Native Americans in the Historical Census: New Data and Applications

Christian Dippel* Dustin Fyre†

March 24, 2020

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-Asssimilation 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.

[1] The 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.

[2] The 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.

[3] The 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. This crosswalk is 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 identity those described on reservations. Second, we geo-locate all reported towns in the Population Census and overlay them with historical reservation maps. This allows us to identify many additional enumeration districts which contain towns that are located in or immediately adjacent to reservations. Combining these two exercises, we can create a stable reservation-to-individual crosswalk that assigns over 90% of individual Native Americans in the 1930 and 1940 Census to their reservation.

We then apply this reservation-to-individual crosswalk to answering two long-standing questions on within-reservation inequality. The first pertains to the differential formation-experience of reservations, namely whether reservations were forcibly formed out of formerly politically disconnected tribal bands. The second pertains to the differential experience at the end of the assimilation era, namely the adoption of the Indian Reorganization Act (IRA).

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). The only exception we are aware of is Akee (2019), who uses county

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

of residence to link Native Americans in 1900 and 1910 to one of two reservations in Minnesota (White Earth and Red Lake).[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 important population.

We being by surveying the historical micro-data that exist for Native Americans in Section 2, before describing the construction of the reservation-to-individual crosswalk in Section 3. In Section 4, we then apply this crosswalk to answering some long-standing questions on within-reservation inequality.

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