Differentially Private Population Quantity Estimates via Survey Weight Regularization

Jeremy Seeman, Yajuan Si, and Jerome P. Reiter

August 15, 2024

Abstract

In general, it is challenging to release differentially private versions of survey-weighted statistics with low error for acceptable privacy loss. This is because weighted statistics from complex sample survey data can be more sensitive to individual survey response and weight values than unweighted statistics, resulting in DP mechanisms that can add substantial noise to the unbiased estimate of the finite population quantity. On the other hand, simply disregarding the survey weights can result in a biased estimate that also underestimates the sampling variance. Thus, the problem of releasing an accurate survey-weighted estimate essentially involves a trade-off among bias, precision, and privacy. We leverage this trade-off to develop a DP method for estimating finite population quantities. The key step is to privately estimate a hyperparameter that determines how much to regularize or shrink survey weights as a function of privacy loss. We illustrate the differentially private finite population estimation using the Panel Study of Income Dynamics. We show that optimal strategies for releasing DP survey-weighted mean income estimates require orders-of-magnitude less DP noise than naively using the original survey weights without modification. We then discuss its implications for integrating DP into survey research.

1 Introduction

As part of efforts to protect data subjects' privacy and confidentiality, data stewards can release statistics that satisfy differential privacy (DP) [9, 10]. To date, typical applications of DP have been based on data from censuses or administrative databases. Often, however, statistics are based on surveys with complex designs, e.g., using multi-stage, unequal probability sampling. It is well known that analysts should account for the design in inferences, which is typically done by using survey weights. Survey-weighted statistics offer unbiased and consistent estimates for finite population quantities, such as population means and totals. Yet, as we describe below, survey weights introduce challenges to implementing DP methods. These challenges motivate our work: how might data stewards apply DP to release survey-weighted statistics from complex sample surveys?

Preliminary work at the intersection of DP and survey statistics has focused on synthetic data generation with weighted distributions [14], estimation under classical sampling designs like stratified sampling [16], and interpretations of survey sampling methods for their privacy amplification properties [6, 13]. Each of these approaches attempts to utilize as much information as possible about the sampling process. However, weighting schemes can cause practical problems for DP. Weighted statistics can have significantly larger sensitivities than their unweighted counterparts, hence requiring substantially more noise to provide the equivalent level of DP protections at the same level of privacy loss [8]. Of course, one could avoid the associated increase in the DP noise variance by disregarding the sampling weights in estimation. Indeed, this can be appealing when the weighted and unweighted estimates are similar, which can occur when the survey weights are uncorrelated with the particular survey variable of interest [17, 5].

However, when this is not the case, the data steward ends up adding noise to a (perhaps severely) biased estimate. Additionally, anomalously large survey weights can substantially increase the variability of survey statistics, prompting approaches to smoothing the estimates or regularizing survey weights [11, 4, 21, 22].

Introducing DP into survey inference suggests that the degree of weight regularization should depend on privacy loss budgets. Just as there is a bias-variance-privacy trade-off for mean estimation with independently identically distributed (iid) records [15], similar three-way trade-offs must be made when analyzing survey-weighted quantities. We propose methods for DP survey-weighted estimates where the optimal degree of regularization depends on the confidential data. In this setting, we must consume privacy loss to estimate this optimal degree of regularization *and* the statistics of interest. Doing so allows us to adaptively consume privacy loss budget for fine-tuning uncertainty quantification when constructing interval estimates and assessing their coverage properties.

1.1 Contributions

We summarize our contributions here for the full paper accompanying this chapter:

1. We analyze the three-way relationship between privacy loss, accuracy, and bias emerging from survey data. To do this, we introduce a regularization parameter $\lambda \in [0, 1]$ that linearly shrinks the survey weights to a constant when $\lambda = 1$. For any survey sample, there exists an "optimal" value λ^* which minimizes DP mean-squared error (for a fixed privacy loss) that depends on the sample size, response range, possible weighting designs, and the difference between the unweighted and weighted mean estimates. We prove that $\lambda^* > 0$ (for any informative sampling design). Similarly, we prove that for any fixed privacy loss, there is a limit to the amount of bias that can be corrected by design-based weight adjustment without requiring DP noise that exceeds said correction.
2. We propose a two-step procedure to estimate survey-weighted population means using ρ-zero-concentrated Differential Privacy ρ-zCDP [7]. First, we use the exponential mechanism to estimate λ^*; then we use this output to shrink the survey weights toward uniform values and estimate the population mean using the Gaussian mechanism. We also provide different asymptotic and finite-sample approaches to quantifying errors due to sampling, weight shrinkage, and DP noise, allowing users to construct DP confidence intervals for our population mean estimates.
3. We demonstrate our methodology on survey microdata from the Panel Study of Income Dynamics (PSID) [24], a longitudinal survey containing family-level statistics on income sources and other sociodemographic information and oversampling from lower income subpopulations. We show how different response variables require different degrees of survey weight regularization, allowing us to more efficiently tailor DP privacy loss budgets when estimating multiple population means for different response variables. We also empirically validate our uncertainty quantification properties, including accuracy and coverage.

1.2 Related Literature

While there is an extensive literature on differentially private statistical analyses (see [23] for a review) and a separate literature on methods for trimming or regularizing survey weights [12], there is little literature at their intersection. Many DP algorithms rely on the "amplification by sub-sampling" property, wherein applying a DP algorithm on a simple random sample without replacement yields smaller privacy loss than the same algorithm applied to the entire population [3]. However, for survey designs besides simple random sampling, this property may not hold [6] nor would it always improve accuracy [13]. We consider a complex sample survey design where the weights themselves contain all relevant sampling information and, therefore, must be protected with DP. We do not consider the release of auxiliary data used to construct final adjustment weights, instead isolating the privacy cost of incorporating survey design exclusively within the weights.

The most direct line of work compared to ours uses methods that generate synthetic data samples containing survey responses and weights [14]. These methods can produce synthetic data that are interoperable with existing analyses and admit combining-rules-based approaches to inferences with synthetic data [20]. Our approach differs in key ways. First, we directly use survey-weighted estimators as opposed to working through multiple synthetic datasets and inferences via combining rules. Second, we provide decision-making guidelines for whether certain kinds of weighting corrections can be sufficiently estimated using DP at a given sample size.

2 Methods

In this section, we summarize our main theoretical findings here. Full derivations and results are available in the technical paper accompanying this book chapter.

2.1 Notation and Problem Definition

We consider a response variable and design-based survey weights $\{(y_i, w_i)\}_{i=1}^N$ lying within bounded intervals $[L_Y, U_Y] \times [L_W, U_W]$ from a population of N observations, where we observe the first $i \in [n]$ units and we do not observe $i \in \{n+1, n+2, \ldots, N\}$. For convenience, we define $\Delta_W \triangleq U_W - L_W$, and without loss of generality, we assume $L_Y = 0$ and $1 \leq L_W \leq U_W$. We will use $\boldsymbol{y}$ and $\boldsymbol{w}$ to correspond to the vector of n observed samples and weights, respectively.

Our goal is to estimate the population mean $\theta \triangleq \frac{1}{N}\sum_{i=1}^N y_i$ using the survey-weighted mean $\hat{\theta}(\boldsymbol{y}, \boldsymbol{w}) \triangleq \frac{1}{N}\sum_{i=1}^n y_i w_i$, assuming we only have access to the survey responses and the weights. The variability of $\hat{\theta}$ about θ depends on our sampling mechanism, which we assume is fully characterized by the survey weights. This allows us to treat the y_is and w_is as fixed constants, making our analysis consistent with DP approaches that treat confidential data entries as constants from a fixed "schema" of possible values.

We focus on the case where the weights correspond to probabilities of inclusion, i.e. $w_i^{-1} \triangleq \pi_i = \mathbb{P}(I_i = 1)$. Our goal is to release a confidence interval around $\hat{\theta}$ that contains θ while satisfying DP, specifically ρ-zero concentrated differential privacy (ρ-zCDP [7]) with respect to both the survey weights and responses. Under ρ-zCDP, for any two datasets that differ on one record's response and survey weight, the distributions of the hypothetically released statistics are close together, with distance measured by the privacy loss parameter ρ. Greater values of ρ correspond to greater disclosure risks and less noise injected into the statistics.

Note that our analysis assumes that survey weights are fixed properties of individual records that do not change depending on which units appear in the realized sample. This helps align our analysis with standard DP analyses that treat observed confidential data (in this case, survey responses and weights) as constants instead of random variables. We additionally assume that the population and sample sizes, N and n, respectively, are public information. While this assumption reflects standard practice for publishing survey metadata, there may be confidentiality concerns if membership in the population under study is privacy-concerning.

The canonical method for satisfying ρ-zCDP requires adding Gaussian noise to statistics whose variance is a function of the statistic's sensitivity, or how much the statistic could maximally change when modifying one possible entry. In this setting, the sensitivity of the weighted estimator is given by

$$\Delta(\hat{\theta}) = \sup_{(\boldsymbol{y},\boldsymbol{w}) \sim_M (\boldsymbol{y}',\boldsymbol{w}')} \left|\hat{\theta}(\boldsymbol{y}, \boldsymbol{w}) - \hat{\theta}(\boldsymbol{y}', \boldsymbol{w}')\right| = \frac{U_W U_Y}{N}. \tag{1}$$

So to satisfy ρ-zCDP, we could release

$$\hat{\theta}^{(\rho)}(\boldsymbol{y}, \boldsymbol{w}) \triangleq \hat{\theta}(\boldsymbol{y}, \boldsymbol{w}) + \varepsilon, \qquad \varepsilon \sim N\left(0, \frac{\Delta(\hat{\theta})^2}{2\rho}\right). \tag{2}$$

This naive approach requires Gaussian noise with variance that grows with U_W^2, which could be prohibitively expensive if U_W is large (i.e., if some units have a significantly larger survey weight than others). Such issues are especially pronounced for surveys, where typical sample sizes are much smaller than those used for DP evaluations [8].

To motivate an alternate approach, let $\hat{\theta}_0$ be the *unweighted* sample mean, allowing us to suggestively rewrite

$$\hat{\theta} = \hat{\theta}_0 + \mathrm{Sign}(\hat{\theta} - \hat{\theta}_0)|\hat{\theta} - \hat{\theta}_0|. \tag{3}$$

The first term in the estimand is the standard, low-sensitivity unweighted mean for which classical DP release mechanisms offer optimal utility guarantees [2]. The second term in the estimand contains two components which we call the *weighting bias sign* and *absolute weighting discrepancy* (AWD), respectively. The AWD's high sensitivity makes DP survey estimation difficult in practice. When survey response variables and survey weights are highly correlated do we see large AWD values; when the two are less correlated, the AWD can be quite small. Therefore we should consider not only whether it's possible to inflate statistic sensitivities to accommodate survey weighting, but whether such an inflation significantly changes our resulting inferences.

We consider a regularization parameter $\lambda \in [0, 1]$ that reduces our estimand's dependence on AWD by shrinking our survey weights towards uniform values. We define this using the function G_λ and the associated biased estimate $\hat{\theta}_\lambda$

$$G_\lambda(\boldsymbol{w}) \triangleq (1 - \lambda)\boldsymbol{w} + \frac{\lambda N}{n}\mathbb{1}_n, \qquad \hat{\theta}_\lambda(\boldsymbol{y}, \boldsymbol{w}) = \hat{\theta}(\boldsymbol{y}, G_\lambda(\boldsymbol{w})). \tag{4}$$

Estimating $\hat{\theta}_\lambda$ is easier under ρ-zCDP because the statistic is less sensitive to changes in survey weights by design. This reduction in sensitivity comes at a cost based on the difference between $\hat{\theta}(\boldsymbol{y}, G_\lambda(\boldsymbol{w}))$ and $\hat{\theta}(\boldsymbol{y}, \boldsymbol{w})$. We give this quantity a name, *mechanism bias*, to quantify bias induced by the DP mechanism, defined as

$$D(\lambda) \triangleq \mathbb{E}_\varepsilon[\hat{\theta}_\lambda^{(\rho)}(\boldsymbol{y}, \boldsymbol{w})] - \hat{\theta} = \lambda(\hat{\theta}_0 - \hat{\theta}). \tag{5}$$

Using this expression for mechanism bias, we can consider the mean-squared error (MSE) about the confidential non-DP estimate

$$\ell(\lambda; \boldsymbol{y}, \boldsymbol{w}) = \mathbb{E}_\varepsilon\left[(\hat{\theta}_\lambda^{(\rho)} - \hat{\theta})^2\right] = \frac{\Delta(\hat{\theta}_\lambda)^2}{2\rho} + D(\lambda)^2. \tag{6}$$

Equation (6) characterizes a three-way "bias-variance-privacy" trilemma for DP survey estimation. As we reduce the mechanism bias of our survey estimates, we require more additive noise to satisfy ρ-zCDP; moreover, this effect becomes more extreme as ρ gets smaller. If we were able to optimally navigate this trade-off for a fixed value of ρ, we could try to minimize ℓ as a function of λ. This yields the following Lemma.

Lemma 1. *Consider minimizing the loss function in Equation* (6)*. Then*

1. *The MSE in Equation* (6) *is minimized by*

$$\lambda^* \triangleq \min\left\{1, \frac{\frac{U_W}{\rho}\left(\frac{U_Y}{N}\right)^2\left(U_W - \frac{N}{n}\right)}{\left[\frac{1}{\rho}\left(\frac{U_Y}{N}\right)^2\left(U_W - \frac{N}{n}\right)^2 + 2(\hat{\theta}_0 - \hat{\theta})^2\right]}\right\}. \tag{7}$$

2. $\lambda^* > 0$ *iff* $U_W > N/n$.

3. $\lambda^* < 1$ *iff*

$$\left|\hat{\theta}_0 - \hat{\theta}\right| > \sqrt{\frac{U_Y^2}{2\rho N n}\left(U_W - \frac{N}{n}\right)}, \tag{8}$$

or, equivalently,

$$\rho > \frac{U_Y^2}{2(\hat{\theta}_0 - \hat{\theta})^2 N n}\left(U_W - \frac{N}{n}\right). \tag{9}$$

Lemma 1 has an interesting interpretation. First, if the weighting scheme is informative (i.e., if $U_W > N/n$), then is it *never* optimal to use survey weights as-is without some degree of regularization (i.e., $\lambda^* > 0$). Second, if the effect of the mechanism bias introduced by shrinking survey weights is not sufficiently large, or if ρ is sufficiently small, then the optimal DP strategy to minimize $\ell(\lambda\cdot; \boldsymbol{y}, \boldsymbol{w})$ is to *ignore* the survey weights entirely (i.e., $\lambda^* = 1$).

The optimal degree of regularization λ^* depends on the confidential data through the AWD, $|\hat{\theta}_0 - \hat{\theta}|$. Therefore, we propose the following two-step approach to estimating $\hat{\theta}$ using $(\rho_1 + \rho_2)$-zCDP:

1. Estimate λ^* while satisfying ρ_1-zCDP by using the exponential mechanism by sampling $\hat{\lambda}^{(\rho_1)}$ from the density

$$f(\lambda) \propto \mathbb{1}_{\{\lambda \in [0,1]\}} \exp\left(-\frac{\sqrt{2\rho_1}}{2\Delta(\ell)}\ell(\lambda; \boldsymbol{y}, \boldsymbol{w})\right), \tag{10}$$

   where we show $\Delta(\ell) = (\Delta(\hat{\theta}) - \Delta(\hat{\theta}_0))^2$.

2. Sample $\hat{\theta}^{(\rho_2)}_{\hat{\lambda}^{(\rho_1)}}$ according to the Gaussian mechanism using weights shrunk with estimated optimal lambda value $\hat{\lambda}^{(\rho_1)}$.

By using $\hat{\lambda}^{(\rho_1)}$ as a noisy proxy for λ^*, we still maintain a high probability of reducing the noise needed to satisfy ρ-zCDP for $\rho = \rho_1 + \rho_2$ when adding Gaussian noise to the weighted mean estimate. We can then spend additional privacy loss budget to calculate DP confidence intervals that simultaneously account for errors due to sampling and DP noise; their interval width and coverage properties are discussed in the technical paper accompanying this chapter.

3 Data Analysis

To demonstrate the methodology, we analyze the PSID data and also simulate data. We use family-level survey records published from 2019, comprising $n = 9420$ families from a population of $N \approx 1.29 \times 10^8$ families. For the purposes of this evaluation, we treat the provided survey weights as design-based (in reality they are adjusted; see [24] for details). Under this weighting scheme, we are setting $U_W = 6 \times 10^4$ as a conservative upper bound on our survey weights. We seek to estimate population means of the variables in Table 1 using ρ-zCDP. In addition to the collected variables in PSID, we include one synthesized random variable, `bern`, which contains iid Bernoulli draws to simulate a random survey response that is theoretically independent of the survey weights by construction.

Variable	Description	U_Y	$\hat{\theta}_0 - \hat{\theta}$
`inc3`	Cube-root-transformed family income	150	-.67
`pov`	1 if family income below poverty line, else 0	1	.022
`nf`	Number of family members	20	.27
`bern`	iid Bernoulli(.5) random draws	1	.004

Table 1: Selected PSID variables and the simulated Bernoulli variable.

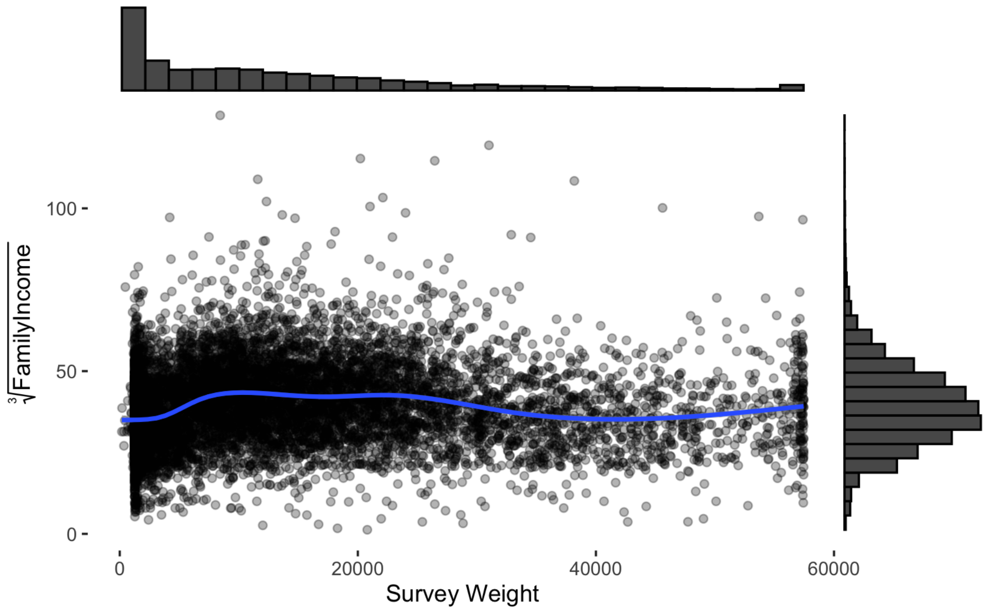

Figure 1: Bivariate scatter plot of survey weights (x-axis) and cube-root-transformed family income values (y-axis), with univariate histograms on the margins and a spline estimate of the central tendency in blue.

PSID facilitates research on employment, income, wealth, health, family and child development, and other sociodemographic and economic topics, with an oversample of lower-income families. We plot the relationship between survey weights and `inc3` in Figure 1. The survey weights are weakly correlated with family income (Spearman's rank correlation of approximately .14). If ignoring survey weights, we would underestimate the national average family income and overestimate the national poverty rate, as expected.

Next, we visualize the privacy-bias-variance trade-off for the population mean estimates of our different variables. Figure 2 shows how the theoretical bias-variance-privacy trade-off manifests for estimating the survey-weighted average cube-root-transformed income (`inc3`) and proportion of families below the 2019 poverty line (`pov`). For comparison purposes, we also include two hypothetical responses: simulated iid Bernoulli(.5) responses (`bern`) and a copy of the survey weights themselves (`wgt`), representing minimal and maximal correlation between survey weights and responses. We plot the noise-to-signal ratio as the theoretical MSE over the weighted mean estimate on the y-axis, with the regularization parameter λ on the x-axis. We see that as the magnitude of the bias decreases (moving from top left subfigure to bottom right subfigure), the optimal MSE is achieved at larger values of λ^* for the same privacy loss budget ρ_2. Moreover, as ρ_2 decreases, λ^* increases for each response variable under consideration. We see that for reasonably small choices of ρ_2, we tend to reject small λ to optimize the bias-variance trade-off at each fixed ρ_2 value.

3.1 End-to-end DP Inferences

In this section, we simulate DP confidence intervals for the survey weighted population mean of `inc3`, assessing their width and coverage properties. We consider $\rho_1, \rho_2, \rho_3 \in [10^{-3}, 10^{-1}]$, which covers the full spectrum of regularization from λ^*, as shown in Figure 2. ρ_1 and ρ_2 correspond to the privacy loss spent on estimating λ^* and $\hat{\theta}_{\lambda^*}$, and ρ_3 corresponds to the privacy loss spent on estimating the confidence interval width. We also vary the interval width upper bound (i.e., the $(1-\alpha_v)$*100% bound, shown by different colors in the following plot) to show trade-offs between coverage and interval width.

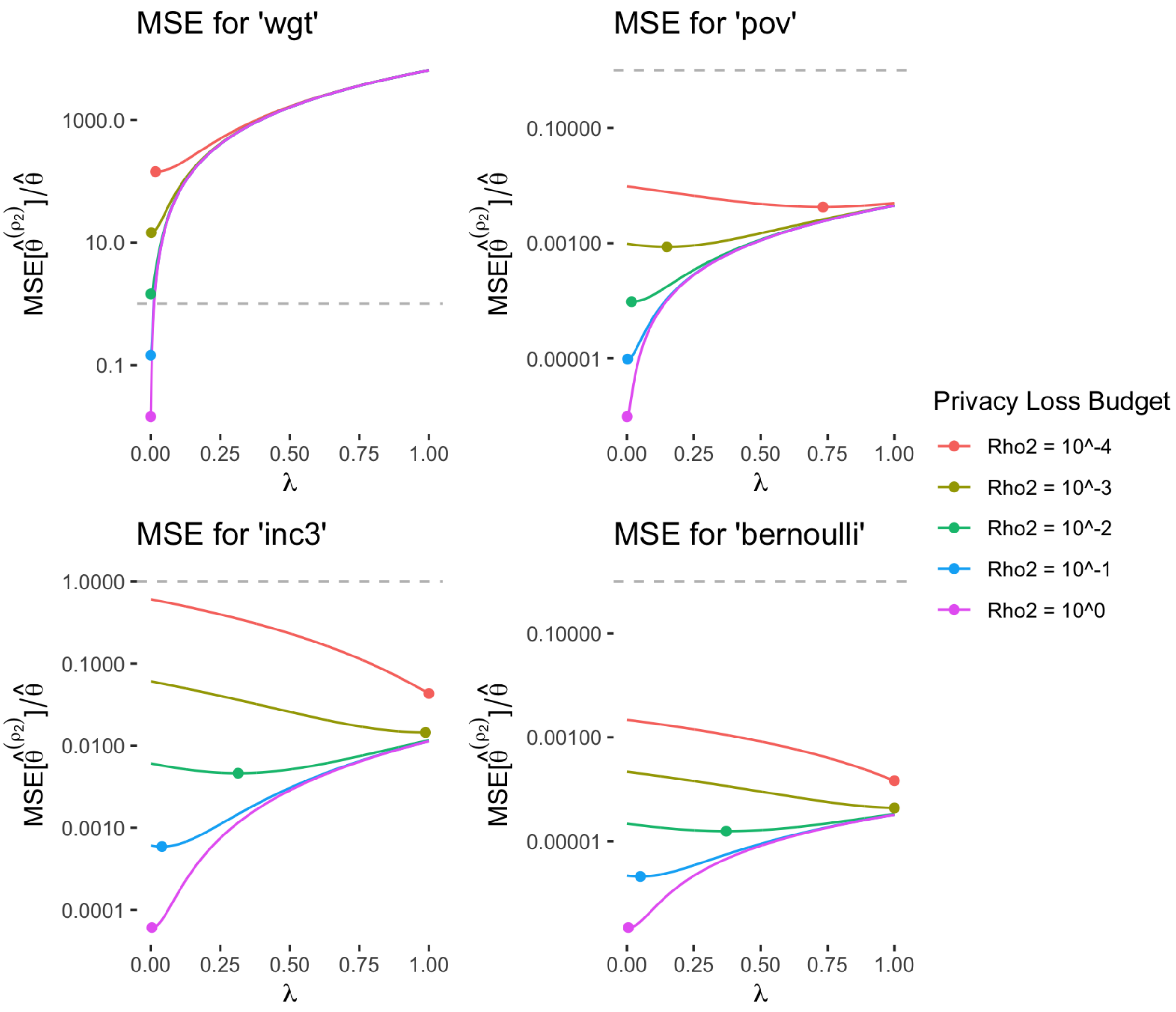

Figure 2: Realized noise-to-signal (DP mean-squared error divided by non-DP mean estimate, y-axis) as a function of λ (x-axis) for different values of privacy loss budget ρ_2 (colored lines). Subplots are ordered with decreasing correlation between response variable and survey weights. Points refer to theoretical minimum values, which depend on confidential data and do not satisfy DP.

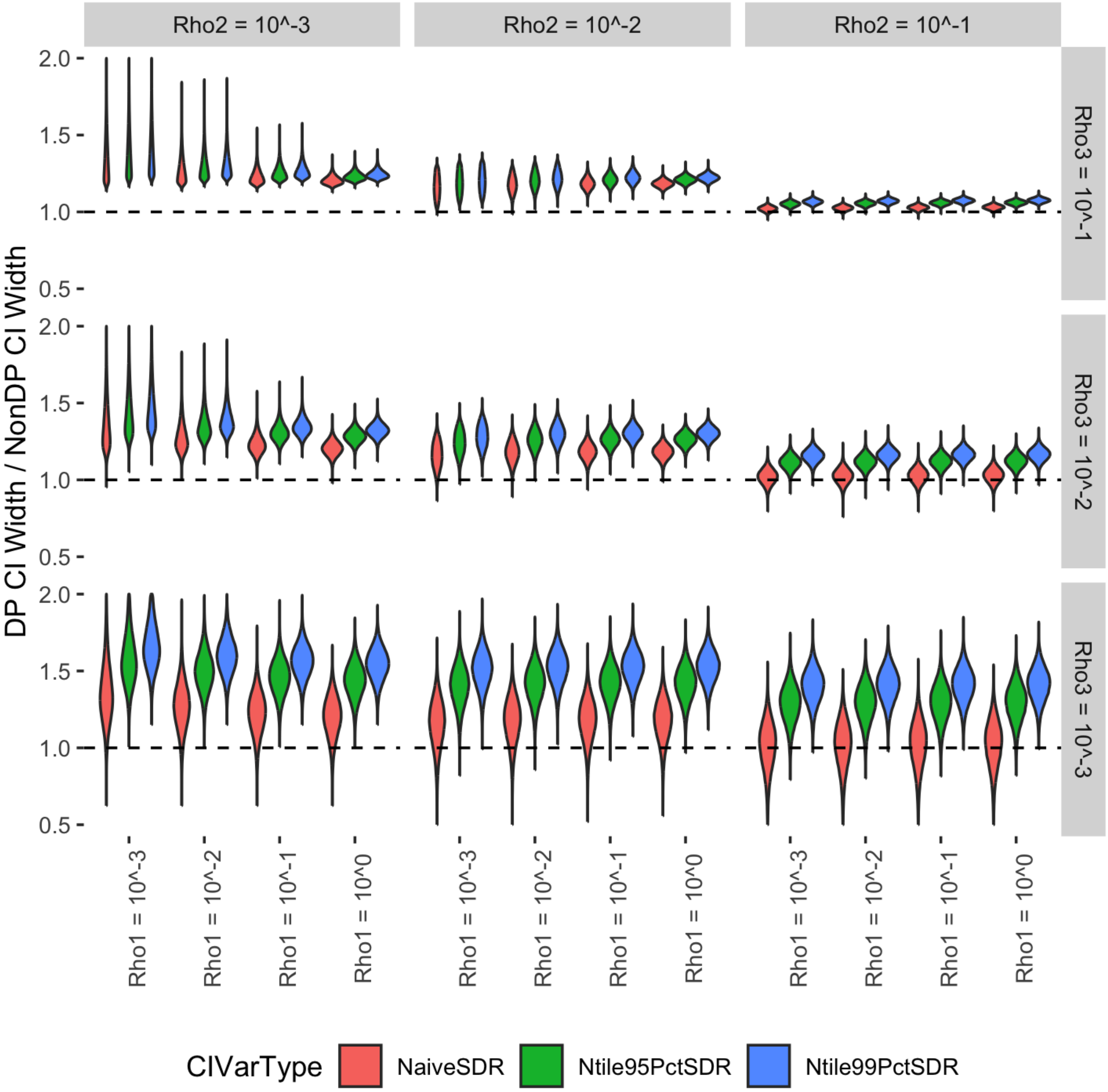

Figure 3: Ratio of DP to non-DP confidence interval widths (y-axis) by values of ρ_1 (x-axis), ρ_2 (subplot columns), ρ_3 (subplot rows), and α_v (colors). Dashed line corresponds to equality (1:1 ratio).

Figure 3 compares the interval widths of our proposed algorithm to their non-DP counterparts. For each violin plot, we show the distribution of the ratio for the DP confidence interval over the non-DP confidence interval. The dashed horizontal line at 1.0 corresponds to equality. As expected, increasing either ρ_1, ρ_2, or ρ_3 decreases the DP confidence interval width relative to the non-DP interval width. Of particular interest is different values for α_v, represented by the different violin plot colors (.5, .05, and .01, respectively). As expected, decreasing α_v gives us wider confidence intervals by accounting for more potential uncertainty in the interval width.

4 Discussion

Our work theoretically and empirically suggests that survey weight regularization, when used appropriately, can reduce the amount of additional noise needed to preserve DP. By adaptively considering how much to shrink weights toward uniform values while satisfying DP, we develop methods that are operationally feasible while allowing uncertainty quantification at different precision levels throughout the entire estimation process.

While our proposed methods can admit the construction of valid finite-sample confidence intervals and asymptotic confidence intervals, different parameter choices may produce intervals that are either too wide or narrow in practice. When selecting privacy loss budgets for each stage of the algorithm, we recommend incorporating as much domain knowledge as possible. For example, by simulating a distribution of plausible AWD values from prior knowledge, one can establish which kinds of survey weighting biases could be correctable at different privacy loss budgets without peeking at the confidential data.

While DP theoretically forbids using data-dependent hyperparameters without DP mechanisms, many commonly used DP algorithms and analyses do not adhere to this rule [1, 25], necessarily yielding additional privacy vulnerabilities in practice [19, 18]. It could be the case that tuning certain hyperparameters could substantially improve the end-to-end usefulness of our estimators at a modest expense to privacy risk. Understanding this would require a much more extensive and nuanced privacy analysis than a simple comparison of privacy loss budgets. Still, such privacy analysis could help illuminate where DP itself fundamentally limits the kinds of statistical validity offered in survey settings, where worst-case data generating scenarios may be unrealistic in practice.

References

[1] Martin Abadi, Andy Chu, Ian Goodfellow, H Brendan McMahan, Ilya Mironov, Kunal Talwar, and Li Zhang. Deep learning with differential privacy. In *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security*, pages 308–318, 2016.

[2] Jordan Awan and Salil Vadhan. Canonical noise distributions and private hypothesis tests. *The Annals of Statistics*, 51(2):547–572, 2023.

[3] Borja Balle, Gilles Barthe, and Marco Gaboardi. Privacy amplification by subsampling: Tight analyses via couplings and divergences. *Advances in neural information processing systems*, 31, 2018.

[4] Jean-François Beaumont. A new approach to weighting and inference in sample surveys. *Biometrika*, 95(3):539–553, 2008. Publisher: Oxford University Press.

[5] Kenneth A Bollen, Paul P Biemer, Alan F Karr, Stephen Tueller, and Marcus E Berzofsky. Are survey weights needed? A review of diagnostic tests in regression analysis. *Annual Review of Statistics and Its Application*, 3:375–392, 2016. Publisher: Annual Reviews.

[6] Mark Bun, Jörg Drechsler, Marco Gaboardi, Audra McMillan, and Jayshree Sarathy. Controlling privacy loss in sampling schemes: An analysis of stratified and cluster sampling. In *3rd Symposium on Foundations of Responsible Computing (FORC 2022)*, 2022.

[7] Mark Bun and Thomas Steinke. Concentrated differential privacy: Simplifications, extensions, and lower bounds. In *Theory of Cryptography Conference*, pages 635–658. Springer, 2016.

[8] Jörg Drechsler. Differential Privacy for Government Agencies—Are We There Yet? *Journal of the American Statistical Association*, 118(541):761–773, 2023. Publisher: Taylor & Francis.

[9] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In *Theory of cryptography conference*, pages 265–284. Springer, 2006.

[10] Cynthia Dwork, Aaron Roth, and others. The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.*, 9(3-4):211–407, 2014.

[11] Andrew Gelman. Struggles with Survey Weighting and Regression Modeling. *Statistical Science*, 22(2):153–164, 2007.

[12] David Haziza and Jean-François Beaumont. Construction of weights in surveys: A review. *Statistical Science*, 32:206–226, 2017.

[13] Jingchen Hu, JÖrg Drechsler, and Hang J Kim. Accuracy Gains from Privacy Amplification Through Sampling for Differential Privacy. *Journal of Survey Statistics and Methodology*, 10(3):688–719, 2022. Publisher: Oxford University Press.

[14] Jingchen Hu, Terrance D Savitsky, and Matthew R Williams. Private tabular survey data products through synthetic microdata generation. *Journal of Survey Statistics and Methodology*, 10(3):720–752, 2022. Publisher: Oxford University Press.

[15] Gautam Kamath, Argyris Mouzakis, Matthew Regehr, Vikrant Singhal, Thomas Steinke, and Jonathan Ullman. A Bias-Variance-Privacy Trilemma for Statistical Estimation. *arXiv preprint arXiv:2301.13334*, 2023.

[16] Shurong Lin, Mark Bun, Marco Gaboardi, Eric D Kolaczyk, and Adam Smith. Differentially Private Confidence Intervals for Proportions under Stratified Random Sampling. *arXiv preprint arXiv:2301.08324*, 2023.

[17] Roderick J.A. Little and S. Vartivarian. Does weighting for nonresponse increase the variance of survey means? *Survey Methodology*, 31(2):161–168, 2005.

[18] Shubhankar Mohapatra, Sajin Sasy, Xi He, Gautam Kamath, and Om Thakkar. The role of adaptive optimizers for honest private hyperparameter selection. In *Proceedings of the aaai conference on artificial intelligence*, volume 36, pages 7806–7813, 2022. Issue: 7.

[19] Nicolas Papernot and Thomas Steinke. Hyperparameter tuning with renyi differential privacy. *arXiv preprint arXiv:2110.03620*, 2021.

[20] Jerome P. Reiter and Trivellore E. Raghunathan. The multiple adaptations of multiple imputation. *Journal of the American Statistical Association*, 102:1462–1471, 2007.

[21] Yajuan Si, Natesh S. Pillai, and Andrew Gelman. Bayesian nonparametric weighted sampling inference. *Bayesian Analysis*, 10(3):605–625, 2015.

[22] Yajuan Si, Rob Trangucci, Jonah Sol Gabry, and Andrew Gelman. Bayesian hierarchical weighting adjustment and survey inference. *Survey Methodology*, 46(2):181–214, 2020.

[23] Aleksandra Slavković and Jeremy Seeman. Statistical data privacy: A song of privacy and utility. *Annual Review of Statistics and Its Application*, 10, 2023. Publisher: Annual Reviews.

[24] Survey Research Center at the Institute for Social Research, University of Michigan Ann Arbor. Panel study of income dynamics, public use dataset, 2019.

[25] Florian Tramer and Dan Boneh. Adversarial training and robustness for multiple perturbations. *Advances in neural information processing systems*, 32, 2019.