Whose Data Is It Anyway?
A Formal Treatment of Differential Privacy for Surveys

James Bailie* & Jörg Drechsler†
*Chalmers University, †German Institute for Employment Research

Adelaide Data Privacy Workshop
26 November 2025

These Slides Are Available Online:

jameshbailie.github.io/talks/

This Presentation Is Based on Two Papers:

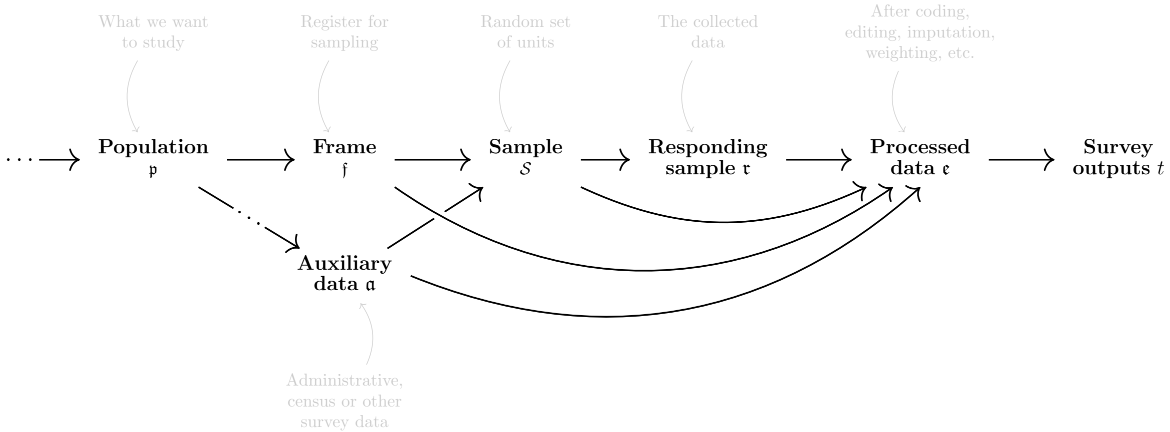
JB and Jörg Drechsler (2024). "Whose Data Is It Anyway? Towards a Formal Treatment of Differential Privacy for Surveys". *NBER Working Paper.*

Jörg Drechsler and JB (2024). "The Complexities of Differential Privacy for Survey Data". To appear in *Data Privacy Protection and the Conduct of Applied Research: Methods, Approaches and Their Consequences.*

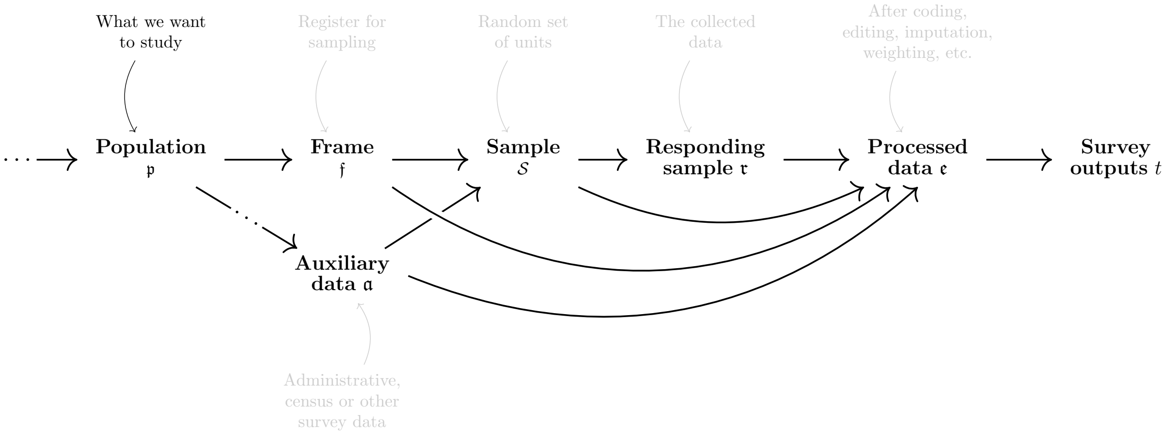
Motivation

- The US Census Bureau has committed to adopting *formal privacy* for all their data products (US Census Bureau 2022).
- Most of their collections are *surveys*.
- Yet the *"science ... does not yet exist"* for a formally private solution to the American Community Survey (for example).
- In implementing differential privacy (DP), surveys come with their own set of *unique challenges and opportunities*.

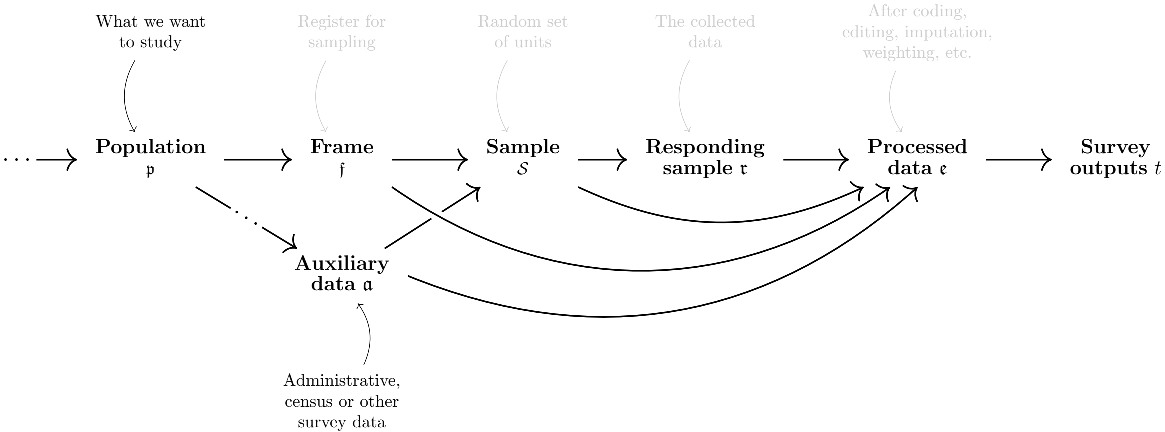
The Survey Data Pipeline



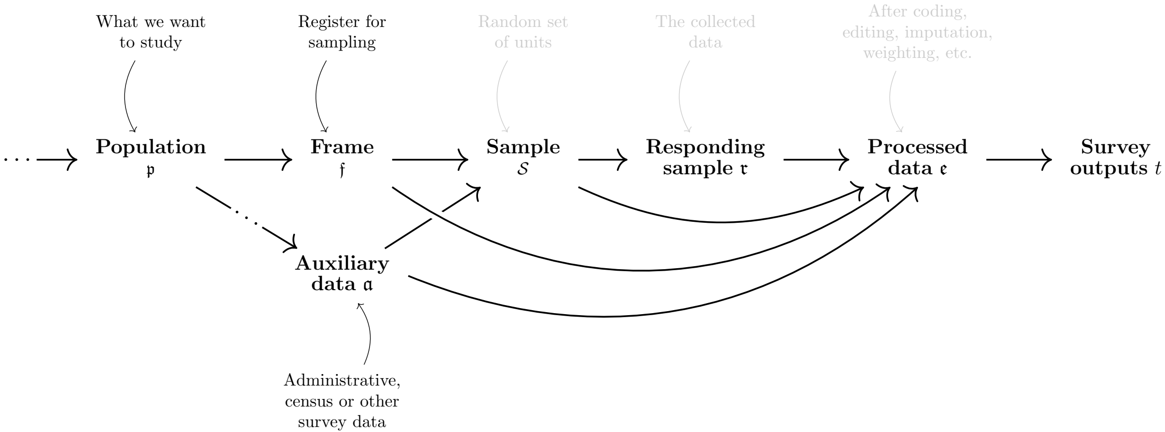
The Survey Data Pipeline



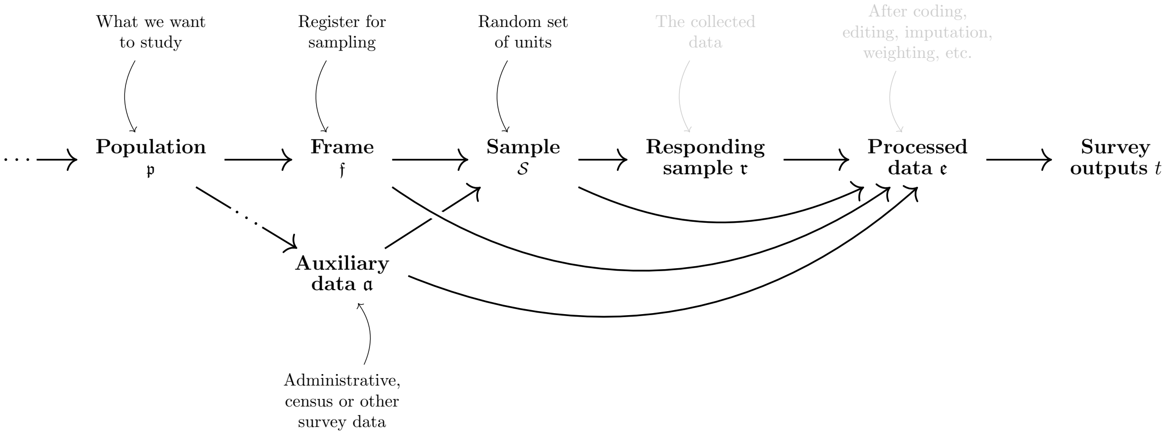
The Survey Data Pipeline



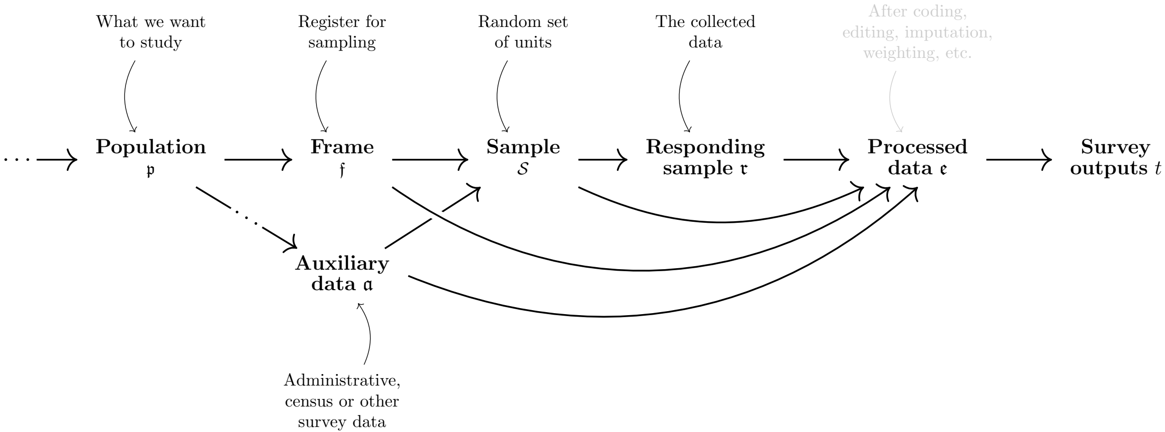
The Survey Data Pipeline



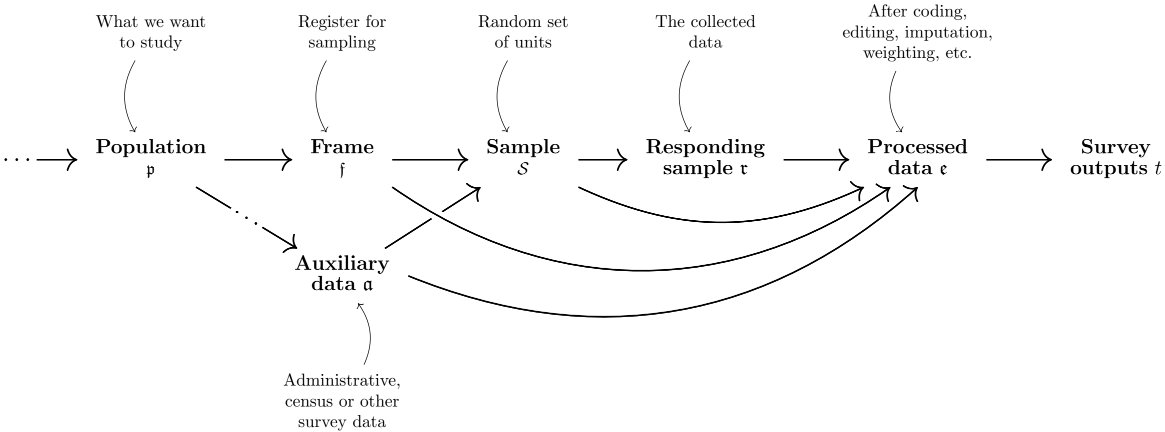
The Survey Data Pipeline



The Survey Data Pipeline



The Survey Data Pipeline



DP Settings for Surveys

$$\cdots \longrightarrow \mathfrak{p} \longrightarrow \mathfrak{f} \longrightarrow \mathfrak{s} \longrightarrow \mathfrak{r} \longrightarrow \mathfrak{e} \longrightarrow t$$

Two considerations

- Where does the DP mechanism *start* in the data pipeline? (What is $\mathcal{X}$?)
- Which of the previous steps are kept *invariant*? (What is $\mathscr{D}$?)

For example,

1)

$$\begin{tikzcd}
& \mathfrak{p} \arrow[r] & \mathfrak{f} \arrow[r] & \mathfrak{s} \arrow[r] & \mathfrak{r} \arrow[r] & \mathfrak{e} \arrow[r] & t \\
\cdots \arrow[ur] \arrow[dr] & & & & & & \\
& \mathfrak{p}' \arrow[r] & \mathfrak{f}' \arrow[r] & \mathfrak{s}' \arrow[r] & \mathfrak{r}' \arrow[r] & \mathfrak{e}' \arrow[r] & t'
\end{tikzcd}$$

2)

$$\begin{tikzcd}
& & & \mathfrak{s} \arrow[r] & \mathfrak{r} \arrow[r] & \mathfrak{e} \arrow[r] & t \\
\cdots \arrow[r] & \mathfrak{p} \arrow[r] & \mathfrak{f} \arrow[ur] \arrow[dr] & & & & \\
& & & \mathfrak{s}' \arrow[r] & \mathfrak{r}' \arrow[r] & \mathfrak{e}' \arrow[r] & t'
\end{tikzcd}$$

Why Does This Matter? One Example

Move $\mathcal{X}$ from the samples $\mathfrak{s}$ to the frames $\mathfrak{f}$ – i.e. start the data-release mechanism one step earlier.

Privacy amplification by sampling

If $T(\mathfrak{s})$ is ε-DP and $\mathcal{S}(\mathfrak{f})$ randomly samples f fraction of the frame $\mathfrak{f}$, then $T' = T \circ \mathcal{S}$ is ε'-DP where $\varepsilon' \approx f\varepsilon$. (Balle et al. 2020)

Take-away: "Privacy for free" – if the sampling procedure is included, less noise is required to achieve the same privacy budget.

Why Does This Matter? One Example

Move $\mathcal{X}$ from the samples $\mathfrak{s}$ to the frames $\mathfrak{f}$ – i.e. start the data-release mechanism one step earlier.

Privacy amplification by sampling

If $T(\mathfrak{s})$ is ε-DP and $\mathcal{S}(\mathfrak{f})$ randomly samples f fraction of the frame $\mathfrak{f}$, then $T' = T \circ \mathcal{S}$ is ε'-DP where $\varepsilon' \approx f\varepsilon$. (Balle et al. 2020)

Take-away: "Privacy for free" – if the sampling procedure is included, less noise is required to achieve the same privacy budget.

Why Does This Matter? One Example

Move $\mathcal{X}$ from the samples $\mathfrak{s}$ to the frames $\mathfrak{f}$ – i.e. start the data-release mechanism one step earlier.

Privacy amplification by sampling

If $T(\mathfrak{s})$ is ε-DP and $\mathcal{S}(\mathfrak{f})$ randomly samples f fraction of the frame $\mathfrak{f}$, then $T' = T \circ \mathcal{S}$ is ε'-DP where $\varepsilon' \approx f\varepsilon$. (Balle et al. 2020)

Take-away: "Privacy for free" – if the sampling procedure is included, less noise is required to achieve the same privacy budget.

Implications for Privacy Semantics and Composition

- Some fundamental results in differential privacy:
  1. *Privacy semantics*: how does DP protect your data from any possible attacker?
  2. *Composition*: how does ε grow as you make more releases?
- Challenges to these results in the survey context (and beyond)

Implications for Privacy Semantics and Composition

- Some fundamental results in differential privacy:
  1. *Privacy semantics*: how does DP protect your data from any possible attacker?
  2. *Composition*: how does ε grow as you make more releases?
- Challenges to these results in the survey context (and beyond)

Implications for Privacy Semantics and Composition

- Some fundamental results in differential privacy:
    1. *Privacy semantics*: how does DP protect your data from any possible attacker?
    2. *Composition*: how does ε grow as you make more releases?
- Challenges to these results in the survey context (and beyond)

Implications for Privacy Semantics and Composition

- Some fundamental results in differential privacy:
    1. *Privacy semantics*: how does DP protect your data from any possible attacker?
    2. *Composition*: how does ε grow as you make more releases?
- Challenges to these results in the survey context (and beyond)

Privacy Semantics

- DP protects against any attacker, *regardless of their auxiliary knowledge.*
- How to formalise this? Model the attacker as a Bayesian agent with prior π.
- Suppose the attacker wants to learn a record x_i.
- Pure ε-DP guarantees that the attacker's prior-to-posterior ratio is bounded by e^{ε}:

$$e^{-\varepsilon} \leq \frac{\pi(x_i \mid T, \boldsymbol{x}_{-i})}{\pi(x_i \mid \boldsymbol{x}_{-i})} \leq e^{\varepsilon}.$$

- There is also posterior-to-posterior semantics, which compare the attacker's posterior to the counterfactual where i didn't contribute their data.

Privacy Semantics

- DP protects against any attacker, *regardless of their auxiliary knowledge.*
- How to formalise this? Model the attacker as a Bayesian agent with prior π.
- Suppose the attacker wants to learn a record x_i.
- Pure ε-DP guarantees that the attacker's prior-to-posterior ratio is bounded by e^{ε}:

$$e^{-\varepsilon} \leq \frac{\pi(x_i \mid T, \boldsymbol{x}_{-i})}{\pi(x_i \mid \boldsymbol{x}_{-i})} \leq e^{\varepsilon}.$$

- There is also posterior-to-posterior semantics, which compare the attacker's posterior to the counterfactual where i didn't contribute their data.

Privacy Semantics

- DP protects against any attacker, *regardless of their auxiliary knowledge.*
- How to formalise this? Model the attacker as a Bayesian agent with prior π.
- Suppose the attacker wants to learn a record x_i.
- Pure ε-DP guarantees that the attacker's prior-to-posterior ratio is bounded by e^{ε}:

$$e^{-\varepsilon} \leq \frac{\pi(x_i \mid T, \boldsymbol{x}_{-i})}{\pi(x_i \mid \boldsymbol{x}_{-i})} \leq e^{\varepsilon}.$$

- There is also posterior-to-posterior semantics, which compare the attacker's posterior to the counterfactual where i didn't contribute their data.

Privacy Semantics

- DP protects against any attacker, *regardless of their auxiliary knowledge.*
- How to formalise this? Model the attacker as a Bayesian agent with prior π.
- Suppose the attacker wants to learn a record x_i.
- Pure ε-DP guarantees that the attacker's prior-to-posterior ratio is bounded by e^{ε}:

$$e^{-\varepsilon} \leq \frac{\pi(x_i \mid T, \boldsymbol{x}_{-i})}{\pi(x_i \mid \boldsymbol{x}_{-i})} \leq e^{\varepsilon}.$$

- There is also posterior-to-posterior semantics, which compare the attacker's posterior to the counterfactual where i didn't contribute their data.

Privacy Semantics

- DP protects against any attacker, *regardless of their auxiliary knowledge.*
- How to formalise this? Model the attacker as a Bayesian agent with prior π.
- Suppose the attacker wants to learn a record x_i.
- Pure ε-DP guarantees that the attacker's prior-to-posterior ratio is bounded by e^{ε}:

$$e^{-\varepsilon} \leq \frac{\pi(x_i \mid T, \boldsymbol{x}_{-i})}{\pi(x_i \mid \boldsymbol{x}_{-i})} \leq e^{\varepsilon}.$$

- There is also posterior-to-posterior semantics, which compare the attacker's posterior to the counterfactual where i didn't contribute their data.

Privacy Semantics

Posterior-to-posterior privacy semantics

What would an attacker learn about a single record if it is included in the input dataset, relative to a counterfactual world in which it is not included?

- If T is ε-DP, then the posterior-to-posterior ratio is in $[e^{-\varepsilon}, e^{\varepsilon}]$. (Kifer et al. 2022)
- What record (in what input dataset) is being protected depends on where T *starts* in the data pipeline; and what counterfactual worlds are possible depends on what steps are *invariant*.

Privacy Semantics

Posterior-to-posterior privacy semantics

What would an attacker learn about a single record if it is included in the input dataset, relative to a counterfactual world in which it is not included?

- If T is ε-DP, then the posterior-to-posterior ratio is in $[e^{-\varepsilon}, e^{\varepsilon}]$. (Kifer et al. 2022)
- What record (in what input dataset) is being protected depends on where T *starts* in the data pipeline; and what counterfactual worlds are possible depends on what steps are *invariant*.

Privacy Semantics

Posterior-to-posterior privacy semantics

What would an attacker learn about a single record if it is included in the input dataset, relative to a counterfactual world in which it is not included?

- If T is ε-DP, then the posterior-to-posterior ratio is in $[e^{-\varepsilon}, e^{\varepsilon}]$. (Kifer et al. 2022)
- What record (in what input dataset) is being protected depends on where T *starts* in the data pipeline; and what counterfactual worlds are possible depends on what steps are *invariant*.

Privacy Semantics

- Suppose $T(\mathfrak{s})$ is ε-DP and $\mathcal{S}(\mathfrak{f})$ randomly samples f fraction of $\mathfrak{f}$.
- $T' = T \circ \mathcal{S}$ is ε'-DP with $\varepsilon' \approx f\varepsilon < \varepsilon$.
- So the posterior-to-posterior ratio of T' should be in the interval $[e^{-\varepsilon'}, e^{\varepsilon'}]$.

Traditional statistical disclosure control attacker models

- The *nosy neighbor:* Knows that a record is in the sample.
- The *journalist:* Wants to learn about *any* record, so picks one in the sample.

- For these attackers, the posterior-to-posterior ratio (or prior-to-posterior ratio) of T' is in the interval $[e^{-\varepsilon}, e^{\varepsilon}]$, *not* the interval $[e^{-\varepsilon'}, e^{\varepsilon'}]$.

Privacy Semantics

- Suppose $T(\mathfrak{s})$ is ε-DP and $\mathcal{S}(\mathfrak{f})$ randomly samples f fraction of $\mathfrak{f}$.
- $T' = T \circ \mathcal{S}$ is ε'-DP with $\varepsilon' \approx f\varepsilon < \varepsilon$.
- So the posterior-to-posterior ratio of T' should be in the interval $[e^{-\varepsilon'}, e^{\varepsilon'}]$.

Traditional statistical disclosure control attacker models

- The *nosy neighbor:* Knows that a record is in the sample.
- The *journalist:* Wants to learn about *any* record, so picks one in the sample.

- For these attackers, the posterior-to-posterior ratio (or prior-to-posterior ratio) of T' is in the interval $[e^{-\varepsilon}, e^{\varepsilon}]$, *not* the interval $[e^{-\varepsilon'}, e^{\varepsilon'}]$.

Privacy Semantics

- Suppose $T(\mathfrak{s})$ is ε-DP and $\mathcal{S}(\mathfrak{f})$ randomly samples f fraction of $\mathfrak{f}$.
- $T' = T \circ \mathcal{S}$ is ε'-DP with $\varepsilon' \approx f\varepsilon < \varepsilon$.
- So the posterior-to-posterior ratio of T' should be in the interval $[e^{-\varepsilon'}, e^{\varepsilon'}]$.

Traditional statistical disclosure control attacker models

- The *nosy neighbor:* Knows that a record is in the sample.
- The *journalist:* Wants to learn about *any* record, so picks one in the sample.

- For these attackers, the posterior-to-posterior ratio (or prior-to-posterior ratio) of T' is in the interval $[e^{-\varepsilon}, e^{\varepsilon}]$, *not* the interval $[e^{-\varepsilon'}, e^{\varepsilon'}]$.

Privacy Semantics

- For these attackers, the posterior-to-posterior ratio (or prior-to-posterior ratio) of T' is in the interval $[e^{-\varepsilon}, e^{\varepsilon}]$, *not* the interval $[e^{-\varepsilon'}, e^{\varepsilon'}]$.
- So DP does *not* provide the nominal protection against an attacker with arbitrary side knowledge.
- This also applies wherever sampling is used for privacy.
  - For example, in DP stochastic gradient descent which is used for privately training neural networks.

Privacy Semantics

- For these attackers, the posterior-to-posterior ratio (or prior-to-posterior ratio) of T' is in the interval $[e^{-\varepsilon}, e^{\varepsilon}]$, *not* the interval $[e^{-\varepsilon'}, e^{\varepsilon'}]$.
- So DP does *not* provide the nominal protection against an attacker with arbitrary side knowledge.
- This also applies wherever sampling is used for privacy.
  - For example, in DP stochastic gradient descent which is used for privately training neural networks.

Privacy Semantics

- For these attackers, the posterior-to-posterior ratio (or prior-to-posterior ratio) of T' is in the interval $[e^{-\varepsilon}, e^{\varepsilon}]$, *not* the interval $[e^{-\varepsilon'}, e^{\varepsilon'}]$.
- So DP does *not* provide the nominal protection against an attacker with arbitrary side knowledge.
- This also applies wherever sampling is used for privacy.
    - For example, in DP stochastic gradient descent which is used for privately training neural networks.

Privacy Semantics

- For these attackers, the posterior-to-posterior ratio (or prior-to-posterior ratio) of T' is in the interval $[e^{-\varepsilon}, e^{\varepsilon}]$, *not* the interval $[e^{-\varepsilon'}, e^{\varepsilon'}]$.
- So DP does *not* provide the nominal protection against an attacker with arbitrary side knowledge.
- This also applies wherever sampling is used for privacy.
  - For example, in DP stochastic gradient descent which is used for privately training neural networks.

Composition

- How does ε grow as you make more releases?
- How to formalise this? Suppose you have two data-releases T_1 and T_2 which are both ε-DP. Then (T_1, T_2) is 2ε-DP.
- This assumes that the randomness in T_1 and T_2 are independent.

Composition

- How does ε grow as you make more releases?
- How to formalise this? Suppose you have two data-releases T_1 and T_2 which are both ε-DP. Then (T_1, T_2) is 2ε-DP.
- This assumes that the randomness in T_1 and T_2 are independent.

Composition

- How does ε grow as you make more releases?
- How to formalise this? Suppose you have two data-releases T_1 and T_2 which are both ε-DP. Then (T_1, T_2) is 2ε-DP.
- This assumes that the randomness in T_1 and T_2 are independent.

Composition

- Statistical agencies often use sample designs which are *dependent*.
    - For example, to reduce respondent burden.
- For $i \in \{1, 2\}$, suppose $T_i(\mathfrak{s})$ is ε-DP, and $T_i' = T_i \circ \mathcal{S}$.
- Privacy loss of the composition (T_1', T_2') is not the sum of T_1' and T_2''s privacy losses.
- More generally, this holds whenever the sampling in T_1 is dependent on the sampling in T_2.
- This will complicate global privacy loss calculations.

Composition

- Statistical agencies often use sample designs which are *dependent*.
  - For example, to reduce respondent burden.
- For $i \in \{1, 2\}$, suppose $T_i(\mathfrak{s})$ is ε-DP, and $T_i' = T_i \circ \mathcal{S}$.
- Privacy loss of the composition (T_1', T_2') is not the sum of T_1' and T_2''s privacy losses.
- More generally, this holds whenever the sampling in T_1 is dependent on the sampling in T_2.
- This will complicate global privacy loss calculations.

Composition

- Statistical agencies often use sample designs which are *dependent*.
  - For example, to reduce respondent burden.
- For $i \in \{1, 2\}$, suppose $T_i(\mathfrak{s})$ is ε-DP, and $T_i' = T_i \circ \mathcal{S}$.
- Privacy loss of the composition (T_1', T_2') is not the sum of T_1' and T_2''s privacy losses.
- More generally, this holds whenever the sampling in T_1 is dependent on the sampling in T_2.
- This will complicate global privacy loss calculations.

Composition

- Statistical agencies often use sample designs which are *dependent*.
  - For example, to reduce respondent burden.
- For $i \in \{1, 2\}$, suppose $T_i(\mathfrak{s})$ is ε-DP, and $T_i' = T_i \circ \mathcal{S}$.
- Privacy loss of the composition (T_1', T_2') is not the sum of T_1' and T_2''s privacy losses.
- More generally, this holds whenever the sampling in T_1 is dependent on the sampling in T_2.
- This will complicate global privacy loss calculations.

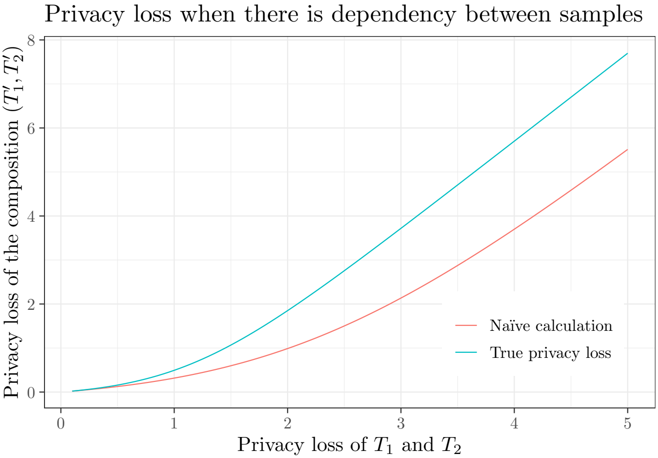
Composition

- Statistical agencies often use sample designs which are *dependent*.
  - For example, to reduce respondent burden.
- For $i \in \{1, 2\}$, suppose $T_i(\mathfrak{s})$ is ε-DP, and $T_i' = T_i \circ \mathcal{S}$.
- Privacy loss of the composition (T_1', T_2') is not the sum of T_1' and T_2''s privacy losses.
- More generally, this holds whenever the sampling in T_1 is dependent on the sampling in T_2.
- This will complicate global privacy loss calculations.

Composition

- Statistical agencies often use sample designs which are *dependent*.
  - For example, to reduce respondent burden.
- For $i \in \{1, 2\}$, suppose $T_i(\mathfrak{s})$ is ε-DP, and $T_i' = T_i \circ \mathcal{S}$.
- Privacy loss of the composition (T_1', T_2') is not the sum of T_1' and T_2''s privacy losses.
- More generally, this holds whenever the sampling in T_1 is dependent on the sampling in T_2.
- This will complicate global privacy loss calculations.

Composition



Utility Considerations (I)

Privacy amplification by sampling

If $T(\mathfrak{s})$ is ε-DP and $\mathcal{S}(\mathfrak{f})$ randomly samples f fraction of the frame $\mathfrak{f}$, then $T' = T \circ \mathcal{S}$ is ε'-DP where $\varepsilon' \approx f\varepsilon$. (Balle et al. 2020)

- *Take-away:* If the sampling procedure is included, less noise is required to achieve the same privacy budget.
- *But* there is little privacy amplification when $\mathcal{S}$ is a complex sampling design. (Bun et al. 2022)

Utility Considerations (I)

Privacy amplification by sampling

If $T(\mathfrak{s})$ is ε-DP and $\mathcal{S}(\mathfrak{f})$ randomly samples f fraction of the frame $\mathfrak{f}$, then $T' = T \circ \mathcal{S}$ is ε'-DP where $\varepsilon' \approx f\varepsilon$. (Balle et al. 2020)

- *Take-away:* If the sampling procedure is included, less noise is required to achieve the same privacy budget.
- *But* there is little privacy amplification when $\mathcal{S}$ is a complex sampling design. (Bun et al. 2022)

Utility Considerations (I)

Privacy amplification by sampling

If $T(\mathfrak{s})$ is ε-DP and $\mathcal{S}(\mathfrak{f})$ randomly samples f fraction of the frame $\mathfrak{f}$, then $T' = T \circ \mathcal{S}$ is ε'-DP where $\varepsilon' \approx f\varepsilon$. (Balle et al. 2020)

- *Take-away:* If the sampling procedure is included, less noise is required to achieve the same privacy budget.
- *But* there is little privacy amplification when $\mathcal{S}$ is a complex sampling design. (Bun et al. 2022)

Utility Considerations (II)

- Surveys use weighted estimators $\sum_{i=1}^{n} w_i x_i$, which have increased sensitivity.
- Unweighted sums $\sum_{i=1}^{n} x_i$ have sensitivity $|\max x_i - \min x_i|$, where the max, min are over all possible values of x_i.
- Weighted estimators can have sensitivity

$$|\max w_i x_i - \min w_i x_i| + (n-1)(\max w_i - \min w_i)(|\max x_i| \vee |\min x_i|),$$

  because changing a record can change the weights of other records.
- Hence, weighted estimators require more noise to achieve the same privacy loss.
- Taking the frame as invariant means that the weights do not change.

Utility Considerations (II)

- Surveys use weighted estimators $\sum_{i=1}^{n} w_i x_i$, which have increased sensitivity.
- Unweighted sums $\sum_{i=1}^{n} x_i$ have sensitivity $|\max x_i - \min x_i|$, where the $\max, \min$ are over all possible values of x_i.
- Weighted estimators can have sensitivity

$$|\max w_i x_i - \min w_i x_i| + (n-1)(\max w_i - \min w_i)(|\max x_i| \vee |\min x_i|),$$

  because changing a record can change the weights of other records.
- Hence, weighted estimators require more noise to achieve the same privacy loss.
- Taking the frame as invariant means that the weights do not change.

Utility Considerations (II)

- Surveys use weighted estimators $\sum_{i=1}^{n} w_i x_i$, which have increased sensitivity.
- Unweighted sums $\sum_{i=1}^{n} x_i$ have sensitivity $|\max x_i - \min x_i|$, where the $\max, \min$ are over all possible values of x_i.
- Weighted estimators can have sensitivity

$$|\max w_i x_i - \min w_i x_i| + (n-1)(\max w_i - \min w_i)(|\max x_i| \vee |\min x_i|),$$

because changing a record can change the weights of other records.
- Hence, weighted estimators require more noise to achieve the same privacy loss.
- Taking the frame as invariant means that the weights do not change.

Utility Considerations (II)

- Surveys use weighted estimators $\sum_{i=1}^{n} w_i x_i$, which have increased sensitivity.
- Unweighted sums $\sum_{i=1}^{n} x_i$ have sensitivity $|\max x_i - \min x_i|$, where the $\max, \min$ are over all possible values of x_i.
- Weighted estimators can have sensitivity

$$|\max w_i x_i - \min w_i x_i| + (n-1)(\max w_i - \min w_i)(|\max x_i| \vee |\min x_i|),$$

  because changing a record can change the weights of other records.
- Hence, weighted estimators require more noise to achieve the same privacy loss.
- Taking the frame as invariant means that the weights do not change.

Utility Considerations (II)

- Surveys use weighted estimators $\sum_{i=1}^{n} w_i x_i$, which have increased sensitivity.
- Unweighted sums $\sum_{i=1}^{n} x_i$ have sensitivity $|\max x_i - \min x_i|$, where the $\max, \min$ are over all possible values of x_i.
- Weighted estimators can have sensitivity

$$|\max w_i x_i - \min w_i x_i| + (n-1)(\max w_i - \min w_i)(|\max x_i| \vee |\min x_i|),$$

  because changing a record can change the weights of other records.
- Hence, weighted estimators require more noise to achieve the same privacy loss.
- Taking the frame as invariant means that the weights do not change.

Additional Complications

- Data-dependent sampling designs are typical; but these pose a challenge unless the frame is fixed.
- Steps of the data release mechanism must be "algorithmised".
- Nonresponse must be included in the mechanism if starting from the sample-level or earlier.
    - In order to satisfy DP, one must assume that the nonresponse indicators are independent.

References I

Balle, Borja, Gilles Barthe, and Marco Gaboardi (Jan. 2020). "Privacy Profiles and Amplification by Subsampling". In: *Journal of Privacy and Confidentiality* 10.1. ISSN: 2575-8527. DOI: 10.29012/jpc.726.

Bun, Mark, Jörg Drechsler, Marco Gaboardi, Audra McMillan, and Jayshree Sarathy (June 2022). "Controlling Privacy Loss in Sampling Schemes: An Analysis of Stratified and Cluster Sampling". In: *Foundations of Responsible Computing (FORC 2022)*, 1:1–1:24.

Kifer, Daniel, John M. Abowd, Robert Ashmead, Ryan Cumings-Menon, Philip Leclerc, Ashwin Machanavajjhala, William Sexton, and Pavel Zhuravlev (Sept. 2022). *Bayesian and Frequentist Semantics for Common Variations of Differential Privacy: Applications to the 2020 Census*. Tech. rep. arXiv:2209.03310. DOI: 10.48550/arXiv.2209.03310. eprint: 2209.03310 (cs, stat). (Visited on 10/23/2022).

References II

US Census Bureau (Dec. 2022). *Disclosure Avoidance Protections for the American Community Survey*. https://www.census.gov/newsroom/blogs/random-samplings/2022/12/disclosure-avoidance-protections-acs.html. (Visited on 12/17/2023).