The Effect of Land Allotment on Native American Households During the Assimilation Era*

Christian Dippel† Dustin Frye‡

October 18, 2019

Abstract

In the early twentieth century, the U.S. government broke up millions of acres of communally owned reservation lands and allotted them to individual Native American households. Households initially received land allotments with limited property rights ('in trust'), and were incentivized to prove themselves "competent" in order to obtain full legal title ('fee simple') after a set period. Indian allotment thus had elements of a conditional cash transfer program aimed at assimilation. The policy was ended suddenly in 1934, locking in-trust land into its status in perpetuity. We link land allotment information to the universe of Native American households in the 1940 U.S. Census. We exploit quasi-random variation in being allotted as well as in securing the allotment in fee simple. Obtaining an allotment significantly increased the likelihood of being a farmer, but also educational attainment, which we interpret as a way of signalling "competency". Obtaining the land in fee simple was associated with increases in home values, a reduction in farming, increases in labor force participation, occupational upgrading, and additional educational attainment of children. The fee-simple effects were more pronounced within tribes whose ancestral tribal norms emphasized private over communal property, indicating a cultural determinant in how the extra wealth transfer was utilized.

Keywords: Indigenous Economic Development, Culture, Assimilation, Property Rights, Conditional Cash Transfers, Record Linkage
JEL Codes: N10, O1, Z1

*We thank Terry Anderson, Leonard Carlson, Marie Duggan, Donna Feir, Miriam Jorgensen, Bryan Leonard, Mindy Miller, Steve Nafziger Dominick Parker, Matthew Snipp, and seminar participants at Hoover Institution's Research Workshop on Renewing Indigenous Economies, the WEAI conference in San Francisco, and the EHA conference in Atlanta for helpful conversations. We thank David Cruse, Michelle Garabetian, Adam Krueger, Jason Kwan, Zach Lewis, Nika McKechnie, Josh Mimura, Lilian Stanwick, Diana (Zhoutong) Wu, Nina Zacharia and especially Chengcheng Zhang for excellent research assistance. Dippel gratefully acknowledges support from a UCLA Senate COR grant.

†University of California, Los Angeles, CCPR, and NBER.
‡Vassar College

1 Introduction

Toward the end of the 19th century, with the end of the Indian Wars and the closing of the frontier, the U.S. government turned its attention towards the cultural assimilation of Native Americans, the overwhelming majority of whom was living on the reservations created during the previous half-century. Assimilation efforts were centered on policies that broke tribally owned reservation lands into individually owned 160-acre land allotments. Bureau of Indian Affairs (BIA) Commissioner Thomas Morgan noted in his annual report of 1891 that "if there were no other reason [for allotment], the fact that individual ownership of property is the universal custom among civilized people of this country would be a sufficient reason for urging the handful of Indians to adopt it." Indian allotment was the cornerstone of federal Indian policy from the passing of the General Allotment Act (or 'Dawes Act') in 1887 to the passing of the Indian Reorganization Act (the IRA, or 'Howard-Wheeler Act') in 1934 (Carlson, 1981, p18).[1]

Indian allotments were first placed in a trust managed by the BIA's local superintendents in charge of reservations (the 'Indian Agents'). In-trust status limited allottees' title to the land and did not grant the full property rights needed to sell or collateralize an allotment. Following a time-window of being held in trust, the local Indian Agent could declare allottees "competent," upon which they were eligible to con vert their land into fee-simple, and coupled with this at the same time became citizens (Banner, 2009; Otis, 2014). It soon became clear that many allottees chose to sell their land upon obtaining full legal title, resulting in worries about the erosion of the remaining Indian land base, (Carlson, 1981, p13). By the late 1920s, concerns also increased that the system was liable to abuse and that it failed to improve the economic position of Indians. These concerns were summarized in an influential 1928 report titled 'The Problem of Indian Administration' (Meriam, 1928). This report led to a fundamental re-think of federal Indian policy, and in 1934, the IRA abruptly ended the issuance of new allotments as well as the transferring of land into fee simple. Allotted lands that had not been transferred into fee simple by 1934 effectively became frozen into its trust status.[2]

The universe of Indian allotments has been digitized by the *Bureau of Land Management* (BLM).

[1] While allotment had been practiced in some places even before 1887, our new data show that it really began in earnest with the passing of the Burke Act in 1906, which explicitly linked landownership to citizenship.

[2] Transfers into fee simple became a bit more common again after 1950, but no new allotments were issued after 1934.

These data include for all allotments their date of issuance, their exact geo-location and the date when (if ever) they were transferred into fee simple. Wealth, income, labor market status and education of Native American individuals and households can be observed in the *Full Count Population Census*, with over 300,000 Native Americans in 1940. There is no information on allotments in the Population Census, and name data in the BLM data is so poorly recorded that it is impossible to link it to the Population Census directly. Fortunately, the BIA collected its own complete census of all Native Americans in the late 1920s and early 1930s, the so-called *Indian Census Rolls* (ICR). The ICR has been published in books which have been scanned and from these scans we have constructed a separate dataset that is a near-complete separate census of Indians on reservations.[3] Unlike the Population Census, the ICR did not record socio-economic outcomes, and were intended only to give a complete listing of people. They contain names, ages, household relations, blood quantum, and—critically—allotment numbers. Using allotment number, we can directly link individual records in the ICR to the BLM data, which contains the allotment information. Linking these treatment data to the outcome data in the Population Census is more difficult, requiring us to move beyond standard linkage approaches because the main linkage variables (names and birth years) are recorded with higher-than-average noise in all Native American records at this time.[4]

We therefore develop a new record linkage algorithm that emphasizes the similarity of *households*, adjusting individual record similarity scores up or down to account for the similarity of individuals' households across data sets. After running this code, we identify questionable links based on a number of criteria (for example, if a husband and wife in a Census household are a husband and wife in two different households in the ICR), and manually check several thousand thus flagged households, hard-coding 'true positives' and 'true negatives'. These hard-coded links then constrain the data when the linkage algorithm is run another time. Previous to the manual step, we link about 8,000 "un-flagged" Native American households from the Census to the ICR.[5]

[3] At this time, the universe of Native Americans on reservations is almost the entire universe of Native Americans.

[4] For one, census enumerators made frequent errors in spelling Indian names. For another, Native Americans were frequently anglicizing their names or re-arranging first and last names in this period. (For example, one individual called 'Little Running Bear' in 1930 had become 'Jim Bear' in 1940.) Stated birth years also varied considerably from data source to data source, especially for older people.

[5] In addition to having individual-level linkage scores above a certain threshold, these households do not throw up any flags along a number of criteria used to identify links that are potentially logically inconsistent. We also link the 1930 Population Census to both the ICR and the 1940 Population Census.

Allotment as a policy was effectively a package of two treatments, an unconditional and a conditional transfer program: The first treatment, receiving an allotment, gave the allottee the unconditional right to use the land for their own purposes, as well as the right to leasing rents. Annual proceeds from leasing out an allotment averaged between 150 and 300 dollars in our data, relative to an average per capita income between 100 and 240 dollars.[6] To be eligible to receive the second treatment—obtaining the land in 'fee simple'— the allottee had to be declared "competent" by the local Indian Agent. this second treatment constituted a major wealth transfer. The average proceeds from selling a 160 acre allotment were 2,800–3,800 dollars in our data, and many, possibly even most allottees chose to sell their land once they had full rights to it (Carlson, 1981, p8,13). It was clear—given the policy's objective of cultural assimilation— that improving one's own education and one's children's education, farming the allotted land, and attending church were all actions that marked one out as competent in the Indian Agents' eyes (Carlson 1981, ch4, Golenko 2010).

The Dawes Act applied to all reservations equally, and every household on allotted reservations was to receive an allotment.[7] However, allotment was phased out sequentially across reservations because it required significant investments into surveying and administration on the BIA's part. As a result, only half of all reservations had been allotted by 1934. The challenge with identifying the effect of allotment is therefore that the BIA phased out allotment in a non-random order across reservations. Two considerations appear to have dominated the government's thinking in this regard: It prioritized allotting reservations of tribes that had shown more organized opposition to the federal government in the past, and it prioritized allotting reservations with more attractive land, primarily because "surplus" tribal lands left after allotment could be opened for white settlement (Carlson 1981, ch1, Leonard, Parker, and Anderson 2018).[8] Our approach to this identification challenge is two-fold: First, we condition all estimations on tribe and state fixed effects, with the reservation-to-tribe mapping being many-to-one.[9] This sweeps out unobserved

[6] See authors' calculations in Table 1.

[7] In practice, because we don't have inheritance information and in some places allotment had been rolled out much earlier, we cannot simply assume every household on an allotted reservation in 1940 was "treated", and need to ICR to identify treatment. In the ICR, we find that about ten percent of households on allotted reservations are unallotted. We omit these from the analysis.

[8] As a result, congress passed a number of supplementary acts for the allotment of specific reservations in advance of the Dawes Act's normal operation (Carlson, 1981, p11). Online Appendix Figure 1 is a sales ad for surplus land.

[9] Dippel (2014) and Feir, Gillezeau, and Jones (2019) use tribe fixed effects to identify reservation treatments.

cultural factors that made the government target certain tribes for allotment. Second, to condition out reservation characteristics that could have influenced both the allotment process and household outcomes in 1940, we include all controls for reservations' land quality from Leonard et al. (2018), and a range of additional controls for the socio-economic environment.

When an Indian was declared competent, their allotment was transferred into fee simple, giving them full property rights. To identify the causal effect of this second treatment, we develop an instrumental variable (IV) strategy based on the exogenous rotation of Indian Agents' across reservations over time. Using a newly constructed reservation-year panel of Indian Agents from 1887–1934, we observe considerable variation in agents' propensity to transfer land into fee simple, and interact this with allotment-specific time-variation in the eligibility for transfer: Agents with high transfer propensity would still only transfer land into fee simple after it had been held in trust for a sufficient period of time, and only once the allottee had come of age. Based on this logic, we construct for each allotment a cumulative probability that it had been transferred into fee simple by 1934.[10]

We find large effects of allotment on households economic activities. Within a tribe, we find that allotted households were about thirty percent more likely to be farmers than households of the same tribe on unallotted reservations.[11] Consistent with the independent small-scale farming prevalent in this setting, we find no effects on labor force participation or unemployment, but a reduction in the likelihood of being employed by someone else. Three-quarters of the employment effect are explained by farm status. We find that wages were 4 percent higher for household heads with allotments. The wage effect is not explained by a household's farm status, suggesting instead a reservation-wage mechanism driven by rents earned from allotments. Once farm status is controlled for, there is no effect of allotments on occupational rank. All of these findings are consistent with the allotment policy's purpose of turning Indians into farmers.

[10] Our use of agents' exogenous rotation is akin to the 'judge fixed effect' literature. The majority of papers in this literature gains identification from the raw-data probability that a specific judge makes a certain decision in a static setup, e.g. in applications ranging from criminal sentencing to patent office decisions (Kling, 2006; Di Tella and Schargrodsky, 2013; Galasso and Schankerman, 2014; Aizer and Doyle Jr, 2015; Melero, Palomeras, and Wehrheim, 2017; Dobbie, Goldin, and Yang, 2018; Frandsen, Lefgren, and Leslie, 2019). Our approach differs somewhat in that the decision to transfer land from trust status to fee simple occurred in a repeated duration setup. We first construct year-specific probabilities of transfer into fee simple, and then aggregate these into a cumulative probability.

[11] This finding seemingly contradicts Carlson, who blames allotment for a decline in Indian farming. However, Carlson bases his thesis on aggregate agricultural output, which may well have declined with allotment even as the number of small independent Indian farmers increased.

What is more surprising is that we find very strong effects of being allotted on educational attainment. Both household heads and spouses have about 2.5 years of extra schooling relative to those on unallotted reservations of the same tribe, and are about ten percent more likely to have some high school. We interpret this education effect as an effort by allottees to improve their chances of being declared competent by their local Indian Agents. Four pieces of evidence are consistent with this 'signalling assimilation' story. One, fifty percent of the differences in educational attainment is explained by farm status, suggesting that households used both together as signals of their assimilation. Two, those households who eventually obtained their land in fee simple had even higher schooling than those who did not, providing prima face evidence that schooling helped in being declared competent. Three, the effect of allotment on the educational attainment on children who have left school but who were of school age when the policy ended in 1934 is less pronounced than the effect on parents, comparing either sons to fathers or daughters to mothers in the same household. We also construct a repeated cross-section of the entire Native American population in the four Population Census waves 1910–1940, and study schooling choices in this data.[12] We find that school attendance increased within reservations over time with the expansion of allotments, conditional on either flexible age polynomials or age fixed effects. One may be concerned that the education effects we find are supply-driven, i.e. if the government pursued education-based assimilation efforts in parallel to land allotments. Several facts speak against this: while the federal government had tried assimilationist educational policies in the form of off-reservation boarding schools, their emergence in 1879 pre-dated the thrust of allotment by almost three decades (Gregg, 2018). As early as 1891, parents had to give their "full consent" for a child to be sent to a boarding school (Adams 1995, 65, Lomawaima 1995, 36), and by around 1910 primary and middle school were fully provided on reservations, while it was a parental choice whether to send their kids to high school, be it a local (mostly non-Indian) public one, or an off-reservation Indian boarding school (Hoxie, 2001, 208). Lastly, the pattern we find likely explains a large part of the secular increase in school attendance documented in the BIA Annual Reports from around 1910 on, and this increase turns out to be entirely driven by local non-boarding schools (Gregg, 2018, Fig1).

[12] Educational attainment is only reported in the 1940 Census so that we can only consider school attendance in previous years.

Next, we investigate the effect of the policy's second treatment arm, the additional wealth transfer from having one's allotment transferred to fee simple. This effect is studied conditional on reservation fixed effect, so that we compare households who obtained their land in fee simple only to households on the same reservation whose allotments had remained in trust. We find that obtaining their allotment in fee simple made a household more likely to leave farming, and increased their wealth as measured by the value of their home. Both findings are consistent with the narrative that many households sold their allotments once they obtained full rights to them. We also find that the fee simple treatment increased the likelihood of working for wages, and led to occupational upgrading, suggesting individuals who sold their land had to substitute independent farming for employment, but could also afford to be more picky with the jobs they took. We also find evidence of a higher reservation wage on the intergenerational transmission channel. Children of households with fee simple land attained higher levels of education, beyond what was explained by their parents' education. As well, they had significantly higher wages once out of school. The identification challenge is that eligibility for transfer into fee simple required the allottee to be deemed competent, and that unobservable characteristics that increased an allottee's chances of being deemed as competent could have also influenced observable socio-economic outcomes.

We also investigate the interaction between allotment and the ancestral cultural norms governing individual property rights in a tribe.[13] We find that the effects of fee simple land on wealth (measured by home values), on parents' occupational upgrading, and on children's wages interact strongly with having ancestral traditions of private property. This suggests that a culture of private property rights allowed individuals to better utilize the income shock from obtaining their land in fee simple.

Our paper lifts the lid on the consequences of the federal government's policy of 'Indian allotment' on affected households. The first treatment arm of this policy was to unconditionally grant the right of usage or the annual rents from 160 acres of land to all households on allotted reservations. The second arm was to grant the much more valuable full property rights over this land, conditional on allottees "proving themselves competent". The unconditional component

[13] Tribes' traditions of private property are coded in the *Ethnographic Atlas* (EA). As an aside, the seminal paper on 'property rights' in the economics literature used the varying property rights among different Native American tribes as an illustration of how such rights may be endogenized (Demsetz, 1967).

of Indian allotment relates our paper to a large literature on the effects of household income on various outcomes, primarily children's health and education. Existing evidence suggests that the causal effect of parents' income on children's health and schooling is generally small in developed economies with universal public good provision, but that it can be large for poor households (Akee, Copeland, Keeler, Angold, and Costello, 2010; Aizer, Eli, Ferrie, and Lleras-Muney, 2016; Cesarini, Lindqvist, Östling, and Wallace, 2016; Bleakley and Ferrie, 2016; Haushofer and Shapiro, 2016; Akee, Copeland, Costello, and Simeonova, 2018). In our setting, treated parents were poor on average, and we find sizable effects of parental wealth on children's education and wages. The conditional component of Indian allotment relates our paper to a large literature on conditional cash programs, to which our setting provides an interesting point of comparison. Modern day programs typically involve ongoing smaller cash payments to families conditional on children's attend school (Schultz, 2004; Barrera-Osorio, Bertrand, Linden, and Perez-Calle, 2011; Behrman, Parker, and Todd, 2011; Parker and Todd, 2017; Barrera-Osorio, Linden, and Saavedra, 2019). By contrast, being declared competent in our setting translated into a large one-time wealth transfer, but this transfer could lie far in the future and was only implicitly linked to educational attainment. Nonetheless, it appears to have provided a powerful incentive for educational attainment in the setting we study.

Our findings speak to a literature on the effects of cultural assimilation policies. Such policies typically work through schooling (Sakalli, 2017; Bandiera, Mohnen, Rasul, and Viarengo, 2018; Fouka, 2019). This is no less true in Indian country, where Indian boarding schools were the primary assimilationist policy in the pre-allotment era.[14] Gregg (2018) shows that reservations that historically sent a higher share of children to boarding schools have more assimilated populations today. At a more micro level, Feir (2016b) studies Canadian residential (boarding) schools , and —using an identification strategy that addresses selection into attending these— shows that they increased the educational attainment and later labor force participation of students. Assimilationist effects show up in a higher likelihood of living off-reserve, and a lower likelihood of speaking an indigenous language. In a follow-up, Feir (2016a) finds broadly negative intergenerational impacts of mothers' residential schooling experience on children's school experiences,

[14] In Canada, where no parallel to allotment was tried, and Indian residential schools continued into the late twentieth century, schooling remained the main assimilationist policy throughout.

possibly reflecting a longer-run form of cultural "backlash" that has also been documented with other assimilationist policies (Sakalli, 2017; Fouka, 2019). In contrast to the assimilationist policies studied in these papers, allotment did not use education to assimilate. Instead, its design as a wealth transfer that was made conditional on signalling one's cultural assimilation appears to have incentivized families indirectly to increase school attendance.

Our paper also contributes to a growing literature emphasizing the importance of cultural norms as drivers of economic outcomes and decision making (Algan and Cahuc, 2010; Fernández, 2011; Nunn, 2012; Costa-Font, Giuliano, and Ozcan, 2018; Dohmen, Enke, Falk, Huffman, and Sunde, 2018; Enke, 2019).[15] We contribute to this literature by showing that ancestral cultural norms of private property were an important mediating channel that determined how successfully households navigated the transition in the property rights regime imposed by the policy of Indian allotment.[16] Giuliano and Nunn (2017) argue that societies that have historically lived in more stable conditions evolved norms that place greater value on tradition. When the conditions drastically change, adherence to tradition can become costly for societies if the traditional cultural norms are ill-suited to dealing with the new conditions. We provide concrete evidence that this was the case. It is clear that Native Americans had norms that placed great value on tradition, and that they faced drastic 'environmental changes' during the reservation and assimilation eras. We show that the cultural groups that had traditions of individual property rights were better able to navigate these changes.[17]

Before proceeding, a comment is in order on what our results imply about the aggregate welfare consequences of Indian allotment. Although we estimate positive effects of allotment on educational attainment and on farming, and positive effects of obtaining land in fee simple on wealth and occupational upgrading, these findings should not be misread as saying that tribes

[15] This literature has also brought to light the tight empirical connection between present-day and ancestral cultural norms. Cultural norms evolve over the long run to optimally regulate a groups' economic and social lives, and they have been shown to be remarkably persistent over time in a variety of contexts (Spolaore and Wacziarg, 2013; Galor and Özak, 2016; Becker, Enke, and Falk, 2018). Certainly, the EA's 'ancestral tribal norms' would have still been very much just the tribes' culture at the time we investigate them between 1906 and 1940.

[16] To some extent this is the flipside of Ostrom (1990) who argued for the importance of culture in making collective property management function without having it deteriorate into the tragedy of the commons.

[17] A related question that we do not explore is whether some tribes' (or even households') cultural norms in relation to private property may have changed in response Indian allotment. Di Tella, Galiant, and Schargrodsky (2007) provide evidence for such a mechanism, exploiting quasi-random variation in giving legal title to squatters in Buenos Aires to identify the effect of property rights on beliefs and norms related to the market.

were better off with allotment, even in a narrow economic sense.[18] There is no doubt that allotment is viewed very negatively in Indian country today, and we do not mean to challenge this view as it prevails by the descendants of the affected population. One obvious reason allotment is viewed in a negative light is that it was a colonial policy aimed at the cultural assimilation of tribes. A second—and we think very important—reason has less to do with allotment itself and more with the ill-conceived inheritance rules that the BIA designed for in-trust land after 1934. These inheritance rules forced in-trust land to be bequeathed to all heirs in *multiple, undivided interests*, which led to an exponential increase in heirs and claimants on these lands over time, and created sizeable inefficiencies in land use for later generations of Native Americans (Leonard and Parker, 2017; Dippel, Frye, and Leonard, 2019). While these problems are clearly a legacy of Indian allotment, they have less to do with allotting per se, and more with not giving full rights to all allottees. They also post-date the setting we study here.

2 Background

2.1 Land Allotment on Reservations

Following the establishment of the reservation system, American Indian reformers considered land allotment as a requisite element in the assimilation of American Indians (Otis, 2014). Early allotment efforts were negotiated into treaties, but legislation implementing a general allotment act stalled over the issues of citizenship, jurisdiction, and whether to immediately grant allotted Indians title to their land.

In 1886, Henry Dawes introduced a modified allotment bill to the Senate. The bill quickly passed before moving to the House, where it passed after the addition of several amendments. On February 8, 1887, President Grover Cleveland signed the Dawes General Allotment Act into law. The Dawes Act authorized the president, through the Office of Indian Affairs, to survey and allot

[18] On the one hand, the seminal work in Ostrom (1990) clearly delineates many successful cases in which communal (tribal) land- or resource ownership can work well. On the other hand, it is not clear that tribal governance over a commonly owned reservation land actually was more consistent with many Native American tribes' traditions and customs. In fact, a number of papers demonstrate that economic under-development on reservations today is partly created by political control over reservation resources by tribal leadership (Anderson and Parker, 2008; Dippel, 2014). Consistent with this, tribal governments on many reservations today are in the process of changing institutions to reduce communal tribal control over resources in order to promote economic development (Regan and Anderson, 2016; Jorgensen, 2007).

reservation lands deemed appropriate (Banner, 2009).[19] Heads of household received 160 acres, single persons over 18 received 80 acres, orphans under 18 received 80 acres. The Dawes Act was amended in 1891 to grant 80 acres to every adult, instead of 160 acres to heads of households. If the land was only suitable for grazing the allotment amounts doubled. If a prior treaty specified larger allotments, the prior treaty acreages were applied. Allotments were mandatory and anyone not selecting an allotment within the first four years, would be assigned a parcel by the Indian Agent. Unallotted reservation land was designated as surplus and made available for outside settlement. The law required tribal approval of ceded surplus land, but tribes were rarely in a position to negotiate (Carlson, 1981).[20] Proceeds from the sales of surplus land were held in trust and appropriated at the discretion of Congress for "education and civilization" (Banner, 2009).

Once selected, allotments were approved by the Secretary of Interior and each Indian was issued a trust patent. This patent held the allotted land in trust for a trust period, during which the Indian or their heirs were the beneficiary of the allotment. Land held in trust could not be alienated or leased and was not subject to state or local taxes. At the end of trust period, the allotment would be transferred to the owner as fee-simple. In 1906, the Burke Act granted the Commissioner of Indian Affairs the authority to shorten or lengthen the 25-year trust period for individual allotments. Shorter trust periods were often at the recommendation of the Indian Agent (Carlson, 1981). The Burke Act gave agents considerable authority over the process of converting land from trust status to fee-simple.

2.2 Administration on the Ground

Implementing the Dawes Act on an individual reservation was a complicated process. First, the BIA agent in charge of the reservation was tasked with determining the list of eligible tribal members entitled to an allotment and the household structure for every household within the reservation (Banner, 2009). These agents were also tasked with managing the surveying of the reservation and its division into parcels. The BIA agents possessed considerable authority over the assignment of allotments, even for those eligible Indians that desired an allotment. (Banner, 2009; Otis, 2014; Carlson, 1981). In cases where the Indian did not select a plot, the agent assigned an allotment.

[19] Tribes in New York and Indian Territory were temporarily exempted from the Dawes Act.
[20] By 1903, tribal approval was no longer necessary.

Each allotment was given an allotment number and a patent was filed with the Government Land Office upon approval by the President. These official patents specified the trustee, the specific plot location, the date, and the unique allotment number. Reservations were either allotted all at once or over a period of several years.

The implementation of the Dawes Act moved quickly on the extensive margin, where nearly four reservations per year were designated for allotment over the first two decades (Otis, 2014). However, constraints over personal and resources slowed surveying and the allocation of allotments in the early years. Following the introduction of the Burke Act, the pace of allotting accelerated considerably.

Concerns about the lack of development of Indian farmers, expansions in leasing, and sales of Indian land to settlers led to a change in public opinion regarding allotment culminating in a review of the current social and economic conditions on reservations by Lewis Meriam of the Institute of Governmental Research. The Meriam Report, published in 1928, was critical of the support provided to Indians by the Office of Indian Affairs (Meriam, 1928). This report led to a shift in federal Indian policy, brought to fruition by President Roosevelt's new Commissioner of Indian Affairs, John Collier. Collier introduced a bill that fundamentally changed Indian policy. In 1934, the Indian Reorganization Act (IRA) ended the allotment of Indian reservations. The IRA returned unallotted lands back to tribal ownership and froze allotted trust land in it's trust status, creating a patchwork of land tenures within Indian reservations.

2.3 Legacy of the Allotment Era

In total, the government extended the Dawes Act to 118 reservations and issued over 245,000 patents covering nearly 41 million acres (Office of Indian Affairs, 1935). On the one hand, Indian allotment resulted in a substantial transfer of revenues that previously would have gone to tribes, and now went to individual tribal members. Table 1 shows that a 160 acre allotment would have generated annual leasing proceeds of between 150 and 300 dollars in our data (e.g., $160 \times 1.73 = 270$ in 1919), relative to average per capita incomes from all sources between 100 and 240 dollars during the same time. At the same time, the average proceeds from selling a 160 acre allotment that was obtained in 'fee simple' would have been between 2,800 and 3,800 dollars in our data (e.g., $160 \times 21.14 = 3382$ in 1919).

Table 1: Leasing Rates and Sales Prices of Indian Allotments

	Leasing Revenue per Acreage				Sales Price per Acreage				Annual Income		
Year	#Leases	Area in Acres	Income	Annual Lease Price per Acre in $	Area	Income	Sale Price per Acre in $	Sales Price / Lease Revenue	Pop.	Total Income	annual per cap income
1920	23,550	2,254,632	4,406,308	$1.95	147,047	3,566,816	$24.26	12.51	306,539	72,696,431	$237
1919	22,658	2,198,753	3,809,291	$1.73	57,947	1,224,823	$21.14	12.55	304,974	53,994,859	$177
1918	19,073	2,145,553	3,067,875	$1.43	74,126	1,541,178	$20.79	14.78	306,755	42,056,070	$137
1917	20,567	2,023,788	2,615,639	$1.29	69,849	1,040,202	$14.89	13.77	309,409	35,867,696	$116
1916	22,612	2,357,542	2,603,498	$1.10	54,959	969,611	$17.64	18.65	307,797	26,489,948	$86
1915	16,500	2,415,794	2,117,166	$0.88	34,429	584,724	$16.98	19.38	309,911	23,193,046	$75

Notes: Leasing Rates and Reservation Incomes Data come from Tables 12 and 11 of the 1920 Annual Report of the Commissioner of Indian Affairs. (The level of detail in these reports was dramatically reduced after 1920, and no longer includes leasing revenues.) Sales prices come from United States National Resources Planning Board 1935 Report *Indian Land Tenure, Economic Status, and Population Trends*

On the other hand, Indian allotment resulted in a substantial transfer of land out of Native and into non-Native hands. Prior to the Dawes Act, Indians controlled over 138 million acres of lands within their reservations. By 1934, Native land holdings had fallen to 52 million acres. Of this 85 million acre reduction, nearly 60 million acres had been ceded as surplus and the remaining were sold as fee-simple or alienated by the Secretary of Interior (Office of Indian Affairs, 1935). Within reservations, the Dawes Act created considerable variation in the status of land tenure between land held in individual trust, and land held in fee simple. Parcels of differing tenure types are often adjacent to one another, creating a checkerboard pattern of land tenure on many reservations. Figure 1 illustrates this pattern on the Pine Ridge Reservation, using the data discussed in Section 3.1.

3 Data Sources

3.1 Allotments in the BLM Land Data

Following approval from the President, each patent issued on the reservation was filed with Government Land Office and was digitized by the BLM. These patents record the transfer of land titles from the federal government to individuals. Each patent contains information regarding the patentee's name, the specific location of the parcel(s), the official signature date, total acreage, and

the type of patent issued. Patent types include cash sales, homestead entries, and Indian allotments. The patent also includes the Indian allotment number associated with the transaction. A nice feature of the BLM data is that we can see exactly the date on which each patent was issued (in trust) and the date on which it transferred into fee simple, if ever.[21]

3.2 Measuring Outcomes: The Full-Count Population Census

The 'rule of 72' dictates that non-anonymized individual-level Census information can be made available 72 years after a Census was published. The 1930 and 1940 Full-Count US Censuses (FCC) each include around 300,000 individual Native Americans living in roughly 80,000 households (with some intermarriage). The FCC data provides us with measures of individuals' and households' incomes, occupational rank, property ownership, and a range of other outcomes. Critically for our purposes, there is no way to directly link the FCC to the BLM land information: the allottee name information in the BLM data is inconsistent, and the FCC contains no information at all related to allotment. Section 3.3 introduces newly collected data that allow us to build a bridge between the BLM data and the FCC.

3.3 Measuring Treatment: The Indian Census Rolls

To be able to link BLM land allotment information to the FCC, we need to link both the BLM and the FCC to the *Indian Census Rolls* (ICR). The ICR were special censuses that were unrelated to the FCC and were collected on reservations consistently from the 1880s to the late 1930s. The ICC did not record any economic outcome data; instead they are simply complete enumerations of all Native American households. Online Appendix Figure 2 depicts a sample page from the ICR. They include only individuals' names, their ages, and their relations within the household (e.g. spouse or son). Critically, they also report whether individuals received land allotments, including a complete listing of unique allotment numbers by individuals. We digitized the entirety of the ICR from the mid-1930s.

The first of two steps is to link the ICR data match uniquely to the BLM data via allotment number. We match about 85% of the allotments in the ICR to the BLM data. The remaining 15%

[21] This ability to "follow the land" and the ability to distinguish parcels that were converted to fee simple is of separate interest because of the large amount of allotted land that became "trapped" in trust in 1934. This topic is the focus of Dippel et al. (2019).

are likely mis-recorded in the ICR. This is illustrated Figure 1 for the the Pine Ridge Reservation. We omit households with unmatched allotment numbers from our analysis. The second step is to link the ICR to the FCC, using record linkage methods described in Section 4.

3.4 The Indian Agents

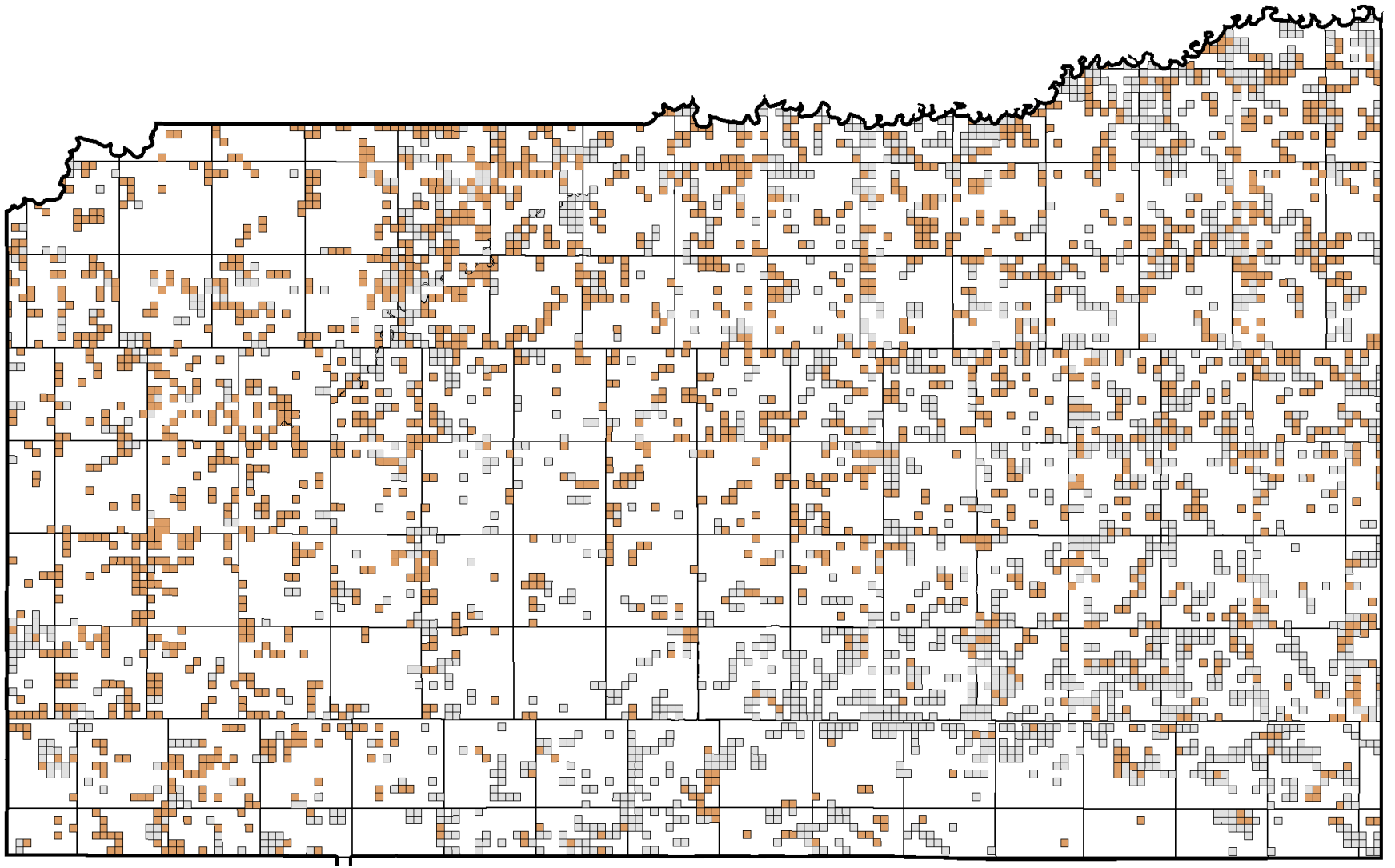
To gain identification on the the causal effect of obtaining land in fee simple, we pursue a strategy based on the exogenous rotation of Indian Agents across reservations, and their varying propensity to transfer land into fee simple. To operationalize this, we construct a complete reservation-year panel of Indian Agents from 1879–1940. Our primary source of agent information is from the Department of Interior employment rosters recorded in the Official Register of the United States (1932).[22] The records provide agent name, birthplace, position title, and annual pay. Each agent is listed by agency and city, which we link to reservations. We supplement these records with two additional resources produced by the Bureau of Indian Affairs. First, we use the agent narratives included in the Bureau of Indian Affairs Reports published annually from 1879 to 1907. Each agent was required to produce an annual summary of agency events. We recorded each agents name from the end of the summary. Second, we use the ICR data, which included the name of the agent in charge of taking the census from the beginning of each census year.

3.5 Ancestral Tribal Norms

The data on ancestral cultural norms comes from the Murdock (1967) *Ethnographic Atlas* (EA). Among many other measures, the EA has tribal information on property rights norms. A second measure that we will include in our analysis when we consider the children of the original allottees is whether a tribe had ancestral norms of having a bride price, i.e. of the groom's family paying a material transfer to the bride's parents. Online Appendix Table 1 provides details on the Ethnographic Atlas's categories for each variable, which we convert into a binary indicator variable for ease of interpretation.

[22] The Official Registers were published biennially from 1879–1940.

Figure 1: Checkerboard Pattern of Land Tenure on the Pine Ridge Reservation



Notes: Distribution of Land tenure on the Pine Ridge reservation by allotment parcel (quarter-section) in the BLM data described in Section 3.1. The reservation is divided into 36-square mile townships. A township itself consists of 36 one-square-mile sections, and a quarter-section corresponds to a 160-acre allotment. Each quarter-section is either an allotment transferred into fee simple (dark/orange), an allotment that remained in trust (light-grey), or else it is tribal land that was never allotted or an allotment that had no match in the ICR (blank).

4 Record Linkage

Measuring treatment and constructing the instrument requires either only the BLM data, or [if allottee characteristics are included in the zero-stage equation (1)] the BLM data and the ICR data. These two data sets are straightforwardly linked by allotment number. Once treatment and instrument are attached, however, the ICR needs to be linked to the Full Count Census (FCC) data for the second-stage analysis, and this record-linkage is difficult. The main individual linkage variables at our disposal in both ICR and FCC are first and last name, birth-year, location, and gender. Fortunately for us, spatial mobility among Native Americans in this period of time was very low, with the vast majority of Native Americans living on reservations or former reservation lands in both 1930 and 1940.[23] We can therefore block the record linkage on location. In our case, because some reservations in the Southwest straddle state borders, we create 40 meta-states so that each reservation is uniquely contained in one, and then block on these. Secondly, we also block on gender.[24]

Within this blocking, individuals are linked by name and birth-year. The challenge in our data is that both of these linkage variables are recorded with a lot of noise. For names, this is partly driven by the fact that Census enumerators made more frequent errors in recording Native American names because they were unfamiliar with them. It is further driven by the fact that there was a strong trend of anglicization of names in this period so that names changed over time. For birth-years, noise is introduced into the records by the fact that most Native Americans at that time did not have birth certificates, and were thus more likely than other populations to revise their stated birth-year as time passed. These challenges make it impossible to obtain a large number of unambiguous individual record links in these data based on only individuals' information. To make progress in this direction, and ultimately enable us to perform the second stage analysis, we had to move beyond existing individual-based record linkage methods and develop a linkage scheme that incorporates the similarity of households' structure across data sets.[25;26]

[23] Substantial off-reservation mobility really took hold in the 1950s during the 'Termination Era', in which the federal government attempted to dissolve reservations.

[24] The FCC does not report reservation or tribe, but county of residence, which is sufficient to uniquely place most individuals in a reservation.

[25] Existing linkage methods are almost entirely focused on individual linkage. The one exception that uses household structure for record linkage that we are aware of is Price, Buckles, Riley, and Van Leeuwen (2019).

[26] Methods of historical record linkage are fast evolving, driven by advances in the access to historical individual level data, computational power needed for linkage algorithms, the ability to scale up manual linkage through online

Table 2: Household-Based Linkage Example

ICR [master dataset]							FCC [using dataset]				
hhid	birthyr	namelast	namefrst	relate	links	score	hhid	birthyr	namelast	namefrst	relate
35545638	1908	SESSPOOCH	WAUN	Head	2	6.3	79055	1907	CESSPOOCH	JUAN	Head
35545638	1901	SESSPOOCH	ELLEN	Wife	2	11.9	79055	1902	CESSPOOCH	ELLEN	Wife
35545638	1934	SESSPOOCH	LOUIS	Son	5	11.2	79055	1933	CESSPOOCH	LEWIS	Son
35545638	1937	SESSPOOCH	DEBOIA	Daughter	3	6.3	79055	1936	CESSPOOCH	DOVELIA	Daughter

Notes: This table shows how household-information can increase the confidence of individual level record linkages. No individual link looks very compelling as not a single person is a perfect match on age and name. Yet, viewed in combination, these two households are clearly the same.

Table 2 provides an example for the potential power of using household information in data like these. The table shows two four-person families (one in the ICR, and one in the 1940 FCC). The noteworthy feature of the table is that not a single person is a perfect match on age and name. The shared last name is different in the two data-sets, all birth-years are off by one year, and only one of four first names is an exact match. Nonetheless, viewing the four records in combination gives a high degree of confidence in this being the same household in the two data sets. This example illustrates the potential usefulness of household structure in determining individual links, when individual linkage variables are measured with a lot of noise. In the following we describe our step-by-step approach to record-linkage:

1) We first use standard individual record linkage methods to establish for each individual in a master data-set (say, the ICR), the set of all potentially linked individuals in a using data-set (say, the 1930 FCC). This can be any of the commonly used approaches discussed in Bailey, Cole, Henderson, and Massey (2017), Abramitzky, Boustan, Eriksson, Feigenbaum, and Pérez (2019), and Online Appendix A.0.1. We pursue a two-step approach: we first apply bigram-indexation to first and last name to define the set of individuals in the using data with similar names.[27] We are 'generous' in this set, in the sense that we set the string-similarity cutoff low enough that each person in the master data has an average of 8 links in the using data.

job platforms, and machine learning capabilities. For a review of the historical record linkage literature, see Online Appendix A.0.1.

[27] We use the `stringsim()` function in R's `RecordLinkage`, and block on gender and on 'meta-state' and gender. (A 'meta-state' is a set of two states whose boundaries are straddled by reservations. This occurs, e.g. in New Mexico and Arizona. We formed 40 meta-states.

We then trim that set to include only potential matches that deviate by at most 4 years in birth-year for master records born after 1910, and by at most 6 years for master records born before 1910.

2) Each individual in the master data and each individual in the using data is associated with a unique household id. Consider a pair of households ICR-HH A in master and FCC-HH B in using. Our approach is to calculate the number of individuals in ICR-HH A who are linked to FCC-HH B in stage 1. This gives a metric of household similarity. We can then upward-adjust the similarity-scores of individual links between individuals in ICR-HH A and FCC-HH B by a percentage for each marginal increase in our metric of household similarity. We also re-generate adjusted similarity-scores between all non-linked individuals in ICR-HH A and FCC-HH B to potentially find new links. The idea is that these new links did not meet the match-threshold in the individual-level linkage in stage 1, but do meet the threshold with the upward-adjusting. Household similarity therefore brings new individual links into the fold.

   The researcher has three levers to control the adjustment process in stage 2: One, household similarity can be calculated in different ways. For example, we use the absolute number of stage 1 individual links as our metric. This inherently creates more adjustments in larger households than in smaller ones. One could alternatively use the share of household members who are linked. As another alternative, one could also create a series of household-pair-specific dummies for binary measures such as whether there is both a male and female head of household, whether a household is a nuclear or an extended family, etc.

   Two, the gradient of how individual similarity is adjusted to household similarity can be controlled. With more aggressive upward-adjusting, the researcher puts more weight on household-similarity. With less aggressive upward-adjusting, the researcher puts more weight on individual similarity.

   Three, one could define a series of explicit rules based on household types in a pair. For example, if the only individual that is linked across ICR-HH A and FCC-HH B, is a 'son' in the ICR, and a 'head' in the FCC, then two households are likely to be otherwise completely disjoint sets, and the researcher might want to prevent the one exiting individual link's simi-

larity score to be downward-adjusted (in relative terms) for the lack of household-similarity. To what extent these levers are used will be dictated by the structure of the specific dataset as well as by computational concerns.

3) Then a "best-off" gridsearch is applied on the adjusted similarity scores created in stage 2. The grid-search maximizes the sum of similarity-scores of all links, with the constraint that each individual record in the master data is linked to at most one record in the using data, and each record in the using data is linked to at most one record in the master data.

4) Relative to individual linkage methods, steps 2 and 3 above have the potential to create additional links to the extent that household similarity can lead individual similarity scores to be scaled up. This is evident in Table 2 where no individual link looks particularly compelling when viewed in isolation, but the links look very compelling when viewed in combination.

   More importantly, the household structure also suggests a range of criteria to generate flags for potentially problematic links, based on reported family relations. For example, while it is reasonable for the husband and the wife in a household in the (later) master data to be linked to records in the (earlier) using data where they are recorded as a son and a daughter, they should then not share a household in the using data, lest they be brother and sister.[28] The final linked data-set will invariably be improved by adding a fourth stage of manually checking links that that throw up flags for family relations that are either logically consistent or unlikely.

Manual linkage after stage 4 is greatly aided by the fact that stages 2–3 have identified the most likely individual link for each member of a given "anchor household" in the master data, so that one can view all members in the anchor household together with all person records belonging to households with any links to that anchor household. In other words, manual links can at this stage be confirmed or changed with all relevant household relations in view. Table 3 shows two

[28] Of course, if master and using data are collected in the same year, we should expect family relations to be exactly identical in the two datasets. Any time gaps between the years in which master and using data were recorded will imply some differences in household structure between the data-sets. For example, the ICR was collected between 1930 and 1937 and is therefore likely to contain more households whose nucleus is unchanged. By contrast, when linking two decadal Census waves like the 1930 and 1940 FCC, a sizeable share of 1930 children will have formed new households in 1940, and become heads or spouses. As well, more 1930 household heads and spouses will have passed away, and two-parent households will thus have turned into single-head households.

examples of how our research assistants manually check flagged households in practice. The flagged household is the "anchor household," and associated with it is "network" of direct and indirect links to person-records in other datasets. In the top-panel, that anchor household has id 15426032 in the 1930 Full Count Census. In the bottom-panel, that anchor household has id 10 in the ICR. A network consists of all individuals that belong to any household linked to the anchor household. For instance, in the top-network, 1930-household 15426032 has one link to an individual in 1940-household 19002788, so that all five members of that household are included in the bloc. Household id, the number of people in a household, and individual id's are reported by data-source; columns 1–3 report on the 1940 FCC, columns 4–6 report on the 1930 FCC, and columns 7–9 report on the ICR. One observations is one person-record. If a person in a 1930-household is individually linked to persons in the 1940 Census and the ICR after stage 2, then there are entries in columns 1–3, 4–6, and 7–9 of that person record.

Columns 16–33 report on six variables that represent the relevant person- and household-information, again separately for each of three data sources. These variables are a person's first and last name, birth-year, gender, their marital status, and their relation to the household head. Based on this information, we can assess whether existing links are correct, and make manual adjustments in columns 10–15, which then get hardcoded into the data before re-running stages 2–3 of the program above. Columns 10-12 are 'true positive' links. A 'true positive' link can confirm an existing link, or it can over-ride an existing link. Columns 13-15 are 'true negative' un-links, which declare that an existing link is wrong. To avoid clutter for illustrative purposes, Table 3 omits entries in column 10–15 that merely confirm existing links, and only includes entries for records where at least one link gets manually over-ridden. In the top-record for example, columns 11-12 confirm the existing link between the 1930 FCC and the ICR, but column 13 unlinks these from their link to a person-record in 1940. In the fourth row, columns 11-12 establish a new link between a 1930-record called Josephine Rose and an ICR-record called Josephine Rose who differs by 11 years in birth-year but shares an ICR household with three others who are linked to 1930-records in Josephine Rose's 1930 household. Column 15 in the same row lists the unlinked previous link.[29] In the bottom-bloc, columns 11-12 link four previously unlinked individuals in

[29] This particular column 15 entry is redundant because columns 11-12 already hard-code a unique 'true positive' link between these records, thus precluding any other possible links.

the 1930 census to four previously unlinked individuals in the ICR. The records were not linked by the algorithm because of unstable last names ('Big' in the ICR and 'Big Knife' in the 1930 FCC) and first names ('Dorothy Ann' in the ICR, and 'Pretty Woman' in the 1930 FCC).

5) Manual changes made in stage 4 are the 'true positives' and 'true negatives' in columns 10–15 of Table 3.

   Hard-coded 'true positives pairs' mean that any other individual links with either record in a pair get deleted. Hard-coded 'true negative pairs' mean that this particular link gets deleted from the stage 1 output.[30] These become hard-coded and are fed into the data-output generated at the end of stage 1.

The researcher can choose the decision rules that determine which records to manually check and link by hand. In our case, we defined a number of different flags for households with inconsistent linkages between the 1940 FCC and the ICR. Possible linkages and their consistency are visualized in Table 4. The side of the table content to the left of the arrow $\Rightarrow$ depicts the 1940 households for which we can study outcomes. Across columns 1–5, the right side of the table content depicts possible ways in which a 1940 household can be linked to ICR households. The top-two panels depict linkages that are logically consistent, firstly for nuclear households with a head and a spouse, and secondly for single-head households. In the top panel, column 1 indicates that both head and spouse in a 1940 FCC household are linked to a head and spouse who are in the same household in the ICR. Column 2 indicates that both 1940 head and spouse are linked to persons in the same household, and that neither of these is head or spouse in the ICR. This captures the relatively common occurrence of young couples living with one spouse's parents. Column 3 indicates that 1940 head and the spouse are linked to persons in the ICR who live in different households and are neither head nor spouse in the ICR. This captures the most common occurrence of new households being formed by young adults. Columns 4–5 capture the case where only one of 1940 head or spouse is linked to a person record in the ICR, and that record is not a head or spouse in the ICR. Column 5 is common because young wives in 1940 who were living with their parents in the ICR are less likely to be linked because of their different last names. The middle panel depicts single-head households in which the head is linked to person-record in the ICR. Such links are

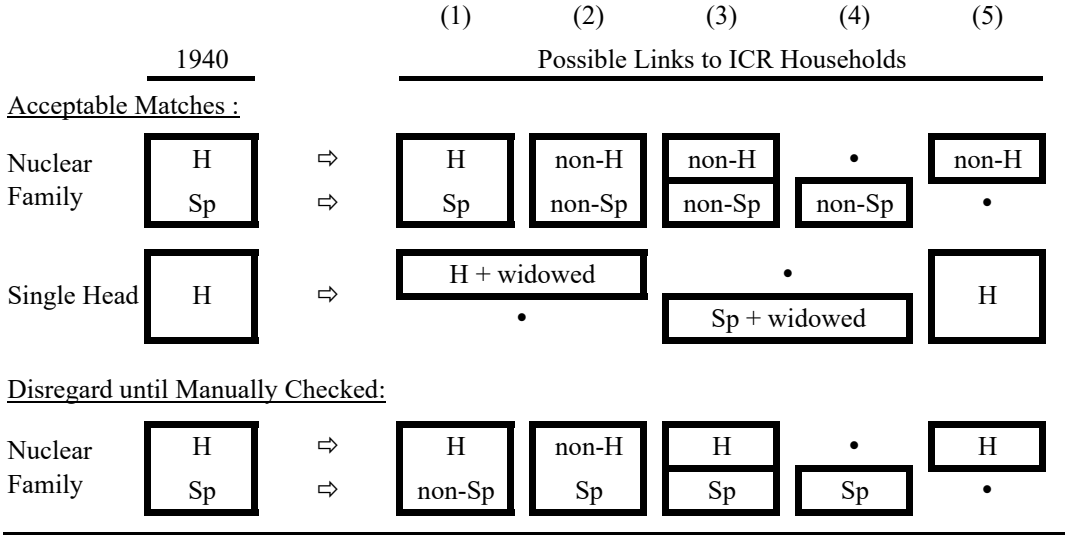
[30] These need to get deleted a second time at the end of stage 2 in case that any are re-introduced during stage 2 by the similarity-score adjustment.

Table 3: Household-Based Linkage Example

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)	(29)	(30)	(31)	(32)	(33)
1940			1930			ICR			true positives			true negatives			1940						1930						ICR					
hhid	#	pid	hhid	#	pid	hhid	#	pid	pid 1940	pid 1930	pid ICR	pid 1940	pid 1930	pid ICR	LN	FN	Birth Y	Sex	Mar ried	Rel	LN	FN	Birth Y	Sex	Mar ried	Rel	LN	FN	Birth Y	Sex	Mar ried	Rel
Anchor-HH : hhid=15426032 in 1930																																
19002788	4	19002788-1	15426032	4	15426032-1	68934	5	9-806-11		15426032-1	9-806-11	19002788-1			RACE	WILLIAM	1891	M	M	Head	ROSE	WILLIAM	1891	M	M	Head	ROSE	WILLIAM	1890	M	M	Head
19051844	5	19051844-2	15426032	4	15426032-4	68934	5	9-807-2							ROCE	FOREST	1925	M		gr-ch	ROSE	FORREST	1925	M		child	ROSE	JAM FORREST	1924	M		child
19051844	5	19051844-3	15426032	4	15426032-3	68934	5	9-807-1							ROCE	JAUNITA	1922	F		gr-ch	ROSE	WANETA	1923	F		child	ROSE	JUANITA	1922	F		child
			15426032	4	15426032-2	67409	3	86-10-2		15426032-2	9-806-12			86-10-2							ROSE	JOSEPHINE	1890	F	M	Sp	SHANE	JOSEPHINE	1891	F	M	Sp
19051844	5	19051844-5	15463552	8	15463552-7	68452	16	9-674-17							REDFOX	CYNTHA	1922	F		gr-ch	RED FOX	CYNTHIA	1922	F		step-ch	RED FOX	CYNTHIA	1921	F		step-ch
19051844	5	19051844-1	15463565	2	15463565-2	68933	2	9-806-10							ROCE	AGNES	1861	F	W	Head	ROSE	AGNES	1862	F		Sp	ROSE	AGNES	1860	F	M	step-ch
19066701	9	19066701-5	15477934	7	15477934-6	67409	3	86-10-3							BEAS CUB	ARCHIE	1929	M		child	BEARCUB	ARCHIE	1928	M		child	ARGROUND	ARCHIBALD	1930	M		Sp
19002788	4	19002788-4				68258	12	9-627-12							BEARDEAU	MELVIN	1928	M		step-ch							BURDEAU	MELVIN JOE	1927	M		child
19002788	4	19002788-3				68258	12	9-627-11							BEARDEAU	JAMES	1926	M		step-ch							BURDEAU	AMES ADAMS	1924	M		child
19002788	4	19002788-2													RACE	MAGGIE	1890	F	M	Sp												
19062368	11	19062368-5				68934	5	9-807-3							LA ROGUE	DOROTHY	1930	F		child							ROSE	DOROTHY	1930	F		child
						67409	3	86-10-1																			ES TO OTHER	GROUND	1870	M	M	Head
						68934	5	9-806-12																			ROSE	OSEPHINE N S	1901	F	M	Sp
Anchor-HH : hhid=10 in ICR																																
19009442	3	19009442-1	15429377	8	15429377-6	10	9	426-617-4							BIG DOY	OLD WOMAN RUTHY	1890	F	W	Head	BIG DOG	OLD WOMAN PRETTY	1868	F	W	Parent	BIG DOG	OLD WOMAN PRETTY	1868	F	W	Head
19009442	3	19009442-2	15429377	8	15429377-7	10	9	426-617-5							BIG DOY	LODGE	1910	F		child	BIG DOG	LODGE	1909	F		Sibling	BIG DOG	LODGE	1908	F		child
			15429377	8	15429377-8	10	9	426-617-6													BIG DOG	ALBERT	1912	M		adop-ch	BIG DOG	ALBERT	1913	M		neph
						10	9	426-617-10																			BIG KNIFE	JOHN	1928	M		child
						10	9	426-617-9																			BIG KNIFE	DOROTHY ANN	1926	F		child
						10	9	426-617-11																			BIG KNIFE	JOSEPH	1933	M		child
						10	9	426-617-8																			BIG KNIFE	FLORA	1907	F	M	Sp
						10	9	426-617-7																			BIG KNIFE		1900	M	M	Head
						10	9	426-617-12																			BIG KNIFE	DOROTHY	1935	F		child
19009442	3	19009442-3													FOUR GAESTO	COLORE THE	1887	F	W	child												
			15429377	8	15429377-1					15429377-1	426-617-7										BIG	KINFE	1901	M	M	Head						
			15429377	8	15429377-4																BIG KNIFE	LEVEL WALING	1928	F		child						
			15429377	8	15429377-3					15429377-3	426-617-9										BIG	PRETTY WOMAN	1926	F		child						
			15429377	8	15429377-2					15429377-2	426-617-8										BIG	FLORA S	1908	F	M	Sp						
			15429377	8	15429377-5					15429377-5	426-617-10										BIG KNIFE	HAD LIGHTENI NG	1929	M		child						

Notes: (*a*) Columns 20, 26, 31: M=married, W=widowed. Columns 21, 27, 32: SP=spouse, step-ch=step-child, adop-child=adopted child, gr-child=grand-child. (*b*) This table reports on two "household networks" Each household network has one "anchor household." In the top-panel, that anchor household has id 15426032 in the 1930 Full Count Census. In the bottom-panel, that anchor household has id 10 in the ICR. Household id, the number of people in a household, and individual id's are reported by data-source; columns 1–3 report on the 1940 FCC, columns 4–6 report on the 1930 FCC, and columns 7–9 report on the ICR. One observations is one person-record. If a person in a 1930-household is individually linked to persons in the 1940 Census and the ICR after stage 2, then there are entries in columns 1–3, 4–6, and 7–9 of that person record. Columns 10-12 are 'true positive' links. A 'true positive' link can confirm an existing link, or it can over-ride an existing link. Columns 13-15 are 'true negative' un-links, which declare that an existing link is wrong. To avoid clutter for illustrative purposes, Table 3 omits entries in column 10–15 that merely confirm existing links, and only includes entries for records where at least one link gets manually over-ridden.

Table 4: Possible Household Links, Grouped by Internal Consistency



Notes: (*a*) The side of the table content to the left of the arrow ⇒ depicts the 1940 households for which we can study outcomes. Across columns 1–5, the right side of the table content depicts possible ways in which a 1940 household can be linked to ICR households. (*b*) The top-two panels depict linkages that are logically consistent, firstly for nuclear households with a head and a spouse, and secondly for single-head households. The bottom panel depicts linkages that appear logically inconsistent and should be checked before using these observations. (*c*) In the top panel: Column 1 indicates that both head and spouse in a 1940 FCC household are linked to a head and spouse who are in the same household in the ICR. Column 2 indicates that both 1940 head and spouse are linked to persons in the same household, and that neither of these is head or spouse in the ICR. This captures the relatively common occurrence of young couples living with one spouse's parents. Column 3 indicates that 1940 head and the spouse are linked to persons in the ICR who live in different households and are neither head nor spouse in the ICR. This captures the most common occurrence of new households being formed by young adults. Columns 4–5 capture the case where only one of 1940 head or spouse is linked to a person record in the ICR, and that record is not a head or spouse in the ICR. Column 5 is common because young wives in 1940 who were living with their parents in the ICR are less likely to be linked because of their different last names. The middle panel depicts single-head households in which the head is linked to person-record in the ICR. Such links are consistent if the ICR person-record was also a single head or was neither head nor spouse (e.g. an adolescent living with parents); they are also consistent if the ICR person-record was part of a nuclear head-spouse pair, but the 1940 person-record is widowed or divorced. In the bottom panel: Columns 1–2 indicate that both head and spouse in a 1940 FCC household are linked to person-records in the same ICR-household but that these person-records had a relation to each other that is inconsistent with their 1940 relation. Column 3 indicates that 1940 head and spouse are both linked to person-records of the same relation but in different households. Columns 4–5 indicate that only of head or spouse are linked to person-records in the ICR, but that they are already a head/spouse.

consistent if the ICR person-record was also a single head or was neither head nor spouse (e.g. an adolescent living with parents); they are also consistent if the ICR person-record was part of a nuclear head-spouse pair, but the 1940 person-record is widowed or divorced. The bottom panel depicts linkages that appear logically inconsistent. In the bottom panel, columns 1–2 indicate that both head and spouse in a 1940 FCC household are linked to person-records in the same ICR-household but that these person-records had a relation to each other that is inconsistent with their 1940 relation. Column 3 indicates that 1940 head and spouse are both linked to person-records of the same relation but in different households. Columns 4–5 indicate that only of head or spouse are linked to person-records in the ICR, but that they are already a head/spouse.

Links depicted in the bottom panel are "flagged" for potential logical inconsistency, and should be omitted from the statistical analysis until they have been manually checked and linked (or de-linked).

5 Identification Strategy

Our identification strategy is anchored on the exogenous rotation of Indian Agents across reservations over time, and their varying propensity to transfer land from trust status into fee simple. To operationalize idea of using this varying propensity, we constructed a panel of the Indian Agents in charge on each reservation in each year reservation-year from the sources described in Section 3.4. We merged this dataset onto a complete panel of all allotments and transfers into fee simple on reservations between the Dawes Act of 1887, and the 1939, the year before the outcomes we observe in the Full-Count Census.

The land transfer data are set up as a duration-analysis dataset: if an allotment i was not transferred into fee simple in year t, then it may be transferred in $t+1$. The key identifying variation in our framework comes from the fixed effect $\mu_{j(rt)}$ of agent j in charge of reservation r in year t.[31] In addition, when allotment i was transferred into fee simple also depended on a number of other factors. On some reservations, the process of land transfer may have been faster so that we include a reservation fixed effect μ_r. At certain times, the process of land transfer was faster faster so that we include a year fixed effect μ_t. Most allotments were held in trust for some

[31] We estimate one fixed effect $\widehat{\mu_{j(\cdot)}}$ per agent j; the notation $j(rt)$ only serves to clarify that agents rotate across reservations over time.

time before they became eligible for transfer so that control for $(t - \tau_i)$, the number of years that had passed since i's initial issuance in year τ_i. Whether a given allotment was transferred into fee simple may have also depended on a reservation's stock of untransferred trust allotments, which we call backlog$_{rt}$. Defining a dummy $D_{i(r)t}$ that tales value 1 if allotment i tied to reservation r was transferred into fee simple in year t, and 0, we run the following duration-style regression

$$\begin{aligned} D_{i(r)t} &= \mu_{j(rt)} + \mu_r + \mu_t \\ &\quad + \beta_\tau \cdot (t - \tau_i) + \beta_b \cdot \text{backlog}_{rt} + \epsilon_{i(r)t}. \end{aligned} \tag{1}$$

We expect $\beta_\tau > 0$, and $\beta_b > 0$. With the estimated coefficients $\{\widehat{\beta_\tau}, \widehat{\beta_b}\}$ and vectors of fixed effects $\{\widehat{\mu_{j(\cdot)}}, \widehat{\mu_r}, \widehat{\mu_t}\}$, we compute an estimated probability of transfer into fee simple $\mathbb{P}(\widehat{D_{i(r)t} = 1})$ for each allotment i in each year t. The key exogenous component in equation (1) are the agent fixed effects $\widehat{\mu_{j(rt)}}$. Other determinants of land transfer in equation (1) will be controlled for in the second stage by conditioning on reservation fixed effects and on the year τ a patent was originally issued.

To turn the results of the estimations outlined here into a cross-sectional instrument Z_i for the final (second-stage) analysis, we calculate the cumulative probability that an allotment was transferred into fee simple between its issuance in year τ and the year 1934 (when the process was terminated), as

$$\begin{aligned} Z_i &= \mathbb{P}(\widehat{D_{i(r),t=1940} = 1}) \\ &= \mathbb{P}(\widehat{D_{i(r),t=\tau} = 1}) \\ &\quad + [1 - \mathbb{P}(\widehat{D_{i(r),t=\tau} = 1})] \cdot \mathbb{P}(\widehat{D_{i(r),t=\tau+1} = 1}) \\ &\quad + [1 - \mathbb{P}(\widehat{D_{i(r),t=\tau+1} = 1})] \cdot \mathbb{P}(\widehat{D_{i(r),t=\tau+2} = 1}) + ... \end{aligned} \tag{2}$$

We implement our IV procedure using standard 2SLS. When estimating 2SLS using a generated regressor like Z_i, under very weak assumptions, the point estimates are consistent and the standard errors and test statistics asymptotically valid. See Pagan (1984) and Wooldridge (2010, pp116–117).

Our identification strategy is akin to the strategies used in the 'judge fixed effect' literature.[32]

[32] See, for example, Kling (2006); Di Tella and Schargrodsky (2013); Galasso and Schankerman (2014); Aizer and Doyle Jr (2015); Melero et al. (2017); Dobbie et al. (2018); Frandsen et al. (2019). Our setup departs from the standard

There are three elements that need to be in place for these strategies: one, the setting needs to allow sufficient discretion that the identity of the decision-maker matters. Two, there needs to be enough statistical power to estimate individual judges' probability of making a certain decision (like a yes/no vote or a guilty/non-guilty verdict). The third element is that the assignment of judges to cases should not be endogenous to the outcome under study. In our setting, this means Indian Agents with a higher propensity to transfer land into fee simple should not be assigned to reservations that are independently more likely to have their land transferred into fee simple. We argue that all three elements are in place in our setting.

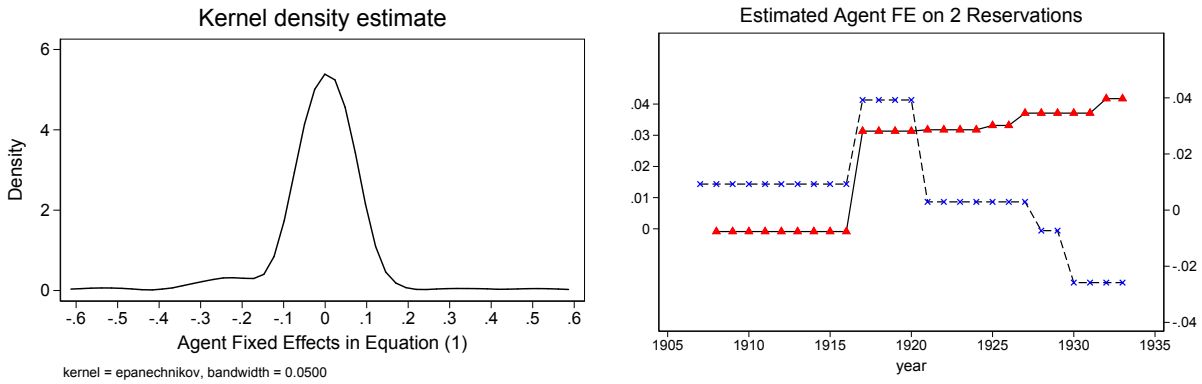
First, the historical and institutional narrative makes it clear that the BIA agents possessed considerable discretionary room over the assignment of allotments (Banner, 2009; Otis, 2014; Carlson, 1981). Second, we are able to estimate an individual BIA agent's probability of transferring an allotment into fee simple because there was frequent rotation of BIA agents across reservations, so that we can separate the agent fixed effect from reservation-traits, and because we have a long panel of annual transfer-decisions that covers all allotted reservations and that spans the entire allotment period. The left panel of Figure 2 shows the distribution of the roughly 450 agent fixed effects $\widehat{\mu_{j(\cdot)}}$ estimated in equation (1). The right panel of Figure 2 shows how the rotation of agents over time induces different time-paths in the propensity to transfer land into fee simple on two different reservations. In the initial years after the Burke Act, Salt River had an Indian Agent whose propensity to transfer land was about average, with a $\widehat{\mu_{j(\cdot)}} \approx 0$ (Charles E. Coe From 1906–1917); but from 1917 until the end of the allotment era in 1934, Spirit Lake, had a series of agents who all had higher than average propensities to transfer land into fee simple (Byron A. Sharp, 1917–1921, Frank A. Virtue, 1921–1925, Charles S. Young 1925–1927, John B. Brown 1927–1932, Arthur J. Wheeler, from 1932). Salt River, by contrast, had agents with a higher propensity to transfer land to fee simple in the early years (Charles M. Ziebach 1906–1917, Samuel A. M. Young, 1917–1921), but then had a succession of three agents with a lower propensity towards the end of the allotment process (William R. Beyer 1921–1928, John S. R. Hammitt 1928–1930, and Orrin C. Gray 1930–1934).

The key aspect is the timing in the rotation. In the early years, most allotments were relatively

'judge fixed effect' setup in that our setup is naturally estimated as a duration analysis because the decision to transfer land from trust status to fee simple was taken *repeatedly*.

newly issued and therefore ineligible for transfer (captured by $\beta_\tau \cdot (t - \tau)$ in equation (1)). Even agents with a high probability to transfer land into fee simple would typically not do it when the allotment was very new. Having agents with a high $\widehat{\mu_{j(\cdot)}}$ in the early year therefore did not make a big difference. In the later years however, most allotments were eligible and the process was primarily determined by the transfer-propensity of agents. In those later years, Salt River's agents had a six percent lower probability to transfer land than Spirit Lake's (e.g. $-0.02 - 0.04$ in the right panel of Figure 2), conditional on reservation and year fixed effects as well as when an allotment was issued. As a result, almost three-quarters of the allotments on Spirit Lake were eventually transferred into fee simple, but less than ten percent were on Salt River.

Figure 2: Distribution of Estimated $\widehat{\mu_{j(\cdot)}}$



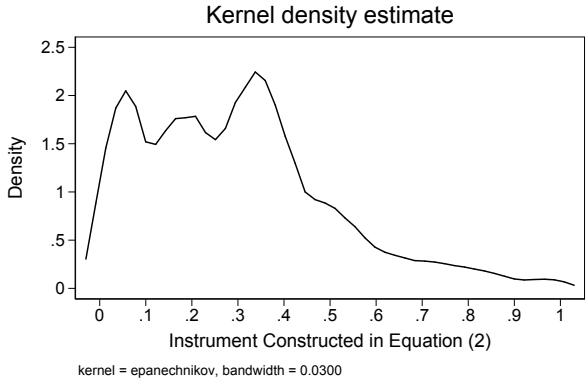
Notes: The left panel of this figure shows the distribution of roughly 450 agent fixed effects $\widehat{\mu_{j(\cdot)}}$ estimated in equation (1). The right panel shows how the rotation of agents over time induces different time-paths in $\widehat{\mu_{j(rt)}}$, i.e. the propensity to transfer land into fee simple, on Spirit Lake (red triangles, solid line) and on Salt River (blue crosses, dashed line)

The key to the instrument we construct in equation (2) is the interaction between a reservation's rotation-induced variation in $\widehat{\mu_{j(rt)}}$ over time, and an allotment's initial issuance year, which creates variation within reservations in whether an allotment was eligible for transfer into fee simple under an agent with a high or a low propensity for land-transferring. Figure 3 shows the distribution of Z_i, where i indexes an allotment and we have just under 110,000 allotments in the data. Given the logic of this identification strategy, it is important to very that Z_i is not simply explained by reservation fixed effects or by the allotments' year of initial issuance. The regression

$$Z_{i(r)} = \mu_r + \beta_\tau \cdot \tau_{i(r)} + \epsilon_{i(r)t} \tag{3}$$

tests this. Results are reported in the right panel of Figure 3, with standard errors clustered at the reservation level. The R-squared of this regression is 0.55; the estimated $\widehat{\beta_\tau}$ suggest that an allotment that was issued 10 years earlier was two percent more likely to be transferred into fee simple eventually, but the effect is not significant, and the F-test for the joint insignificance of all reservation fixed effects is not rejected, with a p-value of 0.36. In combination, this evidence shows that $Z_{i(r)}$ display considerable within-reservation variation, and that this within-reservation variation cannot be explained simply by earlier allotments being more likely to be transferred into fee simple.

Figure 3: Distribution of Instrument Z_i



regression, with outcome Z_i

year(allotment issued)	-0.002
	(0.002)
reservation fixed effects	Y
Observations	108,031
R-squared	0.553
F-stat(rez fixed effects)	1.26

Notes: The left panel of this figure shows the distribution of the over 100,000 allotment-specific instruments Z_i constructed in equation (2). The right panel reports on the results of estimating equation (3).

These first two elements are necessary to be able to separate the agent fixed effect from reservation-traits, but they are not sufficient for statistical identification of $\widehat{\mu_{j(\cdot)}}$ in equation (1). For this, we also require that the assignment of a BIA agent to a reservation was conducted in a manner that was exogenous to the allottees' whose allotments were considered for transfer into fee simple on that reservation. From the perspective of timing, one may worry that BIA agents were rotated only a reservation when the sale of surplus reservation lands was deemed most important by the government. However, the historical record is that rotation was anchored on the federal administration cycle: the majority of BIA agents were rotated with every when a new administration came in at the federal level every four or every eight years.[33]

From the perspective of selection, the ideal institutional setting would be one where BIA agents

[33]On average, BIA agents managed a single reservation for approximately eight years, with the average career length lasting twelve years.

were rotated across reservations via a lottery. Unfortunately, the BIA did not assign agents to reservations via a lottery. One may therefore worry that the BIA allocated agents with a higher proclivity for transferring land into fee simple to reservations with certain characteristics, particularly over land. However, the historical record again suggests that this was not the case. The primary job of BIA agents was to foster education and public health on the reservations, and we argue that any selection on these characteristics would have been orthogonal to the process of allotments. We can statistically test this argument to an extent, based on the idea that if agents were chosen for the purpose of land transfer, then one might expect agent pay to correlate with $\widehat{\mu_{j(\cdot)}}$.

We collected agent salary information from the Official Registers for every agent and year from 1879 to 1940. Average agent salaries were approximately $44,000 in 2018 dollars.[34] To quantify the relationship between agents' pay and the agents' estimated fixed effects, we estimate regression.

$$AgentPay_{jrt} = \mu_r + \delta_t + \beta \cdot \widehat{\mu_{j(\cdot)}} + \epsilon_{jrt}, \tag{4}$$

where $AgentPay_{jrt}$, was collected for each agent, j, located at reservation, r, in year t. Our main coefficient of interest, β, indicates whether or not agents with a higher propensity to transfer land were compensated more. We condition this specification on reservation and year fixed effects and cluster our standard errors at the reservation level. Column (1) of Table 5 reports the results of estimating equation (4). The results indicate that the agent fixed-effect is not significantly correlated with the agent salaries, which we view as evidence against selection of agents on their allotting propensity.

We are also interested in whether agents with a higher propensity to transfer land transferred lower quality land on average. To ask this, we constructed a weighted average of the land quality of reservations that agent j was ever on (weighted by number of years they were on), i.e. $TotalLandQuality_j$. We also construct the average quality of land allotted out of this pool under agent j, i.e. $TrustLandQuality_j$.[35] We then ask whether $\widehat{\mu_{j(\cdot)}}$ correlated with $TrustLandQuality_j$,

[34] This is similar to a current federal employee paid at the General Schedule 8 grade.

[35] We quantify land quality we use the FAO land suitability measure for rain fed wheat. We measure the land quality an agent faced by calculating the weighted average suitability index they faced over their career living at reservations r during years t.

Table 5: Relating Estimated BIA Agent Fixed Effects to Salaries and Land Suitability

	(1) Ln(Agent Salary)	(2) Ln(Trust Land Quality)
Agent Fixed Effect	0.094	0.061
	[0.244]	[0.700]
Ln(Total Land Quality)		1.200***
		[0.000]
reservation fixed effect	Yes	
year fixed effects	Yes	
Observations	8,255	426
R-squared	0.576	0.762

Notes: In this table, we relate the estimated BIA agent fixed effects to agent salaries as well as to the quality of trust land the agent faced during their career. Column (1) reports the results of estimating equation (4). Column (2) reports the results of estimating equation (5). In square brackets are the p-values for the standard errors clustered on the reservation in column (1), and for robust standard errors in column (2); *** p<0.01, ** p<0.05, * p<0.1

conditional on the land quality of the available land pool:

$$TrustLandQuality_j = \beta \cdot \widehat{\mu_{j(\cdot)}} + \gamma \cdot TotalLandQuality_j + \epsilon_j \tag{5}$$

Column (2) of Table 5 reports the results of estimating equation (5). There is no evidence that higher land transfer propensity correlates with the quality of allotted land, relative to what land was available.

References

Abramitzky, R., L. Boustan, and K. Eriksson (2016). To the new world and back again: Return migrants in the age of mass migration. *ILR Review*, 0019793917726981.

Abramitzky, R., L. P. Boustan, and K. Eriksson (2012). Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration. *American Economic Review 102*(5), 1832–56.

Abramitzky, R., L. P. Boustan, K. Eriksson, J. J. Feigenbaum, and S. Pérez (2019). Automated linking of historical data. Technical report, National Bureau of Economic Research.

Adams, D. W. (1995). *Education for Extinction: American Indians and the Boarding School Experience, 1875-1928*. U of Kansas Press.

Aizer, A. and J. J. Doyle Jr (2015). Juvenile incarceration, human capital, and future crime: Evidence from randomly assigned judges. *The Quarterly Journal of Economics 130*(2), 759–803.

Aizer, A., S. Eli, J. Ferrie, and A. Lleras-Muney (2016). The long-run impact of cash transfers to poor families. *American Economic Review 106*(4), 935–71.

Akee, R., W. Copeland, E. J. Costello, and E. Simeonova (2018). How does household income affect child personality traits and behaviors? *American Economic Review 108*(3), 775–827.

Akee, R. K., W. E. Copeland, G. Keeler, A. Angold, and E. J. Costello (2010). Parents' incomes and children's outcomes: a quasi-experiment using transfer payments from casino profits. *American Economic Journal: Applied Economics 2*(1), 86–115.

Algan, Y. and P. Cahuc (2010). Inherited trust and growth. *American Economic Review 100*(5), 2060–92.

Anderson, T. L. and D. P. Parker (2008). Sovereignty, credible commitments, and economic prosperity on american indian reservations. *The Journal of Law and Economics 51*(4), 641–666.

Bailey, M., C. Cole, M. Henderson, and C. Massey (2017). How well do automated linking methods perform in historical samples? evidence from new ground truth. *Unpublished manuscript*.

Bandiera, O., M. Mohnen, I. Rasul, and M. Viarengo (2018). Nation-building through compulsory schooling during the age of mass migration. *The Economic Journal 129*(617), 62–109.

Banner, S. (2009). *How the Indians lost their land: Law and power on the frontier*. Harvard University Press.

Barrera-Osorio, F., M. Bertrand, L. L. Linden, and F. Perez-Calle (2011). Improving the design of conditional transfer programs: Evidence from a randomized education experiment in colombia. *American Economic Journal: Applied Economics 3*(2), 167–95.

Barrera-Osorio, F., L. L. Linden, and J. E. Saavedra (2019). Medium-and long-term educational consequences of alternative conditional cash transfer designs: Experimental evidence from colombia. *American Economic Journal: Applied Economics 11*(3), 54–91.

Becker, A., B. Enke, and A. Falk (2018). Ancient origins of the global variation in economic preferences. Technical report, National Bureau of Economic Research.

Behrman, J. R., S. W. Parker, and P. E. Todd (2011). Do conditional cash transfers for schooling generate lasting benefits? a five-year followup of progresa/oportunidades. *Journal of Human Resources 46*(1), 93–122.

Bleakley, H. and J. Ferrie (2016). Shocking behavior: Random wealth in antebellum georgia and human capital across generations. *The quarterly journal of economics 131*(3), 1455–1495.

Carlson, L. A. (1981). *Indians, bureaucrats, and land: the Dawes Act and the decline of Indian farming*. Number 36. Praeger Pub Text.

Cesarini, D., E. Lindqvist, R. Östling, and B. Wallace (2016). Wealth, health, and child development: Evidence from administrative data on swedish lottery players. *The Quarterly Journal of Economics 131*(2), 687–738.

Costa-Font, J., P. Giuliano, and B. Ozcan (2018). The cultural origin of saving behavior. *PloS one 13*(9), e0202290.

Demsetz, H. (1967). Toward a theory of property rights. *American Economic Review 62*, 347–359.

Di Tella, R., S. Galiant, and E. Schargrodsky (2007). The formation of beliefs: evidence from the allocation of land titles to squatters. *The Quarterly Journal of Economics 122*(1), 209–241.

Di Tella, R. and E. Schargrodsky (2013). Criminal recidivism after prison and electronic monitoring. *Journal of Political Economy 121*(1), 28–73.

Dippel, C. (2014). Forced coexistence and economic development: evidence from native american reservations. *Econometrica 82*(6), 2131–2165.

Dippel, C., D. Frye, and B. Leonard (2019). The costs of tenancy in common: Evidence from indian land allotment. Technical report.

Dobbie, W., J. Goldin, and C. S. Yang (2018). The effects of pretrial detention on conviction, future crime, and employment: Evidence from randomly assigned judges. *American Economic Review 108*(2), 201–40.

Dohmen, T., B. Enke, A. Falk, D. Huffman, and U. Sunde (2018). Patience and comparative development. Technical report, Working Paper.

Enke, B. (2019). Kinship, cooperation, and the evolution of moral systems. *The Quarterly Journal of Economics 134*(2), 953–1019.

Feigenbaum, J. J. (2016). A machine learning approach to census record linking. Technical report, Working Paper.

Feir, D., R. Gillezeau, and M. Jones (2019). The slaughter of the bison and reversal of fortunes on the great plains. Technical report.

Feir, D. L. (2016a). The intergenerational effects of residential schools on childrens educational experiences in ontario and canadas western provinces. *The International Indigenous Policy Journal 7*(3), 5.

Feir, D. L. (2016b). The long-term effects of forcible assimilation policy: The case of indian boarding schools. *Canadian Journal of Economics 49*(2), 433–480.

Fernández, R. (2011). Does culture matter? In *Handbook of social economics*, Volume 1, pp. 481–510. Elsevier.

Ferrie, J. P. (1996). A new sample of males linked from the public use microdata sample of the 1850 us federal census of population to the 1860 us federal census manuscript schedules. *Historical Methods: A Journal of Quantitative and Interdisciplinary History 29*(4), 141–156.

Fouka, V. (2019). Backlash: the unintended effects of language prohibition in us schools after world war i. Technical report.

Frandsen, B. R., L. J. Lefgren, and E. C. Leslie (2019). Judging judge fixed effects. Technical report, National Bureau of Economic Research.

Galasso, A. and M. Schankerman (2014). Patents and cumulative innovation: Causal evidence from the courts. *The Quarterly Journal of Economics 130*(1), 317–369.

Galor, O. and Ö. Özak (2016). The agricultural origins of time preference. *American Economic Review 106*(10), 3064–3103.

Giuliano, P. and N. Nunn (2017). Understanding cultural persistence and change. Technical report, National Bureau of Economic Research.

Golenko, A. (2010). Did allotment contribute to the diminishment of american indian incomes? a study of allotment policies and american indian incomes during 1912-1920.

Gregg, M. T. (2018). The long-term effects of american indian boarding schools. *Journal of Development Economics 130*, 17–32.

Haushofer, J. and J. Shapiro (2016). The short-term impact of unconditional cash transfers to the poor: experimental evidence from kenya. *The Quarterly Journal of Economics 131*(4), 1973–2042.

Hoxie, F. E. (2001). *A final promise: The campaign to assimilate the Indians, 1880-1920*. U of Nebraska Press.

Jorgensen, M. (2007). *Rebuilding native nations: Strategies for governance and development*. University of Arizona Press.

Kling, J. R. (2006). Incarceration length, employment, and earnings. *American Economic Review 96*(3), 863–876.

Leonard, B. and D. Parker (2017). Creating anticommons: Historical land privatization and modern natural resource use.

Leonard, B., D. Parker, and T. Anderson (2018). Poverty from incomplete property rights: Evidence from american indian reservations.

Lomawaima, K. T. (1995). *They called it prairie light: The story of Chilocco Indian school*. U of Nebraska Press.

Melero, E., N. Palomeras, and D. Wehrheim (2017). The effect of patent protection on inventor mobility.

Meriam, L. (1928). *The problem of Indian administration: report of a survey made at the request of honorable Hubert Work, secretary of the interior, and submitted to him, February 21, 1928*. Number 17. Johns Hopkins Press.

Murdock, G. P. (1967). Ethnographic atlas.

Nunn, N. (2012). Culture and the historical process. *Economic History of Developing Regions 27*(sup1), S108–S126.

Office of Indian Affairs (1935). Indian land tenure, economic status, and population trends. In *Part X of the Report on Land Planning*. Washington: Department of Interior.

Ostrom, E. (1990). *Governing the commons*. Cambridge university press.

Otis, D. S. (2014). *The Dawes Act and the allotment of Indian lands*, Volume 123. University of Oklahoma Press.

Pagan, A. (1984). Econometric issues in the analysis of regressions with generated regressors. *International Economic Review*, 221–247.

Parker, S. W. and P. E. Todd (2017). Conditional cash transfers: The case of progresa/oportunidades. *Journal of Economic Literature 55*(3), 866–915.

Price, J., K. Buckles, I. Riley, and J. Van Leeuwen (2019, September). Combining family history and machine learning to link historical records. Technical report, NBER Working Paper No. 26227.

Regan, S. and T. L. Anderson (2016). Unlocking the energy wealth of indian nations. *Unlocking the Wealth of Indian Nations*, 107.

Sakalli, S. O. (2017). Secularization and religious backlash: evidence from turkey. Technical report, mimeo.

Schultz, T. P. (2004). School subsidies for the poor: evaluating the mexican progresa poverty program. *Journal of development Economics 74*(1), 199–250.

Spolaore, E. and R. Wacziarg (2013). How deep are the roots of economic development? *Journal of economic literature 51*(2), 325–69.

United States Government Printing Office (1879-1932). *Official Register of the United States, Executive, Legislative, Judicial*. Department of State.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*.

Online Appendix

to

"The Effect of Land Allotment on Native American Households During the Assimilation Era"

Figure Online Appendix Figure 1: 1910 Advertisement for Reservation Lands Left from Allotment



Notes:

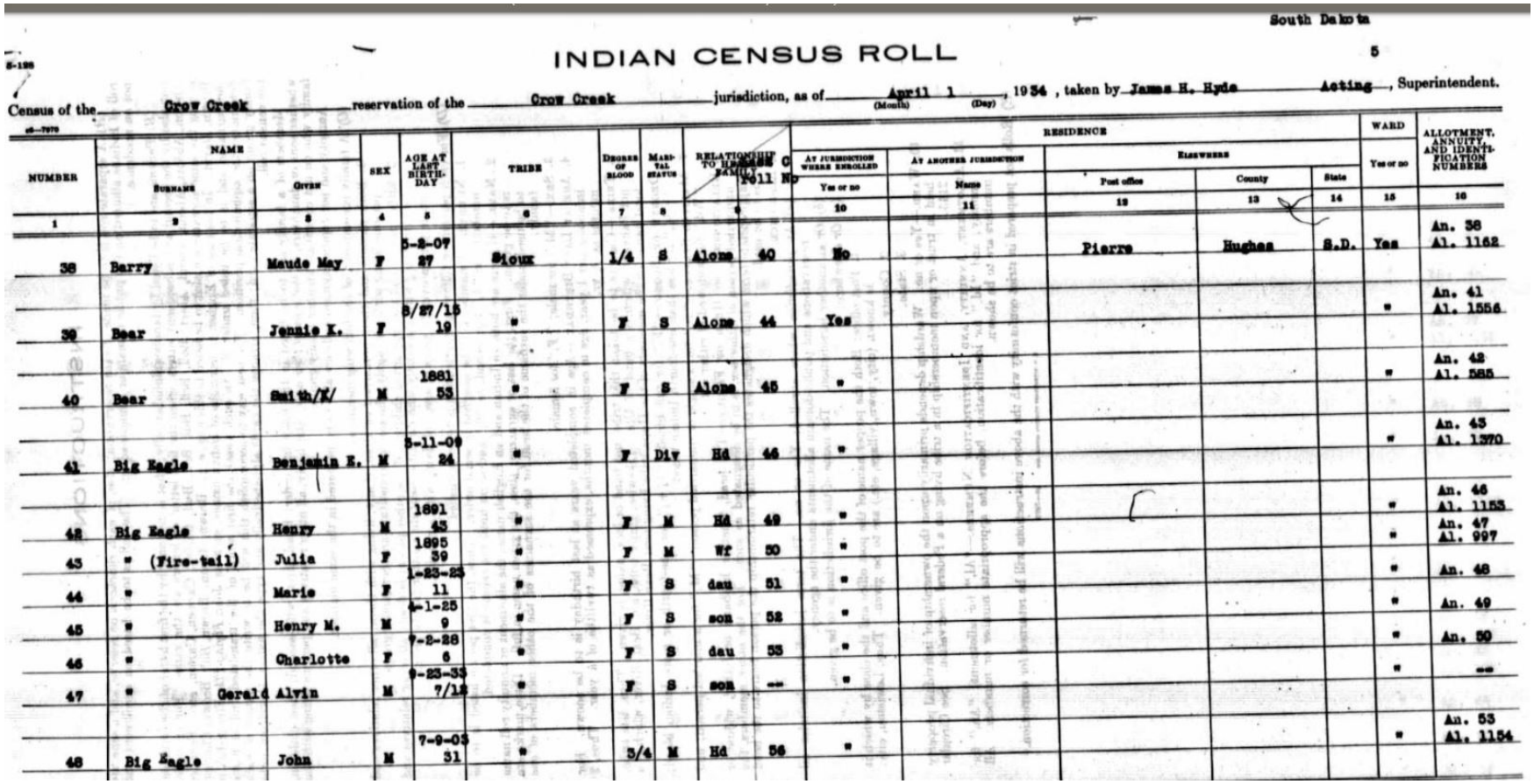
Online Appendix A Online Data Appendix

Online Appendix A.0.1 Linkage

Older record linkage methods used a smaller number of variables for matching, and often used name only, often focusing on samples of people with unusual names in order to reduce false positive matches , e.g. Ferrie (1996). Matching on names is almost always fuzzy matching, i.e. a matching algorithm that allows for typos and mis-spellings. A common approach involves splitting first and last names into substrings ('bigrams'), and to construct a similarity index over all bigrams. A commonly used similarity index is to calculate the 'Jaro Winkler index' between two names. Newer iterations of fuzzy matching have increased the flexibility to include matching on a set of numeric as well as string variables, including distance calipers on numeric variables, e.g. giving a higher match probability when two records' birth-years are one year rather than three years apart. See Abramitzky, Boustan, and Eriksson (2012, 2016) for more recent applications.

More recently, the emergence of machine learning algorithm has given a real boost to the precision of record linkage methods, as it allows for training an algorithm. See for example the method

Figure Online Appendix Figure 2: Sample Page of the Indian Census Rolls



South Dakota

5

INDIAN CENSUS ROLL

Census of the Crow Creek reservation of the Crow Creek jurisdiction, as of April 1 (Month) (Day), 1934, taken by James H. Hyde Acting, Superintendent.

NUMBER	NAME: Surname	NAME: Given	SEX	AGE AT LAST BIRTHDAY	TRIBE	DEGREE OF BLOOD	MARITAL STATUS	RELATIONSHIP TO HEAD OF FAMILY	Roll No.	RESIDENCE: At jurisdiction where enrolled (Yes or no)	RESIDENCE: At another jurisdiction (Name)	RESIDENCE: Elsewhere: Post office	RESIDENCE: Elsewhere: County	RESIDENCE: Elsewhere: State	WARD (Yes or no)	ALLOTMENT, ANNUITY AND IDENTIFICATION NUMBERS
1	2	3	4	5	6	7	8	9		10	11	12	13	14	15	16
38	Barry	Maude May	F	5-8-07 37	Sioux	1/4	S	Alone	40	No		Pierre	Hughes	S.D.	Yes	An. 38 Al. 1162
39	Bear	Jennie K.	F	8/27/13 19	"	F	S	Alone	44	Yes					"	An. 41 Al. 1556
40	Bear	Smith/K/	M	1881 53	"	F	S	Alone	45	"					"	An. 42 Al. 585
41	Big Eagle	Benjamin E.	M	3-11-09 24	"	F	Div	Hd	46	"					"	An. 43 Al. 1370
42	Big Eagle	Henry	M	1891 43	"	F	M	Hd	49	"					"	An. 46 Al. 1153
43	" (Fire-tail)	Julia	F	1895 39	"	F	M	Wf	50	"					"	An. 47 Al. 997
44	"	Marie	F	1-23-23 11	"	F	S	dau	51	"					"	An. 48
45	"	Henry M.	M	4-1-25 9	"	F	S	son	52	"					"	An. 49
46	"	Charlotte	F	7-2-28 6	"	F	S	dau	53	"					"	An. 50
47	" Gerald	Alvin	M	9-23-33 7/12	"	F	S	son	---	"					"	--
48	Big Eagle	John	M	7-9-03 31	"	3/4	M	Hd	56	"					"	An. 53 Al. 1154

Notes:

Online Appendix – Not for Publication

Table Online Appendix Table 1: Ethnographic Atlas Variables 'Property Rights' and 'Bride Price'

D(Property Rights)		EA: 76. Inheritance Rule for Movable Property	D(Bride Price)		EA: 6. Mode of Marriage (Primary)
0	1	Absence of individual property rights or rules	1	1	Bride Price or wealth, to bride's family
1	2	Matrilineal (sister's sons)	1	2	Bride Service, to bride's family
1	3	Other matrilineal heirs (e.g., younger brothers)	1	3	Token bride price
1	4	Children, with daughters receiving less	0	4	Gift Exchange, reciprocal
1	5	Children, equally for both sexes	0	5	Sister or female relative exchanged for bride
1	6	Other patrilineal heirs (e.g., younger brothers)	0	6	Absence of consideration
1	7	Patrilineal (sons)	0	7	Dowry, to bride from her family

Notes: TBA

outlined in Feigenbaum (2016). There is an active and ongoing debate on the choice of methods. Bailey et al. (2017) review several of these methods and show that all algorithm may produce samples that are not fully representative of the underlying population. This includes linking records by hand, although this method is favored by Bailey et al. (2017). By contrast, Abramitzky et al. (2019) show that a range of automated linkage methods on a range of standard linkage sample perform as well as manual linkage can be expected to.[36]

[36] For another summary, Ran Abramitzky's website at https://people.stanford.edu/ranabr/matching-codes.