

Native Americans in the Historical Census: New Data and Applications

Christian Dippel*

Dustin Frye[†]

June 1, 2021

PRELIMINARY DRAFT FOR CONFERENCE SUBMISSION

Abstract

The digitized historical Full Count Census waves from 1900–1940 are a rich source of information for individual- or household-level quantitative research on the Native American population, with the average census wave containing more than 300,000 Native American individuals. Without the missing information on reservation, however, there is no treatment variation in any of the major historical policies that Native Americans were exposed to, such as Indian boarding schools and land allotment. We describe the construction of a stable reservation-to-individual crosswalk that assigns a reservation to over ninety percent of individuals in the historical Native American population, and apply this crosswalk to answering some long-standing questions on within-reservation inequality.

Keywords:

JEL Codes:

*University of California, Los Angeles, and NBER.

[†]Vassar College.

1 Introduction

American Indians are the poorest Census-defined racial group in the United States; they have the highest rates of violent victimization ([Bureau of Justice Statistics, 2019](#)), and they have the poorest health outcomes, including the worst substance abuse problems ([Center for Disease Control, 2019](#)).¹ A large body of literature in institutional, cultural, and development economics and related fields suggests that these socio-economic outcomes (income, health, and well-being) are shaped by historical factors. In short, history matters.

In the Native American context, there is great potential for quantitative historical research using the historical waves of the 1900–1940 Full Count Census waves.² The 1900–1940 Census waves are a rich data source. The average Census wave contains more than 300,000 individuals whose race is enumerated as being Native American, and the data include rich demographic, and in later years also economic, information. Furthermore, 1900–1940 covers almost the entirety of the “Assimilation Era,” during which Native American tribes were exposed to a range of assimilationist policies such as boarding schools and land allotment, the consequences of which reverberate to the present day ([Treuer, 2012, 2019](#)).³

Unfortunately, the micro data contain no information on the reservation an individual Native American belonged to. This is a critical limitation because reservations varied dramatically in the experience of their formation ([Dippel, 2014](#); [Feir, Gillezeau, and Jones, 2019](#); [Anderson, 2020](#)), in their exposure to different policies during the Assimilation Era ([Carlson, 1978, 1981](#); [Golenko, 2010](#); [Feir, 2016](#); [Gregg, 2018](#); [Leonard, Parker, and Anderson, 2020](#)), and in their political recognition and post-Assimilation era experience ([Anderson and Parker, 2008](#); [Akee, Jorgensen, and Sunde, 2015](#); [Frye and Parker, 2019](#)). It merits clarifying that today’s federally or state-recognized ‘administrative tribes’ are almost always reservations, with the exception of a few reservations that jointly form a federally recognized tribe, and a few recognized tribes without a reservations.

¹The percentage of Native Americans who experienced violence in 2013 was 2.8%, relative to 1.3% for both blacks and Hispanics, and 1.1% for whites.

²The 1900 starting year is determined by the fact that “Indians not taxed” (i.e. the vast majority who were living on reservations at this time) were not enumerated in the Population Census before 1890, and the 1890 Census was lost to a fire. By historical estimates, these constituted roughly ninety percent of the Native American population in 1860, 1870 and 1880. See [National Archives \(2019\)](#). According to the “rule of 72”, the 1950 wave will gradually be made available to researchers starting in 2022.

³The beginning of the era is usually marked by the General Allotment (or ‘Dawes’) Act of 1887, and its end by the Indian Reorganization (or ‘Howard-Wheeler’) Act of 1934.

Reservations can almost always be clearly traced to a single cultural/ancestral tribe (e.g., White Earth is an Ojibwe reservation, Rosebud is a Sioux reservation), for which the Census equally had no information. The absence of reservation/tribe information is a combination of two factors: In most Census waves, reservation/tribe was never enumerated in the first place; and in the 1930 wave where it was enumerated, it is not among the variables that have up to now been digitized in the Full Count micro data.⁴

We address this problem through the construction of a stable reservation-to-individual crosswalk based on a combination of spatial matching techniques using Census-enumerated variables. First, we do text-analysis of the official descriptions of ‘enumeration districts’ to identify districts within reservations. Second, we geolocate all reported places in the Population Census and from the partitioned enumeration district location descriptions. We overlay both sets of locations with historical reservation maps to identify places within reservations. For places outside of reservation boundaries, we calculate the distance from each place to the nearest reservation to determine whether individuals are immediately adjacent to reservations. Combining these two exercises, we create a stable reservation-to-individual crosswalk for each decade from 1900 to 1940 that assigns over 90% of individual Native Americans to their reservation.

This paper’s contribution is to the field of quantitative research on Native American history. This research area has been slow to adopt the micro-data based research approaches that are becoming more common in the quantitative study of other historical populations; instead relying almost exclusively on reservation-aggregates that were reported in Annual Reports by the Bureau of Indian Affairs (BIA) from the 19th century until the Indian Reorganization Act (IRA) of 1934. Examples of such reservation-aggregate studies include [Carlson \(1978, 1981\)](#); [Anderson and Parker \(2008\)](#); [Golenko \(2010\)](#); [Dippel \(2014\)](#); [Akee et al. \(2015\)](#); [Frye and Parker \(2019\)](#); [Leonard et al. \(2020\)](#); and [Feir et al. \(2019\)](#). Exceptions we are aware of include [Akee \(2019\)](#), who uses county of residence to link Native Americans in 1900 and 1910 to one of two reservations in Minnesota (White Earth and Red Lake), as well as [Dippel and Frye \(2021\)](#).⁵ By clarifying the main challenges that researchers will face when using historical micro-data on Native Americans, and proposing solutions to these challenges, we hope to enable more micro-based research on this

⁴Digitization efforts prioritize having a full count of people over having a wide data-set with all variables contained in the hard-copies.

⁵[Miller \(2015, 2016\)](#) also uses micro-data, but her focus is primarily on freed black slaves within the Cherokee nation.

important population.

2 Native Americans in the Historical Censuses

Beginning with the first U.S. census in 1790, enumerators were explicitly instructed not to enumerate “Indians not taxed.” This excluded every Native American living in tribal communities or unsettled territories, including those relocated to reservations. Native Americans were first enumerated as a separate racial group in 1860, but the census continued excluding “Indians not taxed” until 1890, which constituted over 90 percent of Native Americans.⁶

Beginning in 1900, all Native Americans should be enumerated in the Census, but records from the Office of Indian Affairs (OIA) report large population discrepancies between the decennial censuses and OIA tabulations. Table 1 shows population estimates for Native Americans from 1860 – 1940 reported separately by the Census Bureau and the Office of Indian Affairs. The large systemic discrepancy between the census and the OIA from 1860–1880 is explained by the exclusionary enumeration instructions in the early census waves.

From 1900 – 1920, the Office of Indian Affairs consistently reported larger Native American populations than the Full Count Census estimates from the Census Bureau. The Office of Indian Affairs (OIA) noted the discrepancies in population counts between their office and the Census Bureau ([Office of Indian Affairs, 1935](#), p. 62–73). They believed that the enumerations in 1900 and 1910, which featured additional Indian supplements and included personnel from the OIA that were familiar with reservations and Native Americans, resulted in a more accurate count. They attributed the general under counting to resource constraints in reaching remote Native American communities and felt that the large difference in 1920 population is attributed to the exclusion of OIA personnel from the enumeration process, resulting in Native Americans being enumerated as non-Native. These constraints improved by 1930, when the population counts between the census and OIA converge. This paper focuses on the available census sample of individual Native Americans and their household members from 1900 to 1940.

⁶The 1850 census enumerated all members of listed households, but did not systematically identify Native Americans. In 1880, special census enumerators visited a few reservations ([National Archives, 2019](#)).

Table 1: Population Counts of American Indians by Decade

Decade	Census Bureau	Office of Indian Affairs
1860	44,021	339,421
1870	25,731	313,712
1880	65,407	306,543
1890	248,253	249,278
1900	237,196	270,544
1910	265,683	304,950
1920	244,437	336,337
1930	332,397	340,541
1940	333,969	327,958

Notes: Census Bureau figures from the [National Archives \(2019\)](#) and [US Census Bureau \(1951\)](#). Office of Indian Affairs figures from [Office of Indian Affairs \(1935\)](#) and [Office of Indian Affairs \(1941\)](#).

3 Determining Reservation of Residence

Over the first half of the twentieth century, roughly 90 percent of Native Americans resided on or near reservations, but the census rarely collected tribal or reservation affiliations for Native Americans. Not having this information severely limits the questions that can be asked in the micro-data, because nearly all federal policies were administered at the reservation-level. Special schedules in 1900 and 1910 included questions about tribal affiliation and reservation specific living conditions, however the enumerated tribal affiliation was more akin to an individual’s ethnic tribe than to their affiliation with a specific reservation.⁷ This is less useful in the sense that federal policy varied by reservation, not tribe, and because it was reservations that, after 1934, would become the polity that is today called a recognized tribe.

The most readily available spatial information in the historical Census is county. Unfortunately, reservations tend to be relatively tightly clustered in parts of the Pacific Northwest, the Southwest, the Plains, and around the border area of Minnesota and Michigan. These within-county clusters prohibit accurate matching of individuals to reservations.⁸ Instead of county, we rely on reported location from the census and enumeration district descriptions. Enumeration districts have not been geo-referenced, and we therefore rely on the text of their official descrip-

⁷The 1930 regular census schedule also recorded tribal identity, with a similar interpretation.

⁸An additional concern is that county boundaries were also in flux over this period. Between 1900 and 1940 there were nearly 750 county boundary adjustments, with many of these occurring in western states with American Indian reservations ([Siczewicz, 2011](#)).

tions to make use of them. We web-scraped these descriptions from [Morse and Weintraub \(2019\)](#). These descriptions often specifying multiple narrow geographies within a single description. We partition each description into individual, searchable locations, giving us granular location information for every individual in the census. Both sources provide detailed geographic information that we spatially match to contemporary reservation boundaries.⁹

Determining Reservations: First, we manually search for reservation names within the enumeration district descriptions that we web-scraped from [Morse and Weintraub \(2019\)](#). If the description identifies a single reservation we directly assign all individuals living within that enumeration district to that reservation.¹⁰

Second, we clean the census locations and ED location information to create a searchable place with town, county, and state. We geolocate these places through the Google Maps API, which delivers specific coordinates and the level of precision associated with each found location. We exclude any API match where the geolocated county does not match the county identified in the census and any API match where the geographic precision is broader than a locality or reservation. From the set of unmatched locations, we identify two select cases. First, a small set of locations are identified by their placement in the Public Land Survey System (PLSS). For this subset, we assign the centroid coordinates of the surveyed township. Second, we manually search for the set of excluded locations that do not meet our API quality criteria to determine their latitude and longitude and add these to the geolocated places. After compiling the API matched locations, PLSS locations, and manually searched locations we have approximate location information for every member of a Native American household.

Next, we determine whether each location is within a modern reservation boundary, and if not, how far away the location is from the nearest reservation border.¹¹ We assign individuals to a reservation if they are living in a location within a reservation boundary or if their location is within 40 kilometers of a reservation border.

Finally, the spatial information in the 1930 census is better than in the 1940 census. Therefore,

⁹We include the full set of federally recognized reservations, along with Oklahoma Tribal Statistical Areas (OTSAs) and State Designated Tribal Statistical Areas (SDTSAs).

¹⁰In the 1900 and 1910 waves, some counties uniquely identify reservations, and we perform a similar manual assignment.

¹¹To determine reservation boundaries, we use the 2010 reservation boundary file from [Manson, Schroeder, Riper, Kugler, and Ruggles \(2020\)](#), which includes federally recognized reservations, OTSAs, and SDTSAs.

to assign reservations in 1940 we construct a crosswalk between enumeration districts in 1930 and 1940 with content web-scraped from [Morse and Weintraub \(2019\)](#). This `ENUMDIST-1930-to-ENUMDIST-1940` crosswalk allows us to map the 1930 reservation assignment at the enumeration district level to enumeration districts in 1940.

Reservation Assignment: The process resulted in a reservation assignment for roughly 90 percent of individuals living in households with a Native American. Table 2 shows the distribution of reservation assignment by source. The reservation population comes from over 300 different reservations. The table highlights the decade over decade stability in the location of Native Americans over the first half of the twentieth century. Consistent with Table 1, the total enumerated population of Native Americans changes considerably between the early decades and later decades. Despite these differences, we find roughly 70 percent of Native Americans are living within modern reservation boundaries and around 20 percent are within 40 kilometers of modern reservations. The decade over decade stability is consistent with low levels of off-reservation migration prior to the World War II and the Urban Indian Relocation Program of the 1950s.

Table 2: Reservation Assignment by Source and Decade

Panel A: Total Population

Source	1900	1910	1920	1930	1940
Enumeration District Descrip	63,202	75,679	67,225	86,621	75,426
PLSS Within Reservation	0	0	10,568	20,528	20,871
API Located Within Reservation	40,096	107,673	100,888	166,517	146,585
Handlinked Within Reservation	33,411	25,284	16,928	7,211	13,781
API 40 km Radius	48,401	53,402	51,669	86,927	63,335
Handlinked 40 km Radius	899	9,454	4,103	2,884	3,343
Total Reservation Population	186,009	271,492	251,381	370,688	323,341
Off-Reservation Population	15,518	17,112	19,459	31,524	23,719
Total Population	201,527	288,604	270,840	402,212	347,060

Panel B: Population Shares

Source	1900	1910	1920	1930	1940
Enumeration District Descrip	31.4%	26.2%	24.8%	21.5%	21.7%
PLSS Within Reservation	0.0%	0.0%	3.9%	5.1%	6.0%
API Located Within Reservation	19.9%	37.3%	37.3%	41.4%	42.2%
Handlinked Within Reservation	16.6%	8.8%	6.3%	1.8%	4.0%
API 40 km Radius	24.0%	18.5%	19.1%	21.6%	18.2%
Handlinked 40 km Radius	0.4%	3.3%	1.5%	0.7%	1.0%
Total Reservation Population	92.3%	94.1%	92.8%	92.2%	93.2%
Off-Reservation Population	7.7%	5.9%	7.2%	7.8%	6.8%
Total Population	100.0%	100.0%	100.0%	100.0%	100.0%

Panel C: Reservation Coverage

Source	1900	1910	1920	1930	1940
Enumeration District Descrip	51	61	49	61	45
PLSS Within Reservation	0	0	5	9	9
API Located Within Reservation	55	45	61	60	77
Handlinked Within Reservation	9	11	19	12	12
API 40 km Radius	213	169	192	190	172
Handlinked 40 km Radius	3	31	17	15	17
Total Unique Reservations	331	317	343	347	332

References

- Akee, R. (2019). Land titles and dispossession: Allotment on american indian reservations. *Journal of Economics, Race, and Policy*, 1–21.
- Akee, R., M. Jorgensen, and U. Sunde (2015). Critical Junctures and Economic Development—Evidence from the Adoption of Constitutions Among American Indian Nations. *Journal of Comparative Economics* 43(4), 844–861.
- 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.
- Anderson, W. (2020). Politics, Economics, and Native American Conflicts. *University of Michigan mimeo*.
- Bureau of Justice Statistics (2019). American Indians and Crime. <https://www.bjs.gov/content/pub/pdf/aic.pdf>. Accessed: 2019-11-08.
- Carlson, L. A. (1978). The Dawes Act and the Decline of Indian Farming. *The Journal of Economic History* 38(1), 274–276.
- Carlson, L. A. (1981). *Indians, Bureaucrats, and Land: the Dawes Act and the Decline of Indian Farming*. Number 36. Praeger Pub Text.
- Center for Disease Control (2019). Health of American Indian or Alaska Native Population. <https://www.cdc.gov/nchs/fastats/american-indian-health.html>. Accessed: 2019-11-08.
- Dippel, C. (2014). Forced Coexistence and Economic Development: Evidence from Native American Reservations. *Econometrica* 82(6), 2131–2165.
- Dippel, C. and D. Frye (2021). The Effect of Land Allotment on Native American Households During the Assimilation Era. *mimeo*.
- Feir, D., R. Gillezeau, and M. Jones (2019). The Slaughter of the Bison and Reversal of Fortunes on the Great Plains. *R&R at the Review of Economic Studies*.
- Feir, D. L. (2016). The long-term effects of forcible assimilation policy: The case of indian boarding schools. *Canadian Journal of Economics* 49(2), 433–480.
- Frye, D. and D. Parker (2019). Local versus Central Governance: Long Run Effects of Federal Oversight over American Indian Reservations. *Vassar College mimeo*.
- 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.
- Leonard, B., D. P. Parker, and T. L. Anderson (2020). Land quality, land rights, and indigenous poverty. *Journal of Development Economics* 143.
- Manson, S., J. Schroeder, D. V. Riper, T. Kugler, and S. Ruggles (2020). Ipums national historical geographic information system: Version 15.0 [dataset]. doi: 10.18128/d050.
- Miller, M. (2015). Dawes cards and indian census data. *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 48(4), 214–229.
- Miller, M. C. (2016). “the righteous and reasonable ambition to become a landholder”: Land and racial inequality in the postbellum south. *Review of Economics and Statistics*, 1–45.
- Morse, S. P. and J. D. Weintraub (2019). The Unified Census Enumeration District Finder 1880–1940. <https://stevemorse.org/census/unified.html>. Accessed: 2019-11-08.
- National Archives (2019). Native Americans in the Census, 1860-1890. <https://www.archives.gov/publications/prologue/2006/summer/indian-census.html>. Accessed: 2019-11-08.
- 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.
- Office of Indian Affairs (1941). *Statistical Supplement to the Annual Report of the Commissioner of Indian Affairs for the Fiscal Year Ended June 30, 1941*. Department of Interior.
- Siczewicz, P. (2011). U.S. Historical Counties. Available online from <https://publications.newberry.org/ahcbp/>.
- Treuer, D. (2012). *Rez Life: An Indian's Journey Through Reservation Life*. Grove/Atlantic, Inc.
- Treuer, D. (2019). *The Heartbeat of Wounded Knee: Native America from 1890 to the Present*. Riverhead Books.
- US Census Bureau (1951). Statistical abstract of the united states. *US Department of Commerce, Washington: Government Printing Office*.