DIFFERENTIAL PRIVACY GUARANTEES IN SMALL AREA ESTIMATION

Soumojit Das[1,2] *and Jörg Drechsler*[1,3]

[1] University of Maryland, College Park, MD [2]Washington State University, Pullman, WA
[3]Institute for Employment Research, Germany

Motivation. Statistical agencies face growing challenges protecting respondent confidentiality while releasing survey estimates. The database reconstruction theorem (Dinur and Nissim, 2003) shows that simple heuristics (e.g., minimum cell sizes) cannot guarantee privacy, while differential privacy (DP) often requires noise injection that degrades utility (Barrientos et al., 2024). This tension is particularly acute for small area estimation (SAE), where hierarchical models borrow strength across areas to produce reliable estimates for sparsely sampled subpopulations. Staff at U.S. Census Bureau research data centers currently struggle to assess disclosure risks of SAE outputs, as no formal privacy framework exists for these methods.

Research Question. Does the Bayesian Fay-Herriot model (Fay III and Herriot, 1979)—the most widely used area-level SAE method—provide inherent differential privacy guarantees through posterior sampling, without requiring additional noise injection?

Contributions. (1) We provide the *first formal DP analysis* of Bayesian SAE, deriving closed-form privacy bounds as functions of model and survey design parameters. (2) We show pure ε-DP cannot be achieved due to unbounded Gaussian support, but derive meaningful bounds under zero-concentrated DP (zCDP) (Bun and Steinke, 2016) and Rényi DP (Mironov, 2017). (3) We validate empirically using ACS microdata (ground-truth parameters) and SAIPE county poverty estimates (policy relevance).

Methodology. Under the Fay-Herriot model, the direct survey estimate y_i for area i has sampling variance $\sigma^2_{y,i}$, and the true area mean follows $\theta_i = \mathbf{x}_i^\top \boldsymbol{\beta} + v_i$ with $v_i \sim N(0, \sigma_v^2)$. The posterior distribution is Gaussian: $\theta_i \mid \mathcal{D} \sim N(\mu_i, \sigma_i^2)$, where $\mu_i = (1 - B_i)y_i + B_i \mathbf{x}_i^\top \hat{\boldsymbol{\beta}}$ and $B_i = \sigma^2_{y,i}/(\sigma^2_{y,i} + \sigma_v^2)$ is the shrinkage factor (exact form of the posterior variance is omitted for brevity). For neighboring databases differing in one record (samples within one area), the expected privacy loss is bounded: $D_{\mathrm{KL}}(P_{\mathcal{D}} \| P_{\mathcal{D}'}) \leq S_y^2/(2\sigma^2_{\min})$, where $S_y = \max_{i,r}(w_{ir} R_y / N_i)$ is the sensitivity depending on survey weights w_{ir}, outcome range R_y, and population size N_i. This yields ρ-zCDP with $\rho = S_y^2/(2\sigma^2_{\min})$, convertible to (ε, δ)-DP via $\varepsilon = \rho + 2\sqrt{\rho \log(1/\delta)}$.

Results. Figure 1 shows the key finding: privacy scales as $\varepsilon \propto 1/n_i$–larger sample sizes provide stronger inherent protection. Table 1 summarizes privacy parameters for binary outcomes under realistic survey configurations. For $n_i = 100$ samples per area and variance ratio $\sigma_v^2/\sigma_y^2 = 1$, we obtain per-area $\varepsilon = 0.068$ at $\delta = 10^{-5}$, competitive with purpose-built DP mechanisms. When releasing estimates for m areas, zCDP composition yields $\varepsilon \propto \sqrt{m}$, substantially tighter than naive ε-DP composition.

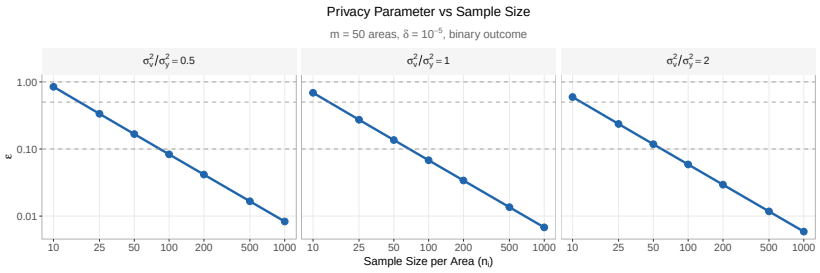

Figure 1. (Simulation Study) Privacy parameter ε versus sample size for binary outcomes ($\delta = 10^{-5}$). The log-log relationship confirms $\varepsilon \propto 1/n_i$.

Table 1. Privacy parameters for binary outcomes ($\delta = 10^{-5}$, variance ratio $\sigma_v^2/\sigma_y^2 = 1$).

Sample Size (n_i)	ρ-zCDP	ε (per area)	$\varepsilon_{\text{composed}}$ (100 areas)
10	1.0×10^{-2}	0.69	7.79
100	1.0×10^{-4}	0.068	0.69
500	4.0×10^{-6}	0.014	0.14

Empirical Validation. Analysis of 2022 ACS PUMS (3.19M person records, 2,462 PUMAs) confirms the theoretical framework with exact ground-truth parameters: median per-area $\varepsilon = 0.23$, with 99.9% of PUMAs achieving $\varepsilon < 1$. Application to SAIPE county poverty estimates (3,143 counties; Figure 2) yields median $\varepsilon = 0.77$, with 89.6% of counties satisfying $\varepsilon < 1$. The bounds are robust to variance estimation errors: 20% error in σ_v^2 produces only 2.5% change in privacy parameters when shrinkage is low.

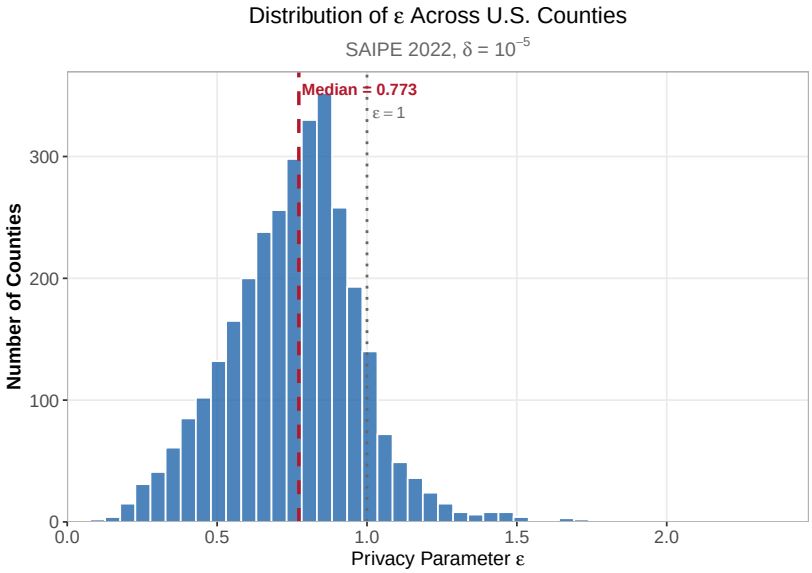

Figure 2. Distribution of ε across 3,143 U.S. counties from SAIPE 2022 ($\delta = 10^{-5}$). Median $\varepsilon = 0.77$; 89.6% of counties achieve $\varepsilon < 1$.

Implications. Our results suggest that routine SAE outputs may satisfy meaningful privacy standards without modification. This has immediate practical relevance: agencies can quantify inherent privacy protection for existing releases, identify areas requiring additional scrutiny (those with small samples or extreme weights), and integrate SAE outputs into formal privacy accounting frameworks. The *shrinkage-privacy duality*–areas with high uncertainty benefit from both Bayesian hierarchical borrowing and privacy protection. While composed ε for all areas can be substantial ($\varepsilon \approx 18$ for 2,462 PUMAs), sublinear zCDP composition ($\varepsilon \propto \sqrt{m}$) provides meaningful individual protection.

Broader Context. Our work complements recent findings that adding DP noise to regression destroys utility at meaningful privacy levels (Barrientos et al., 2024). In contrast, we show that Bayesian SAE provides "privacy for free" (Wang et al., 2015)—the posterior sampling mechanism inherently limits information leakage without sacrificing the utility gains from hierarchical modeling. This positions SAE favorably for agencies seeking to balance privacy protection with the production of reliable small area statistics for policy applications such as federal funding allocation.

References

Barrientos, A. F., A. R. Williams, J. Snoke, and C. M. Bowen (2024). A feasibility study of differentially private summary statistics and regression analyses with evaluations on administrative and survey data. *Journal of the American Statistical Association 119*(545), 52–65.

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

Dinur, I. and K. Nissim (2003). Revealing information while preserving privacy. In *Proceedings of the twenty-second ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, pp. 202–210.

Fay III, R. E. and R. A. Herriot (1979). Estimates of income for small places: an application of james-stein procedures to census data. *Journal of the American Statistical Association 74*(366a), 269–277.

Mironov, I. (2017). Rényi differential privacy. In *2017 IEEE 30th computer security foundations symposium (CSF)*, pp. 263–275. IEEE.

Wang, Y.-X., S. Fienberg, and A. Smola (2015). Privacy for free: Posterior sampling and stochastic gradient monte carlo. In *International Conference on Machine Learning*, pp. 2493–2502. PMLR.