

A Public Use Microdata Sample of the 1860 Census of Slave Inhabitants

J. TRENT ALEXANDER
*Minnesota Population Center
University of Minnesota*

SEAN CONDON
*Department of History
Adrian College*

JASON CARL DIGMAN
*Minnesota Population Center
University of Minnesota*

J. DAVID HACKER
*Department of History
Binghamton University
State University of New York*

Abstract. The authors describe a public use microdata sample of the 1860 slave population of the United States recently created at the Minnesota Population Center. They discuss the key substantive issues that quantitative historians are likely to address with the data set, such as the demography of slavery, patterns in slaveholding, and miscegenation. They outline the sample design, data-entry procedures, variable availability, and documentation of the final data set.

Keywords: data set, demography, slaveholding, slavery

This article describes a new resource for the study of slavery in the United States: a public use microdata sample of the 1860 slave population. This data set was created in conjunction with a public use sample of the 1860 free population, and both samples are part of the Integrated Public Use Microdata Series (IPUMS). The IPUMS is a nationally representative data set of individual- and household-level census information spanning the years 1850 to 2000. The article discusses the key substantive issues that researchers are likely to address with the 1860 slave data set and also details the sample design and variable availability.

The 1860 census of the United States included two population schedules: Schedule 1, which enumerated the nation's 27 million free inhabitants, and Schedule 2, which enumerated the nation's 4 million slave inhabitants. Unfortunately, the data collected on the slave population were more limited than those collected on the free population. The slave schedules lack critical information such as name, occupation, and birthplace. As a result, most researchers studying antebellum social and economic history have relied solely on data sets drawn from Schedule 1. The existing 1860 IPUMS sample, for example, was drawn solely from the schedule of free inhabitants (Ruggles and Sobek 1997).

Despite their deficiencies, the slave schedules contain a host of individual-level data including the age, sex, and "color" (e.g., black or mulatto) of each slave, whether they were "deaf," "dumb," "blind," "insane," or "idiotic," and the

names of slaves over 100 years of age. The slave schedules also contain valuable information about each slaveholding, including the name of each slaveholder, the size and composition of the holding, the number of slaves in the holding who were fugitives from the state, the number of slaves who were manumitted in the year preceding the census, and the number of houses provided to slaves.

The Minnesota Population Center is constructing several historical census samples that will enhance the usefulness of the 1860 slave sample. We are currently working on an NSF-funded project to complete a 1-in-20 sample of the 1850 slave population ("Public Use Microdata Samples of the 1850 Slave Population," NSF SES-0214300). We plan to release that sample in late 2003. In addition, we will construct a second sample of the 1850 slave population by linking slaveholding individuals in the existing 1850 IPUMS sample to their slaveholdings in the census of slave inhabitants. We hope to release the linked sample in late 2004. Together, the 1850 and 1860 slave samples will allow researchers to examine changes in the slave population and slaveholdings in the critical decade preceding the American Civil War.

Research Applications

In the last three decades, the focus of research on American slavery has shifted dramatically. Instead of evaluating the institution of slavery as a whole, scholars have increasingly centered their analyses on slave life and culture. Unfortunately, there are few sources of individual-level data on slaves. Those seeking quantitative evidence have been faced with two choices: they can either study a limited geographic area and create small samples of microdata from the slave schedules, or they must rely on limited aggregate data from the Census Office's published summary tables. In short, scholars have been forced to choose between individual-level detail and geographically representative coverage.

There are many advantages to microdata from the census of slave inhabitants. Individual-level census files have proven to be an indispensable resource for social scientists because they allow researchers to make tabulations tailored to their specific research questions. In addition, census microdata enable researchers to move beyond simple tabular analysis and apply increasingly sophisticated multivariate techniques. In the following discussion, we outline some of the most significant issues that scholars are likely to address using the added power of individual-level data from the 1860 slave sample.

Demography of the Slave Population

A sample of the slave population in 1860 offers new opportunities for the study of slave demography. First, the detailed age structure of the slave population in 1860 can be examined with microdata. Given the paucity of data available on slaves, the age distribution published by the Census Office has served as the principal means to study slave fertility, mortality, and migration (Menard 1975; McClelland and Zeckhauser 1982; McDaniel and Grushka 1995). Unfortunately, published census data are available only in large age-intervals. In 1860, the Census Office reported the number of slaves under age 1, between the ages of 1 and 4, in 5-year intervals through age 20, and in 10-year intervals at ages 20 and higher. Using the 1860 slave sample, researchers can tabulate slaves by single years of age. Single-year age distributions will also allow investigators to estimate—and possibly correct for—age-heaping errors.

Second, the 1860 slave sample allows researchers to study how the demographic behavior of slaves was related to various aspects of slavery. Using a sample of 57 southern counties in 1860, for example, Richard Steckel (1985) demonstrated that slave fertility varied by size of the slaveholding. Slave women who lived with a relatively small number of other slaves had higher fertility (measured by child-woman ratios) than did women on larger holdings, suggesting that the upward trend in mean plantation size between 1850 and 1860 contributed to declining slave fertility. Other factors, including cropping patterns (tobacco, cotton, rice, sugar, and the like) and the sex ratio were also correlated with this tendency. Alternatively, demographic historians of the colonial south report that fertility increased with plantation size (Kay and Cary 1995; Menard 1995), suggesting that the relationship between plantation size, cropping patterns, adult sex ratios, marriage opportunities, and fertility was complicated and perhaps changed over time. The public use sample of the 1860 slave population allows the relationships between fertility and plantation size to be more fully modeled and tested.

In addition to fertility, the 1860 slave sample facilitates the investigation of a wide range of demographic topics, including analysis of sex- and age-specific residence and

migration patterns, estimation of potential spouse availability within slaveholdings, identification of factors related to manumission, and various other individual- and holding-level analyses.

Patterns in Slaveholding

In addition to providing key demographic information on slaves, the 1860 slave sample permits detailed study of the structure of slaveholdings. Since the 1960s, numerous studies of American slavery have focused on the development of an autonomous slave community (Blassingame 1972; Genovese 1972; Gutman 1976; White 1985). The issue of holding size has emerged as central to these discussions; critics of the idea of an autonomous slave community have pointed out that the vast majority of slaves lived on small holdings, often in relatively isolated areas (Elkins 1976; Shore 1986; Wyatt-Brown 1988; Kolchin 1993). Published figures from the 1860 census provide aggregate data only on holding size by county; the slave data sets will offer much more. For example, published figures group farm and plantation holdings into broad categories; one category ranges from 50 to 69 slaves, one from 70 to 99, and another from 100 to 199. The 1860 slave sample allows researchers to produce smaller categories for analysis and to estimate median holding sizes. The data set also permits the study of interstate and intrastate patterns of holding size, comparison of regions that cut across state boundaries such as the cotton south, the tobacco south, and Appalachia, and comparison of urban areas with their surrounding hinterlands. Using the 1860 slave data set in conjunction with the 1850 data set currently being produced, scholars will also be able to investigate how the composition of the shrinking urban slave population changed between 1850 and 1860 (Wade 1964; Starobin 1970; Fields 1984; Goldin 1976; Fogel 1989).

Unlike the 1850 data set, however, the 1860 slave sample will further allow researchers to analyze patterns in ownership type. Census instructions in 1860 noted that marshals were to record the names of corporations and whether or not the “owner” was holding slaves in trust for minors. Marshals were also instructed to designate persons employing slaves as “employers,” with the name of the bona fide owner returned as a “proprietor” (U.S. Census Office 1860). Although the degree to which marshals followed these instructions is unclear, about 7 percent of slaveholders in the data set were designated as corporations, employers, trustees, proprietors, and minor heirs. Because the Census Office recorded ownership types much less systematically in 1850, the 1860 slave sample presents a unique opportunity to investigate the diversity of slaveholding in the antebellum United States.

Miscegenation

Scholars have long discussed the prevalence of interracial relationships between slaveholders and their slaves (Gen-

ovese 1972; Fogel and Engerman 1974; Malone 1992). Indeed, the topic generated extensive dialogue during the last decades of slavery as abolitionists charged widespread “miscegenation” between slaveholders and their subordinate female slaves (Douglass 1845). Census data cannot answer unequivocally the question of how often interracial relationships took place, nor can they help scholars characterize the nature of those relationships, but they can provide a number of important clues to its incidence and correlates through use of the “color” designation present in the slave schedules. The published figures on the color of slaves, for example, indicate a marked growth in the number of mulatto slaves in the 1850s.¹ Scholars have observed this pattern and noted some specific facets of the mulatto population in 1860, such as their increased likelihood of living in urban areas (Steckel 1980). The 1850 and 1860 slave samples will make it possible for the first time for researchers to undertake a systematic study of change in the overall characteristics of the mulatto slave population across time and space.

In addition to allowing scholars to contribute to a broad range of historical debates regarding slavery and slaveholding, the 1860 slave sample will add a critical population to the larger IPUMS database. At the most basic level, IPUMS users will be able to generate representative estimates for the entire 1860 population including both its slave and free components. When combined with the existing IPUMS samples and the forthcoming 1850 slave sample, the 1860 slave sample will open up many lines of inquiry into the study of slavery and its impact on nineteenth-century United States.

Sample Design and Variable Availability

The 1860 slave sample was drawn from the manuscript records of the Census of Slave Inhabitants conducted by the U.S. Census Office. The complete manuscript record of the slave population consists of approximately 50,000 census forms, with a maximum of 80 lines per form. The enumeration covered the population of all slaveholding states, territories, and the District of Columbia. These records are contained on 94 reels of microfilm from the National Archives.²

Creation of the 1860 slave sample proceeded in two phases. In an initial pilot project, we entered, cleaned, and coded the complete information for all slaveholdings found in 160 counties (14 complete microfilm reels).³ On the basis of the results of the pilot project, we decided to create a 1-in-20 sample of the slave population in the remaining slaveholding counties by randomly sampling slaveholdings. We later sampled the slaveholdings obtained from the “complete-count” counties in the pilot project and combined them with the slaveholdings from the sampled counties to yield a 1-in-20 sample of the slave population for the entire nation. The final sample contains 195,270 slaves enumerated on 20,078 slaveholdings.

Sample Design

We relied on slaveholding as the basic sample unit. In the 1860 census, all slaves were assigned to one or more *slave owners*, a term that included owners, agents, corporations, estates, trustees, guardians, overseers, and employers. We prefer *slaveholder*, the term the Census Office used when it published the aggregate statistics. Regardless of whether the named slaveholder was an actual owner, all slaves in the holding were numbered in order of sex and age, usually one slave per manuscript line. Each slave was described by age, sex, color, and whether deaf, dumb, blind, insane, or idiotic. Slaveholders' names were listed on the same line as the information for the first slave in the holding; slaveholdings with multiple slaveholders were named on subsequent lines or in the margins. The number of slaves fugitive from the state (i.e., runaways) and the number of slaves manumitted in the previous year was entered on the first line across from the first slaveholder's name.⁴ The number of slave houses was typically entered across from the last slave in the holding.

The sample design is analogous to the existing 1860 IPUMS sample of the free population, which is simultaneously a sample of households and of free individuals (Hacker et al. 1999). Instead of the household, however, the sample unit was the slaveholding. Some variables, such as holding size and state and county of residence, are common to all slaves within the holding. Other variables, such as age, sex, color, and the various disabilities, are specific to individual slaves (for a full list of variables, see tables 1 and 2). Although most slaves in a given holding share the same slaveholder information, there were a significant number of exceptions. One slave in a holding, for example, may be listed as held in trust by the slaveholder for an unnamed minor child; another may be listed as employed on the holding and owned by a nonresident proprietor. We therefore decided to treat owner name, owner type, and number of owners as individual slave-level data. Inclusion of all slaves in a holding allows the construction of intraholding variables, such as sex ratios, the number of adult male slaves, and child-woman ratios.

Sampling Strategy

The sampling strategy was based on the manuscript page from the slave schedules. Each page contains 80 lines, each of which describes a slave of a particular age, sex, color, and disability. We randomly generated four consecutive sample points on every manuscript page (e.g., lines 24, 25, 26, and 27). To ensure that holdings had an equal probability of being included in the sample regardless of their size, they were selected only if a sample point fell on the line containing the first slave in the holding (i.e., the line containing the slaveholder's name). When the sample point fell on any other line, the holding was skipped. If the holding

was selected, all slaves in the holding were entered. Under this procedure, each holding and slave in the population had an equal probability of inclusion (Ruggles 1995).

Data Entry

Data entry was carried out at a workstation equipped with a computer and microfilm reader. We created a custom data-

entry program based on Microsoft Access. The program integrated the sample selection and data-entry procedures, incorporated interactive data consistency checking, and included context-sensitive help.

The data-entry program presented the data-entry operator with a close facsimile of the census form. The operator read the page and line number of each sample point from a pop-up window on the computer screen and then located the cor-

TABLE 1. Variable Availability for the 1860 Slave Public Use Microdata Sample

Order	Variable name	Description
1	YEAR	Census year
2	DATANUM	Data-set number
3	SERIAL	Slaveholding serial number
4	SLAVENUM	Slave number within holding
5	WEIGHT	Sample weight
6	REEL	Microfilm reel number
7	PAGESEQ	Microfilm page sequence
8	PAGENO	Microfilm page number
9	LINE	Microfilm line number
10	STATEICP	State ICPSR code
11	COUNTY	County ICPSR code
12	CITY	City name code
13	CITYPOP	City population (incorporated places)
14	ENUMDAY	Day of enumeration
15	ENUMMO	Month of enumeration
16	SIZEHOLD	Holding size
17	NUMTAKN	Number taken
18	SAMPTYPE	Sample type
19	PGCNTY	Parker-Gallman county number
20	AGE	Age of slave
21	AGEMNTH	Age of slave in months
22	SEX	Sex
23	COLOR	Color/race
24	FUGITIVE	Fugitive slave (identified on slave record)
25	MANUMIT	Manumitted slave (identified on slave record)
26	BLIND	Blind
27	DEAF	Deaf
28	IDIOTIC	Idiotic
29	INSANE	Insane
30	NOHOLDRS	Number of slaveholders
31	NFUGITVS	Number of slaves fugitive from holding
32	NMANUMTS	Number of slaves manumitted in holding
33	NHOUSES	Number of slave houses on holding
34	SH1TYPE	Slaveholder number 1 type
35	SH2TYPE	Slaveholder number 2 type
36	SH3TYPE	Slaveholder number 3 type
37	SH4TYPE	Slaveholder number 4 type
38	SH5TYPE	Slaveholder number 5 type
39	SH6TYPE	Slaveholder number 6 type
40	SH7TYPE	Slaveholder number 7 type
41	SH8TYPE	Slaveholder number 8 type
42	QAGE	Quality flag for age
43	QSEX	Quality flag for sex
44	QRACE	Quality flag for color

Note. ICPSR = Inter-university Consortium for Political and Social Research.

TABLE 2. Variable Availability for the 1860 Slave Public Use Microdata Sample Supplemental File

Order	Variable name	Description	Variable type
1	YEAR	Census year	Numeric
2	DATANUM	Data-set number	Numeric
3	SERIAL	Slave serial number	Numeric
4	SLAVENUM	Slave number within holding	Numeric
5	REEL	Microfilm reel number	Numeric
6	PAGESEQ	Microfilm page sequence	Numeric
7	PAGENO	Microfilm page number	Numeric
8	LINE	Microfilm line number	Numeric
9	PLACE	Detailed place information	String
10	O1LN	Slaveholder 1 last name	String
11	O1FN	Slaveholder 1 first name	String
12	O2LN	Slaveholder 2 last name	String
13	O2FN	Slaveholder 2 first name	String
14	O3LN	Slaveholder 3 last name	String
15	O3FN	Slaveholder 3 first name	String
16	O4LN	Slaveholder 4 last name	String
17	O4FN	Slaveholder 4 first name	String
18	O5LN	Slaveholder 5 last name	String
19	O5FN	Slaveholder 5 first name	String
20	O6LN	Slaveholder 6 last name	String
21	O6FN	Slaveholder 6 first name	String
22	O7LN	Slaveholder 7 last name	String
23	O7FN	Slaveholder 7 first name	String
24	O8LN	Slaveholder 8 last name	String
25	O8FN	Slaveholder 8 first name	String
26	SLAVNAME	Slave name	String

responding page and line number on the microfilm. If the sample point was valid, operators transcribed all the information for the holding verbatim, including owner names, and the state, county, and "place" information that appears at the top of the census form. All alphabetic fields were entered as written, except that certain common ownership types and disabilities were assigned standard abbreviations. We programmed the data-entry software to perform a variety of interactive logical checks for internal consistency and for valid entries in most fields.

Variable Availability

Numeric codes have been substituted for all alphabetic fields except for slaveholder names, the names of slaves over 100 years of age, and detailed "place" information. Coding is fully compatible with the IPUMS database whenever possible. State and county codes rely on the Inter-university Consortium for Social and Political Research (ICPSR) standards, allowing users to attach selected ecological characteristics of the state or county to each record, such as population characteristics, type and volume of manufacturing, and type and volume of agriculture. These variables will allow multilevel

analysis of the effects of contextual variables on individual-level behavior. We have also included the county code used in the 1860 Parker-Gallman (1976) sample of southern farms to facilitate comparison of the two data sets.

Substitution of numeric codes for most alphabetic strings dramatically decreased the size of the database. Unfortunately, the sample remained very large because of the large number of slaveholders' names. As many as 8 unique slaveholder names existed for individual slaves, requiring 16 separate slaveholder name fields in the sample. The addition of detailed place information and the names of slaves over age 100 also increased the size of the sample. Because these variables will be of little analytical use to most users, we decided to remove them from the sample and include them in a separate supplemental file. Interested users can download the supplemental file and attach it to the sample by matching the reel, page sequence, page, line, and slave numbers.

As in previous census samples, we relied on hot-deck allocation procedures to impute a small number of missing or illegible values for age, color, and sex. We relied on holding size, position in holding, and the age, color, and sex of the previous slave in the holding as criteria for matching a

“donor” record used to impute the missing or inconsistent value. To allow researchers to identify altered variables, allocated data items are identified by a data-quality flag. A list of variables included in the sample are shown in table 1. Variables included in the supplemental file are shown in table 2.

Documentation and Dissemination of Samples

As part of the project, we produced a user's guide that includes a general description of the 1860 Census of Slave Inhabitants, a complete list of enumerator instructions, details of how the sample was created, and guidelines for its use. The user's guide also includes a data dictionary with descriptions and frequency distributions for most variables. Users can browse the documentation on the IPUMS Web site or download the entire document in Adobe PDF format. The data are in raw ASCII format, and we provide the necessary syntax file to convert the ASCII file to SPSS. The syntax file includes variable names, sample weights, and labels. All files can be downloaded from the IPUMS Web site at <http://www.ipums.org>.

In the summer of 2003, we will also release the complete data collected in the pilot project for the 1860 slave sample. This data set will include all information for the 539,509 slaves enumerated in 53,790 slaveholdings in the 160 selected counties. We also plan to release an enhanced version of the 1860 slave sample that includes an oversample of slaves in large slaveholdings (holdings with 100 or more slaves). Interested users should check the IPUMS Web site for release dates, documentation, and special notes on how to use these samples.

NOTES

The sample was created at the Minnesota Population Center of the University of Minnesota with funding provided by the National Institute of Child Health and Human Development (R01 HD34572). Steven Ruggles was the principal investigator. J. David Hacker was the project manager, Cuong Nguyen was the programmer, and Sean Condon was the graduate research assistant. Dianne Star, Cindy Hackbarth, and Beth Gipsky were the data-entry operators. J. Trent Alexander, Jason Carl Digman, and J. David Hacker cleaned the data, constructed and coded variables, designed the record layout, and wrote the user's guide.

1. There is strong evidence that the free African American population in the antebellum South was also becoming more “light-skinned” than it had been earlier (Berlin 1974; Williamson 1980).

2. No manuscript records survive for five counties: Hernando, Florida; Bienville, Louisiana; Hancock, Mississippi; Blanco, Texas; and Tarrant, Texas. The Census Office estimated the total slave population in these counties to be 7,014 (less than 0.2 percent of the total slave population in 1860).

3. No attempt was made in the pilot study to obtain a representative sample. We purposely selected reels containing the records of Washington, D.C., and Richmond, Virginia, to make certain we encountered small slaveholdings with multiple owners and owner types. In toto, the complete-count data set included information on 539,509 slaves in a variety of slaveholding counties. Because of user interest, we will release the data for these counties in the near future. Users should be aware, however, that these data are nonrepresentative of the national slave population.

4. Occasionally, census marshals noted which individual slaves in the holding were fugitive from the state and which slaves were manumitted in the previous year. Although we have retained this individual-level detail in

the data set, users should be aware that this information was not systematically recorded.

REFERENCES

- Berlin, I. 1974. *Slaves without masters: The free Negro in the antebellum south*. New York: Pantheon Books.
- Blassingame, J. W. 1972. *The slave community: Plantation life in the antebellum south*. New York: Oxford University Press.
- Douglass, F. 1845. *Narrative of the life of Frederick Douglass, an American slave, written by himself*. Boston: American Anti-Slavery Society.
- Elkins, S. M. 1976. *Slavery: A problem in American institutional and intellectual life*. Chicago: University of Chicago Press.
- Fields, B. J. 1984. *Slavery and freedom on the middle ground: Maryland during the nineteenth century*. New Haven: Yale University Press.
- Fogel, R. W. 1989. *Without consent or contract: The rise and fall of American slavery*. New York: W. W. Norton.
- Fogel, R. W., and S. L. Engerman. 1974. *Time on the cross: The economics of American Negro slavery*. Boston: Little, Brown.
- Genovese, E. D. 1972. *Roll, Jordan, roll: The world the slaves made*. New York: Vintage Books.
- Goldin, C. D. 1976. *Urban slavery in the American south, 1820–1860: A quantitative history*. Chicago: University of Chicago Press.
- Gutman, H. G. 1976. *The black family in slavery and freedom, 1725–1925*. New York: Vintage Books.
- Hacker, J. D., S. Ruggles, A. R. Foroughi, B. D. Jarvis, and W. L. Sargent. 1999. Public use microdata samples of the 1860 and 1870 U.S. censuses of population. *Historical Methods* 32: 125–33.
- Kay, M. L., and L. L. Cary. 1995. *Slavery in North Carolina, 1748–1775*. Chapel Hill: University of North Carolina Press.
- Kolchin, P. 1993. *American slavery, 1619–1877*. New York: Hill and Wang.
- Malone, A. P. 1992. *Sweet chariot: Slave family and household structure in nineteenth-century Louisiana*. Chapel Hill: University of North Carolina Press.
- McClelland, P. D., and R. J. Zeckhauser. 1982. *Demographic dimensions of the new republic: American interregional migration, vital statistics, and manumissions, 1800–1860*. Cambridge: Cambridge University Press.
- McDaniel, A., and C. Grushka. 1995. Did Africans live longer in the antebellum United States? The sensitivity of mortality estimates of enslaved Africans. *Historical Methods* 28: 97–105.
- Menard, R. R. 1975. The Maryland slave population, 1658 to 1730: A demographic profile of blacks in four counties. *William and Mary Quarterly* 32: 29–54.
- . 1995. Slave demography in the low country, 1670–1740: From frontier to plantation regime. *South Carolina Historical Magazine* 96: 280–303.
- Parker, W. N., and R. E. Gallman. 1976. *Southern farms study, 1860* [computer file]. Ann Arbor, Mich.: Inter-university Consortium for Political and Social Research [producer and distributor].
- Ruggles, S. 1995. Sample designs and sampling errors. *Historical Methods* 28: 40–46.
- Ruggles, S., and M. Sobek. 1997. *Integrated public use microdata series: Version 2.0*. Minneapolis: Historical Census Project, University of Minnesota.
- Shore, L. 1986. The poverty of tragedy in historical writing on southern slavery. *South Atlantic Quarterly* 85: 147–64.
- Starobin, R. S. 1970. *Industrial slavery in the old south*. New York: Oxford University Press.
- Steckel, R. H. 1980. Miscegenation and the American slave schedules. *Journal of Interdisciplinary History* 11: 251–63.
- . 1985. *The economics of U.S. slave and southern white fertility*. New York: Garland.
- U.S. Census Office. 1860. *Eighth census of the United States, instructions to enumerators*. Washington, D.C.: George W. Bowman.
- Wade, R. C. 1964. *Slavery in the cities: The south, 1820–1860*. New York: Oxford University Press.
- White, D. G. 1985. *Ar'n't I a woman? Female slaves in the plantation south*. New York: W. W. Norton.
- Williamson, J. 1980. *New people: Miscegenation and mulattoes in the United States*. New York: Free Press.
- Wyatt-Brown, B. 1988. The mask of obedience: Male slave psychology in the old south. *American Historical Review* 93: 1228–52.