Science
AAAS

Supplementary Materials for

The U.S. Census Bureau's Ex Post Confidentiality Analysis of the 2010 Census Data Publications

Authors: Authors: John M. Abowd[1]*, Tamara Adams[1], Robert Ashmead[1], David Darais[2], Sourya Dey[2], Simson L. Garfinkel[3]†, Nathan Goldschlag[1], Daniel Kifer[4], Philip Leclerc[1], Ethan Lew[2], Scott Moore[2], Rolando A. Rodríguez[1], Ramy N. Tadros[5]‡, Lars Vilhuber[1]§

Corresponding author: john.maron.abowd@census.gov

The PDF file includes:

Materials
Supplementary Text
Figs. S1
Tables S1 to S13
References 43 to 63

Other Supplementary Materials for this manuscript include the following:

Data S1 to S4

Materials

We provide a high-level overview of the databases that underlie the confidential and published versions of the 2010 Census. We start by describing the internal, confidential Census Bureau databases from which the public Summary File 1 is created, then describe the Summary File 1 (SF1) itself. The process of creating SF1 from the internal databases is illustrated in Figure S1. We also describe the features and limitations of the commercial databases used in the analysis. We deliberately abstract from some of the complexity of these databases to focus on the characteristics that are salient to our reconstruction and reidentification attacks.

2010 Census Internal Databases

The U.S. Constitution mandates that a census of population be conducted every ten years. Since the 1970 Census, these enumerations have collected primarily self-reported information on households and the individuals in those households. For the 2010 Census, the confidential data are stored in several databases. The raw responses are stored in the 2010 Census Unedited File (CUF), which contains features collected directly from respondents and nonresponse follow-up fieldwork. There are different record types in the CUF; however, this paper uses only the person record and the associated address from the housing unit or group quarters record. The 2010 Census Edited File (CEF) constitutes the final, fully edited, permanent electronic record of the responses to the 2010 Census. The application of confidentiality protections, primarily household swapping of geographic identifiers, group quarters synthetic data, and tabulation recodes applied to the CEF produces the confidential Hundred-percent Detail File (HDF). It is the HDF microdata that are aggregated to create the SF1. Importantly, residential geographic location is coded to the 2010 Census tabulation schema described below.

While the feature sets for the full hierarchical CUF, CEF, and HDF are larger, salient features of the CUF for the person table are:

$$P_{\text{CUF}} := \{\textit{name}, \textit{address}, \textit{sex}, \textit{age}, \textit{race}, \textit{ethnicity}\}.$$

As part of the internal confidentiality safeguards and to support person-level record linkage with other Census Bureau data assets, the *name* feature was replaced with a unique person identifier called a Protected Identification Key (PIK) using the production household record-linkage system called the Person Identification Validation System (PVS). The PVS has evolved over time. The application of the production record linkage was completed contemporaneously with the 2010 Census data processing using the 2010-vintage version of the PVS. For details on PIK assignment, see *(33)*. The *address* feature was replaced by a unique internal address identifier called the MAFID.

The use of the Census Bureau's production record-linkage system, and the selection of the 2010 Census vintage allowed us to do record linkage on name and address without having to design our own linkage system, which is outside the scope of this project. We ensured that the same vintage of PVS was used for the commercial data we discuss later in this section. If the PVS recognized a person in the 2010 Census, it is extremely likely that the same vintage of the software would recognize the same person in the commercial data, thus assigning the same PIK.

All record linkage systems are subject to false positive and negative linkages. By employing the same system on all data used for this paper, we accept linkages based on PIKs with the error properties denoted in *(33)*.

Not all records in the CEF have a PIK. In some cases, the same PIK appears on multiple records because the PVS was not designed to unduplicate the input data set. For example, within members of the same household, incompleteness in name and birth date information can lead to the PVS assigning the same PIK to two or more of these persons. For the purposes of this paper, we refer to the subset of records in the CEF with a distinct PIK within the record's census block as the data-defined population. To create the data-defined population, if there were multiple records with the same PIK within a block, one record was randomly chosen. The data-defined population is 276.0 million records (89.4% of all person records in the CEF). Records with duplicate PIKs within a single block appeared in 15% of blocks. In total, 1.0% of records with a PIK were removed for duplication. The remainder of the difference between the total population and the data-defined population are incomplete or imputed census responses for which the PVS cannot assign a PIK; these include all whole-person census imputations—individuals from whom no response data were collected at the address—and some partial responses—individuals who supplied too few of the *name*, *sex* and *birthdate* features required by PVS. The SF1 tables contain sufficient block-level information to reconstruct the variables defining whole-person census imputations and incomplete sex and age information, but such an exercise is outside the scope of our study because it would only confirm that the 10.6% of the population that is not data-defined was not at risk for a record linkage attack. Instead, we assume that is the case.

The MAFID is geocoded into the 2010 Census tabulation geography. In this paper, we distinguish two components of the 15-digit tabulation geography—census tract (11-digit concatenation of FIPS state, county equivalent, and tract) and census block (15-digit concatenation of FIPS state, county equivalent, tract, and block). FIPS stands for Federal Information Processing Standards (FIPS) and refers to numeric and two-letter alphabetic codes defined in U.S. FIPS Publication 5-2. FIPS 5-2 was superseded by ANSI standard INCITS 38:2009. For details, see *(43)*. A census block is a statistical definition of geography, not the commonly used "city block," with complete coverage of the entire territory of the United States. Census blocks and are the atoms in the Census Bureau's geographic lattice that are used to build all other geographic summary levels, such as census tracts or counties. Census blocks are defined in terms of territory, not population, and tessellate the entire United States. Some blocks may therefore be uninhabited (even underwater), others may have a very large population. See *(44)* for an overview. For more details on the person and geography features, see *(34)*. Subject to these qualifications the person tables in CEF and HDF have the feature sets:

$$P_{\text{CEF}} := \{pik, block, sex, age, race, ethnicity\} \text{ and}$$

$$P_{\text{HDF}} := \{block, sex, age, race, ethnicity\}.$$

In our evaluations, we make comparisons with both the CEF and HDF to show the success of reconstruction and the degree to which the distortions due to statistical disclosure limitation (SDL) applied to the HDF impact the results. The confidential databases share the same schema and feature sets: one column for the census block, one column each for the person

identifier (PIK, appears only in the CEF), sex, age, and ethnicity (Hispanic/Latino or not Hispanic/Latino), and six columns for the 1997 Office of Management and Budget (OMB)-defined race categories. For details on the person attributes, see *(34)*. For background on the definitions of ethnicity and race categories, see *(45)*. The category "some other race" is mandated by law, not statistical policy. Persons may self-declare multiple race categories; hence the binary race variables are not mutually exclusive. In practice, most 2010 Census respondents only identified with a single race. Age is recorded in integer values. If a census response is missing, the process that creates the CEF performs imputation. There are no missing data in the CEF and, in particular, at least one of the six race categories must be selected. Excluding PIK, which is standing in for name, we define all valid combinations of *block*, *sex*, *age*, *race*, and *ethnicity* as the feature space for CEF and HDF, χ. There are approximately 161×10^9 (161 billion) such combinations which gives the cardinality $|\chi|$. Cardinality $|\chi| = 161{,}109{,}592{,}812 = 6{,}207{,}027 \times 2 \times 103 \times 63 \times 2$, where 6,207,027 is the number of inhabited blocks in the 2010 Census, 103 is the number of single-year age categories 0 to 99 plus grouped ages 100-104, 105-109, 110+ allowed in the published tables, and 63 is the number of allowable race combinations. Note that, for technical reasons, when we implement the reconstructions of HDF, we modify the feature set to eliminate the age binning for ages 100+.

2010 Summary File 1

The most extensive and widely used 2010 Census data product is Summary File 1 *(34)*. SF1 contains counts of persons, households/families, group quarters residents, and housing units tabulated at the census block, tract, and county-equivalent geographic levels. SF1 also includes the tables released separately as the Redistricting Data (P.L. 94-171) Summary File and Advanced Group Quarters Summary File, which form the basis for redistricting every legislative body in the United States and are normally released by March 31st of the year following the decennial census, several months before SF1. The populations used for apportionment include a limited number of U.S. citizens and their families living abroad. These persons do not have records in the CEF, and their total for each state is added to the resident population for that state prior to apportionment, see Appendix G of *(34)*. The 2010 Redistricting Data (P.L. 94-171) Summary File and 2010 Advanced Group Quarters Summary File tables are renumbered in 2010 SF1 but otherwise identical to the original release.

The SF1 and other published tables are created by tabulating the HDF according to various combinations of geographic and demographic detail. All published data from the 2010 Census used the same geographic hierarchy. See Appendix A of *(34)* for more details. The census block is the most detailed geographic category. There were 11,078,297 blocks defined for 2010 of which 6,207,027 were occupied. Whether a block was inhabited or not was published without any statistical disclosure limitation in 2010. The census blocks aggregate into 73,057 defined census tracts of which 72,531 were inhabited. The census tracts, in turn, aggregate into 3,143 county equivalents, all of which were inhabited.

Within this hierarchy, tables of varying demographic and household detail are created. In this paper, we focus on census block and census tract-level tabular summaries of persons using only the tables shown in Table 2. The block-level tables, labeled Px in Panel A of Table 2, provide detailed information on sex and race, but with coarsened age information for those age

22 and over. The tract-level summaries, labeled PCTx in Panel B of Table 2, report most of the detail in block-level tables in addition to reporting more detailed age. The tract-level schema for race is less detailed than the block-level schema; however, this does not affect our reconstructions because we never use the tract-level data alone, and its race and ethnicity schema is nested within the block-level schema. We note for completeness that the 2010 Census Public Use Microdata Sample (PUMS) was also created from the HDF. The confidential 2010 HDF itself, but not the 2010 CEF, can be used by external researchers with approved projects in the Federal Statistical Research Data Centers.

The Treatment of Age in Summary File 1

The 2010 SF1 tabulated age differently depending on the specific table and the level of geographic detail. At the census tract level and above (e.g., SF1 PCT12), age was tabulated in single years from 0 to 99 years, then binned into 100-104 years, 105-109 years, and 110 years and over. At the block level in most tables (e.g., SF1 P12) age was binned into the following ranges: 0-4; 5-9; 10-14; 15-17; 18-19; 20; 21; 22-24; 25-29; 30-34; 35-39; 40-44; 45-49; 50-54; 55-59; 60-61; 62-64; 65-66; 67-69; 70-74; 75-79; 80-84; and 85+. Also at the block level, SF1 P10 and P11 selected only persons ages 18 and older. Finally, the block-level SF1 P14 selected only individuals ages 20 or younger and encoded age in single years. Combining the different age binning and universe selection rules applied at the block level defines the most detailed age schema that these tables can support. That schema has 38 age groups: single year of age from 0 to 21, then: 22-24; 25-29; 30-34; 35-39; 40-44; 45-49; 50-54; 55-59; 60-61; 62-64; 65-66; 67-69; 70-74; 75-79; 80-84; and 85+. We use this 38-bin age schema, designated as the feature *agebin* in the main text, as well as the exact *age* feature in our assessments of agreement of the reconstructed HDF with the HDF and CEF. We show in the reidentification experiments that the 38-bin age schema provides sufficient uniqueness for persons at the block level to enable reconstruction-abetted reidentification at scale. Resolution to single years of age is unnecessary for this purpose. In our matching algorithms, we distinguish between matches based on exact age, denoted by the feature *age*, and those based on binned age.

Circa 2010 Commercial Databases

The commercial database (COMRCL) used for reidentification experiments was created by combining data extracts originally purchased in support of the 2010 Census evaluations from four commercial providers between 2009 and 2011. The four commercial databases were from Experian Marketing Solutions Incorporated, Infogroup Incorporated, Targus Information Corporation, and VSGI LLC. The databases used are the same as in *(33)* except that we did not use a fifth commercial database from the Melissa Data Corporation for technical reasons. These commercial databases serve as the background knowledge of the attacker. While the database schema and the purposes for which these data were originally collected vary, they all share certain features. All have basic personally identifiable information (PII) including names, addresses, sex, and birth dates. The vintage 2010 versions of these databases that we used did not include self-reported race and ethnicity data. Race and ethnicity data are modeled in some of these commercial databases *(35)*.

We harmonized the feature sets for the commercial data to

$$P_{\text{COMRCL}} := \{pik, block, sex, age\},$$

and harmonized the schema for each variable to match the schema defined for the CEF. Name and address were mapped to PIK and MAFID, respectively. The MAFIDs were originally geocoded in 2009-2011, when these commercial databases were acquired, because the final 2010 tabulation geography schema was not available at that time. We remapped the MAFID to final 2010 Census tabulation blocks in early 2019. PII was standardized and mapped to PIK using the same 2010-vintage PVS that was used for the 2010 CEF. Table S1 shows that there were 289,100,000 records with a valid {*pik*, *block*, *sex*, *age*} in the commercial database. Among those commercial database records, only 106,300,000 (or 37%) matched a CEF record on {*pik*, *block*, *sex*, *agebin*}. See Table S1 for details.

Supplementary Text

Supplementary Methods Details

Reconstruction Overview

We define database reconstruction as any attempt to re-create the record-level image of the database from which a set of published tabulations were originally calculated; that is the confidential HDF in this case. Database reconstruction attempts to reverse engineer the confidential HDF records that were the input data used in a tabulation system with the goal of making the reconstruction as close as possible to these confidential data. The reconstruction described here is not the most powerful reconstruction that is possible. Only a relatively small subset of SF1 tables was used, there was no attempt to reconstruct households or relationships within a household, and we did not use statistical modeling to improve reconstruction. Because block populations and the age indicator "less than 18 years old" versus "age 18 and older," called "voting age" in the SF1 documentation *(34)*, were published as enumerated, every reconstruction starts with the correct number of total records 308,745,538, and each one of these records has the correct census block identifier and voting age variable. There is never any solution variability on those features. We did not attempt to elaborate the reconstruction to undo the geographic identifier swapping.

For our experiment, we used data for the 50 states and the District of Columbia. Use of the term "state" in this document refers to these 51 political divisions. Puerto Rico was excluded from these experiments because the Census Bureau's production name and address record linkage system does not work as well for this commonwealth. We used a subset of the statistical tabulations published in SF1, namely tables labeled Px and PCTx as shown in Table 2 for any universe that was part of the total population. We did not use tables where the universe was households, which means that we did not use the "relationship to the householder" information to reconstruct households in addition to persons. The tables shown in Table 2 in the main text, computed from the HDF person table, are multidimensional marginal counts related to age, sex, race, and ethnicity by census block and tract. Our reconstructed microdata contain these same variables and only these variables, and in the main text we denote them with the feature set: {*block, sex, age, race, ethnicity*}. The reconstructed data, therefore, necessarily reflect the schemas used for SF1 and are only informative about variables, in particular age, in the schemas used for publication, as described in materials section on the treatment of age in SF1.

Our experiments used two different subsets of SF1 data as shown in Table 2. The first reconstruction, which we denote as $\text{rHDF}_{b,t}$, uses both tract and block-level SF1 tables (Panels A and B of Table 2), taking advantage of both the geographic and racial detail in the block summaries and the age detail in the tract summaries. The second reconstruction, rHDF_{b}, uses only the block-level tables (Panel A of Table 2). Thus, the second reconstruction removes the more granular age information found in the tract-level tables while retaining the full race and ethnicity schema used in the block-level data. The comparison of results from $\text{rHDF}_{b,t}$ to those from rHDF_{b} tests two effects: (1) the loss of precision in the reconstruction from removing some tables and (2) the imprecision from using only aggregated age groups in the reidentification. Given the structure of our reconstruction algorithm, adding sets of tables is substantially more difficult than removing tables from the subset of SF1 tables used in our most detailed reconstruction, $\text{rHDF}_{b,t}$; however, they are logically equivalent. Therefore, comparing rHDF_{b} to $\text{rHDF}_{b,t}$ shows the effects of adding tract-level tables to the reconstruction for variables supported by either schema.

We first present our reconstruction model using linear algebra. This representation is useful for understanding the structure of the problem. Then, we provide a description of the integer program (IP) setup used to generate the solutions, which does not convey the high-level structure of the problem as simply but closely follows our software implementation of the reconstruction. The inputs to the reconstruction are the database schema for the tabulation features P_{HDF} := {*block, sex, age, race, ethnicity*}; the vector of all published statistics in the appropriate order; and the matrix workload that maps the histogram representation of the HDF onto the published statistics. Note that in the Census Bureau's geographic entity identifier, the code for a census tract is embedded in the code for a census block as the first 11 characters of the 15-character block code. We represent two-dimensional tables of the record-level data as the vectorized fully saturated contingency table (alternatively called a histogram in computer science) where every cell corresponds to a possible record type in χ and its value is the number of records of that type. Thus, instead of a multidimensional array (common for contingency tables), the contingency table is flattened into a vector in some arbitrary predefined order. Let $\mathbf{x}$ represent the contingency table vector. Database reconstruction consists of finding at least one non-negative integer solution for $\mathbf{x}$ in the equation system

$$\mathbf{Ax} = \mathbf{c}, \qquad x_i \in \mathbb{Z}^+ \text{ for all } I, \tag{1}$$

where $\mathbf{c}$ is $K \times 1$ column vector of the $K = 5.0$ billion statistics extracted from SF1 for a given reconstruction, $\mathbf{A}$ is the matrix that defines a set of linear queries on $\mathbf{x}$ resulting in $\mathbf{c}$, and $\mathbb{Z}^+$ is the set of non-negative integers. Each row of $\mathbf{A}$ and each component of $\mathbf{c}$, therefore, represent a single query and a single statistic from SF1, respectively.

Although there are fewer statistics published per block than points in the sample space ($|\chi| > K$), in many small blocks, there are very few non-zero entries in the many published tables suggesting that the system may not be as underdetermined as it appears. In our IP setup, we modify slightly the definition of χ to allow more flexibility in setting up the integer programs that accommodate the multiple age schemas discussed above. For the purposes of this high-level explanation, coarsening of the most general SF1 age schema used to define χ is accomplished by

the matrix **A**. This means that in some cases there is a unique non-negative, integer-valued solution to equation system (1), or at least blocks of **A** that have a unique solution. There may also be blocks that have multiple solutions. SF1 statistics that take the value zero often eliminate many infeasible solutions. We are guaranteed that at least one solution exists because SF1 was tabulated from a single real database with the schema encoded in χ. Because the x_i are integer-valued, we used IP algorithms in Gurobi™ to produce solutions. See below for details. Given a solution $\hat{x}$ to equation (1), another question of interest is the variability of other solutions. If $\hat{x}$ is the only solution, then it represents an exact reconstruction of the microdata used to tabulate SF1, namely HDF, with certainty. Moreover, if the absence of solution variability can be determined from the published inputs to equation (1), then this fact implies that those reconstructed records were in HDF. To examine how often the solution $\hat{x}$ was strongly constrained, we built a second IP model and solved it to search for a second solution $\tilde{x}$ that maximized the L_1 distance to $\hat{x}$. This problem is described below.

Reconstruction Integer Programs

In this section we describe how the IP is implemented as a mathematical programming problem. We begin with the basic schema used to define the binary variables that form the solution space for our integer programs in both the $\text{rHDF}_{b,t}$ and rHDF_b reconstructions:

Reconstruction Schema with Exact Age (2)

W = {White = 1, Not White = 0}
BL = {Black or African American = 1, Not Black or African American = 0}
AIAN = {American Indian or Alaskan Native = 1,not American Indian or Alaskan Native = 0}
ASIAN = {Asian = 1, n Asian = 0}
NHOPI = {Native Hawaiian or Other Pacific Islander = 1, not Native Hawaiian or Other Pacific Islander = 0}
SOR = {Some Other Race = 1, not Some Other Race = 0}
HISP = {Hispanic/Latino = 1, not Hispanic/Latino = 0}
AGE = {0, 1, 2, ..., 110}
SEX = {Male = 0, Female = 1}.

The notation in equation (2) describes nine feature sets that expand the basic five used in the main text to facilitate building binary variables for the integer programs: W, BL, AIAN, ASIAN, NHOPI, SOR, HISP, SEX, and AGE. The elements of these feature sets are the values allowed for each feature.

For the demographic variables we denote the indices with lower case letters shown here in the same order as the feature sets of equation (2) as (*w, bl, aian, asian, nhopi, sor, hisp, a, s*). In a block *b*, we create a binary variable for each potential record. Since there could be many records with the same demographic type, we create multiple such variables $B_{(i,b,w,bl,aian,asian,nhopi,sor,hisp,a,s)}$ for $i = 0, 1,2, \ldots$. To determine how many binary variables of each type to create, we use SF1 P12 (age group by sex for each census block, see Table S2 for the age groupings used in SF1 P12). For example, for the demographic type of a 22-year-old Asian Hispanic female in a block *b*, SF1 P12 for that block might say there are 50 females in the age group 22-24; therefore 50 is an upper bound on the number of potential records for 22-year-old Asian Hispanic females in the block. Thus, we create the 50 variables $B_{(i,b,w=0,bl=0,aian=0,asian=1,nhopi=0,sor=0,hisp=1,a=22,s=1)}$ for $i = 0,\ldots,49$. These variables are binary, with a

value of 1 indicating the presence of the potential record in a candidate microdata reconstruction. Once the binary variables are assigned values, the summation

$$\sum_{i=0}^{49} \mathcal{B}_{(i,b,w=0,bl=0,aian=0,asian=1,nhopi=0,sor=0,hisp=1,a=22,s=1)}$$

represents the number of 22-year-old female Asian Hispanic people in that block. More formally, these variables are created as follows.

- Define the SF1 P12 binning function AGEBINP12, as in Table S2, such that for any age $a \in$ AGE, AGEBINP12(a) returns the bin containing that age from SF1 Table P12.
- For any age a, sex s, and block b, define the constant $c^{P12}_{b,AGEBINP12(a),s}$ to be the count in SF1 P12 of the number of people in block b, sex s, and age that is in the same age bin as age a. This is the upper bound discussed above.
- For every block b, age a, sex s, and value of the race/ethnicity variables *w*, *bl*, *aian*, *asian*, *nhopi*, *sor*, *hisp*, define the binary variables:

$$\mathcal{B}_{(i,b,w,bl,aian,asian,nhopi,sor,hisp,a,s)} \in \{0,1\} \qquad (3)$$

$$i = 0, \ldots, c^{P12}_{b,\mathrm{AGEBINP12}(a),s} - 1$$

We emphasize that SF1 P12 is not coarsening our age schema. It is used to determine the maximum number of variables the integer program needs for each demographic type in the full schema.

Let T represent the set of all tract indices, and B_t represent the set of all block indices in tract t. Notice that lowercase italic letters are used for indices, lowercase letters as well as the numbers 0 and 1 for elements of the feature sets, and capital letters for set and function names. To further understand the structure of these variables, refer to the file recon/s3_pandas_synth_lp_files.py in the replication archive that defines the inputs to the Gurobi™ solver. Functions makeSimpleProductConstraints and makeTwoPlusRaceConstraints are responsible for identifying the binary variable indices relevant to constraints for representing a target tabulation, using summary information initially computed in the function get_constraint_summary.

Each SF1 table adds additional constraints on the $B_{(i,b,w,bl,aian,asian,nhopi,sor,hisp,a,s)}$ variables. For example, consider the tract-level SF1 PCT12I, which encodes the tabulation sex by age group in each tract for people who are "White Alone" and "Not Hispanic or Latino." The age binning used by this table is {0, 1, ..., 99,100−104,105−109,110+}, so let AGEBINPCT12I(a) = z be the function that returns the age bin z for a given age a. For each tract t, sex s, and PCT12I age bin z let $c^{PCT12I}_{t,s,z}$ be the corresponding count in table PCT12I. Then for each t, s, z we add the following constraint:

$$\sum_{b \in B_t} \sum_{\mathrm{a:\ AGEBINPCT12I}(a)=z} \sum_i \mathcal{B}_{(i,b,w=1,bl=0,aian=0,asian=0,nhopi=0,sor=0,hisp=0,a,s)} = c^{PCT12I}_{t,s,z} \quad (4)$$

where the summation over i uses the upper bound on the number of $\mathcal{B}$ variables constructed as in equation (3), i.e., $i = 0, \ldots, c^{P12}_{b,\text{AGEBINP12}(a),s} - 1$.

We use the shorthand $\mathcal{T}_t[tabname] = c_t^{tabname}$ to represent all such constraints created by SF1 tract-level table *tabname* (e.g., PCT12I) for tract *t*. In the case of PCT12I, this means an application of equation (4) for each age bin *z* and sex *s*. The tract-level tables are listed in Panel B of Table 2. Similarly, we use the shorthand $\mathcal{T}_b[tabname] = c_b^{tabname}$ for the block-level constraints created by SF1 block-level table *tabname* (Panel A of Table 2). The IP for rHDF$_{b,t}$ is given by

$$\begin{aligned} &\max 0 \\ &s.t. \mathcal{T}_t[tabname] = c_t^{tabname} \ \ \forall\, tabname \in \text{Panel B Table 2} \\ &\quad \mathcal{T}_b[tabname] = c_b^{tabname} \ \ \forall\, tabname \in \text{Panel A Table 2} \ \ \forall b \in \text{B}_t\,. \end{aligned} \tag{5}$$

The "max 0" indicates that any feasible solution that satisfies the constraints should be returned; i.e., no statistical modeling should be used to return the most plausible solution if multiple feasible solutions exist. The optimization variables are the $\mathcal{B}_{(i,b,w,bl,aian,asian,nhopi,sor,hisp,a,s)}$ defined above. Once a feasible solution is obtained, for every $\mathcal{B}_{(i,b,w,bl,aian,asian,nhopi,sor,hisp,a,s)}$ that is set to 1 a corresponding record for that block and demographic type is added to the reconstructed dataset. Thus, upon completion, the IP from equation (5) yields a reconstructed HDF that contains one record for each person in the 2010 Census with the following features:

$$P_{\text{rHDF}_{b,t}} := \{block, sex, age, race, ethnicity\},$$

where *age* is single year of age 0 , ... ,110 (exact age), *race* is defined by concatenating the six variables *w*, *bl*, *aian*, *asian*, *nhopi*, *sor*, and *ethnicity* = *hisp*. Because in SF1 persons 100 years of age or older were always tabulated into these age groups: 100-104, 105-109, and 110 and over, single-year-of age tabulations for ages greater than 99 are never available. The IP used for the rHDF$_{b,t}$ reconstruction schema nevertheless finds solutions for the single years of age 0,...,110, which means that there is inherent solution variability for the oldest sub-populations.

Next, we present the IP for the reconstruction rHDF$_b$, which uses only SF1 block-level tables shown in Panel A of Table 2. Block-level tables use several age binning schemes, but their intersection is not exact age. Instead, their intersection is the age binning shown in Table S3, which has 38 age bins (feature *agebin* in the main text). Since this is the most fine-grained age resolution that can possibly be obtained from block-level tables, rHDF$_b$ only has age reconstructed up to this age binning.

For programming purposes, we re-use, the same $\mathcal{B}_{(i,b,w,bl,aian,asian,nhopi,sor,hisp,a,s)}$ variables defined above for the tract problem. In particular, the binary variables use exact age. After the reconstructed microdata are created, we recode the *age* feature into the 38-bin *agebin* feature. The IP for the block-level reconstruction is like the tract-level reconstruction. For each block *b*, solve

$$\max 0 \tag{6}$$

$$s.t. \mathcal{T}_b[tabname] = c_b^{tabname} \ \forall\, tabname \in \text{Panel A of Table 2.}$$

Again, no statistical modeling is used to return the most plausible solution if multiple feasible solutions exist. The optimization variables are the $\mathcal{B}_{(i,b,w,bl,aian,asian,nhopi,sor,hisp,a,s)}$ defined in equation (3). Once a feasible solution is obtained, for every $\mathcal{B}_{(i,b,w,bl,aian,asian,nhopi,sor,hisp,a,s)}$ that is set to 1 a corresponding record for that block and demographic type is created, the age is binned to the appropriate bin in Table S3, and the resulting record is added to the reconstructed dataset rHDF_b. Thus, the reconstruction produces results in records with the feature set:

$$P_{\text{rHDF}_b} := \{block, sex, agebin, race, ethnicity\},$$

where the variable *agebin* is encoded as *agebin* = AGEBINBLOCK(*a*) for $a \in \{0, ..., 110\}$ defined in Table S3.

We note that the IP description of the reconstruction model is mathematically equivalent to the algebraic model in the overview. The only difference is that the variables $\mathcal{B}_{(i,b,w,bl,aian,asian,nhopi,sor,hisp,a,s)}$ in the IP implemented in our code correspond to potential records, while the vector **x** in equation (1) is the vectorized fully saturated contingency table. The code in the replication archive implements the integer programs in equations (5) and (6), and these equations are useful for reading the software implementation in the replication archive, whereas the histogram representation in equation (1) is useful for understanding the high-level structure of the problem. For reconstruction outcomes, only run-time, not the space of feasible solutions (except in syntactic form), is affected by the choice of which representation to implement. It is not obvious *a priori* which representation should solve more quickly. We have direct experience only with solving the IP representation, and we found it consistently solved very quickly at default Gurobi™ settings for the set of tabulations we implemented. Equation (1) is succinct, easy to represent in any matrix programming language that implements sparse matrix storage and Mixed Integer Linear Programming (MILP) solvers and yields a model with considerably fewer variables. On the other hand, the IP representation, while it produces less succinct models, uses only binary variables in the solution set, rather than general integers, and binary variables are usually processed more efficiently in modern MILP solvers. We switch between the two representations to permit clarity of expression (equation (1)) versus fidelity to the details of our implemented reconstruction code (equations (5) and (6)).

Solution Variability Model

There can be multiple, distinct $\text{rHDF}_{b,t}$ databases that satisfy the constraints imposed by the published tables. As such, it is useful to consider the uniqueness of records for a given reconstruction. With a uniqueness measure, one could determine bounds on the level of confidence associated with the likelihood that reconstructed feature values of any given record correctly reflect a record in the HDF. If the reconstructed solution for a given block is unique– there exists only one set of microdata records consistent with the published tables for that block– then we can be certain that the resulting microdata exactly match the HDF. An attacker would also be more confident about linking these reconstructed records to other data sources, since

swapping and record-level synthetic data are the only remaining disclosure avoidance techniques that could cause these reconstructed HDF records to differ from their CEF counterparts.

For any reconstruction rHDF and geographic region g, such as a specific census tract or block, let rHDF(g) represent the subset of records in rHDF that belongs to geographic region g. If there are two feasible reconstructions, rHDF* and rHDF$'$, we consider rHDF*(g) and rHDF$'$(g) to be equivalent for region g if the only difference is the ordering of the records; that is, if one is a permutation of the other. Two reconstructions are distinct on g if and only if their corresponding fully saturated contingency tables for region g are different. Let Hist($\cdot$) represent the operator that converts a reconstruction into a fully saturated contingency table (also known as a histogram). Thus, two feasible reconstructions rHDF*(g) and rHDF$'$(g) in region g are distinct whenever Hist(rHDF*(g)) and Hist(rHDF$'$(g)) are different.

Let i = 1, ..., k index the k cells of the histograms. Measure the difference between two histograms using the L_1 norm:

$$L_1\left(\text{Hist}\left(\text{rHDF}^*(g)\right), \text{Hist}\left(\text{rHDF}'(g)\right)\right) = \sum_{i=1,\ldots,k} \left|\text{Hist}\left(\text{rHDF}^*(g)\right)_i - \text{Hist}(\text{rHDF}'(g))_i\right|.$$

Note that the cells depend on the reconstruction schema. For example, when reconstructing rHDF$_b$, the histograms use the 38-bin age grouping shown in Table S3 (feature *agebin* in the main text). Also note that L_1(Hist(rHDF*(g)), Hist(rHDF$'$(g))) $\leq 2N_g$, where N_g is the total population of geographic unit g, and equality is achieved if and only if the two histograms completely disagree on the types of records present in g. The $2N_g$ bound holds because the total population in the block is one of the constraints in IP (5) and (6) as the overall margin of SF1 table P12. So, both solutions must have the same total population. On the other hand, the L_1 norm is 0 if and only if the two histograms agree exactly on the types of records present and on their multiplicities.

Given a feasible solution rHDF*, we define solution variability (the statistic *solvar* in the main text) for geographic unit g relative to solution rHDF* as the largest value of

$$L_1\left(\text{Hist}\left(\text{rHDF}^*(g)\right), \text{Hist}\left(\text{rHDF}'(g)\right)\right) / \left(2N_g\right) \qquad (7)$$

among all feasible solutions rHDF$'$ to the IP in equation (5) (when g is a tract) or equation (6) (when g is a block) for an arbitrary rHDF* solution.

When solution variability in equation (7) is 0, there is a single, unique solution to the reconstruction problem in geographic unit g, and the rHDF* records in g must exactly match the HDF for the variables present in records of both data sets and under the schema used to create rHDF*. When solution variability in equation (7) is 1, the rHDF* is a poor reconstruction of HDF, and any agreement between the two may be happenstance. Since solution variability is strictly bounded between 0 and 1, it can be interpreted as the percentage of records in rHDF* that could possibly have changed if a different solution to the reconstruction IP had been used. For example, if the solution variability is 0.1, then at most one-tenth of the records in rHDF* could change if an alternative solution to the reconstruction problem in either equation (5) or (6) were

used, relative to the schema used to encode the variables of the reconstruction IP. Computing the solution variability in geography *g* is a harder problem than solving for an initial rHDF*. To control the run-time costs of these problems, we worked over census blocks rather than tracts to limit the problem size. That is, rather than using $\text{rHDF}_{b,t}$ as the base solution, we used rHDF_b (though we retained tract-level constraints in the solution variability constraint set, converted from equalities to inequalities). This implies that solution variability is measured with respect to the 38-bin age groupings in Table S3 (feature *agebin* in the main text). A solution variability of 0 therefore means that a block has a unique solution in terms of the binned age schema, not the exact age schema shown in equation (2). We would have to solve analogous but larger and more difficult optimization problems to attain exact-age solution variability results.

Changing from $\text{rHDF}_{b,t}$ to rHDF_b to assess solution variability at the block-level requires converting tract-level tabulations to inequality constraints and yields a looser bound on the underlying variability. However, block-level solution variability still upper bounds the maximum distance between rHDF* and any feasible, alternative solution rHDF′, for the binned-age schema in use. When solution variability is 0, this block must have a unique reconstruction solution. A technical caveat is appropriate. As an engineering quirk, the process we followed to produce the solution variability estimates involved *re-solving* the initial reconstruction optimization problems, rHDF_b, not directly re-loading the original reconstruction solutions. The Gurobi™ software that we used for reconstruction is largely, but not completely, deterministic at default settings. This nondeterminism could have caused the rHDF* used in solution variability problems to differ from the original reconstructed solutions in some cases, although we suspect this would be unusual. For readers concerned about this issue, we note that a triangle-inequality argument shows that twice the solution variability is a bound on the maximum L_1 distance between *any* two possible reconstruction solutions. In particular, when the solution variability produced by the IP in equation (8) is 0, the bound on the distance between two arbitrary reconstruction solutions is also 0; i.e., the reconstruction solution is indeed unique, regardless of the starting point.

Solution variability uses an output of the basic reconstruction model as one of its inputs. In practice, we found that solving the solution variability model was often computationally taxing. Concerted research on this problem class is likely to yield faster solutions. However, using binned-age, block-level solution variability models was sufficient to avoid prohibitive computational cost. For a given block *b* in a tract *t*, we solved the following problem:

$$\max_{\text{rHDF}'(b)} \left\{ L_1 \left(Hist\left(\text{rHDF}^*(b)\right), Hist\left(\text{rHDF}'(b)\right) \right) / (2N_b) \right\} \qquad (8)$$

$$s.t. \mathcal{T}_t[tabname] \leq c_t^{tabname} \; \forall tabname \in \text{Panel B Table 2}$$
$$\mathcal{T}_b[tabname] = c_b^{tabname} \; \forall tabname \in \text{Panel A Table 2,}$$

where rHDF* is the initial solution. The optimization variables are $\mathcal{B}_{(i,b,w,bl,aian,asian,nhopi,sor,hisp,a,s)}$ are defined for block *b*. Hist(rHDF′(*b*)) is defined in terms of those variables. This means that the constraint $\mathcal{T}_t[tabname] \leq c_t^{tabname}$ uses only those optimization variables and ignores the variables for other blocks in tract *t*, thus providing a relatively loose constraint.

Using only block-level solutions reduced the model's computational burden sufficiently that, at default settings, Gurobi™ could quickly solve the solution variability problem for each

block in the U.S. The resulting solution variability values are not as tight as those that could be achieved using the tract-level model; specifically, summing over our block-level solution variability measures for a given tract will generally yield larger tract-level solution variability upper bounds than would solving the tract-level problem directly. Hence, strengthened solution variability bounds could be produced in the future. However, since summing the block-level solutions to equation (8) is also an upper bound on solution variability for the tract-level problem, the large number of blocks with 0 block-level solution variability still contribute 0 to the tract-level solution variability in the 38-bin age schema. Thus, the variability in these blocks is already bounded by 0 and cannot get tighter. Tracts with 0 tract-level solution variability found in solving the optimization problem (8) must also have 0 solution variability in the stronger tract-level IP. 946 tracts have an upper bound of zero on their solution variability.

We found that solution variability could be readily computed using the optimization in equation (8) for all 6,207,027 blocks with positive population in the 2010 Census tabulations. This demonstrates that attackers have a publicly computable method for independently identifying blocks for which aggregation into tables introduces no additional uncertainty about the underlying microdata beyond the SDL measures that were applied at the record level to generate the HDF. No access to confidential data is required to perform these solution variability calculations.

Reconstruction Agreement Match

Algorithm 1 is the basic matching algorithm used generically as part of the agreement, putative, and confirmation matches. Given two databases and a set of common features, Algorithm 1 matches records exactly on the set of features without replacement by iterating over the rows in the Left database (*L*) and searching in order over the rows in the Right database (*R*) to look for the first (if any) record that matches on all the selected features. If a matching record is found, the matching records in the *L* and *R* databases are both removed and the algorithm continues to the next record in the *L* database, looking for a match in the remaining *R* records.

Algorithm 1 Match

Require: Data *L, R*, and a set of features *P*, where *p* = *dim*(*P*), that *L* and *R* have in common.
Returns: Index *M* of link records
Returns: Data *L, R* reduced to non-matches
Returns: Count of matches

```
procedure MATCH(L, R, P)
    Match ← 0
    for l ← 1 to rows(L) do
        for r ← 1 to rows(R) do
            if L[l,{1,...,p}] = R[r,{1,...,p}] then          ▷ MATCH = TRUE
                pop(l); pop(r)                               ▷ Remove records indexed by l and r
                M ← (M,{l, r})                               ▷ Append to link index
                BREAK                                        ▷ Break out of r loop
            end if
        end for
    end for
    Match ← rows(M)                                          ▷ Count
    return L, R, Match, M                                    ▷ L, R have been reduced by Match records
end procedure
```

The rHDF$_{b,t}$, rHDF$_b$, CEF, and HDF have overlapping feature sets that support the schema in equation (2), namely {*block*, *sex*, *age*, *race*, *ethnicity*}, as well as the schema in Table S3 supporting IP (6) where *age* is replaced by *agebin*. In order to measure how well the reconstructed records match the underlying confidential data (both CEF and HDF), we used Algorithm 2 to match the reconstructed microdata to our confidential databases on common features. Algorithm 2 works block-by-block implementing Algorithm 1 in two passes. First it finds all matches using {*block*, *sex*, *age*, *race*, *ethnicity*} and then any remaining matches using {*block*, *sex*, *agebin*, *race*, *ethnicity*}. The algorithm returns the unmatched records, counts of the matched records in both passes, and indices of the matched records in the original database for both passes by block.

Algorithm 2 Agreement

Require: Data *L*, with *n* records and features *P* = {*block, sex, race, ethnicity, age, agebin*}
Require: Data *R*, with *m* records and features *P* = {*block, sex, race, ethnicity, age, agebin*}
block is the geographic identifier on *L* and *R*

```
procedure AGREEMENT(L, R)
    P1 ← {block, sex, race, ethnicity, age}                  ▷ Matching Features Pass 1
    P2 ← {block, sex, race, ethnicity, agebin}               ▷ Matching Features Pass 2
    for block ∈ L do
        L_block ← Select block ∈ L
        R_block ← Select block ∈ R
        ExactAgeMatch[block] ← 0                             ▷ Indexed by [block]
        BinAgeMatch[block] ← 0                               ▷ Indexed by [block]
        L', R', ExactAgeMatch[block], M' ← MATCH(L_block, R_block, P1)   ▷ Pass 1
        L'', R'',BinAgeMatch[block],M'' ← MATCH(L', R', P2)              ▷ Pass 2
    end for
    return M', M'', ExactAgeMatch, BinAgeMatch
end procedure
```

Reidentification Details

To assess the quality of the reconstructed microdata and execute the reidentification experiments we perform several different types of matching between various data sources. Fig. 1 provides an overview of the processing beginning with the SF1 tables and finishing with confirmed reidentifications.

As described in reconstruction details, we reconstruct the HDF as either $\text{rHDF}_{b,t}$ using block and tract-level tables or rHDF_{b} using only block-level tables from the published SF1. Our first-order assessment of the quality of the reconstructed microdata matches these reconstructions directly to both the HDF and the CEF. We label this step the *agreement match* because it provides measures of agreement in record-level characteristics between $\text{rHDF}_{b,t}$ (or rHDF_{b}) and the confidential HDF and CEF. After the agreement match, we link $\text{rHDF}_{b,t}$ and rHDF_{b} separately to both commercial data (COMRCL) and a specially constructed extract from the CEF called CEF_{atkr} that includes linking variables from the schema in equation (2) feature set {*block, sex, age*} and the person identifier *pik* but no other features–specifically, neither *race* nor *ethnicity*. These linkages generate potential or putative reidentifications. By linking on *block, sex,* and *age* (or *agebin*), we create putative reidentifications. We then attach the data on *race* and *ethnicity* from $\text{rHDF}_{b,t}$ (or rHDF_{b}) to the *pik, block, sex,* and *age* information in the attacker's database (either the COMRCL or CEF_{atkr}). Finally, to evaluate the accuracy of the putative reidentifications, and classify confirmed reidentifications, we match putative reidentifications to the full CEF, linking by *pik, block, sex,* and *age* (or *agebin*), then comparing the *race* and *ethnicity* inferred from the reconstructed microdata, and attached to a putatively reidentified person, to that person's actual census responses in the CEF. We label the reidentification confirmed, when *pik, block, sex, age* (or *agebin*), *race,* and *ethnicity* all match in either the schema of equation (2) for *age* or Table S3 for *agebin.*

Reidentification Match

The reconstruction-abetted reidentification attack uses the common features between the reconstructed database and the attacker database (either COMRCL or CEF_{atkr}) to attach a previously unknown feature set (in this case {*race, ethnicity*}) to the attacker database by linking on the common attributes P_C = {*block, sex, age*}. The attacker then learns information about the database members that was previously not available. To evaluate the strength of the inference an attacker might achieve from access to improved auxiliary data, we compare the results from commercial data that were acquired by the Census Bureau contemporaneously with the 2010 Census with an improved attacker database formed from extracting {*pik, block, sex, age*} directly from the CEF, called CEF_{atkr}. When CEF_{atkr} is the attacker database, we exclude *pik* from the putative match linkage, using only {*block, sex, age*}, as we do with the commercial data. In general, we denote the attacker's external database as D_X and the reconstructed database as D_R. Note that D_X may have incomplete coverage, rows(D_X) < rows(D_R) and may contain incorrect information relative to the confidential data.

A successful match between records in D_R and D_X, based on the features in P_C, is called a putative reidentification, since the attacker must collect additional information, possibly through independent field work, to verify that the putative match is correct. The record linkage algorithm

for generating putative matches is shown in Algorithm 3. Like the agreement match algorithm, Algorithm 3 consists of two matching passes that use *age* then *agebin* to find matches. The algorithm returns an enhanced attacker external database D_{X+} consisting of records from D_X for which a match was found in the reconstructed database with the additional sensitive features {*race*, *ethnicity*} attached.

Algorithm 3 Putative reidentification using D_R

Require: Data $L = D_R$, with n records and features P_R = {*block, sex, race, ethnicity, age, agebin*}
Require: Data $R = D_X$, with m records and features P_X = {*pik, block, sex, age, agebin*}
block is the geographic identifier on L and R

procedure PUTATIVE(L, R)
  $P_1 \leftarrow$ {*block, sex, age*}. ▷ Pass 1 matching features
  $P_2 \leftarrow$ {*block, sex, agebin*}. ▷ Pass 2 matching features
  P_S = {*race, ethnicity*} ▷ Sensitive features
  for *block* $\in L$ **do**
    $L_{block} \leftarrow$ **Select** *block* $\in L$
    $R_{block} \leftarrow$ **Select** *block* $\in R$
    ExactAgeMatch[*block*] $\leftarrow 0$ ▷ Indexed by [*block*]
    BinAgeMatch[*block*] $\leftarrow 0$ ▷ Indexed by [*block*]
    L', R', ExactAgeMatch[*block*], $M' \leftarrow$ MATCH(L_{block}, R_{block}, P_1) ▷ Pass 1
    L'', R'', BinAgeMatch[*block*], $M'' \leftarrow$ MATCH(L', R', P_2) ▷ Pass 2
  end for
  return M',M'', ExactAgeMatch, BinAgeMatch
end procedure
Attach sensitive features
$D_X^1 \leftarrow D_X \cap_r M' \cap_l D_R[l, P_S]$ ▷ Exact age matches
$D_X^2 \leftarrow D_X \cap_r M'' \cap_l D_R[l, P_S]$ ▷ Binned age matches
$D_{X+} \leftarrow (D_X^1, D_X^2)$

Given the enhanced attacker external database D_{X+}, we next check if the *race* and *ethnicity* derived from the reconstructed data both match the confidential census responses in the CEF. Algorithm 4 gives this procedure. Like the agreement and putative reidentification algorithms, Algorithm 4 consists of two matching passes which use *age* and *agebin* to find matches. Records that meet the matching criteria are called confirmed reidentifications.

Algorithm 4 Confirmed reidentification using D_{X+}

Require: Data $L = D_{X+}$ with features P_{CEF} = { *pik, block, sex, race, ethnicity, age, agebin*}
Require: Data $R = D_{CEF}$, with n records and features P_{CEF}

procedure CONFIRMATION(L, R)
  $P_1 \leftarrow$ {*pik, block, sex, race, ethnicity, age*} ▷ Pass 1 matching features
  $P_2 \leftarrow$ {*pik, block, sex, race, ethnicity, agebin*}. ▷ Pass 2 matching features
  for *block* ∈ L **do**
    $L_{block} \leftarrow$ **Select** *block* ∈ L
    $R_{block} \leftarrow$ **Select** *block* ∈ R
    ExactAgeMatch[*block*] ← 0 ▷ Indexed by [*block*]
    BinAgeMatch[*block*] ← 0 ▷ Indexed by [*block*]
    L', R', ExactAgeMatch[*block*], $M' \leftarrow$ MATCH(L_{block}, R_{block}, P_1) ▷ Pass 1
    L'', R'', BinAgeMatch[*block*], $M'' \leftarrow$ MATCH(L', R', P_2) ▷ Pass 2
  end for
  return M', M'', ExactAgeMatch, BinAgeMatch
end procedure

Statistical Baseline Details

To capture privacy-violating inferences, rather than statistical or generalizable inferences, that is, to distinguish generalizable leave-one-out inferences from those that are not, the results of the reidentification attack must be compared to inferences that would be possible in a privacy-preserving counterfactual setting in which the same data are provided, except that a target individual's record has been removed. In this case, we would compare inferences made about the that person from the published 2010 Census data to inferences that would be made in the counterfactual world in which that person's record was removed. Exact comparisons of this sort are difficult because they involve removing a target individual, re-swapping and re-tabulating the data, performing reconstruction and reidentification to make inferences about the individual, and then repeating this process for every person in the United States.

In lieu of explicitly leaving each record out, as in the first-best approach, we limit our focus to small populations, and emphasize inference on persons not matching the modal *race* × *ethnicity* in their block. In this exercise, we are isolating the sensitive populations as enumerated in the confidential 2010 Census responses. Modal *race* × *ethnicity* within a census block is defined as follows. Using the CEF, we form the 126-cell *race* × *ethnicity* for all persons in the block (not just data-defined persons), all persons in the block group, and all persons in the nation. The modal *race* × *ethnicity* at the block level is the cell with most persons. If there is a tie, or if the modal population count is one, then we use the cell with the most persons in the block group containing the block. If there is still a tie, or the block group modal value is one, we use the cell in the national table with most persons (White-alone, not Hispanic or Latino).

Records with nonmodal *race* and *ethnicity* and which are unique on {*block*, *sex*, *age*} or {*block*, *sex*, *agebin*}, according to the operative schema, are of particular interest. We use the full population of the CEF, not just data-defined persons, to determine these population unique cases. For population unique persons, the reconstruction-abetted reidentification attack could not even have identified a corresponding record as a putative reidentification had the target record been absent from the CEF. These cases are not rare–fully 44% of all persons in the CEF are unique on {*block*, *sex*, *age*}. We posit two statistical guessing baselines that generate inferences using either the modal *race* × *ethnicity* in a census block (MDG), or proportional to the frequency of

each *race* × *ethnicity* in a census block (PRG) and compare the performance of these statistical guessers to the performance achieved by the reconstruction-abetted reidentification attack. Such statistical guessers simulate the counterfactual world in which statistical (non-privacy-violating) methods based on the CEF are used to generate an inference about a person that had been hypothetically removed from the CEF.

We emphasize analysis of small populations that differ significantly from the people around them because inference on these target populations is likely to be especially sensitive to the presence or absence of a target person's record. We attempt to approximate performance in the counterfactual world (removing the target person's record) by comparing to attackers armed with statistical information. By considering only a small subset of possible inferences and by allowing these statistical guessers to use information that implicitly treats the release of the modal *race* × *ethnicity* even in very small blocks as statistical rather than privacy-breaching, this approach likely underestimates the true extent of privacy violations. However, it is computationally inexpensive and helps to identify a class of inferences and group of target persons for which it is difficult to view the resulting inferences as anything but privacy violations.

We illustrate this idea with an example. Given the homogeneity of *race* and *ethnicity* within census blocks, it might be reasonable to use the {*race*, *ethnicity*} of other individuals in a *block* to make inferences about the target person. For example, suppose a *block* consisted of 10 people $\{r_1, ..., r_{10}\}$, with the first 9 being White-alone, and the 10th person being Asian-alone. All are non-Hispanic. In Census Bureau nomenclature, White-alone means the individual responded White in the WHITE set of equation (2) and did not select any of the other five choices. Similarly, the Asian-alone respondent selected Asian from the ASIAN feature and did not select any of the other five choices. All 10 respondents selected not Hispanic or Latino in the HISP feature. When the target person is r_1, the attacker in the counterfactual world (r_1's record is removed) would see 8 White-alone individuals and 1 Asian-alone. A modal guesser (MDG) would predict that the target person is White while a proportional guesser (PRG) would guess randomly in proportion to each type, assigning non-Hispanic White-alone with probability 8/9 and non-Hispanic Asian-alone with probability 1/9. Alternatively, if the target person is r_{10}, the attacker in the counterfactual world would see 9 non-Hispanic White-alone individuals and both the modal and proportional guessers would incorrectly guess non-Hispanic White-alone. Repeating such an exercise across all individuals would result in a modal guesser achieving an accuracy of 90% (the only mistake coming when the target individual is non-Hispanic Asian-alone) while the expected accuracy of the proportional guesser is approximately 81.1%. To simplify the calculations of these baselines, we use upper bounds. An upper bound on the accuracy of the modal guesser is the fraction of the block's population that reports the modal race and ethnicity in the block. The upper bound on the accuracy of the proportional guesser is $\sum_i p_i^2$, where $\{p_1, p_2, ...\}$ are the positive-only proportions of the block's population of each *race* × *ethnicity* cell.

Note that the modal guesser is targeting overall accuracy and would perform poorly when guessing the race and ethnicity of minorities within a census block. The proportional guesser would perform better with minorities at the expense of overall accuracy, and so both are reasonable to consider. Accuracy comparisons between reidentification experiments in small,

nonmodal populations involving the reconstructed microdata ($rHDF_{b,t}$ and $rHDF_b$) and the MDG and PRG give an estimate of the improved inference about these individuals due to the use of their actual census responses in the data (the privacy cost of the individual's participation in the census).

There are two additional points worth making. The first is that *race* × *ethnicity* can also be inferred statistically from the name of the target individual. However, since the reconstruction and reidentification do not model this interaction, it makes sense to omit it from the baseline. These properties of the matched experiments are salient: (1) all else being equal, they measure the privacy cost of being included in the data ($rHDF_{b,t}$ or $rHDF_b$) versus being excluded (MDG or PRG), and (2) using each pair of matched experiments ($rHDF_{b,t}$ vs. MDG, and so forth)provides a lower bound on this privacy cost (the actual privacy cost could be larger but not smaller). Furthermore, had the reconstruction used additional variables in the published tables (including household composition), the reconstructed data could have included additional features that are much harder to predict using only statistical information than *race* and *ethnicity*, for example, same-sex couples, household racial composition, or number of occupants for renters. In those cases, the gap between inference due to reidentification and statistical inference would be much larger. This is another sense in which the experiments we performed understate the true privacy cost.

To assess the relative accuracy of reidentifications using either the MDG or PRG, we generate two databases, one containing the modal guess of *race* × *ethnicity* and the other containing proportional guesses. Specifically, we use the HDF to create a frame of {*block*, *sex*, *agebin*} to which we attach the statistical baseline guesses of {*race*, *ethnicity*}. Since we are simulating an attacker making guesses using published tabulations, we use the HDF (the input file for SF1) to compute the {*race*, *ethnicity*} for each census block and block group. The national mode is White-alone not Hispanic or Latino. For the MDG database, we compute, for each *block*, the modal *race* × *ethnicity* and attach it to each record in the *block*. For the PRG database, we randomly select a {*race*, *ethnicity*} pair for each record, guessing each {*race*, *ethnicity*} in proportion to its relative frequency within the *block*. Given a tie in the block-level modal {*race*, *ethnicity*} or a block population of 1, we attempt to assign the block-group-level modal value. In the rare event that no block-group mode can be assigned, either because of a block-group-level tie or a block-group population of 1, we assign the block the national modal {*race*, *ethnicity*}. An identical exercise could be performed using published tables. One would use the tabulations in SF1 P12 and P14 to create a frame of microdata records containing {*block*, *sex*, *agebin*}, then use SF1 P8 and P9 to compute the census block-level and block-group level {*race*, *ethnicity*} tables from which to select modal and proportionally guessed {*race*, *ethnicity*}. We substitute each of our statistical baselines for the reconstructed HDF in the reidentification experiments to generate the baseline statistics. Note that, by construction, the $rHDF_b$, $rHDF_{b,t}$, MDG, and PRG databases have identical putative match rates using the binned age schema in Table S3 since all rely on an identical frame of {*block*, *sex*, *agebin*} and both reconstructions perfectly replicate this frame because SF1 P12 and P14 are fully saturated for {*sex*, *agebin*}, meaning there is never any solution variability for {*block*, *sex*, *agebin*} in either reconstructed HDF.

Distinguishing privacy breaches from generalizable inferences

Privacy and confidentiality protections are subtle concepts that give rise to many misunderstandings. It is a common (but incorrect) belief that a breach of confidentiality, also known as a privacy breach in privacy policy domain, occurs when a dataset is used to make any harmful or undesirable inference about an individual. Such an error has made its way into countless peer-reviewed papers. To see where the problem lies, we first discuss the canonical "smoking causes cancer" thought experiment *(46)* and then discuss possible confidentiality concerns in the 2020 Census data. A more complete version of this argument can be found in *(47)*.

The first Cancer Prevention Study, also known as CPS-I, followed a cohort of volunteers from 1959 to 1972 and conclusively established the link between smoking tobacco cigarettes as a cause of death from lung cancer and coronary heart disease *(48, 49)*. As a result of this study, we know that persons who smoke have a much higher risk of developing cancer. Such inferences are potentially harmful as they may result in higher health and life insurance premiums for smokers. Persons born after 1972 may be subject to this potentially harmful inference caused by the study; however, since their data could not have been used for the study (they were not born until after it was completed), the study cannot possibly be considered a privacy breach of their data. For those people, one would say that the inference is purely statistical in nature. The link between smoking and lung cancer is a population property or "statistical" use (in the sense of the 2018 Confidential Information Protection and Statistical Efficiency Act, Title III of the 2018 Foundations of Evidence-based Policymaking Act; 44 U.S. Code § 3561(12)) that was uncovered with the help of the data set. This is exactly why data are collected and published.

In contrast, a harmful inference is a privacy breach when it is specifically caused by the inclusion of the individual's information in the dataset from which the inference was made, what we called in the main text non-leave-one-out (LOO) inference. Now consider a hypothetical CPS-I study participant named Charlie who was a lifelong smoker. Suppose that, because of the study, Charlie's insurance company decided to ask enrollees whether or not they smoked and charge a higher premium if they did. As a result, Charlie was harmed by the result of the study in which he was a participant. Is this a privacy breach? To answer that question, we turn to causal reasoning, and specifically consider a counterfactual in which Charlie had not participated in the study; that is, compare the non-LOO inference with a properly computed LOO inference. Would the outcome of the study have been different enough without Charlie's participation to change the findings–thus changing whether the insurance company enacted its premium surcharge for smokers, which harmed Charlie? Given the strength of the findings in CPS-I *(49)*, we can be reasonably confident that the study would have drawn the same conclusions even if Charlie had not participated. Therefore, this example would also not be considered a privacy breach.

What would be considered a privacy breach for our hypothetical participant Charlie? Suppose that the CPS-I data were released publicly and that an external agent could reidentify Charlie's record in those data. The record shows that Charlie was one of the participants who had developed lung cancer. In this case, Charlie's insurance company would not need to ask if Charlie was a smoker–the insurance company would only need to check the public data to learn not only that he was a smoker but that he *already had* lung cancer. The knowledge that Charlie

has cancer, learned from the public CPS-I data, is an example of a privacy breach because the harmful inference (Charlie has lung cancer) was only possible because of Charlie's participation in the study.

One also must be careful about the conflation of harmful inferences with privacy violations for another reason, which is also related to privacy and trust in statistical agencies, again as they are expressed in 44 U.S. Code § 3561-4 and *(50, p. 47)*. Specifically, sometimes a statistical inference should not be allowed even if it is not privacy violating. A statistical agency might be asked to produce data on a particular sensitive population that could reasonably expect harm from those data even if they passed the LOO inference test. For example, same-sex married couples in states that did not permit such weddings in 2010 might expect harm if even a statistical summary were published for their state from the 2010 Census. Whether or not the government has a legitimate statistical interest in this population is indeed a policy question, but the policy concerns ingesting the data with the intention to publish summaries in the first place. Barring the collection or publication of data on the grounds that even the statistical inference may be harmful is a privacy and trust concern. In this paper, we presume that the agency's legitimate interest in supporting statistical inference has already been determined; that is, the collection of data in the decennial census and their publication for statistical purposes are authorized by Congress and undertaken consistent with the required trust and confidentiality policies.

In practice, the distinction between privacy-violating and generalizable statistical inferences is not always so clear-cut as in our smoking example. Information obtained from individuals is often aggregated, fields are suppressed (e.g., identifiers), and additional SDL beyond aggregation may also be used. Still, it is important to disentangle what could be learned from statistical inference versus what can be specifically learned (or caused by) a person's participation/inclusion in a study/dataset. This idea is central to differential privacy *(14)*.

Legal Authorities and Confidentiality Protections for the U.S. Census Bureau

Title 13 of the U.S. Code mandates that information gathered from individuals and establishments remain confidential. Specifically, Section 8(b) allows the Census Bureau to "furnish copies of tabulations and other statistical materials which do not disclose the information reported by, or on behalf of, any particular respondent," and Section 9 prohibits the release of "any publication whereby the data furnished by any particular establishment or individual under this title can be identified." First and foremost, the Census Bureau is required to protect the confidentiality of respondents by law. Additionally, it is in the best interest of data quality that the public trust the Census Bureau to protect their data so that truthful responses are given, especially to sensitive questions *(15, 16, 51, 52)*.

There is a common misconception that there is nothing sensitive in decennial census data. One of the reasons for this belief is that potentially harmful inferences are often about how an individual differs from a reference population. Hence, people who belong to demographic majorities in their area may have fewer or no concerns about harmful inferences. However, there are many situations in which individuals may feel uncomfortable sharing their true data:

- Age, sex, race, and ethnicity data about children are often missing in commercial databases due to legal restrictions, as detailed information about children is generally considered more sensitive.
- Household composition may be a sensitive subject in some areas and the decision to reveal this information in identifiable form should be up to the householder and not the Census Bureau according to the principles guiding statistical agencies *(50, p.3)*. This includes the detailed (census block-level) location of same-sex spouses, unmarried partners, mixed-race households, households with adopted children, older individuals living alone, etc. Thus, to encourage accurate reporting, the Census Bureau should protect the confidentiality of those responses.
- Individuals, especially those who are demographic minorities in their region, may believe that commercial databases should not collect detailed information about them without consent. Race and ethnicity information are often missing or inaccurate in commercial data but are much more accurate in census data because they are mandatorily self-reported on the questionnaire.
- Residents of rented properties in which the occupancy capacity is exceeded may wish this information to be protected.

Another misconception is that strong confidentiality protections for census data are not required because is that there is so much personal information data "out there" that the census data does not pose an incremental risk. While there are large amounts of data available externally, the accuracy of this information is generally unknown, and our experiments demonstrate that some commercial data are indeed inaccurate or at least very noisy on so-called pseudo-identifiers like age and sensitive variables like race or ethnicity. Additionally, if an agency/organization adopts the policy that data they collect should not be protected because it is already "out there," then survey response rates would drop: "why should I fill out the survey if my data is already out there, just use that and don't bother me?" But even when this position is accepted by the statistical agency, confidentiality statutes require that the census data be protected. One might be able to learn respondent-specific information collected on the census from other sources, but the census publications should not facilitate this learning, a policy that Statistics Canada makes explicit *(53)*.

Statistical Disclosure Limitation Applied to the 2010 Census

In the disclosure avoidance system for the 2010 Census, confidentiality protections for the tabular data were provided mainly by the following techniques:

- Data swapping of the geographic identifiers for the population in households;
- Partially synthetic data for the population in group quarters;
- Age and race coarsening depending upon the geographic level of the table
- Aggregation of records in the universe into summary tables.

The first of these, noise infusion via targeted data swapping, was the primary SDL method for tabular output. Households deemed at high risk for reidentification were swapped with a higher probability, but all records that were not entirely imputed had some chance of being swapped *(2)*. High-risk households were those in smaller blocks or those who had a

member with a unique race category in the block. Pairs of swapped households matched on two key demographic variables: the total number of persons and the number of adults (persons of voting age) living in the household. Once swap pairs were determined, the geographic identifiers were swapped, effectively relocating the two records in different geographies from their 2010 Census values. The swapped file was used to produce all data products.

Rules were different for 2010 Census microdata publications (39). For public-use microdata samples the disclosure avoidance included:

- Sampling, only a fraction of the population records could be released (10% for the 2010 Public Use Microdata Sample);
- Minimum population size for the geographic area identifier (100,000 persons for the Public Use Microdata Area)
- Minimum national population for one-way marginal of demographic characteristics (10,000 persons for *age*, *race*, and *ethnicity*).

The reconstructed microdata (either $rHDF_{b,t}$ or $rHDF_b$) do not satisfy these criteria. The minimum population size for a geographic feature on a microdata release of 100,000 persons is clearly violated by the *block* feature. There is no sampling; the attacker can reconstruct all 308,745,538 records with at most 10% differing on a single bit from the HDF images of the {*block, sex, agebin, race, ethnicity*} feature set. The attacker learns that an individual is a population unique for both the $rHDF_{b,t}$ and $rHDF_b$ feature sets with perfect reliability for 97 million persons and with reliability no worse than 90% for the balance of the population. Many one-way marginal distributions (*age* and *race*, in particular) have cells with fewer than 10,000 persons nationally. No reidentification, nor access to the confidential data, nor additional fieldwork is required for these inferences. Neither the reconstructed microdata nor the {*block, sex, agebin, race, ethnicity*} values from the HDF itself would not have been approved for release under the 2010 Census disclosure avoidance rules. However, the point of our reconstruction is that they were released, implicitly, in the SF1 tables because those tables are the only source for $rHDF_{b,t}$ and $rHDF_b$. The tabular and microdata rules thus do not compose. Abiding by only the tabular rule allowed reconstructed microdata that violate the microdata rules. Hence, neither product's confidentiality protections are valid. The released 2010 Public Use Microdata Sample (PUMS) was not used in the reconstruction. However, the confidentiality protections inherent in the specification for the PUMS were compromised because the reconstructed microdata $rHDF_{b,t}$ permit reidentifications alleged to be protected in the PUMS. *It is the assumption that separate rules can be applied to tabular and microdata releases without proving that those rules compose that is the vulnerability exposed by our simulated attack.*

Related Reconstruction Studies and the Relationship to Differential Privacy

In addition to the theoretical demonstration of reconstruction described in *(3)*, there are now many documented disclosure limitation failures; see the thorough reviews in *(54, 55)*. Recent work shows that practical examples of large-scale reconstruction can be paired with re-identification through auxiliary data, including in foreign-trade statistics published by Brazilian government agencies *(56)* and genomic data *(57)*. Our attack is closest to *(56)*, but our application addresses the problem of reconstruction-abetted reidentification at massive scale,

using the flagship publication of a major statistical agency. Similarly, *(31)* recently proposed a complementary attack strategy to the one described in the main text, outlining another approach to launching a reconstruction attack on census products, including both the Decennial Census of Population and Housing and the American Community Survey.

The attack we describe in the main text motivated the adoption of the differential privacy framework by the U.S. Census Bureau, but the agency's implementation of differential privacy does not render such attacks obsolete, for several reasons. First, the parameters used in real-world implementations of differentially private mechanisms tend to be large enough that the privacy guarantees they provide are not compelling viewed with a purely theoretical lens. This is also true of most present-day production implementations of differential privacy of which we are aware. See, for example, Table 2 of *(58)* and the privacy-loss budget expenditures reviewed in *(59)*. With few exceptions, even daily privacy-loss budgets generally involve expenditure of $\epsilon > 1$. It is also unclear how reasonable the independence assumptions required to justify focusing on budget for single days (or even, in some cases, individual user actions) are. Similarly, total budget expended on the 2020 Census is already in the range $\epsilon > 52$ at $\delta = 10^{-10}$ (actually, $\rho > 15$ under zero-concentrated differential privacy) for the Demographic and Housing Characteristics File, which includes all redistricting data, and is likely to increase considerably as more products are released based on the 2020 Census Edited File.

These overall budgets can be interpreted by converting to statements like: "an attacker trying to build a hypothesis test for determining whether a target person's record was in or not in the original data set could only achieve a significance level (probability of a Type 1 error) of 0.01 if the test's power (1 – probability of a Type 2 error) were no greater than β." When privacy-loss budgets are small, β is close to 0.01, and this statement is compelling. However, with budgets in the ranges used in practice, the power bound in this statement will generally be very close to 1.0 (i.e., nearly trivial). Similar statements with more quantitatively meaningful bounds can be made *(47)*, but only at the expense of focusing on specific attributes and making additional assumptions about the information available for relating different attributes. Empirical attacks like the reconstruction-abetted reidentification attack reported on here, therefore, serve five important purposes, even after the adoption of differential privacy.

First, if the empirical attack fails, it provides a direct justification for parameter choices, by showing that at least this known, important attack vector is properly protected. If the attack succeeds, it shows that the privacy-loss budget is either too large or must be distributed across the features differently. The Census Bureau actively used the attacks reported here during the development of the 2020 Census DAS *(29)*. When theoretical statements are focused on specific attributes considered in isolation, empirical attacks help motivate the choice of attributes—for example, the reconstruction-abetted reidentification attack investigated in this paper illustrates that census block is an important feature for linking to auxiliary data; hence, strong protection of the *block* feature (total privacy-loss budget expended on that feature of ρ=0.14 *(41)* for the 2020 Census) is essential.

Second, although the Census Bureau responded to the reconstruction-abetted reidentification attack described in the main text by adopting the differential privacy framework for SDL in the 2020 Census, many products produced by the agency continue to use traditional

SDL frameworks, and traditional SDL is common in many other national statistical agencies as well. For example, like the decennial census techniques used in the 2010 Census, for which the main text reports a successful simulated attack, the American Community Survey continues to use household-swapping of geographic identifiers as the primary SDL framework *(60, 61)* for tabular data but a non-composing framework for microdata releases, putting both tabular and public-use microdata products at risk of confidentiality breaches. The Census Bureau has announced that ACS will not transition to a differentially private framework until at least 2025 *(62)*. Similarly, the Office for National Statistics of the United Kingdom implemented confidentiality protection for its 2021 Census of England and Wales respondents through a combination of geographic identifier swapping and noise infusion heuristics *(63)* that are not differentially private but produce tables with many of the same properties as the tables from the 2020 Census: random noise added to every cell, including zeros, for example. Thus, although swapping is no longer in use for disclosure avoidance in the 2020 Census, demonstrations of the vulnerability of swapping-based methods continue to be highly relevant in practical contexts. The Office of National Statistics in the U.K. also performed limited-scale simulated reconstruction-abetted reidentification attacks in preparing the SDL. It issued a Request for Proposals to find an external entity to perform large-scale reconstruction experiments, but no bidder was chosen.

Third, the attack we simulated is deliberately generic. It can be applied to simulate an attack against any system for which the primary output is unweighted tabulations based on microdata. Nothing about our attack approach exploits features specific to swapping. That our attack is generic in this sense is precisely why, for example, it could be—and was—applied, without modification, to assess the strength of the parameter choices made for the differentially private 2020 Census Disclosure Avoidance System, which, like its 2010 Census swapping predecessor, also outputs microdata, over which tabulations are then computed and published.

Fourth, we demonstrate that an attacker could show that, using the binned age feature (*agebin* in the main text), there is a provably unique microdata set consistent with the published tabulations for 70% of census blocks. We show that an attacker could confirm the exact reconstruction using only public data and describe how our technique also provides a bound on how different distinct solutions to the reconstruction problem could be, even in areas where the solution is not unique. To the best of our knowledge, this additional feature of our attack is novel. Through access to the underlying confidential data, we show that an attacker exploiting this additional attack strategy to target specific areas can significantly exacerbate disclosure risk.

Fifth, and lastly, although the Census Bureau has adopted formally private methods to mitigate the concerns raised in this paper, external stakeholders still exert considerable pressure to publish very detailed statistical summaries on very small populations with very limited disclosure avoidance applied. For example, *(6)* argues that preliminary reports on the results of our simulated attack are not compelling evidence of widespread privacy vulnerabilities, and *(4, 5)* argue that variability in the space of possible solutions to the reconstruction problem similarly implies that evidence of privacy vulnerabilities is weak. Persistent statements like these are often coupled with requests to delay or reverse the course of disclosure avoidance modernization at the agency. This situation makes the current paper especially important; it is the primary public

documentation of both the mechanics of the attack itself and our methodology for interpreting the results.

To briefly summarize: the attack we illustrate on the Census Bureau's flagship publication motivated the agency to adopt a new, differentially private framework to protect respondent confidentiality for the 2020 Census. However, the implementation of differential privacy used parameters that are large enough to warrant investigating the efficacy of empirical attacks like the one described in the main text. Even if the differential privacy parameters were small, traditional disclosure avoidance systems also continue to be in widespread use. Our attack is generic enough that it can be used to help interpret privacy-loss budgets and assess vulnerabilities in traditional systems. Lastly, our solution variability analysis is novel, and demonstrates a concrete way an attacker could substantially improve confidence in, and success of, their attack.

Fig. S1.

Summary of the creation of Summary File 1. Collected census responses are stored in the Census Unedited File (CUF) for the complete in-scope living quarter domain. CUF is edited and all missing responses are allocated or imputed to produce the Census Edited File (CEF). Record-level confidentiality protections and tabulation edits are applied to create the Hundred-percent Detail File (HDF), which is then aggregated into tables producing Summary File 1 and other publications.

Table S1.

	Records in CEF	Records not in CEF	Total
Records in COMRCL	106,300	182,900	289,100
Records not in COMRCL	169,700		
Total	276,000		

Notes: Counts rounded to four significant digits to conform to current disclosure avoidance rules. Shown are counts of records that do and do not match exactly on the feature set {*pik, block, sex, agebin*} between CEF and COMRCL along with the total number of data-defined records in CEF that are unique within {*pik, block*} and the total number of data-defined records in COMRCL unduplicated.

Overlap of data-defined persons in CEF and COMRCL databases.

Table S2.

a	0-4	5-9	10-14	15-17	18-19	20	21	22-24	25-29	30-34	35-39
z	0	1	2	3	4	5	6	7	8	9	10
a	40-44	45-49	50-54	55-59	60-61	62-64	65-66	67-69	70-74	75-79	80-84
z	11	12	13	14	15	16	17	18	19	20	21
a	85-110										
z	22										

Index mapping for the 23-bin age grouping in Summary File 1 Table P12, function z = AGEBINP12(a).

Table S3.

a	0	1	2	3	4	5	6	7	8	9	10
z	0	1	2	3	4	5	6	7	8	9	10
a	11	12	13	14	15	16	17	18	19	20	21
z	11	12	13	14	15	16	17	18	19	20	21
a	22-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-61	62-64	65-66
z	22	23	24	25	26	27	28	29	30	31	32
a	67-69	70-74	75-79	80-84	85-110						
z	33	34	35	36	37						

Index mapping for the 38-bin age grouping in Summary File 1 census block-level tables, function z = AGEBINBLOCK(a). The table also defines the mapping that defines the feature *agebin* in relation to the feature *age*.

Table S4.

Cumulative Percentage	Quantile	Population in Interval	Cumulative Population	Solution Variability	Cumulative Solution Variability
5	0	6,405	6,405	0	0.0
10	0	6,376	12,781	0	0.0
15	0	6,390	19,172	0	0.0
20	0	6,379	25,550	0	0.0
25	0	6,382	31,932	0	0.0
30	0	6,388	38,320	0	0.0
35	0	6,387	44,707	0	0.0
40	0	6,369	51,075	0	0.0
45	0	6,367	57,442	0	0.0
50	0	6,357	63,799	0	0.0
55	0	6,403	70,202	0	0.0
60	0	6,369	76,571	0	0.0
65	0	6,379	82,949	0	0.0
70	0	14,288	97,238	196	0.1
75	1.7	34,272	131,509	2,415	0.9
80	4.7	28,282	159,791	5,768	1.8
85	7.3	28,774	188,565	10,843	2.9
90	10.5	30,318	218,884	18,411	4.2
95	14.6	36,466	255,350	31,308	6.1
100	21.1	53,395	308,746	61,987	10.0

Notes: Solution variability is the statistic *solvar* defined in the main text. Cumulative solution variability is the sum of solvar up to and including the displayed cell. Population and cumulative population counts are thousands. This table is based entirely on public data. No rounding has been applied to the population counts. Block ties in the definition of quantiles were broken randomly. Running the replication package for this table will result in minor variations in the population and cumulative population columns because of this computational randomness.

Selected empirical quantiles for census block-level solution variability.

Table S5.

Block Population Range	All Blocks			Zero Solution Variability Blocks			
	Block Count	Population	Unique Percentage	Block Count	Population	Percentage of Blocks in Row	Percentage of Population in Row
1-9	1,823,665	8,070	81.7	1,815,218	8,011	99.5	99.3
10-49	2,671,753	67,598	50.3	2,096,508	48,409	78.5	71.6
50-99	994,513	69,073	26.9	324,641	21,474	32.6	31.1
100-249	540,455	80,021	12.2	67,884	9,156	12.6	11.4
250-499	126,344	42,911	3.4	3,718	1,174	2.9	2.7
500-999	40,492	27,029	0.8	308	196	0.8	0.7
1000+	9,805	14,043	0.2	105	169	1.1	1.2

Notes: Population counts shown in thousands. This table is based entirely on public data. No population counts have been rounded. Unique percentage shows the percent of the total population in the row that are data-defined and unique on the feature set {*block, sex, agebin*}, where *agebin* is the 38-bin feature defined in the main text.

Population uniques and zero solution variability by census block size.

Table S6.

Data (L in Algorithm 3)	Block Population Range	Attacker (R in Algorithm 3): COMRCL			Attacker (R in Algorithm 3): CEF_{atkr}		
		Putative	Confirmed	Precision Percentage	Putative	Confirmed	Precision Percentage
CEF	All	167,500	82,760	49.4	276,000	237,500	86.1
HDF	All	166,100	80,540	48.5	267,800	228,400	85.3
$rHDF_{b,t}$	All	166,100	68,480	41.2	267,800	208,500	77.9
$rHDF_{b}$	All	166,100	67,450	40.6	267,800	203,100	75.9
MDG	All	166,100	76,270	45.9	267,800	205,100	76.6
PRG	All	166,100	66,260	39.9	267,800	177,700	66.3
CEF	1-9	3,470	2,976	85.8	7,373	7,357	99.8
HDF	1-9	2,862	2,303	80.5	5,517	5,427	98.4
$rHDF_{b,t}$	1-9	2,862	2,201	76.9	5,517	5,402	97.9
$rHDF_{b}$	1-9	2,862	2,196	76.7	5,517	5,395	97.8
MDG	1-9	2,862	2,228	77.9	5,517	5,193	94.1
PRG	1-9	2,862	2,170	75.8	5,517	5,054	91.6
CEF	10-49	34,250	23,300	68.0	61,820	60,350	97.6
HDF	10-49	33,730	22,430	66.5	58,630	56,840	96.9
$rHDF_{b,t}$	10-49	33,730	19,600	58.1	58,630	54,540	93.0
$rHDF_{b}$	10-49	33,730	19,430	57.6	58,630	53,840	91.8
MDG	10-49	33,730	20,870	61.9	58,630	50,620	86.3
PRG	10-49	33,730	19,010	56.4	58,630	46,250	78.9
CEF	50-99	38,160	21,200	55.6	62,760	58,450	93.1
HDF	50-99	37,980	20,780	54.7	60,970	56,440	92.6
$rHDF_{b,t}$	50-99	37,980	17,120	45.1	60,970	51,230	84.0
$rHDF_{b}$	50-99	37,980	16,860	44.4	60,970	49,880	81.8
MDG	50-99	37,980	18,980	50.0	60,970	48,170	79.0
PRG	50-99	37,980	16,500	43.5	60,970	41,990	68.9
CEF	100-249	45,320	20,440	45.1	71,410	61,190	85.7
HDF	100-249	45,250	20,240	44.7	70,400	60,030	85.3
$rHDF_{b,t}$	100-249	45,250	16,660	36.8	70,400	52,650	74.8
$rHDF_{b}$	100-249	45,250	16,350	36.1	70,400	50,890	72.3
MDG	100-249	45,250	18,940	41.9	70,400	52,190	74.1
PRG	100-249	45,250	16,120	35.6	70,400	44,310	62.9
CEF	250-499	24,680	8,923	36.2	37,660	28,480	75.6
HDF	250-499	24,670	8,878	36.0	37,400	28,140	75.3
$rHDF_{b,t}$	250-499	24,670	7,608	30.8	37,400	24,940	66.7

rHDF$_b$	250-499	24,670	7,459	30.2	37,400	24,090	64.4
MDG	250-499	24,670	8,886	36.0	37,400	26,520	70.9
PRG	250-499	24,670	7,409	30.0	37,400	22,050	59.0
CEF	500-999	15,090	4,416	29.3	23,270	15,250	65.5
HDF	500-999	15,090	4,398	29.2	23,180	15,110	65.2
rHDF$_{b,t}$	500-999	15,090	3,898	25.8	23,180	13,700	59.1
rHDF$_b$	500-999	15,090	3,809	25.3	23,180	13,230	57.1
MDG	500-999	15,090	4,664	30.9	23,180	15,370	66.3
PRG	500-999	15,090	3,748	24.9	23,180	12,430	53.6
CEF	1000+	6,510	1,510	23.2	11,670	6,432	55.1
HDF	1000+	6,510	1,505	23.1	11,650	6,403	55.0
rHDF$_{b,t}$	1000+	6,510	1,380	21.2	11,650	5,986	51.4
rHDF$_b$	1000+	6,510	1,339	20.6	11,650	5,779	49.6
MDG	1000+	6,510	1,699	26.1	11,650	7,056	60.6
PRG	1000+	6,510	1,305	20.1	11,650	5,576	47.9

Notes: Counts in thousands rounded to four significant digits to conform to current disclosure avoidance rules. COMRCL and CEF$_{atkr}$ use only data-defined records.

Putative reidentifications, confirmed reidentifications, and precision rates for all data-defined persons by census block size.

Table S7.

Data (L in Algorithm 3)	Population	Putative Reidentifications	Putative Percentage	Confirmed Reidentifications	Confirmed Percentage	Precision Percentage
	Panel A: All data-defined records in CEF					
CEF	289,100	167,500	57.9	82,760	28.6	49.4
HDF	289,100	166,100	57.5	80,540	27.9	48.5
$rHDF_{b,t}$	289,100	166,100	57.5	68,480	23.7	41.2
$rHDF_{b}$	289,100	166,100	57.5	67,450	23.3	40.6
MDF	289,100	112,100	38.8	32,790	11.3	29.3
$rMDF_{b,t}$	289,100	112,100	38.8	32,380	11.2	28.9
MDG	289,100	166,100	57.5	76,270	26.4	45.9
PRG	289,100	166,100	57.5	66,260	22.9	39.9
	Panel B: Records in COMRCL that match CEF on *{pik, block, sex, agebin}*					
CEF	106,300	92,950	87.5	82,760	77.9	89.0
HDF	106,300	91,080	85.7	80,540	75.8	88.4
$rHDF_{b,t}$	106,300	83,180	78.3	68,480	64.5	82.3
$rHDF_{b}$	106,300	82,750	77.9	67,450	63.5	81.5
MDF	106,300	46,130	43.4	32,790	30.9	71.1
$rMDF_{b,t}$	106,300	46,000	43.3	32,380	30.5	70.4
MDG	106,300	91,080	85.7	76,270	71.8	83.8
PRG	106,300	91,080	85.7	66,260	62.4	72.8

Notes: Counts in thousands rounded to four significant digits to conform to current disclosure avoidance rules. Panel A shows counts using all data-defined records in COMRCL. Panel B limits to the set of records in COMRCL that match by {*pik, block, sex, agebin*} to CEF. Population in Panel A putative and confirmed reidentification rates is the count of records in CEF with {*pik, block, sex, age*}; i.e., data-defined records at risk of reidentification from COMRCL. Population in Panel B is the count of records in COMRCL that match CEF on {*pik, block, sex, agebin*}; that is, CEF records actually at risk of reidentification. Putative and confirmation counts are derived from the execution of Algorithms 3 and 4 in Materials and Methods. Putative matches attach the value of {*race, ethnicity*} to a *pik* by matching on {*block, sex, age*} or {*block, sex, agebin*}. Confirmed matches are those where {*pik, block, sex, age, race, ethnicity*} or {*pik, block, sex, agebin, race, ethnicity*) match between the attacker's putative matches and the CEF. Precision is the ratio of confirmed to putative matches stated as a percentage. The row labeled CEF is the upper bound. The CEF was substituted for the reconstructed HDF as the L file in the putative match Algorithm 3. The rows labeled $rHDF_{b,t}$ and $rHDF_{b}$ are the reconstructed HDF defined in the main text. The rows labeled $rMDF_{b,t}$ and MDF (shaded) are the analogues of $rHDF_{b,t}$ and HDF when the 2020 Census Disclosure Avoidance System, based on a differentially private framework, is applied to the 2010 Census Edited File as explained in the main text. The rows labeled MDC and PRG are the statistical benchmarks defined in the main text.

Putative reidentification, confirmed reidentification, and precision rates for all data-defined COMRCL records and for all data-defined records matching the CEF.

Table S8.

Data (*L* in Algorithm 3)	Population	Attacker (*R* in Algorithm 3): COMRCL Putative	Confirmed	Precision Percentage	Attacker (*R* in Algorithm 3): CEF_{atkr} Putative	Confirmed	Precision Percentage
		Panel A: Nonmodal Persons					
CEF	6,517	1,993	1,633	81.9	6,517	5,727	87.9
HDF	6,517	1,716	1,278	74.5	5,189	4,209	81.1
$rHDF_{b,t}$	6,517	1,638	1,009	61.6	5,305	3,634	68.5
$rHDF_{b}$	6,517	1,598	914	57.2	5,304	3,369	63.5
MDG	6,517	1,716	33	1.9	5,189	88	1.7
PRG	6,517	1,716	274	16.0	5,189	890	17.2
$rMDF_{b,t}$	6,517	649	103	15.8	1,866	352	18.9
MDF	6,517	649	103	15.8	1,866	353	18.9
		Panel B: Nonmodal Uniques on *{block, sex, agebin}*					
CEF	3,364	856	834	97.4	3,364	3,364	100.0
HDF	3,364	673	625	93.0	2,418	2,311	95.6
$rHDF_{b,t}$	3,364	614	560	91.2	2,418	2,301	95.2
$rHDF_{b}$	3,364	593	537	90.6	2,418	2,237	92.5
MDG	3,364	673	21	3.2	2,418	61	2.5
PRG	3,364	673	136	20.2	2,418	488	20.2
$rMDF_{b,t}$	3,364	212	42	19.8	671	147	21.8
MDF	3,364	212	42	19.7	671	147	21.8

Notes: Counts in thousands rounded to four significant digits to conform to current disclosure avoidance rules. Zero solution variability is explained in the main text. The rows labeled $rMDF_{b,t}$ and MDF (shaded) are the analogues of $rHDF_{b,t}$ and HDF when the 2020 Census Disclosure Avoidance System, based on a differentially private framework, is applied to the 2010 Census Edited File as explained in the main text.

Putative reidentifications, confirmed reidentifications, and precision rates for nonmodal persons in census blocks with zero solution variability.

Table S9.

Data (L in Algorithm 3)	Block Size	Attacker (R in Algorithm 3): COMRCL			Attacker (R in Algorithm 3): CEF_{atkr}		
		Putative	Confirmed	Precision Percentage	Putative	Confirmed	Precision Percentage
CEF	All	80,970	73,010	90.2	210,100	195,400	93.0
HDF	All	79,670	71,570	89.8	205,200	190,300	92.7
$rHDF_{b,t}$	All	73,020	62,330	85.4	204,900	181,300	88.5
$rHDF_{b}$	All	72,910	61,940	85.0	204,900	179,100	87.4
MDG	All	79,670	76,100	95.5	205,200	204,500	99.7
PRG	All	79,670	63,790	80.1	205,200	167,700	81.7
CEF	1-9	2,854	2,833	99.3	6,893	6,886	99.9
HDF	1-9	2,265	2,222	98.1	5,241	5,177	98.8
$rHDF_{b,t}$	1-9	2,177	2,126	97.6	5,239	5,163	98.5
$rHDF_{b}$	1-9	2,176	2,123	97.6	5,239	5,159	98.5
MDG	1-9	2,265	2,220	98.0	5,241	5,173	98.7
PRG	1-9	2,265	2,141	94.5	5,241	4,966	94.8
CEF	10-49	21,840	21,050	96.4	52,270	51,630	98.8
HDF	10-49	21,400	20,570	96.1	50,660	49,930	98.6
$rHDF_{b,t}$	10-49	19,440	18,170	93.5	50,550	48,760	96.5
$rHDF_{b}$	10-49	19,410	18,090	93.2	50,550	48,450	95.8
MDG	10-49	21,400	20,810	97.3	50,660	50,440	99.6
PRG	10-49	21,400	18,570	86.8	50,660	44,660	88.2
CEF	50-99	19,930	18,450	92.6	49,020	47,250	96.4
HDF	50-99	19,760	18,250	92.4	48,180	46,380	96.3
$rHDF_{b,t}$	50-99	17,690	15,430	87.3	48,070	44,010	91.5
$rHDF_{b}$	50-99	17,660	15,340	86.8	48,080	43,460	90.4
MDG	50-99	19,760	18,940	95.8	48,180	48,010	99.7
PRG	50-99	19,760	15,930	80.6	48,180	39,880	82.8
CEF	100-249	20,040	17,640	88.0	52,740	48,730	92.4
HDF	100-249	19,970	17,540	87.9	52,200	48,150	92.3
$rHDF_{b,t}$	100-249	18,240	14,960	82.0	52,140	45,000	86.3
$rHDF_{b}$	100-249	18,220	14,860	81.6	52,150	44,330	85.0
MDG	100-249	19,970	18,900	94.7	52,200	52,060	99.7
PRG	100-249	19,970	15,410	77.2	52,200	41,510	79.5
CEF	250-499	9,482	7,860	82.9	26,690	23,200	86.9
HDF	250-499	9,468	7,831	82.7	26,520	23,000	86.7
$rHDF_{b,t}$	250-499	8,903	6,921	77.7	26,510	21,610	81.5

rHDF$_b$	250-499	8,889	6,868	77.3	26,510	21,280	80.3
MDG	250-499	9,468	8,874	93.7	26,520	26,470	99.8
PRG	250-499	9,468	7,049	74.5	26,520	20,460	77.1
CEF	500-999	4,995	3,879	77.7	15,430	12,510	81.1
HDF	500-999	4,991	3,866	77.5	15,370	12,420	80.8
rHDF$_{b,t}$	500-999	4,793	3,516	73.4	15,370	11,760	76.5
rHDF$_b$	500-999	4,781	3,477	72.7	15,370	11,570	75.3
MDG	500-999	4,991	4,658	93.3	15,370	15,340	99.8
PRG	500-999	4,991	3,512	70.4	15,370	11,310	73.6
CEF	1000+	1,822	1,293	71.0	7,066	5,219	73.9
HDF	1000+	1,821	1,290	70.8	7,056	5,200	73.7
rHDF$_{b,t}$	1000+	1,780	1,206	67.8	7,056	4,969	70.4
rHDF$_b$	1000+	1,772	1,185	66.9	7,056	4,868	69.0
MDG	1000+	1,821	1,697	93.2	7,056	7,048	99.9
PRG	1000+	1,821	1,183	65.0	7,056	4,859	68.9

Notes: Counts in thousands rounded to four significant digits to conform to current disclosure avoidance rules. COMRCL and CEF$_{atkr}$ use only data-defined records.

Putative reidentifications, confirmed reidentifications, and precision rates for all modal persons by census block size.

Table S10.

Data (*L*-*R* in Algorithm 2)	Block Population Range	Population	Agreement		Agreement Percentage	
			Exact Age	Binned Age	Exact Age	Binned Age
HDF-CEF	All	308,746	297,200	297,600	96.3	96.4
rHDF$_{b,t}$-CEF	All	308,746	143,800	283,600	46.6	91.9
rHDF$_{b}$-CEF	All	308,746	132,200	276,900	42.8	89.7
rMDF$_{b,t}$-CEF	All	308,746	58,520	113,100	18.9	36.6
MDF-CEF	All	308,746	75,950	113,300	24.6	36.7
HDF-CEF	1-9	8,070	5,866	5,973	72.7	74.0
rHDF$_{b,t}$-CEF	1-9	8,070	2,325	5,968	28.8	74.0
rHDF$_{b}$-CEF	1-9	8,070	2,419	5,971	30.0	74.0
rMDF$_{b,t}$-CEF	1-9	8,070	232	647	2.9	8.0
MDF-CEF	1-9	8,070	276	647	3.4	8.0
HDF-CEF	10-49	67,598	63,460	63,580	93.9	94.1
rHDF$_{b,t}$-CEF	10-49	67,598	29,500	62,870	43.6	93.0
rHDF$_{b}$-CEF	10-49	67,598	28,990	62,330	42.9	92.2
rMDF$_{b,t}$-CEF	10-49	67,598	4,999	12,330	7.4	18.2
MDF-CEF	10-49	67,598	6,216	12,320	9.2	18.2
HDF-CEF	50-99	69,073	66,560	66,630	96.4	96.5
rHDF$_{b,t}$-CEF	50-99	69,073	30,600	63,130	44.3	91.4
rHDF$_{b}$-CEF	50-99	69,073	31,280	64,330	45.3	93.1
rMDF$_{b,t}$-CEF	50-99	69,073	8,350	18,830	12.1	27.3
MDF-CEF	50-99	69,073	10,670	18,820	15.5	27.2
HDF-CEF	100-249	80,021	78,370	78,420	97.9	98.0
rHDF$_{b,t}$-CEF	100-249	80,021	36,840	73,810	46.0	92.2
rHDF$_{b}$-CEF	100-249	80,021	34,690	71,940	43.4	89.9
rMDF$_{b,t}$-CEF	100-249	80,021	15,030	30,810	18.8	38.5
MDF-CEF	100-249	80,021	19,750	30,790	24.7	38.5
HDF-CEF	250-499	42,911	42,320	42,340	98.6	98.7
rHDF$_{b,t}$-CEF	250-499	42,911	20,750	39,240	48.3	91.4
rHDF$_{b}$-CEF	250-499	42,911	18,170	37,960	42.3	88.5
rMDF$_{b,t}$-CEF	250-499	42,911	12,220	22,570	28.5	52.6
MDF-CEF	250-499	42,911	16,290	22,600	38.0	52.7
HDF-CEF	500-999	27,029	26,720	26,740	98.9	98.9
rHDF$_{b,t}$-CEF	500-999	27,029	11,380	23,480	42.1	86.9
rHDF$_{b}$-CEF	500-999	27,029	14,220	24,550	52.6	90.8
rMDF$_{b,t}$-CEF	500-999	27,029	10,280	17,210	38.0	63.7

MDF-CEF	500-999	27,029	13,540	17,310	50.1	64.0
HDF-CEF	1000+	14,043	13,930	13,940	99.2	99.3
$rHDF_{b,t}$-CEF	1000+	14,043	8,835	12,870	62.9	91.7
$rHDF_{b}$-CEF	1000+	14,043	6,009	12,120	42.8	86.3
$rMDF_{b,t}$-CEF	1000+	14,043	7,407	10,670	52.8	76.0
MDF-CEF	1000+	14,043	9,204	10,820	65.5	77.0

Notes: Counts in thousands. 2010 Census exact block populations are public data; therefore, the population column is not rounded. Other counts rounded to four significant digits to conform to current disclosure avoidance rules. Agreement percentages use the population in the census blocks included for that row. Rows in gray shade are output from the 2020 Disclosure Avoidance System, based on the differential privacy framework. Unshaded rows are identical to Table 3.

Selected reconstruction agreement statistics with comparisons to output from the 2020 Census Disclosure Avoidance System using the 2010 Census as input.

Table S11.

Data (L in Algorithm 3)	Block Population Range	Attacker (R in Algorithm 3): COMRCL			Attacker (R in Algorithm 3): CEF_{atkr}		
		Putative	Confirmed	Precision Percentage	Putative	Confirmed	Precision Percentage
CEF	All	16,340	9,746	59.7	65,850	42,070	63.9
HDF	All	15,780	8,967	56.8	62,530	38,120	61.0
$rHDF_{b,t}$	All	14,990	6,147	41.0	62,810	27,180	43.3
$rMDF_{b,t}$	All	9,632	1,795	18.6	35,360	7,192	20.3
MDF	All	9,631	1,775	18.4	35,380	7,108	20.1
CEF	1-9	148	143	97.0	480	472	98.3
HDF	1-9	92	81	87.7	276	249	90.2
$rHDF_{b,t}$	1-9	90	75	83.7	279	240	86.0
$rMDF_{b,t}$	1-9	19	5	24.4	56	14	25.5
MDF	1-9	19	5	24.3	56	14	25.4
CEF	10-49	2,621	2,242	85.6	9,545	8,713	91.3
HDF	10-49	2,337	1,865	79.8	7,971	6,904	86.6
$rHDF_{b,t}$	10-49	2,199	1,435	65.3	8,084	5,777	71.5
$rMDF_{b,t}$	10-49	813	160	19.7	2,404	538	22.4
MDF	10-49	813	160	19.7	2,405	535	22.3
CEF	50-99	3,745	2,748	73.4	13,740	11,200	81.5
HDF	50-99	3,600	2,527	70.2	12,790	10,060	78.6
$rHDF_{b,t}$	50-99	3,355	1,688	50.3	12,900	7,221	56.0
$rMDF_{b,t}$	50-99	1,697	322	19.0	5,147	1,087	21.1
MDF	50-99	1,698	321	18.9	5,148	1,079	21.0
CEF	100-249	4,775	2,797	58.6	18,670	12,460	66.8
HDF	100-249	4,718	2,700	57.2	18,210	11,870	65.2
$rHDF_{b,t}$	100-249	4,457	1,706	38.3	18,260	7,645	41.9
$rMDF_{b,t}$	100-249	2,893	538	18.6	9,668	1,981	20.5
MDF	100-249	2,893	533	18.4	9,673	1,958	20.2
CEF	250-499	2,531	1,063	42.0	10,970	5,275	48.1
HDF	250-499	2,523	1,047	41.5	10,870	5,142	47.3
$rHDF_{b,t}$	250-499	2,430	687	28.3	10,880	3,337	30.7
$rMDF_{b,t}$	250-499	1,950	354	18.2	7,502	1,493	19.9
MDF	250-499	1,950	350	17.9	7,506	1,471	19.6
CEF	500-999	1,677	537	32.0	7,846	2,735	34.9

HDF	500-999	1,675	532	31.8	7,812	2,690	34.4
$rHDF_{b,t}$	500-999	1,635	383	23.4	7,813	1,945	24.9
$rMDF_{b,t}$	500-999	1,469	268	18.2	6,361	1,234	19.4
MDF	500-999	1,469	262	17.9	6,364	1,217	19.1
CEF	1000+	840	216	25.8	4,602	1,213	26.4
HDF	1000+	840	215	25.6	4,595	1,203	26.2
$rHDF_{b,t}$	1000+	827	174	21.0	4,595	1,017	22.1
$rMDF_{b,t}$	1000+	791	148	18.8	4,224	844	20.0
MDF	1000+	790	144	18.3	4,225	834	19.7

Notes: Counts in thousands rounded to four significant digits to conform to current disclosure avoidance rules. The row $MDF_{b,t}$ (shaded) is the full reconstruction-abetted re-identification attack on 2010 Census data that have been processed through the 2020 Census Disclosure Avoidance System (DAS) using final production parameters and reported as tabular summaries with the same schema as the 2020 Census Demographic and Housing Characteristics (DHC) File, the successor to the 2010 Census Summary File 1, comparable to the tables in Table 2 in the main text. The row MDF (shaded) is only the re-identification attack run against the Microdata Detail File (MDF), the input to the tabular DHC publications. A version of the MDF, called the Privacy-Protected Microdata File (PPMF) will be released for the 2020 Census as a successor to the 2010 Census Public-Use Microdata Sample. Unshaded rows are identical to entries in Table 5 in the main text.

Putative reidentifications, confirmed reidentifications, and precision rates for all nonmodal persons using the 2020 Disclosure Avoidance System by census block size.

Table S12.

Data (L in Algorithm 3)	Block Population Range	Attacker (R in Algorithm 3): COMRCL			Attacker (R in Algorithm 3): CEF_{atkr}		
		Putative	Confirmed	Precision Percentage	Putative	Confirmed	Precision Percentage
CEF	All	167,500	82,760	49.4	276,000	237,500	86.1
HDF	All	166,100	80,540	48.5	267,800	228,400	85.3
$rHDF_{b,t}$	All	166,100	68,480	41.2	267,800	208,500	77.9
$rMDF_{b,t}$	All	112,100	32,790	29.3	130,800	82,510	63.1
MDF	All	112,100	32,380	28.9	130,800	81,670	62.5
CEF	1-9	3,470	2,976	85.8	7,373	7,357	99.8
HDF	1-9	2,862	2,303	80.5	5,517	5,427	98.4
$rHDF_{b,t}$	1-9	2,862	2,201	76.9	5,517	5,402	97.9
$rMDF_{b,t}$	1-9	894	260	29.1	785	584	74.5
MDF	1-9	894	260	29.1	785	584	74.4
CEF	10-49	34,250	23,300	68.0	61,820	60,350	97.6
HDF	10-49	33,730	22,430	66.5	58,630	56,840	96.9
$rHDF_{b,t}$	10-49	33,730	19,600	58.1	58,630	54,540	93.0
$rMDF_{b,t}$	10-49	15,470	4,966	32.1	14,950	10,770	72.0
MDF	10-49	15,470	4,939	31.9	14,950	10,720	71.7
CEF	50-99	38,160	21,200	55.6	62,760	58,450	93.1
HDF	50-99	37,980	20,780	54.7	60,970	56,440	92.6
$rHDF_{b,t}$	50-99	37,980	17,120	45.1	60,970	51,230	84.0
$rMDF_{b,t}$	50-99	22,360	6,954	31.1	23,180	15,570	67.2
MDF	50-99	22,360	6,896	30.8	23,180	15,450	66.7
CEF	100-249	45,320	20,440	45.1	71,410	61,190	85.7
HDF	100-249	45,250	20,240	44.7	70,400	60,030	85.3
$rHDF_{b,t}$	100-249	45,250	16,660	36.8	70,400	52,650	74.8
$rMDF_{b,t}$	100-249	32,330	9,745	30.1	36,710	23,600	64.3
MDF	100-249	32,330	9,628	29.8	36,710	23,370	63.7
CEF	250-499	24,680	8,923	36.2	37,660	28,480	75.6
HDF	250-499	24,670	8,878	36.0	37,400	28,140	75.3
$rHDF_{b,t}$	250-499	24,670	7,608	30.8	37,400	24,940	66.7
$rMDF_{b,t}$	250-499	20,800	5,974	28.7	25,640	15,840	61.8
MDF	250-499	20,800	5,875	28.3	25,640	15,640	61.0
CEF	500-999	15,090	4,416	29.3	23,270	15,250	65.5
HDF	500-999	15,090	4,398	29.2	23,180	15,110	65.2
$rHDF_{b,t}$	500-999	15,090	3,898	25.8	23,180	13,700	59.1

$rMDF_{b,t}$	500-999	13,920	3,540	25.4	18,820	10,700	56.9
MDF	500-999	13,920	3,463	24.9	18,820	10,540	56.0
CEF	1000+	6,510	1,510	23.2	11,670	6,432	55.1
HDF	1000+	6,510	1,505	23.1	11,650	6,403	55.0
$rHDF_{b,t}$	1000+	6,510	1,380	21.2	11,650	5,986	51.4
$rMDF_{b,t}$	1000+	6,299	1,346	21.4	10,700	5,450	50.9
MDF	1000+	6,299	1,314	20.9	10,700	5,357	50.1

Notes: Counts in thousands rounded to four significant digits to conform to current disclosure avoidance rules. The row $MDF_{b,t}$ (shaded) is the full reconstruction-abetted re-identification attack on 2010 Census data that have been processed through the 2020 Census Disclosure Avoidance System (DAS) using final production parameters and reported as tabular summaries with the same schema as the 2020 Census Demographic and Housing Characteristics (DHC) File, the successor to the 2010 Census Summary File 1, comparable to the tables in Table 2 in the main text. The row MDF (shaded) is only the re-identification attack run against the Microdata Detail File (MDF), the input to the tabular DHC publications. A version of the MDF, called the Privacy-Protected Microdata File (PPMF) will be released for the 2020 Census as a successor to the 2010 Census Public-Use Microdata Sample. Unshaded rows are identical to entries in Table S6. Row populations are shown in Table 3 in the main text.

Putative reidentifications, confirmed reidentifications, and precision rates for all data-defined persons using the 2020 Disclosure Avoidance System by census block size.

Table S13.

Data (L in Algorithm 3)	Block Population Range	Attacker (R in Algorithm 3): COMRCL			Attacker (R in Algorithm 3): CEF_{atkr}		
		Putative	Confirmed	Precision Percentage	Putative	Confirmed	Precision Percentage
CEF	All	80,970	73,010	90.2	210,100	195,400	93.0
HDF	All	79,670	71,570	89.8	205,200	190,300	92.7
$rHDF_{b,t}$	All	73,020	62,330	85.4	204,900	181,300	88.5
$rMDF_{b,t}$	All	40,110	30,990	77.3	95,410	75,320	78.9
MDF	All	40,000	30,600	76.5	95,400	74,560	78.2
CEF	1-9	2,854	2,833	99.3	6,893	6,886	99.9
HDF	1-9	2,265	2,222	98.1	5,241	5,177	98.8
$rHDF_{b,t}$	1-9	2,177	2,126	97.6	5,239	5,163	98.5
$rMDF_{b,t}$	1-9	337	256	75.8	729	570	78.2
MDF	1-9	337	255	75.7	729	569	78.1
CEF	10-49	21,840	21,050	96.4	52,270	51,630	98.8
HDF	10-49	21,400	20,570	96.1	50,660	49,930	98.6
$rHDF_{b,t}$	10-49	19,440	18,170	93.5	50,550	48,760	96.5
$rMDF_{b,t}$	10-49	6,109	4,805	78.7	12,540	10,230	81.6
MDF	10-49	6,103	4,779	78.3	12,540	10,190	81.2
CEF	50-99	19,930	18,450	92.6	49,020	47,250	96.4
HDF	50-99	19,760	18,250	92.4	48,180	46,380	96.3
$rHDF_{b,t}$	50-99	17,690	15,430	87.3	48,070	44,010	91.5
$rMDF_{b,t}$	50-99	8,521	6,632	77.8	18,030	14,480	80.3
MDF	50-99	8,510	6,575	77.3	18,030	14,370	79.7
CEF	100-249	20,040	17,640	88.0	52,740	48,730	92.4
HDF	100-249	19,970	17,540	87.9	52,200	48,150	92.3
$rHDF_{b,t}$	100-249	18,240	14,960	82.0	52,140	45,000	86.3
$rMDF_{b,t}$	100-249	11,810	9,207	78.0	27,050	21,620	79.9
MDF	100-249	11,780	9,095	77.2	27,040	21,410	79.2
CEF	250-499	9,482	7,860	82.9	26,690	23,200	86.9
HDF	250-499	9,468	7,831	82.7	26,520	23,000	86.7
$rHDF_{b,t}$	250-499	8,903	6,921	77.7	26,510	21,610	81.5
$rMDF_{b,t}$	250-499	7,236	5,620	77.7	18,140	14,350	79.1
MDF	250-499	7,208	5,525	76.6	18,130	14,170	78.2
CEF	500-999	4,995	3,879	77.7	15,430	12,510	81.1
HDF	500-999	4,991	3,866	77.5	15,370	12,420	80.8
$rHDF_{b,t}$	500-999	4,793	3,516	73.4	15,370	11,760	76.5

$rMDF_{b,t}$	500-999	4,376	3,272	74.8	12,450	9,468	76.0
MDF	500-999	4,354	3,201	73.5	12,450	9,319	74.9
CEF	1000+	1,822	1,293	71.0	7,066	5,219	73.9
HDF	1000+	1,821	1,290	70.8	7,056	5,200	73.7
$rHDF_{b,t}$	1000+	1,780	1,206	67.8	7,056	4,969	70.4
$rMDF_{b,t}$	1000+	1,723	1,198	69.5	6,477	4,606	71.1
MDF	1000+	1,713	1,169	68.3	6,476	4,523	69.8

Notes: Counts in thousands rounded to four significant digits to conform to current disclosure avoidance rules. The row $MDF_{b,t}$ (shaded) is the full reconstruction-abetted re-identification attack on 2010 Census data that have been processed through the 2020 Census Disclosure Avoidance System (DAS) using final production parameters and reported as tabular summaries with the same schema as the 2020 Census Demographic and Housing Characteristics (DHC) File, the successor to the 2010 Census Summary File 1, comparable to the tables in Table 2 in the main text. The row MDF (shaded) is only the re-identification attack run against the Microdata Detail File (MDF), the input to the tabular DHC publications. A version of the MDF, called the Privacy-Protected Microdata File (PPMF) will be released for the 2020 Census as a successor to the 2010 Census Public-Use Microdata Sample. Unshaded rows are identical to entries in Table S9.

Putative reidentifications, confirmed reidentifications, and precision rates for all modal persons using the 2020 Disclosure Avoidance System by census block size.

Data S1. uscensusbureau_recon_replication.xlsx
Instructions for accessing the complete replication archive.

Data S2. rhdf_bt_0solvar_extract.xlsx
Data for 29 reconstructed census tracts from the 2010 Census based on $rHDF_{b,t}$. All census blocks in these tracts have zero solution variability.

Data S3. rmdf_bt_0solvar_extract.xlsx
Data for the same 29 reconstructed census tracts as in Data S3 from the 2010 Census based on $rMDF_{b,t}$ for comparison purposes. The tables reconstructed in $rMDF_{b,t}$ were protected by the 2020 Census Disclosure Avoidance System, which is based on a differential privacy framework.

Data S4. 20230502-All tables with summary data.xlsx
Excel workbook with Tables 1-S13 in electronic format. Calculations in the main text based on these tables are shown in the sheet associated with each table.