

Evaluating the Impacts of Swapping on the US Decennial Census*

María Ballesteros
Harvard University
Cambridge, MA, USA

Cynthia Dwork
Harvard University
Cambridge, MA, USA

Gary King
Harvard University
Cambridge, MA, USA

Conlan Olson
Columbia University
New York, NY, USA

Manish Raghavan
MIT
Cambridge, MA, USA

Abstract

To meet its dual burdens of providing useful statistics and ensuring privacy of individual respondents, the US Census Bureau has for decades introduced some form of “noise” into published statistics. Initially, they used a method known as “swapping” (1990–2010). In 2020, they switched to an algorithm called TopDown that ensures a form of Differential Privacy. While the TopDown algorithm has been made public, no implementation of swapping has been released and many details of the deployed swapping methodology deployed have been kept secret. Further, the Bureau has not published (even a synthetic) “original” dataset and its swapped version. It is therefore difficult to evaluate the effects of swapping, and to compare these effects to those of other privacy technologies. To address these difficulties we describe and implement a parameterized swapping algorithm based on Census publications, court documents, and informal interviews with Census employees. With this implementation, we characterize the impacts of swapping on a range of statistical quantities of interest. We provide intuition for the types of shifts induced by swapping and compare against those introduced by TopDown. We find that even when swapping and TopDown introduce errors of similar magnitude, the direction in which statistics are biased need not be the same across the two techniques. More broadly, our implementation provides researchers with the tools to analyze and potentially correct for the impacts of disclosure avoidance systems on the quantities they study.

CCS Concepts

- **Security and privacy** → Social aspects of security and privacy;
- **Social and professional topics** → **Computing / technology policy**;
- **Applied computing** → Law.

Keywords

US Decennial Census, data swapping, privacy

*A full version of this paper can be found at <https://arxiv.org/pdf/2502.01320>. This work was supported in part by Sloan Foundation grant G-2017-9890, a Google Research Scholar award, and the Center for Research on Computation and Society (CRCS), Harvard.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CSLAW '25, München, Germany

© 2025 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 979-8-4007-1421-4/25/03
<https://doi.org/10.1145/3709025.3712210>

ACM Reference Format:

María Ballesteros, Cynthia Dwork, Gary King, Conlan Olson, and Manish Raghavan. 2025. Evaluating the Impacts of Swapping on the US Decennial Census. In *Symposium on Computer Science and Law (CSLAW '25)*, March 25–27, 2025, München, Germany. ACM, New York, NY, USA, 13 pages. <https://doi.org/10.1145/3709025.3712210>

1 Introduction

Census statistics are used for a wide variety of planning, administrative, and redistricting purposes. In order to meet its dual burdens of providing useful statistics and protecting the privacy of individual respondents, the US Census Bureau has for decades introduced noise into published statistics. From 1990–2010, the decennial census used a form of “swapping” [13, 28]. The details of the method are kept confidential.

In 2010, the Bureau determined that the existing disclosure avoidance system (DAS) was vulnerable to reconstruction attacks (see [4] for a detailed description of how the Bureau launched a linkage-assisted reconstruction attack on the 2010 data). In 2020, the Decennial Census deployed a new DAS known as TopDown, which seeks to provide formal differential privacy guarantees [2]. Unlike swapping, the TopDown algorithm is public.¹

At a high level, swapping works by choosing households that are at a particularly high disclosure risk—say, a household that is racially unique in its neighborhood—and “swapping” each such household with another household in a different region, “hiding” the unique household in an area where it may be less identifiable. The precise instantiation of this framework used by the Bureau is kept secret, and indeed, we do not even know whether any single algorithm was used across all states in the 1990–2010 decennial censuses.

Prior to 2020, although it was generally known that swapping was deployed, consumers of Census statistics tended to treat published data as ground truth. This became impossible with the announcement of the deployment of differential privacy, the publication of demonstration data and, eventually, the TopDown algorithm itself.² In a nutshell, consumers were forced to confront the “ground truth” fiction. Each DAS, swapping and TopDown, has both privacy properties and impacts on data analysis. We know that the Bureau chose TopDown because it can provide better privacy protection than that obtained with the level of swapping used in the 2010 census. In this work, we turn to the impact of these disclosure avoidance systems on data analysis.

¹https://github.com/uscensusbureau/DAS_2020_Redistricting_Production_Code

²Differential privacy does not rely on secrecy of the algorithm to maintain privacy for respondents.

Studying the impacts of disclosure avoidance systems is challenging due to a lack of both data and information. In the case of TopDown, the Census Bureau has provided the details of the algorithm and a series of “before and after” demonstration data releases to facilitate research (see Section 2 for details). No analogous releases exist for swapping. In this work, we facilitate research into the effects of swapping through a careful implementation of a swapping algorithm based on information released by the Census Bureau (see Section 3); specifically, we draw from a variety of publications, presentations, and statements by Bureau employees, as well as court documents, to create an approximation to the Bureau’s practices. For details that remain unknown, we conduct robustness checks to verify that our conclusions hold across a range of reasonable choices.

We use this implementation to study a variety of impacts that swapping has on the Census data. In Section 4, to investigate the basic statistical properties of swapping, we measure the errors that swapping introduces to population counts by race, and compare them to those of TopDown and a simple variant, ToyDown, created by Cohen et al. [12], at a specific setting and apportioning of the privacy parameter ϵ . We further examine detailed properties of swapping with respect to household size, mixed-race households, and racial entropy across geographies. Surprisingly, despite the fact that swapping targets households that are racially unique in their current location and “moves” them elsewhere, we find that swapping *increases* racial entropy. We provide a detailed analysis of why this is the case.

We next consider the impact of swapping on downstream uses of Census data for both public policy and social science (Section 5). Prior work has found that differential privacy biases statistics derived from ecological regression (ER) analyses, which are used in cases involving the Voting Rights Act [12]. We find that the same is true for swapping, but, crucially, swapping appears to bias these statistics *in the opposite direction*. We again provide intuition for why this might be the case. Thus, viewing swapping as a similarly noisy predecessor to differential privacy may obscure key differences in the direction of effects of that noise.

We further compare swapping to TopDown in terms of impact on social science applications, where recent work has shown that the noise introduced by the TopDown algorithm can lead to substantial variance in downstream analyses [20, 29, 33, 34, 37]. In particular, Mueller and Santos-Lozada [29] show that TopDown can introduce large errors in rural areas. We find that the same is true for swapping. We conclude by proposing a more general framework for using our methods to estimate the impacts of swapping on downstream statistics of interest.

To summarize, the key contributions of this work are as follows:

- We implement a swapping algorithm that we believe closely resembles the process used by the US Census Bureau based on a careful analysis of public-domain documents. Code for this implementation can be found at <https://github.com/ceolson/census-swapping>.
- Using this implementation, we provide a detailed description of the impacts of swapping and how those impacts compare to the impacts of TopDown and ToyDown.

- We highlight several important findings, including an increase in racial entropy due to swapping, biases in opposite directions to those induced by differential privacy, and the ability to replicate studies describing errors from differential privacy.
- More broadly, we offer a framework to estimate and correct for the effects of disclosure avoidance systems in downstream analyses.

2 Related Work

Disclosure Avoidance in the US Decennial Census. The US Census Bureau has a statutory obligation to protect the privacy of respondents under Title 13 of the U.S. Code. To do so, they have deployed a number of Disclosure Avoidance Systems (DASs) over the years [28]. Of note, for the purposes of this work, are two methods in particular: data swapping, used in 1990, 2000, and 2010; and TopDown, the Bureau’s implementation of differential privacy, used in 2020.

Data swapping, introduced by Dalenius and Reiss [13], works by exchanging the values contained in particular cells in a dataset while leaving aggregate statistics unchanged. For a survey on variants of swapping over the years, see Fienberg and McIntyre [17]. At a high level, the swapping algorithm used by the Census Bureau swapped the records associated with one household with those of another household in another location. While the exact details of this procedure are unknown, we compile a combination of reports and statements issued by the Census Bureau and related personnel [3, 19, 21, 28, 31, 36, 38–41] that collectively provide enough information to implement a close approximation to the original Census procedure. We provide further details in Section 3.

Beginning in 2020, the Census Bureau began using a new DAS based on differential privacy [2, 16]. Instead of swapping records, the Bureau’s algorithm (known as TopDown) adds noise to tabulated counts, providing formal privacy guarantees, and then, in a *post-processing* phase, solves a constrained optimization problem to find “nearby” noisy counts satisfying hierarchical consistency and other requirements (e.g., non-negativity) [1]. Unlike swapping, the algorithm itself is public knowledge; the realized values of the added noise are kept secret to protect privacy.

Our focus in this work is on the Decennial Census. Disclosure avoidance is also used in the American Community Survey (ACS), which collects more detailed information on a sample of the population (see Zayatz et al. [41]). The Census Bureau uses a variety of disclosure avoidance techniques to release ACS data, including swapping [30, Section 13.8], [9, 32].

Analyzing Census DASs. Research on the effects of swapping in the Census, especially in comparison with differential privacy, is closely related to our work [6, 11, 26]. Kim [26] implements a version of swapping for the 2007–2011 American Community Survey in Alleghany County, PA. Kim [26] provides results using both synthetic data at the tract level and privileged access to the true (unswapped) microdata through the Census Bureau. In contrast, our methods make analyses like this possible *without* privileged access, which may benefit future researchers, and they are sufficiently general to be used across the country. Christ et al. [11] compare the effects of swapping and differential privacy on both utility and privacy. To do so, they create synthetic data by combining blocks

from nine counties across the country from the 2010 Census. They provide two plausible implementations of swapping which, like Kim [26], operate by swapping individual records with an external pool of candidates that is representative of the US population. In contrast, our swapping implementation operates on *households*, as opposed to *individuals*, which is consistent with practices in the 1990–2010 Decennial Censuses.

Bailie et al. [6] give a swapping algorithm that provides a differential privacy guarantee and analyze it on Census data. In contrast, the swapping procedure used by the Census, which we aim to reconstruct here, has no formal differential privacy guarantees. Because swapping requires access to household-level data at the block level, Bailie et al. [6] evaluate their approach on the 1940 Census, for which the original microdata have been made public. Our approach instead generates synthetic microdata for the 2010 Census [15], to which we then apply our swapping procedure.

While it is still fairly new, the TopDown algorithm and its close variants have been the subject of a number of studies [12, 23, 25, 38]. Both Cohen et al. [12] and Kenny et al. [23] consider how TopDown will impact analyses related to voting rights. Because the original TopDown algorithm is difficult to run reliably, we adopt a simpler variant (“ToyDown”) introduced by Cohen et al. [12], similar to the “Reasonably Good Algorithm” of Long [27]. Wright and Irinata [38] analyze the reliability of TopDown outputs for redistricting applications. Kenny et al. [25] characterize the amount of noise introduced by the overall TopDown process. We build upon some of their techniques to calibrate a key parameter of our algorithm against released TopDown data. Beyond the uncertainty introduced by the DAS, Steed et al. [35] and Boyd and Sarathy [7] detail how other sources of error lead to inaccurate data and downstream policy impacts.

Synthetic data generation. Our swapping implementation requires access to household-level microdata, which are in general only made public by the Census Bureau after 72 years. Instead, we rely on *synthetic* microdata for the 2010 Census, derived from a combination of Census data products using techniques from Dwork et al. [15]. Synthetic data are not perfect: while they match tabular data released by the Census, they will differ in distribution and locations of households. Moreover, disclosure avoidance techniques have already been used prior to the release of the data products on which they are based, meaning that when we run our swapping implementation, it will effectively be on *already-swapped* data. Still, prior work has had to rely on a variety of heuristics to obtain household-level datasets [11, 12], and the implementation provided by Dwork et al. [15] provides a principled way to sample synthetic data. We provide further details on this approach in Figure 1.

3 Swapping Implementation

Swapping works by exchanging the details of pairs of households within a state. To a first approximation, the swapping “algorithm” is as follows: for each household in the state, with some probability, swap the details of that household with a randomly selected household in the same state. The exact implementation of this process is unknown; the Census Bureau has not publicly released the details of their swapping implementation. Indeed, the “swapping

algorithm” is not necessarily an entirely algorithmic procedure; the level of human involvement is unknown.

Nevertheless, we provide an implementation of swapping that is faithful to public disclosures made by the Census Bureau and its representatives [3, 21, 28, 31, 39–41]. All of these sources confirm the following framework for swapping:

- (1) Within a state, households are assigned a probability of being swapped based on their local uniqueness, which is determined by “flagging” variables [36].
- (2) If selected to be swapped, a household is paired with another household that matches on certain “key” variables [36]. The matching household is chosen based on its geographic proximity: neither too close nor too far [28].
- (3) The maximum number of swaps is determined by the “swap rate” [5, 36], which is the fraction of households in the state that are swapped.

Within this framework, “[t]he decision variables in the swapping program are the swap rate, the flagging variables, the matching key variables and the ranking of pairs” [31]. We consider these four decisions, plus additional implementation details, in what follows.

Swap rate. The swap rate is not public, but we have some information about reasonable ranges. Bailie et al. [6] claim that the swap rate is 2–4%. John Abowd, the former Chief Scientist at the U.S. Census Bureau, states that the swap rates that the Bureau considered when studying the feasibility of using swapping in 2020 were higher than the swap rates used in 2010 [3]. An example of the swap rates that the Bureau considered for 2020 appears in Hawes and Rodríguez [22] (also written by Bureau officials), in which the authors present swap rates in the range of 5–50%. Based on this information, we use 2% as a reasonable low swap rate and 10% as a reasonable high swap rate. In some places, we also present results for higher swap rates.

Flagging variables. Multiple sources [31, 41] highlight the importance of demographic uniqueness within a small geographic region for classifying households into these tiers. While the precise set of flagging variables used to determine reidentification risk are not made public, sources indicate that “key demographic variables” serve as flagging variables [41]. In this work, we choose flagging variables to be counts of individuals by race, the number of Hispanic individuals, the total number of persons, and the number of adults. To sort households by reidentification risk, we sort them by the number of other households in the same block that have the same values across all of the flagging variables.

Matching key variables. Documentation on matching variables is inconsistent. According to Zayatz [39], “The key variables were number of people in the household of each race by Hispanic/Non-Hispanic by age group (<18,18+), number of units in building, rent/value of home, and tenure.” Griffin et al. [19] states that the planned swapping methodology matches on “number of persons in household, population characteristics of race, Hispanic origin and age, and on housing characteristics of units at building, rent/value and tenure.” If true, however, this would guarantee that block-level counts by race and ethnicity are invariant to swapping, which goes against assumptions commonly held in the literature [e.g., 11, 25]. Moreover, Steel and Zayatz [36] appear to contradict this in their

evaluation of the impacts of swapping, writing “Our procedure targeted households with unique characteristics, and this increased the number of zero cells (for these characteristics) because it tended to draw people from cells with values 1, 2, and 3 (presumably where the entire contribution to the cell is from one household and hence selected for high disclosure risk) and swap them into cells with larger values.”

Abowd [3] (the chief scientist at the Census Bureau) offers an even more clear contradiction in a sworn declaration: “For the 2000 and 2010 censuses, those key matching characteristics were (1) the whole number of persons in the household, and (2) the whole number of persons aged 18 or older in the household.” Wright and Irimata [38] corroborate this, writing, “the SWA [swapped] TOTAL counts and the corresponding CEF [unswapped] TOTAL counts at the block level were the same in 2010. The same is true for TOTAL18 counts for the 18 years and over population at the block level.” Kenny et al. [24] draw the same conclusion from Wright and Irimata [38]. One potential explanation for this apparent contradiction is that the extended set of matching characteristics described by Zayatz [39] and Griffin et al. [19] were used for the 1990 Census, and a smaller set of matching characteristics were used in 2000 and 2010 as per Abowd [3]. In this work, we take Abowd’s definition of key matching characteristics: number of persons and number of persons aged 18 or older.

Ranking of pairs. According to McKenna [28], “A household record is typically swapped with another household within a large area but in a different smaller area within the larger one, for example, across tracts but within the same county.” We were unable to find more specific information than this on how match partners were chosen. In our implementation, we require households to be swapped with a household from a different tract in the same state. Combining this requirement with the requirement that swap partners match on key variables, we can define the set of potential swap partners to be the households that match key variables, are in a different tract, and are in the same state. Among these potential swap partners, we prioritize the geographically closest households.³

Additional implementation details. There are more minor implementation design decisions that we describe in the supplementary materials. Since we have no way of knowing which decisions the Census Bureau made, we study two variants of our swapping implementation. We use one “standard” implementation of swapping to conduct the experiments reported in this work. We also make an alternative implementation of swapping, which we call our “high variance” implementation because it is tuned to have a higher variance than our standard implementation. The details of the differences between these variants are in the supplementary materials.

We run all of our experiments using the high variance implementation of swapping and find that our conclusions are largely unchanged. Results for the high variance implementation are not included in this work (because they are similar to the results for the standard implementation), with the exception of the variance

estimation results, which we report for both our standard and high variance implementations.

Many of the choices we make are heuristics. As such, the analyses we provide in the remainder of the paper should not be viewed as authoritative claims about the exact impacts of swapping. Instead, we view our implementation as approximately correct in its overall approach, and the analyses we produce are reasonable estimates in terms of both magnitude (as a function of the swap rate) and direction.

4 Characterizing the Effects of Swapping

With this implementation, we are ready to analyze the effects of swapping. In this section, we focus on basic descriptive statistics of the changes induced by swapping to the microdata. In particular, we focus on race and household size. Section 5 considers the use of Census data in specific downstream applications.

4.1 Setup and notation

We begin by introducing our notation and set up. We view a database as a list of households $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$. From the perspective of our swapping implementation, each household \mathbf{x}_i is described by the count of people of each race, the count of Hispanic people, the count of people over 18, and the Census block where the household is located.

Let \mathbf{X}_{CEF} be the Census Edited File (CEF), which is the unprotected Census microdata (which we never observe). Let $\mathbf{X}_{\text{Released}}$ be the released Census data (which is the result of applying swapping and other disclosure avoidance methods to the CEF). Let $\mathbf{X}_{\text{Synthetic}}$ be the unmodified synthetic database. Note that $\mathbf{X}_{\text{Released}}$ and $\mathbf{X}_{\text{Synthetic}}$ match on any statistic aggregated to the Census block level by construction [15].

By running our swapping implementation on $\mathbf{X}_{\text{Synthetic}}$, we can produce swapped databases. In this section, we consider swapping at a 2% swap rate and a 10% swap rate (representing the limits of the range of what we believe are realistic swap rates). Let $\mathbf{X}_{\text{Swapped2}}$ be the result of applying swapping at a 2% swap rate to $\mathbf{X}_{\text{Synthetic}}$ and let $\mathbf{X}_{\text{Swapped10}}$ be the result of applying swapping at a 10% swap rate to $\mathbf{X}_{\text{Synthetic}}$.

By running ToyDown, the simplified version of TopDown implemented by Cohen et al. [12], on our synthetic data, we can also produce $\mathbf{X}_{\text{ToyDown}}$. To produce this database, we run ToyDown with $\epsilon = 3.26$ and an equal split of privacy budget across geographic levels. This value of ϵ was chosen by calibrating the variance of ToyDown against the variance of TopDown in demonstration data (see Section 4.3 for details). Technically, our implementation of ToyDown differs from the one used in [12] because we do not treat Hispanic as a race category to be consistent with our swapping algorithm.

Finally, we have data produced by TopDown, the Census Bureau’s real differential privacy-based disclosure avoidance system, from demonstration data released to help consumers prepare for the new DAS. This demonstration data included P.L. 94-171 data from two independent runs of the TopDown algorithm under production settings, one released in 2021 [8] and the other released in 2023 [10].

³Each household is assigned a geographic location based on the block in which it appears.

We call these two databases $X_{\text{TopDown2021}}$ and $X_{\text{TopDown2023}}$. Figure 1 shows the databases that we consider and the relationships between them.

For a race $r \in \mathcal{R}$ and a geographic level Γ (like the set of Census blocks in a state s , which we denote \mathcal{B}_s , or the set of counties in a state s , which we denote C_s) we can aggregate households to obtain counts of that race for each geographic unit in that geographic level. For $r \in \mathcal{R}$, $\gamma \in \Gamma$, and a database X , we write the number of people of race r in region γ according to the database X as $c_\gamma^r(X)$.

In our experiments, we study five states: Alabama, Wisconsin, Texas, Vermont, and Nevada. These states were selected to cover a wide range of sizes, demographic characteristics, and geographic locations. When showing detailed results about one state, we use Alabama. In many cases, corresponding results for other states are in the supplementary materials.

4.2 Quantifying the Errors Introduced by Swapping

We begin by studying the bias of swapping on race counts. It is important to remember that the *state-wide average* bias of the count of any race is zero, since swapping (unlike TopDown) does not change the statewide counts of each race. Therefore, we study the error in the individual counts at lower geographic levels. These errors will average to zero across the entire state, but we are interested in their magnitudes.

For a race r and a geographic region γ , we can define the error

$$\text{Error}_\gamma^r(X_1, X_2) \triangleq c_\gamma^r(X_1) - c_\gamma^r(X_2).$$

Ideally, we would estimate quantities $\text{Error}_\gamma^r(X_{\text{CEF}}, X_{\text{Released}})$, which represent the error in race counts due to the Bureau’s implementation of swapping applied to the underlying Census data (shown in blue in Figure 1). This, however, is impossible without privileged access to confidential Census microdata. Instead, we study the quantity $\text{Error}_\gamma^r(X_{\text{Synthetic}}, X_{\text{Swapped}})$, which we can define for swapping at a 2% swap rate or a 10% swap rate. Note that this, too, is the result of applying the swapping to an underlying dataset. However, it differs from our ideal quantity of interest $\text{Error}_\gamma^r(X_{\text{CEF}}, X_{\text{Released}})$ in two key ways: (1) the “unswapped” data we use are synthetic microdata, and (2) the swapping algorithm used is our own implementation as opposed to the Bureau’s. Our conclusions are valid insofar as these two components (the synthetic data and the swapping algorithm) are sufficiently close to their real-world analogues (the true microdata and the Bureau’s swapping procedure).

For comparison, we also consider two baselines using differential privacy. The first is a direct comparison with ToyDown, which we apply to the same synthetic data: $\text{Error}_\gamma^r(X_{\text{Synthetic}}, X_{\text{ToyDown}})$. In addition, we compare the TopDown demonstration data against the released 2010 Census data, yielding $\text{Error}_\gamma^r(X_{\text{Released}}, X_{\text{TopDown2021}})$. We caution that $\text{Error}_\gamma^r(X_{\text{Released}}, X_{\text{TopDown2021}})$ should not be directly compared with the three other quantities $\text{Error}_\gamma^r(X_{\text{Synthetic}}, X_{\text{Swapped2}})$, $\text{Error}_\gamma^r(X_{\text{Synthetic}}, X_{\text{Swapped10}})$, and $\text{Error}_\gamma^r(X_{\text{Synthetic}}, X_{\text{ToyDown}})$; these three quantities measure the error introduced by applying a single process (swapping or ToyDown, respectively) to a dataset, while $\text{Error}_\gamma^r(X_{\text{Released}}, X_{\text{TopDown2021}})$ measures the error between the results of *two different* processes—the Census Bureau’s swapping method and TopDown—being applied to the

same underlying dataset.⁴ See Figure 1 for further intuition for these comparisons.

Figure 2 shows $\text{Error}_\gamma^r(X_{\text{Synthetic}}, X_{\text{Swapped2}})$, $\text{Error}_\gamma^r(X_{\text{Synthetic}}, X_{\text{Swapped10}})$, $\text{Error}_\gamma^r(X_{\text{Synthetic}}, X_{\text{ToyDown}})$, and $\text{Error}_\gamma^r(X_{\text{Released}}, X_{\text{TopDown2021}})$ at the county level in Alabama. As noted earlier, the errors for swapping (at both 2% and 10%) are both mean zero.⁵ We observe that errors due to swapping (Panels A and B) are generally greater than errors due to ToyDown (Panel C). Also, swapping results in higher errors on large racial groups, while ToyDown distributes errors evenly across racial groups. The picture for TopDown (Panel D) may superimpose the errors due to swapping and TopDown, which is consistent with the fact that both swapping and TopDown stand between X_{TopDown} and X_{Released} .

For downstream applications (for example, any application using the proportion of each race) errors relative to population size are also important. To evaluate relative errors, we use the definition

$$\text{RelativeError}_\gamma^r(X_1, X_2) \triangleq \frac{2}{1 + \frac{c_\gamma^r(X_1)}{c_\gamma^r(X_2)}}$$

as a measure of the effect of swapping on the count of race r in region γ , normalized by the size of the count. If the count increases by a relatively large amount from X_1 to X_2 , $\text{RelativeError}_\gamma^r(X_1, X_2)$ will be close to 2. If the count decreases by a relatively large amount, $\text{RelativeError}_\gamma^r(X_1, X_2)$ will be close to 0. If the count is unchanged, $\text{RelativeError}_\gamma^r(X_1, X_2) = 1$. If both counts are 0, we treat the relative error as 1. In practice, the way we treat zeros is not very important because the only county-level race counts that are zero are the number of Hawaiian or Pacific Islander people in five counties (before swapping).

Again, our ideal comparison is $\text{RelativeError}_\gamma^r(X_{\text{CEF}}, X_{\text{Released}})$. However, since we do not observe X_{CEF} , we instead measure $\text{RelativeError}_\gamma^r(X_{\text{Synthetic}}, X_{\text{Swapped}})$ (defined for swapping at a 2% swap rate and a 10% swap rate), as well as $\text{RelativeError}_\gamma^r(X_{\text{Synthetic}}, X_{\text{ToyDown}})$ and $\text{RelativeError}_\gamma^r(X_{\text{Released}}, X_{\text{TopDown2021}})$. These are shown in Figure 3. We observe that 2% swapping has smaller relative errors than ToyDown, but 10% swapping has comparable relative errors to ToyDown. Both swapping and ToyDown exhibit very low relative error on large racial groups and much larger relative error on small racial groups (especially with a 10% swap rate, but often with a 2% swap rate as well). In the case of ToyDown, this is consistent with the fact that ToyDown distributes *absolute* errors relatively evenly across racial groups, so groups with smaller sizes have larger *relative* errors. In the case of swapping, it appears that, while swapping has larger absolute errors on larger racial groups, this is more than canceled out by the effect of the group sizes. The overall effect is that swapping, at least at a 10% swap rate, introduces larger relative errors to smaller groups.

⁴As an illustrative note, a direct comparison to $\text{Error}_\gamma^r(X_{\text{Released}}, X_{\text{TopDown2021}})$ using synthetic data would be the quantity $\text{Error}_\gamma^r(X_{\text{Swapped}}, X_{\text{ToyDown}})$.

⁵The swap errors do not necessarily have *median* zero; the blue lines in the boxplots are the medians.

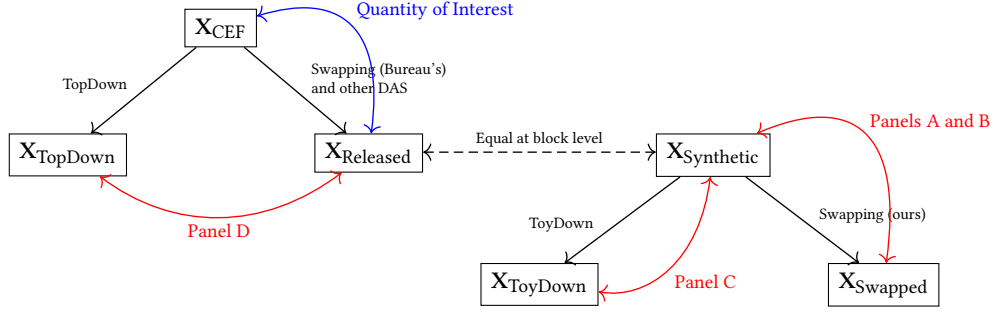


Figure 1: Various versions of the Census data and the relationships between them. Note that there are two releases of X_{TopDown} , which we call $X_{\text{TopDown}2021}$ and $X_{\text{TopDown}2023}$. We also produce multiple versions of X_{Swapped} : $X_{\text{Swapped}2}$ is produced by swapping at a 2% swap rate and $X_{\text{Swapped}10}$ is produced by swapping at a 10% swap rate. The red arrows show the comparisons made in the panels in Figures 2 and 3. The blue arrow shows the comparison of interest (which we cannot directly observe).

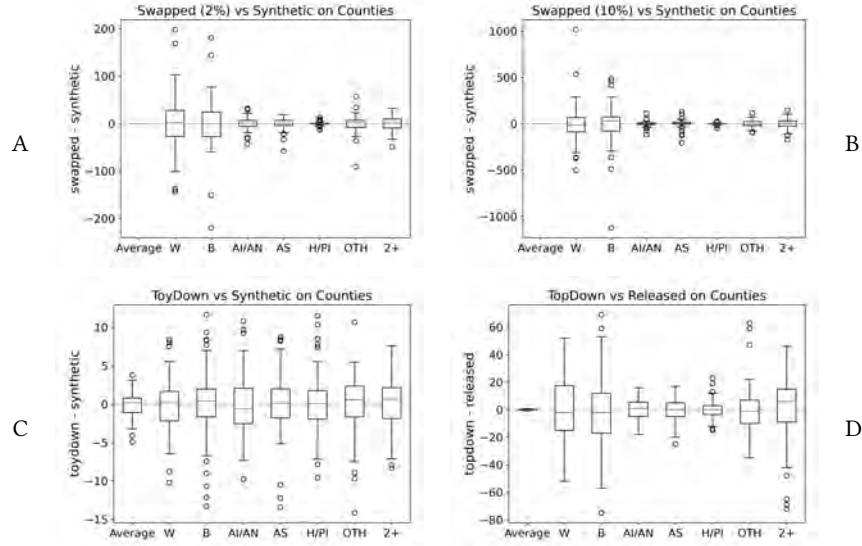


Figure 2: Plots of the four quantities $\text{Error}_\gamma^r(X_{\text{Synthetic}}, X_{\text{Swapped}2})$, $\text{Error}_\gamma^r(X_{\text{Synthetic}}, X_{\text{Swapped}10})$, $\text{Error}_\gamma^r(X_{\text{Synthetic}}, X_{\text{ToyDown}})$, $\text{Error}_\gamma^r(X_{\text{Released}}, X_{\text{TopDown}2021})$ for all $\gamma \in C_{\text{Alabama}}$, where C_{Alabama} is the set of all counties in Alabama. To generate the “Average” plot, for each county, the errors for each race were averaged. Note that the scale of the y-axes are different in different plots.

4.3 Quantifying the Variance of Swapping

In addition to biasing quantities of interest, because they are randomized algorithms, both swapping and TopDown introduce variance into their outputs. Here, we compare the variance they introduce into block-level counts by race. Our techniques draw inspiration from Kenny et al. [25], who establish estimators for the bias and variance introduced by TopDown. Bailie et al. [6] also compare swapping at different rates against TopDown, but with one key difference: while we compare the data *accuracy* of swapping and TopDown, they compare the formal *privacy guarantees* of swapping and TopDown. Though we do not study formal privacy guarantees, our study of variance is motivated by its connection to privacy. Noise is necessary to ensure privacy [14], and the variance of a statistical process is a measure of how much noise it introduces.

Similarly to Kenny et al. [25], define

$$\varepsilon_\gamma^r(X_{\text{Swapped}}) \triangleq c_\gamma^r(X_{\text{Swapped}}) - \mathbb{E}_{X \sim \text{SWAPPING}} [c_\gamma^r(X)],$$

where $X \sim \text{SWAPPING}$ indicates that X is generated by running swapping with fresh randomness. With this, the variance of the counts by race at the block level can be written

$$V_{\text{swap}} \triangleq \frac{1}{|\mathcal{B}_s| |\mathcal{R}|} \sum_{\gamma \in \mathcal{B}_s} \sum_{r \in \mathcal{R}} \text{Var}(\varepsilon_\gamma^r(X_{\text{Swapped}})),$$

where \mathcal{B}_s is the set of blocks in a state s and \mathcal{R} is the set of Census race categories.

We do not observe the expected counts $\mathbb{E}_{X \sim \text{SWAPPING}} [c_\gamma^r(X)]$, so we can't directly compute $\varepsilon_\gamma^r(X_{\text{Swapped}})$. However, we can produce an unbiased estimate of V_{Swapped} by running swapping twice to

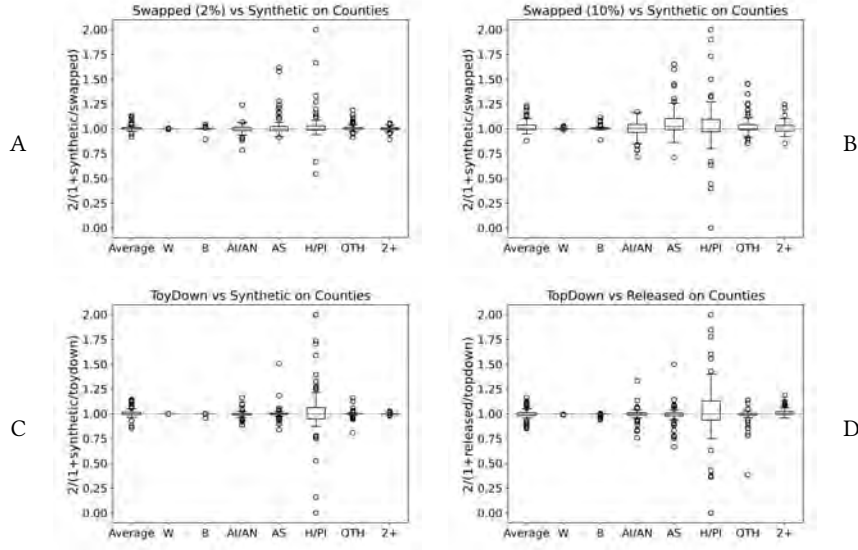


Figure 3: Plots of the four quantities $\text{RelativeError}_\gamma^r(\mathbf{X}_{\text{Synthetic}}, \mathbf{X}_{\text{Swapped}2})$, $\text{RelativeError}_\gamma^r(\mathbf{X}_{\text{Synthetic}}, \mathbf{X}_{\text{Swapped}10})$, $\text{RelativeError}_\gamma^r(\mathbf{X}_{\text{Synthetic}}, \mathbf{X}_{\text{ToyDown}})$, and $\text{RelativeError}_\gamma^r(\mathbf{X}_{\text{Released}}, \mathbf{X}_{\text{TopDown}2021})$ for all $\gamma \in \mathcal{C}_{\text{Alabama}}$, where $\mathcal{C}_{\text{Alabama}}$ is the set of all counties in Alabama. To generate the “Average” plot, for each county, the relative errors for each race were averaged.

produce $\mathbf{X}_{\text{Swapped}}$ and $\mathbf{X}'_{\text{Swapped}}$. This yields the estimator

$$\widehat{V}_{\text{Swapped}} \triangleq \frac{1}{2|\mathcal{B}_s||\mathcal{R}|} \sum_{\gamma \in \mathcal{B}_s} \sum_{r \in \mathcal{R}} \left(c_\gamma^r(\mathbf{X}_{\text{Swapped}}) - c_\gamma^r(\mathbf{X}'_{\text{Swapped}}) \right)^2.$$

Note that $\mathbb{E}[\widehat{V}_{\text{Swapped}}] = V_{\text{Swapped}}$ since

$$\begin{aligned} & \mathbb{E} \left[\left(c_\gamma^r(\mathbf{X}_{\text{Swapped}}) - c_\gamma^r(\mathbf{X}'_{\text{Swapped}}) \right)^2 \right] \\ &= \mathbb{E} \left[\left(\varepsilon_\gamma^r(\mathbf{X}_{\text{Swapped}}) - \varepsilon_\gamma^r(\mathbf{X}'_{\text{Swapped}}) \right)^2 \right] \\ &= 2 \mathbb{E} \left[\varepsilon_\gamma^r(\mathbf{X}_{\text{Swapped}})^2 \right] - 2 \mathbb{E} \left[\varepsilon_\gamma^r(\mathbf{X}_{\text{Swapped}}) \varepsilon_\gamma^r(\mathbf{X}'_{\text{Swapped}}) \right] \\ & \quad (\varepsilon_\gamma^r(\mathbf{X}_{\text{Swapped}}), \varepsilon_\gamma^r(\mathbf{X}'_{\text{Swapped}}) \text{ are i.i.d.}) \\ &= 2 \text{Var}(\varepsilon_\gamma^r(\mathbf{X}_{\text{Swapped}})). \end{aligned}$$

Following a similar approach for TopDown, we can define

$$\varepsilon_\gamma^r(\mathbf{X}_{\text{TopDown}}) \triangleq c_\gamma^r(\mathbf{X}_{\text{TopDown}}) - \mathbb{E}_{\mathbf{X} \sim \text{TopDown}} [c_\gamma^r(\mathbf{X})],$$

where $\mathbf{X} \sim \text{TopDown}$ indicates that \mathbf{X} is generated by running TopDown with fresh randomness. Then, we have

$$V_{\text{TopDown}} \triangleq \frac{1}{|\mathcal{B}_s||\mathcal{R}|} \sum_{\gamma \in \mathcal{B}_s} \sum_{r \in \mathcal{R}} \text{Var}(\varepsilon_\gamma^r(\mathbf{X}_{\text{TopDown}}))$$

Following [25], we use the two publicly available TopDown protected datasets $\mathbf{X}_{\text{TopDown}2021}$ and $\mathbf{X}_{\text{TopDown}2023}$ to form the estimator

$$\widehat{V}_{\text{TopDown}} \triangleq \frac{1}{2|\mathcal{B}_s||\mathcal{R}|} \sum_{\gamma \in \mathcal{B}_s} \sum_{r \in \mathcal{R}} \left(c_\gamma^r(\mathbf{X}_{\text{TopDown}2021}) - c_\gamma^r(\mathbf{X}_{\text{TopDown}2023}) \right)^2.$$

While we are unable to precisely characterize the variance of the TopDown estimator $\widehat{V}_{\text{TopDown}}$ (since we only observe two samples), we can estimate $\text{Var}(\widehat{V}_{\text{Swapped}})$ since we can sample it as many times as we like (to obtain each sample, we run swapping twice to generate two independent swapped datasets). In Figure 4, we show the estimated variance of swapping for different swap rates. Because the variance of swapping is a function of many parameters of the swapping algorithm, we report results for both our standard implementation of swapping and our high variance implementation of swapping. While we believe these two versions of swapping are realistic, we emphasize that our variance estimates are not necessarily close to the variance of the Bureau’s swapping method.

For comparison, we also plot the estimated variance of TopDown as a horizontal line. First, we observe that the variance estimates for both processes are different across different states. Generally, all variance estimates are higher in larger states. Second, in all states, it takes a high swap rate to reach the same variance as TopDown (in some states, it appears there is no swap rate resulting in variance as high as TopDown). Swapping at the rates that we believe the Census realistically used has lower variance than TopDown. We caution that we do not claim that the swap rate should be the rate that yields equal variance to that of TopDown. This is because (1) it is not clear that the Bureau was interested in equalizing the variances of the 2010 DAS and the 2020 DAS, (2) our variance estimates of swapping do not necessarily estimate the variance of the Bureau’s swapping algorithm, and (3) our two versions of swapping (standard and high variance) yield significantly different swap rates at which swapping and TopDown have comparable variances.

Given the connection between variance and privacy described above, these observations suggest that our swapping algorithm with realistic swap rates likely has lower variance, and therefore

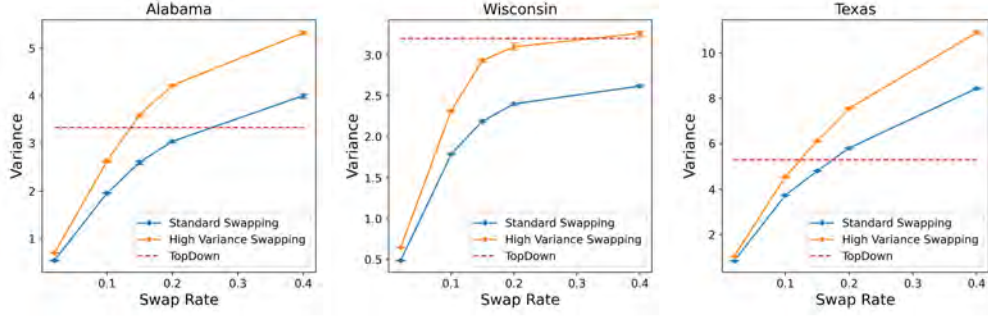


Figure 4: The variance of swapping at different swap rates. Error bars show the minimum and maximum of 5 runs of our estimator. The horizontal dotted lines show the variance estimate for TopDown. Note that we do not characterize the error in our estimate for TopDown. Plots for two more states appear in the supplementary materials.

lower privacy, than TopDown. This observation is consistent with statements made by Abowd [3].

We also use the above method for variance estimation to calibrate ϵ for the ToyDown data used in Section 4.2 to match the variance of TopDown. For different values of ϵ , we run ToyDown twice, producing $\mathbf{X}_{\text{ToyDown},\epsilon}$ and $\mathbf{X}'_{\text{ToyDown},\epsilon}$. Then, we measure

$$\widehat{V}_{\text{ToyDown}}(\epsilon) \triangleq \frac{1}{2|\mathcal{B}_s||\mathcal{R}|} \sum_{\gamma \in \mathcal{B}_s} \sum_{r \in \mathcal{R}} (c_Y^r(\mathbf{X}_{\text{ToyDown},\epsilon}) - c_Y^r(\mathbf{X}'_{\text{ToyDown},\epsilon}))^2.$$

We find that $\widehat{V}_{\text{ToyDown}}(3.26) \approx \widehat{V}_{\text{TopDown}}$, suggesting that ToyDown with $\epsilon = 3.26$ provides roughly the same amount of variance as the TopDown implementation run by the Census Bureau.

4.4 Effect of Swapping on Racial Entropy

Here, we consider how swapping impacts the racial distribution of the population. We use entropy as our primary measure of interest here. Higher racial entropy would indicate that populations are more geographically segregated. Here, we focus our analysis on Alabama. Formally, define the racial entropy of a geographic area to be the entropy of the empirical distribution of race in that area; i.e., for a geographic area γ ,

$$\text{RacialEntropy}_\gamma(\mathbf{X}) = - \sum_{r \in \mathcal{R}} \frac{c_Y^r(\mathbf{X})}{\sum_{r' \in \mathcal{R}} c_Y^{r'}(\mathbf{X})} \log \left(\frac{c_Y^r(\mathbf{X})}{\sum_{r' \in \mathcal{R}} c_Y^{r'}(\mathbf{X})} \right).$$

Because swapping specifically targets rare households and moves them to a different tract, we might expect it to have a homogenizing effect, resulting in reduced tract-level racial entropy. Surprisingly, we find that, on average, swapping increases the racial entropy of each tract. The average entropy increases from unswapped data to swapped data at a 2% rate and again from swapped data at a 2% rate to swapped data at a 10% rate. Table 1 shows the average entropy of tracts before and after swapping.

To see why this is the case, observe that a household can be targeted for swapping on the basis of things besides having uncommon races.

- (1) Household size is a non-racial factor to swapping, and it turns out to be quite influential in determining which households are swapped.
- (2) Households may be swapped not because they have uncommon races, but because their exact composition of races is uncommon. One reflection of this is that mixed race households are over-represented among the households involved in swapping.

The more independently of race that swapping happens, the more it increases racial entropy instead of decreasing it. Race is still important in swapping, but does not fully determine who is swapped. The following statistic illustrates this: in Alabama, white people make up 69% of the total population and 39% of the people targeted in swapping at a 10% swap rate. We see that white people (i.e. people with the statewide majority race) are indeed under-represented among the population targeted for swapping, but they are still swapped often. This is because they may live in households of unusual sizes or with unusual racial compositions.

Another reason that swapping increases entropy, on average, is because many components of swapping are targeted only minimally. Conceptually, we can break each swap down into four steps: (1) the targeted household leaving its original location, (2) the targeted household arriving at its new location, (3) the partnered household leaving its original location, and (4) the partnered household arriving at its new location. The intuition that swapping decreases entropy by removing uncommon households only holds for the first step. This is illustrated in Figure 5, which shows the effect on entropy of each step of swapping. The effect of step (1) (removing all the targeted households) indeed decreases racial entropy. Step (2) (removing all the partner households) also slightly decreases racial entropy. However, steps (3) and (4) (adding all the targeted households to their partners' tracts and adding all the partner households to the tracts of the targets) both have the effect of increasing racial entropy. The overall effect is to increase racial entropy.

We study several other interesting impacts of swapping in the supplementary materials. These include the rates of swapping on different household sizes, the effects of swapping on the distribution of people by race, and statistics describing which races are swapped for which.

State	Average tract entropy before swapping	Average tract entropy after 2% swapping	Average tract entropy after 10% swapping
AL	0.598	0.603	0.624
WI	0.431	0.434	0.447
TX	0.807	0.810	0.819
NV	0.997	0.999	1.006
VT	0.232	0.233	0.241

Table 1: Average racial entropy at the tract level before and after swapping.

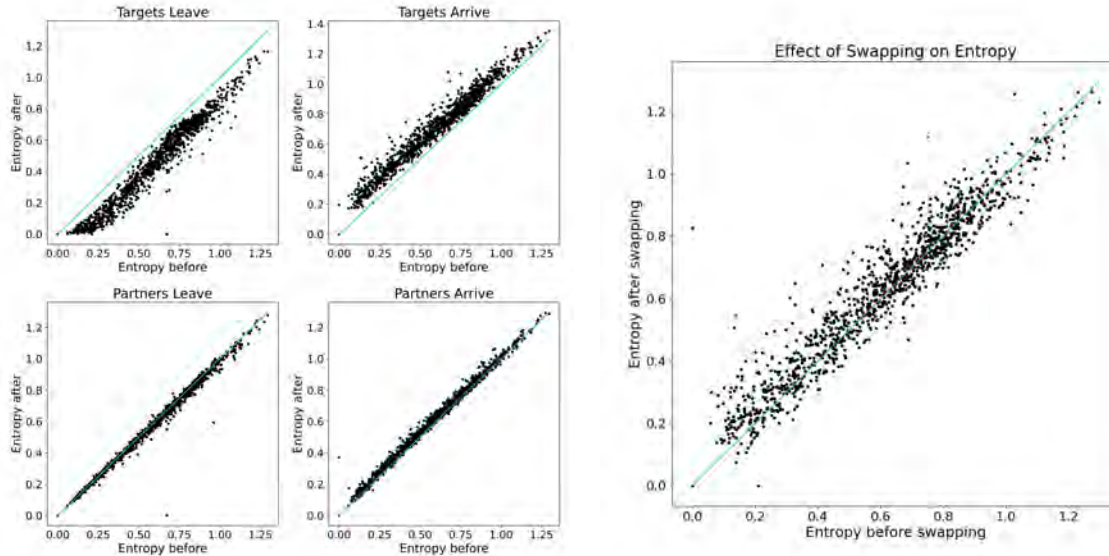


Figure 5: The effect of swapping on racial entropy at the tract level for a 10% swap rate in Alabama. The left panel shows the effect of each step of swapping on entropy. The right panel shows the overall effect. Corresponding plots for 2% and 10% swap rates in other states appear in the supplementary materials.

5 Downstream Impacts of Swapping

Having established basic characteristics of swapping, we next turn to its impacts on downstream analyses that rely on Census data. We consider two prototypical analyses in public policy (Section 5.1) and social science (Section 5.2) settings. Our results suggest that swapping can be a significant source of error in downstream analyses, and we suggest a general framework using our swapping implementation to estimate how swapping will impact other analyses in Section 5.3.

5.1 Voting

The estimation of voting patterns by race is crucial for applications such as enforcing the Voting Rights Act requirement that voting districts are drawn in a way that allows minority voters to elect their preferred candidates [18]. Given many sets of marginal counts of votes and racial groups (one for each voting district or precinct), two common methods are used to infer voting preferences: ecological regression (ER) and ecological inference (EI). Following Cohen et al. [12], we focus on ER.

In ecological regression, the support of each racial group for each candidate is learned separately. To learn the support of racial group r for candidate c , ecological regression is a simple linear regression with the percent of race r in the voting area as the explanatory variable and the percent of votes for c as the response variable. This regression line is used to predict the percent of votes for candidate c in a hypothetical voting district composed of all people of race r , which represents the support of race r for candidate c .

In Section 4.4, we observed that swapping has the effect of increasing racial entropy. If we compare ecological regression before and after swapping, swapping will tend to move the values for percent of race r away from 0 and 1 and towards the middle. This increases the magnitude of the slope of the regression line, resulting in greater apparent polarization (i.e., the support of different racial groups for a given candidate will vary more from the overall support for that candidate).

This effect is illustrated in Figure 6, which is the result of running ER on Alabama using election data from the 2018 lieutenant gubernatorial general election between Will Ainsworth, a white Republican candidate, and Will Boyd, a Black Democratic candidate. Since ER requires extrapolating to a hypothetical voting district

with 100% of race r , it is only reliable for racial groups r that have reasonably high percentages in many of the real voting districts. Therefore, Figure 6 only shows white and Black voter support for the Democratic and Republican candidates. We include the result of ER for all combinations of races and candidates in the supplementary materials.

The regression lines based on swapped data (magenta) have slopes of greater magnitude than the regression lines based on unswapped data (blue). This means that estimates of racial support based on swapped data are slightly more polarized than estimates of racial support based on unswapped data. For Figure 6, the data was swapped at a 10% swap rate.

On the other hand, the regression lines based on ToyDown protected data (orange) have slopes of slightly smaller magnitude than those based on unswapped data. This replicates the effect found by Cohen et al. [12], and occurs because the addition of noise to the covariates reduces the correlation between the covariates and the outcome variable, resulting in a regression line with a slope of smaller magnitude. To create Figure 6, we used ToyDown with $\epsilon = 0.25$ to exaggerate the effect of ToyDown. Since ϵ and the swap rate affect the magnitude of the effects, it is not meaningful to compare the relative magnitudes of the effect due to swapping versus the effect due to ToyDown. Our key observation is simply that the effect of swapping on the slope of ER regression lines is in the opposite direction from the effect of ToyDown on the slope of ER regression lines.⁶

Like Cohen et al. [12], we observe that weighting ER by the population of the voting districts reduces the effect of ToyDown to almost nothing. In contrast, weighting does not appear to reduce the effect of swapping in the same way. The result of ER with weighting appears in the supplementary materials. This suggests that interventions that mitigate errors due to differential privacy may not hold for swapping, and perhaps vice versa. Our results indicate that the effects of swapping can differ substantially from those of differential privacy.

5.2 Social science

The use of DP in the 2020 Census disclosure avoidance system has caused concern about the impact of TopDown on important analyses in social sciences [20, 29, 33, 34, 37]. Here, we choose one such study and investigate whether the concerns it raises also hold for swapping. In particular, we build upon a study by Mueller and Santos-Lozada [29], who observe that TopDown introduces significant errors in population and race counts in rural areas.

With our methods, we can carry out an analogous analysis on the impact of swapping, shown in Figure 7. We replicate the qualitative observations of Mueller and Santos-Lozada [29] and make new observations about swapping.⁷ We conclude that, while differential privacy indeed introduces problematic relative errors in small racial groups in rural areas, so does swapping. Even at a 2% swap rate

⁶We compare swapping to ToyDown to align our results with those of [12], but we also ran ER with swapping compared to TopDown and observed qualitatively similar effects.

⁷Our results are based only on Alabama, Nevada, Texas, Vermont, and Wisconsin, while the results in Mueller and Santos-Lozada [29] are based on the whole U.S. While we limit our analyses to those states for computational reasons, in principle, it would be possible to use our methods to swap the whole country.

(which we believe to be conservative), the relative errors introduced by swapping in counts of small racial groups in rural areas are as great or greater than those introduced by TopDown.⁸ Note that total population counts are not affected by swapping, so the errors introduced by swapping are always zero (compared to small, but nonzero, errors introduced by TopDown).

Our results have broader implications for social science that relies on Census data. In general, these data already contain a variety of sources of error or noise including nonresponse and misreporting [7, 35]. While differential privacy is an additional source of noise for 2020 and beyond, swapping may have been an equally significant (albeit harder to quantify) source of error for analyses based on 1990–2010 censuses. We conclude with a discussion of how researchers might seek to account for these errors using our implementation.

5.3 A General Framework

Here, we offer a general recipe for researchers seeking to characterize the effects of swapping on a particular statistic. Suppose one is interested in a statistic s that depends on Census data $\mathbf{X}_{\text{Released}}$ and some auxiliary information \mathbf{Z} . Note that for Census years 1990–2010, $\mathbf{X}_{\text{Released}}$ has already been swapped by the Census Bureau prior to its release. A researcher might be interested in $s(\mathbf{X}_{\text{CEF}}, \mathbf{Z})$, but instead can only compute $s(\mathbf{X}_{\text{Released}}, \mathbf{Z})$. To investigate the effects of swapping on s , one could follow our recipe:

- (1) Generate a synthetic database $\mathbf{X}_{\text{Synthetic}}$ (for example, using the methods of Dwork et al. [15]).
- (2) Use our implementation to generate multiple datasets $\mathbf{X}_{\text{Swapped}}$.
- (3) Estimate the bias and variance of the random variable $\Delta \triangleq s(\mathbf{X}_{\text{Swapped}}, \mathbf{Z}) - s(\mathbf{X}_{\text{Synthetic}}, \mathbf{Z})$ over the randomness of our swapping algorithm. This will require multiple independent runs of swapping.

Under the assumption that our implementation of swapping behaves similarly to the Census Bureau’s true swapping procedure, $s(\mathbf{X}_{\text{Released}}, \mathbf{Z}) - s(\mathbf{X}_{\text{CEF}}, \mathbf{Z})$ behaves like a draw from the distribution of Δ . If Δ has high variance, we cannot reliably measure our statistic of interest s using a single Census release. If, on the other hand, Δ has nonzero mean and low variance, we might “de-bias” s as $s(\mathbf{X}_{\text{CEF}}, \mathbf{Z}) \approx s(\mathbf{X}_{\text{Released}}, \mathbf{Z}) - \mathbb{E}[\Delta]$. We suggest reporting Δ across a range of parameter values (in particular, the swap rate). Using the same method with ToyDown instead of swapping, one can study the effect of a TopDown-like process. In this way, researchers can characterize the effects of both swapping-based and DP-based disclosure avoidance systems.

6 Discussion and Limitations

Reliable Census statistics are essential to a range of research and public policy applications. With the introduction of differential privacy into published US Census data, stakeholders are understandably concerned about noisier estimates and potentially worse decisions. However, prior Census releases have been subject to

⁸Note that, to match Mueller and Santos-Lozada [29], we actually compare TopDown to the released 2010 data, which is swapped (this corresponds to the arrow marked “Panel D” in Figure 1). As mentioned before, this is a comparison between two different processes applied to the same underlying data, not a direct comparison between the underlying data and the processed data.

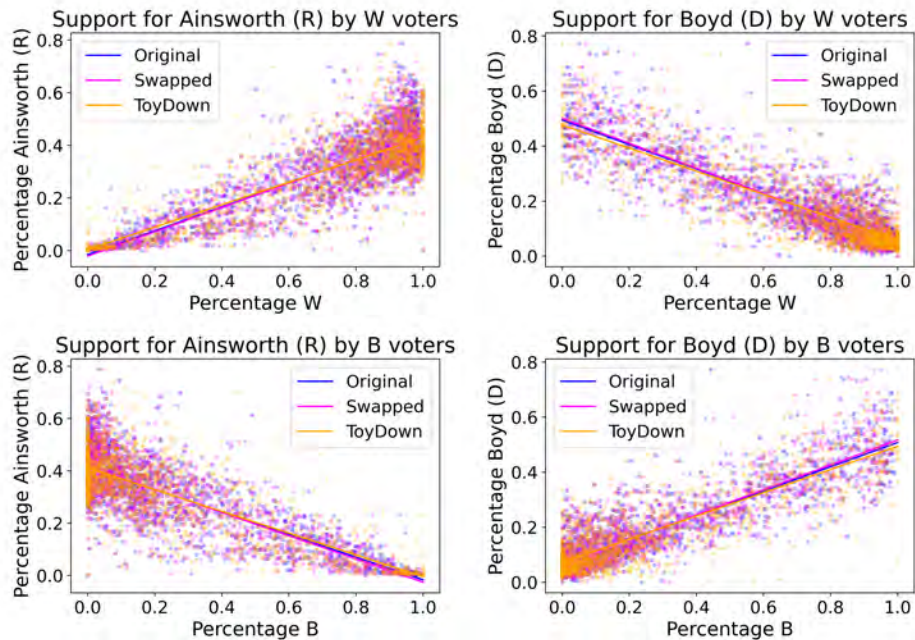


Figure 6: Ecological regression run on unswapped (in blue), swapped (in magenta), and ToyDown protected (in orange) data. The outputs of ER (the estimates of supports of racial group r for candidate c) are the intersections of the regression lines with the line $x = 1$. In this figure, ER is not weighted by population.

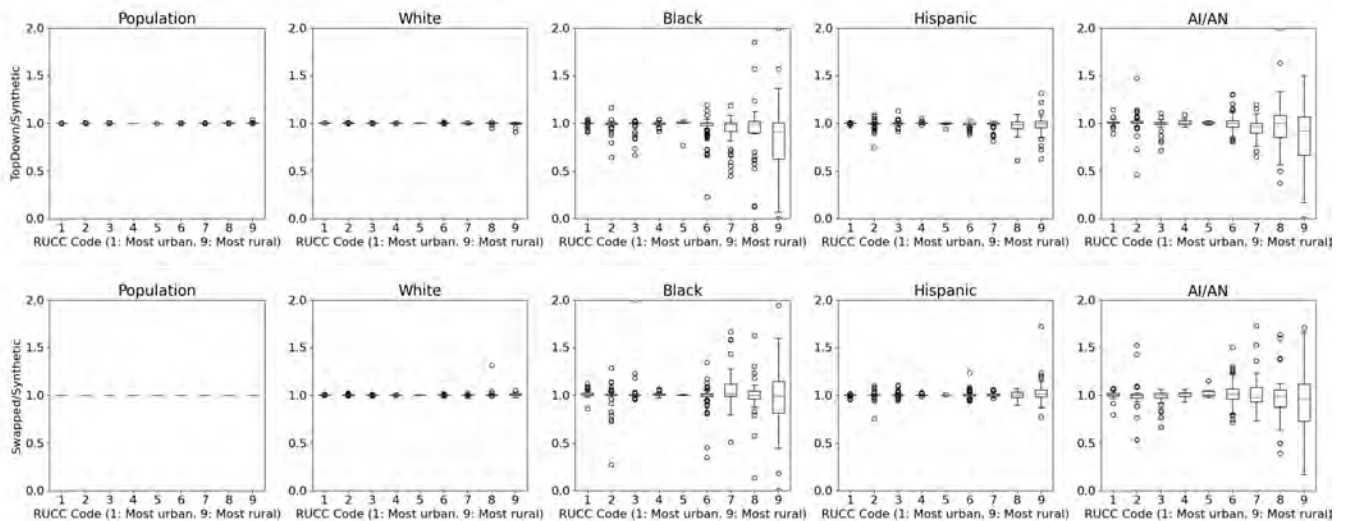


Figure 7: In the top panel, we create a similar figure to [29, Figure 2]. The vertical axis is the TopDown count divided by the count in the released summary data for 2010. The horizontal axis groups each county by its RUCC number, which represents how urban or rural it is (1 is very urban, 9 is very rural). Infinite or 0/0 values are left off of the plot. In the bottom panel, we create the analogous figure for swapping. Note that all vertical axes are scaled from 0 to 2 for readability (some outliers are cut out of the graph).

swapping, a different (and more opaque) form of disclosure avoidance that has been hard to quantify thus far. Our work seeks to close this gap by characterizing the effects of swapping and providing others with the tools to analyze its effects in their own contexts.

We find that swapping can have similarly significant impacts on downstream analyses as differential privacy. In a sense, this is to be expected: if swapping seeks to provide comparable privacy protections to differential privacy, it must inject randomness on a similar scale. More surprisingly, we find that swapping cannot simply be treated as a like-for-like replacement for differential privacy in terms of the *types* of errors it introduces. Swapping can bias estimates in opposite directions as differential privacy, and remedies to make analyses robust to differential privacy may not work for swapping. We hope that the methods and tools we provide will give researchers a principled way to reason about and potentially mitigate errors introduced by swapping.

Our work carries a number of significant limitations. While we give our best effort to provide an implementation of swapping that is consistent with the Bureau's, there will undoubtedly be differences. Our robustness checks seek to address this, but do not rule out the possibility that some of the effects we describe are artifacts of our implementation. Similarly, the synthetic data we generate will in general differ from the true microdata, and our conclusions are only valid insofar as the effects of swapping on this synthetic data are comparable to its effects on the true microdata. The magnitudes of the effects we observe are in general hard to quantify without more knowledge about the actual parameter values (such as the swap rate) used by the Bureau. And finally, swapping is not the only disclosure avoidance technique used in the 1990–2010 censuses. Some entries were censored, rounded, top-coded, or otherwise altered to meet confidentiality requirements; we are unable to evaluate these in this work.

Our work is intended to be a starting point for future research on the impacts of disclosure avoidance systems on Census data and downstream analyses. When analyzing Census data, researchers can and should use simulations like ours to determine how privacy preserving techniques are likely to affect their estimates. We seek to do this for swapping; future work could do the same TopDown. Moreover, future work could extend our methods to other Census data products like the American Community Survey. Our hope is that this line of work ultimately enables better research and decision-making based on data generated via fully- or partially-known random processes.

References

- [1] John Abowd, Robert Ashmead, Garfinkel Simson, Daniel Kifer, Philip Leclerc, Ashwin Machanavajjhala, and William Sexton. 2019. Census topdown: Differentially private data, incremental schemas, and consistency with public knowledge. *US Census Bureau* (2019).
- [2] John M. Abowd. 2018. The US Census Bureau adopts differential privacy. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*. 2867–2867.
- [3] John M. Abowd. 2021. Declaration of John Abowd. State of Alabama et al. v. United States Department of Commerce et al.
- [4] John M. Abowd, Tamara Adams, Robert Ashmead, David Darais, Sourya Dey, Simson L. Garfinkel, Nathan Goldschlag, Daniel Kifer, Philip Leclerc, Ethan Lew, Scott Moore, Rolando A. Rodriguez, Ramy N. Tadros, and Lars Vilhuber. 2023. The 2010 Census Confidentiality Protections Failed, Here's How and Why. <https://www2.census.gov/library/working-papers/2023/adrm/ces/CES-WP-23-63.pdf>. (2023).
- [5] John M. Abowd, Ian M. Schmutte, and Lars Vilhuber. 2018. *Disclosure Limitation and Confidentiality Protection in Linked Data*. Technical Report. U.S. Census Bureau Center for Economic Studies.
- [6] James Bailie, Ruobin Gong, and Xiao-Li Meng. 2023. Can Swapping be Differentially Private? A Refreshment Stirred, not Shaken. *NBER working paper* (2023).
- [7] danah boyd and Jayshree Sarathy. 2022. Differential perspectives: Epistemic disconnects surrounding the US Census Bureau's use of differential privacy. *Harvard Data Science Review* (2022).
- [8] U.S. Census Bureau. 2021. 2010 Demonstration Privacy-Protected Microdata Files: Production Settings 2021-08-12. https://www2.census.gov/programs-surveys/decennial/2020/program-management/data-product-planning/2010-demonstration-data-products/01-Redistricting_File--PL_94-171/2021-06-08_ppmf_Production_Settings/.
- [9] U.S. Census Bureau. 2022. Public Use Microdata Sample (PUMS) Accuracy of the Data (2022). https://www2.census.gov/programs-surveys/acs/tech_docs/pums/accuracy/2022AccuracyPUMS.pdf.
- [10] U.S. Census Bureau. 2023. 2010 Demonstration Data Products Suite: Redistricting and DHC. https://www2.census.gov/programs-surveys/decennial/2020/program-management/data-product-planning/2010-demonstration-data-products/04-Demonstration_Data_Products_Suite/2023-04-03/.
- [11] Miranda Christ, Sarah Radway, and Steven M Bellovin. 2022. Differential privacy and swapping: Examining de-identification's impact on minority representation and privacy preservation in the US Census. In *2022 IEEE Symposium on Security and Privacy (SP)*. IEEE, 457–472.
- [12] Aloni Cohen, Moon Duchin, JN Matthews, and Bhushan Suwal. 2021. Census TopDown: The Impacts of Differential Privacy on Redistricting. In *2nd Symposium on Foundations of Responsible Computing (FORC 2021)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik.
- [13] Tore Dalenius and Steven P Reiss. 1982. Data-swapping: A technique for disclosure control. *Journal of statistical planning and inference* 6, 1 (1982), 73–85.
- [14] Irit Dinur and Kobbi Nissim. 2003. Revealing information while preserving privacy. In *Proceedings of the Twenty-Second ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems* (San Diego, California) (PODS '03). Association for Computing Machinery, New York, NY, USA, 202–210. doi:10.1145/773153.773173
- [15] Cynthia Dwork, Kristjan Greenewald, and Manish Raghavan. 2024. Synthetic Census Data Generation via Multidimensional Multiset Sum. In *Foundations of Responsible Computing*.
- [16] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. 2006. Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography: Third Theory of Cryptography Conference, TCC 2006, New York, NY, USA, March 4-7, 2006. Proceedings 3*. Springer, 265–284.
- [17] Stephen E Fienberg and Julie McIntyre. 2004. Data swapping: Variations on a theme by dalenius and reiss. In *Privacy in Statistical Databases: CASC Project Final Conference, PSD 2004, Barcelona, Spain, June 9-11, 2004. Proceedings*. Springer, 14–29.
- [18] David A. Freedman, Stephen P. Klein, Jerome Sacks, Charles A. Smyth, and Charles G. Everett. 1991. Ecological Regression and Voting Rights. *Evaluation review* 15, 6 (1991), 673–711.
- [19] Richard A Griffin, Linda Flores-Baez, and Alfredo Navarro. 1989. *Disclosure avoidance for the 1990 Census*. US Census Bureau [custodian].
- [20] Mathew E Hauer and Alexis R Santos-Lozada. 2021. Differential privacy in the 2020 census will distort COVID-19 rates. *Socius* 7 (2021), 2378023121994014.
- [21] Sam Hawala. 2003. Microdata disclosure protection research and experiences at the US Census Bureau. In *Workshop on Microdata: 21-22 August 2003, Stockholm*.
- [22] Michael Hawes and Rolando Rodríguez. 2021. Determining the Privacy-loss Budget. <https://www2.census.gov/about/training-workshops/2021/2021-06-04-das-presentation.pdf>.
- [23] Christopher T Kenny, Shiro Kuriwaki, Cory McCartan, Evan Rosenman, Tyler Simko, and Kosuke Imai. 2021. The use of differential privacy for Census data and its impact on redistricting: The case of the 2020 U.S. Census. *Science Advances* 7 (2021). Issue 41.
- [24] Christopher T Kenny, Shiro Kuriwaki, Cory McCartan, Evan TR Rosenman, Tyler Simko, and Kosuke Imai. 2023. Comment: The Essential Role of Policy Evaluation for the 2020 Census Disclosure Avoidance System. *Harvard Data Science Review* (2023).
- [25] Christopher T Kenny, Shiro Kuriwaki, Cory McCartan, Tyler Simko, and Kosuke Imai. 2023. Evaluating Bias and Noise Induced by the US Census Bureau's Privacy Protection Methods. *arXiv preprint arXiv:2306.07521* (2023).
- [26] Nicolas Kim. 2015. The effect of data swapping on analyses of American Community Survey data. *Journal of Privacy and Confidentiality* 7, 1 (2015).
- [27] Gordon Long. 2022. *Consistency of Data Products and Formal Privacy Methods for the 2020 Census*. Technical Report. JASON, the Mitre corporation.
- [28] Laura McKenna. 2019. *Disclosure Avoidance Techniques Used for the 1960 Through 2010 Decennial Censuses of Population and Housing Public Use Microdata Samples*. Technical Report. U.S. Census Bureau.

- [29] J. Tom Mueller and Alexis R. Santos-Lozada. 2022. The 2020 US Census Differential Privacy Method Introduces Disproportionate Discrepancies for Rural and Non-White Populations. *Population research and policy review* 41, 4 (2022), 1417–1430.
- [30] E. Poehler. 2022. American Community Survey and Puerto Rico Community Survey Design and Methodology. https://www2.census.gov/programs-surveys/acs/methodology/design_and_methodology/2022/acs_design_methodology_report_2022.pdf.
- [31] Asoka Ramanayake and Laura Zayatz. 2010. Balancing disclosure risk with data quality. *Statistics* 4 (2010).
- [32] Rolando A. Rodríguez. 2023. Confidentiality for American Community Survey Respondents. <https://acsdatacommunity.prb.org/discussion-forum/m/2023-acs-conference-files/204/download>.
- [33] Steven Ruggles, Catherine Fitch, Diana Magnuson, and Jonathan Schroeder. 2019. Differential privacy and census data: Implications for social and economic research. In *AEA papers and proceedings*, Vol. 109. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, 403–408.
- [34] Alexis R Santos-Lozada, Jeffrey T Howard, and Ashton M Verdery. 2020. How differential privacy will affect our understanding of health disparities in the United States. *Proceedings of the National Academy of Sciences* 117, 24 (2020), 13405–13412.
- [35] Ryan Steed, Terrance Liu, Zhiwei Steven Wu, and Alessandro Acquisti. 2022. Policy impacts of statistical uncertainty and privacy. *Science* 377, 6609 (2022), 928–931.
- [36] Phil Steel and Laura Zayatz. 2003. *The Effects of the Disclosure Limitation Procedure on Census 2000 Tabular Data Products (Abridged)*. Technical Report. U.S. Census Bureau Statistical Research Division.
- [37] Richelle L Winkler, Jaclyn L Butler, Katherine J Curtis, and David Egan-Robertson. 2021. Differential privacy and the accuracy of county-level net migration estimates. *Population Research and Policy Review* (2021), 1–19.
- [38] Tommy Wright and Kyle Irimata. 2021. *Empirical study of two aspects of the TopDown Algorithm output for redistricting: reliability & variability (August 5, 2021 Update)*. Technical Report. U.S. Census Bureau Center for Statistical Research & Methodology.
- [39] Laura Zayatz. 2002. *SDC in the 2000 U.S. Decennial Census*. Springer Berlin Heidelberg, Berlin, Heidelberg, 193–202. doi:10.1007/3-540-47804-3_16
- [40] Laura Zayatz. 2007. Disclosure Avoidance Practices and Research at the U.S. Census Bureau: An Update. *Journal of Official Statistics* 23, 2 (2007).
- [41] Laura Zayatz, Jason Lucero, Paul Massell, and Asoka Ramanayake. 2009. *Disclosure Avoidance for Census 2010 and American Community Survey Five-year Tabular Data Products*. Technical Report. U.S. Census Bureau Statistical Research Division.