Statistica Sinica Preprint No: SS-2022-0187	
Title	Assessing Statistical Disclosure Risk for Differentially Private, Hierarchical Count Data, With Application to the 2020 U. S. Decennial Census
Manuscript ID	SS-2022-0187
URL	http://www.stat.sinica.edu.tw/statistica/
DOI	10.5705/ss.202022.0187
Complete List of Authors	Zeki Kazan and Jerome P. Reiter
Corresponding Authors	Zeki Kazan
E-mails	zekican.kazan@duke.edu

ASSESSING STATISTICAL DISCLOSURE RISK FOR DIFFERENTIALLY PRIVATE, HIERARCHICAL COUNT DATA, WITH APPLICATION TO THE 2020 U.S. DECENNIAL CENSUS

Zeki Kazan and Jerome P. Reiter

Duke University

Abstract: We propose Bayesian methods to assess statistical disclosure risks for count data released under zero-concentrated differential privacy, focusing on settings with a hierarchical structure. We discuss applications of these risk assessment methods to differentially private data releases from the 2020 U.S. decennial census and perform empirical studies using public individual-level data from the 1940 U.S. decennial census. Here, we examine how the data holder's choice of privacy parameters affects disclosure risks and quantify the increases in risk when an adversary incorporates substantial amounts of hierarchical information.

Key words and phrases: confidentiality, privacy, re-identification

1. Introduction

To protect individual respondents' privacy in the 2020 decennial census public release data products, the U.S. Census Bureau applies a variant of differential privacy (Abowd et al. (2022)). The Census Bureau decided to use a differentially private method, which they call the TopDown algorithm (TDA), because they determined that the risks of re-identifications when using methods employed in previous censuses, namely data swapping, are too great (Abowd (2018)). However, some stakeholders have questioned whether or not differential privacy is necessary, for example, Ruggles et al. (2019) and Kenny et al. (2021). This has led to calls from outside groups, such as JASON (2022), for the Census Bureau to evaluate the disclosure risks of the TDA. Such risk evaluations are particularly prudent because the Census Bureau quotes an $\varepsilon = 17.91$ to generate the perturbed counts. This exceeds the recommendations of Dwork (2008), who suggests $\varepsilon \leq \ln(3) \approx 1.10$ for typical applications, and is also likely outside the range of reasonable ε suggested by Lee and Clifton (2011).

In this article, we present and illustrate methods for evaluating statistical disclosure risks for differentially private count data nested in hierarchies. The core idea is to compute Bayesian posterior probabilities of disclosure, given the released counts and assumptions about adversaries' knowledge. We do so for counts generated using the differentially private mechanisms employed in the 2020 decennial census products, except that, unlike the TDA, we do not require released counts to be non-negative or to sum consistently across geographical hierarchies; we provide rationales for this choice in Section 2.3.

The remainder of this article is organized as follows. In Section 2, we present relevant background on differential privacy, statistical disclosure risks, and the 2020 decennial census application. In Section 3, we describe the Bayesian methods for assessing disclosure risks, distinguishing between settings with and without hierarchical information. In Section 4, we apply these methods to data from the 1940 decennial census using a mechanism that satisfies zero-concentrated differential privacy (Bun and Steinke (2016)), showing how the disclosure risks vary as a function of the privacy parameters and released counts. Finally, in Section 5, we conclude the paper.

2. Background

In this section, we review background material relevant to our methods, beginning with the privacy definitions that we reference throughout.

2.1 Differential Privacy

A mechanism $\mathcal{M}(\cdot)$ for releasing statistics satisfies differential privacy (henceforth, DP) if for any two data sets D and D' that differ in only one row, $\mathcal{M}(D)$ and $\mathcal{M}(D')$ are "similar." How similar $\mathcal{M}(D)$ and $\mathcal{M}(D')$ must be is determined by two privacy parameters, namely, $\varepsilon \in (0, \infty)$ and $\delta \in [0, 1)$. This idea is formalized as follows, drawing from Dwork et al. (2006).

Definition 1 (Differential Privacy)**.** A mechanism $\mathcal{M}$ satisfies (ε, δ)-DP if for input data sets D, D' that differ in only one row and $S \subseteq \mathit{Range}(\mathcal{M})$, $\mathbf{P}[\mathcal{M}(D) \in S] \leq e^{\varepsilon}\mathbf{P}[\mathcal{M}(D') \in S] + \delta$.

When $\delta = 0$, the guarantee is referred to as pure DP, and when $\delta > 0$, the guarantee is referred to as approximate DP. For approximate DP, typical values of δ in practice are extremely small; for example, the Census Bureau uses $\delta = 10^{-10}$ in their application.

With small ϵ (and δ), DP provides a formal guarantee that any one person's participation in D does not affect the results sufficiently for an adversary to detect that the person in question participated. In effect, this protects against an adversary who possesses information about all but one individual in D.

Mechanisms that satisfy DP can have undesirably long tails. Because of this drawback, the Census Bureau uses a variant of DP, called zero-concentrated differential privacy (henceforth, zCDP). The following definition draws from Bun and Steinke (2016).

Definition 2 (Zero-Concentrated Differential Privacy)**.** A mechanism $\mathcal{M}$ satisfies ρ-zCDP if for inputs D, D' that differ in only one row and all $\alpha \in (1, \infty)$, $D_\alpha(\mathcal{M}(D) \parallel \mathcal{M}(D')) \leq \rho\alpha$, where $D_\alpha(\mathcal{M}(D) \parallel \mathcal{M}(D'))$ is the α-Rényi divergence between the distributions of $\mathcal{M}(D)$ and $\mathcal{M}(D')$.

Note that zCDP allows for mechanisms in which randomness is added using a Gaussian distribution, which is not possible under pure DP and has undesirable properties under approximate DP (Bun and Steinke (2016)).

Because DP is a stronger criterion than zCDP, any mechanism that satisfies pure DP also satisfies zCDP (see Proposition 1.4 in Bun and Steinke (2016)). Converting from zCDP to DP is less straightforward; there is no guarantee that a mechanism satisfying zCDP will satisfy pure DP. Instead, for any $\delta > 0$, one can convert from zCDP to approximate DP, using the following theorem (Proposition 1.3 in Bun and Steinke (2016)).

Theorem 1. *If $\mathcal{M}$ satisfies ρ-zCDP, then for any $\delta > 0$, $\mathcal{M}$ satisfies (ε, δ)-DP for $\varepsilon = \rho + 2\sqrt{\rho \log(1/\delta)}$.*

The Census Bureau uses Theorem 1 to report the guarantees of their mechanisms in terms of approximate DP.

Finally, we define a mechanism for releasing integer-valued statistics under zCDP. Because the mechanism involves adding noise from a discrete Gaussian distribution (see Canonne et al. (2020)), we refer to it as the discrete Gaussian mechanism.

Definition 3 (Discrete Gaussian Mechanism)**.** Let $x \in \mathbb{Z}_{\geq 0}$ be a count statistic and suppose we wish to release a noisy count $X^* \in \mathbb{Z}$ satisfying ρ-zero concentrated differential privacy. The discrete Gaussian mechanism accomplishes this by producing a count centered at x, with noise from a discrete Gaussian distribution with parameter $1/(2\rho)$. That is,

$$\mathbf{P}[X^* = x^*] = \frac{e^{-\rho(x^*-x)^2}}{\sum_{\tilde{x}=-\infty}^{\infty} e^{-\rho(x^*-\tilde{x})^2}}, \qquad x^* \in \mathbb{Z}. \tag{2.1}$$

Notationally, $X^* \sim \mathsf{DG}(x, 1/(2\rho))$. Canonne et al. (2020) present an efficient algorithm for sampling from a discrete Gaussian distribution.

2.2 Statistical Disclosure Risk

For a given data set, the statistical disclosure risk (henceforth, used interchangeably with disclosure risk and risk) is a measure of the risk to respondent confidentiality as a consequence of releasing the data set to the public

(Domingo-Ferrer and Torra (2004); Duncan and Keller-McNulty (2000)). Duncan and Lambert (1986) propose measuring disclosure risk by directly modeling the behavior of a Bayesian adversary with a certain target in a released data set. Fienberg and Sanil (1997) present a similar approach, focusing on how the results change when bias or noise is present in the data. Reiter (2005) extends these methods to account for the technique used to perturb the data using legacy disclosure control methods, such as top-coding variables and data swapping.

There are several studies on statistical disclosure risks specifically under differential privacy. Lee and Clifton (2011) use a hypothetical adversary's posterior beliefs to provide a method for selecting ε for an arbitrary data set. Abowd and Vilhuber (2008) and McClure and Reiter (2012) focus on generating synthetic data with binary variables. In particular, Abowd and Vilhuber (2008) demonstrate that by measuring the disclosure risk using the posterior odds ratio between two data sets that differ in only one row, ε-DP provides a bound on the disclosure risk for that row. McClure and Reiter (2012) hypothesize a Bayesian adversary, make assumptions about what information the adversary has available a priori, and compare the adversary's posterior probability of determining the true values for a particular target with the corresponding prior probability.

Table 1: PL 94-171 tables for 2020 decennial census data.

P1	Race
P2	Hispanic or Latino, and not Hispanic or Latino by Race
P3	Race for the Population 18 Years and Over
P4	Hispanic or Latino, and not Hispanic or Latino by Race for the Population 18 Years and Over
P5	Group Quarters Population by Major Group Quarters Type
H1	Occupancy Status (Housing)

We expand on the methodologies of the aforementioned works, adapting the framework of McClure and Reiter (2012) and applying it to more general settings, including categorical variables with more than two levels and data sets with a hierarchical structure, such as the 2020 U.S. decennial census data.

2.3 DP in the 2020 Census Data Products

To describe the DP methods used by the Census Bureau for the 2020 decennial census, we draw from Abowd et al. (2022). We focus on the release of the PL 94-171 file, which comprises 2020 census data used for redistricting. The file includes the six summary tables listed in Table 1. These tables are produced across six levels of geographic hierarchy: blocks, optimized block groups, tracts, counties, states, and the nation. We focus on the persons tables P1 through P5 in Table 1.

To create these tables, the Census Bureau begins with a detailed his-

Table 2: Allocations of ρ in the 2020 decennial census. The second column displays the total proportion of ρ allocated to each geographic level, and the third column displays the proportion of ρ at that level allocated to the GVHR query. Values are taken from Section 8.2 of Abowd et al. (2022).

Geographic Level	Total Prop. ρ	GVHR Prop. ρ	GVHR $\approx \rho$
United States	104/4,099	189/241	0.0509
State	1,440/4,099	230/4,097	0.0505
County	447/4,099	754/4,097	0.0514
Tract	687/4,099	241/2,051	0.0504
Optimized Block Group	1,256/4,099	1,288/4,099	0.2465
Block	165/4,099	3,945/4,097	0.0992

togram of the Group quarters, Voting age, Hispanic, and Race variables—comprising a total of 2,016 cells—at each location at each level of the hierarchy. We refer to this histogram as the GVHR query. The Census Bureau first generates differentially private counts (satisfying zCDP) for this histogram using the discrete Gaussian mechanism. Then, they apply a post-processing algorithm to force counts at each level of the hierarchy to sum to the counts in the level above and to ensure that no counts are negative. This post-processing starts at the national level and works down the hierarchy. Finally, the Census Bureau aggregates the histograms to produce the summary tables listed in Table 1 and releases them to the public.

The Census Bureau assigned a global ρ of 2.56 that was divided among the levels of the geographic hierarchy, as displayed in the second column of Table 2. Each census block, for example, is allocated a ρ of $2.56 \times$

165/4,099 $\approx$ 0.103. Within each level of the hierarchy, the ρ budget is further divided among 11 queries. The allocations for the GVHR query are displayed in the third column of Table 2. The remainder of the allocations can be found in Section 8.2 of Abowd et al. (2022). The GVHR query in each census block, for example, is allocated a ρ of $0.103 \times 3{,}945/4{,}097 \approx 0.099$. The approximate ρ for each level is presented in the final column of Table 2. The Census Bureau also reports the DP guarantee, as computed using Theorem 1: the global $\rho = 2.56$ corresponds to $\varepsilon = 17.91$ for $\delta = 10^{-10}$.

The Census Bureau does not release the noisy GVHR query; rather, they release post-processed and aggregated counts. Several experts in DP have argued that the Census Bureau should additionally release the noisy counts without any post-processing (e.g., Dwork et al. (2021), Seeman et al. (2020)). Doing so could allow researchers using the data to avoid biases resulting from post-processing and estimate uncertainty properly. However, JASON, an advisory group for the U.S. government on issues related to science and technology, raises the concern that such a release could introduce disclosure risks. In their report (JASON (2022)), they recommend that the Census Bureau should "release all noisy measurements that are used to produce a published statistic when doing so would not incur undue disclosure risk" (p. 64), and should evaluate the risk that "the released data allows an adversary to make inferences about an individual's characteristics with more accuracy and confidence than could be done without the data released by the Census Bureau" (p. 114). The JASON recommendations, coupled with the calls to release the noisy counts without post-processing, motivate our methodological developments, which we now describe.

3. Methods for Assessing Disclosure Risks

We begin with methods that do not use information from further up the hierarchy, and then discuss accounting for extra information from higher levels.

3.1 Setting Without Hierarchical Information

Under DP and zCDP, the random noise protects against an adversary who possesses information about all but one individual in the data. For applications with data nested in geographic hierarchies, such as the 2020 census, this adversary seems unlikely. The adversary would have to possess information on all but one individual in the entire United States. To assess disclosure risks, we instead consider an attack scenario in which the adversary possesses information on all but one individual in a census block (more generally, in the lowest level of the hierarchy). For example, the adversary could be a landlord who owns all property in a block or an administrator for a group quarters institution that makes up an entire block. This set of assumptions represents a type of "worst case" scenario; the adversary possesses the most possible information about the individuals in the block, without possessing information about the target. We discuss and evaluate other attack scenarios, for example, in which the adversary knows information at the block and block-group levels, in Section S1 of the Supplementary Material.

To formalize, suppose the data comprise solely categorical variables and are organized in a hierarchy with h levels. We focus on a particular group, g_1, at the lowest level of the hierarchy, comprising n_1 individuals. In the census application, g_1 is a census block with n_1 persons. The adversary possesses complete data for $n_1 - 1$ of these individuals. We seek to assess the disclosure risk for the remaining individual, henceforth referred to as individual t (the targeted individual). Let c be the characteristics of individual t; that is, we label the combination of the variables in individual t's row as c. Let X_1 be the random variable from the adversary's perspective, representing the count of individuals in g_1 with characteristics c, and let $x_{1,-t}$ be the count of individuals in g_1 with characteristics c, excluding individual t; the adversary knows $x_{1,-t}$ a priori. The support of X_1 is $\{x_{1,-t}, x_{1,-t} + 1\}$.

Let $p \in (0,1)$ be the adversary's prior probability of assigning individual t to the correct category. That is, $\mathbf{P}[X_1 = x_{1,-t} + 1] = p$. Finally, let x_1 be the true count of individuals in g_1 with characteristics c. Here, x_1 is unknowable to the adversary but is known by the data holder (e.g., the Census Bureau). Here, we present results for the case $x_1 = x_{1,-t} + 1$.

As in the 2020 census application, we use zCDP and the discrete Gaussian mechanism. Let η_1 be the added noise such that $\eta_1 \sim \mathsf{DG}(0, 1/(2\rho_1))$. Let $X_1^* = x_1 + \eta_1$ be the random variable representing the differentially private value of x_1 and let x_1^* be its realized outcome. The adversary's posterior probability of correctly concluding that $X_1 = x_1 = x_{1,-t} + 1$ is

$$\mathbf{P}[X_1 = x_1 \mid X_1^* = x_1^*] = \frac{pe^{-\rho_1(x_1^* - x_1)^2}}{pe^{-\rho_1(x_1^* - x_1)^2} + (1-p)e^{-\rho_1(x_1^* - x_{1,-t})^2}}. \tag{3.1}$$

We define the ratio of the posterior and prior probabilities as

$$R'(x_1^*) = \frac{\mathbf{P}[X_1 = x_1 \mid X_1^* = x_1^*]}{\mathbf{P}[X_1 = x_1]}. \tag{3.2}$$

For the data holder, a relevant measure of disclosure risk averages (3.1) over possible realizations of X_1^*, given the true counts. We write this as

$$\mathbf{P}[X_1 = x_1 \mid x_1 = x_{1,-t}+1] = \sum_{x_1^*=-\infty}^{\infty} \mathbf{P}[X_1 = x_1 \mid X_1^* = x_1^*] \\ \mathbf{P}[X_1^* = x_1^* \mid x_1 = x_{1,-t}+1]. \quad (3.3)$$

All probabilities written as conditional on x_1 indicate that they depend on the true count. Dividing (3.3) by the adversary's prior $\mathbf{P}[X_1 = x_1]$, define

$$\begin{aligned} R &= \frac{\mathbf{P}[X_1 = x_1 \mid x_1 = x_{1,-t}+1]}{\mathbf{P}[X_1 = x_1]} \\ &= \sum_{x_1^*=-\infty}^{\infty} R'(x_1^*)\, \mathbf{P}[X_1^* = x_1^* \mid x_1 = x_{1,-t}+1]. \end{aligned} \quad (3.4)$$

We also examine how the adversary can use the posterior probabilities to make decisions. Let $\hat{X}_1$ be the adversary's point estimate of X_1 (either $x_{1,-t}$ or $x_{1,-t}+1$). Here, $\hat{X}_1$ is a function of the observed noisy counts, $\mathcal{D} = \{x_1^*\}$. We assume the adversary decides the value of $\hat{X}_1$ by minimizing a loss function $\mathcal{L}(\hat{X}_1, X_1)$. We use the zero-one loss, $\mathcal{L}(\hat{X}_1, X_1) = \mathbf{1}[X_1 \neq \hat{X}_1]$, where $\mathbf{1}[\cdot]$ is an indicator function. The Bayes estimator is then

$$\operatorname*{argmin}_{\hat{X}_1} \mathsf{E}[\mathcal{L}(\hat{X}_1, X_1) \mid \mathcal{D}] = \operatorname*{argmax}_{\hat{X}_1} \mathbf{P}[X_1 = \hat{X}_1 \mid \mathcal{D}]. \quad (3.5)$$

That is, the adversary's point estimate for X_1 is whichever of $x_{1,-t}$ and $x_{1,-t}+1$ has the higher posterior probability. Because $\hat{X}_1$ is a deterministic function of x_1^*, the probability of the adversary correctly selecting $\hat{X}_1 = x_1$ is

$$\begin{aligned}\mathbf{P}[\hat{X}_1 = x_1] &= \sum_{x_1^*=-\infty}^{\infty} \mathbf{P}[X_1^* = x_1^* \mid X_1 = x_1]\mathbf{1}\left[\mathbf{P}[X_1 = x_1 \mid X_1^* = x_1^*] > \frac{1}{2}\right] \\ &= \mathbf{P}[X_1^* \geq \tilde{x}_1^* \mid X_1 = x_1], \qquad (3.6)\end{aligned}$$

where $\tilde{x}_1^*$ is the smallest value of x_1^* such that $\mathbf{P}[X_1 = x_1 \mid X_1^* = x_1^*] > \frac{1}{2}$.

3.2 Incorporating Hierarchical Information

It makes sense that, for example, if $x_{1,-t} = 5$ and we observe that the count of people with the characteristics c one level up (the block group level) is five, then it should be more likely that $x_1 = 5$ than $x_1 = 6$. Thus, using hierarchical information should improve the adversary's inference. However, if the noisy count one level up is 100, then the hierarchical information is probably not very useful in deciding between $x_1 = 5$ and $x_1 = 6$. In this section, we formalize this intuition for the case in which we use two levels of the hierarchy; generalizing to more levels is straightforward conceptually.

Let g_2 be the group at the second level of the hierarchy containing t, and let n_2 be the number of individuals in g_2. In the census application, g_2 is the block group—comprising n_2 individuals—that contains census block g_1. Let X_2 be the random variable from the adversary's perspective representing the count of individuals in g_2 with characteristics c, and let x_2 be the true count. Let $\eta_2 \sim \mathsf{DG}(0, 1/(2\rho_2))$ be noise added using the differentially private mechanism. Let $X_2^* = x_2 + \eta_2$ be the random variable representing the differentially private version of x_2, and let x_2^* be the realization that is released. Let $Y_1^{(1)}, \ldots, Y_1^{(d)}$ be random variables representing the counts of individuals with characteristics c in the other d groups at the lowest level. That is, $X_2 = X_1 + \sum_{i=1}^d Y_1^{(i)}$. Let $\eta_1^{(1)}, \ldots, \eta_d^{(1)} \overset{iid}{\sim} \mathsf{DG}(0, 1/(2\rho_1))$ be the perturbations from the privacy mechanism. Set $Y_1 = \sum_{i=1}^d Y_1^{(i)}$ and $Y_1^* = \sum_{i=1}^d (Y_1^{(i)} + \eta_1^{(i)}) = Y_1 + \sum_{i=1}^d \eta_1^{(i)}$. The observed data are $\mathcal{D} = \{x_1^*, x_2^*, y_1^*\}$.

We first briefly describe an approximation that the adversary (or data holder) can make to simplify the computations. Because $Y_1^* = Y_1 + \sum_{i=1}^d \eta_1^{(i)}$, it follows that Y_1^* is an approximation of Y_1, with noise from a sum of discrete Gaussian distributions. The noise term is difficult to handle when performing inference, so we assume that the adversary makes the approximation

$$\eta_1^{(i)} \overset{iid}{\sim} \mathsf{DG}(0, 1/(2\rho_1)) \implies \sum_{i=1}^d \eta_1^{(i)} \approx \mathsf{DG}(0, d/(2\rho_1)). \tag{3.7}$$

This approximation improves as $\rho \to 0$ and $d \to \infty$ and is very accurate for the ρ and d used in Section 4. See Section S2 in the Supplementary Material for details.

The adversary can use Bayesian inference to compute posterior probabilities and disclosure risks akin to those in Section 3.1 using information from the second level of the hierarchy. Specifically, the adversary runs a Gibbs sampler to sample from the posterior distribution of (X_1, X_2) given $\mathcal{D}$ and uses the marginal posterior for X_1 to estimate the posterior probability that $X_1 = x_{1,-t} + 1$. Then the risk computation and point estimation techniques described in Section 3.1 can be applied. For the adversary's prior distribution for $X_2 \mid X_1 = k_1$, one reasonable choice is a uniform distribution on the set $\{k_1, k_1 + 1, \ldots\}$. In our numerical experiments, the results were minimally affected when imposing finite upper bounds on the support of a uniform prior distribution. However, the results can change substantially when using highly informative prior distributions that are poorly specified. Section S4 of the Supplementary Material provides details of these experiments.

The full conditional for X_1 is, for $k_1 \in \{x_{1,-t}, x_{1,-t}+1\}$ and $k_1 \leq k_2$,

$$\begin{aligned}\mathbf{P}[X_1 = k_1 \mid X_2 = k_2, X_1^* = x_1^*, X_2^* = x_2^*, Y_1^* = y_1^*]\\ \propto \exp\left\{-\frac{d+1}{d}\rho_1\left[k_1 - \frac{dx_1^* + (k_2 - y_1^*)}{d+1}\right]^2\right\}\mathbf{P}[X_1 = k_1]. \quad (3.8)\end{aligned}$$

The full conditional for X_2 is, for $k_2 \in \{k_1, k_1+1, \ldots\}$,

$$\begin{aligned}\mathbf{P}[X_2 = k_2 \mid X_1 = k_1, X_1^* = x_1^*, X_2^* = x_2^*, Y_1^* = y_1^*]\\ \propto \exp\left\{-\left(\rho_2 + \frac{\rho_1}{d}\right)\left[k_2 - \frac{\rho_2 x_2^* + \frac{\rho_1}{d}(y_1^* + k_1)}{\rho_2 + \frac{\rho_1}{d}}\right]^2\right\}. \quad (3.9)\end{aligned}$$

Thus, the full conditional for X_1 has a Bernoulli distribution, and the full conditional for X_2 can be sampled easily over a grid. Section S3 of the Supplementary Material includes details of the derivations.

4. Empirical Applications with 1940 Census Data

In this section, we demonstrate the methodology from Section 3 using data from the 1940 U.S. decennial census, investigating whether the Census Bureau's choice of privacy parameters in 2020 would have led to unacceptably high disclosure risks for the GVHR query. In Section 4.1, we describe the 1940 decennial census data and how they differ from the 2020 decennial census data. In Section 4.2, we consider the case in which an adversary uses only the released counts from the lowest level of the hierarchy. In Section 4.3, we extend this to the case in which the adversary leverages the released counts from a second level of the hierarchy, examining how the hierarchical information affects risks and how this effect depends on x_2 and ρ_2.

4.1 The 1940 Census

Every 72 years, the Census Bureau is permitted to release record-level data collected in the decennial census, without any redaction for privacy protection. Thus, the 1940 census data are an excellent testbed for our methods.

The structure of the 1940 census data differs from that of the 2020 census data. The 1940 census data have a smaller hierarchy, comprising enumeration districts within counties, states, and the country. Per the national archives, enumeration districts could be "covered by a single enumerator or census taker in one census period that lasted several weeks" (National Archives and Records Administration, 2022) and so can vary substantially in size. For illustrative purposes, we focus on a small set of enumeration districts in North Carolina that are roughly the size of the average modern-day census block.

Another important difference is the number of categories for the vari-

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0187

Table 3: Histogram for the GVHR query for enumeration district 28-21 from the 1940 decennial census.

HHGQ	VOTINGAGE	HISPANIC	CENRACE	Count
Household	Of Voting Age	Not Hispanic	White	34
Household	Not Of Voting Age	Not Hispanic	White	10
Household	Of Voting Age	Not Hispanic	Black	1

ables of interest. In the 2020 census, the GVHR query produces a histogram with 2,016 possible levels—the product of two levels for each of the Hispanic and Voting Age variables, eight for the Group Quarters variable, and 63 for the Race variable. In 1940, this query had only 864 possible levels—the product of two for Voting Age, six for Hispanic, eight for Group Quarters, and nine for Race. The 1940 census includes the exact age of each individual, which we transform to indicate whether or not the person is of voting age, in order to match the 2020 census data releases for the PL 94-171 file.

4.2 Setting Without Hierarchical Information

We consider enumeration district 28-21 in Dare County, North Carolina, in 1940. Table 3 displays the counts for its GVHR query. This enumeration district was similar in size to a modern-day census block and consisted of 45 nonHispanic individuals residing in households. One individual has unique characteristics and so is particularly vulnerable to an attack.

We assume the adversary possesses complete information about all non-

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0187

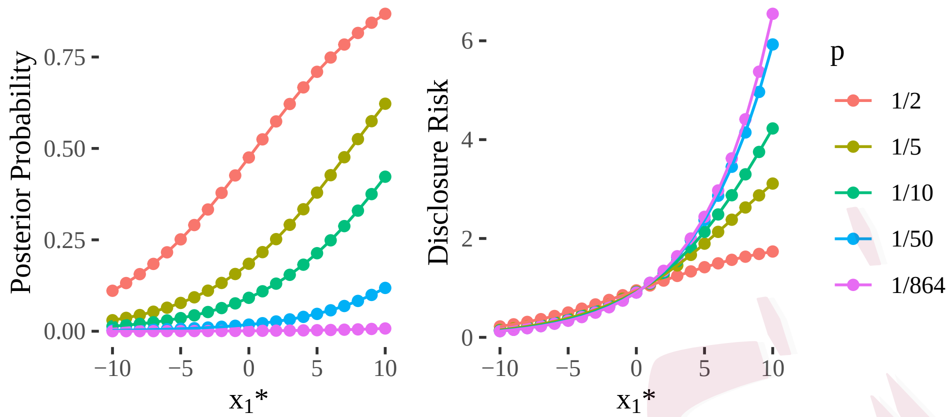

Figure 1: The left panel plots the posterior probability that the adversary makes the correct decision, $\mathbf{P}[X_1 = 1 \mid X_1^* = x_1^*]$, as a function of x_1^*. The right panel plots the implied disclosure risk, $R'(x_1^*)$, as a function of x_1^*. The colors correspond to different adversary prior beliefs. We set $\rho_1 = 0.099$.

unique residents in enumeration district 28-21 (g_1) and wishes to target the resident (t), who is uniquely a householder of voting age, nonHispanic, and black (c). The adversary's known count is $x_{1,-t} = 0$. We consider five adversaries, who differ only in the prior probability they assign to the event that $X_1 = 1$; the first assigns probability $p = 1/2$, the second $p = 1/5$, the third $p = 1/10$, the fourth $p = 1/50$, and the fifth $p = 1/864$ (equal prior probability on the 864 possible levels of the GVHR query in 1940).

We first examine the relationship between the posterior probability and the noisy released count, x_1^*, when $\rho_1 \approx 0.099$, as in the 2020 census application. Figure 1 displays a plot of the posterior probability that the adversary assigns to the correct conclusion that $x_1 = 1$ as a function of x_1^* and the implied risk $R'(x_1^*)$ for each of the five adversaries. For each adversary, both the posterior probability and the risk are monotonically increasing functions of the released x_1^*. Notably, the posterior probability is greater than the prior probability p (and, thus, the risk is greater than one) if $x_1^* \geq 1$, whereas the posterior is less than p (and the risk is less than one) when $x_1^* < 1$. This makes intuitive sense: if the released statistic is one or more, this is evidence of the true count being one, and the posterior probability that $X_1 = 1$ will increase relative to the prior probability. If the released statistic is zero or less, this is evidence of the true count being zero and the posterior probability that $X_1 = 1$ will decrease relative to the prior probability.

In general terms, the risk is acceptable if it is near one or, equivalently, if the posterior probability that the adversary makes the correct decision is near p. Thus, we should be concerned if the probability of releasing a statistic that produces a risk much greater than one is high. Table 4 presents the risks and posterior probabilities implied by several values of x_1^*, along with the probability of observing that value. For $x_1^* \geq 4$, the risk substantially exceeds one, especially for adversaries with low p. A statistic of $x_1^* = 4$ will occur only 3.6% of the time, and a statistic of $x_1^* = 5$ only

Table 4: For several values of x_1^*, the values of the probability mass $\mathbf{P}[X_1^* = x_1^* \mid X_1 = 1]$, the posterior probability that the adversary makes the correct decision $\mathbf{P}[X_1 = 1 \mid X_1^* = x_1^*]$, and the disclosure risk $R'(x_1^*)$. We set $\rho_1 = 0.099$. Posterior probabilities for $p = 1/864$ are quite small and so are omitted.

		Posterior Probability				Disclosure Risk				
x_1^*	Mass	$p = \frac{1}{2}$	$\frac{1}{5}$	$\frac{1}{10}$	$\frac{1}{50}$	$\frac{1}{2}$	$\frac{1}{5}$	$\frac{1}{10}$	$\frac{1}{50}$	$\frac{1}{864}$
1	0.161	0.525	0.216	0.109	0.022	1.05	1.08	1.09	1.10	1.10
2	0.119	0.574	0.252	0.130	0.027	1.15	1.26	1.30	1.34	1.35
3	0.073	0.622	0.291	0.154	0.032	1.24	1.46	1.54	1.62	1.64
4	0.036	0.667	0.334	0.182	0.039	1.33	1.67	1.82	1.96	2.00
5	0.015	0.710	0.379	0.213	0.047	1.42	1.90	2.13	2.37	2.44

1.5% of the time. While it is unlikely that the released statistic will produce a risk of this magnitude in any particular census block, we observe such risks in a sizable number of the millions of census blocks in the United States. Whether or not a risk of this magnitude is unacceptable is a decision for policymakers.

We next examine the expected value of the disclosure risks over realizations of the released differentially private counts, and how this quantity varies with ρ_1. The first panel of Figure 2 plots the posterior probability that the adversary correctly concludes $x_1 = 1$, marginalizing over X_1^*, as a function of ρ_1. The second panel of Figure 2 plots the implied risk, R. As expected, both quantities increase monotonically with ρ_1, with the posterior probability increasing from p at ρ_1 near zero to one for ρ_1 very large, and the risk increasing from 1 to $1/p$ along the same range. Across the various

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0187

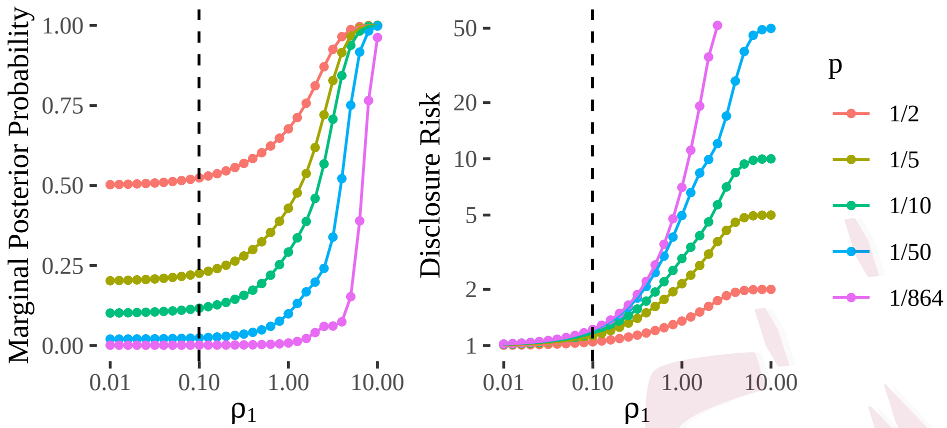

Figure 2: The left panel plots the marginal posterior probability that the adversary makes the correct decision, $\mathbf{P}[X_1 = 1 \mid x_1 = 1]$, as a function of ρ_1. The right panel plots the implied disclosure risk, R (defined in Equation 3.4), as a function of ρ_1. The colors correspond to different adversary prior beliefs, and the dashed line corresponds to $\rho_1 = 0.099$. Note the log scales.

p, for $\rho_1 < 0.25$, the posterior probability is approximately p and the risk is approximately one, indicating that the adversary gleans little from the released statistic. For $0.25 \leq \rho_1 \leq 5$, the posterior probability increases rapidly from p to 1 and the risk from 1 to $1/p$, indicating that the choice of ρ_1 is crucial in this range. Even small increases in ρ_1 can cause large increases in a hypothetical adversary's posterior probability. For $\rho_1 > 5$, the posterior probability is approximately one and the risk is approximately $1/p$, indicating that the parameters in this range are too high for practical settings.

Figure 2 also demonstrates possible implications of the above discussion

Table 5: The marginal posterior probability that the adversary makes the correct decision and the implied risk, R, for each prior probability p when $\rho_1 = 0.099$.

p	Posterior Probability	Disclosure Risk
1/2	0.524	1.05
1/5	0.225	1.13
1/10	0.117	1.17
1/50	0.024	1.21
1/864	0.0014	1.22

for the 2020 decennial census. The Census Bureau's chosen $\rho_1 \approx 0.099$ is at the high end of the range where the posterior probability is approximately equal to the prior probability and the implied risk is approximately one. As shown in Table 5, the risks are between 1.05 and 1.21, indicating that the Census Bureau's choice of ρ_1 is at a reasonable level for the release of the query. If the Census Bureau were to increase ρ_1 by any substantial amount, the risk would increase to levels likely deemed unacceptable. On the other hand, if they were to decrease ρ_1 slightly, the decrease in risk would be minimal.

We next examine the adversary's behavior from a decision-theoretic standpoint. Figure 3 displays a plot of the probability that the adversary correctly selects $\hat{X}_1 = 1$ as a function of ρ_1 for the five adversary prior probabilities. For $p = 1/2$, any $x_1^* \geq 1$ leads to a posterior probability greater than 0.5. Thus, the curve when $p = 1/2$ is simply a plot of $\mathbf{P}[X_1^* \geq$

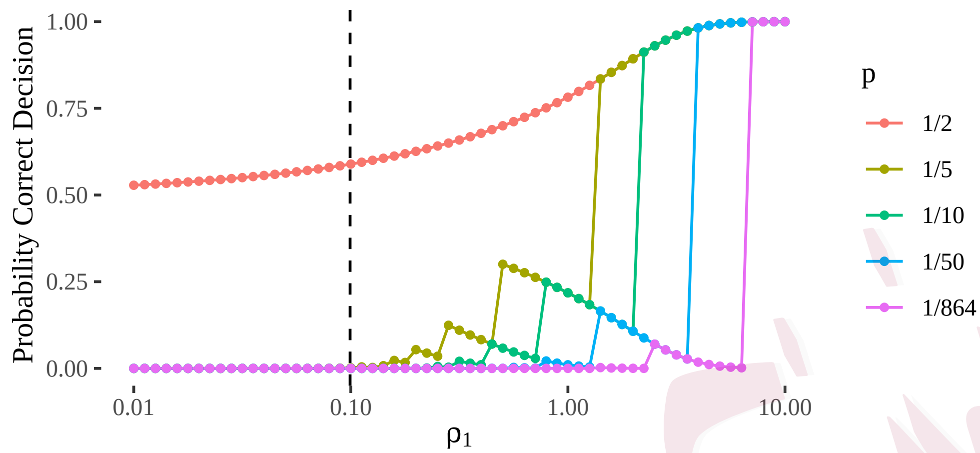

Figure 3: The probability that the adversary makes the correct decision under a 0–1 loss as a function of ρ_1. Colors correspond to different adversary prior beliefs, and the dashed line presents $\rho_1 = 0.099$. The x-axis is in log scale.

$1 \mid X_1 = 1]$ as a function of ρ_1. For $p < 1/2$, this need not be the case. For example, when $p = 1/5$ and $\rho_1 = 0.5$, $\mathbf{P}[X_1 = 1 \mid X_1^* = x_1^*] > 0.5$ if and only if $x_1^* \geq 2$, giving $\mathbf{P}[\hat{X}_1 = 1] = \mathbf{P}[X_1^* \geq 2 \mid X_1 = 1] \approx 0.30$. If, however, we increase to $\rho_1 = 0.6$, still $\mathbf{P}[X_1 = 1 \mid X_1^* = x_1^*] > 0.5$ if and only if $x_1^* \geq 2$, but now $\mathbf{P}[\hat{X}_1 = 1] = \mathbf{P}[X_1^* \geq 2 \mid X_1 = 1] \approx 0.28$. That is, because the probability of extreme values decreases as ρ_1 increases, it is possible for the probability of making the correct decision to decrease as well. This leads to the jaggedness in the plot for $p \leq 1/5$ (and for other $p < 1/2$) when ρ_1 is not sufficiently high that observing $x_1^* = 1$ will give the correct decision.

As evident in Figure 3, the implications for the 2020 census application are very different for $p = 1/2$ and for $p \leq 1/5$. For $p = 1/2$, the probability that the adversary makes the correct decision is 0.59, which is near 0.5. For $p \leq 1/5$, however, the probability that the adversary makes the correct decision is a fraction of a percent. Apparently, the assumptions about the adversary's prior knowledge are extremely important for this metric. An adversary who believes there is a 50–50 chance that the target has characteristics c will make the correct decision a substantial proportion of the time, but if that prior probability decreases even slightly, the probability of the adversary making the correct decision under a 0–1 loss decreases significantly.

4.3 Incorporating Hierarchical Information

In this section, we explore how incorporating hierarchical information affects the disclosure risks. To do so, we consider enumeration district 39-14 in Granville County, North Carolina, one of 28 enumeration districts in the county. This district includes a white Hispanic individual who resided in an institution for the elderly, handicapped, and poor and was not of voting age. This person was unique at both the enumeration district level (which contained 209 people) and at the county level (which contained 29,364 peo-

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0187

Table 6: One sample of possible $\mathcal{D} = \{x_1^*, x_2^*, y_1^*\}$ with ρ_1 and ρ_2 from the Census application.

x_1	x_1^*	x_2	x_2^*	y_1	y_1^*
1	2	1	1	0	-1

Table 7: Adversary's posterior distribution for X_1 given $\mathcal{D}$ from Table 6, $p = 1/2$, and 10^4 MCMC draws.

k_1	0	1
$\mathbf{P}[X_1 = k_1 \mid \mathcal{D}]$	43.0%	57.0%

Table 8: Adversary's empirical posterior distribution for X_2 given the data in Table 6, $p = 1/2$, and 10^4 MCMC draws.

k_2	0	1	2	3	4	5	6	7
$\mathbf{P}[X_2 = k_2 \mid \mathcal{D}]$	11%	40%	30%	14%	4.6%	0.7%	0.08%	0.01%

ple). Thus, the true data is $(x_1, x_2, y_1) = (1, 1, 0)$, where we define y_1 as the actual count of Y_1. Table 6 displays one realization of the discrete Gaussian mechanism using the Census Bureau's chosen $\rho_1 \approx 0.099$ and $\rho_2 \approx 0.246$. The noisy counts are $\mathcal{D} = (x_1^*, x_2^*, y_1^*) = (2, 1, -1)$

Given $\mathcal{D}$, the adversary would sample from the posterior distribution of (X_1, X_2). We draw 10,000 posterior samples from the MCMC sampler described in Section 3.2, assuming that the adversary places a prior probability $p = 1/2$ on $X_1 = 1$. Table 7 summarizes the adversary's marginal posterior distribution for X_1. Because $x_1^* = 2$, the posterior distribution places more weight on $X_1 = 1$ than it does on $X_1 = 0$, but the difference is only slight because of the low ρ values. If the adversary does not use hierarchical information, the posterior probability that $X_1 = 1$ is 57.4%, indicating a minimal change from using the hierarchical information in this case (in fact, the hierarchical information slightly decreases the adversary's posterior probability of the correct choice). The adversary can also examine the posterior distribution of X_2, although this is of less practical interest; Table 8 presents the posterior summaries. This posterior distribution places 40% of the mass on $X_2 = 1$, and the probabilities for the right tail decline to zero relatively quickly.

For the remainder of this section, we consider the disclosure risks at different values of $\mathcal{D}$. We assume throughout that the adversary sets $p = 1/2$. We focus on the case in which the targeted individual is unique at the lowest level of the hierarchy; see Section S1.1 of the Supplementary Material for a discussion of extending this to non-unique individuals.

We begin by examining when the hierarchical information affects the adversary's decision for $\rho_1 \approx 0.099$ and $\rho_2 \approx 0.247$, as in the 2020 decennial census. To do so, we enumerate all reasonable combinations of x_2^* and y_1^*, which we choose as $-3 \leq x_2^* \leq 5$ and $-30 \leq y_1^* \leq 30$, respectively; this region contains over 99% of the probability mass. Figure 4 displays the decisions for these combinations. Without hierarchical information, the adversary always decides that $x_1 = 1$ when $x_1^* \geq 1$ and $x_1 = 0$ when $x_1^* \leq 0$. Using the hierarchical information does not change the adversary's decision when $x_1^* \geq 2$ or $x_1^* \leq -1$. However when $x_1^* \in \{0, 1\}$, the hierarchical

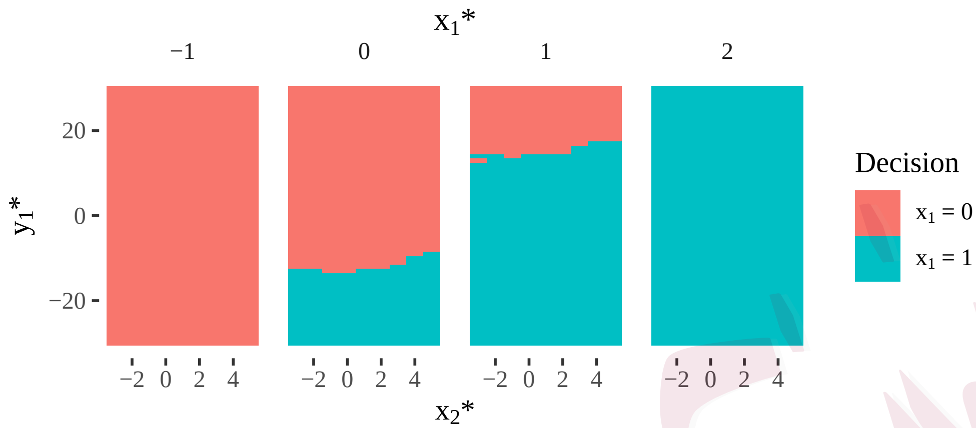

Figure 4: Adversary's decision under 0–1 loss for each combination of x_1^*, x_2^*, and y_1^*. Privacy parameters are set as in the census application, $p = 1/2$, and $d = 27$. 10^3 MCMC draws are taken for each combination in most cases. When the posterior probability $x_1 = 1$ is close to 0.5, the number of MCMC draws is increased to 2.5×10^5.

information can change the adversary's decision. When $x_1^* = 1$ and y_1^* is positive and large, the adversary decides that $x_1 = 0$, whereas they choose $x_1 = 1$ without the hierarchical information. Similarly, when $x_1^* = 0$ and y_1^* is negative and large in absolute value, the adversary decides that $x_1 = 1$, whereas they choose $x_1 = 0$ without the hierarchical information. Exactly how large y_1^* must be depends on the observed x_2^*. The tiles in Figure 4 where the hierarchical information causes the adversary to correctly change their decision correspond to 2.48% of the probability mass, whereas the tiles where the reverse occurs correspond to 1.83% of the probability mass. Thus, the adversary has a slight net gain from the hierarchical information,

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0187

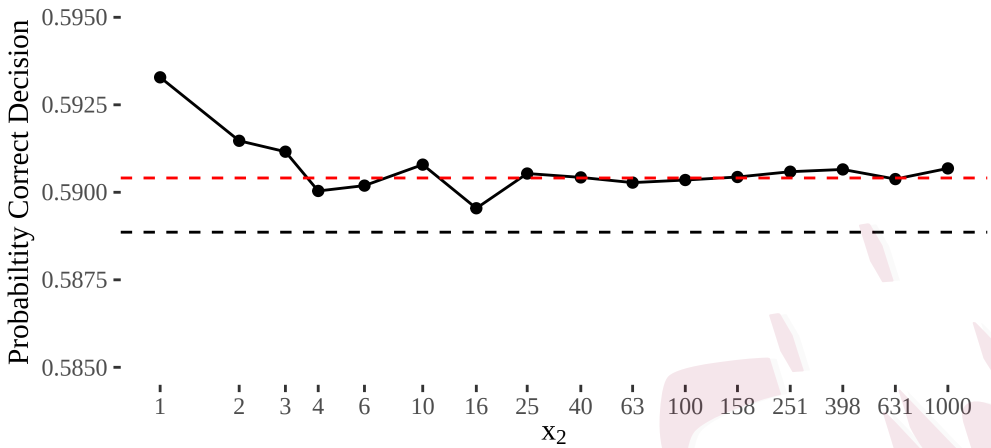

Figure 5: Proportion of times the adversary correctly concludes that $x_1 = 1$, as a function of x_2. Proportions are over 10^6 random draws of $\mathcal{D}$. For each draw, 10^3 MCMC samples are used to estimate the posterior. The dashed black line is the corresponding probability when ignoring hierarchical information, and the dashed red line is the average proportion over $x_2 \geq 4$. We set $\rho_1 = 0.099$, $\rho_2 = 0.247$, $p = 1/2$, and $d = 27$.

increasing the probability of a correct decision by 0.65%.

We next examine how sensitive the disclosure risks are to the true count at the second level, x_2. How would the results change if the target was unique at the lowest level, but not at the second level? To examine this, we alter the true counts, adding individuals to group g_2 with characteristics c. Figure 5 displays the proportion of times the adversary makes the correct decision as a function of x_2. Note that the y-axis has a small range; even with 10^6 samples per point, the Monte Carlo noise obscures the relationship. Regardless, $x_2 = 1$ is a clear outlier; the probability that

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0187

the adversary makes the correct decision is not only higher than the corresponding probability when ignoring hierarchical information, but is higher than the analogous probabilities for $x_2 > 1$. For $x_2 > 3$, the probabilities of making the correct decision are centered around 59.05%, with a small amount of Monte Carlo error.

Applying these findings to the 2020 census application, the main implication is that the conclusions when $x_2 = 1$ generalize to $x_2 > 1$, although the increase in the probability that the adversary makes the correct decision is less pronounced. Therefore, the discussion about the trade-off between privacy and accuracy in terms of the Census Bureau's choice of ρ_2 applies, no matter the value of x_2. A remarkable feature of Figure 5 is how small the effect is of changes in x_2.

Finally, we examine the effect of changes in the second-level privacy parameter, ρ_2. Figure 6 displays the probability that the adversary makes the correct decision, as a function of ρ_2, for $\rho_1 \approx 0.099$ and a few values of x_2. The effect of $x_2 = 1$ observed previously—increasing the adversary's probability of making the correct decision from when $x_2 > 1$—lessens as ρ_2 increases. For $\rho_2 > 1$, the probabilities for $x_2 \in \{1, 10, 100\}$ are quite similar. The most striking feature of Figure 6 is that even for extremely large values of ρ_2, the probability does not noticeably increase for any of

Statistica Sinica: Preprint
doi:10.5705/ss.202022.0187

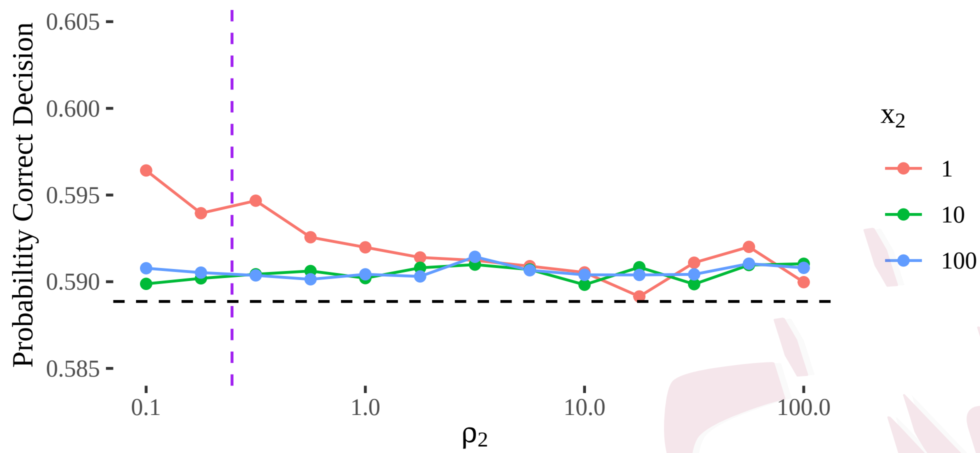

Figure 6: Plot of the proportion of the time the adversary correctly concludes that $x_1 = 1$, as a function of ρ_2, colored for a few selected x_2. The proportions are over 5×10^5 MC draws for $\mathcal{D}$. For each draw, 10^3 MCMC samples are used to estimate the posterior. The dashed black line is the corresponding probability when ignoring hierarchical information, and the dashed purple line is $\rho_2 = 0.247$. We set $\rho_1 = 0.099$, $p = 1/2$, and $d = 27$.

the selected x_2. When ignoring hierarchical information, the corresponding probability is 58.89%. Even when $\rho_2 = 100$, the probability only increases to 59.2%. This finding is a feature of the specific attack scenario. Because the adversary knows only $x_{1,-t}$, and not the counts in the other blocks, knowing x_2 nearly exactly does not tell them which blocks making up g_2 contain the individuals with characteristics c. Thus, their probability of making the correct decision with respect to block g_1 improves as the uncertainty in X_2 decreases, but does not go to one. Mathematically, in (3.9) as $\rho_2 \to \infty$, $\mathbf{P}[X_2 = x_2 \mid X_1 = k_1, \mathcal{D}] \to 1$, causing the full conditional for X_1 in (3.8) to converge to a constant function of x_2 (but not to one unless $\rho_1 \to \infty$). In the Supplementary Material, we show that hierarchical information can increase risks more noticeably in other attack scenarios, for example, when the intruder knows g_2.

Overall, the results suggest that the adversary gains little from using the hierarchical information in this attack scenario. This may provide evidence that the Census Bureau's choice of ρ_2 is reasonable from a disclosure risk perspective. It also suggests that the Census Bureau could increase ρ_2—and likely the ρ values at higher levels of the hierarchy—without significantly compromising the disclosure risk, at least under this attack scenario. Thus, more accurate statistics could be released at the upper levels of the hierarchy. Essentially, if we assume the adversary only has complete information about everyone except the target at the lowest level of the hierarchy, then only the choice of ρ_1 has a meaningful effect on the probability that the adversary makes the correct decision.

5. Conclusion

We provide a methodology to compute statistical disclosure risks for categorical data with many level,s released under zCDP, while incorporating hierarchical information. Following the suggestion in the JASON report, we demonstrate how to conduct empirical analyses that could be used to evaluate the effect on disclosure risks of releasing the GVHR query at the census block level, prior to post-processing. In our studies of the 1940 census data, we find that, when assuming the adversary possesses information about all but one individual at the lowest level of the hierarchy, the main factor affecting the disclosure risk is ρ_1, the privacy parameter at the lowest level. The hierarchical information does not have an appreciable affect on the accuracy of the adversary's posterior inference under these assumptions.

The redistricting files are only one set of counts from the 2020 census released by the Census Bureau; others are released over time. Thus, from the lens of differential privacy, it is reasonable to question the value of assessing disclosure risks for specific attack scenarios at a point in time. We believe such assessments have a useful role to play. First, at the stage of algorithm design, they can help the data holder, including decision-makers who may comprehend Bayesian probabilities more readily than bounds on Rényi divergences, understand the risks inherent in different choices of privacy parameters. Similarly, they can help the data holder explain the privacy protection to the public, because posterior probabilities and posterior-to-prior ratios can be more interpretable than guarantees expressed in terms of privacy parameters (Hotz et al. (2022)). We also note that the risk measures have desirable composition properties in settings in which the same attack is applied to sequential releases, with the total disclosure risk from m releases being equivalent to the product of the risk from each release (see Section S5 of the Supplementary Material for details). Note, however, that settings in which the output of one release is used as side information for a different attack strategy are less straightforward; analyses of the risk composition in such settings is left to future research.

Naturally, our findings are specific to a particular attack scenario—an adversary with complete information about everyone in g_1, except the target, about whom they know nothing. We can modify these assumptions in a number of ways, and still use the same approach; see Section S1.2 of the Supplementary Material for a few examples. However, if we make significant changes to the adversary's assumed knowledge, the results of the analysis may change. Section S1.3 of the Supplementary Material includes an example of this, whereby we assume the adversary knows the target is unique at multiple levels of the hierarchy. Here, the effect of the hierarchical information can be stronger. An adversary with this type of knowledge may or may not be realistic in some settings; whether this is the case for the decennial census is a decision for the Census Bureau.

As noted previously, we do not consider the TDA's post-processing step or population invariants. The addition of a post-processing step by itself, that is, absent invariants derived directly from the confidential data, does not affect the formal privacy guarantee and also should not increase the statistical disclosure risks. Either the post-processing step is invertible, in which case the risk analysis does not change, or it is not invertible, in which case the computation of disclosure risks is far more uncertain and computationally difficult (Gong and Meng, 2020). We illustrate this in Section S6 of the Supplementary Material. The population invariants include the total populations of each state, total number of housing units in each census block, and number of occupied group quarters of each type in each census block (Abowd et al. (2022)). It is unclear how possessing these quantities would affect our disclosure risk measures, because they are not easily related to the counts that an adversary considers in our methods. Future work could provide a more formal analysis of these points, examining in more detail how much extra protection could be offered by the post-processing step (it is highly unlikely to be invertible) and how adversaries could use population invariants in their prior or data distributions.

This work also points to other directions for future research. One involves relaxing the assumption that the adversary possesses complete information about all but one individual. For example, it may be possible to adapt methods used in McClure and Reiter (2016) to examine inferences when the adversary possesses information about all but two—or all but n—individuals. Another direction involves using disclosure risk measures to approximate an "empirical" DP bound for a data set released under DP or zCDP. For the 2020 decennial census, the Census Bureau quotes a total $\varepsilon = 17.91$, computed by composing the ρ allocations at the six levels of the hierarchy and using Theorem 1. Our results indicate that reporting the privacy guarantee in this way may understate the degree of privacy. Other works, for example the partial DP of Ghazi et al. (2022), examine how to produce a more meaningful parameter for interpretation when the ε from DP is large. A method based on disclosure risks also may be possible and useful in practice.

Supplementary Material

The online Supplementary Material contains a proposition on the sums of discrete Gaussians, a derivation of the full conditionals, and an analysis of prior sensitivity. It also contains extensions to the methods and analysis that incorporate other attack scenarios, post-processing, and sequential releases.

Acknowledgments

This research was supported by NSF grant SES-2217456.

References

Abowd, J. M. (2018). The U.S. Census Bureau adopts differential privacy. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, KDD '18, New York, NY, USA, pp. 2867. Association for Computing Machinery.

Abowd, J. M., R. Ashmead, R. Cumings-Menon, S. Garfinkel, M. Heineck, C. Heiss, R. Johns, D. Kifer, P. Leclerc, A. Machanavajjhala, et al. (2022). The 2020 census disclosure avoidance system TopDown algorithm. *arXiv preprint arXiv:2204.08986*.

Abowd, J. M. and L. Vilhuber (2008). How protective are synthetic data? In *International Conference on Privacy in Statistical Databases*, pp. 239–246. Springer.

Bun, M. and T. Steinke (2016). Concentrated differential privacy: Simplifications, extensions, and lower bounds. In *Theory of Cryptography Conference*, pp. 635–658. Springer.

Canonne, C. L., G. Kamath, and T. Steinke (2020). The discrete Gaussian for differential privacy. *arXiv preprint arXiv:2004.00010*.

Domingo-Ferrer, J. and V. Torra (2004). Disclosure risk assessment in statistical data protection. *Journal of Computational and Applied Mathematics 164*, 285–293.

Duncan, G. and S. Keller-McNulty (2000). Bayesian insights on disclosure limitation: Mask or impute? In *Sixth World Meeting of the International Society for Bayesian Analysis*.

Duncan, G. T. and D. Lambert (1986). Disclosure-limited data dissemination. *Journal of the American Statistical Association 81*(393), 10–18.

Dwork, C. (2008). Differential privacy: A survey of results. In *International Conference on Theory and Applications of Models of Computation*, pp. 1–19. Springer.

Dwork, C., R. Greenwood, and G. King (2021). There's a simple solution to the latest census fight. *Boston Globe*.

Dwork, C., F. McSherry, K. Nissim, and A. Smith (2006). Calibrating noise to sensitivity in private data analysis. In *Theory of Cryptography Conference*, pp. 265–284. Springer.

Fienberg, S. E. and A. P. Sanil (1997). A Bayesian approach to data disclosure: Optimal intruder behavior for continuous data. *Journal of Official Statistics 13*(1), 75.

Ghazi, B., R. Kumar, P. Manurangsi, and T. Steinke (2022). Living with large ε via partial differential privacy. TPDP 2022.

Gong, R. and X.-L. Meng (2020). Congenial differential privacy under mandated disclosure. In *Proceedings of the 2020 ACM-IMS on Foundations of Data Science Conference*, pp. 59–70.

Hotz, V. J., C. R. Bollinger, T. Komarova, C. F. Manski, R. A. Moffitt, D. Nekipelov, A. Sojourner, and B. D. Spencer (2022). Balancing data privacy and usability in the federal statistical system. *Proceedings of the National Academy of Sciences 119*(31), e2104906119.

JASON (2022). Consistency of data products and formal privacy methods for the 2020 census.

Technical report, The MITRE Corporation.

Kenny, C. T., S. Kuriwaki, C. McCartan, E. Rosenman, T. Simko, and K. Imai (2021). The impact of the US census disclosure avoidance system on redistricting and voting rights analysis. *arXiv preprint arXiv:2105.14197*.

Lee, J. and C. Clifton (2011). How much is enough? choosing ε for differential privacy. In *International Conference on Information Security*, pp. 325–340. Springer.

McClure, D. and J. P. Reiter (2012). Differential privacy and statistical disclosure risk measures: An investigation with binary synthetic data. *Transactions on Data Privacy 5*(3), 535–552.

McClure, D. and J. P. Reiter (2016). Assessing disclosure risks for synthetic data with arbitrary intruder knowledge. *Statistical Journal of the IAOS 32*(1), 109–126.

National Archives and Records Administration (2022). Enumeration district and related maps, 1880 - 1990. `https://catalog.archives.gov/id/821491`. Accessed: May 27, 2022.

Reiter, J. P. (2005). Estimating risks of identification disclosure in microdata. *Journal of the American Statistical Association 100*(472), 1103–1112.

Ruggles, S., C. Fitch, D. Magnuson, and J. Schroeder (2019). Differential privacy and census data: Implications for social and economic research. In *AEA Papers and Proceedings*, Volume 109, pp. 403–08.

Seeman, J., A. Slavkovic, and M. Reimherr (2020). Private posterior inference consistent with public information: A case study in small area estimation from synthetic census data. In

International Conference on Privacy in Statistical Databases, pp. 323–336. Springer.

Zeki Kazan, Duke University

E-mail: zekican.kazan@duke.edu

Jerome Reiter, Duke University

E-mail: jreiter@duke.edu