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

Christian Dippel† Dustin Frye‡

June 10, 2019

PRELIMINARY DRAFT FOR CONFERENCE SUBMISSION

Abstract

Between 1906 and 1924, the U.S. government broke up millions of acres of communally owned reservation lands under the authority of the Dawes and Burke Acts, and allotted them to individual Native American households. We exploit quasi-random variation in the legal title with which land was granted to individual households: ‘fee simple’ gave full ownership rights while ‘in trust’ status did not. We link land allotment information to the universe of Native Americans in the 1940 U.S. Census, and find that households with full property rights over their allotted land had higher wealth, income, and occupational status, and their children had higher schooling attainment. We find that these differences were more pronounced within reservations whose ancestral tribal norms emphasized private property.

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

We thank Terry Anderson, Donna Feir, Miriam Jorgensen, Bryan Leonard, Dominick Parker, Matthew Snipp, Sheldon Spotted Elk, and seminar participants at the Hoover Institute’s Research Workshop on Renewing Indigenous Economies for helpful conversations. We thank Zach Lewis and Gwyneth Teo for valuable research assistance.

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

Native American reservations were formed in the U.S. in the second half of the 19th century. By the beginning of the 20th century, with the end of the Indian Wars and the closing of the frontier, the U.S. government turned its attention towards the assimilation of the Native Americans now living on reservations, and paving a path towards their citizenship. ‘Indian allotment’, which broke reservation lands into individually owned land allotments, was a major component of this strategy. Indian allotment had theoretically started with the passing of the Dawes Act in 1887. De facto, however, it only began in earnest with the Burke Act in 1906, which explicitly linked landownership to citizenship. All Indian allotments were first placed in a trust managed by the Bureau of Indian Affairs (BIA) local superintendents (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, allottees who were declared “competent” were eligible to convert their trust land to fee-simple, and coupled with this at the same time become citizens (Carlson, 1981; Banner, 2009; Otis, 2014). In 1924, at the height of the process of allotment, the Indian Citizenship Act (ICA) extended citizenship to all Native Americans, which abruptly throttled the process of allotment and transfers into fee simple. In 1934, allotment ended entirely with the passing of the Indian Reorganization (or ‘Howard-Wheeler’) Act (IRA). The aggregate flow of allotments and transfers is depicted in Figure 1. This paper studies the consequences for Native American households of this policy.

We link the universe of Indian land allotments to the universe of Native American households in the 1940 Full Count Census. Almost all Native Americans in 1940 lived on reservations or former reservation lands, and the focus of our analysis will be to compare households whose allotted land was transferred into fee simple to households on the same reservation whose allotted land remained in trust. Our analysis will not estimate the aggregate effect of Indian allotment on the allotted reservations. It is difficult to generate such an estimate because only about half of all reservations had been allotted when the policy ended,1 and whether reservations were ever allotted was endogenous: Leonard, Parker, and Anderson (2018) show that reservations with better land were more likely to be allotted, ostensibly because of the second motive of the Burke

---

1 On allotted reservations every household received an allotment.
Act (possibly the primary motive for some legislators) to free up the surplus tribal land left after allotment for white settlement.\footnote{As a colonial policy that was imposed on tribes by Congress, Indian allotment is viewed very negatively by many Native Americans to the present day. Although sales of un-allotted land in the end remained small in overall magnitude, they contributed considerably to this negative view, as Online Appendix Figure 1 suggests it would.}

Figure 1: Flow of Allottments and Transfers into Fee Simple

Notes: This figure tracks the flow of total acres that were allotted and the flow of acres subsequently transferred into fee simple in the BLM data.

Statistical identification for within-reservation variation in individual households’ property rights over allotted land comes from three sources. One source is time-variation in initial allotment across reservations: reservations had to be surveyed and parceled in order to be allotted, and the BIA’s limited resources for this introduced a strongly sequential pattern into allotment. The 1924 ICA and 1934 IRA meant not only that half of all reservations were never allotted, but also that reservations who were allotted later had less land transferred into fee simple because the required ‘trust-period’ extended beyond 1934. Importantly, we know for every land allotment when it was originally made, and when, if ever, it was transferred into ‘fee simple’. A second source comes from interacting this time variation with age variation in individual allottees: young allottees on a reservation that was allotted late had not come of age by the time the policy of Indian allotment ended. A third source of identification comes from the individual Indian Agents who were tasked by the BIA with governing the reservations: the Indian Agents were appointed to perform a range
of tasks on reservations, and we show that they varied considerably in their propensity to transfer land into fee simple, conditional on other determinants of being eligible for transfer. We construct a complete reservation-year panel of Indian Agents from 1887–1934 and use the agents’ exogenous rotation for an identification strategy that is akin to those used in the ‘judge fixed effect’ literature.3

Our analysis proceeds in two steps. First, we investigate the baseline effect of having one’s allotted land transferred into fee simple, comparing only allotted households on the same reservation. We find that households whose allotted land was transferred into fee simple had higher wealth, income, and occupational status by 1940, relative to households on the same reservation whose allotted land remained in trust. We also find that the children of households whose allotted land was transferred into fee simple had obtained higher levels of education by 1940, again relative to children of households on the same reservation whose allotted land had remained in trust. We think of these effects as ‘medium run’ effects in the sense that we observe households about 10-30 years after allotment, enough time to change occupation or to accumulate some wealth.

The second step of our analysis is to investigate the interaction between households’ land allotments and the ancestral cultural norms governing individual property rights in a tribe, which we glean from the *Ethnographic Atlas* (EA). Almost all reservations are uniquely associated with one ancestral tribe, and thus one set of ancestral tribal norms. We find that the differences in 1940 income and occupational status between households with fee simple and in-trust land allotments were more pronounced on reservations whose ancestral tribal norms allowed for private property in addition to communal property rights.4;5

One contribution of our paper is to lift the lid on the consequences of one of the largest policy interventions ever targeted at indigenous peoples.6 The Dawes and Burke Acts are much

---

3 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 dynamic duration setup, where a reservation’s time-varying backlog of untransferred trust allotments, an individual allotment’s year of initial issuance, and an individual allottee’s time-varying age all played a role in the decision.

4 It could reasonably be argued that these ancestral norms had developed partly as an adjustment to European settlement. For example, the classic treatment of property rights by Demsetz (1967) includes as a case study for the emergence of property rights exactly this argument, building on anthropological work Speck (1915); Leacock (1954).

5 We also find that the differences in children’s 1940 educational attainment across households were more pronounced on reservations whose ancestral tribal traditions included the norm of having a ‘bride price’, consistent with economic theory and with previous evidence (Ashraf, Bau, Nunn, and Voena, 2019).

6 We exclude warfare and genocide from the category of policy interventions.
maligned in Indian Country to the present day, partly because it was a colonial policy that was imposed on tribes, and partly because of the decision to lock individual land ‘in trust’ indefinitely with the 1934 IRA (but effectively already in 1924). Much of this land remains in trust to the present day, under an ill-conceived design whereby all descendants of the original allottees have claims on the land, and have to agree unanimously to any activity on the land, including activities such as leasing as well as the later transfer into fee-simple. Evidence is mounting on the long-run inefficiencies created by this arrangement and the opportunity costs for Native American tribes who had been subject to allotment (Leonard and Parker, 2017; Dippel, Frye, and Leonard, 2019). Our paper cannot answer the question whether Indian land allotment would have had aggregate long-run welfare benefits if all land had been transferred to households into fee simple.\(^7\) It does, however, show that even in the relative near term that we capture with the 1940 outcomes (where the problem of multiple heirs was not important), being given land with full property rights (i.e. in ‘fee simple’) was far better for Native American households than being given the same land without the ability to collateralize or sell it.

A second contribution of our paper is to the literature on culture and its impact on economic choices and outcomes. This growing literature emphasizes 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).\(^8\) 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.\(^9\) Giuliano and Nunn (2017) argue that societies that have historically lived in more stable condi-

---

\(^7\) 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, and 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, there are today many successful efforts by tribal governments to change institutions of tribal governance to promote economic development (Regan and Anderson, 2016; Jorgensen, 2007).

\(^8\) 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.

\(^9\) To some extent this is the flipside of Ostrom (1990) who argued for the importance of culture in making collective property management function without it deteriorate into the tragedy of the commons.
tions 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.

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 to Indian allotment. Di Tella, Galiani, 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.

---

10 Feir, Gillezeau, and Jones (2019) provide a particularly striking example of this, focusing on the bison cultures in the great plains, and the long run consequences for them of the sudden slaughter of the bison in the 1880s.
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). Congress experimented with allotment clauses in treaties governing individual reservations starting in the 1850s. These treaties formed the basis for the first general allotment acts, discussed by Congress and endorsed by the Office of Indian Affairs in the 1870s (Otis, 2014). Legislation stalled over the issues of citizenship, jurisdiction, and whether to immediately grant allotted Indians title to their land. In 1880, Senator Coke of Texas introduced a bill calling for the allotment of all reservation lands. This bill provided for the allotment of Indian lands following the designation of the president and a 2/3 majority vote in favor by the adult male members of the tribe Prucha (2014). The allocation of land varied by household type and age and allotments were inalienable for 25 years. The bill also allowed for the tribe to sell excess reservation lands to the government (Prucha, 2014). The bill found heavy support within Congress, the Office of Indian Affairs, and among reformers. The Coke bill passed in the Senate, but failed in the House due to opposition from western legislators that felt the bill was too generous to Indians (Carlson, 1981).

Henry Dawes introduced a modified allotment bill to the Senate in 1886. 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 reservation lands deemed appropriate (Banner, 2009). Heads of household received 160 acres, single persons over 18 received 80 acres, orphans under 18 received 80 acres, and other single individuals under 18 received 40 acres. 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.

Once selected, allotments were approved by the Secretary of Interior and each Indian was

---

11 By the late 1870s, nearly seventy treaties included clauses regarding the allotment of Indian lands (Prucha, 2014).

12 Tribes in New York and Indian Territory were temporarily exempted from the Dawes Act.
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. 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).13 Proceeds from the sales of the surplus land were held in trust and appropriated at the discretion of Congress for “education and civilization” (Banner, 2009). The Dawes Act was amended in 1891 to grant 80 acres to every adult, instead of 160 acres to heads of households. It also gave the Secretary of Interior the authority to establish leasing regulations for allotted lands. Leasing agreements were managed by Indian Agents and the leasing of allotted lands was widely adopted across reservations (Carlson, 1981). Gradually the Dawes Act was amended to relax constraints on sales of trust lands. In 1903, the Office of Indian Affairs was authorized to sell allotments of deceased allottees with multiple heirs, and in 1907, they were authorized to sell allotments of original allottees under special circumstances. 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 allotting 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 surveying and dividing the reservation into parcels. The agent possessed considerable authority over the assignment of allotments, even for those eligible Indians that desired an allotment. There are numerous accounts of outside settlers influencing agents to set aside the highest quality land for surplus (Banner, 2009; Otis, 2014; Carlson, 1981). In cases where the Indian did not select a plot, the agent assigned an allot-
ment. 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.

2.3 The End of the Allotment Era

The allotment process moved quickly. On the extensive margin, nearly four reservations per year were designated for allotment over the first two decades. On the intensive margin, the government averaged nearly 4,500 allotments per year over the same period (Otis, 2014). The rapid expansion of allotment and 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. These concerns culminated in a review of the current social and economic conditions on reservations by Lewis Meriam of the Institute of Governmental Research in 1928. 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 its trust status, creating a patchwork of land tenures within Indian reservations.

2.4 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). The policy resulted in a substantial transfer of land out of Native ownership. 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. Nearly 60 million acres were 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. Often parcels of differing tenure types are adjacent to one another, creating a checkerboard pattern on
reservations. Figure 2 illustrates this pattern on the Pine Ridge Reservation.

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.¹⁴

3.2 The Full-Count Population Census (Measuring Outcomes)

The ‘rule of 72’ dictates that non-anonymized individual-level Census information can be made available 72 years after a Census was published. The 1940 Full-Count US Census (FCC) includes 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 spotty, 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 the BLM data and the FCC.

3.3 The Indian Census Rolls (Measuring Treatment)

To link land allotment information to the FCC, we use the Indian Census Rolls (ICR). The ICR were special censuses that were unrelated to the FCC and were collected on reservations in the 1920s

¹⁴ 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).
Figure 2: 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. This image is taken from Dippel et al. (2019). 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.
and early 1930s. The ICC did not record any economic outcome data; instead they are simply complete enumerations of all Native American households. 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 ICC. These data match uniquely to the BLM data via allotment number. We then link the ICC to the FCC using record linkage methods described in Section 4.

3.4 The Indian Agents

Our identification strategy will combine information on when an allotment was issued (which we glean from the BLM data) with information on allottees, which we glean from the ICR data. Conditional on these pieces of information, we then use the exogenous variation generated by the rotation of Indian Agents across reservations, and their varying propensity to transfer land into fee simple to estimate an exogenous probability that a given allotment will be transferred into fee simple in any given year. To do 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). The Official Registers were published biennially from 1879–1940. 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 Indian Census Rolls described in section 3.2. We collected 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, and mar-
riage and domestic organization, and inheritance customs.

4 Linkage

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

Linking Native Americans across historical records is both easier and harder than linking other populations. On the one hand, the more unusual names that Native Americans typically have make it in principal easier to uniquely track individuals across historical records based on only name and a few more characteristics like age and birthplace. In addition, Native Americans in the 1920s to 1940s were not very mobile: the vast majority lived on or very near their tribe’s reservation in both the ICR and the FCC. On the other hand, the more unusual names of Native Americans meant Census enumerators made more frequent errors in recording names. In addition, there was a strong trend of anglicization of names during the period in which we link samples, so that both first and last name of the same person were more likely to be anglicized in the FCC than in the ICR. Ages may also have been less precisely recoded in the ICR than in the FCC.

A key feature of both the ICR and the FCC is that they are organized by household and contain information on household relations (e.g. spouse, son, etc). To improve the frequency and accuracy of record links, we expanded existing record linkage methods to take explicit account of household structure within our linkage algorithm.

- We first used the R’s RecordLinkage package to perform an initial linkage, blocking 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.) Similarity between individual records was based on bigram-indexation of first and last name (using the stringsim() function), and allowing an age difference of no more than two years.

- Rather than keeping only the best possible match (which is the default), we kept all potential matches that passed a certain threshold of similarity, which was decided based on the distribution of similarity scores across the dataset.
• Records from the using data (the FCC) that have only one match are placed directly into the final output.

• For records from the using data that are matched to several different individuals in the master data (the ICR), we compared matches by calculating the similarity between the record’s household in the using data, and the household of all potential records in the master data. Household characteristics include the size of the household and the gender distribution, but could be expanded (at a computational cost) to include the distribution of ages and relationships.

• We expanded this approach to account for cases where an individual match could only be true if the individual had changed households. For instance, an individual link between a ‘child’ in the ICR and a ‘head’ or ‘spouse’ in the FCC can only be true if the person who was a child in the 1920s had formed their own household by 1940. If this was true for the highest-score match, we placed this match directly into the final output and disregarded the others. If the highest highest-score match had a consistent household relation (e.g. ‘child’ in the master and using data), then we compared it only to potential matches that also had consistent household relations.

Table 1 shows an example of how this approach improves confidence in individual matches. Observations 7, 10, 11, and 12 are all examples of linkages that look very uncertain on their own, but look very certain in the context of the household-linkage we perform after the individual linkage. (As well, observation 2 constitutes an example of an individual who grew up and formed her own household between the 1920s ICR and the 1940 FCC.)
Table 1: Household-Information-Augmented Record Linkage Example

| HH-id | birthyear | namelast | namefrst | relate | sex   | HH-id | birthyear | namelast | namefrst | relate | sex   |
|-------|-----------|----------|----------|--------|-------|-------|-----------|----------|----------|--------|-------|-------|
| 1     | 79206     | 1887     | MART     | FRED   | Head  | 35545693| 1889     | MART     | FRED    | Head  |
| 2     | 79206     | 1920     | MART     | ROSE P | Daughter| 355454873| 1920     | MART     | ROSE   | Spouse|
| 3     | 79206     | 1902     | MART     | NELLIE | Wife  | 35545693| 1901     | MART     | NELLIE | Spouse|
| 4     | 79206     | 1922     | MART     | LILA   | Daughter| 35545693| 1926     | MART     | LAURA  | Child |
| 5     | 79206     | 1924     | MART     | VIOLET | Daughter| 35545693| 1924     | MART     | VIOLET | Child |
| 6     | 79206     | 1926     | MART     | LAURA  | Daughter| 35545693| 1926     | MART     | LAURA  | Child |
| 7     | 79206     | 1930     | MART     | FELIX ALTON | Son | 35545693| 1930     | MART     | TELEY  | Child |
| 8     | 79206     | 1932     | MART     | AL SMITH | Son | 35545693| 1932     | MART     | AL SMITH | Child |

| HH-id | birthyear | namelast | namefrst | relate | sex   | HH-id | birthyear | namelast | namefrst | relate | sex   |
|-------|-----------|----------|----------|--------|-------|-------|-----------|----------|----------|--------|-------|-------|
| 9     | 79199     | 1929     | LUCERO   | FRED   | Son   | 35545587| 1930     | LUCERO   | FRED   | Child |
| 10    | 79199     | 1911     | LUCERO   | ROSELYA | Wife | 35545587| 1911     | LUCERO   | ROSELYA| Spouse|
| 11    | 79199     | 1936     | LUCERO   | PETE ROBERT | Son | 35545587| 1936     | LUCERO   | PETE   | Child |
| 12    | 79199     | 1933     | LUCERO   | BERNARD | Son | 35545587| 1933     | LUCERO   | BERNARD| Child |
| 13    | 79199     | 1931     | LUCERO   | MANUEL | Son | 35545587| 1932     | LUCERO   | MANUEL | Child |

Notes: This table shows how household-information can increase the confidence of individual level record linkages. Observations 7, 10, 11, and 12 are all examples of linkages that look very uncertain on their own, but look very certain in the context of the household-linkage we perform after the individual linkage. Observation 2 is an individual who formed her own household between the 1920s ICR and the 1940 FCC.
References


Online Appendix

to

“The Effect of Land Allotment on Native American Households During the Assimilation Era”
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
outlined in Feigenbaum (2016). There is an active and ongoing debate on the choice of methods. Bailey, Cole, Henderson, and Massey (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, Boustan, Eriksson, Feigenbaum, and Pérez (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.\footnote{For another summary, Ran Abramitzky’s website at https://people.stanford.edu/ranabr/matching-codes.}