 miles of track. At the same time, the US population as measured by the Census increased from 12.8 million to almost 63 million. The simultaneous occurrence of these two shifts have not gone unnoticed by economists and historians, who have long linked the two and attributed the nation's growth to the locomotive to varying degrees Allen (2019).

Railroad access dramatically changed the way in which goods and people moved in the US. As Donaldson and Hornbeck (2016) note, prior to railroads transcontinental journeys would have used a combination of ships on water (including steamships along rivers) and wagons along roads. Wagons, in particular, were costly, slow, and uncomfortable. The spread of the railroad meant that people could now move across large distances much more rapidly in a comfortable and relatively affordable manner. This vast improvement was encapsulated by the inaugural trans-continental rail trip from New York to San Francisco on June 4, 1876, taking a mere 83 hours. In comparison, President Jefferson's 225 mile trip from his home in Virginia to

Philadelphia took 10 days.[3] The railroad therefore represented a quantum leap in transportation technology for the time.

In addition to greatly easing the movement of people across space, the railroad also connected the frontier to previously distant markets. W.W. Rostow, in his 1960 book,[4] argues that the US railroad network "performed the Smithian function of widening the market". The network facilitated the movement of goods from agricultural regions to major cities on both coasts and to ports for export, particularly in the case of cotton. It also facilitated movements of goods for consumption towards the interior. Research on the "market access" channel suggests that it is a key way through which railroads have affected the economic growth of the United States during this period.[5]

3 Mechanisms

Donaldson and Hornbeck (2016) provide a theoretical framework to consider how market access might affect Indigenous peoples' land dispossession. Their model is based on a standard theoretical approach built from the framework of Eaton and Kortum (2002). In their model, which they use to describe the effects of railroad induced market access increases on US county-level agricultural land values, Donaldson and Hornbeck (2016) find that market access increased land values from 1870 to 1890. The channel through which this occurred was through increased trade opportunities with other counties, leading to more agricultural output and specialization. Our measure of market access in this paper, which is adopted from the data made available publicly by Donaldson and Hornbeck (2016), follows their approach and is theoretically consistent with their model.

Donaldson and Hornbeck (2016) do not factor Indigenous peoples in their analyses, but one could imagine several different ways in which they could be differentially affected by increased market access. For one, while it is possible that Indigenous communities may have seen some gains from better market access, it is likely that Indigenous peoples would not have benefited from increased trade opportunities with settler

[3] History.com editors (2009)

[4] Rostow (1960).

[5] Donaldson and Hornbeck (2016), for example, find that increased market access from railroad construction during the 1870-1890 period led to an increase in agricultural land values; they estimate that the removal of the railroad network as it existed in 1890 would have slashed agricultural land values by 60 percent.

populations as much as other settlers would have. A host of reasons could drive this difference, such as discrimination against Indigenous peoples in the form of refusing to do business with them, settlers charging higher prices for goods purchased by Indigenous peoples and only accepting lower prices for goods sold by Indigenous peoples relative to other settlers. If Indigenous peoples are also less able to access and use rail for the shipping of goods and people (due to discrimination or institutional barriers), then the benefits of railroad induced market access will be lower for them relative to settlers. One natural consequence of this heterogeneous effect of market access will be that Indigenous lands that see improved market access will be increasingly attractive to settlers, with their Indigenous inhabitants subject to higher pressure to cede the land.

Another mechanism at play could be one where market access leads to higher settler populations in nearby areas. Chan (2022), for example, finds that US (settler) populations increased in counties that experienced increases in market access between 1870 and 1890. Increases in market access, therefore, lead to closer proximity to settler populations for Indigenous peoples. Increased proximity, unfortunately, can lead to increased frictions between settlers and Indigenous peoples and more instances of squatting, where settlers move into sovereign Indigenous territory without any legal justification. This squatting can often serve as a pretext for more formal land dispossession, with the pressure exerted by these squatters forcing the hand of the rightful landowners to agree to land cessions.

4 Data and Methodology

4.1 Data

We employ a variety of historical and spatial data sources to evaluate the relationship between market access, land dispossession, and post-dispossession outcomes. For convenience, summary statistics of the key variables for the estimation sample of cession-years is provided in Table 1. The summary statistics of key variables for the cross-sectional data used in the analysis are reported in Table 2. For all GIS-based calculations, we employ the North American equal area conic projection for area-based calculations and the North American equidistant conic projection for distance-based calculations.

4.1.1 Land Cessions

Our primary outcome variable is the occurrence of Indigenous land cessions. GIS data on the shape and location of Indigenous land cessions are taken from the United States Forest Service, which hosts a digitized map of all ceded Indigenous lands in the United States up to 1894.[6] This information underlying the map are originally from the *Eighteenth Annual Report of the Bureau of American Ethnology to the Secretary of the Smithsonian Institution, 1896-1897*; the specific part of this publication used (Royce, 1897) is "Indian Land Cessions in the United States." In addition to the extent of historical land cessions, the Forest Service has also provided a linkage from each cession to the modern-day, federally-recognized nations that are associated with that cession. We use this linkage to connect historical cessions with the modern-day locations of these nations in our consideration of reservation land quality.

4.1.2 Homelands

In many of our short and long run results, and to test for the robustness of our main results to an alternative geography, we make use of boundaries of the traditional territories of Indigenous peoples as of 1600 using a map from Martin and O'Leary (1990). The shapefile of these territories is provided in the public replication package of Feir et al. (2023b).

4.1.3 AIANNH Areas

To link the above historical land cession data with the modern locations of Indigenous nations, we use the 2020 US Census version of the American Indian/Alaska Native/Native Hawaiian (AIANNH) Areas. The shapefile delineating the extent of these areas is available from the US Census website. These areas include the reservations, as well as lands held in trust for all federally and state recognized tribes as of the 2020 Census.

4.1.4 Market Access

To construct the market access measures, we use county-level population data from NHGIS (Manson et al. (2021)) for each Census year from 1830 to 1910. We combine these population data with time-varying county-to-county transport cost data from Donaldson and Hornbeck (2016). One caveat of this data is that,

[6]The Forest Service data includes both cessions and reservations, but no identifier to clearly separate the two. In the analysis, we distinguish cessions from reservations by using data on reservation status from Claudio Saunt's *Invasion of America* digital project.

as it is based on the county as the unit of analysis, unorganized territories such as Indian Territory do not show up in the Donaldson and Hornbeck (2016) data. This means that any geographies which overlap with unorganized Indian Territory (or any other non-county areas) have their market access calculated using only the areas that do overlap with counties. Since the transport cost data are defined using 1890 county definitions, we convert the population data to 1890 county boundaries using population-weighted crosswalks available from Ferrara et al. (2022).[7]

Figures 4 to 6 present maps of the cessions in our sample and the change in market access for each cession for the following three time periods, respectively: 1830-1890, 1830-1860, and 1860-1890. For comparison, depictions of the state of the US railroad network in 1830, 1860, and 1890 are presented in Figures 2, 3, and 4, respectively. The maps show that as the railroad networks expanded westward, so too did the cessions facing the largest increases to market access.

4.1.5 Conflicts

Data on conflicts between the US and Indigenous peoples are obtained from Anderson (2024).[8] The data he compiles are from three different sources. The first is the *Chronological list of actions, &c., with Indians from January 15, 1837 to January, 1891* (Adjutant-General's Office (1979)). This source contains the fewest conflicts and thus is labelled as "small" in both Anderson (2024) and our analysis. The second source (labelled as "medium" in our analysis), which contains the next-largest number of conflicts recorded, is *Chronological list of engagements between the Regular Army of the United States and various tribes of hostile Indians which occurred during the years 1790 to 1898, inclusive* (Webb (1976)). Finally, Anderson (2024) uses *Indian battles and Skirmishes on the American frontier, 1790–1898* (Peters (1966)), which is labelled as "large" due to its having the most number of conflicts. Anderson (2024) codes the conflicts to the state-year level, which we similarly adopt.[9] In our sample, we keep only states for which at least one conflict occurs in our sample period.

[7] More specifically, we use the M2 version of their weights. Crosswalk data from Ferrara et al. (2022) are available from https://doi.org/10.3886/E150101.

[8] We are very grateful to R. Warren Anderson for generously sharing his data with us.

[9] The geocoding and location descriptions for conflicts from these sources is often sparse and does not provide sufficient detail to allow for a more disaggregated geocoding to occur without additional sources.

4.1.6 Crop Suitability

We obtain measures of crop suitability for cotton, maize, and wheat from version 4 of FAO-GAEZ. Specifically, we use their predicted crop yield measures derived from their agroclimatic model using an intermediate level of technology and inputs and rain-fed irrigation; these levels have been previously used in Fiszbein (2022) and Chan (2022) in the same time period and setting to proxy for crop suitability. FAO-GAEZ's model then also inputs geographic factors such as soil conditions, climate, and wind and generates crop suitability measures for crops at a 0.083 by 0.083 degree cell resolution for the surface of the Earth. We take these measures and calculate mean crop suitability for individual land cessions and AIANNH Areas. In some cases, the size of either cessions or Areas were smaller than the resolution of the cells in the FAO-GAEZ data and would have resulted in a missing value for that geography. In these cases, which occur in only the smaller cessions and Areas, we impute the crop suitability by taking the mean of the 5 closest crop suitability cells.

The predicted crop yields from FAO-GAEZ are commonly used across multiple fields of economics as a convenient and arguably exogenous measure of productivity in different crops for small-scale geographies. In addition to the aforementioned Fiszbein (2022) and Chan (2022), Adamopoulos and Restuccia (2014) use the FAO-GAEZ data as measures of land productivity to facilitate cross-country comparisons of land quality and to study the aggregate productivity consequences of resource misallocation across farms. Another example is Cherniwchan and Moreno-Cruz (2019), who use predicted crop yields in maize to show that the introduction of maize to Africa increased population density in maize-productive areas. These settings and usages vary greatly in their time periods of focus, from the pre-colonial to the modern eras. One might therefore be concerned that the FAO-GAEZ predicted yields are not an adequate measure of crop productivity for all these time periods, especially since the measures were calculated relatively recently and as such many implicitly rely on modern-day agricultural methods or crops. We make some adjustments to mitigate such concerns but also believe these concerns not to be at odds with our intended use of the FAO-GAEZ data. First, we select the level of inputs to be low, so as to best remove the effects of endogenous input selection in crop suitability. We also choose irrigation to be provided only via rain-feeding; this removes the possibility that productivities can be driven by irrigation technologies. This mitigates concerns that improvements to irrigation methods could introduce wedges between the true underlying productivity from our setting and

what is measured in FAO-GAEZ.

We also emphasize that our objective in using the FAO-GAEZ data is not necessarily to obtain a quantitatively correct measure of crop productivity across geographies; such a measure might be more important in papers estimating the maximum potential output or conducting other similar quantitative exercises. Instead, our paper uses the predicted crop yields only as a proxy for how productive each area is in producing different crops relative to other areas. In addition, given that Chan (2022) finds that the FAO-GAEZ measures are strongly predictive of actual crop output by county in the US from 1870 to 1890 (a similar time period to our own), we are comfortable in asserting that the FAO-GAEZ data provide a reasonable facsimile of underlying productivities.

4.1.7 Railroads

To construct measures for distance from the nearest railroad as well as whether a cession has a railroad within its boundaries, we make use of the railroads shapefile from Atack (2016), which covers the extent of the US continental railroad network from 1826 to 1911. Figures 1 to 3 show the extent of the railroad network in 1830, 1860, and 1890 (the last year in the sample). The figures illustrate the rapid expansion of the network from very limited coverage only in the northeast in 1830 to a cross-continental network by 1890.

4.2 Short and Long Run Outcomes Data

The following data sources are used in our analyses of whether market access-induced land loss led to differences in outcomes in the short and long run for Indigenous peoples. These datasets are obtained from the replication files of Feir et al. (2023b) (with the exception of the CICD and AIANNH data), and are matched to homelands using the crosswalks from Feir et al. (2023b). For more information on the matching between homelands and the Feir et al. (2023b) data, see the Appendix.

4.2.1 IPUMS Indian Schedules

We use US Census microdata, available from IPUMS (Ruggles et al. (2024)), which contains tribal identifiers that we link to our homelands; historically, these identifiers were available only for 1900, 1910,

and 1930. For 1900 and 1910, this identifier, and more generally microdata for Indigenous peoples, were available only for those enumerated using the so-called Indian schedules of the corresponding decennial Censuses. In 1930 the tribe question was asked of Indigenous peoples who were enumerated using the normal Census enumeration form. Unfortunately there is subsequently a data gap in being able to identify Indigenous peoples and their tribes of origin until 1990, when the question is asked once again. We use the available years for analyses on short-run post-dispossession outcomes. Our analysis dataset is obtained in the replication files of Feir et al. (2023b). Tribal identifiers are matched to homelands using the linkage from Feir et al. (2023b).

4.2.2 Boas Heights Data

We use data on heights for Indigenous peoples originally collected between 1888 and 1903 by Franz Boas and other anthropologists. The data includes a range of biological information such as height, age, gender, and racial composition. The data were previously used by Feir et al. (2023b) (from whom we obtain the data used in this paper), and more information on the heights data can be found in Jantz (1995), who digitized the Boas data for use. Tribal identifiers are matched to homelands using the linkage from Feir et al. (2023b), as in the IPUMS data.

4.2.3 Long-Run Reservation Incomes

For our analysis of incomes through the 20th century, we use the updated panel of reservation-level income data in Feir et al. (2023b) , originally from Leonard et al. (2020). Leonard et al. (2020) compile their data from a range of sources, such as the American Community Survey and the Bureau of Indian Affairs. Dollar values have been converted to 2010 USD. Readers interested in additional details are referred to Feir et al. (2023b). Linkages to homelands are those from Leonard et al. (2020).

4.2.4 Data on Income Sources, 1915-1920

We use information on the percentage of incomes derived from various sources such as land sales, wages, and crop sales from digitized Bureau of Indian Affairs Annual Reports from 1915 to 1920. These digitized tables were originally processed by Feir et al. (2023b), who also linked the groups in this data to homelands used in our paper.

4.2.5 CICD Data

As an alternative method to examine modern-day outcomes without having to make use of homeland geographies and the necessary assumption that Indigenous groups were truly, consistently linked to the geographic areas in the homeland map, we also use data from the Center for Indian Country Development (CICD) at the Minneapolis Federal Reserve, from their Native Community Data Profiles.[10] The CICD compiles this data from a variety of publicly available sources, chiefly the 5-year American Community Survey data produced by the US Census. The data contain, for each AIANNH Area, a variety of variables such as population, median household income, share of the population for various education levels, poverty shares, labor force participation, and other outcomes of interest.

We take these AIANNH Area level variables and calculate the mean value of these variables for all federally recognized tribes (that have currently have lands affiliated with them) associated with a given cession. The crosswalking procedure is the same as that for the calculation of the modern-day area and crop suitabilities, described earlier. The only exception is that due to more frequent missing values in the CICD data, we calculate mean values for each cession using only tribes for which data is non-missing; if all tribes associated with a cession have missing data, then that cession's value for that given variable will also be treated as missing.

4.3 Market Access Measure

The measure of market access used in this paper is identical to the functional form used in Donaldson and Hornbeck (2016). Specifically, we use the following definition of market access:

$$MA_{ct} = \sum_{d \in C_{-c}} \tau_{cdt}^{-\theta} * N_{dt} \tag{1}$$

For a given county c in year t, market access is defined by the sum (over the set of all other counties C_{-c}) of each other county d's population (N_{dt}) multiplied by the inverse of the transport cost from d to the origin county c. The transport cost is assumed to be 1 plus the freight rate of the lowest cost route between counties c and d in year t, divided by the average price of transported goods. This average price is assumed to be 35, as in Donaldson and Hornbeck (2016) and based on the Fogel (1964) approach. The inverse of the transport

[10] These data were accessed in June 2024.

cost is also taken to the power of θ, a parameter from Donaldson and Hornbeck (2016) that represents the trade elasticity in their model. θ is assumed to be 8.22, as in Donaldson and Hornbeck (2016), who use data to estimate this elasticity using data from nineteenth century America.

To construct a cession level version of market access, we take the county-level measure above and use an area-based crosswalk to construct a weighted mean market access measure for each cession, where the mean and weights are based on the counties that fall within each given cession and the share of the cession's area that is made up of that county.

Figures 4 to 6 provide a visual illustration of the change in cession level market access (transformed via inverse hyperbolic sine) over 3 different periods: 1830-1890, 1830-1860, and 1860-1890. Similarly, Figures A1 to A3 in the Appendix provide visual illustrations of the change in transformed market access over the same three time periods.

4.4 Specifications

Using the above market access measure, we then estimate equations of the following form:

$$\mathbb{1}(signed_{ct}) = \beta_0 + \beta_1 a(MA_{ct}) + \gamma_c + \delta_t + \epsilon_{ct} \tag{2}$$

The coefficient of interest is β_1, which is an estimate of the effect of market access on the probability of cession signing. One important note is that market access is only available in decennial years due to the transport cost database from Donaldson and Hornbeck (2016) only being constructed for those years. To map our cession signing dates, we therefore consider whether a cession was signed by that decennial year. Whether a cession is signed in that decennial year is therefore our main left hand side measure in our baseline results.

We also explore specifications where we control for railroad access and distance to nearest railroad (as measured from a cession's centroid) in that censal year t, although we omit these in most specifications due to multicollinearity.[11] Cession fixed effects and year/decade fixed effects are also included.

[11] Railroad access and distance from the nearest railroad are strongly correlated with each other, and also with our measure of market access.

To examine whether market access's effects on Indigenous land dispossession occurred through the mechanism of increased armed conflict between the US and Indigenous peoples, we use state-level conflicts data obtained from Anderson (2024). We construct a state-level version of our market access measure to do so,[12] and estimate the following:

$$(battles_{st}) = \beta_0 + \beta_1 a(MA_{st}) + \gamma_s + \delta_t + \epsilon_{st} \tag{3}$$

Our specification regresses the number of conflicts occurring in the 10 years during and after the decade of market access being observed against our market access measure at the state level. We use both the count of battles (transformed by inverse hyperbolic sine) as well as a dummy variable equal to one if there are any battles in that subsequent decade for that state. Standard errors in this specification are clustered by state, and we include state fixed effects and year fixed effects.

We use individual-level data from the US Censuses of 1900, 1910, and 1930 to study the effects of market access at time of signing on family and labour market outcomes. In these exercises, we estimate specifications of the form:

$$y_{igt} = \beta_0 + \beta_1 asinh(MA_g) + X'_{igt}\beta_2 + \epsilon_{igt} \tag{4}$$

These regressions estimate the effect on market access at time of session signing for homeland g on individual level outcomes for individual i from homeland g during Census year t. We control for a large set of possible confounders, such as age, age squared, bison loss,[13], decade of cession signing fixed effects, region fixed effects, and year fixed effects. We also include a large set of control variables obtained from Feir et al. (2023b) that account for cross-homeland differences in geography and culture.[14] Standard errors are clustered by homeland, and observations are weighted using the person-level sample weights provided by IPUMS.

We next turn to the Boaz height data to estimate a difference-in-differences specification that leverages

[12] Specifically, we take the weighted average of the default county-level market access measure from Chan (2022) and Donaldson and Hornbeck (2016) for each state-year, where the weights are given by each county's land area share within a given state.

[13] This is the binary treatment variable taken from Feir et al. (2023b). For the remainder of this paper, we use bison loss to refer to this variable.

[14] For a full accounting of these control variables, see the Appendix.

cross-homeland variation in market access at time of cession signing and cohort exposure based on year of birth. Specifically, we designate cohorts born after a homeland's average year of cession signing as being treated. This leads us to specifications of the form:

$$height_{igt} = \beta_0 + \beta_1 asinh(MA_g) * treated_{ig} + \beta_2 treated_{ig} + X'_{igt}\beta_3 + \epsilon_{igt} \tag{5}$$

Our dependent variable is the height in centimeters of an individual i from homeland g as measured in year t. Our main regressor of interest interacts the mean market access of a tribe at time of cession signing with a treated variable that equals one if an individual was born after cession signing; this cohort variation allows us to include homeland fixed effects. We also control for year fixed effects, age, age squared, and a gender dummy. We include a dummy equal to 1 if the individual was born after bison loss, plus its interaction with a dummy for whether the homeland was affected by bison loss; these variables control for the effects of the slaughter of the bison. In most of our regressions, we also control for a set of geographic and cultural factors; the full details for these control variables can be seen in the Appendix. Standard errors are clustered by homeland.

We track the evolution of market access-induced land dispossession's effect on incomes over the 20th and 21st centuries using data from Leonard et al. (2020) and Feir et al. (2023b). These regressions make use of cross-homeland variation in market access at time of cession signing:

$$y_{gt} = \beta_0 + \beta_1 asinh(MA_g) + X'_{gt}\beta_2 + \epsilon_{gt} \tag{6}$$

The dependent variable in these regressions is reservation-level income for a homeland g in year t, as measured in various sources. We regress this income measure against the market access at time of cession signing for homeland g to determine the effect of market access on long-run incomes. We control for decade of cession signing fixed effects, and year fixed effects in all these regressions. In some columns, we control for an assortment of other confounders, such as bison loss. We also explore the time-varying effects of market access over our sample period by allowing the market access coefficient to vary by year in some additional reported estimates. Standard errors are clustered by homeland.

To explore whether tribes dispossessed due to market access had different income sources in the medium run, we use tribe-level income data by source from Feir et al. (2023b), sourced originally from *Bureau of Indian Affairs Annual Reports*. We use these data to estimate the following:

$$y_g = \beta_0 + \beta_1 asinh(MA_g) + \beta_2' X_g + \epsilon_g \tag{7}$$

Our dependent variable is the proportion of income coming from a given source, across the nine categories available in our data. Due to the smaller sample sizes in this dataset and the fact that our data only provides cross-sectional variation over what is essentially a single time period, we include only fixed effects for decade of cession signing as controls in these regressions. Standard errors are clustered by homeland.

One final specification examines the characteristics of each cession, linking each to the value of market access that was experienced by that cession during the decade of treaty signing:

$$char_c = \beta_0 + \beta_1 * a(MA_{c,t(c)}) + \delta_{t(c)} + \epsilon_c \tag{8}$$

Each observation now represents a single cession, measured in the decade that the cession signing occurs. This cross-sectional specification otherwise largely mirrors the main cession-year regressions from before. The market access measure is measured in the decade of signing for cession c, $t(c)$. Finally, we include time fixed effects for decade of signing.

5 Results

5.1 Main Analysis

Table 3 presents the main results on how market access affects the signing of cessions. The dependent variable in all columns is a dummy variable equal to one if the given cession was signed over by that year. In column 1, we first only include, along with the fixed effects specified in the earlier section, our market access measure. The coefficient suggests that market access increases the probability of a cession being signed; in other words, increasing an area's access to more distant, populated locations raises the likelihood of that area being ceded by Indigenous peoples. In column 2, we include a control for whether that cession area has

railroad access in that given year; this additional control helps mitigate concerns that increases to market access could be driven by local improvements to transport infrastructure or railroads, which could then be correlated to other positive economic shocks. Column 2's coefficient on market access remains similar in magnitude to that in column 1, suggesting that market access's effect is robust to such concerns. Column 3 further takes such local shocks into consideration by instead including a control for a cession's distance to the nearest railroad (transformed by the inverse hyperbolic sine). Again, the results for market access are very similar to the original coefficient in column 1. Column 4 includes both the control for railroad access and distance to the nearest railroad. Once again, the coefficient on market access is very similar. The coefficient of 0.0570 is large and significant, and is very similar to the coefficient of 0.0558 without any additional controls. For this reason, and to mitigate issues of multicollinearity between market access and the railroad variables, we take column 1 (without any additional controls) to be our preferred specification. The coefficient from column 1 implies that a one standard deviation increase in market access increases the probability of cession signing by 16 percentage points, or 0.32 standard deviations. This large coefficient implies that market access plays an economically meaningful role in determining when land is ceded by Indigenous peoples.

One concern readers might have is that using the land cession as the unit of observation is unnatural, and that the tribe or Indigenous group should be the unit of analysis instead. While this has significant appeal, not the least of which is simplicity, the use of cessions is the more reasonable choice in the main analysis. This is because many cessions in actuality are signed by multiple tribes or groups. In addition, many tribes signed multiple cessions; while this could be the US carving away successive parts of this tribe's territory, it could also reflect groups being pushed increasingly further away from American territory and off their traditional homelands. This implies that the use of a static, time-invariant homeland would misclassify such cases as belonging to the incorrect group.

Nonetheless, in Table 4 we conduct a homeland-level analysis similar to that in our complementary paper, Chan et al. (2024). We use homeland boundaries from the Smithsonian Handbook of the North American Indian, which has been linked to data in the Ethnographic Atlas (Murdock (1967)).[15] In this new analysis, we focus on three homeland-level outcomes: 1) a dummy variable equal to one if a homeland has some of

[15]These data were originally cleaned and used in Feir et al. (2023a).

its territory overlapping with a signed cession in a given year, 2) the amount of land (transformed by inverse hyperbolic sine) in a homeland overlapping with a signed cession in a given year, and 3) the share of land in that homeland overlapping with a signed cession in a given year. The resulting coefficients on the market access measure, presented in Table 4, are all positive and strongly statistically significant. This confirms that even at the homeland level, market access has a strong, positive effect on cession signing.

In Table 5, we probe the channels through which market access affects cession signing. In particular, we note that market access is a combination of both market size of other counties (as proxied by population) and the transport costs to those counties. In principle, one of these two channels could be the chief source of variation through which market access affects land dispossession. In Chan (2022), for example, market access affects agricultural output mainly through reductions in transport costs to other counties. In column 2 of Table 5, therefore, we construct and use a measure of market access where only the transport costs vary; in column 3, we hold the transportation costs constant and only allow the populations of other counties to vary.[16] We find that both versions of market access continue to exert an effect on cession signing, suggesting that both the size of US markets and the transport costs to these markets matters in inducing Indigenous peoples to cede their lands.

Column 3's estimate is consistent with the historical narrative that the increasing proximity and number of settlers near Indigenous lands placed pressure on Indigenous peoples to be displaced. To test this mechanism more specifically, we construct measures of population proximity. In columns 4 and 5, we construct measures of the US population within 100 and 200 kilometers of a cession.[17] Consistent with the settler narrative, we find that cessions which had increasingly more settlers nearby were more likely to be signed. We therefore view columns 4 and 5 of Table 5 as providing evidence to support the role of settler pressure on Indigenous land dispossession.

Another related mechanism through which market access could have affected Indigenous land dispossession is in the increasing probability of armed conflict between Indigenous peoples and the United States. To

[16] Column 1 of Table 5 reproduces our preferred estimate from Table 3 for reference.

[17] We construct these measures similarly to our main market access measures. More specifically, we construct the population proximity measures at the US county level and then take the same weighted average approach to our market access measures to construct cession level equivalents.

test this mechanism, we use conflicts data from Anderson (2024) and a state-year specification. The results for this analysis are reported in Table 6. In columns 1 to 3, we report, for each of the three categories of sources reported in Anderson (2024) (small, medium, or large), the effects of market access on the number of conflicts. We do not find any statistically significant effect, with our coefficients being fairly small in magnitude. For reference, using column 1's estimate a one standard-deviation increase in transformed market access is associated with a 0.16 standard deviation decrease in transformed battles, using the "small" source for conflicts data.[18] We find similarly insignificant and small coefficients when turning to columns 4 to 6, which again examine conflicts from the three different sources of conflict data but use a dummy variable for the occurrence of any conflicts during the period. Given the consistency of the lack of significance and the size of coefficients, we conclude that market access does not appear to induce more conflicts between the US and Indigenous peoples and therefore do not find that increases in the likelihood of armed conflict is a channel through which land dispossession was hastened.

We then turn to analyzing whether market access at the time of cession signing is related to any cession level characteristics. We do this to explore whether cessions signed at times of high market access are different in any way to cessions signed that faced lower market access. Table 7 explores whether, conditional on cession signing, the market access at time of signing is correlated with cession characteristics.

Columns 1 to 3 in Table 7 show that cessions signed when market access is higher tend to have higher predicted yields in wheat, maize, and cotton. This suggests that market access tends to disproportionately affect high-fertility areas. In column 4, we show that areas that Indigenous nations signed away as market access increased tended to be smaller in area than those ceded when market access was lower. The results in column 5 suggest that increased market access did not induce more radical displacement of Indigenous peoples. While the railway may have hastened cessions, the lands to which Indigenous nations were relegated were no further away. One potential explanation for the area estimate is that Indigenous peoples may have leveraged settlers' desire for their valuable lands to negotiate more favourable terms, giving away less land.

[18] For summary statistics of the conflicts state-year dataset we use for this analysis, see Table A3 in the Appendix.

5.2 Short and Long-Run Outcomes

To trace the effects of market access induced land dispossession on Indigenous peoples, we study a broad range of outcomes with a number of datasets. First, in Tables 8 and 9, we make use of US Census micro-data from 1900, 1910, and 1930 to investigate whether those who experienced larger market access in the period when their land was dispossessed differed in terms of their family compositions and labour market outcomes. For a table of summary statistics of our sample used for this analysis (i.e. Tables 11, 10, and 12), as well as for our other analyses relying on data from Feir et al. (2023b), see Table A4 in the Appendix.

Table 8 focuses on outcomes related to family composition and background. Our results are consistent with small or no effect of market access on the proportion of children who survive (Column 1) or in the likelihood of gender composition of children being different for those children born after cession signing (Column 3). In contrast, we find a modestly sized, but precisely estimated, effect of market access at time of session signing on Indigenous peoples' fertility, as measured in the number of children ever born. The first three columns suggest that there is some effect on family composition for Indigenous individuals affected by market access induced land dispossession. In our final two columns of Table 8, we explore whether market access at time of signing induced more Indigenous people to assimilate into settler populations by intermarrying and having children who are part-white. As in column 3, we focus our attention only on those individuals born after cession signing to isolate only variation that could have arisen from market access induced land dispossession. Our estimate in column 4, where we use the proportion of white blood in a respondent as the outcome variable, shows a strong, positive association between market access and white blood share. Similarly, in column 5 we instead use a dummy variable equal to one if the respondent has zero percent white blood and find similar results, with a strongly significant negative coefficient. This is consistent with the idea that market access may have induced Indigenous peoples to integrate into settler populations.

Table 9 then turns to Indigenous individuals' labour market outcomes. More specifically, we make use of a measure of occupational income (standardized to a z-score), a dummy for having an occupation of employment recorded, a dummy for having an industry of employment recorded, a dummy for being employed in agriculture, and a dummy for being employed in a capital intensive industry. Across all these measures,

we do not find any strong or statistically robust evidence that market access at the time of dispossession affects Indigenous individuals' labour market outcomes in the medium run.

Next, we employ a cohort-based difference-in-differences design as outlined in our methodology section using the Boas height data to determine whether tribes whose lands were ceded while market access exposure was high had differential heights in those born after cession. These results are reported in Table 11. In column 1, we employ our fixed effects but omit our vector of geographic and cultural controls. In this specification, we estimate a negative coefficient on the interaction between market access and being in a treated cohort that is born post-cession, although the coefficient is not statistically significant at conventional levels. When adding a set of control variables in column 2, however, we find a negative and now statistically significant effect of market access at time of signing on heights for Indigenous individuals born post-cession. This also holds when the sample is divided into male and female individuals in columns 3 and 4, respectively. Both male and female Indigenous individuals are relatively shorter if born to treated cohorts and from homelands which had higher average market access at the time of cession signing. In columns 5 and 6, we explore whether treated cohorts born nearer to the cession signing timing were more affected by market access. Specifically, in column 5 we drop all treated cohorts that were born more than 10 years post-cession; our results are strongly significant and remain negative. In contrast, in column 6, when we drop all treated cohorts born within 10 years of cession signing, our estimated treatment effect becomes much smaller in magnitude and is now insignificant at conventional levels. Given that the sample size of this latter group is much larger, we conclude that columns 5 and 6 are consistent with market-access induced land dispossession creating worse conditions for individuals born in the immediate aftermath of the cession, leading to lower observed heights.

Table 10 uses data, obtained from Feir et al. (2023b), from Bureau of Indian Affairs Annual Reports from 1915-1920 to dissect reservation-level incomes into different sources. Specifically, income is broken down into the following origins, using the consistent categories in Feir et al. (2023b): land sales, leases, wages, livestock sales, rations, crops, traditional industries, and treaties and interest. We also examine the effects of market access at time of cession on per capita income, transformed by inverse hyperbolic sine. Our results suggest that there is minimal impact on per capita incomes overall, although there is some compositional shifting occurring across sources. For example, our estimates suggest that Indigenous peoples who had

higher market access in their lands at time of cession signing relied relatively more on leases and crops as sources of income, and correspondingly relied less on land sales, rations, and income from treaties and interest. One potential implication of these results is that market access may have shifted Indigenous peoples to income sources that are arguably more independent and less tied on the US government or any treaties or agreements.

We next turn to using the reservation-level incomes from Leonard et al. (2020) and Feir et al. (2023b) from 1945 to the 21st century to look at whether market access during the time of cession signing had adverse effects on Indigenous groups over the long run. Panel A of Table 12 first estimates a single market access coefficient and pools all years of our data together. Each column of the table includes a different set of control variables; varying the set of controls allows us to better understand the mechanism through which Indigenous peoples are affected by market access. Panel A, when considering all estimates presented, provides some evidence that market access-induced dispossession modestly increases reservation-level incomes over the long run. When looking at some of the control variable sets, some groups appear to drive the market access coefficient closer towards zero than others. In particular, of the control variables that are arguably more endogenous and less time-invariant (like the geographic controls), we find that the inclusion of governance and disruption based controls is particularly impactful.[19] This implies that one channel through which market access in ceded lands may have transmitted through to Indigenous peoples in the long run may be through increased conflict, worsened governance, and a higher likelihood of forced co-existence in the modern era.

We then further disentangle the effects of market access in Panel B by estimating year-specific coefficients. The year-specific results largely corroborate the pooled estimate from Panel A, which show that there may be a modestly positive effect of market access at time of cession signing on long-run incomes. Many coefficients are not precisely estimated, although the coefficients for 1969, 1979 and 1999 are somewhat more statistically significant but of the same general magnitude as the others. We therefore conclude, based on the cumulative evidence shown in Table 12, that market access may have induced slightly higher reservation-level incomes. We caution, however, that this effect may be driven in some part by differential compositional

[19] Governance variables include whether a nation had a constitution under the Indian Reconstitution Act and a measure of self-governance. Our disruption measures include a measure of the number of battles within a nation's territory before the bison slaughter and a measure of forced co-existence with other groups taken from Dippel (2014).

shifts in Indigenous groups whose lands were ceded as a result of increased market access: in other words, market access may have led to changes in the demographics and other characteristics of Indigenous peoples. This is particularly plausible given our finding with the IPUMS data that showed that market access may have increased the likelihood of having part-white children.

Table 13 links the historical cessions made by Indigenous peoples to their 2020 lands by using data from the US Census Bureau. We then explore whether a cession's market access at the time of signing affects the cession signees' 2020 land characteristics. Specifically, we examine whether there is a relationship between market access and four 2020 land characteristics: wheat suitability, maize suitability, cotton suitability, and the size of the modern-day lands in square kilometers. In the first four columns, we include only the market access variable on the right hand side, as in the initial analyses. We find that market access is associated with higher wheat suitability and higher maize suitability of present-day reservations.

Columns 5 to 8 of Table 13 estimate the same specifications as in columns 1 to 4 but additionally control for the land characteristics of the land already ceded by the Indigenous nation; these specifications therefore allow us to study whether market access affects modern-day land characteristics while separating out any potential correlation between historical and modern land characteristics. With the cession controls, we find that market access of a cession at the time of signing increases the wheat suitability of the associated tribes' 2020 lands.

The result that market access induced cession signings have higher wheat suitability is initially surprising but likely is an artifact of Indigenous peoples being moved from their homelands to areas initially designated for resettlement such as the so-called Indian Territory (modern-day Oklahoma) which may have had higher wheat suitability. This geographical pattern of forced resettlement, coupled with wheat not being a major exported commodity of the United States until the late 1860s, suggests that the viability of growing wheat was not a prime consideration in what lands market access affected in terms of displacement.

We also perform some additional analyses to test the robustness of the wheat result and the crop suitability analysis in Table 13 more generally, using a dummy variable equal to one if the cession suitability is higher than the modern-day land suitability as the dependent variable instead; these results are reported in

Appendix Table A1. The coefficients reported for wheat suitability are still statistically significant, although less precisely estimated.

We next turn to studying whether market access for Indigenous peoples at the time of their land cessions' signing affects their modern-day outcomes. To do this, we link each of the federally recognized tribes for each cession to the US Census AIANNH Area for that tribe and then subsequently to Area-level variables from CICD. In Table 14, we first study the effects of market access at the time of cession on present-day population and population composition. In columns 1 and 2, we find that there is a modest negative effect on total and Indigenous populations.[20] These effects, however are not significant at conventional levels. In column 3, however, the population effect appears to be scaled to the land area of the modern-day Area since population density is unaffected by market access at time of signing. We similarly find no meaningful or statistical effect on the median age for either the population as a whole or of Indigenous peoples. We find a modest negative effect on those with less than a high school education, but only for the total population in column 6; the coefficient for the share of Indigenous peoples with less than a high school education is also negative (in column 7), but is not significant. Finally, we do not find any statistically significant effects for the share of those with a bachelor degree or higher, either for the total or Indigenous populations. Taken together, Table 14 suggests that there are no large effects of historical market access on modern-day population or composition; at most, one can conclude that market access at the point of cession is negatively correlated with the share of the total population with less than a high-school education.

Next, Table 15 examines the effects of cession-level market access at time of signing on labor force and income-related outcomes. We divide the table into two panels that separately present results for both the Indigenous population only (in Panel A) and for the population of the modern-day Areas as a whole (in Panel B). In summary, we find some modest, positive effects of market access on several outcomes. We find, for example, a slight increase in the employment-to-population ratio and the labor force participation rate. While the number of unemployed appears to weakly increase in column 2, this is likely due to the increased labor force participation found in column 3. We also find a slight increase in median household income and a corresponding decrease in the poverty population share for both population types. We also examine

[20] We use the AIANa definition of Indigenous peoples available in the data, which are those who report in the ACS that they are of American Indian or Alaska Native origin.

the share of those living in various housing arrangements like those in owner-occupied housing, in crowded housing, or in mobile homes; we do not find any meaningful or significant effects for these outcomes. We therefore conclude, on the basis of the estimates in Table 15 suggest some positive effects of cession-level market access on modern-day outcomes for Areas associated with that cession.

There are some caveats to keep in mind in relation to the results in Tables 14 and 15. First, the modern-day outcomes are far removed from the timing of the cessions themselves. This means that our modern-day analysis ignores any potential dynamics or negative effects in the intervening years. We also do not have data on all tribes or peoples associated with a given cession; the analysis links only currently federally recognized tribes that have land and outcomes captured by the CICD data. All other peoples, including some which may have suffered the most, would not be included in our analysis. Finally, the analysis potentially includes changes in the composition of peoples living in the modern-day Areas which could shrink negative effects or even result in a positive association between outcomes and market access. For example, if the worst-affected peoples move away from Areas that had higher market access in the associated cessions, this could result in a compositional shift which could improve aggregate outcomes.

5.3 Robustness Checks

Table 16 contains several robustness checks for the main results in Table 3.[21] First, in column 1 we expand the sample to include cessions made prior to 1830 and after 1890. We find that the resulting coefficients are similar, maintaining the same sign and level of statistical significance from the main table of results. In column 2, we change the level of clustering for the standard errors from the cession level to the tribe level. Since tribes, or groupings of tribes, may have signed multiple land cessions, the original level of clustering may have ignored possible correlations in errors between these types of cessions as well. Using the data's classification of tribes,[22] The estimates from column 2 remain statistically significant at conventional levels, suggesting that different choices in the level of clustering do not affect the results in this paper. In column 3, we replace all variables transformed using the inverse hyperbolic sine with the log-transformed equivalents instead. The results remain of the same sign, magnitude, and significance as in the original results from

[21] We also re-run our robustness checks from Table 16, with the exception of the tribe clustering, on the homeland-level sample. These results are reported in Table A2, and similarly corroborate our main estimates.

[22] Multiple tribes are classified as a single "tribe" in the data if they jointly were associated with a given cession. Since these same groupings of tribes often gave up multiple land cessions, we use these tribe groupings as tribes as in the data.

Table 3, suggesting that our results remain robust to alternative functional form specifications.

In columns 4 and 5 of Table 16, we replace the market access variable with a single-lagged version of market access (in column 4) or with a lead version of market access (in column 5). The results in the lagged specification are of roughly similar magnitude and are of similar significance to the main results, although the coefficient is slightly smaller. The lead variable's coefficient, although significant at the 10% level, is even smaller than the lagged version and is less than half the size of the original market access coefficient; Given the size of the coefficient, and the fact that the lead version of market access is highly correlated with the original measure, we therefore conclude that our main results are not chiefly driven by future shocks.

One potential explanation for our main results is that land cessions could simply be preceding eventual railroad access, as lands are seized in order to facilitate railroad construction. In column 6 of Table 16, we therefore restrict our sample to cessions that do not receive railroad access by 1890, the end of our sample period. This restricted sample should not be susceptible to concerns of land seizures being an initial necessary step in railroad construction. This restriction, however, reduces our sample to 357 cession-years, which is less than a quarter of our original sample size. Nonetheless, the coefficient on market access is very similar in magnitude and significance to the main results; this similarity implies that estimated effect of market access is not reliant on railroad access itself. Rather, our results are more consistent with market access operating via increased connectivity to US markets over time.

In column 7 we instrument market access using a measure of market access that is constructed using only 1830 populations and counties that are at least 500 kilometers away. This instrumental variable approach was previously used in Chan (2024); for a more in-depth discussion of the construction of the instrument, see the Appendix. The instrument itself is very strongly predictive of market access, with a first-stage F-statistic for the excluded instrument of 25467.[23] The resulting coefficient in column 7 is very similar to our main estimates in sign, magnitude, and significance, suggesting that our results are not reliant on variation in populations or in market access from nearby counties.

[23]The coefficient on the first-stage for the instrument is 0.99.

6 Conclusion

In this paper, we find that increasing access to US markets caused by the expansion of the railroad network affected whether Indigenous lands were ceded to the United States. We also find evidence that market access-induced cessions had higher predicted crop yields than lands ceded with lower market access, suggesting that they may have been targeted especially because of their attractiveness to settlers. We also show, using a variety of data sources, that market access at the time of land dispossession is associated with lower heights for those born after dispossession, particularly for those born closest to the time of dispossession. Market access is also associated with having more children who are part white in the early 20th century, and modestly higher reservation incomes in the long run. Finally, we link market access faced at the time of each land cession to the associated Indigenous peoples' 2020 lands and find some evidence that modern-day land characteristics are also affected. Taken together, the evidence in this paper suggests that market access not only affects the taking of Indigenous peoples' lands but also their long-run outcomes.

References

Adamopoulos, T. and D. Restuccia, "The Size Distribution of Farms and International Productivity Differences," *American Economic Review*, 2014, *104* (6), 1667–1697.

Adams, H., *Report on transportation business in the United States at the eleventh census: 1890*, Government Printing Office, 1894.

Adjutant-General's Office, *Chronological list of actions, c., with Indians from January 15, 1837 to January, 1891*, Old Army Press, 1979.

Allen, Douglas W, "Homesteading and Property Rights; Or," How the West Was Really Won"," *The Journal of Law and Economics*, 1991, *34* (1), 1–23.

_ , "Establishing economic property rights by giving away an empire," *The Journal of Law and Economics*, 2019, *62* (2), 251–280.

_ **and Bryan Leonard**, "Late Homesteading: Native Land Dispossession through Strategic Occupation," *American Political Science Review*, 2021, pp. 1–15.

Anderson, W.A., "Politics, economics and Native American conflicts," *Asia-Pacific Economic History Review*, 2024, *64* (1), 10–33.

Atack, J., "Historical Geographic Information Systems (GIS) database of U.S. Railroads for 1826-1911," 2016.

Banner, S., *How the Indians Lost their Land: Law and Power on the Frontier*, Harvard University Press, 2005.

Bazzi, S., M. Fiszbein, and M. Gebresilasse, "Frontier Culture: The Roots and Persistence of "Rugged Individualism" in the United States," *Econometrica*, 2020, *88* (6), 2329–2368.

Bleakley, H. and J. Lin, "Portage and Path Dependence," *Quarterly Journal of Economics*, 2012, *127*, 587–644.

Breyer, Stephen, "The Cherokee Indians and the Supreme Court," *Journal of Supreme Court History*, 2000, *25* (3), 215–227.

Carlos, Ann M., Donna L. Feir, and Angela Redish, "Indigenous Nations and the Development of the U.S. Economy: Land, Resources, and Dispossession," *The Journal of Economic History*, 2022, *82* (2), 516–555.

Chan, J., "Farming Output, Concentration, and Market Access: Evidence from the Nineteenth Century American Railroad Expansion," *Journal of Development Economics*, 2022, *157*.

_ , "The long-run effects of childhood exposure to market access shocks: Evidence from the US railroad network expansion," *Explorations in Economic History*, 2024, *91*.

_ **, E. Essaji, and R. Gillezeau**, "Railroads, Treaty-making, and Indigenous Nations: Evidence from Canada," 2024.

Cherniwchan, J. and J. Moreno-Cruz, "Maize and Precolonial Africa," *Journal of Development Economics*, 2019, *136*, 137–150.

Dippel, Christian, "Forced Coexistence and Economic Development: Evidence From Native American Reservations," *Econometrica*, 2014, *82* (6), 2131–2165.

Donaldson, D. and R. Hornbeck, "Railroads and American Economic Growth: A "Market Access" Approach," *Quarterly Journal of Economics*, 2016, *131* (2), 799–858.

Eaton, J. and S. Kortum, "Technology, Geography, and Trade," *Econometrica*, 2002, *70* (5), 1741–1779.

Feir, D., R. Gillezeau, and M. Jones, "Institutional Drift, Property Rights, and Economic Development: Evidence from Historical Treaties," 2023. NBER Working Paper No. 31713.

_ , _ , and _ , "The Slaughter of the Bison and Reversal of Fortunes on the Great Plains," *Review of Economic Studies*, 2023.

Ferrara, A., P. Testa, and L. Zhou, "New Area- and Population-based Geographic Crosswalks for U.S. Counties and Congressional Districts, 1790-2020," 2022.

Fiszbein, M., "Agricultural Diversity, Structural Change and Long-Run Development: Evidence from the U.S.," *American Economic Journal: Macroeconomics*, 2022, *14* (2).

Fogel, Robert William, *Railroads and American economic growth*, Johns Hopkins Press Baltimore, 1964.

Gregg, M., "The long-term effects of American Indian boarding schools," *Journal of Development Economics*, 2018, *130*, 17–32.

History.com editors, "Express train crosses the nation in 83 hours," 2009.

Hornbeck, Richard, "The Enduring Impact of the American Dust Bowl: Short- and Long-Run Adjustments to Environmental Catastrophe," *American Economic Review*, 2012, *102* (4), 1477–1507.

Jantz, R., "Franz Boas and Native American biological variability," *Human Biology*, 1995, *67* (3), 345–353.

Leonard, B., D. Parker, and T. Anderson, *Journal of Development Economics*, 2020, *143*, 102435.

Manson, S., J. Schroeder, D. Van Riper, T. Kugler, and S. Ruggles, "IPUMS National Historical Geographic Information System: Version 16.0 [dataset]," 2021.

Martin, M. and T. O'Leary, *Supplement to the 4th Edition, Volume 1: Indexes of Ethnographic Bibliography of North America*, Human Relations Area Files Press, 1990.

Murdock, George P., *Ethnographic Atlas*, University of Pittsburgh Press, 1967.

Peters, J., *Indian battles and skirmishes on the American Frontier, 1790–1898*, Argonaut Press LTD, 1966.

Rostow, W., *The Stages of Economic Growth*, The University Press, 1960.

Royce, C., *Indian Land Cessions in the United States*, US Government Printing Office, 1897.

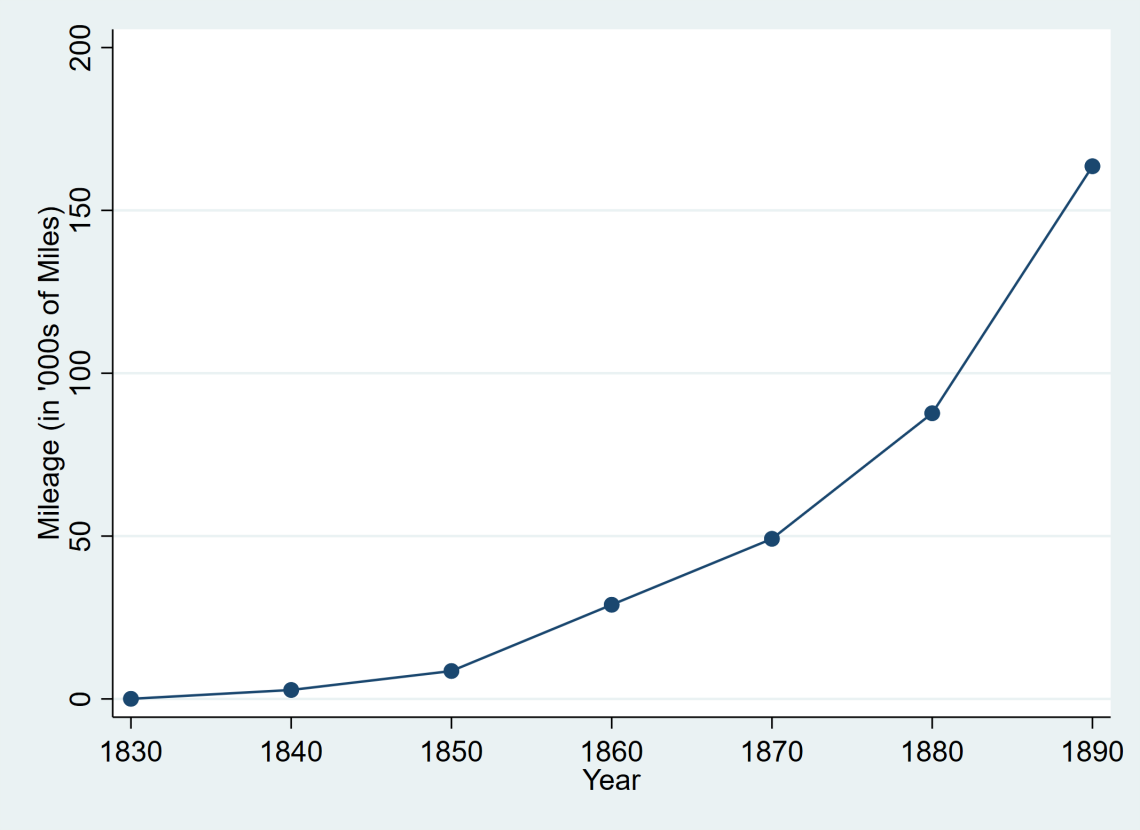
Ruggles, S., S. Flood, M. Sobek, D. Backman, A. Chen, G. Cooper, S. Richards, R. Rodgers, and M. Schouweiler, "IPUMS USA: Version 15.0 [dataset]," 2024.

Saunt, Claudio, *Unworthy Republic*, W.W. Norton, 2020.

Webb, G., *Chronological list of engagements between the Regular Army of the United States and various tribes of hostile Indians which occured during the years 1790 to 1898, inclusive*, W.W. Norton Co., 1976.

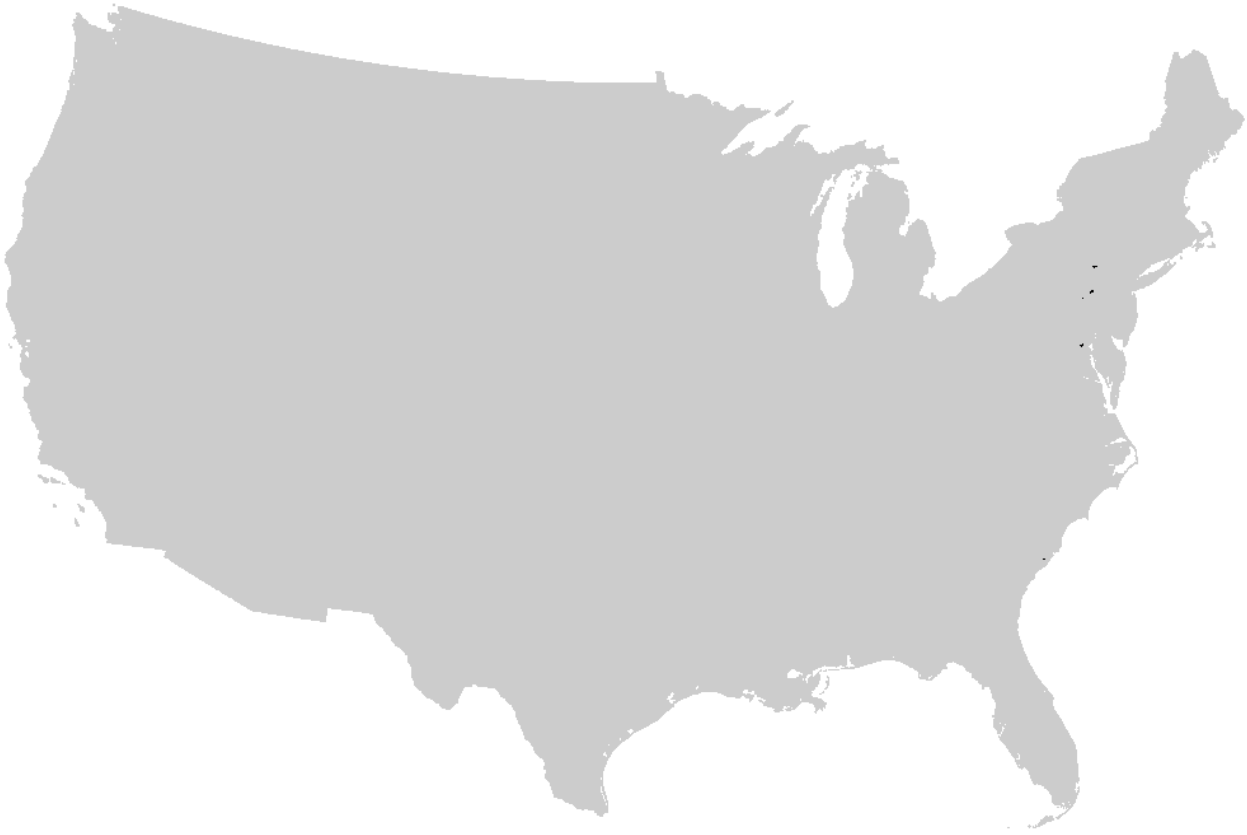
You, X., D. Bogart, E. Alvarez, M. Satchell, and L. Shaw-Taylor, "Transport Development and Urban Population Change in the Age of Steam: A Market Access Approach," 2023.

Figure 1: US Railroad Mileage,1830-1890



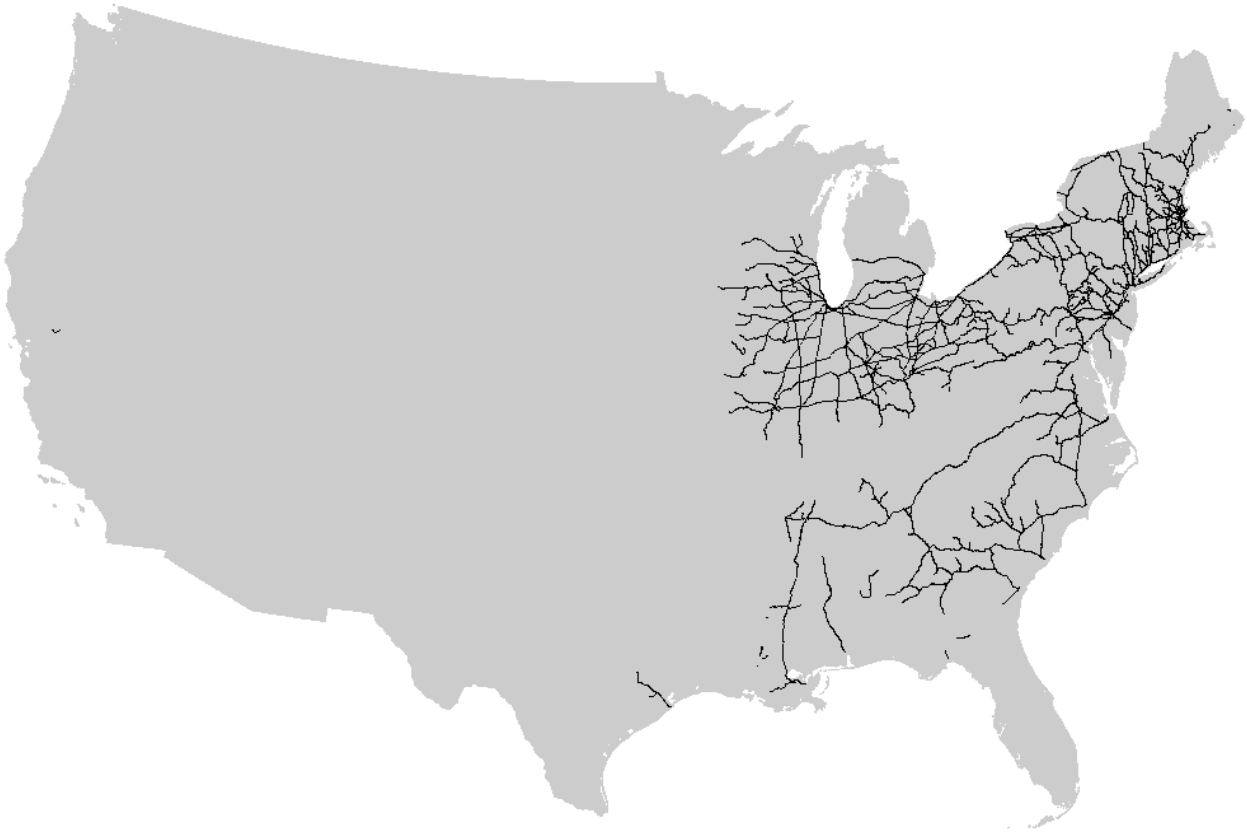
Source: Adams (1894).

Figure 2: US Railroads, 1830



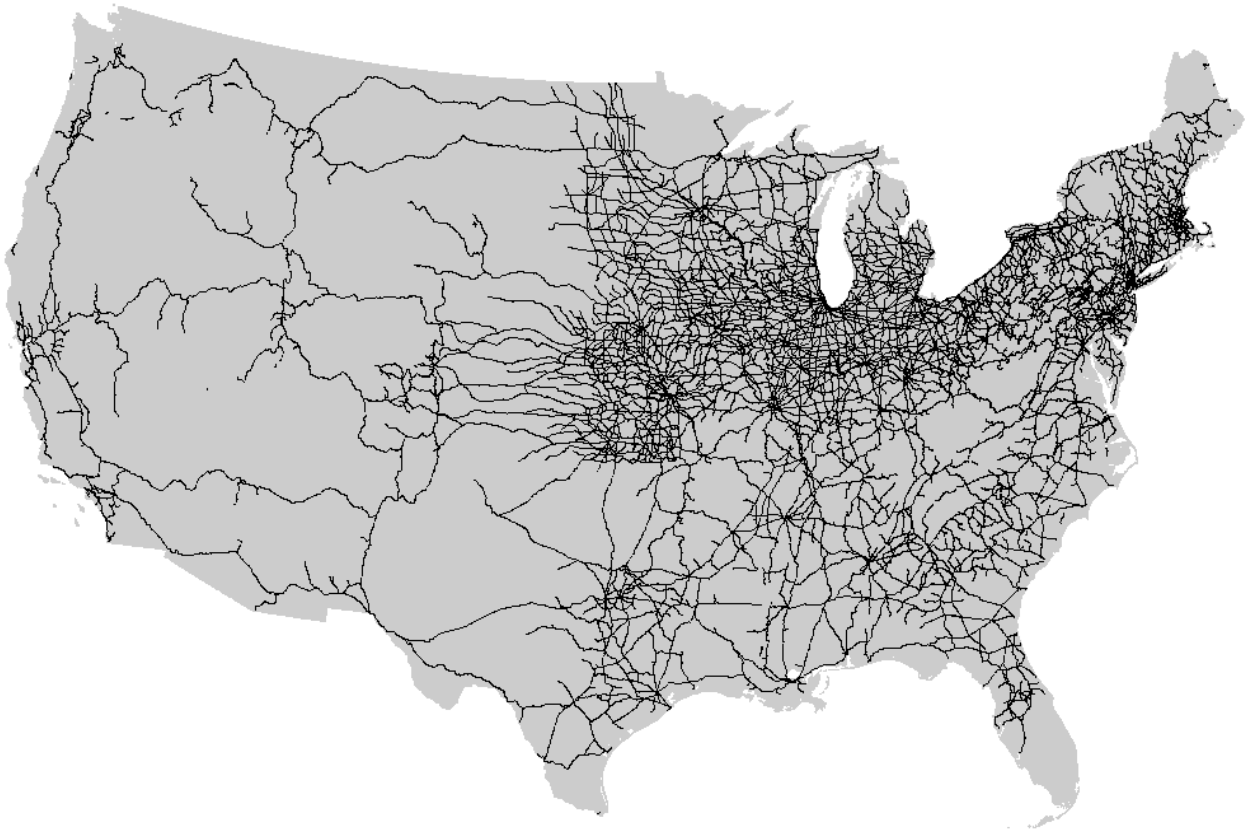
Notes: black lines displayed are railroads in operation by the date in question. Railroad data are obtained from Atack (2016).

Figure 3: US Railroads, 1860



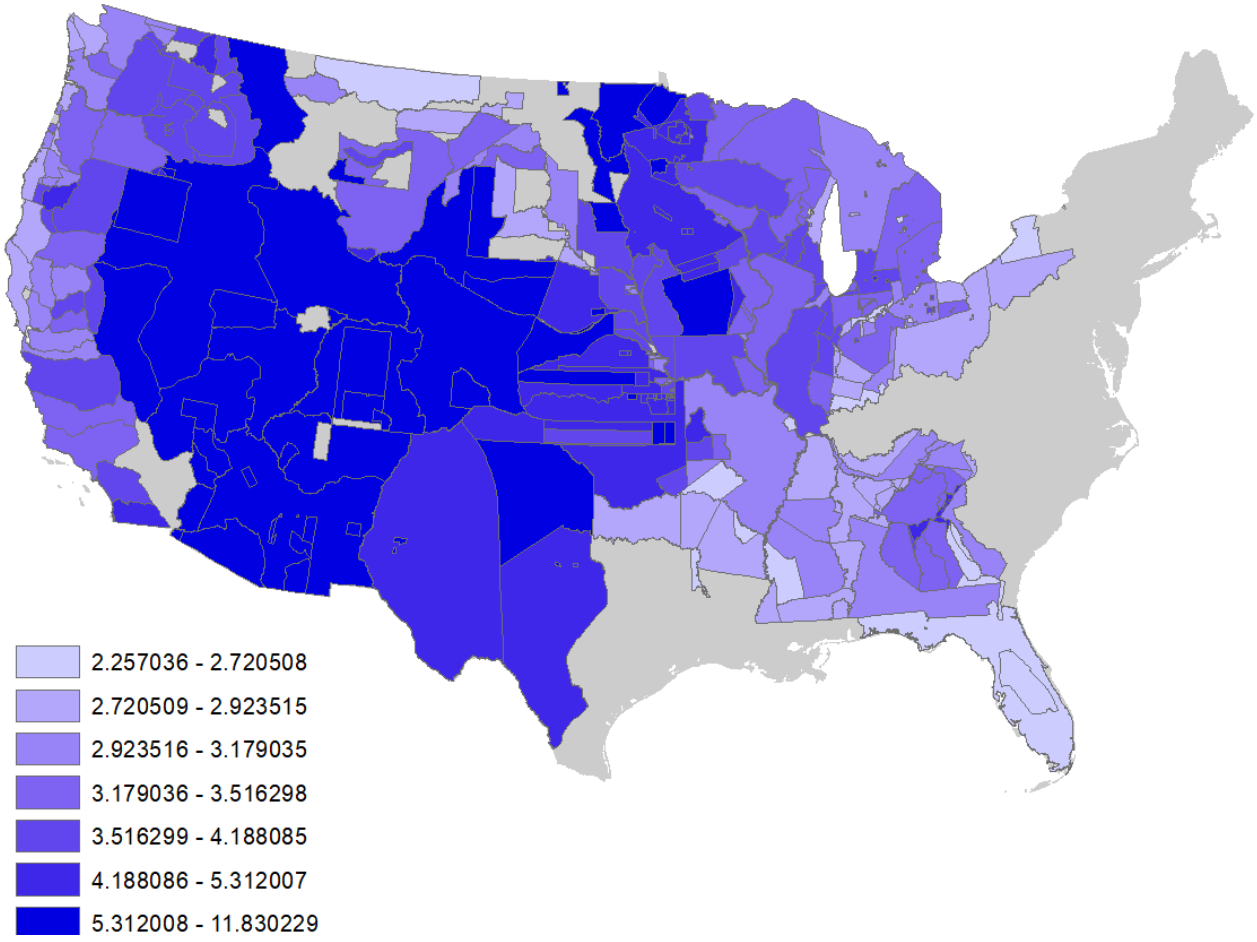
Notes: black lines displayed are railroads in operation by the date in question. Railroad data are obtained from Atack (2016).

Figure 4: US Railroads, 1890



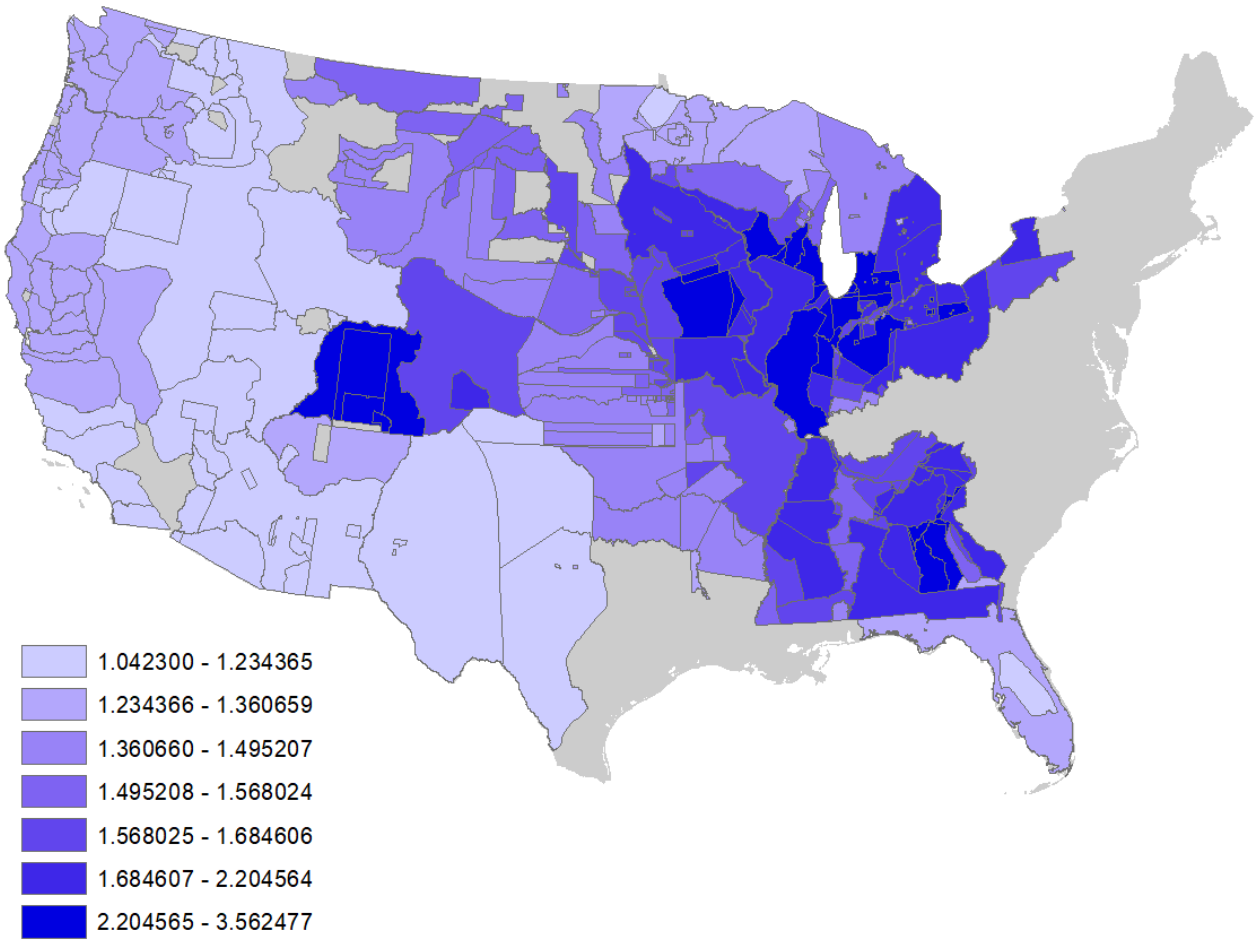
Notes: black lines displayed are railroads in operation by the date in question. Railroad data are obtained from Atack (2016).

Figure 5: Change in Market Access, 1830-1890



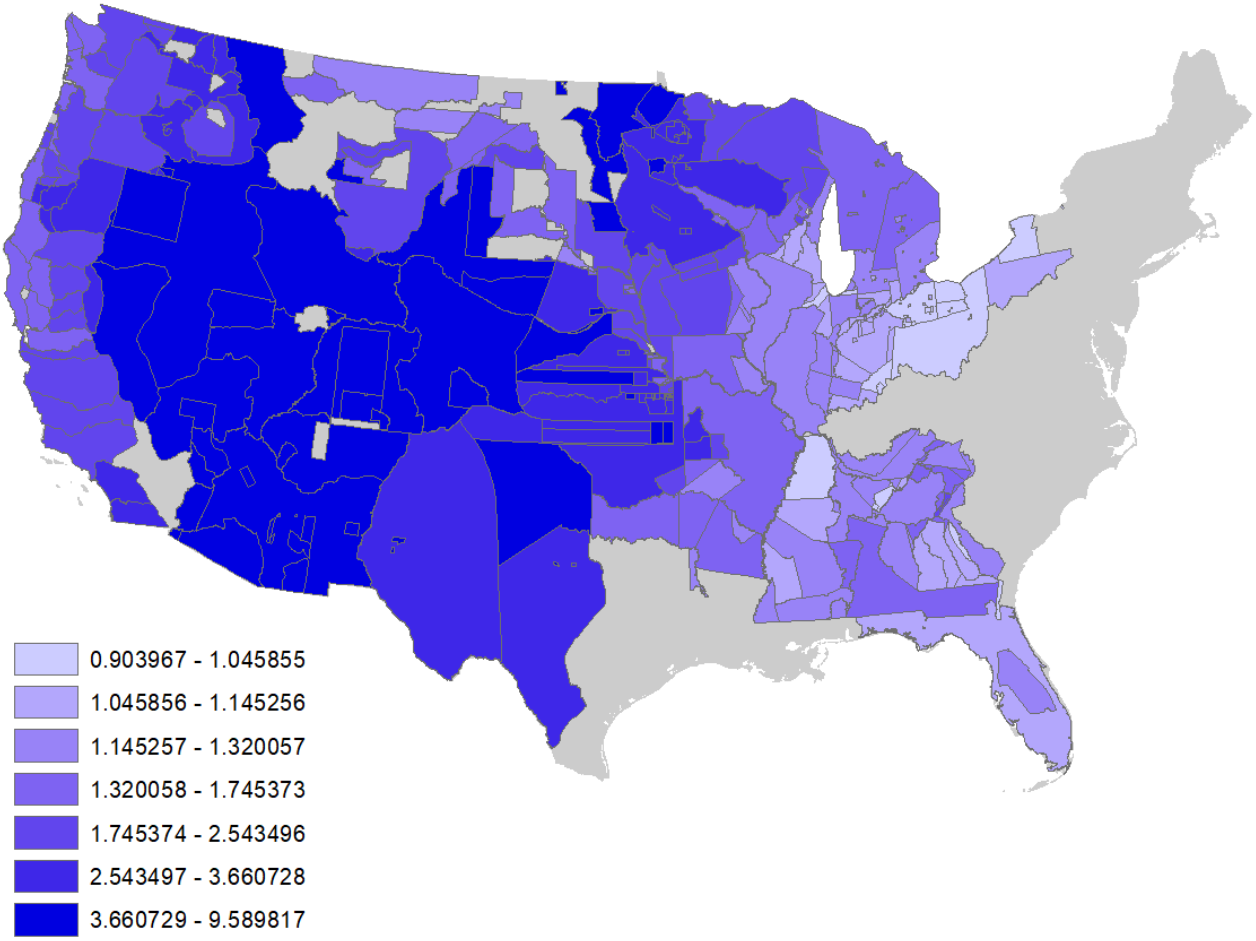
Notes: each colour represents one of 7 quartiles of market access exposure during this period, with darker colours representing larger (positive) changes in market access. The units in the legend are changes in the inverse hyperbolic sine of market access during the period in question, as calculated by the authors.

Figure 6: Change in Market Access, 1830-1860



Notes: each colour represents one of 7 quartiles of market access exposure during this period, with darker colours representing larger (positive) changes in market access. The units in the legend are changes in the inverse hyperbolic sine of market access during the period in question, as calculated by the authors.

Figure 7: Change in Market Access, 1860-1890



Notes: each colour represents one of 7 quartiles of market access exposure during this period, with darker colours representing larger (positive) changes in market access. The units in the legend are changes in the inverse hyperbolic sine of market access during the period in question, as calculated by the authors.

Table 1: Summary Statistics, Cession Panel

Variable	Mean	Std. Dev.	Min.	Max.	N
$\mathbb{1}$(signed)	0.559	0.497	0	1	1547
asinh(MA)	13.172	2.868	3.667	17.453	1547
RR access	0.228	0.42	0	1	1547
asinh(RR dist)	6.122	2.224	0.014	8.978	1547

Notes: summary statistics calculated on main sample of cession-year observations that transition from not being signed to being signed during the course of the sample time period.

Table 2: Summary Statistics, Cross-Sectional

Variable	Mean	Std. Dev.	Min.	Max.	N
asinh(MA)	13.396	1.427	8.356	16.257	221
RR access	0.088	0.285	0	1	226
asinh(RR dist)	6.399	1.099	1.553	7.834	226
asinh(wheat, cession)	7.84	1.452	0	8.622	226
asinh(cotton, cession)	2.183	2.654	0	6.314	226
asinh(maize, cession)	7.224	2.585	0	9.137	226
asinh(wheat, AIANNH)	8.231	0.702	0.061	8.613	209
asinh(cotton, AIANNH)	3.033	2.886	0	6.289	209
asinh(maize, AIANNH)	8.018	1.73	0	9.137	209
Modern-day outcomes from CICD:					
asinh(total population)	9.393	1.614	4.27	13.588	209
asinh(AIANa population)	8.127	1.24	4.026	12.663	209
asinh(population density, per square mile)	4.161	1.182	1.418	7.889	209
asinh(median age, total population)	4.273	0.175	3.757	4.858	209
asinh(median age, AIANa population)	4.131	0.194	3.709	4.905	209
share of total population with less than HS education	0.131	0.049	0.057	0.352	209
share of AIANa population with less than HS education	0.141	0.06	0	0.387	209
share of total population with a bachelor degree or higher	0.151	0.048	0.003	0.301	209
share of AIANa population with a bachelor degree or higher	0.109	0.041	0	0.242	209
employment-to-population ratio, total population	0.5	0.054	0.341	0.675	209
asinh(unemployed, total population)	0.088	0.043	0	0.218	209
LF participation rate, total population	0.548	0.054	0.369	0.687	209
asinh(HH median income, total population)	11.569	0.252	10.021	12.298	209
asinh(per capita income, total population)	10.832	0.226	10.176	11.913	209
share in poverty, total population	0.223	0.083	0.075	0.683	209
share living in owner-occupied housing, total population	0.634	0.151	0.119	0.936	209
share living in crowded housing, total population	0.055	0.041	0	0.226	209
share living in mobile homes, total population	0.127	0.075	0.007	0.443	209
employment-to-population ratio, AIANa population	0.452	0.071	0.267	0.689	209
asinh(unemployed, AIANa population)	0.131	0.06	0	0.278	209
LF participation rate, AIANa population	0.522	0.078	0.28	0.719	209
asinh(HH median income, AIANa population)	11.413	0.286	9.999	12.335	207
asinh(per capita income, AIANa population)	10.655	0.324	10.09	12.537	209
share in poverty, AIANa population	0.293	0.093	0.047	0.714	209
share living in owner-occupied housing, AIANa population	0.546	0.156	0	1	209
share living in crowded housing, AIANa population	0.079	0.055	0	0.253	209
share living in mobile homes, AIANa population	0.123	0.086	0	0.551	209

Notes: summary statistics presented only for estimation sample in cross-sectional regressions. AIANa refers to those in each modern-day AIANNH Area that identified as being solely of American Indian or Alaska Native origin. Total refers to all individuals living within a modern-day AIANNH Area.

Table 3: Market Access and Cessions

	(1)	(2)	(3)	(4)
Dep. Var.:	𝟙(signed)	𝟙(signed)	𝟙(signed)	𝟙(signed)
asinh(MA)	0.0558***	0.0572***	0.0551***	0.0570***
	(0.00636)	(0.00624)	(0.00646)	(0.00638)
RR access		-0.0808***		-0.176***
		(0.0275)		(0.0310)
asinh(RR dist)			-0.0225**	-0.0542***
			(0.0102)	(0.0117)
Observations	1,547	1,547	1,547	1,547
R-squared	0.712	0.714	0.713	0.719

Notes: *** p< 0.01, **p< 0.05, *p< 0.1. All specifications include cession fixed effects and time fixed effects. asinh() refers to the inverse hyperbolic sine transformation. Standard errors are clustered by cession. All columns estimated using OLS.

Table 4: Homeland-Level Main Results

	(1)	(2)	(3)
Dep. Var.:	1(cession overlapping)	asinh(area covered by cessions)	share of land covered by cessions
asinh(MA)	0.0592***	0.664***	0.0648***
	(0.00714)	(0.0797)	(0.00900)
Observations	1,302	1,302	1,302
R-squared	0.733	0.779	0.744

Notes: *** p< 0.01, **p< 0.05, *p< 0.1. All specifications include homeland fixed effects and year fixed effects. asinh() refers to the inverse hyperbolic sine transformation. Standard errors are clustered by homeland. All columns estimated using OLS.

Table 5: Market Access and Cessions, Disentangling Market Size and Transport Cost

Dep. Var.:	(1) 𝟙(signed)	(2) 𝟙(signed)	(3) 𝟙(signed)	(4) 𝟙(signed)	(5) 𝟙(signed)
asinh(MA)	0.0558***				
	(0.00636)				
asinh(MA), constant populations		0.0529***			
		(0.00633)			
asinh(MA), constant transport costs			0.135***		
			(0.0411)		
asinh(pop within 100km)				0.0370***	
				(0.00354)	
asinh(pop within 200km)					0.0270***
					(0.00298)
Observations	1,547	1,547	1,547	1,547	1,547
R-squared	0.712	0.711	0.703	0.721	0.711

Notes: *** p< 0.01, **p< 0.05, *p< 0.1. All specifications include cession fixed effects and time fixed effects. asinh() refers to the inverse hyperbolic sine transformation. Standard errors are clustered by cession. All columns estimated using OLS.

Table 6: Market Access and Conflicts

Dep. Var.:	(1) asinh(battles)	(2) asinh(battles)	(3) asinh(battles)	(4) 1(battle occurrence)	(5) 1(battle occurrence)	(6) 1(battle occurrence)
asinh(MA)	-0.0529	-0.0707	-0.0530	-0.00278	-0.00548	-0.000696
	(0.0781)	(0.0790)	(0.0820)	(0.0440)	(0.0454)	(0.0451)
Source	small	medium	large	small	medium	large
Observations	98	98	98	98	98	98
R-squared	0.465	0.500	0.512	0.445	0.479	0.492

Notes: *** p< 0.01, **p< 0.05, *p< 0.1. All specifications include state fixed effects and time fixed effects. asinh() refers to the inverse hyperbolic sine transformation. Standard errors are clustered by state. All columns estimated using OLS. The "small" source refers to battles in Adjutant-General's Office (1979). The "medium" source refers to battles in either Adjutant-General's Office (1979) or Webb (1976). The "large" source refers to battles in Peters (1966).

Table 7: Market Access and Cession Characteristics at Signing

Dep. Var.:	(1) asinh(wheat, cession)	(2) asinh(maize, cession)	(3) asinh(cotton, cession)	(4) asinh(area, cession)	(5) asinh(dist)
asinh(MA)	0.156***	0.562***	0.0378	-0.396***	-0.0636
	(0.0455)	(0.104)	(0.139)	(0.123)	(0.0617)
Observations	221	221	221	221	204
R-squared	0.297	0.379	0.035	0.164	0.133

Notes: *** p< 0.01, **p< 0.05, *p< 0.1. All specifications include decade-of-signing fixed effects. asinh() refers to the inverse hyperbolic sine transformation. Standard errors are clustered by cession. All columns estimated using OLS.

Table 8: Family Composition

Dep. Var.:	(1) prop. child. survive	(2) children ever born	(3) 1(female)	(4) prop. white blood	(5) 1(full blood status)
asinh(MA)	0.00370	0.0499***	-0.00247	2.964***	-0.0467***
	(0.00403)	(0.0178)	(0.00236)	(0.535)	(0.0105)
Observations	10,024	29,642	40,151	33,523	40,151
R-squared	0.125	0.601	0.001	0.266	0.317

Notes: *** p< 0.01, **p< 0.05, *p< 0.1. asinh() refers to the inverse hyperbolic sine transformation. All columns control for region fixed effects, Census year fixed effects, decade of first cession signing fixed effects, age, age squared, bison loss, and a large set of cultural and geographical controls; see the Appendix for more details. Regressions in columns 3 and 4 estimated only on those born after cession signing. All columns estimated using OLS. Standard errors clustered by homeland. Observations weighted using person-level sampling weights from IPUMS.

Table 9: Occupation Scores and Industrial Composition

Dep. Var.:	(1) asinh(occ. score, std.)	(2) 1(has occ)	(3) asinh(occ. score, std.)	(4) 1(has ind)	(5) 1(agricul-ture)	(6) 1(capital intense ind)
asinh(MA)	0.0253	0.0188	0.00990	0.0214**	0.0106	0.00172
	(0.0220)	(0.0177)	(0.00718)	(0.00921)	(0.0102)	(0.00474)
Observations	7,312	7,312	5,889	18,254	18,254	18,254
R-squared	0.152	0.184	0.108	0.102	0.095	0.101

Notes: *** p< 0.01, **p< 0.05, *p< 0.1. asinh() refers to the inverse hyperbolic sine transformation. All columns control for region fixed effects, Census year fixed effects, decade of first cession signing fixed effects, age, age squared, bison loss, and a large set of cultural and geographical controls; see the Appendix for more details. All columns estimated using OLS. Standard errors clustered by homeland. Observations weighted using person-level sampling weights from IPUMS.

Table 10: 1915-1920 Income Composition Results

Dep. Var.:	(1) asinh(total income per capita)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		percent of income from given source, where source is:							
		land sales	leases	wages	livestock sales	rations	crops	traditional industries	treaties and interest
asinh(MA)	-0.0109	-1.514**	1.667*	0.325	-0.443	-0.853***	3.088*	0.217	-3.164**
	(0.0557)	(0.601)	(0.877)	(1.206)	(0.735)	(0.226)	(1.713)	(0.710)	(1.274)
Observations	59	59	59	59	59	59	59	59	59
R-squared	0.060	0.181	0.198	0.136	0.121	0.318	0.196	0.037	0.260

Notes: *** p< 0.01, **p< 0.05, *p< 0.1. asinh() refers to the inverse hyperbolic sine transformation. All regressions control for decade of first cession signing fixed effects. All columns estimated using OLS. Standard errors clustered by homeland.

Table 11: Heights

Dep. Var.:	(1) height	(2) height	(3) height	(4) height	(5) height	(6) height
asinh(MA) x born post cession	-0.649	-0.786**	-0.679**	-1.516***	-0.547***	-0.0937
	(0.402)	(0.350)	(0.340)	(0.419)	(0.154)	(0.594)
born post cession	11.34**	12.53***	11.51***	21.26***	7.283***	10.89
	(4.727)	(3.829)	(3.357)	(4.958)	(1.820)	(6.624)
sample:	full	full	male	female	<= 10 years	> 10 years
Observations	7,123	7,123	4,496	2,627	1,867	6,234
R-squared	0.786	0.788	0.796	0.750	0.603	0.807

Notes: *** p< 0.01, **p< 0.05, *p< 0.1. asinh() refers to the inverse hyperbolic sine transformation. All columns control for year fixed effects, homeland fixed effects, age, age squared, female dummy, born post-bison loss dummy, bison treatment dummy interacted with post-bison loss dummy, and for columns 2-6 a large set of geographic and cultural control variables; see the Appendix for more details on these controls. Sample in column 5 keeps all non-treated cohorts but keeps only treated cohorts born within 10 years of cession. Sample in column 6 keeps all non-treated cohorts but keeps only treated cohorts born after 10 years of cession. All columns estimated using OLS. Standard errors clustered by homeland.

Table 12: Long-Run Income Results

Dep. Var.:	(1) asinh(real per capita income)	(2) asinh(real per capita income)	(3) asinh(real per capita income)	(4) asinh(real per capita income)	(5) asinh(real per capita income)	(6) asinh(real per capita income)	(7) asinh(real per capita income)	(8) asinh(real per capita income)	(9) asinh(real per capita income)
Panel A: Main Results									
asinh(MA)	0.0435***	0.0267	0.0419**	0.0153	0.0175	0.0242	0.0355**	0.0344*	0.0157
	(0.0135)	(0.0176)	(0.0176)	(0.0182)	(0.0150)	(0.0183)	(0.0178)	(0.0202)	(0.0162)
Observations	676	676	676	676	676	676	676	676	676
R-squared	0.609	0.621	0.656	0.664	0.637	0.623	0.632	0.632	0.651
Panel B: Results, by Year									
asinh(MA) x 1945	0.0439	0.0243	0.0564	-0.00686	0.0160	0.0223	0.0457	0.0352	0.0108
	(0.0316)	(0.0333)	(0.0361)	(0.0328)	(0.0320)	(0.0348)	(0.0341)	(0.0352)	(0.0307)
asinh(MA) x 1969	0.0568***	0.0408*	0.0696***	0.0216	0.0357*	0.0380*	0.0542**	0.0494**	0.0211
	(0.0185)	(0.0210)	(0.0239)	(0.0213)	(0.0196)	(0.0207)	(0.0221)	(0.0243)	(0.0191)
asinh(MA) x 1979	0.0523***	0.0352*	0.0586***	0.0214	0.0275	0.0324*	0.0463**	0.0437**	0.0205
	(0.0151)	(0.0183)	(0.0203)	(0.0195)	(0.0174)	(0.0184)	(0.0187)	(0.0213)	(0.0165)
asinh(MA) x 1989	0.0302*	0.0130	0.0305	0.00414	0.00504	0.0106	0.0214	0.0218	0.000979
	(0.0158)	(0.0191)	(0.0186)	(0.0197)	(0.0170)	(0.0189)	(0.0189)	(0.0221)	(0.0178)
asinh(MA) x 1999	0.0603***	0.0444**	0.0548***	0.0346	0.0322*	0.0418*	0.0512**	0.0508**	0.0343*
	(0.0178)	(0.0213)	(0.0201)	(0.0213)	(0.0185)	(0.0217)	(0.0215)	(0.0226)	(0.0197)
asinh(MA) x 2010	0.0296*	0.0129	0.0162	0.00957	0.00216	0.00998	0.0173	0.0211	0.00662
	(0.0155)	(0.0193)	(0.0172)	(0.0208)	(0.0164)	(0.0195)	(0.0194)	(0.0211)	(0.0179)
asinh(MA) x 2019	0.0311	0.0151	0.0142	0.0162	0.00382	0.0129	0.0168	0.0208	0.0117
	(0.0204)	(0.0249)	(0.0202)	(0.0260)	(0.0221)	(0.0257)	(0.0255)	(0.0255)	(0.0240)
Observations	676	676	676	676	676	676	676	676	676
R-squared	0.612	0.624	0.662	0.666	0.640	0.625	0.635	0.634	0.652
control variables:									
bison loss		x	x	x	x	x	x	x	x
geographic			x						
culture				x					
disruption					x				
Dust Bowl						x			
prime land							x		
General Allotment Act								x	
governance									x

Notes: *** p< 0.01, **p< 0.05, *p< 0.1. asinh() refers to the inverse hyperbolic sine transformation. All columns, in addition to the control variable set specified, control for decade of first cession signing fixed effects, and year fixed effects. All columns estimated using OLS. Standard errors clustered by homeland.

Table 13: Differences, AIANNH Areas and Cessions

Dep. Var.:	(1) asinh(wheat, AIANNH)	(2) asinh(maize, AIANNH)	(3) asinh(cotton, AIANNH)	(4) asinh(area, AIANNH)	(5) asinh(wheat, AIANNH)	(6) asinh(maize, AIANNH)	(7) asinh(cotton, AIANNH)	(8) asinh(area, AIANNH)
asinh(MA)	0.122***	0.300***	-0.0878	-0.119	0.113***	0.174	-0.125	-0.103
	(0.0404)	(0.110)	(0.152)	(0.118)	(0.0371)	(0.113)	(0.108)	(0.120)
asinh(wheat)					0.0549			
					(0.0464)			
asinh(maize)						0.237**		
						(0.0916)		
asinh(cotton)							0.696***	
							(0.0516)	
asinh(area)								0.0397
								(0.0565)
Observations	204	204	204	204	204	204	204	204
R-squared	0.180	0.160	0.098	0.100	0.188	0.237	0.484	0.103

Notes: *** p< 0.01, **p< 0.05, *p< 0.1. All specifications include decade-of-signing fixed effects. asinh() refers to the inverse hyperbolic sine transformation. Standard errors are clustered by cession. All columns estimated using OLS.

Table 14: Modern Outcomes: Population

Dep. Var.:	(1) asinh(population, total)	(2) asinh(population, AIANa)	(3) asinh(population density, square miles)	(4) asinh(median age, total)	(5) asinh(median age, AIANa)	(6) share of total population with less than HS education	(7) share of AIANa population with less than HS education	(8) share of total population with bachelor degree or higher	(9) share of AIANa population with bachelor degree or higher
asinh(MA)	-0.123	-0.0820	0.0395	-0.0117	-0.0130	-0.00551**	-0.00375	0.00323	0.00105
	(0.100)	(0.0839)	(0.0509)	(0.0100)	(0.0114)	(0.00247)	(0.00317)	(0.00302)	(0.00214)
Observations	204	204	204	204	204	204	204	204	204
R-squared	0.077	0.074	0.144	0.034	0.076	0.107	0.088	0.037	0.064

Notes: *** p< 0.01, **p< 0.05, *p< 0.1. All specifications include decade-of-signing fixed effects. asinh() refers to the inverse hyperbolic sine transformation. AIANa refers to those who identified as "American Indian or Alaska Native" alone, without any additional racial composition reported. Total refers to the entire population residing within a given geography. Standard errors are clustered by cession. All columns estimated using OLS.

Table 15: Modern Outcomes: Labor Force and Income

Dep. Var.:	(1) employment to population ratio	(2) asinh(unemployed)	(3) LF participation rate	(4) asinh(median HH income)	(5) asinh(per capita income)	(6) share in poverty	(7) share living in owner-occupied housing	(8) share living in crowded housing	(9) share living in mobile homes
Panel A: AIANa									
asinh(MA)	0.00553	0.00511*	0.00955**	0.0307*	0.0124	-0.00952*	0.00216	0.00591	-0.00518
	(0.00385)	(0.00302)	(0.00466)	(0.0168)	(0.0150)	(0.00505)	(0.00803)	(0.00374)	(0.00423)
Observations	204	204	204	202	204	204	204	204	204
R-squared	0.067	0.156	0.059	0.098	0.061	0.219	0.037	0.069	0.014
Panel B: Total									
asinh(MA)	0.00624*	0.00254	0.00819**	0.0379**	0.0174	-0.0108**	0.00703	0.00193	-0.00539
	(0.00320)	(0.00249)	(0.00344)	(0.0159)	(0.0106)	(0.00518)	(0.00756)	(0.00247)	(0.00377)
Observations	204	204	204	204	204	204	204	204	204
R-squared	0.044	0.062	0.062	0.090	0.060	0.114	0.025	0.028	0.027

Notes: *** p< 0.01, **p< 0.05, *p< 0.1. All specifications include decade-of-signing fixed effects. asinh() refers to the inverse hyperbolic sine transformation. AIANa refers to those who identified as "American Indian or Alaska Native" alone, without any additional racial composition reported. Total refers to the entire population residing within a given geography. Standard errors are clustered by cession. All columns estimated using OLS.

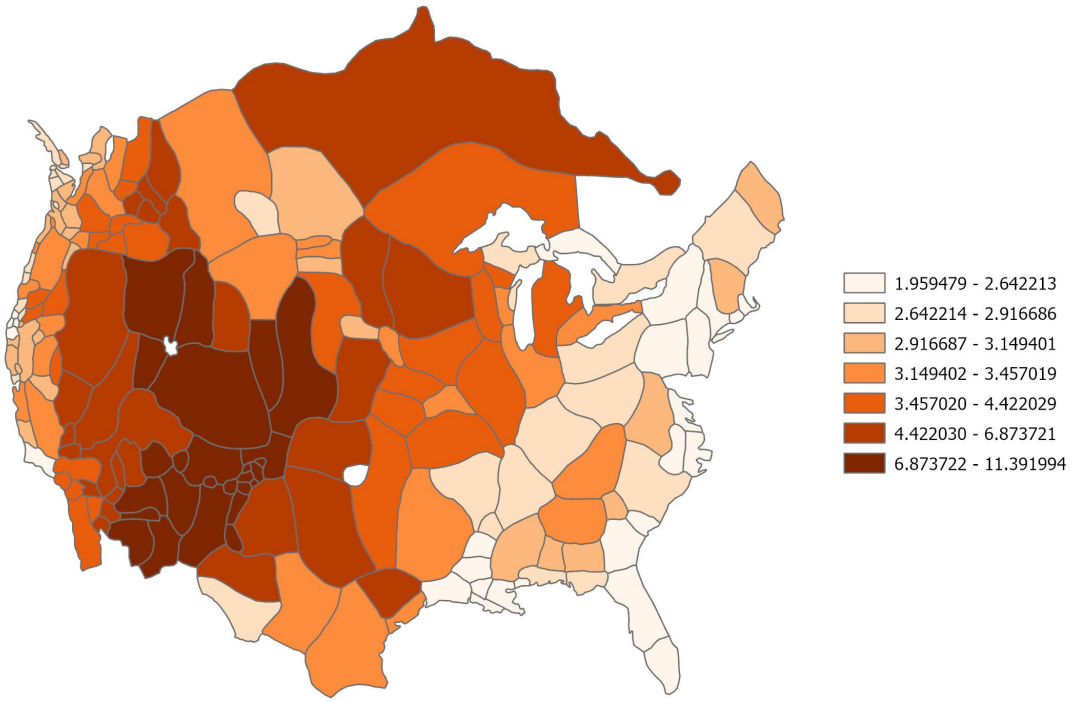
Table 16: Robustness Checks, Cession Level

Dep. Var.:	(1) $\mathbb{1}$(signed)	(2) $\mathbb{1}$(signed)	(3) $\mathbb{1}$(signed)	(4) $\mathbb{1}$(signed)	(5) $\mathbb{1}$(signed)	(6) $\mathbb{1}$(signed)	(7) $\mathbb{1}$(signed)
asinh(MA)	0.106***	0.0558***				0.0435***	0.0534***
	(0.00667)	(0.00947)				(0.0134)	(0.00627)
ln(MA)			0.0558***				
			(0.00636)				
asinh(MA_{t-1})				0.0375***			
				(0.00798)			
asinh(MA_{t+1})					0.0165*		
					(0.00878)		
Observations	2,933	1,547	1,547	1,326	1,326	357	1,547
R-squared	0.659	0.712	0.712	0.666	0.693	0.722	0.040

Notes: *** p< 0.01, **p< 0.05, *p< 0.1. All specifications include cession fixed effects and time fixed effects. asinh() refers to the inverse hyperbolic sine transformation. Standard errors are clustered by cession. All columns estimated using OLS, except the last column. Column 1 uses a fuller sample of all cessions. Column 2 clusters by tribe as denoted in the original Forest Service data. Column 3 uses natural logs to replace all instances of inverse hyperbolic sine transformations. Column 4 includes a lagged market access measure. Column 5 includes a lead market access measure. Column 6 uses only cessions that do not receive direct railroad access in our sample period. Column 7 uses an instrumental variable constructed using other counties at least 500km away and using fixed 1830 populations.

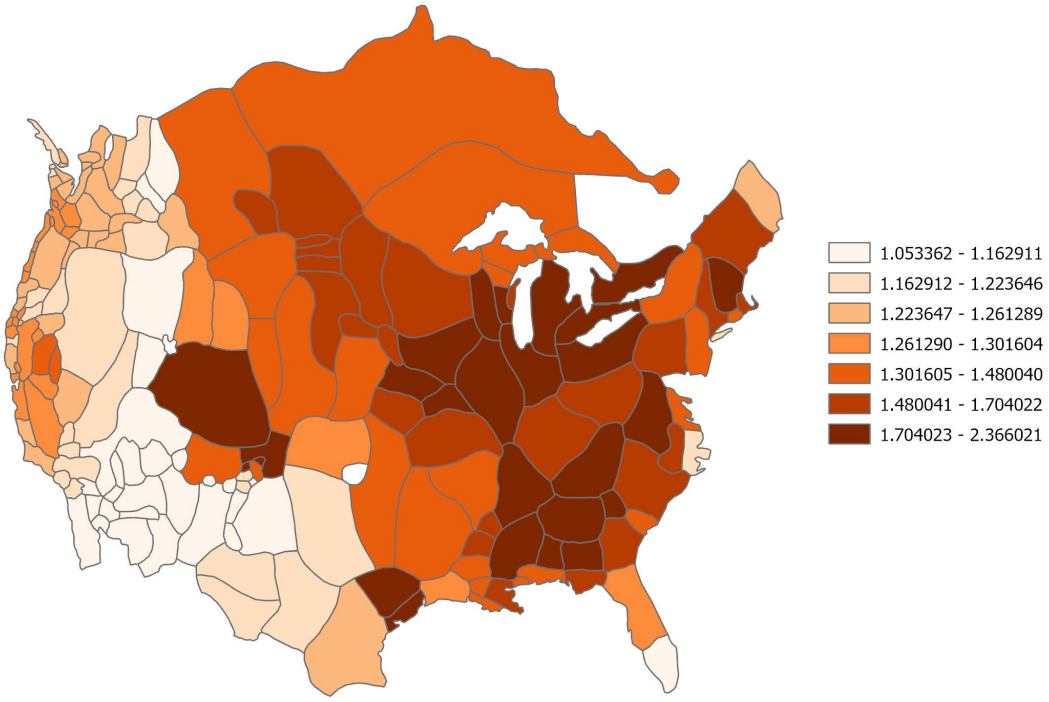
A Appendix: Additional Figures and Tables

Figure A1: Change in Homelands' Market Access, 1830-1890



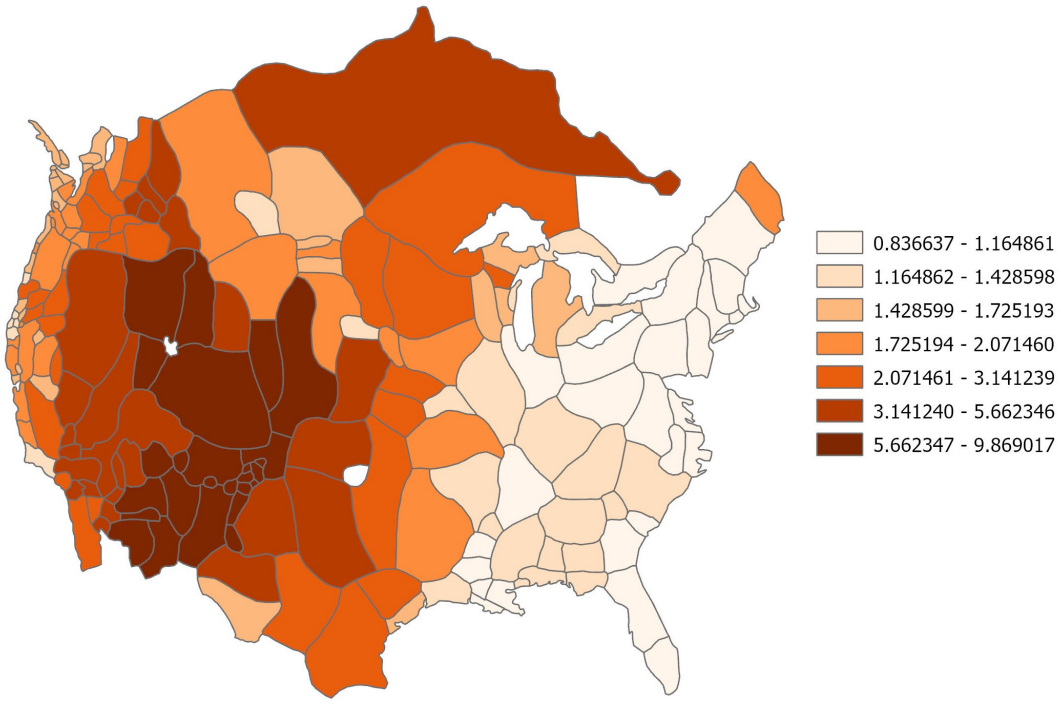
Notes: each colour represents one of 7 quartiles of market access exposure during this period, with darker colours representing larger (positive) changes in market access. The units in the legend are changes in the inverse hyperbolic sine of market access during the period in question, as calculated by the authors.

Figure A2: Change in Homelands' Market Access, 1830-1860



Notes: each colour represents one of 7 quartiles of market access exposure during this period, with darker colours representing larger (positive) changes in market access. The units in the legend are changes in the inverse hyperbolic sine of market access during the period in question, as calculated by the authors.

Figure A3: Change in Homelands' Market Access, 1860-1890



Notes: each colour represents one of 7 quartiles of market access exposure during this period, with darker colours representing larger (positive) changes in market access. The units in the legend are changes in the inverse hyperbolic sine of market access during the period in question, as calculated by the authors.

Table A1: Differences, AIANNH Areas, Dummies

Dep. Var.:	(1) 𝟙(wheat, cession > wheat, AIANNH)	(2) 𝟙(maize, cession > maize, AIANNH)	(3) 𝟙(cotton, cession > cotton, AIANNH)
asinh(MA)	0.0430*	0.0414	0.0281*
	(0.0248)	(0.0251)	(0.0153)
asinh(wheat)	0.0340**		
	(0.0133)		
asinh(maize)		0.0259**	
		(0.0103)	
asinh(cotton)			0.0550***
			(0.00932)
Observations	221	221	221
R-squared	0.252	0.136	0.233

Notes: *** p< 0.01, **p< 0.05, *p< 0.1. All specifications include decade-of-signing fixed effects. asinh() refers to the inverse hyperbolic sine transformation. Standard errors are clustered by cession. All columns estimated using OLS.

Table A2: Homeland Robustness Checks

Dep. Var.:	(1) 1(cession overlapping)	(2) 1(cession overlapping)	(3) 1(cession overlapping)	(4) 1(cession overlapping)	(5) 1(cession overlapping)	(6) 1(cession overlapping)
asinh(MA)	0.0592***				0.0224***	0.0282***
	(0.00714)				(0.00771)	(0.00731)
ln(MA)		0.0323***				
		(0.00723)				
asinh(MA), lag			0.0319***			
			(0.00707)			
asinh(MA), lead				0.0195*		
				(0.0115)		
Observations	1,302	1,043	894	894	182	1,043
R-squared	0.733	0.749	0.726	0.753	0.835	0.022

Notes: *** p< 0.01, **p< 0.05, *p< 0.1. All specifications include homeland fixed effects and year fixed effects. asinh() refers to the inverse hyperbolic sine transformation. ln() refers to the natural log transformation. Column 1 uses a larger sample of all homelands. Column 2 uses the log transformation. Column 3 uses a lag specification. Column 4 uses a lead specification. Column 5 uses only homelands that never receive railroads. Column 6 uses an instrumental variable approach. Standard errors are clustered by homeland. All columns estimated using OLS.

Table A3: Summary Statistics, Conflicts

Variable	Mean	Std. Dev.	Min.	Max.
asinh(MA)	12.011	3.026	5.154	16.716
asinh(battles), source = "small"	0.534	1.023	0	4.026
asinh(battles), source = "medium"	0.579	1.073	0	4.026
asinh(battles), source = "large"	0.596	1.083	0	4.061
1(battles), source = "small"	0.265	0.444	0	1
1(battles), source = "medium"	0.286	0.454	0	1
1(battles), source = "large"	0.296	0.459	0	1

Table A4: Summary Statistics, Feir et al. (2023b) Datasets

Variable	Mean	Std. Dev.	Min.	Max.	N
Panel A: Heights					
height (cm)	153.937	19.298	63.1	195.6	7128
after cession x asinh(MA)	10.285	4.66	0	14.834	7128
after cession	0.854	0.353	0	1	7128
asinh(MA)	11.923	2.126	6.584	14.834	7128
Panel B: IPUMS Data					
asinh(MA)	12.125	2.391	6.584	14.834	81518
prop. of children born who survived	0.832	0.224	0.091	1	11462
children ever born	1.783	2.905	0	21	29642
female	0.491	0.5	0	1	41367
proportion of white blood	20.493	32.933	0	100	40665
full blood status	0.669	0.471	0	1	40665
asinh(occ. score, standardized)	0.303	0.570	-0.645	2.514	18254
has occupation	0.815	0.388	0	1	18254
has industry	0.778	0.416	0	1	18254
agriculture	0.577	0.494	0	1	18254
capital intensive industry	0.077	0.266	0	1	18254
Panel C: 1915-1920 Indian Affairs Reports					
asinh(MA)	11.389	2.318	6.584	14.805	59
asinh(total income per capita)	9.082	0.793	6.825	10.493	59
% income from land sales	5.095	9.707	0	38.173	59
% income from leases	9.372	14.608	0	54.586	59
% income from wages	16.32	16.501	0.386	77.674	59
% income from livestock sales	9.011	10.235	0	48.296	59
% income from rations	1.965	2.915	0	11.863	59
% income from crops	36.955	25.588	0	145.335	59
% income from traditional industries	7.017	13.273	0	72.313	59
% income from treaties and interest	11.06	13.133	0	50.156	59
Panel D: Leonard et al Income Data					
asinh(MA)	11.465	2.261	6.584	14.805	676
asinh(per capita income)	9.893	0.584	7.31	11.623	676

B Appendix: Additional Data Construction Details

B.1 Instrumental Variable

The instrument is adapted from the empirical methods in Chan (2024). Specifically, the instrumental variable, at the county level, is of the form:

$$MAIV_{ct} = \sum_{d \in C_{-c, \geq 500km}} \tau_{cdt}^{-\theta} * N_{d,1830}$$

The instrument is similar to the market access measure but with two key differences. First, only counties at least 500 kilometers away are used in the calculation of the IV. Second, the populations of the other counties are fixed at their 1830 levels. The instrument is then converted to the cession level using the same procedure as outlined in the main text for the market access measure.

B.2 Converting Counties to 1890 Counties

In order to convert all county-level data to 1890 county boundary definitions, we use the crosswalks developed by Ferrara et al (2022). These crosswalks improve upon area-based crosswalks by also incorporating population into the weighting procedure. The crosswalks therefore divide each county into different pieces that fit into different 1890 counties, with the weighting of each piece based on the population that lives in that piece.

B.3 Converting 1890 County Market Access to Cessions

The market access measures must be constructed at the level of 1890 counties, since the transport cost database from Donaldson and Hornbeck (2016) is available only for that geography. To construct cession level measures, we therefore must develop a custom crosswalk. We do so by taking the boundaries of each cession and breaking each apart into separate pieces for each 1890 county that intersects with it; this is done using the Intersect tool in ArcGIS. We drop all pieces that are less than 1% of a cession's total area, to minimize issues arising from imperfections in the shapefiles coupled with extreme values of market access driving the calculation of market access for any given cession. For each remaining piece, we then take that 1890 county's market access measure and assign it a weight, with the weight being based on the proportion of that cession's total area that this piece comprises. The weighted values of market access are then summed

to form a weighted mean value of market access for that cession. We also employ this methodology to construct the population proximity measures used in our analysis (i.e. for the population within 100 and 200 kilometers of a cession).

B.4 Control Variables, Short and Long Run Outcomes

In our specifications using the IPUMS US Census data from 1900, 1910, and 1930, we include a set of control variables for geographic and cultural factors. Geographic control variables we include are: latitude, log ruggedness, and log area. For cultural controls, we include: the logged population in 1500 (constructed from HYDE), a dummy for if the peoples were nomadic, the historic centralization of the group, wealth inequality, and three dummies flagging if each of the latter three variables are missing. All cultural controls, other than the logged population, are from the Ethnographic Atlas. We also include the same control variables but interacted with birth year, as in Feir et al. (2023b).

In the Leonard et al. (2020) long-run income regressions, we include numerous sets of control variables one-by-one to help tease out potential mechanisms. Our geographic control variable set includes: latitude, log ruggedness, and log area. Our culture controls include: the logged population in 1500 (constructed from HYDE), the logged population in 1600 (constructed from HYDE), a dummy for if the peoples were nomadic, wealth inequality, and two dummies for each of the latter two variables are missing. The disruption controls are: battles fought within the territory pre-bison slaughter and a measure of whether the group was forced to co-exist with others (from Dippel (2014)). Our dust bowl controls include: the fraction of a territory subject to medium erosion due to the Dust Bowl and the fraction of a territory subject to high erosion due to the Dust Bowl (both obtained from Hornbeck (2012)). Our prime land controls include: percent of a reservation on prime land and its square (from Leonard et al. (2020)), as well as flag dummies for whether these variables were missing. Our controls for being affected by the General Allotment Act are: the degree of land ownership fractionation and whether a reservation was allotted, as well as flag dummies for whether these variables were missing. Our governance controls include: whether a nation had a constitution under the Indian Reorganization Act and a measure of self-governance taken from Gregg (2018), as well as flag dummies for whether these variables are missing. For all these controls, we also include interactions with linear trends.

In our regressions using the Boaz height data, we include control variable sets for cultural and geographic factors. Our cultural set includes: the logged population in 1500 (constructed from HYDE), a dummy for if the peoples were nomadic, the historic centralization of the group, wealth inequality, and three dummies flagging if each of the latter three variables are missing. Our geographic controls include: latitude, the log of ruggedness, and the log of the area. We also include, for all these controls, versions of each variable that are interacted with birth year.

B.5 Matching Homelands to Feir et al (2023) Data

We make use of the replication files publicly available for Feir et al. (2023b) in many of the short and long run outcomes results in this paper. The homelands in our paper are broadly consistent with the homeland/tribe definitions used in Feir et al. (2023b). In a minority of cases in each dataset from Feir et al. (2023b), the names of homelands diverge, due to a variety of reasons such as aggregation issues and mismatches in tribe names between the specific dataset and the homeland-based names. In order to maintain as much consistency as possible in our analysis, we opt to keep a consistent definition of homelands across all of our analyses and naming convention of homelands as much as possible and match only those homelands/tribes in the Feir et al. (2023b) datasets for which we find exact name matches. In all cases, this still leads to the majority of observations being matched in the outcome datasets.

B.6 Linking Cessions to 2020 AIANNH Areas

To link historical cessions to modern-day 2020 AIANNH Areas, we make use of the cession data's identifier for modern-day tribes (which are recognized by the federal government) involved in each historical cession. This allows us to link tribes involved in each cession to the tribes located in each AIANNH Area. One complication is that multiple tribes could be associated with each cession, and multiple tribes could also be associated with each Area.

To address the issue of multiple tribes potentially sharing a single AIANNH Area (which occurs principally with off-reservation trust lands), we simply assume that each tribe splits a given area equally with the other tribes in that same area. That area's total landmass is therefore apportioned in a 1/N proportion to each tribe, where N is the total number of tribes sharing that area.

To calculate the average distance from a given historical cession-tribe combination, we then take a weighted average of the distance from that cession to each AIANNH Area that the tribe is associated with, with the weights being given by the share of landmass that the Area contributes to the total landmass associated with that tribe over all Areas. To calculate a mean value of distance for a given historical cession, we then take a simple mean of the distance across all tribes associated with that cession.

The final measure of distance therefore measures, for a given cession, the mean distance from that cession to all the AIANNH Areas associated with the tribes in that cession. To construct modern-day measures of AIANNH Area characteristics, we implement the same procedure but replace distance with that characteristic, similarly taking means and weighted means as specified above.