IPUMS Complete-Count Datasets of the Slave Inhabitants and Slave Owners, 1850-1860: Dataset Construction and Analysis *

J. David Hacker, Matt A. Nelson, Matthew Sobek, and Lap Huynh

Paper presented at the annual meeting of the Social Science History Association, Chicago, Illinois, November 18, 2022.

Abstract: IPUMS has finalized databases for the censuses of slave inhabitants for the 1850 and 1860 censuses. Owners listed in the slave censuses have been linked to the population censuses, allowing analysis of owner characteristics. We discuss the construction of the datasets, their potential uses, and some preliminary analysis. The data will be distributed via IPUMS and is now available for researchers to use.

* This research was supported in part by funding from the Minnesota Population Center (P2C HD041023) and by a grant from the Eunice Kennedy Shriver National Institute for Child Health and Human Development (R01-HD082120-01). Many individuals at the Minnesota Population Center contributed to the project, but we wish to especially acknowledge the work of Ron Goeken, who contributed a significant amount of effort to all aspects of the project over a five-year period. Bigyan Khanal provided some initial programming. Some of the linking information in this paper was originally presented as Ron Goeken, J. David Hacker, Lap Huynh, and Bigyan Khanal, "An Inquiry into Some Points of Record Linkage: Linking 19th Century U.S. Decennial Population Records to the Mortality and Slave Schedules," presented at "Putting the Pieces Together: Promise, Programs and Pitfalls in Linking Historical and Contemporary Records," at the Center for Economic History, Northwestern University, May 17-19, 2019.

IPUMS Complete-Count Datasets the 1850 and 1860 Censuses of Slave Inhabitants

This paper describes new public- and restricted-use datasets documenting the slave population of the United States in 1850 and 1860. The datasets were constructed at the University of Minnesota as part of the IPUMS historical census projects with funding from the National Institutes of Health and are distributed for research via the IPUMS website (http://ipums.org) and data extraction system. We anticipate these datasets, which include individual-level data on xxxx slaves in 1850 and 3,9xx,xxx slaves in 1860, will be valuable resources for the study of the slave population. Approximately 80% of owners named on the 1850 and 1860 slave schedules were located in the IPUMS complete-count datasets for the 1850 and 1860 free populations, allowing slaves to be attached to owners' households and facilitating the study of slave ownership. Our intention here is to provide background information of the slave censuses and describe the datasets, variable availability, and potential concerns for users to consider in their analyses. We also suggest some potential lines of new research these new datasets might support.

Background

The U.S. Constitution required the nation to conduct a census every ten years to determine apportionment of Congress. According to Article 1, Section x of the Constitution, representatives were to be "apportioned among the several States … according to their respective Numbers, which shall be determined by adding to the whole Number of free Persons, including those bound to Service for a Term of Years, and excluding Indians not taxed, three fifths of all other Persons." "All other persons" was an oblique reference to the nation's enslaved population, which at the first census in 1790 totaled 697,624 individuals. From the beginning, Congress required the census to collect more information than simple counts of the free and slave populations. By 1840, for example, the census counted for each household the number of free white male and female household members in 13 different age groups, the number of persons in seven occupation groups, the number of white persons age 20 years or older who could not read and write, and information of the number of white persons who were deaf and dumb, blind, insane or idiotic. Much less data were collected for the free black and enslaved population. In 1840, for example, the number of free blacks and slaves were collected in just six different age groups for both sexes.

In 1850, when the census was changed from a household-level survey to an individual-level survey (albeit individuals nested within dwellings and families), the nation's enslaved population was enumerated on a separate schedule. Although owners' names were recorded, the change had the effect of removing slaves from owner's households. Again, the census collected less information on the slave population An advisory board of leading statisticians had recommended that the census collect the name, age, sex, color, and place of birth of each slave; whether the slave was deaf, dumb, blind, insane, idiotic, or a fugitive; the number of children each enslaved woman had borne and the number of those children still surviving; and the names of each slaves' owners. Forms were printed, but the proposal ran into trouble in Congress, which was concurrently debating sensitive sectional issues regarding slavery, including the Fugitive Slave Act and the possible extension of slavery into California and new territories recently acquired in the Mexican American War. Southern Congressmen, fearing the possibility that information collected by the census could be used to attack their "peculiar institution," successfully reduced the number of questions included on the slave schedule (Anderson 1988, 37-41). The final slave schedule in 1850 recorded only the names of slave owners (not the names of the slaves, the age, sex, and color of each slave ("Black" or "Mulatto"), the number of slaves fugitive in the previous year who had not been returned to their owners, the number of slaves manumitted in the year prior to the census, and

whether the slave was deaf, blind, insane or idiotic. Slaves were grouped by slaveholdings, with no designations within the holding of family groupings or relationships. A similar slave schedule was used in 1860, with the addition of a column for the total number of slave houses in the holding, which were to be shown on the line for the last slave in the holding. A header on each page recorded the place, county, and state, the day of the enumeration, and the name of the assistant marshal conducting the enumeration. See figure 1 for a an example of the slave schedule used in 1860.

Although the census forms were nearly identical in 1850 and 1860, instructions to enumerators were revised in 1860 in an attempt to get a better count of the number of slave owners. In 1850, the instructions under heading 1, "Names of Owners," specified that "where there are several owners to a slave, the name of one only need be entered, or when owned by a corporation or trust estate, the name of the trustee or corporation." Ten years later, the instructions called for assistant marshals to "insert, in proper consecutive order, the names of all owners of slaves. When slaves are the property of a corporation, enter the name of the corporation. If held in trust for persons who have attained to their majority, whose names as owners do not elsewhere appear, the names of such persons way be entered, or their number, as "John Smith and two others;" always provided that the "others" do not appear as owners in other places. If held in trust for minors, give the number of such minors. The desire is to obtain a true return of the number of owners."

The slave schedules were copied and sent to the Census Office in Washington, D.C., where they were hand tabulated by several hundred clerks. The labor-intensive task meant that cross-tabulations were few. The 1850 returns counted the number of slaves and free blacks in each state, cross tabulated by color (Black and Mulatto), and cross-tabulations by age group (not individual ages), sex, and state. The Census Office also counted the number of slave owners by state and crosstabulated the number owners in each state by the size of slave holdings. In total, the 1850 census counted xxx slaves held by 347,325 owners, with the greatest number of owners in Virginia (55,063). Nearly one-half of slaves lived in slave holdings of five or fewer slaves. (Compendium, pp. 82-95)

The number of owners counted by the 1860 census increased between 1850 and 1860. But whether than number reflected increasing ownership of slaves or simply the revised enumeration instructions, which encouraged all owners to be named, is impossible to determine. Some historians (e.g. xxxx), have assumed that slaveholding became more difficult. The 1860 Census Office made similar cross-tabulations….

Although limited, the slave censuses and the published aggregate statistics published by the Census Office includes useful information for the study of the slave population, which like the nation's white population, was experiencing rapid natural population growth, despite the abolition of the international slave trade in 1808 (Hacker 2020). Between 1800 and 1860, the slave population increased xxx to yyyy. The age structure of the population is consistent with very high fertility, with most estimates suggesting a crude birth rates in excess of 50 per thousand population. [Examples of research]. Census data are also critical for tracing the effects of the internal slave trade on the geographic distribution of the population. The predominant patters resulted in slaves being sold (or migrating with their owners) from tobacco growing areas in the upper South to owners in cotton and sugar-growing regions in the lower South and South West Central region. Economic historians have estimated that the slave trade disrupted about 40% of slave marriages.

Although manumission resulted in some attrition from the slave population, the free Black population (which totaled about 10% of the total Black population of free Blacks and slaves) rarely had the means to leave the United States. About xxx,xxx Blacks lived in Canada in 1861, representing less than xx % of North America's Black population. Researchers disagree about the size of slave smuggling in the years between the 1808 and 1865. But despite some known smuggling—the slave ship Clotilda, for example, smuggled 110 African slaves to Alabama in 1859 or 1860—evidence suggests the numbers were low overall. The 1870 census, for example, recorded an African or Caribbean birthplace for only xxxx Blacks, less than xx percent of the total. For practical purposes, therefore, the Black population of the United States can be considered closed to migration, allowing the estimation of mortality and life expectancy using two census methods.

Because the slave population

IPUMS microdata of the slave censuses increases the research potential of these data.

The IPUMS complete-count datasets of the 1850 and 1860 slave censuses

In 2001, the Minnesota Population Center received a grant from the National Science Foundation.

IPUMS microdata of the slave censuses increases the research potential of these data. Users, for example, can use the IPUMS 1850 and 1860 slave datasets to create custom tabulations (e.g., the age structure of the slave population by single years of age or the proportion of mixed race "mulatto" slaves recorded by enumerators in each county) and construct custom variables. Although enumerators did not identify slave dwellings, slave households, or slaveholdings with unique serial numbers as they did for dwellings and families on the free population schedules, the IPUMS datasets identify slaves in the same "slaveholding" using owners' names, which typically appeared on the record of the first slave in the holding. Using a holding identification number constructed by the IPUMS project, researchers can construct holding-level variables (e.g., the child to woman ratio or adult sex ratio) and conduct holding-level analyses.

Table 1 shows the information collected by the census, together with

Despite their larger size and comprehensiveness, we anticipate that IPUMS complete-count slave datasets will not completely replace earlier IPUMS samples constructed from the slave censuses two decades ago. In an earlier publication, Alexander *et al.* (2003) described the original IPUMS 1860 slave sample, including its design and method of construction, availability of variables, and potential uses for the study of slave demography, patterns of slaveholding, and the spatial distribution of the mixed-race

population. Although this earlier dataset and a subsequent dataset constructed for the 1850 slave census was a random sample of only 5 percent of nation's slaves, these datasets were constructed entirely by the IPUMS project. Care was taken with all information recorded on the manuscript returns, with concurrent data entry, checking, verification, and error correction, resulting in the highest possible data quality.[1]

In contrast, the complete-count slave datasets, while including individual records for all slaves, rely on data entered by Ancestry.com without IPUMS oversight. We found that many fields were poorly entered, most notably the geographic place names, number of slave houses, manumissions, and fugitives in each holding, and the disability variables for individual slaves. Stray marks and numbers written in a different hands–such as running totals made by Census Office staff when tabulating the returns–were sometimes entered into fields for individual slaves. In addition, Ancestry.com recorded a maximum of three owners for each slave holding (the earlier sample datasets recorded information for up to eight owners). Identifying information for the type of owner (e.g., whether each "owner" named in the census was an owner, renter, overseer, trustee, guardian, business partnership, etc.) was frequently not recorded. Given the large number of records in the complete-count dataset–more the seven million slaves were enumerated in the 1850 and 1860 slave schedules–the IPUMS project lacked the resources to re-enter all these data. The project did correct some critical errors in the recording of slaves ages (data entry operators for Ancestry.com often entered a month of age for slaves aged 0 and 1 as a year of age, which we corrected), added information on the type of owner in 1860, and identified and corrected errors in the identification of different slave holdings (e.g., when more than one slave holder was listed on two lines of a slave holding it erroneously appeared to be multiple holdings instead of one holding with multiple owners or an owner and employer).

Unsurprisingly, given these problems, we found that many of the variables were unreliable and overall totals differed markedly from published statistics. Users needing individual-level data are therefore encouraged to rely on the earlier sample datasets where possible. Nonetheless, the complete-count datasets should prove the better choice for many research tasks where low sample densities will not suffice.

We anticipate that one use of the slave datasets will be to investigate slave ownership (e.g, Olsen 1972; Oakes 1982: 245-250). In addition to the problems in the data entry and identification of owners noted above, we caution that differences in census instructions makes the comparison of slave ownership rates between 1850 and 1860 problematic. In contrast to 1850, where the instructions indicated that "the principal object [was] to get the number of slaves, and not that of masters or owners," the 1860 instructions emphasized getting an accurate count of the number of owners. The instructions stated that "The person in whose charge, or on whose plantation the slave is to be found to be employed may return all slaves in his charge (although they may be owned by other persons) provided they are not returned by their proper owner. The name of the bona fide owner should be returned as proprietor, and the name of the person having them in charge as employer."[2] Thus each holding should have an owner's name on the slave schedule (and the owner will usually also be the employer) while other holdings will have both an owner and employer name. And, in the latter case, it was up to the enumerator whether the slaves would

[1] These datasets are available for free public download at https://usa.ipums.org/usa/slavepums/data/data.html .

[2] See https://usa.ipums.org/usa/voliii/inst1860.shtml. See also https://www.afrigeneas.com/library/slave_schedule2.html.

be enumerated at the time of the owner's or employer's visitation (i.e., when the owner or employer was enumerated on the population schedule).

Because of the importance of the name fields for linking (discussed below), the project spent a fair amount of effort correcting the name information in the original Ancestry data. In addition to improving name transcription accuracy, we also restored non-name information describing the relationship between the holding names and the slaves.

VI. Linked Datasets

The research potential for the censuses of mortality and slave inhabitants is increased substantially by linking decedents to their household of origin and slaves to their owners in the population censuses. In this section we discuss the construction of linked mortality datasets (the 1870 and 1880 mortality censuses linked to the corresponding population censuses) and linked slave datasets (slaves with owners in the 1850 and 1860 slaves censuses linked to owners in the corresponding population censuses).

VI - b. Linking the 1850 and 1860 slave datasets to the 1850 and 1860 IPUMS complete-count datasets

Linking slaveholders and their slaves to the population record was difficult because owners' names on the slave schedules and individuals' names on the population schedule are the only explicit linkage variables. As noted above, the slave schedules contained limited information on the slaves themselves (age, sex, and color, but not their names). The only names on the form were for slaveholders. Unfortunately, with the exception of their names, no other slaveholder characteristics—like age, sex or place of birth—that are typically used to link census population records were recorded. See Figure 2 for an example. The schedules contained reliable information for state and county, however, which we use for blocking during the potential links generation process. In addition, we rely on the order slaveholdings were recorded in the manuscript pages. As noted above in the description of the mortality censuses, a feature of the various decennial schedules (i.e., free, slave and mortality) is that they were enumerated by the same enumerator during the same visitation (at least in theory). Thus the holdings and population records should follow the same general order in their respective enumerations.[3]

For linking to the population data, we reduced the slave data to unique combinations of holdings and owner/employer names. We then constructed a potential links file by comparing the holding data to individuals' names in the population data (blocking by state and county). We wrote out potential links if both the given name and surname comparisons had a Jaro-Winkler score of at least 0.8.[4] In addition, we only accepted potential links if the population record was 18 or older and had personal property of at least $100.[5]

[3] However, we assumed that enumerators occasionally did not fill out the free and slave schedules during the same visitation. We also had to deal with the possibility that the pages of the respective schedules were microfilmed out of their original order.

[4] For the preliminary links we did not use given name standardizations. In addition, if either a holding or population given name consisted of a single letter, then we hardcoded the Jaro-Winkler score at 0.8 for any match to the first letter of a full given name. For example, a comparison of "J" to "John" will return a score of 0.8.

[5] The value of slaves was supposed to be included in an individual's personal property value on the population schedule; thus slaveowners would be expected to typically have more than $100 in personal property (although this would not necessarily be true if the holding name is an employer rather than an owner). We assumed that the

For 1850 slave linking, which began with an earlier project, Jaro Winkler scores were calculated using FEBRL. The majority of cases were machine linked, but approximately 12% of the slave owners were hand linked. Hand links were sometimes necessary because slave owners did not always live in the county where the slaves they owned were enumerated. 1860 slave linking was entirely machine linked over multiple rounds. We assigned a score of 0.8 to the given name when one side contains a single letter and matches the other schedule's first letter in the name (e.g. 'J' and 'John'). Subsequent rounds included some cleaning of bad links and lowering the Jaro Winkler threshold to 0.65 for both given and surname. In the final round of linking, we blocked by state and raised the Jaro Winkler threshold to 0.9 for the last name.

For illustration, we show an example. Figure 3 shows a partial transcription for the ten slaveholdings shown in Figure 2 for Lafayette County, Mississippi, together with IPUMS-generated variables "Holding_Serial," "Holding_Order," "Holding_Type1" (for the first slaveholder named) and "Holding_Type2" (for the second slaveholder named). On line 19 of the manuscript the enumerator listed "R Hewlett" and appears to continue on line 20 with "for H Mullin." We've interpreted this as a new slaveholding beginning on line 19, increasing the holding serial number at that line, and including R Hewlett's type as an employer and H. Mullin's type as the owner. The holding continues through line 22 (four slaves in total), when the entry of a different owner on line 23, "John Burgess" indicates a new holding. In Figure 4, we show all slaveholders named on the slave schedule. The second holding number repeats, because the enumerator gave information for the owner and employer (holding 308: "R Hewitt" and "H Mullin" respectively).

Figure 5 contains the same holdings along with their potential links from the population data in Lafayette County. The figure also shows age and wealth information along with Jaro-Winkler name similarity scores for each potential link. The bolded lines indicate the link with the highest name similarity. However, we assumed that the slave linking process would be iterative and that relying solely on name similarity would introduce an unacceptable number of false positives (i.e., we anticipated multiple passes and a process of elimination, and we wanted to minimize false positives in the first linking pass).

We augmented the name similarity scores by determining whether slave-to-population links were "neighbors" with other slave-to-population links. The basic idea was to search a grid consisting of a range of slaveholding serial numbers and a corresponding range of population household serial numbers and count the number of likely slaveholding-to-population links that existed in that space. The center of the grid was determined by a specific slaveholder potential link (i.e., the combination of slaveholder serial and population household serial) and was constructed for each potential link. Figure 6 gives an example with the Thomas Williams/Thomas William potential link as the reference point. For this specific potential link the slaveholding serial was equal to 314 and the population household serial in the complete-count dataset was 690. For this example the range was defined as +/- 5 slaveholdings and +/- 25 population households and we wanted to know if there are any (high quality) potential links where the slaveholding serial was between 309 and 319 (inclusive) and the population household serial was between 665 and 715 (inclusive). We determined that the Thomas Williams/Thomas William potential link had five

various thresholds (especially names) prevented the identification of some true links. However, we expected the linking to be a multi-pass process, and assumed that any true links missed in the first pass will be identified in subsequent passes.

neighbors (that also appear to be plausible holding-to-population links), which are shown in neighbor records.

We used the neighbor count and the potential link's location in the population file to reject potential links (i.e., potential links that we would select on the basis of name similarity to the slaveholder). Figure 7a shows the set of 10 holdings after we reduced to potential links with the highest name similarity. Again, two of the holdings did not have any potential links, but we also rejected two of our original (top name similarity) links because of the absence of neighbors and improbable location in the population data (i.e., the population page information is out of order for holding 311 J B Caruthers and holding 315 Alfred Browning). In this example both of the rejected links have relatively low name similarity. But in general, our approach here is fairly conservative; we were willing to reject high name similarity potential links on the basis of implausible location in the population file.

The current process linked over 60 percent of the slaveholdings. Although it is impossible to evaluate the accuracy without external data, we are optimistic about the quality of the links. Our optimism is based on having links with reasonable name similarity that also meet minimum requirements for age and personal property wealth, along with the presence of slaveholder-to-population neighbors and logical order in the population files. We assumed that the currently unlinked holdings fell into two general categories. Absentee owners will (possibly) not have slave-to-population neighbors (according to our constructed neighbor count) and will usually not be in the same population order as adjacent slaveholders. This group will be difficult to link accurately because links would depend largely on name similarity to identify the correct link. The second category will consist of holdings where there was insufficient name similarity to qualify as a potential link. We believe that many of these will be identifiable at lower name similarity thresholds.

For example, we can identify three of the four unlinked holdings in Figure 7a. The first unlinked holding is listed as an owner, and we see little evidence of "H Mullin" in the population data for Lafayette County (and it is possible he resided in another county). But the true links for the other three unlinked holdings were located in the population data (in close proximity to the true links for their slaveholding neighbors, but were not included in the original set of true links because of low name similarity; see Figure 7b). Depending on future funding, we hope to automate this part of the process and release revisions to the dataset (i.e., using lower name similarity links after we have established where we expect to find the true link in the population file).[6]

VI. Conclusion

New nineteenth-century population, mortality, and slave complete-count datasets from IPUMS provide researchers with new resources to study socioeconomic and demographic outcomes in the United States. Although work on the data, coding and linking is complete, the final datasets need to be processed

[6] Part of this process will involve identifying "likely" slaveholders in the population data who are currently unlinked. Identifying the true population link for absentee owners (who will have to be linked using name only) will remain difficult, however. In addition, we expect to revise the breaks between specific holdings based on our preliminary linkage results. For example, the first two holdings in Figure 1 have Rueben Howlett and R Howlitt (owner/employer for holding 307 and employer of holding 308), and both link to the same population record (Reuben Hewlett). Thus it seems plausible to interpret lines 1 through 22 (first column) of Figure 1 as comprising a single holding, with Rueben Hewlett employing all 22 slaves, but only owning the slaves on lines 1 through 18 (with H Mullins owning the slaves on lines 19 through 22).

by the IPUMS project before being made accessible on the IPUMS website for public downloading (ipums.org). Preliminary full count population data are available now. We anticipate that all final datasets will be available in the next few months.

Because the data are complete, researchers can now study small subpopulations and sub-county geography. Researchers can also calculate various aggregated summary statistics. Because the nineteenth-century Census Office tabulated results by hand–the 1860 Census Office required the work of 400 clerks–virtually nothing beyond basic population counts was published for sub-county areas, and only basic counts of broad categories of characteristics were tabulated for counties. The few statistics that were provided frequently used categories that are incompatible with those of later census years. Using the complete count datasets allows researchers to create aggregated summary files for sub-county areas.

Numerous other applications are possible, including analyses of slaveholding, slave population processes, segregation, and correlates of morbidity and mortality. Because of a few data transcription issues noted above, however, researchers should consider whether the samples remain more appropriate for particular analyses.

Restricted versions of the data are available for each nineteenth century census. The restricted versions include names, street address when available, and the input data (strings and coded). Accessing these data does require specific stipulations in order to use, and interested users should contact ipums@umn.edu or ipumsres@umn.edu to request access to these data.

Users interested in linking individuals across censuses can use IPUMS MLP data (Helgertz et al 2020). While IPUMS MLP has similar representative issues regarding non-whites and migrants as other linked methods, IPUMS MLP links women far more effectively than other methods and obtains lower false positive match rates than other methods (Helgertz et al 2022). Some work that has used IPUMS MLP includes the effect of kin proximity on married women's fertility and mortality (Hacker et al 2021), educational opportunities during Reconstruction on Black occupational outcome and literacy (Jones and Schmick 2022), and childhood growth and later life socioeconomic status (Roberts, Helgertz & Warren 2022).

Citations

Alexander, Trent, Condon, Sean, Digman, Jason and Hacker, J. David. 2003. Public Use Microdata Samples of the 1860 Census of Slave Inhabitants. *Historical Methods* 36:1, 21-26.

Anderson, Margo J. 1988. The American Census: A Social History (New Haven: Yale University Press).

Anderton, Douglas L and Leonard, Susan Hautaniemi. 2004. Grammars of Death: An Analysis of Nineteenth-Century Literal Causes of Death from the Age of Miasmas to Germ Theory. *Social Science History* 28:1, 111-143.

Condran, G. A. and E. Crimmins. 1979. A description and evaluation of mortality data in the federal census: 1850-1900. *Historical Methods* 12: 1-23.

J.D.B. Debow. 1853 The Seventh Census of the United States: 1850. Washington, D.C.: *Robert Armstrong, Public Printer*.

Ferrie, J. P. 1996. A new sample of males linked from the public use microdata sample of the 1850 U.S. federal census of population to the 1860 U.S. federal census manuscript schedules. Historical Methods 29: 141-156.

Ferrie, J. P. 2003. The rich and the dead: Socioeconomic status and mortality in the United States, 1850-1860. In Health and Labor Force Participation over the Life Cycle: Evidence from the Past, edited by D. L. Costa, 11-50. Chicago, IL: The University of Chicago Press.

Gardner, T.K. 1998. The Metropolitan Fringe: Suburbanization in the United States before World War II. University of Minnesota Dissertation, 1998.

Glover, J. W. 1921. United States Life Tables, 1890, 1901, 1910, and 1901-1910. Washington, D.C.: GPO.

Goeken, Ron, Huynh, Lap, Lynch, T. A., and Vick, Rebecca Vick. 2011. New Methods of Census Record Linking, *Historical Methods*, 44:1, 7-14. DOI: 10.1080/01615440.2010.517152

Hacker, J. David. 2010. Decennial Life Tables for the White Population of the United States, 1790-1900. *Historical Methods* 43:3, 45-79.

Hacker, J. David. 2020. From '20 and odd' to 10 million: The Growth of the Slave Population in the United States. *Slavery and Abolition* 41:4, 840-855.

Hacker, J. David, Helgertz, Jonas, Nelson, Matt A. and Roberts, Evan. 2021. The Influence of Kin Proximity on the Reproductive Success of American Couples, 1900-1910. *Demography* 41:6, 2337-2364.

Haines, M. R. 1979. The use of model life tables to estimate mortality for the United States in the late nineteenth century. Demography 16: 289-312.

Haines, M. R. 1998. Estimated life tables for the United States, 1850-1910. Historical Methods 31: 149-169.

Haines, M.R., and Inter-university Consortium for Political and Social Research. 2010. Historical, Demographic, Economic, and Social Data: The United States, 1790-2002. Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/ICPSR02896.v3

Helgertz, Jonas, Price, Joseph, Wellington, Jacob, Thompson, Kelly J, Ruggles, Steven & Fitch, Catherine A. 2022. A new strategy for linking U.S. historical censuses: A case study for the IPUMS multigenerational longitudinal panel. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 55:1,12-29, DOI: 10.1080/01615440.2021.1985027

Jones, Daniel and Schmick, Ethan, Reconstruction-Era Education and Long-Run Black-White Inequality (July 18, 2022). Available at SSRN: https://ssrn.com/abstract=4165983 or http://dx.doi.org/10.2139/ssrn.4165983

Logan, Trevon D., and Parman, John M. 2017. The national rise in residential segregation, *The Journal of Economic History*, 77, 127-170. https://doi.org/10.1017/S0022050717000079

Olsen, Otto H. 1972. Historians and the Extent of Slave Ownership in the Southern United States. *Civil War History*, 18:2, 101-116.

Roberts, Evan, Helgertz, Jonas & John Warren, John. 2022. Childhood growth and socioeconomic outcomes in early adulthood evidence from the inter-war United States. *The History of the Family*, DOI: 10.1080/1081602X.2022.2034658

Vinovskis, M. A. 1978. The Jacobson life table of 1850: A critical re-examination from a Massachusetts perspective. Journal of Interdisciplinary History 8: 703-724.