Whose Data Is It Anyway?
Flavors of Differential Privacy for Survey Sampling

Jörg Drechsler & James Bailie

December 2023

1 Introduction

The survey is the workhorse of statistical agencies. For example, the U. S. Census Bureau (USCB) conducts more than 100 surveys annually [U.S. Census Bureau, 2023] including key data collections such as the American Community Survey (ACS), the Current Population Survey (CPS), the Survey of Income and Program Participation (SIPP) and the new Annual Business Survey (ABS). The data gathered through these surveys provide invaluable information on the economy and on society more generally. They are used by various stakeholders – for example, businesses, researchers, local and federal governments, media and not-for-profits – to make investment decisions, to inform policy and to allocate government funding, among many other uses.

On the other hand, national statistical organizations (NSOs) have acknowledged for decades their obligation to maintain the confidentiality of survey respondents due to legal and ethical considerations, but also to safeguard institutional trust and thus sustain the quality of their data products. To address these conflicting goals, various methods have been proposed over the years to protect the confidentiality of survey respondents while still maintaining the value of the data for the different stakeholders involved. In the last two decades, a new framework for assessing the privacy of statistical data products has emerged: *differential privacy* (DP) [Dwork et al., 2006b]. This framework is mathematically appealing as it offers a formal guarantee: any single unit's influence on the probability of observing a specific output is bounded. This guarantee translates into quantifiable measures of protection against an adversary seeking to learn the confidential responses, although not without some complications [Cuff and Yu, 2016, Tschantz et al., 2020, Kifer and Machanavajjhala, 2011, Bailie et al., 2023+b]. Beyond the formal guarantees, DP is attractive as it allows for full transparency of the methods that were used to protect the output, without further loss to respondents' privacy beyond that associated with publishing the outputted statistics and the

DP specification. This is in contrast to many methods that are currently employed at statistical agencies, which rely on hiding some of the parameters of the protection mechanism (such as the variance of the noise term when noise addition is used to protect a continuous attribute) to ensure privacy. With DP, agencies typically release all details about the mechanism including the levels of the privacy parameters. This implies that at least in principle users of the protected data will be able to account for the additional uncertainty that is introduced through the protection step (although this turns out to be difficult for many of the algorithms used in practice so far). Finally, DP offers several additional properties such as immunity to postprocessing and composition of privacy budget (see for example Dwork and Roth [2014]). This second property makes the DP framework specifically interesting for statistical agencies as it allows for the quantification of the privacy loss over multiple data releases.

These attractive features motivated the adoption of DP in the private sector [Erlingsson et al., 2014, Apple's Differential Privacy Team, 2017, Ding et al., 2017, Messing et al., 2020, Uber Security, 2017], as well as at some NSOs such as the USCB [Machanavajjhala et al., 2008, Foote et al., 2019, Abowd and Hawes, 2023]. Still, all deployments of DP so far have focused on situations in which the data to be protected coincided with the population of interest. As pointed out above, this is rarely the case for data collected by NSOs. Except for censuses – which are typically only conducted every five to ten years – and some administrative databases, most data at statistical agencies are collected via probability surveys. In the survey context, information is only gathered for a small fraction of the population, but the careful design of the selection process and several adjustment steps after the survey has been conducted (such as weighting, editing and imputation) ensure that the resulting data can be used to obtain approximately unbiased estimates for the population of interest. (We will offer a more detailed review of the survey process in Section 2.2.) However, how to properly account for these particularities within the framework of DP is currently poorly understood (see Reiter [2019] and Drechsler [2023] for an in-depth discussion of the challenges that will arise in this context). Gaining a better understanding is especially critical as the USCB have publicly committed to adopting DP for all their data products [US Census Bureau, 2018] – a resolution that has been recently reaffirmed [US Census Bureau, 2022]. (In fact, the USCB only recommitted to adopting "formal privacy"; however we are not aware of any other formal privacy frameworks except for DP.) In the same 2022 press release, the USCB concluded that "the science does not yet exist" to implement DP for their flagship survey – the ACS – highlighting the need for additional research in this area.

We are aware of only a few papers that address DP in the survey context and, moreover, all these papers only focus on specific aspects of this process. Lin et al. [2023] study how to estimate the mean of a binary variable under DP assuming stratified sampling using proportional allocation and simple random sampling within the strata. Bun et al. [2022] investigate if the complex sam-

pling designs commonly used in the survey context can offer increased privacy protection building on previous results showing that simple sampling procedures such as simple random sampling or Poisson sampling will amplify the privacy protection [Balle et al., 2020]. We will summarize their findings in Section 4.1. Finally, in some preliminary work, Das et al. [2022] study the effects of imputation. They find that if DP is only considered when analyzing the imputed data, the required privacy loss budget can increase linearly with the number of missing cases. They also show that this problem can be avoided – at least for certain imputation schemes – if DP is already considered during imputation.

This paper aims to contribute to the growing literature on DP in the survey context by discussing the implications of (for example) whether the privacy guarantees should hold only for the sampled units or the entire population. We identify ten settings that vary in their assumptions about the data at different levels (the responding sample, the selected sample, the sampling frame, and the target population). Building on the framework introduced in Bailie et al. [2023+a], we formalize the DP flavors for these settings and discuss their implications on both data utility and privacy.

2 Background

2.1 Differential Privacy

DP studies data-release mechanisms – functions T which take as input the confidential object $\boldsymbol{x} \in \mathcal{X}$ and a random seed $\omega \in \Omega$, and output a stochastic summary $T(\boldsymbol{x}, \omega)$ of $\boldsymbol{x}$.

Definition 2.1. A *data-release mechanism* is a function $T : \mathcal{X} \times \Omega \to \mathcal{T}$ along with a probability measure on Ω and a σ-algebra on $\mathcal{T}$ such that $T(\boldsymbol{x}, \cdot)$ is measurable for all $\boldsymbol{x} \in \mathcal{X}$.

It is convenient to think of a data-release mechanism as a function $\boldsymbol{x} \mapsto \mathsf{P}_{\boldsymbol{x}}(T \in \cdot)$. Here the probability distribution $\mathsf{P}_{\boldsymbol{x}}(T \in \cdot)$ of the summary $T(\boldsymbol{x}, \omega)$ is the push-forward measure induced by the distribution of the random seed ω, taking $\boldsymbol{x}$ as fixed:

$$\mathsf{P}_{\boldsymbol{x}}(T \in B) := \Pr(\{\omega \in \Omega : T(\boldsymbol{x}, \omega) \in B\}),$$

where B is any measurable subset of the output space $\mathcal{T}$. DP is the condition that the data-release mechanism is Lipschitz continuous – i.e. that the distance between outputs $\mathsf{P}_{\boldsymbol{x}}$ and $\mathsf{P}_{\boldsymbol{x}'}$ is at most a multiplicative factor of the distance between the corresponding inputs $\boldsymbol{x}$ and $\boldsymbol{x}'$.

Example 2.2. For pure ε-DP, as defined in Dwork et al. [2006b], the multiplicative factor is ε; the distance between inputs $\boldsymbol{x}$ and $\boldsymbol{x}'$ is the Hamming distance; and the distance between outputs $\mathsf{P}_{\boldsymbol{x}}$

and $\mathsf{P}_{\boldsymbol{x}'}$ is the multiplicative distance

$$\textsc{Mult}(\mathsf{P}_{\boldsymbol{x}}, \mathsf{P}_{\boldsymbol{x}'}) = \sup_E \left| \ln \frac{\mathsf{P}_{\boldsymbol{x}}(T \in E)}{\mathsf{P}_{\boldsymbol{x}'}(T \in E)} \right|,$$

(where the supremum is over all measurable sets E). (For readers that are familiar with the definition of pure ε-DP in terms of neighboring datasets $\boldsymbol{x}$ and $\boldsymbol{x}'$, the Lipschitz condition for non-neighbors is implied by the condition for neighbors by group privacy. Hence, the neighbor definition of pure ε-DP is the equivalent to the above definition.)

For approximate (ε, δ)-DP [Dwork et al., 2006a], the multiplicative factor is again ε; the distance between inputs is given by

$$d_{\mathcal{X}}^{\text{neighbors}}(\boldsymbol{x}, \boldsymbol{x}') = \begin{cases} 0 & \text{if } \boldsymbol{x} = \boldsymbol{x}', \\ 1 & \text{if } \boldsymbol{x} \text{ and } \boldsymbol{x}' \text{ are neighbors,} \\ \infty & \text{otherwise;} \end{cases}$$

and the distance between outputs is given by

$$\textsc{Mult}^{\delta}(\mathsf{P}_{\boldsymbol{x}}, \mathsf{P}_{\boldsymbol{x}'}) = \sup_E \left\{ \ln \frac{[\mathsf{P}_{\boldsymbol{x}}(E) - \delta]^+}{\mathsf{P}_{\boldsymbol{x}'}(E)}, \ln \frac{[\mathsf{P}_{\boldsymbol{x}'}(E) - \delta]^+}{\mathsf{P}_{\boldsymbol{x}}(E)}, 0 \right\},$$

(where $[x]^+ = \max\{x, 0\}$ and the supremum is over all measurable sets E). Note that $d_{\mathcal{X}}^{\text{neighbors}}$ and $\textsc{Mult}^{\delta}$ are not distances in the mathematical sense of a metric; we will instead refer to them as *divergences* from herein. Since $\textsc{Mult}^{\delta}$ does not satisfy the triangle inequality, approximate (ε, δ)-DP's group privacy budget does not increase linearly with the group size; hence we cannot replace $d_{\mathcal{X}}^{\text{neighbors}}$ with the Hamming distance, as we did for pure ε-DP.

A data-release mechanism T satisfies DP if and only if it is Lipschitz continuous. There are different *flavors* (i.e. types or versions) of DP; each of these flavors corresponds to different ways to specify continuity. For our purposes, there are four components to the specification of Lipschitz continuity. Most obviously, there are the divergences $d_{\mathcal{X}}(\boldsymbol{x}, \boldsymbol{x}')$ and $D_{\mathrm{Pr}}(\mathsf{P}_{\boldsymbol{x}}, \mathsf{P}_{\boldsymbol{x}'})$. These divergences measure the 'distance' between any two inputs $\boldsymbol{x}$ and $\boldsymbol{x}'$, or between any two output probabilities $\mathsf{P}_{\boldsymbol{x}}$ and $\mathsf{P}_{\boldsymbol{x}'}$. Secondly, there is the domain $\mathcal{X}$ of the data-release mechanism, which – as we shall see – serves as the parameter space of the attacker's inferential model.[1] Finally, there is the data multiverse $\mathscr{D}$, which allows the data custodian to restrict the Lipschitz continuity condition to certain pairs of inputs – as is often desirable in practice. For example, we may only want to compare samples drawn from the same population. This restriction is achieved by specifying the data multiverse $\mathscr{D}$.

[1] This discussion is still missing at this point, but will be included in the final version of the paper.

Definition 2.3 (Bailie et al. [2023+a])**.** A *differential privacy flavor* is a quadruple $(\mathcal{X}, \mathscr{D}, d_{\mathcal{X}}, D_{\mathrm{Pr}})$ where:

1. The *protection domain* $\mathcal{X}$ is a set, which is typically the *data space* – the space of all (theoretically-possible) datasets. More generally, $\mathcal{X}$ is a set of *protection objects*, which are denoted by $\boldsymbol{x}$ or $\boldsymbol{x}'$.
2. The *data multiverse* $\mathscr{D} \subset 2^{\mathcal{X}}$ is a set of *data universes*, which are denoted by $\mathcal{D}$ or $\mathcal{D}'$.
3. The *measure of input variations* $d_{\mathcal{X}}$ is a divergence on $\mathcal{X}$ – i.e. a function $\mathcal{X} \times \mathcal{X} \to \mathbb{R}^{\geq 0}$ such that $d_{\mathcal{X}}(\boldsymbol{x}, \boldsymbol{x}) = 0$ for all $\boldsymbol{x} \in \mathcal{X}$.
4. The *measure of output variations* D_{Pr} is a divergence on the space of all probability distributions – i.e. a function $\mathcal{P}^2 \to \mathbb{R}^{\geq 0}$ on pairs $(\mathsf{P}, \mathsf{Q}) \in \mathcal{P}^2$ of probabilities P, Q defined on the same measurable space such that $D_{\mathrm{Pr}}(\mathsf{P}, \mathsf{P}) = 0$ for all probabilities P.

Once we have specified the four components for Lipschitz continuity via a DP flavor, we also need to specify the multiplicative constant (known as the Lipschitz constant) which controls the rate between output and input variations. Together, choices for these five components are called a DP specification.

Definition 2.4. A *differential privacy specification* is a quintuple $(\mathcal{X}, \mathscr{D}, d_{\mathcal{X}}, D_{\mathrm{Pr}}, \varepsilon_{\mathcal{D}})$ consisting of a DP flavor $(\mathcal{X}, \mathscr{D}, d_{\mathcal{X}}, D_{\mathrm{Pr}})$ and a privacy-loss budget $\varepsilon_{\mathcal{D}} : \mathscr{D} \to \mathbb{R}^{\geq 0}$. We denote a DP specification by $\varepsilon_{\mathcal{D}}$-DP$(\mathcal{X}, \mathscr{D}, d_{\mathcal{X}}, D_{\mathrm{Pr}})$.

A data-release mechanism $T : \mathcal{X} \times \Omega \to \mathcal{T}$ *satisfies* the DP specification $\varepsilon_{\mathcal{D}}$-DP$(\mathcal{X}, \mathscr{D}, d_{\mathcal{X}}, D_{\mathrm{Pr}})$ if, for all data universes $\mathcal{D} \in \mathscr{D}$, and all $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{D}$,

$$D_{\mathrm{Pr}}\big[\mathsf{P}_{\boldsymbol{x}}(T \in \cdot), \mathsf{P}_{\boldsymbol{x}'}(T \in \cdot)\big] \leq \varepsilon_{\mathcal{D}} d_{\mathcal{X}}(\boldsymbol{x}, \boldsymbol{x}'). \tag{2.1}$$

Definition 2.5 (Comparability)**.** Given a DP flavor $(\mathcal{X}, \mathscr{D}, d_{\mathcal{X}}, D_{\mathrm{Pr}})$, two protection objects (e.g. datasets) $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{X}$ are *comparable* when $d_{\mathcal{X}}(\boldsymbol{x}, \boldsymbol{x}') < \infty$ and there exists a data universe $\mathcal{D} \in \mathscr{D}$ such that $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{D}$.

When $\boldsymbol{x}$ and $\boldsymbol{x}'$ are comparable, a DP specification restricts the divergence between $\mathsf{P}_{\boldsymbol{x}}$ and $\mathsf{P}_{\boldsymbol{x}'}$. On the other hand, a DP specification places no restrictions on the output variations for non-comparable $\boldsymbol{x}, \boldsymbol{x}'$, except perhaps implicitly through comparable $\boldsymbol{x}, \boldsymbol{x}'$.

Definition 2.6 (Completeness)**.** A DP flavor $(\mathcal{X}, \mathscr{D}, d_{\mathcal{X}}, D_{\mathrm{Pr}})$ is *complete* if $d_{\mathcal{X}}(\boldsymbol{x}, \boldsymbol{x}') < \infty$ for all $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{D}$ and all $\mathcal{D} \in \mathscr{D}$.

Completeness means that every pair of protection objects $\boldsymbol{x}, \boldsymbol{x}'$ in the same data universe are comparable. In order to simplify the semantics of DP, we assume throughout this article that all

DP flavors under consideration are complete. This is a trivial assumption because, given any flavor, we can always construct a strictly equivalent, complete flavor (see Proposition 2.11 below).

Definition 2.7 (Stronger DP flavors)**.** A DP flavor $(\mathcal{X}, \mathscr{D}, d_{\mathcal{X}}, D_{\mathrm{Pr}})$ is *stronger* than another DP flavor $(\mathcal{X}', \mathscr{D}', d_{\mathcal{X}'}, D'_{\mathrm{Pr}})$ with *accounting function* $f_{\mathrm{acct}} : \mathbb{R}^{\geq 0} \to \mathbb{R}^{\geq 0}$ if there exists an injection $\iota : \mathcal{X}' \to \mathcal{X}$ such that

1. If $\boldsymbol{x}, \boldsymbol{x}' \in \mathcal{X}'$ are comparable under $(\mathcal{X}', \mathscr{D}', d_{\mathcal{X}'}, D'_{\mathrm{Pr}})$, then $\iota(\boldsymbol{x}), \iota(\boldsymbol{x}') \in \mathcal{X}$ are comparable under $(\mathcal{X}, \mathscr{D}, d_{\mathcal{X}}, D_{\mathrm{Pr}})$; and
2. For all budgets $\varepsilon_{\mathcal{D}} : \mathscr{D} \to \mathbb{R}^{\geq 0}$, and all mechanisms $T : \mathcal{X} \to \mathcal{T}$ which satisfy $\varepsilon_{\mathcal{D}}$-DP$(\mathcal{X}, \mathscr{D}, d_{\mathcal{X}}, D_{\mathrm{Pr}})$, the mechanism $T \circ \iota$ satisfies $\varepsilon'_{\mathcal{D}'}$-DP$(\mathcal{X}', \mathscr{D}', d_{\mathcal{X}'}, D'_{\mathrm{Pr}})$, where

$$\varepsilon'_{\mathcal{D}'} = \inf\{f(\varepsilon_{\mathcal{D}}) : \text{there exists } \boldsymbol{x}, \boldsymbol{x}' \in \mathcal{D}' \text{ such that } d_{\mathcal{X}'}(\boldsymbol{x}, \boldsymbol{x}') < \infty \text{ and } \iota(\boldsymbol{x}), \iota(\boldsymbol{x}') \in \mathcal{D}\},$$

   for all $\mathcal{D}' \in \mathscr{D}'$.

Moreover, if f_{acct} is the identity function, then we say that $(\mathcal{X}, \mathscr{D}, d_{\mathcal{X}}, D_{\mathrm{Pr}})$ is *strictly stronger* than $(\mathcal{X}', \mathscr{D}', d_{\mathcal{X}'}, D'_{\mathrm{Pr}})$. If a DP flavor A is stronger than another flavor B, then we write $A \leq_{f_{\mathrm{acct}}} B$ and say that A implies B.

Remark 2.8. Condition 1. in the above definition ensures that the resulting budget $\varepsilon'_{\mathcal{D}'}$ is finite, for all $\mathcal{D}' \in \mathscr{D}'$.

We motivate the above definition via two examples, which will become important in Section 5 when we discuss the composition of data-release mechanisms with different DP flavors. In the first example, suppose that two datasets $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}$ – each with different variables – were collected from the same group of units. Information from both datasets $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}$ were published via DP data-release mechanisms $T^{(1)}$ and $T^{(2)}$ respectively. We wish to study the privacy loss associated with the composition $\boldsymbol{T} = (T^{(1)}, T^{(2)})$, which is a function of the linked database $\boldsymbol{x}^{(1\vee 2)}$ containing both sets of variables. Since the datasets have different variables, the protection domains of $T^{(1)}$, $T^{(2)}$ and $\boldsymbol{T}$ are all different. Denoting these protection domains by $\mathcal{X}^{(1)}, \mathcal{X}^{(2)}$ and $\mathcal{X}^{(1\vee 2)}$ respectively, one can define injections $\iota^{(k)} : \mathcal{X}^{(k)} \to \mathcal{X}^{(1\vee 2)}$, for $k = 1, 2$ and prove that the DP flavor $(\mathcal{X}^{(1\vee 2)}, \{\mathcal{X}^{(1\vee 2)}\}, d_{\mathcal{X}^{(1\vee 2)}}, D_{\mathrm{Pr}})$ is strictly stronger than $(\mathcal{X}^{(k)}, \{\mathcal{X}^{(k)}\}, d_{\mathcal{X}^{(k)}}, D_{\mathrm{Pr}})$, where we define

$$d_{\mathcal{X}^{(k)}}(\boldsymbol{x}, \boldsymbol{x}') := d_{\mathcal{X}^{(1\vee 2)}}\big(\iota^{(k)}(\boldsymbol{x}), \iota^{(k)}(\boldsymbol{x}')\big).$$

For the second example, suppose instead that the two datasets $\boldsymbol{x}^{(1)}$ and $\boldsymbol{x}^{(2)}$ were collected from two mutually-disjoint samples $S^{(1)}$ and $S^{(2)}$. Again, the resulting DP flavors will have different protection domains but the DP flavor associated with $\boldsymbol{T}$ will be strictly stronger, since there are natural injections from each sample $S^{(k)}$ into their union $S^{(1)} \cup S^{(2)}$.

Definition 2.9 (Equivalent DP flavors)**.** Two DP flavors A and B are *equivalent* if there exist accounting functions f_{acct} and f'_{acct} such that

1. $B \leq_{f_{\text{acct}}} A$; and
2. $A \leq_{f'_{\text{acct}}} B$.

Moreover, these flavours are *strictly equivalent* if the accounting functions f_{acct} and f'_{acct} are both the identity function.

Definition 2.10 (Completion of a DP flavor)**.** Given a DP flavor $(\mathcal{X}, \mathscr{D}, d_{\mathcal{X}}, D_{\Pr})$, denote the protection objects *connected* to $\boldsymbol{x} \in \mathcal{X}$ by

$$\mathcal{D}_{\boldsymbol{x}} = \{\boldsymbol{x}' \in \mathcal{X} : d_{\mathcal{X}}(\boldsymbol{x}, \boldsymbol{x}') < \infty\}.$$

Then the *completion* $\overline{\mathscr{D}}$ of the data multiverse $\mathscr{D}$ is defined as

$$\overline{\mathscr{D}} = \{\mathcal{D} \cap \mathcal{D}_{\boldsymbol{x}} : \mathcal{D} \in \mathscr{D}, \boldsymbol{x} \in \mathcal{D}\}.$$

Proposition 2.11. *Given a DP flavor $(\mathcal{X}, \mathscr{D}, d_{\mathcal{X}}, D_{\Pr})$, the completion $(\mathcal{X}, \overline{\mathscr{D}}, d_{\mathcal{X}}, D_{\Pr})$ is complete and strictly equivalent to $(\mathcal{X}, \mathscr{D}, d_{\mathcal{X}}, D_{\Pr})$.*

(For a proof see Bailie et al. [2023+c].)

2.2 Survey Sampling

Surveys are conducted to learn some characteristics of a well-defined population by collecting information from a random subset of this population. Any survey sampling process relies on three key ingredients: the *target population* of interest; the *sampling frame* from which the random sample to be surveyed is drawn; and the *sampling design* for drawing this sample.

The target population (also known as the universe) is the scope of the survey; it is the population the survey is aiming to learn about. It is typically defined conceptually, while the sampling frame F, on the other hand, is an existent register containing names (or other identifiers) and contact information (postal or physical address, email, and/or telephone number) of survey units. The sampling frame F serves as the source from which the sample is drawn. Ideally, it covers all the units – and only those units – from the target population, although in practice this is rarely true. For the discussions in the remainder of the paper it is important to clearly distinguish between the target population and the sampling frame F. While the sampling frame aims to cover all the units from the target population, it might include units that are not part of the target population (*overcoverage*), and it might also miss units that should be included (*undercoverage*).

Finally, the sampling design is defined as a probability measure π on the set of all possible samples S. For each subset $S \subset F$, $\pi(S)$ is the probability that the realised sample is S. Sometimes the frame contains basic demographic information on the survey units, which can be used to construct the sample selection probabilities $\pi(S)$. The sampling designs used in practice are often complex multi-stage designs, with different sampling strategies (e.g. cluster sampling, stratified sampling, probability-proportional-to-size (PPS) sampling) for each of the different stages.

To illustrate the relevance of this discussion, we look at the Current Population Survey (CPS) conducted by the USCB for the Bureau of Labor Statistics (BLS). The target population of the CPS is the civilian noninstitutionalized population in the U.S, or, more exactly,

> "all people residing in the 50 states [of the U.S] and the District of Columbia who are not confined to institutions such as nursing homes and prisons, and who are not on active duty in the U.S. Armed Forces. Included are citizens of foreign countries who reside in the United States but do not live on the premises of an embassy. The civilian noninstitutional population ages 16 and older is the base population group used for CPS statistics" [U.S. Bureau of Labor Statistics, 2018].

The survey uses two different sampling frames: one for households and one for group quarters. Both are derived from the master address file (MAF) of the USCB: "The MAF is a national inventory of addresses that is continually updated by the U.S. Census Bureau to support its decennial programs and demographic surveys" [U.S. Census Bureau, 2019]. The CPS uses a stratified two-stage sampling design. In the first stage, the population is divided into geographical clusters and one cluster is sampled within each stratum using PPS sampling. A small group of households is selected in the second stage using systematic sampling based on a list sorted by demographic composition and geographic proximity. (See Section 2.2 in U.S. Census Bureau [2019] for a full description of the sampling methodology.)

2.3 Survey Weights

A distinctive feature of survey data is that they typically contain survey weights. Survey weights are provided by statistical agencies as a convenient tool to account for the sampling design and additional data preparation steps such as nonresponse adjustments when analyzing the data. Because complex sampling designs are often used (as we described in the previous section) and because not all sampled units actually respond to the survey, the resulting dataset cannot be treated as a simple random sample from the target population. Most estimators need to be adjusted to take these complications into account. For example, the (unweighted) sample mean can no longer be treated as an unbiased estimator for the mean in the population if the probability of being included in the responding-sample varies between the units. Instead, it is typical to use weighted estimators,

where individual data points are weighted according to their survey weights.

Survey weights are typically generated in three stages. In the first stage, *design weights* are generated that reflect the sampling design. In the second stage, *non-response adjustment weights* are used to account for different response propensities in different subgroups of the population. Finally, *calibration weights* try to correct for any deficiencies in the sampling frame and also help to reduce the variance of the final estimates.

In the remainder of the paper, we only consider the impacts of the sampling design on the privacy guarantees, that is, we treat the design weights as the final weights and ignore any adjustments to these weights to account for nonresponse or due to calibration. We note that these adjustment steps will always be data dependent and thus need to be taken into account in practice when assessing the privacy guarantees of the data-release mechanism. We leave this topic for future research (but see Drechsler [2023] for some high level thoughts regarding these challenges).

The design weights w^D are defined as the inverse of the probability of selection: $w_i^D = 1/\pi_i$, where π_i is the probability that unit i is selected into the sample (for $i = 1, \ldots, N$, with N the size of the sampling frame).

3 DP Flavors for Survey Statistics

As we have seen in Subsection 2.2, there are multiple phases in the creation of survey statistics: defining the target population, compiling the sampling frame, selecting the sample according to the sampling design, and collecting data from the responding units. (From herein, we use the term 'target sample' to refer to the sample of units selected by the sampling design from the sampling frame, in order to differentiate this sample with the responding sample – the set of units which were selected and responded.) The data output by each phase of this pipeline is feed into the subsequent phase as input. For example, data about the target population is used to compile the frame and data on the frame is used to select the sample.

The data custodian (e.g. NSO) could plausibly start the data-release mechanism T at any point along this data pipeline. That is, the data-release mechanism could take as input the dataset corresponding to any of the various phases. Moreover, the custodian could also plausibly condition on previous phases in the data pipeline (taking their data as invariant). Thus, the data custodian is faced with two decisions: what should the protection domain $\mathcal{X}$ be? And what should the data multiverse $\mathscr{D}$ be?

In this section, we formalize the various options for these two decisions in terms of their corresponding DP flavors. In Sections 4 and 5, we show why these two decisions are important by describing the consequences of each option on both data utility and privacy.

We need the following technical definition:

Definition 3.1 (Coarsening, Refinement, Coarser, Finer)**.** A data multiverse $\mathscr{D} \subset 2^{\mathcal{X}}$ is a *coarsening* of another multiverse $\mathscr{D}' \subset 2^{\mathcal{X}}$ from the same protection domain $\mathcal{X}$ if, for all $\mathcal{D}' \in \mathscr{D}'$ there exists $\mathcal{D} \in \mathscr{D}$ with $\mathcal{D}' \subset \mathcal{D}$.

When $\mathscr{D}$ is a coarsening of $\mathscr{D}'$, we write $\mathscr{D}' \leq \mathscr{D}$ and say that $\mathscr{D}'$ is a *refinement* of $\mathscr{D}$, or that $\mathscr{D}'$ is *finer* than $\mathscr{D}$, or that $\mathscr{D}$ is *coarser* than $\mathscr{D}'$.

If $\mathscr{D}$ is a coarsening of $\mathscr{D}'$ then there are more comparable datasets in $\mathscr{D}$ than in $\mathscr{D}'$ and hence $(\mathcal{X}, \mathscr{D}, d_{\mathcal{X}}, D_{\mathrm{Pr}})$ is strictly stronger than $(\mathcal{X}, \mathscr{D}', d_{\mathcal{X}}, D_{\mathrm{Pr}})$.

Let $\mathcal{X}_{\mathrm{pop}}$ be the set of all possible populations; $\mathcal{X}_{\mathrm{fr}}$ the set of all possible frames (from all possible populations); $\mathcal{X}_{\mathrm{samp}}$ the set of all possible target samples (from all possible frames); and $\mathcal{X}_{\mathrm{resp}}$ the set of all possible responding samples (from all possible target samples). We say that a DP flavor $(\mathcal{X}, \mathscr{D}, d_{\mathcal{X}}, D_{\mathrm{Pr}})$ is *population-level* if $\mathcal{X} = \mathcal{X}_{\mathrm{pop}}$. The definitions of *frame-level*, *(target-)sample-level* and *responding-sample-level* DP flavors are analogous.

Definition 3.2 (Primitive data multiverses)**.** Define the primitive data multiverses:

1. $\mathscr{D}_{\mathrm{fr|pop}} = \{\mathcal{D}_{\boldsymbol{x}} : \boldsymbol{x} \in \mathcal{X}_{\mathrm{pop}}\}$, where $\mathcal{D}_{\boldsymbol{x}}$ is the set of all possible frames drawn from the population $\boldsymbol{x} \in \mathcal{X}_{\mathrm{pop}}$;
2. $\mathscr{D}_{\mathrm{samp|pop}} = \{\mathcal{D}_{\boldsymbol{x}} : \boldsymbol{x} \in \mathcal{X}_{\mathrm{pop}}\}$, where $\mathcal{D}_{\boldsymbol{x}}$ is the set of all possible target samples drawn from the population $\boldsymbol{x} \in \mathcal{X}_{\mathrm{pop}}$.
3. $\mathscr{D}_{\mathrm{samp|fr}} = \{\mathcal{D}_{\boldsymbol{x}} : \boldsymbol{x} \in \mathcal{X}_{\mathrm{fr}}\}$, where $\mathcal{D}_{\boldsymbol{x}}$ is the set of all possible target samples drawn from the frame $\boldsymbol{x} \in \mathcal{X}_{\mathrm{fr}}$.
4. The data multiverses $\mathscr{D}_{\mathrm{resp|pop}}, \mathscr{D}_{\mathrm{resp|fr}}$ and $\mathscr{D}_{\mathrm{resp|samp}}$ can be defined analogously, as the set of data universes $\mathcal{D}_{\boldsymbol{x}}$, with $\mathcal{D}_{\boldsymbol{x}}$ the set of all possible responding samples drawn from, respectively, the population, frame, or target sample $\boldsymbol{x}$.

Definition 3.3 (Population-, frame- and sample-invariance)**.** A DP flavor $(\mathcal{X}, \mathscr{D}, d_{\mathcal{X}}, D_{\mathrm{Pr}})$ is:

1. *population-invariant* if $\mathscr{D} \leq \overline{\mathscr{D}}_{\cdot|\mathrm{pop}}$, where
   - for frame-level flavors: $\mathscr{D}_{\cdot|\mathrm{pop}} = \mathscr{D}_{\mathrm{fr|pop}}$,
   - for sample-level flavors: $\mathscr{D}_{\cdot|\mathrm{pop}} = \mathscr{D}_{\mathrm{samp|pop}}$, and
   - for responding-sample-level flavors: $\mathscr{D}_{\cdot|\mathrm{pop}} = \mathscr{D}_{\mathrm{resp|pop}}$;
2. *frame-invariant* if $\mathscr{D} \leq \overline{\mathscr{D}}_{\cdot|\mathrm{fr}}$, where
   - for sample-level flavors: $\mathscr{D}_{\cdot|\mathrm{fr}} = \mathscr{D}_{\mathrm{samp|fr}}$, and

- for responding-sample-level flavors: $\mathscr{D}_{\cdot|\text{fr}} = \mathscr{D}_{\text{resp}|\text{fr}}$;

3. *sample-invariant* if $\mathscr{D} \leq \overline{\mathscr{D}}_{\text{resp}|\text{samp}}$ (for responding-sample-level flavors).

The intuition behind these definitions is very simple. The idea is to restrict the comparable datasets. Population-invariance means that comparable frames (or samples or responding samples) must be from the same population. (That is, a pair of frames are comparable only if they are drawn from the same population.) Analogously, frame-invariance means that comparable samples must be drawn from the same frame.

If a DP flavor is not population-invariant (resp. frame-invariant or sample-invariant), then we say it is *population-agnostic* (resp. *frame-agnostic* or *sample-agnostic*). Since we assume that the DP flavor is complete, frame-agnosticism implies that there are two comparable samples which are drawn from different frames.

Because invariance at one level implies invariance at previous data pipeline phases, we identify ten settings (which together exhaust the potential options for where the DP mechanism starts and which phases are taken as invariant): one setting for population-level flavors; two for frame-level flavors; three for sample-level flavors; and four for responding-sample-level flavors.

4 Utility considerations

In this section we consider the possible implications of the different DP flavors on the achievable level of accuracy of the noisy outcome given a desired level of privacy (expressed by fixing the privacy parameters). Two components are relevant when evaluating the accuracy for DP estimates from survey data: the privacy amplification effects from sampling, which imply that less noise needs to be infused to achieve a given privacy level and the increased sensitivity of the weighted estimator (where weights are included to account for the sampling design), which typically implies that more noise is required. We discuss the effects of the different flavors on both components in the following chapters. We emphasize again that we ignore all the extra complexities that arise if the additional weighting adjustments for nonresponse would also be taken into account when analyzing the responding sample.

4.1 Privacy Amplification via Sampling

Previous research has shown that simple sampling designs offer privacy amplification, that is, the privacy offered when running a DP algorithm on a random subset of the population is higher than if the same algorithm with the same privacy parameters is run on the full population. Balle et al. [2018] proof the following theorem for simple random sampling with replacement (they also obtain similar results for Poisson sampling and simple random sampling without replacement):

Theorem 4.1. *Balle et al. [2018] Let $\mathcal{C}$ be a sampling scheme that uniformly randomly samples n values out of N possible values without replacement. Given an (ε, δ)-bounded differentially private mechanism $\mathcal{M}$, we have that $\mathcal{M}_{\mathcal{C}}$ is (ε', δ')-bounded differentially private for $\varepsilon' = \log(1 + \frac{n}{N}(e^{\varepsilon} - 1))$ and $\delta' = \frac{n}{N}\delta$.*

In this theorem bounded differential privacy refers to the scenario in which neighboring datasets are obtained by changing the values of one record in the data while keeping the size of the data fixed. Note that for small ε and small sampling rates this implies that $\varepsilon' \approx n/N\varepsilon$, i.e., the amplification is proportional to the sampling rate. Based on these results Bun et al. [2022] studied to what extent privacy amplification can also be achieved for the more complex sampling designs commonly used at statistical agencies. Their findings can be summarized as follows:

- Cluster sampling using simple random sampling without replacement to draw the clusters offers negligible amplification in practice except for small ε (less than 0.5) and very small cluster sizes (less than 15 units).
- With minor adjustments, stratified sampling using proportional allocation can provide privacy amplification.
- Data dependent allocation functions such as Neyman allocation for stratified sampling will likely result in privacy degredation (the effects will depend on the sensitivity of the allocation function).
- With pps sampling at the individual level, the privacy amplification will linearly depend on the maximum probability of inclusion (for small ε).
- Systematic sampling will only offer amplification if the ordering of the population is truly random. In all other cases, systematic sampling will suffer from the same effects as cluster sampling leading to no amplification (assuming the ordering is known to the attacker).

In practice this implies that for the multi-stage sampling designs that typically start with (multiple stages of) stratified cluster sampling, amplification effects can generally only be expected from those stages at which individual units are selected (typically the last stage of selection).

4.2 Privacy Amplification for Different DP Flavors

Before discussing the implications of the DP flavors introduced in Section 3, it is important to consider at which stages of the data production pipeline amplification effects could occur. Conceptually, three different sampling steps can be defined when moving from the population to the responding sample. The most obvious step (and the only one that is fully controlled) is the selection of the target sample from the sampling frame. However, if nonresponse is treated as a stochastic process (as is commonly done in the survey literature), moving from the target to the responding sample

can be interpreted as another sampling step. The same is true when moving from the population to the sampling frame if we assume that each unit in the population has a certain probability to be included in the frame. Still, the amplification effects of these two steps are difficult to take into account in practice as the inclusion probabilities are unknown and would need to be estimated. Errors when modeling these probabilities would lead to invalid statements regarding the amplification effects. Besides, the amplification effects when moving from the population to the sampling frame will typically be negligible given that the probability to be included in the frame should be well above 90% for high quality frames.

Considering the DP flavors, we can distinguish three scenarios: If the target sample is treated as invariant, the DP mechanism can only be applied at the responding sample level (as there are no neighboring frames to compare at the target sample level). This scenario boils down to the standard setting considered in most DP papers. There is no (sub)sampling step within the data release mechanism T and thus there is no amplification effect. Interestingly, this scenario offers the same privacy guarantees as the more restrictive assumption that the attacker knows who participated in the survey. In all other scenarios, privacy is amplified through the (sub)sampling process. In the second scenario, the DP flavor is frame-invariant but target sample-agnostic. In this scenario, amplification can only arise from the subsampling step when moving from the target sample to the responding sample. As response rates are often less than 20% in practice, this subsampling might offer some privacy amplification. However, as mentioned earlier, quantifying this effect will be difficult in practice as response probabilities are unknown and will likely differ between the units. In the third scenario, only the population is considered invariant. This scenario will offer privacy amplification as discussed in Bun et al. [2022] in addition to the theoretical amplification offered from nonresponse. Finally, if the DP flavor is agnostic at all levels, a third layer of amplification is possible by moving from the population to the sampling frame. As discussed above, this layer will typically be negligible for sampling frames commonly used in practice.

4.3 Weighting

Using weighted estimators generally increases the amount of noise that needs to be added to achieve a desired level of privacy protection. This is because the sensitivity of the result, i.e., the maximum possible change in the result when changing a single record, increases when incorporating the survey weights. To illustrate, we can consider the simple example of a counting query. A counting query simply counts the number of units in a database that satisfy a given set of conditions, for example, the total number of unemployed men between 30 and 40. Counting queries are attractive under DP as they have low sensitivity and thus require limited amounts of noise to achieve DP (as the noise scales with the sensitivity of the query). Under unbounded DP (i.e., defining neighboring datasets by adding or removing one record) the sensitivity of a counting query is 1.

In the survey literature a counting query is called a total and the most convenient way to estimate a total for complex sampling designs is to use the Horvitz-Thompson estimator [Horvitz and Thompson, 1952], which provides approximately unbiased estimates for most sampling designs. The Horvitz-Thompson estimator for a total is given as $\hat{t}_x = \sum_{i=1}^{n} w_i x_i$, where $\hat{t}_x$ is the estimated total in the population for the target variable x, n is the sample size and w_i is survey weight for unit i. In our example, x_i is a binary indicator which equals 1 if unit i satisfies the conditions of interest (i.e. unit i is unemployed, male, and between 40 and 50 years old) and is zero otherwise. Using the Horvitz-Thompson estimator, the sensitivity increases to $\max(w_i)$ (where the maximum is taken either over the records in the sample (under target sample invariance), the frame (under frame invariance) or over the entire population (under population invariance)).

Since the amount of noise that is required typically scales with the sensitivity of the output, this implies that much more noise needs to be added when trying to protect a weighted survey estimate. However, the considerations so far assume that the weights can be considered fixed. This assumption is only justified, if changing one record in the database does not change the probability of inclusion for any of the records in the population. Whether this is a realistic assumption will depend on the DP flavor to be considered but also on the properties of the sampling design.

In general, the design weights can only be treated as fixed under the frame-invariant or target sample invariante scenario. In all other scenarios the weights will typically change. How much the weights will change will typically depend on the sensitivity of the sampling design, which in turn depends on the question, how data dependent the sampling design is. To illustrate, data dependence will be small for single stage cluster sampling designs especially if the clusters are selected using simple random sampling (such a design is used for example for the German Microcensus). For such a design, the probability of selection does not change over neighboring frames (as long as the definition of the clusters does not change). This would still hold if this design would be extended to a multi-stage design and simple random sampling would be used on all of the following stages. On the other hand, PPS sampling will generally be highly data dependent as the probability of selection directly depends on some features of the data. This will be less problematic if PPS sampling is used to select the clusters as the probability of selection will only depend on the size of the clusters and these sizes will only change by one record over neighboring databases. However, if PPS sampling is used to select individual units, the probabilities of selection can change arbitrarily over neighboring datasets. Thus, for these designs the sensitivity of the final estimate might increase considerably and it seems difficult to correctly quantify this sensitivity in practice.

Table 4.1 summarizes the implications of the different DP flavors considered in this paper. It highlights the inherent trade-off between the various flavors. For example, considering the frame as invariant implies that no utility improvements through amplification can be achieved. On the other hand, treating the weights as fixed will generally reduce the sensitivity of the final estimates and thus

DP Setting	Amplification from	Effects on Weights
Target sample invariance	No	Can be treated as fixed
Frame Invariance	NR	Can be treated as fixed
Population Invariance	NR&S	Sensitivity needs to be considered
Population agnostic	FR&S&NR	Sensitivity needs to be considered

Table 4.1: Overview of the implications of various DP flavors (The abbreviations for the privacy amplification are NR=nonresponse,S=sampling,FR=frame).

the noise that needs to be added to ensure privacy. For the other flavors utility improvements could be achieved through privacy amplification, but this benefit comes at the cost that the sensitivity of the weights needs to be considered. This increase might outweigh the benefits of amplification from sampling, especially since as Bun et al. [2022] have shown, the amplification effects tend to be small for sampling designs commonly used in practice. Which of the flavors will be most attractive from a utility perspective will crucially depend on the sampling design in practice as the design will have an effect on both, the amplification and the sensitivity of the weights. It will also depend on the question whether response probabilities can be determined reliably.

5 Privacy Considerations

Note to the conference organizers and paper reviewers: This section is currently in preparation. In its place, we are submitting an outline of the main discussion points for the section:

- What are the privacy semantics associated with the 10 settings (population-, frame- and sample-level; and population-, frame- and sample-invariance)?
  - Explain how the protection domain $\mathcal{X}$ serves as the parameter space for the attacker's inferential model.
  - Link the choice of $\mathcal{X}$ and $\mathscr{D}$ to the various attacker models in traditional statistical disclosure control (SDC):
    - the 'nosy neighbor', who knows the identity of a single individual in the sample (but not the rest of the sample);
    - the 'investigative journalist', who does not know the identities of anyone in the sample;
    - the 'infiltrator': who knows the identities of everyone in the sample.
    - How do the semantics change under each of these models and for different choices of $\mathcal{X}$ and $\mathscr{D}$?

– Using $D_{\mathrm{Pr}} = \textsc{Mult}$ (for simplicity), provide a couple of concrete examples of privacy semantics in terms of prior-to-posterior learning:

* When $\mathcal{X}$ is sample-level, the Bayesian attacker must assume that the unit of interest i is in the sample in order to recover the traditional prior-to-posterior bound. Conditioning on the unit of interest i not being in the sample, the prior-to-posterior is vacuously 1. Hence

$$\begin{aligned}\pi(x_i \mid \boldsymbol{x}_{-i}, T) &= \pi(x_i \mid \boldsymbol{x}_{-i}, T, i \in S)\pi(i \in S) + \pi(x_i \mid \boldsymbol{x}_{-i}, T, i \notin S)\pi(i \notin S) \\ &\leq e^{\varepsilon}\pi(x_i \mid \boldsymbol{x}_{-i}, i \in S)\pi(i \in S) + \pi(x_i \mid \boldsymbol{x}_{-i}, i \notin S)\pi(i \notin S) \\ &< e^{\varepsilon}\pi(x_i \mid \boldsymbol{x}_{-i}).\end{aligned}$$

* In contrast, when $\mathcal{X}$ is population-level, the standard prior-to-posterior results hold: $\pi(x_i \mid \boldsymbol{x}_{-i}, T) \leq e^{\varepsilon}\pi(x_i \mid \boldsymbol{x}_{-i})$.

* This describes a phenomenon which is the inverse of privacy amplification by subsampling. The Bayesian attacker can have strictly less learning for sample-level DP than for population-level DP.

- What are the privacy properties associated with the 10 settings? In particular, how does composition work, when the two data-release mechanisms satisfy DP flavors with different $\mathcal{X}$ and/or different $\mathscr{D}$?

  – This question is important as an NSO considers quantifying the privacy leakage globally, across their multiple data collections.

  – In order to describe composition, is it necessary to pull $\mathcal{X}$ back to the population-level? (What would happen to the data multiverse?) (By the definitions in Subsection 2.1, it is sufficient to pull $\mathcal{X}$ back to the population level.)

- The impact of nuisance parameters in privacy?

  – When the protection object $\boldsymbol{x}$ has a higher dimension (such as when $\boldsymbol{x}$ is a population, instead of a sample), then there are more nuisance parameters $\boldsymbol{x}_{-i}$ for the attacker to account for.

  – Plausibly, having more nuisance parameters should increase the amount of privacy provided to $\boldsymbol{x}_i$, since nuisance parameters make inference harder.

  – Example: It is easy to determine x_i when the attacker is differentiating between a true count of 1 versus a true count of 2. It is much harder to determine x_i when differentiating between a true count of 100,000 versus a true count of 100,001.

- This nuance is not captured within the DP framework. The measure of privacy loss does not depend on the dimension of the dataset $\boldsymbol{x}$ being protected.
- However, this is an important consideration when determining if the DP flavor should be at the population-level versus the sample-level. When the DP flavor is at the population-level, there are more nuisance parameters, since $\boldsymbol{x}$ is larger. Is there some way to augment the DP framework to account for this?

6 Discussion

This paper develops theory for understanding and implementing differential privacy in the context of survey statistics. By recognizing the major phases in the survey-data pipeline, we identified ten different settings of DP. These settings corresponding to different choices for 1) where the DP data-release mechanism starts in this pipeline; and for 2) which of the previous phases are taken as invariant. Section 3 formalized these ten settings into ten different conditions on the DP flavor.

Sections 4 and 5 show that the choice of the setting makes significant impacts in terms of both privacy and utility. Therefore, while DP is invariant to post-processing, pre-processing steps matter. Moreover, the data custodian must necessarily choose a setting – they cannot implement DP without first deciding (perhaps implicitly) where the DP mechanism starts and which pre-processing steps are taken as invariant. Hence, contrary to commonly-held beliefs, DP does make important assumptions on the data and on the attacker, because the data custodian's decision impacts both the utility and privacy semantics of the DP-outputted data.

Acknowledgements

JB gratefully acknowledges partial financial support from the Australian-American Fulbright Commission and the Kinghorn Foundation.

References

J. M. Abowd and M. B. Hawes. Confidentiality protection in the 2020 us census of population and housing. *Annual Review of Statistics and Its Application*, 10:119–144, 2023.

Apple's Differential Privacy Team. Learning with privacy at scale. *Apple Machine Learning Journal*, 1(8), 2017.

J. Bailie, R. Gong, and X.-L. Meng. Can swapping be differentially private? A refreshment stirred, not shaken. *In preparation*, 2023+a.

J. Bailie, R. Gong, and X.-L. Meng. A statistical appreciation and assessment of differential privacy. *in preparation*, 2023+b.

J. Bailie, R. Gong, and X.-L. Meng. On the uniqueness of differential privacy. *in preparation*, 2023+c.

B. Balle, G. Barthe, and M. Gaboardi. Privacy amplification by subsampling: Tight analyses via couplings and divergences. *Advances in neural information processing systems*, 31, 2018.

B. Balle, G. Barthe, and M. Gaboardi. Privacy profiles and amplification by subsampling. *Journal of Privacy and Confidentiality*, 10(1), Jan. 2020. ISSN 2575-8527. doi: 10.29012/jpc.726.

M. Bun, J. Drechsler, M. Gaboardi, A. McMillan, and J. Sarathy. Controlling privacy loss in sampling schemes: An analysis of stratified and cluster sampling. In *Foundations of Responsible Computing (FORC 2022)*, page 24, June 2022.

P. Cuff and L. Yu. Differential privacy as a mutual information constraint. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, pages 43–54, 2016.

S. Das, J. Drechsler, K. Merrill, and S. Merrill. Imputation under differential privacy. http://arxiv.org/abs/2206.15063, July 2022.

B. Ding, J. Kulkarni, and S. Yekhanin. Collecting telemetry data privately. In *Advances in Neural Information Processing Systems*, pages 3571–3580, 2017.

J. Drechsler. Differential privacy for government agencies—Are we there yet? *Journal of the American Statistical Association*, 118(541):761–773, Jan. 2023. ISSN 0162-1459. doi: 10.1080/01621459.2022.2161385. https://doi.org/10.1080/01621459.2022.2161385.

C. Dwork and A. Roth. The algorithmic foundations of differential privacy. *Foundations and Trends® in Theoretical Computer Science*, 9(3–4):211–407, 2014.

C. Dwork, K. Kenthapadi, F. McSherry, I. Mironov, and M. Naor. Our data, ourselves: Privacy via distributed noise generation. In S. Vaudenay, editor, *Advances in Cryptology - EUROCRYPT 2006*, Lecture Notes in Computer Science, pages 486–503, Berlin, Heidelberg, 2006a. Springer. ISBN 978-3-540-34547-3. doi: 10.1007/11761679_29.

C. Dwork, F. McSherry, K. Nissim, and A. Smith. Calibrating noise to sensitivity in private data analysis. In *Theory of cryptography conference*, pages 265–284. Springer, 2006b.

Ú. Erlingsson, V. Pihur, and A. Korolova. Rappor: Randomized aggregatable privacy-preserving ordinal response. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, pages 1054–1067, 2014.

A. D. Foote, A. Machanavajjhala, and K. McKinney. Releasing earnings distributions using differential privacy: Disclosure avoidance system for post-secondary employment outcomes (pseo). *Journal of Privacy and Confidentiality*, 9(2), 2019. doi: 10.29012/jpc.722. URL https://journalprivacyconfidentiality.org/index.php/jpc/article/view/722.

D. G. Horvitz and D. J. Thompson. A generalization of sampling without replacement from a finite universe. *Journal of the American statistical Association*, 47(260):663–685, 1952.

D. Kifer and A. Machanavajjhala. No free lunch in data privacy. In *Proceedings of the 2011 International Conference on Management of Data - SIGMOD '11*, pages 193–204, Athens, Greece, 2011. ACM Press. ISBN 978-1-4503-0661-4. doi: 10.1145/1989323.1989345.

S. Lin, M. Bun, M. Gaboardi, E. D. Kolaczyk, and A. Smith. Differentially private confidence intervals for proportions under stratified random sampling. *arXiv preprint arXiv:2301.08324*, 2023.

A. Machanavajjhala, D. Kifer, J. Abowd, J. Gehrke, and L. Vilhuber. Privacy: Theory meets practice on the map. In *Proceedings of the 2008 IEEE 24th International Conference on Data Engineering*, pages 277–286. IEEE Computer Society, 2008.

S. Messing, C. DeGregorio, B. Hillenbrand, G. King, S. Mahanti, Z. Mukerjee, C. Nayak, N. Persily, B. State, and A. Wilkins. Facebook Privacy-Protected Full URLs Data Set. https://doi.org/10.7910/DVN/TDOAPG, 2020.

J. P. Reiter. Differential privacy and federal data releases. *Annual review of statistics and its application*, 6:85–101, 2019.

M. C. Tschantz, S. Sen, and A. Datta. SoK: Differential privacy as a causal property. In *2020 IEEE Symposium on Security and Privacy (SP)*, pages 354–371. IEEE, 2020.

Uber Security. Uber releases open source project for differential privacy. https://medium.com/uber-security-privacy/differential-privacy-open-source-7892c82c42b6, 2017.

U.S. Bureau of Labor Statistics. Current Population Survey: Concepts. https://www.bls.gov/opub/hom/cps/concepts.htm, Apr. 2018.

US Census Bureau. Protecting the confidentiality of America's statistics: Adopting modern disclosure avoidance methods at the Census Bureau. https://www.census.gov/newsroom/blogs/research-matters/2018/08/protecting_the_confi.html, Aug. 2018.

U.S. Census Bureau. Design and methodology: Current Population Survey. *Technical Paper 77*, 2019.

US Census Bureau. Disclosure avoidance protections for the American Community Survey. https://www.census.gov/newsroom/blogs/random-samplings/2022/12/disclosure-avoidance-protections-acs.html, Dec. 2022.

U.S. Census Bureau. List of Surveys. https://www.census.gov/programs-surveys/surveyhelp/list-of-surveys.html, 2023. Accessed: 12/12/2023.