The Thirty-Ninth AAAI Conference on Artificial Intelligence (AAAI-25)

Fairness Issues and Mitigations in (Differentially Private) Socio-Demographic Data Processes

Joonhyuk Ko[1], Juba Ziani[2], Saswat Das[1], Matt Williams[3], Ferdinando Fioretto[1]

[1]University of Virginia
[2]Georgia Institute of Technology
[3]RTI International
tah3af@virginia.edu, jziani3@gatech.edu, duh6ae@virginia.edu,
mrwilliams@rti.org, fioretto@virginia.edu

Abstract

Statistical agencies rely on sampling techniques to collect socio-demographic data crucial for policy-making and resource allocation. This paper shows that surveys of important societal relevance introduce sampling errors that unevenly impact group-level estimates, thereby compromising fairness in downstream decisions. To address these issues, this paper introduces an optimization approach modeled on real-world survey design processes, ensuring sampling costs are optimized while maintaining error margins within prescribed tolerances. Additionally, privacy-preserving methods used to determine sampling rates can further impact these fairness issues. This paper explores the impact of differential privacy on the statistics informing the sampling process, revealing a surprising effect: not only is the expected negative effect from the addition of noise for differential privacy negligible, but also this privacy noise can in fact reduce unfairness as it positively biases smaller counts. These findings are validated over an extensive analysis using datasets commonly applied in census statistics.

1 Introduction

Statistical agencies across various countries gather, anonymize, and disseminate socio-demographic data, which is foundational to high-impact applications such as policy development, urban planning, and public health initiatives (USCB 2023b; FHWA 2023; USCB 2022). In the United States, for example, Census Bureau data guide more than \$2.8 trillion in federal funding yearly (U.S. Census Bureau 2023; Tran et al. 2021). Major surveys such as the American Community Survey (ACS) (USCB 2024), the Current Population Survey (CPS) (USCB 2023a), and the National Health Interview Survey (NHIS) (CDC 2024) are central to gathering essential demographic data. The ACS, for example, annually collects data from approximately 3.54 million housing unit addresses across the United States (about a 1% sample of the U.S. population). This sample-based approach allows the ACS to provide detailed insights into the population's living conditions, educational attainment, employment, and health status, among other factors.

The accuracy of these demographic reports is thus crucial to ensure that resources and policy measures are effectively targeted toward appropriate population segments. However, despite their critical role, the collection of these statistics typically involves surveying a small fraction of the population, inherently introducing sampling errors. While these surveys strive to provide estimates with controlled error rates and confidence intervals, such control is typically applied across the entire survey population. However, this approach *can lead to varying error rates among population groups*, particularly those distinguished by ethnicity, introducing biases in critical downstream tasks relying on this data.

Therefore, the *first* major contribution of this paper is to address the need for developing sampling schemes that not only aim to reduce costs but also meet acceptable errors *within* each demographic group. The approach explored recasts the sampling process as an optimization program, ensuring that statistical accuracy is maintained across diverse sub-populations within prescribed confidence errors.

The *second* major contribution of this work is to analyze the impacts of privacy-enhancing technologies on the biases of demographic data. In particular, we focus on *differential privacy* (DP) (Dwork et al. 2006) as implemented by the U.S. Census Bureau. Interestingly, contrary to prevailing intuition in the literature (Fioretto et al. 2022), our findings suggest that these privacy measures do not necessarily exacerbate disparities under Laplace mechanism. In fact, it is possible that differential privacy can surprisingly mitigate the observed unfairness by boosting the representation of smaller populations (minorities) due to positive biases introduced during the DP post-processing phase. This observation provides a novel perspective on the tradeoffs between privacy and fairness, demonstrating that DP can contribute positively to fairness when implemented in contexts similar to those analyzed in this work. A schematic illustration of the proposed framework is provided in Figure 1.

Contributions. This paper makes several contributions:

1. First, we show that conventional sampling strategies may overlook potential disparate impacts on crucial ethnic demographic groups. This aspect, illustrated in more detail in Section 3, provides the basis for the proposed work.
2. We then introduce an optimization approach aimed at mitigating these disparate errors, detailed in Section 4. The proposed approach is modeled on real-world survey design processes, such as those employed by the ACS, which

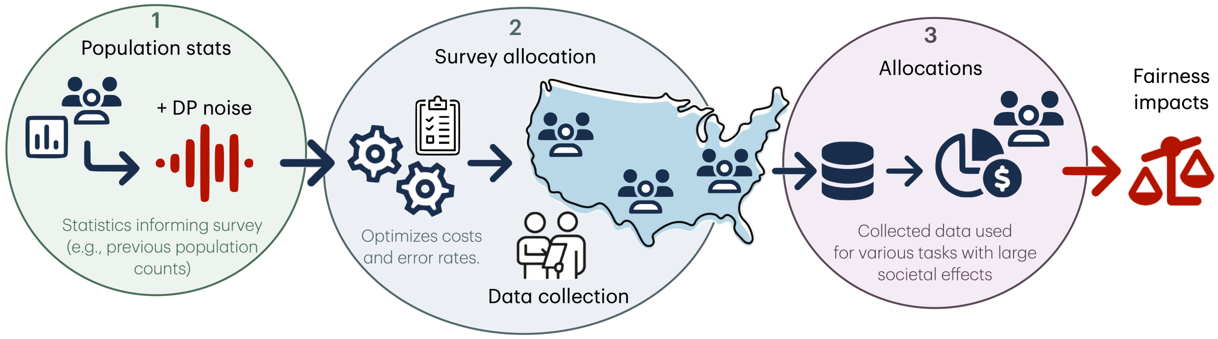


Figure 1: **1.** Population statistics from previous years are often used to inform the survey design process; Differential privacy can be used at this stage to protect sensitive information (e.g., population counts). **2.** The survey process includes selecting the amount of the population to sub-sample as well as collecting information from individuals in multiple phases (e.g., phone calls and in-person interviews). **3.** The collected data is used for important tasks, such as the allocation of funds or the release of migration patterns. *The paper studies the fairness impacts of this pipeline (steps 1 and 2) on multiple population segments.*

involve two phases: remote communications (e.g., phone calls, emails) and door-to-door, geographically targeted interventions. Our optimization framework is designed to optimize sampling costs while ensuring that error margins are within the prescribed error tolerance with a high probability for each population segment.

3. Next, Section 5 explores the impact of differential privacy on the statistics informing the sampling process. Since the noise adopted by differential privacy can influence the estimation of group sizes, we ask if it may negatively affect the reliability of the error bounds established in our model and exacerbate unfairness. *Surprisingly*, we found that on real U.S. survey data, not only is the impact of this noise negligible, but also due to an intriguing by-product of positive bias induced by DP post-processing on small counts, the resulting sampling process exhibits *reduced* unfairness across various population segments.

2 Preliminaries and Goals

This paper considers a target population, such as the U.S. population, segmented into G distinct groups characterized by race, socio-economic status, and other demographic factors. Let N represent the total population size, with N_i indicating the size of each group $i \in [G]$. We examine the population statistics $\boldsymbol{\theta}(N)$, such as average income or poverty levels, and aim to estimate these via subsampling. The subsample, of size n where $n \ll N$, is used to derive the estimates $\hat{\boldsymbol{\theta}}(n)$. In particular, this analysis extends to group-specific statistics, where $\hat{\boldsymbol{\theta}}(n_i)$ represents estimates from a subsample of size n_i (the number of individuals sampled from group i), and $\boldsymbol{\theta}(N_i)$ represents the actual statistics for group i.

To ease notation, in the discussions that follow, $\boldsymbol{\theta}$ and $\hat{\boldsymbol{\theta}}$ will be used to represent the true population statistics and their estimates from the subsample, respectively, when clear from the context. Similarly, within each group $i \in [G]$, $\boldsymbol{\theta}_i$ and $\hat{\boldsymbol{\theta}}_i$ will denote the actual statistics for the population and their corresponding estimates from the subsample.

Accuracy and fairness. The accuracy of these estimates is evaluated through their error and variance, defined for group i as $\text{Err}(\hat{\boldsymbol{\theta}}_i) = |\hat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}_i|$, and $\text{Var}(\hat{\boldsymbol{\theta}}_i) = \mathbb{E}[\hat{\boldsymbol{\theta}}_i^2] - (\mathbb{E}[\hat{\boldsymbol{\theta}}_i])^2$, respectively. The primary goal is to devise sampling strategies that minimize the *sampling cost*—defined in subsequent sections—while ensuring that the probability of an estimator's error exceeding a certain threshold (γ_i) for each group i:

$$\Pr(|\text{Err}(\hat{\boldsymbol{\theta}}_i)| > \gamma_i) \leq \alpha, \quad \forall i \in [G],$$

remains less than α.

*Un***fairness** in this context is quantified by the maximum discrepancy in estimator's variances between any two groups,

$$\xi_{\text{Var}} = \max_{i,j \in G} |\text{Var}(\hat{\boldsymbol{\theta}}_i) - \text{Var}(\hat{\boldsymbol{\theta}}_j)|,$$

since the goal of the survey process is controlling confidence intervals across various populations.

In the first part, this paper will discuss the development of optimal sampling schemes that balance the reduction of sampling costs against the constraints on estimator's accuracy for each group. Subsequently, we will examine the impact of privacy on the biases and variance of the various subpopulations, specifically when privacy-preserving counts $\tilde{N}_i$ are used instead of the actual group population counts N_i. We next discuss the notion of privacy adopted in this study.

Differential privacy. Differential Privacy (DP) (Dwork et al. 2006) is a rigorous privacy notion that characterizes the amount of information of an individual's data being disclosed in a computation. Formally, a randomized mechanism $\mathcal{M} : \mathcal{X} \to \mathcal{R}$ with domain $\mathcal{X}$ and range $\mathcal{R}$ satisfies (ϵ, δ)-*differential privacy* if for any output $O \subseteq \mathcal{R}$ and datasets $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{X}$ differing by at most one entry (written $\boldsymbol{x} \sim \boldsymbol{x}'$),

$$\Pr[\mathcal{M}(\boldsymbol{x}) \in O] \leq \exp(\epsilon) \Pr[\mathcal{M}(\boldsymbol{x}') \in O] + \delta. \quad (1)$$

Intuitively, DP states that specific outputs to a query are returned with a similar probability regardless of whether the data of any individual is included in the dataset. Parameter $\epsilon > 0$ describes the *privacy loss* of the mechanism, with values close to 0 denoting strong privacy. When $\delta = 0$, mechanism $\mathcal{M}$ is said to achieve ϵ- or pure-DP.

A function f from a dataset $\boldsymbol{x} \in \mathcal{X}$ to an output set $O \subseteq \mathbb{R}^n$ can be made differentially private by injecting random noise onto its output. The amount of noise relies on the notion of *global sensitivity* $\Delta_f = \max_{\boldsymbol{x} \sim \boldsymbol{x}'} \|f(\boldsymbol{x}) - f(\boldsymbol{x}')\|_p$, for $p \in \{1, 2\}$. In particular, the *Laplace mechanism* for histogram data release (sensitivity $\Delta_f = 1$), defined by $\mathcal{M}_{\text{Lap}}(\boldsymbol{x}) = \boldsymbol{x} + \text{Lap}(1/\epsilon)$, where $\text{Lap}(\eta)$ is the Laplace distribution centered at 0 and with scaling factor η, satisfies $(\epsilon, 0)$-DP.

Post-processing. DP satisfies several important properties. Notably, *post-processing immunity* ensures that privacy guarantees are preserved by arbitrary post-processing steps. More formally, let $\mathcal{M}$ be an (ϵ, δ)-DP mechanism and g be an arbitrary mapping from the set of possible output sequences to an arbitrary set. Then, $g \circ \mathcal{M}$ is (ϵ, δ)-differentially private.

3 Real-World Impact: The ACS Case

Next, the paper looks at the implications of sampling strategies in the American Community Survey (ACS), the largest sampling effort in the U.S. carried out by the U.S. Census Bureau. The ACS samples approximately 3.54 million housing unit addresses annually, representing about 1% of the U.S. population, and of those approximately 1.98 million result in successful samples (USCB 2023c). This relatively small sample size introduces inherent uncertainties, termed *sampling errors*, which are critical in understanding the limitations and accuracies of the data collected. The Census Bureau addresses these uncertainties by calculating standard errors and publishing margins of error at a 90 percent confidence level.

The sampling process in data collection efforts such as the ACS introduces at least two fairness issues: *disparate error rates* across different populations and *disparate impact of privacy-preserving mechanisms* on sampling errors. Figure 2 illustrates the former issue, and we will focus on the latter in Section 5. The figure shows the simulation results using 2021 data for Nebraska with a 1% sampling rate obtained from IPUMS (Ruggles et al. 2024). The lines represent the errors attained while estimating the *population income*, within 6 distinct sub-populations (x-axis). The red dotted horizontal lines illustrate the target error rates. Mimicking real-world behaviors, surveys are uniformly distributed so each group receives a proportionate number based on its size.

Notice that, while the overall population errors (rightmost bars, in dark-blue colors) are well within the prescribed confidence errors, minority groups, such as *Native* (American Indians), *Black*, and *Asian*, experience systematically larger errors compared to the *White* population. This disparity arises primarily due to smaller sample sizes among minority groups, which result in higher margins of error (Kalsbeek 2003). Additionally, when analyzing sub-populations by demographic groups, these discrepancies reveal that they often do not adhere to the 90 percent confidence levels established by the Census Bureau.

These disparities can have important repercussions given the role of these estimators in driving key policy decisions and beyond. Crucially, these behaviors are not well documented and the next section delves into our first key contribution: an optimization-based mitigation strategy.

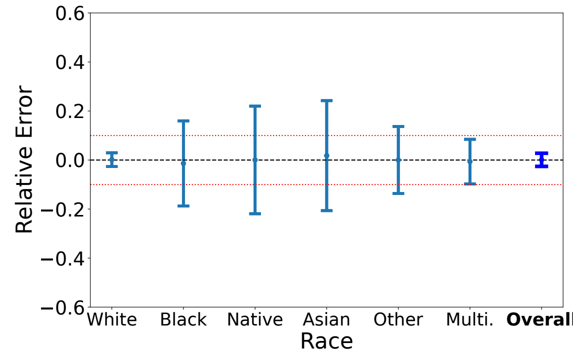


Figure 2: Disparate errors when allocating a proportional number of surveys to each racial group in Nebraska using 2022 ACS data. 2021 ACS data is used to compute the proportional allocation which subsample 1% of the total population.

4 The Optimal Sampling Design Problem

The proposed approach casts the sampling process for target estimation as an optimization process, going beyond typical cost minimization of large-scale surveys, to ensure that error rates within each sub-population are met with high probability. This section first outlines the optimization problem, considering real-world survey constraints, and then quantifies the error of estimators for each sub-population, enabling efficient implementation of the optimization model.

4.1 Modeling Real-World Sampling Processes

Large survey processes typically involve two phases. The first phase adopts various *remote* data collection modes; for example, the ACS has used internet interviews since 2013 and computer-assisted phone interviewing until 2017 (Poehler 2022). The second phase relies on in-person, *door-to-door interviews*, requiring the physical allocation of survey workers. Although more expensive, this phase aims to improve data completeness and reliability, especially in environments where remote methods are less effective.

The efficacy of each phase is distinguished by distinct *failure rates*– the likelihood that an individual, once contacted, does not contribute data. These rates are denoted as F_i^1 and F_i^2 for the first and second phases respectively and vary across different population segments i. The costs associated with each contact attempt are denoted by c_1 and c_2 for the first and second phases, respectively. Typically, the cost-efficiency trade-off is clear: remote methods (Phase 1) are cheaper but often less effective ($F_i^1 > F_i^2$), while in-person interventions (Phase 2) yield higher success rates at a higher cost ($c_1 < c_2$). We define g^r as the targeted or feasible sampling rate in region r once selected for phase 2. Further, γ_i represents the upper bound of acceptable error for population segment i (i.e., error rates depicted by the red dotted lines in Figure 2), and α as the probability that this limit is exceeded.

We define the following program to optimize this process:

$$\underset{\boldsymbol{p},\boldsymbol{z}}{\text{minimize}} \; \underbrace{c_1 \left(\sum_{i \in [G]} p_i N_i \right)}_{\text{1}^{\text{st}} \text{ phase cost}} + \underbrace{c_2 \left(\sum_{r \in R} z_r \right)}_{\text{2}^{\text{nd}} \text{ phase cost}} \tag{2a}$$

$$\texttt{s.t.} \; n_i = \underbrace{p_i N_i \left(1 - F_i^1\right)}_{\text{1}^{\text{st}} \text{ phase samples}} + \overbrace{\sum_{r \in R} z_r g^r N_i^r \left(1 - F_i^2\right)}^{\text{2}^{\text{nd}} \text{ phase samples}} \; \forall i \in [G] \tag{2b}$$

$$\Pr(|\text{Err}(\hat{\boldsymbol{\theta}}_i(n_i))| > \gamma_i) \leq \alpha, \;\; \forall i \in [G], \tag{2c}$$

$$0 \leq p_i \leq 1 \;\; \forall i \in [G], \; z_r \in \{0, 1\} \;\; \forall r \in R. \tag{2d}$$

The goal is to minimize the costs associated with contacting individuals through the two sampling phases described above, denoted by c_1 and c_2 (objective (2a)), while ensuring that the error across each demographic group i does not exceed γ_i with a probability greater than α (constraint (2c)). Decision variables p_i model the fraction of group i contacted in the remote Phase 1, and z_r, a binary variable, determines whether workers are deployed in region r (where R is the set of all regions) during Phase 2 (constraint (2d)). In the minimizer notation $\boldsymbol{p}$ and $\boldsymbol{z}$ are used as shorthand for the vectors $(p_i)_{i \in [G]}$ and $(z_r)_{r \in R}$, respectively. Constraint (2b) defines n_i, the average number of individuals that *respond* to the survey across Phases 1 and 2. In population of size N_i, the surveyor contacts $p_i N_i$ individuals in Phase 1; $p_i N_i (1 - F_i^1)$ is then the rate at which individuals respond *in expectation*. In Phase 2, in each region r, the surveyor reaches $z_r g^r N_i^r$ individuals, making $z_r N_i^r g^r (1 - F_i^2)$ the expected numbers of responses from population i in region r.

4.2 Tractable Error Quantification

A key challenge with solving Program (2) is Constraint (2c), which involves a probability estimation. The lack of a closed-form expression for this probability hinders the direct integration of this constraint into the optimization. To address this, this section provides a tractable upper bound to be used in place of the probability in Constraint (2c).

Note that, using Chebyshev's inequality, the probability of the estimator's error exceeding γ_i is bounded above by:

$$\Pr(|\text{Err}(\hat{\boldsymbol{\theta}}_i)| > \gamma_i) = \Pr(|\hat{\boldsymbol{\theta}}_i - \boldsymbol{\theta}_i| > \gamma_i) \leq \frac{\sigma^2(\hat{\boldsymbol{\theta}}_i)}{\gamma_i^2}, \tag{3}$$

where $\sigma^2(\hat{\boldsymbol{\theta}}_i)$ represents the variance of the estimator $\hat{\boldsymbol{\theta}}_i$. This variance can then be estimated empirically, as done in practice (Poehler 2022), using prior data releases. This creates a statistical proxy, which is discussed in the Section 4.3.

For a given confidence level α, from (3), we can replace Constraint (2c) by the *stronger* constraint $\frac{\sigma^2(\hat{\boldsymbol{\theta}}_i)}{\gamma_i^2} \leq \alpha$, and obtain a closed-form approximation for the threshold γ_i as:

$$\sigma^2(\hat{\boldsymbol{\theta}}_i) \leq \alpha \gamma_i^2. \tag{4}$$

This new constraint strengthens the program by enforcing $\Pr(|\text{Err}(\hat{\boldsymbol{\theta}}_i)| > \gamma_i) \leq \frac{\sigma^2(\hat{\boldsymbol{\theta}}_i)}{\gamma_i^2} \leq \alpha$, which restricts the likelihood that the error in group i exceeds the desired γ_i threshold, thus tightening the optimization.

The variance of the estimator $\sigma^2(\hat{\boldsymbol{\theta}}_i)$ can thus be expressed as

$$\sigma^2(\hat{\boldsymbol{\theta}}_i) = \frac{C_i}{n_i}, \tag{5}$$

where C_i is a constant that depends on the variance for group i. This follows from the variance of the estimator $\sigma^2(\hat{\boldsymbol{\theta}}_i)$ being inversely proportional to the sample size n_i.[1] Thus, by substituting this expression into Equation (4), constraint (2c) can be replaced by the following tractable form:

$$n_i \geq \frac{C_i}{\alpha \gamma_i^2}, \quad \forall i \in [G]. \tag{2$\bar{c}$}$$

4.3 Empirical Variance Estimation

In the above expression, C_i is a constant that depends on the variance of the population for group i (see Equation (5)) and can be estimated by approximating the variance of the target estimates across a range of sampling rates. Figure 3 (left) illustrates such an approach, showing how the variance of each subgroup changes across sampling rates within practical (e.g., budget) constraints that limit sampling to no more than 10% of the population. The middle and right figures also report this effect when privacy is considered to protect the sub-population counts N_i, as discussed in Section 5.

Approximating the variance $\sigma^2(\boldsymbol{\theta}_i)$ relies on fitting a curve of the form $\frac{a_i}{x}$ for each population group i, where a_i is the constant to be estimated and $x = \frac{n_i}{N_i} \in [0, 1]$ represents the sampling rate. These curves, referred to as *proxy functions*, estimate the variance based on sampling rates rather than absolute sample sizes, enabling a direct integration in our error quantification constraints. Finally, to translate these proxy functions into a usable format within the constraint (2$\bar{c}$), we equate C_i with $a_i \, N_i$ as follows:

$$\sigma^2(\hat{\boldsymbol{\theta}}_i) = \frac{C_i}{n_i} = \frac{a_i}{x} = \frac{a_i}{n_i / N_i} = \frac{a_i N_i}{n_i}.$$

5 Private Sampling Scheme

The sampling design discussed above assumes accurate knowledge of population sizes; however, the confidentiality of collected micro-data is often legally mandated. For example, it is regulated by Title 13 (U.S. Congress 1954) in the U.S., and, to comply with it, the U.S. Census Bureau used differential privacy for their 2020 decennial census release (Abowd 2018). However, the use of DP mechanisms introduces perturbations in the data that may disproportionately affect smaller populations (Tran et al. 2021; Zhu, Fioretto, and Hentenryck 2022). This section studies how privacy-protected statistics could influence fairness in data collection.

More precisely, we consider population sizes N_i^r released differentially-privately for each group $i \in [G]$ and region $r \in R$, emulating the census data release. Therefore, instead of having access to the exact N_i^r, the survey designer only has access to imperfect, noisy estimates given by:

$$\tilde{N}_i^r = \max\left(0, N_i^r + Lap(\Delta x / \varepsilon)\right), \tag{6}$$

[1] For n iid random variables $x_i \sim \mathcal{N}(\mu, \sigma^2)$ and estimator $\hat{x} = \frac{\sum_i^n x_i}{n}$, $\text{Var}(\hat{x}) = \text{Var}(\frac{1}{n} \sum_i^n x_i) = \frac{1}{n^2} \text{Var}(\sum_i^n x_i) = \frac{1}{n^2} n \sigma^2 = \frac{\sigma^2}{n}$.

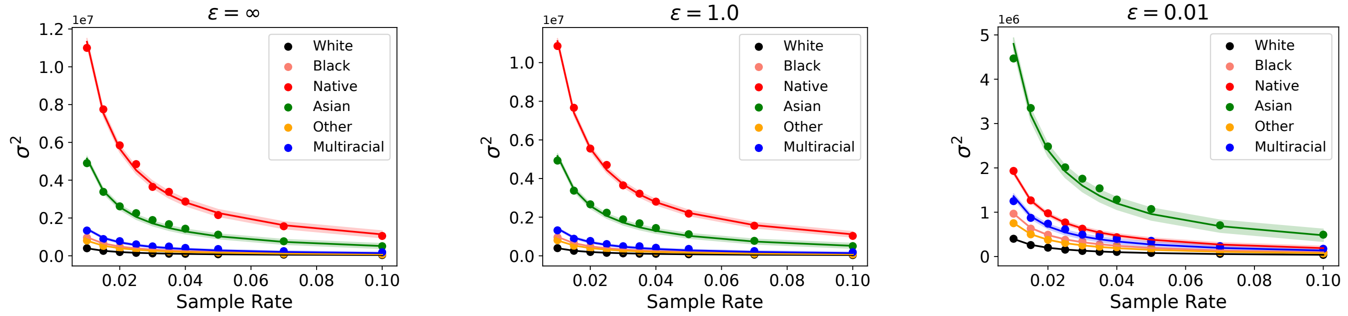


Figure 3: Estimating the variance of mean *income* in Connecticut using *race* as a subgroup with different privacy budget ε. **Points**: actual estimator measurement, **curves**: proxy function fitting. Results averaged over 200 trials and 200 data points.

where $\Delta x = 1$ (sensitivity of the count query). We further note that the noisy counts are post-processed to ensure non-negativity (as is done in the U.S. Census (Spence 2023)), here using the $\max(0, .)$ operator.

The challenge in this context is that noise can distort estimates of population sizes N_i^r, influencing the number of individuals n_i who respond to the survey in each group i. This distortion affects errors in Constraint (2b) and compromises achieving the desired error targets and confidence levels. This section outlines our second key contribution: we offer theoretical insights into the biases introduced by using $\tilde{N}_i^r$ instead of N_i^r, while Section 6 will offer a practical analysis of these impacts. Our main result is a *closed-form* expression for the bias of the estimate $\tilde{N}_i^r$, showing that this bias is invariably positive

Theorem 1. *For all $i \in [G]$, $r \in R$, the bias of estimate $\tilde{N}_i^r$ is given in closed-form by:*

$$\mathcal{B}(\tilde{N}_i^r) = \mathbb{E}\left[\tilde{N}_i^r\right] - N_i^r = \frac{\Delta x}{2\varepsilon} \exp\left(-\frac{N_i^r \varepsilon}{\Delta x}\right) > 0.$$

Observe that not only is the bias term always positive (ε is always positive), but also for fixed privacy budget ε, this bias is *higher* on groups with small N_i^r. This implies that minority populations, such as Native Americans, are more likely to be overestimated. Further, if all other parameters are fixed, when ε increases, the bias *decreases*; in extreme cases, when $\varepsilon \to \infty$ (no privacy), the bias converges to 0, and when $\varepsilon \to 0$ (perfect privacy), the bias grows large. This implies an interesting effect: the bias-induced overestimation of minority populations and its beneficial effects in correcting for under-allocations increase as ε decreases. Surprisingly, and contrary to much of previous known effects (Fioretto et al. 2022), in the context studied here, enforcing stronger privacy induces *less unfairness towards minority populations!*

This result implies the following corollary deriving the bias of the aggregated (e.g., state level) counts on $\tilde{N}_i = \sum_{r \in R} \tilde{N}_i^r$ given the various $\tilde{N}_i^r$ (e.g., county level), highlighting once again a more pronounced effect on minority populations:

Corollary 1. *The bias of the aggregated counts for each subgroup on the* state *level is*

$$\mathcal{B}(\tilde{N}_i) = \mathbb{E}\left[\tilde{N}_i\right] - N_i = \sum_{r \in [R]} \frac{\Delta x}{2\varepsilon} \exp\left(\frac{-N_i^r \varepsilon}{\Delta x}\right) > 0.$$

This further highlights that while this positive bias will have a major relative impact on minority populations, overestimating minority populations allows a standard sampling scheme to allocate more surveys to minorities, thus reducing their relative errors. An empirical analysis of this phenomenon is provided in Section 6.2.

It is important to note that the U.S. Census Bureau implemented the TopDown Algorithm for the 2020 Census, a differential privacy mechanism designed to ensure hierarchical consistency across varying geographical units. Rather than introducing a positive bias across all populations, the algorithm introduces a positive bias towards minority groups and a negative bias towards majority groups to maintain consistency in the hierarchical statistics (Michael Hawes 2020). Here, we use a differentially private mechanism without hierarchical consistency for simplicity and clearer analysis.

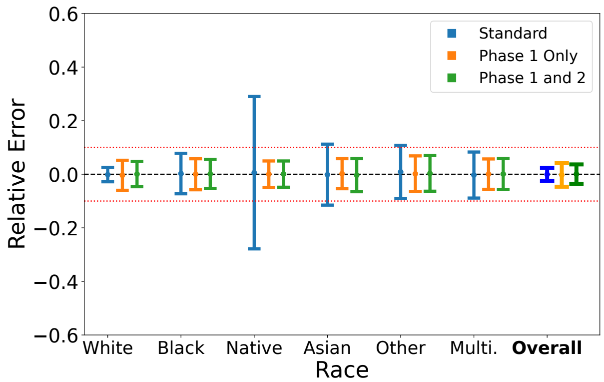
6 Experimental Results

Next, the paper provides empirical evidence for the efficacy of the proposed optimization method on real-world data and settings first without and then with privacy considerations at hand. The experiments examine survey costs, group fairness, and utility offered by the proposed fairness-aware method.

Datasets and settings. The experiments use ACS data from IPUMS (Ruggles et al. 2024) for 2021 and 2022, leveraging 2021 data for estimating the various N_i^r and 2022 data as ground truth for sampling and assessing target estimators. We divide geographical units based on Census Tract-level data, each containing about 4,000 individuals. The focus is on estimating annual total pre-tax personal income across different ethnic and educational groups as defined by IPUMS and the Census Bureau. We focus on Connecticut as the primary state for analysis here.

Algorithms. This analysis evaluates various survey allocation mechanisms, comparing their efficiency, fairness, and effectiveness in achieving desired confidence levels, not only at the entire state level but also at the sub-population levels:

- **Standard Allocation:** This baseline method, also known as *proportional stratified random sampling*, allocates surveys to each population group i in proportion to their size. This approach is a *stronger* baseline than simple random sampling for two key reasons: it provides more precise



Method	Sampling Rate	Survey Cost	ξ_{Var}	# of Violations
Standard	0.964%	100%	0.260	3/6
Phase 1 Only	0.964%	100%	**0.023**	**0/6**
Phase 1 and 2	**0.770%**	**86%**	0.030	**0/6**

Figure 4: Relative group errors from estimating mean income in Connecticut.

population estimates by reducing variance within each subgroup[2], and it ensures that the sample proportions are representative of the overall population (Lohr 1998).

- **Optimization: Phase 1 Only:** This variant applies the optimization from Program (2), assuming the survey is conducted by only using the first phase. More concretely, the program excludes the *2nd phase* components in (2a) and (2b). This model mirrors the proportional stratified random sampling by optimizing survey numbers within a single operational phase.
- **Optimization: Phase 1 and 2:** This approach uses both phases as outlined in optimization (2), aligning closely with practical survey methodologies. Note that the choice of failure rates and costs influences the optimization outcomes. In particular, high failure rates (F_i^1 and F_i^2) or low error tolerances (α and γ) increase the total survey cost due to more failures and tighter constraints. The default confidence constraints are set at $\alpha = 0.1$ and $\gamma_i = 10\%$ of the mean income for each subgroup i. Default failure rates are $F_i^1 = 0.60$ and $F_i^2 = 0.20 \, \forall i \in [G]$, and the cost of surveying a region in phase 2 is set to be 500 times more expensive than the cost of reaching out to an individual with phone calls in phase 1. Finally, the sampling rates for geographies are set as $g^r = 0.1 \, \forall \, r \in R$. This translates to sampling 400 people per selected region.

Evaluation metrics. The evaluation of these mechanisms focuses on three primary metrics:

1. *Survey cost*: Measured as a percentage of the cost reference used by the standard allocation.
2. *Fairness of variance*: Assesses the equitable distribution of survey errors across different groups.
3. *Confidence compliance*: Evaluates the ability to meet the prescribed confidence errors (γ_i) at a 10% threshold, which aligns with the current standards of the ACS (Poehler 2022), and setting $\alpha = 0.1$.

[2]The number of surveys per subgroup has no variance.

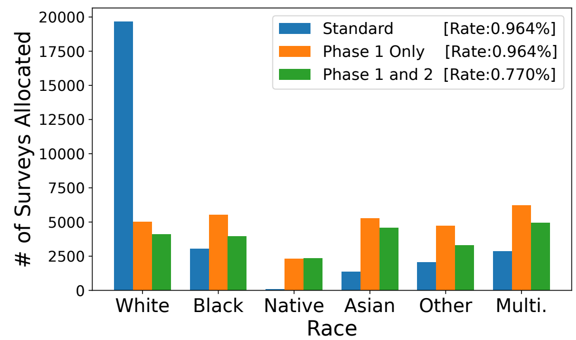


Figure 5: Number of surveys allocated for each subgroup in the experiments reported in Figure 4.

Further exploration of the impact of privacy on these metrics is detailed in Section 6.2.

6.1 Optimized Sampling: Errors and Fairness

We start by assessing the performance of two variants of our method ("*Phase 1 Only*" vs. "*Phase 1 and 2*") against the standard allocation mechanism, without DP considerations.

The results, summarized in Figure 4, show that the *Standard Allocation* method yields the lowest variance of error when estimating the overall population's income. However, this method disproportionately affects minorities, who receive fewer surveys and experience a higher variance of error at the group level. This discrepancy results in the worst fairness of variance (ξ_{Var}) observed (refer to table under Figure 4), and minority groups even fail to meet the confidence constraints set for their estimations!

In contrast, the *Phase 1 Only* optimization approach achieves a more uniform error variance across all subgroups while using *the same* budget used in the *Standard Allocation* method. Inspecting the optimization solutions, it can be observed that equity is achieved by allocating a similar number of surveys to each subgroup, irrespective of their population size. Figure 5 reports the number of survey allocations by race and by each method, and provides a clear view of the nature of the disparities. This redistribution significantly lowers the error variance for minorities (including Native, Black, Asian, Other, and Multi.), while slightly increasing it for the majority (White). Importantly, this approach *enhances fairness and ensures all groups meet the confidence thresholds*, addressing the main drawback of the *Standard Allocation*.

Next, we focus on our main approach. As discussed in Section 4, phase 2 is characterized by a higher success rate ($F_i^1 > F_i^2$) at a greater cost ($c_1 < c_2$). Despite its higher per-survey cost, phase 2's low failure rate results in a higher number of *successful* samples for the same overall cost, making the *Phase 1 and 2* method *substantially cheaper* (86% of Phase 1 Only cost) (see table under Figure 4). However, once regions are selected for phase 2, simple random sampling is executed at a 10% rate (g^r) from each chosen region. This method introduces some uncertainty in the number of successful samples for each subgroup, although the optimizer prioritizes regions with high densities of the targeted population.

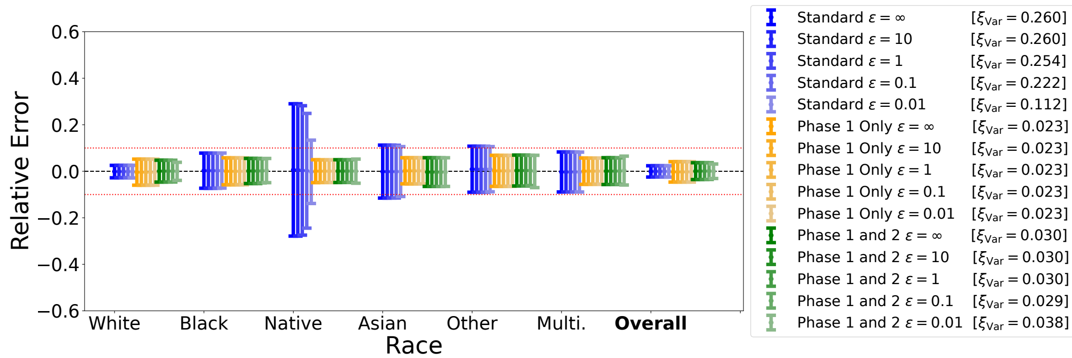


Figure 6: Relative errors from estimating mean income using DP-noised N_i^r in Connecticut. Each region used in the phase 2 contains approximately 4,000 people, similar to the size of Census Tracts.

ε \ Race	White	Black	**Native**	Asian	Other	Multi.	Total
∞	2,039,731	315,568	**7,571**	143,584	215,150	295,844	3,017,448
10	2,039,355	315,182	**7,524**	143,226	214,766	295,482	3,015,535
1	2,039,298	315,199	**7,699**	143,260	214,736	295,495	3,015,687
0.1	2,038,844	315,320	**10,681**	143,846	214,555	295,705	3,018,951
0.01	2,034,218	321,513	**42,068**	158,498	222,160	304,533	3,082,990

Table 1: Impact of DP on estimated population size for each race in Connecticut using prior (e.g., ACS 2021 dataset).

This slightly reduces the performance and fairness of variance compared to the *Phase 1 Only* optimization. Nonetheless, the *Phase 1 and 2* method meets (by construction) the confidence constraints for every group, as empirically demonstrated.

6.2 DP-Sampling: Errors and Fairness

Next, we focus on the setting *with differential privacy*, employing the privately adjusted counts $\tilde{N}_i^r$ as described in Equation (6). The results are reported in Figure 6, again for the state of Connecticut.

The first surprising result comes when analyzing the *Standard Allocation* approach. While one might expect the added noise to exacerbate errors for minorities, here *adding more noise reduces the variance of errors for minorities*! This interesting behavior occurs because the induced strong positive bias overestimates the minority population size resulting in a higher allocation of surveys to these groups, as observed by our theoretical analysis in Section 5. Table 1 summarizes this effect, where it is possible to observe how much the smallest group (*Native*) size is conflated with the addition of noise (smaller ε). This increased allocation not only reduces the error variance but also improves fairness, countering the typical expectation that more noise increases error.

On the other hand, the *Phase 1 Only* optimization appears to be insensitive in the variance of errors with respect to ε. This stability arises due to C_i, which determines the required number of samples, does not depend on group size. Thus, noise added to the population count does not impact survey distribution, maintaining consistent error variance across varying noise levels.

In contrast, the *Phase 1 and 2* method experiences slight changes in the variance of errors with added noise. This is due to how noise affects the selection of regions for Phase 2, which relies on the population composition from prior data. More noise increases the probability of incorrect region selection, altering survey distribution and consequently, error variance.

The observed higher positive bias in minorities as ε decreases is explained by Corollary 1: it notes that a smaller N_i^r leads to larger positive biases. This implies that the region size used in phase 2 directly influences the level of positive bias.

Finally, note that the larger positive bias observed as $\varepsilon \to 0$ in Table 1 led to less discrepancy between the population sizes of different subgroups by overestimating the minorities. This results in a more uniform allocation of surveys across subgroups, thereby improving fairness. This suggests that allocating an equal number of surveys to each group, irrespective of population size, may result in roughly the same relative errors.

7 Conclusion

This work was motivated by the observations of unfairness in large survey efforts of critical importance for driving many policy decisions and allocations of large amounts of funds and benefits. This paper showed that in surveys like the American Community Surveys, traditional sampling methods disproportionally affect minority groups, leading to biased statistical outcomes. To address these issues, we introduced an optimization-based framework to ensure fair representation in error margins in each population segment while minimizing the total sampling costs. Additionally, this paper examined the effects of differential privacy on the accuracy and fairness of the realized surveys. Contrary to common intuitions, our findings reveal that differential privacy can reduce unfairness by introducing positive biases beneficial to underrepresented populations. These findings are validated through rigorous and comprehensive experimental analysis using real-world data, demonstrating the effectiveness of the proposed optimization-based strategies in terms of enhancing fairness without compromising data utility and costs.

We believe that these results may have significant implications for policy formulation and resource allocation with critical societal and economic impacts.

Acknowledgements

This research is partly funded by NSF grants SaTC-2133169, RI-2232054, and CAREER-2143706. The views and conclusions of this work are those of the authors only. The authors are also thankful to Christine Task for early discussion and feedback regarding this topic.

References

Abowd, J. M. 2018. The U.S. Census Bureau Adopts Differential Privacy. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD '18, 2867. New York, NY, USA: Association for Computing Machinery. ISBN 9781450355520.

CDC. 2024. National Center for Health Statistics. *CDC - National Center for Health Statistics*.

Dwork, C.; McSherry, F.; Nissim, K.; and Smith, A. 2006. Calibrating noise to sensitivity in private data analysis. *Theory Of Cryptography, Proceedings*, 3876: 265–284. 3rd Theory of Cryptography Conference ; Conference date: 04-03-2006 Through 07-03-2006.

FHWA, U. D. 2023. U.S. Department of Transportation - Federal Highway Administration. *Census Transportation Planning Products*.

Fioretto, F.; Tran, C.; Hentenryck, P. V.; and Zhu, K. 2022. Differential Privacy and Fairness in Decisions and Learning Tasks: A Survey. In *International Joint Conference on Artificial Intelligence*, 5470–5477. ijcai.org.

Kalsbeek, W. D. 2003. Sampling minority groups in health surveys. *Statistics in Medicine*, 22(9): 1527–1549.

Lohr, S. 1998. *Sampling: Design and Analysis*. Boca Raton, FL: CRC Press.

Michael Hawes. 2020. Differential Privacy and the 2020 Decennial Census. URL: https://www.baltometro.org/sites/default/files/bmc_documents/committee/presentations/cfg/CFG200226pres_Differential-Privacy-Census.pdf. Accessed: 2024-12-16.

Poehler, E. 2022. American Community Survey and Puerto Rico Community Survey Design and Methodology.

Ruggles, S.; Flood, S.; Sobek, M.; Backman, D.; Chen, A.; Cooper, G.; Richards, S.; Rodgers, R.; and Schouweiler, M. 2024. IPUMS USA: Version 15.0.

Spence, M. 2023. What to Expect: Disclosure Avoidance and the 2020 Census Demographic and Housing Characteristics File. *U.S. Census Bureau*.

Tran, C.; Fioretto, F.; Van Hentenryck, P.; and Yao, Z. 2021. Decision Making with Differential Privacy under a Fairness Lens. In *International Joint Conference on Artificial Intelligence*, 560–566. ijcai.org.

U.S. Census Bureau. 2023. Census Bureau Data Guide More Than $2.8 Trillion in Federal Funding in Fiscal Year 2021. Accessed: 2024-12-17.

U.S. Congress. 1954. Title 13 - Census. *The United States Code*.

USCB. 2022. Health. *U.S. Census Bureau*.

USCB. 2023a. Current Population Survey (CPS).

USCB. 2023b. Decennial Census P.L. 94-171 Redistricting Data.

USCB. 2023c. Sample size and data quality.

USCB. 2024. American Community Survey (ACS).

Zhu, K.; Fioretto, F.; and Hentenryck, P. V. 2022. Post-processing of Differentially Private Data: A Fairness Perspective. In *International Joint Conference on Artificial Intelligence*, 4029–4035. ijcai.org.