FROM ERASURE TO LEAKAGE: NONRESPONSE UNDER LOCAL DIFFERENTIAL PRIVACY

Anonymous authors
Paper under double-blind review

ABSTRACT

Nonresponse is often treated simply as data loss, but under local differential privacy (LDP), its impact depends on the response mechanism. We study categorical distribution estimation with an explicit nonresponse symbol and distinguish two cases. When nonresponse is independent of the private category, the mechanism is exactly k-ary randomized response followed by erasure. This gives a simple statistical interpretation: several information measures contract by $1-\delta$, and the squared risk of the raw inverse estimator is inflated asymptotically by $1/(1-\delta)$. The situation is different when response probabilities vary across private categories. The decision to respond then carries information about the private value and also changes the composition of the observed sample. We derive the exact privacy cost of this effect and characterize the resulting selection bias. If the true response propensities are known, the population distribution can be recovered with an unbiased estimator. If they are latent, the released distribution generally does not identify the population separately from the response process. Experiments on synthetic distributions and empirical categorical distributions from four UCI datasets closely match the theoretical predictions. They also show that two mechanisms with the same average nonresponse rate can have very different privacy and estimation costs.

1 INTRODUCTION

Local differential privacy (LDP) protects each user before their data reaches the analyst. Randomized response is the canonical example: instead of reporting a sensitive value directly, the user releases a randomized version whose likelihood ratio is bounded across possible inputs (Warner, 1965; Duchi et al., 2013). For categorical data, this basic idea has led to a large family of frequency and distribution estimators, including optimized unary encoding, local hashing, subset mechanisms, and Hadamard response (Wang et al., 2017; Ye & Barg, 2018; Acharya et al., 2019; Cormode et al., 2021).

One assumption is often left implicit: every user eventually produces an informative privatized report. Real data collection is less tidy. A user may refuse, abstain, or answer "don't know". In evidential LDP, for example, an explicit ignorance outcome is part of the released alphabet (Li et al., 2022). The presence of such an outcome raises a basic question: *is nonresponse merely lost information, or can the act of responding itself become informative?*

This distinction also matters in deployed local privacy systems. Methods such as RAPPOR and private telemetry place local randomization inside a broader data collection process (Erlingsson et al., 2014; Ding et al., 2017). It is therefore useful to separate the decision to produce a report from the mechanism that privatizes its contents.

The answer depends on the response mechanism. Suppose first that every private category has the same nonresponse probability δ. Then nonresponse is statistically simple: run ordinary k-ary randomized response and erase the result independently with probability δ. This exact factorization has a strong consequence. Because the nonresponse channel is a postprocessing of ordinary randomized response, it is Blackwell less informative (Blackwell, 1953). Independent nonresponse therefore cannot improve the optimal risk of any downstream statistical decision problem at the same privacy level. It only removes usable information.

The statistical cost is equally transparent. Independent nonresponse removes an informative report with probability δ, so pairwise f divergences, mutual information, and Fisher information all contract

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

by $1 - \delta$. For categorical distribution estimation, we derive the exact finite sample squared ℓ_2 risk of the raw inverse estimator. For any fixed $\delta \in [0, 1)$,

$$\frac{R_{n,\delta}}{R_{n,0}} \longrightarrow \frac{1}{1-\delta}. \tag{1}$$

Thus $\delta = 0.1$ requires asymptotically about 11.1% more users to attain the same squared error, while $\delta = 0.2$ requires about 25% more. This is the familiar loss of effective sample size in a particularly simple categorical LDP channel. Related minimax results for locally private estimation with missing data were studied by Ono (2025).

A different phenomenon appears when the response probability depends on the private category. Let δ_i be the probability of nonresponse for category i. Now observing only whether a report was returned changes the posterior belief about the private value. The response pattern becomes a privacy channel in its own right. We derive its exact contribution to the LDP parameter and show that heterogeneous response probabilities consume part of a fixed privacy budget. If the informative randomized response step already uses the full budget, any nonconstant vector $(\delta_1, \ldots, \delta_k)$ necessarily raises total privacy loss.

The same dependence also creates a selection problem. Among respondents, the private categories are distributed according to $\pi_i(1 - \delta_i)$ after renormalization, rather than according to the population distribution π_i. Dropping nonresponses and applying an ordinary randomized-response inverse therefore targets the respondent population. When the response propensities are known, we derive an unbiased correction and its exact covariance. Externally estimated or calibrated propensities can instead be used in a plug-in correction, whose accuracy depends on calibration error. When the propensities are latent, the released law generally identifies only the products $\pi_i(1 - \delta_i)$, so the original population need not be identifiable. The main point is therefore not another estimator for missing data, but that the same observable response process can simultaneously change the privacy guarantee, the population represented by the reports, and what can be identified from those reports.

Contributions. For category independent nonresponse, we show that the mechanism is exactly k-ary randomized response followed by erasure. This yields a Blackwell ordering, exact information contraction identities, and an exact finite sample risk expression for the raw inverse estimator.

The main contribution concerns category dependent response. We derive its exact LDP cost, characterize the resulting respondent selection, and determine what can be recovered from the released distribution. Known response propensities give an unbiased estimator with exact covariance, whereas latent propensities leave only $w_i = \pi_i(1 - \delta_i)$ identifiable in general. The experiments test these channel specific predictions on synthetic and empirical categorical distributions.

2 Related Work

Local privacy and categorical estimation. Randomized response was introduced for sensitive surveys by Warner (1965). Modern LDP theory formalized the statistical cost of local privacy and established minimax rates for private estimation (Duchi et al., 2013). For discrete distribution estimation, Kairouz et al. (2016) analyzed k-RR and RAPPOR across privacy regimes, while Wang et al. (2017) studied practical frequency-estimation protocols including unary encoding and local hashing. Ye & Barg (2018) characterized optimal schemes for discrete distribution estimation, and Acharya et al. (2019) developed Hadamard response. Broader comparisons and information-theoretic analyses appear in Cormode et al. (2021); Pastore & Gastpar (2021). These results are also why we do not claim global optimality for k-RR: it is the base channel for which the nonresponse mechanism considered here admits a particularly transparent exact characterization.

Missingness and nonresponse under privacy. Missing or withheld reports have appeared in several privacy settings. Sun et al. (2018) treat NULL as part of the randomized answer domain in sparse crowdsourcing and note that participation patterns can themselves be sensitive. Sun et al. (2020) study missing data under LDP, while Ono (2025) analyze locally private estimation under several missingness mechanisms and derive minimax consequences. More recently, Hu & Liu (2026) study censoring under asymmetric local privacy.

Randomized response and nonresponse also have a longer history in survey sampling; for example, Barabesi et al. (2014) study estimation and variance correction when randomized response is combined with nonresponse. These works establish that missingness can affect both information and inference. Survey methodology provides a useful statistical perspective on the same issue. Response rate alone is not a reliable measure of nonresponse bias; what matters is whether the propensity to respond is associated with the quantity being measured (Groves, 2006; Groves & Peytcheva, 2008). Self-selection can therefore distort estimates in web surveys (Bethlehem, 2010), and privacy concerns themselves have been associated with participation decisions (Singer et al., 2003).

These results are close in spirit to the distinction studied here, but they do not account for the additional randomization introduced by an LDP mechanism. Our focus is the interaction between the response process and the privacy channel.

Explicit ignorance under LDP. The closest predecessor in terms of the released alphabet is Li et al. (2022), who include an explicit "don't know" outcome in evidential LDP. For a generalized binary Warner mechanism, they derive an unbiased estimator and a conditional variance expression and show that more "don't know" mass worsens the privacy and utility tradeoff within that family.

Our independent response analysis should therefore not be read as the first use of a nonresponse outcome under LDP. What we add is the categorical channel interpretation: independent nonresponse is exactly an erasure of k-RR, which yields a Blackwell ordering, several exact information identities, and a finite sample risk expression. The category dependent case then goes further by making the response process itself part of the privacy and selection problem.

Selection and identifiability. The distinction between ignorable and nonignorable missingness, and the need for additional assumptions under MNAR-type mechanisms, are classical in the missing-data literature (Little & Rubin, 2019). Our identifiability claim is therefore channel specific: for the observable category-dependent nonresponse channel in Eq. (5), the released distribution identifies the products $w_i = \pi_i(1 - \delta_i)$ when $\alpha > 0$, but generally cannot separate prevalence from latent response propensity. Combined with the exact LDP calculation, this yields a joint privacy–selection–identification characterization for this categorical channel.

Our analysis is deliberately restricted to singleton informative outputs plus one complete-ignorance symbol. Arbitrary belief-function focal sets, unobserved dropout from an unknown sampling frame, and general optimal LDP mechanism design are outside its scope.

3 Problem Setup

Let $X \in [k] = \{1, \dots, k\}$ denote a private categorical variable with population distribution $\boldsymbol{\pi} = (\pi_1, \dots, \pi_k) \in \Delta_k$, where

$$\Delta_k = \left\{ \boldsymbol{p} \in \mathbb{R}^k_{\geq 0} : \sum_i p_i = 1 \right\}.$$

Each sampled user releases one symbol in $\mathcal{Y} = [k] \cup \{\Omega\}$. The symbol Ω records an explicit nonresponse. This observation model is important. We assume that the analyst knows that a sampled user produced a report, even when that report is Ω. We do not consider users who disappear from the sampling frame without leaving any observable record.

Definition 1 (Local differential privacy)**.** A channel Q is ε-locally differentially private if, for every $x, x' \in [k]$ and every output event A,

$$Q(A \mid x) \leq e^{\varepsilon} Q(A \mid x'). \tag{2}$$

We use the inequality form because it remains well defined when some outputs have zero probability.

As the informative channel, we use k-ary randomized response. For $\alpha \geq 0$,

$$K_\alpha(j \mid i) = \begin{cases} p_\alpha, & j = i, \\ r_\alpha, & j \neq i, \end{cases} \qquad p_\alpha = \frac{e^\alpha}{e^\alpha + k - 1}, \quad r_\alpha = \frac{1}{e^\alpha + k - 1}. \tag{3}$$

It is convenient to write

$$\gamma_\alpha = p_\alpha - r_\alpha = \frac{e^\alpha - 1}{e^\alpha + k - 1}. \tag{4}$$

The response process is modeled separately from randomized response. For a user whose private category is i, let $a_i = 1-\delta_i$ be the probability that the user provides an informative report. Conditional on responding, the report is generated by K_α. Otherwise the released symbol is Ω:

$$Q_{\alpha,\boldsymbol{\delta}}(y \mid i) = \begin{cases} a_i K_\alpha(y \mid i), & y \in [k], \\ \delta_i, & y = \Omega. \end{cases} \tag{5}$$

The model separates response behavior from the privacy mechanism. This matters because the a_i play different statistical roles depending on what the analyst knows about them. They may be fixed by design, estimated from external data, or completely latent. The estimator in Section 5 is exactly unbiased only when the true a_i are known; estimated propensities lead instead to a plug-in correction.

4 Category Independent Nonresponse

We begin with the case $\delta_i = \delta$ for every category. This is the cleanest setting because the decision to respond carries no information about X. When the informative channel uses privacy parameter ε, we denote the resulting mechanism by $Q_{\varepsilon,\delta}$.

4.1 Erasure and Blackwell Ordering

Let T_δ preserve a singleton output with probability $1-\delta$ and replace it by Ω with probability δ.

Theorem 1 (Erasure representation)**.** *For $k \geq 2$, $\varepsilon \geq 0$, and $0 \leq \delta \leq 1$,*

$$Q_{\varepsilon,\delta} = T_\delta K_\varepsilon. \tag{6}$$

The value of this representation is that it tells us what nonresponse can and cannot do in this regime. The mechanism first produces the same information as ordinary k-RR and then discards part of it independently. Nonresponse therefore does not create a new source of privacy protection or a new statistical signal. It only removes observations that would otherwise have been available.

Corollary 1 (Blackwell dominance)**.** *For any number of independent users, any prior on the private data, any action space, and any loss function, the minimum Bayes risk achievable from $Q_{\varepsilon,\delta}$ observations cannot be smaller than that achievable from K_ε observations.*

This is stronger than a comparison between two particular estimators. A projected estimator may occasionally have a smaller realized error after nonresponse because of finite sample randomness, but no decision rule can gain information from the erasure itself. The Blackwell ordering therefore gives a clean benchmark for interpreting the experiments later in the paper.

4.2 Exact Information Contraction

The erasure representation also shows how much information is lost. Let K_i and Q_i denote the conditional output laws under input i. For distinct i, j,

$$D_{\mathrm{TV}}(K_i, K_j) = \gamma_\varepsilon, \qquad D_{\mathrm{TV}}(Q_i, Q_j) = (1-\delta)\gamma_\varepsilon. \tag{7}$$

More generally, for any f divergence for which the expressions are defined,

$$D_f(Q_i \| Q_j) = (1-\delta) D_f(K_i \| K_j). \tag{8}$$

For KL divergence,

$$D_{\mathrm{KL}}(Q_i \| Q_j) = (1-\delta)\varepsilon\gamma_\varepsilon. \tag{9}$$

If Z is the k-RR output and Y is the released output after possible nonresponse, then for any prior on X,

$$I(X;Y) = (1-\delta) I(X;Z), \tag{10}$$

and, for any regular interior parameterization of the population distribution,

$$\mathcal{I}_{\mathrm{NR}} = (1-\delta)\mathcal{I}_{\mathrm{RR}}. \tag{11}$$

The repeated factor $1-\delta$ is not a coincidence. Every information measure above is seeing the same erasure event: with probability δ, an otherwise informative report is replaced by a symbol whose distribution does not depend on the private category.

4.3 Distribution Estimation and Exact Risk

The same structure gives a direct interpretation of estimation error. Let N_j be the number of singleton reports equal to j and let $M = \sum_j N_j$. Since each informative report is retained with probability $1 - \delta$,

$$M \sim \text{Binomial}(n, 1 - \delta).$$

Conditional on $M = m > 0$, the retained reports are simply an ordinary k-RR sample of size m. If $F_j = N_j/m$ and $\boldsymbol{u} = (1/k, \ldots, 1/k)$, then for $\varepsilon > 0$ the raw inverse is

$$\widetilde{\boldsymbol{\pi}} = \boldsymbol{u} + \frac{\boldsymbol{F} - \boldsymbol{u}}{\gamma_\varepsilon}. \tag{12}$$

It is conditionally unbiased. Define

$$C_{\varepsilon,k}(\boldsymbol{\pi}) = \frac{1 - 1/k}{\gamma_\varepsilon^2} - \|\boldsymbol{\pi} - \boldsymbol{u}\|_2^2. \tag{13}$$

Then

$$\mathbb{E}\big[\|\widetilde{\boldsymbol{\pi}} - \boldsymbol{\pi}\|_2^2 \mid M = m\big] = \frac{C_{\varepsilon,k}(\boldsymbol{\pi})}{m}. \tag{14}$$

Conditioning on M removes the effect of nonresponse entirely: one is left with an ordinary k-RR sample of size M. The statistical price is the random reduction from n users to M informative reports.

There is one finite sample detail that must still be handled. With probability δ^n, no informative report is observed. We define the raw estimator to be the uniform distribution on that event, $\widetilde{\boldsymbol{\pi}} = \boldsymbol{u}$. The resulting unconditional risk is

$$R_{n,\delta} = \delta^n \|\boldsymbol{\pi} - \boldsymbol{u}\|_2^2 + C_{\varepsilon,k}(\boldsymbol{\pi}) \sum_{m=1}^{n} \frac{1}{m} \binom{n}{m} (1 - \delta)^m \delta^{n-m}. \tag{15}$$

At $\delta = 0$, $R_{n,0} = C_{\varepsilon,k}(\boldsymbol{\pi})/n$. For every fixed $\delta < 1$,

$$R_{n,\delta} = \frac{C_{\varepsilon,k}(\boldsymbol{\pi})}{n(1 - \delta)} + O(n^{-2}), \tag{16}$$

which gives Eq. (1).

The asymptotic factor $1/(1 - \delta)$ therefore has a simple meaning: in this regime, nonresponse behaves like losing a fraction δ of the sample. It does not introduce a new bias in the target distribution. Projection onto the simplex can only reduce Euclidean error pointwise, but the exact expression above is for the raw estimator with the stated fallback.

5 Category Dependent Nonresponse

The situation changes when the response probability depends on the private category. A user who responds and a user who does not respond are no longer drawn from the same mixture of categories. Since Ω is observed, the response decision itself now carries information about X. This creates two effects at once: it changes the privacy accounting and it changes the population represented by the informative reports.

5.1 Exact Privacy Cost

Theorem 2 (Exact privacy cost)**.** *Assume $k \geq 2$, $\alpha \geq 0$, and $0 < \delta_i < 1$ for all i. Let $\delta_{\min} = \min_i \delta_i$ and $\delta_{\max} = \max_i \delta_i$. The exact LDP parameter of $Q_{\alpha,\boldsymbol{\delta}}$ is*

$$\varepsilon_Q = \max\left\{\log \frac{\delta_{\max}}{\delta_{\min}},\ \alpha + \log \frac{1 - \delta_{\min}}{1 - \delta_{\max}}\right\}. \tag{17}$$

The two terms correspond to two different ways of learning about the private category. The first comes from observing Ω: if some categories refuse more often than others, nonresponse alone changes the likelihood ratio. The second comes from an informative report, where the usual k-RR privacy loss is combined with the fact that some categories are more likely to respond.

This distinction matters in practice because a privacy budget that is safe for the content channel alone may no longer be safe once the response behavior is included.

Corollary 2 (No free heterogeneous nonresponse)**.** *If the informative channel already uses the full target budget, $\alpha = \varepsilon$, then $Q_{\alpha,\boldsymbol{\delta}}$ can remain ε-LDP only when all δ_i are equal.*

If the informative channel already uses the full privacy budget, any heterogeneity in response probabilities necessarily increases the overall privacy loss. Maintaining the same target ε therefore requires a smaller privacy parameter for the informative channel. Define

$$\tau_{\mathrm{resp}} = \log \frac{1-\delta_{\min}}{1-\delta_{\max}}. \tag{18}$$

Provided that $\log(\delta_{\max}/\delta_{\min}) \le \varepsilon$ and $\varepsilon - \tau_{\mathrm{resp}} \ge 0$, the largest feasible informative parameter is

$$\alpha^\star = \varepsilon - \tau_{\mathrm{resp}}. \tag{19}$$

We refer to this reduction as the response tax. It measures how much of the available privacy budget is absorbed by category dependent response behavior.

5.2 Information Carried by the Response Decision

The exact privacy parameter gives the worst case likelihood ratio. A complementary question is how the information in the released observation is split between the decision to respond and the content that is reported. Let $R = \mathbf{1}\{Y \neq \Omega\}$ and $a_i = 1 - \delta_i$.

Theorem 3 (KL information decomposition)**.** *For private categories i, j,*

$$D_{\mathrm{KL}}(Q_i \| Q_j) = D_{\mathrm{KL}}(\mathrm{Ber}(a_i) \| \mathrm{Ber}(a_j)) + a_i D_{\mathrm{KL}}(K_\alpha(\cdot \mid i) \| K_\alpha(\cdot \mid j)). \tag{20}$$

Let $\bar{a} = \sum_i \pi_i a_i$ and $\pi_i^{(R)} = \pi_i a_i / \bar{a}$. Then

$$I(X;Y) = I(X;R) + \bar{a}\, I_{\boldsymbol{\pi}^{(R)}}(X;Z), \tag{21}$$

where Z denotes the informative k-RR output conditional on $R = 1$.

These identities make the new source of information explicit. The response indicator contributes nothing when all categories respond at the same rate. Once the rates differ, part of the information about X is already present before the privatized content is examined. This is precisely what is absent in the independent regime.

5.3 Selection Bias and Known Propensity Correction

The same response behavior also changes the estimand seen by a responder only analysis. Conditioning on an informative report gives

$$\pi_i^{(R)} = \frac{\pi_i a_i}{\bar{a}}, \qquad \pi_i^{(R)} - \pi_i = \frac{\pi_i (a_i - \bar{a})}{\bar{a}}. \tag{22}$$

Thus, after removing Ω, the remaining users are generally not distributed according to $\boldsymbol{\pi}$. Applying the ordinary k-RR inverse to them is not merely noisy; it estimates a different population, $\boldsymbol{\pi}^{(R)}$.

If the response propensities are known, this selection can be corrected. Let

$$D_a = \mathrm{diag}(a_1, \ldots, a_k),$$

and let $\boldsymbol{s}$ denote the unconditional probabilities of the singleton outputs. Then

$$\boldsymbol{s} = K_\alpha^\top D_a \boldsymbol{\pi}. \tag{23}$$

For $\alpha > 0$ and $a_i > 0$, define

$$\widetilde{\boldsymbol{\pi}}_{\text{known}} = D_a^{-1}(K_\alpha^\top)^{-1}\widehat{\boldsymbol{s}}, \qquad \widehat{\boldsymbol{s}} = \frac{1}{n}(N_1, \ldots, N_k)^\top. \tag{24}$$

This estimator is unbiased and has exact covariance

$$\mathrm{Cov}(\widetilde{\boldsymbol{\pi}}_{\text{known}}) = \frac{1}{n} D_a^{-1}(K_\alpha^\top)^{-1}[\mathrm{Diag}(\boldsymbol{s}) - \boldsymbol{s}\boldsymbol{s}^\top]K_\alpha^{-1}D_a^{-1}. \tag{25}$$

Its exact squared ℓ_2 risk is the trace of Eq. (25) and therefore scales as $1/n$.

Known propensities resolve a basic ambiguity: a category may appear rare because it is genuinely uncommon, or because its members respond less often. The correction separates these two effects.

5.4 What If the Propensities Are Unknown?

Let $w_i = \pi_i a_i$. For $\alpha > 0$, the released law identifies $\boldsymbol{w}$ because K_α is invertible. The data therefore reveal the amount of observed mass contributed by each category, but they do not reveal how much of that mass comes from prevalence and how much comes from willingness to respond.

Theorem 4 (Nonidentifiability)**.** *Assume $k \geq 2$ and $\alpha > 0$. If category dependent response propensities are latent, the output law identifies $w_i = \pi_i(1 - \delta_i)$ but does not, in general, identify $\boldsymbol{\pi}$ and $\boldsymbol{a}$ separately.*

For example, any alternative probability vector $\boldsymbol{\pi}'$ with $\pi'_i \geq w_i$ can reproduce the same $\boldsymbol{w}$ by taking $a'_i = w_i/\pi'_i$ whenever $\pi'_i > 0$, and $a'_i = 0$ when $\pi'_i = w_i = 0$. Under interior conditions, one can choose $\pi'_i > w_i$ for every i, so the same ambiguity remains even when all response probabilities lie strictly between zero and one.

The ambiguity is structural rather than a finite sample effect. Even with unlimited data, distinct pairs $(\boldsymbol{\pi}, \boldsymbol{a})$ can induce the same released distribution. Recovering $\boldsymbol{\pi}$ therefore requires information outside the LDP reports, such as known propensities, calibration data, or assumptions on the response process.

6 Experiments

The experiments test the exact identities and illustrate the finite sample behavior of the two response regimes. Complete repetition counts, random seeds, and preprocessing details are given in Appendix B.

6.1 Exact Identities and Independent-Nonresponse Risk

The analytical identities are recovered numerically to floating point precision over the validation grid. The exact finite sample risk also agrees closely with Monte Carlo simulation. At $\delta = 0.1$, for example, the empirical raw risk is 0.002716, compared with the exact value 0.002700. Full numerical checks and the corresponding risk curve are reported in Appendix Table 2 and Figure 2.

6.2 Synthetic Benchmark

Table 1 reports projected ℓ_1 error at $\varepsilon = 1$. Independent nonresponse does not systematically beat k-RR. For $\delta = 0.1$, paired 95% intervals for the difference from k-RR contain zero for the binary and Zipf-5 settings, while Sparse-10 and Zipf-20 show positive penalties. At larger alphabets, OUE and OLH can outperform both k-RR and its nonresponse extension. This is consistent with our scope: the theory characterizes the cost of adding nonresponse to k-RR rather than claiming global optimality among LDP mechanisms.

Table 1: Mean projected ℓ_1 error on synthetic distributions at $\varepsilon = 1$. Lower is better.

Scenario	k-RR	OUE	OLH	HR-C1C2	NR $\delta = .1$	NR $\delta = .2$
Binary skew $(0.7, 0.3)$	0.025162	0.031934	0.034979	0.037904	0.023529	0.027746
Zipf $k = 5$	0.092063	0.100788	0.103609	0.104706	0.091075	0.095736
Sparse $k = 10$	0.211500	0.191884	0.192346	0.204724	0.222266	0.230053
Zipf $k = 20$	0.379524	0.280030	0.283734	0.291678	0.402545	0.422936

6.3 The Privacy Tax of Heterogeneous Response

We next hold the population-average nonresponse rate at 0.2 while increasing a heterogeneity amplitude h. The construction keeps $\sum_i \pi_i \delta_i = 0.2$ exactly; its formula is given in Appendix B. If the informative channel is left at $\alpha = 1$, total privacy increases with h. At $h = 0.10$, the category-specific nonresponse probabilities range from 0.157662 to 0.30, and the total privacy loss becomes 1.185101. Enforcing overall $\varepsilon = 1$ requires $\alpha^\star = 0.814899$. At $h = 0.14$, the corresponding informative budget is 0.736153. The full trajectory is reported in Appendix Table 4 and Figure 3.

6.4 Selection, Correction, and Exact Risk

At $h = 0.10$, the naive responder only inverse closely tracks the responder distribution rather than the original population. In contrast, the known propensity estimator returns to the population target. The full category level values are reported in Appendix Table 5 and Figure 4.

The Monte Carlo risks also agree closely with the exact covariance prediction over sample sizes from 500 to $20{,}000$. The largest relative discrepancy among the reported configurations is below 0.4%, and nR_n remains constant, as predicted by the exact $1/n$ scaling.

6.5 Nonidentifiability and the Cost of Heterogeneity

The constructive example in Theorem 4 is directly visible in the released law. We use two population distributions whose ℓ_1 distance is 0.14 and choose different response propensities so that the product vector $\boldsymbol{w}$ is identical. Their privatized output laws differ by at most 1.11×10^{-16} numerically. The analyst therefore cannot distinguish the two populations from the released distribution alone when the propensities are latent.

Finally, Figure 1 holds both the average nonresponse rate (0.2) and total privacy budget $(\varepsilon = 1)$ fixed. Only the heterogeneity of response propensities changes. Along this controlled path, the exact known-propensity risk ratio rises to 1.741536 at $h = 0.10$ and 2.296712 at $h = 0.14$. These are configuration-specific values, not universal bounds. They show why an average nonresponse rate alone is insufficient to predict utility once response depends on the private category.

6.6 Real Categorical Data

The four UCI tasks show the same qualitative pattern. At $\delta = 0.1$, paired intervals include zero for Adult and Car Evaluation, whereas Mushroom and Nursery show positive penalties. The small negative mean on Adult is consistent with finite sample variation after projection, not an information gain. Full results and paired intervals are reported in Appendix Tables 8 and 12.

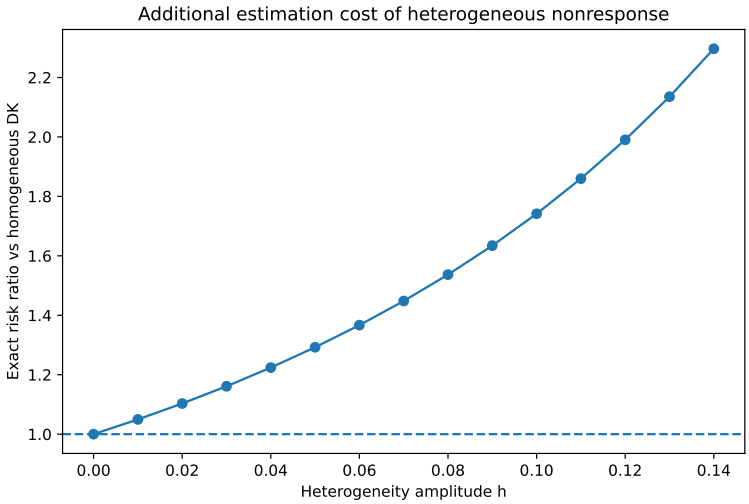

Figure 1: Same average nonresponse, different statistical cost. The average nonresponse rate and total privacy budget remain fixed, while increasing category dependence raises the exact known-propensity risk along the controlled heterogeneity trajectory.

7 Discussion and Limitations

From a practical perspective, an overall participation or nonresponse rate may obscure response patterns that are critical for inference. When response propensity depends on the private category, the response process affects both privacy accounting and the composition of the observed sample. Accordingly, estimating or externally calibrating category-specific response propensities may be as important as specifying the privacy mechanism itself.

Our analysis is limited to categorical data, k-RR as the informative channel, singleton reports, and an explicit nonresponse symbol observed by the analyst. Extensions to unobserved dropout, richer evidential outputs, multiple missingness processes, and optimal design over general LDP mechanisms are left for future work.

8 Conclusion

Independent nonresponse acts as an erasure mechanism, reducing statistical information without otherwise altering the structure of the observed data. Category-dependent response is fundamentally different: the response decision itself becomes informative, changes the composition of the respondent population, and may preclude separate identification of population prevalence and response propensities. Consequently, the average nonresponse rate alone is insufficient to characterize either the privacy guarantees or the statistical utility of the mechanism.

AI Use Statement

Generative AI tools were used during the preparation of this work for literature search assistance, proof checking, experimental design, code generation and debugging, and language editing. All theoretical claims, numerical results, citations, and manuscript text were reviewed and verified by the authors before submission. The authors take full responsibility for the submitted work.

Reproducibility Statement

The theorem assumptions and complete derivations are provided in the appendix. The experimental code uses fixed random seeds and exports run manifests, package versions, trial-level results, summary

tables, and publication figures. Appendix B records the synthetic configurations, Monte Carlo counts, confidence-interval procedure, heterogeneous-response construction, and real-data preprocessing used for the reported results. The UCI datasets supply empirical categorical distributions; the response and nonresponse mechanisms are simulated rather than observed in those datasets.

REFERENCES

Jayadev Acharya, Ziteng Sun, and Huanyu Zhang. Hadamard response: Estimating distributions privately, efficiently, and with little communication. In *Proceedings of the Twenty-Second International Conference on Artificial Intelligence and Statistics*, volume 89 of *Proceedings of Machine Learning Research*, pp. 1120–1129. PMLR, 2019.

Lucio Barabesi, Giancarlo Diana, and Pier Francesco Perri. Horvitz-thompson estimation with randomized response and nonresponse. *Model Assisted Statistics and Applications*, 9(1):3–10, 2014.

Barry Becker and Ronny Kohavi. Adult. UCI Machine Learning Repository, 1996. Dataset.

Jelke Bethlehem. Selection bias in web surveys. *International statistical review*, 78(2):161–188, 2010.

David Blackwell. Equivalent comparisons of experiments. *The Annals of Mathematical Statistics*, 24 (2):265–272, 1953. doi: 10.1214/aoms/1177729032.

Marko Bohanec. Car evaluation. UCI Machine Learning Repository, 1988. Dataset.

Graham Cormode, Samuel Maddock, and Carsten Maple. Frequency estimation under local differential privacy. *Proceedings of the VLDB Endowment*, 14(11):2046–2058, 2021. doi: 10.14778/3476249.3476261.

Bolin Ding, Janardhan Kulkarni, and Sergey Yekhanin. Collecting telemetry data privately. *Advances in neural information processing systems*, 30, 2017.

John C. Duchi, Martin J. Wainwright, and Michael I. Jordan. Local privacy and minimax bounds: Sharp rates for probability estimation. In *Advances in Neural Information Processing Systems*, volume 26, pp. 1529–1537, 2013.

Úlfar Erlingsson, Vasyl Pihur, and Aleksandra Korolova. Rappor: Randomized aggregatable privacy-preserving ordinal response. In *Proceedings of the 2014 ACM SIGSAC conference on computer and communications security*, pp. 1054–1067, 2014.

Robert M. Groves. Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly*, 70(5):646–675, 01 2006. ISSN 0033-362X. doi: 10.1093/poq/nfl033. URL `https://doi.org/10.1093/poq/nfl033`.

Robert M Groves and Emilia Peytcheva. The impact of nonresponse rates on nonresponse bias: a meta-analysis. *Public opinion quarterly*, 72(2):167–189, 2008.

Qirui Hu and Yi Liu. Censoring with plausible deniability: Asymmetric local privacy for multi-category cdf estimation. In *Proceedings of the 43rd International Conference on Machine Learning*, 2026. Accepted poster.

Peter Kairouz, Keith Bonawitz, and Daniel Ramage. Discrete distribution estimation under local privacy. In *Proceedings of the 33rd International Conference on Machine Learning*, volume 48 of *Proceedings of Machine Learning Research*, pp. 2436–2444. PMLR, 2016.

Qiyu Li, Chunlai Zhou, Biao Qin, and Zhiqiang Xu. Local differential privacy for belief functions. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pp. 10025–10033, 2022. doi: 10.1609/aaai.v36i9.21241.

Roderick J. A. Little and Donald B. Rubin. *Statistical Analysis with Missing Data*. Wiley, 3 edition, 2019. doi: 10.1002/9781119482260.

Hajime Ono. Theoretical lower bounds for one-dimensional locally private estimations with missing data: Proofs with three types of missing mechanisms. *International Journal of Information Security*, 24(1), 2025. doi: 10.1007/s10207-024-00973-2. Article 58.

Adriano Pastore and Michael Gastpar. Locally differentially-private randomized response for discrete distribution learning. *Journal of Machine Learning Research*, 22(132):1–56, 2021.

Vladislav Rajkovic. Nursery. UCI Machine Learning Repository, 1989. Dataset.

Eleanor Singer, John Van Hoewyk, and Randall J Neugebauer. Attitudes and behavior: The impact of privacy and confidentiality concerns on participation in the 2000 census. *Public Opinion Quarterly*, 67(3):368–384, 2003.

Haipei Sun, Boxiang Dong, Hui Wendy Wang, Ting Yu, and Zhan Qin. Truth inference on sparse crowdsourcing data with local differential privacy. In *2018 IEEE International Conference on Big Data (Big Data)*, pp. 488–497. IEEE, 2018. doi: 10.1109/BigData.2018.8622635.

Lin Sun, Xiaojun Ye, Jun Zhao, Chenhui Lu, and Mengmeng Yang. Bisample: Bidirectional sampling for handling missing data with local differential privacy. In *Database Systems for Advanced Applications*, volume 12112 of *Lecture Notes in Computer Science*, pp. 88–104. Springer, 2020. doi: 10.1007/978-3-030-59410-7_6.

UCI Machine Learning Repository. Mushroom, 1981. Dataset, DOI: 10.24432/C5959T.

Tianhao Wang, Jeremiah Blocki, Ninghui Li, and Somesh Jha. Locally differentially private protocols for frequency estimation. In *26th USENIX Security Symposium (USENIX Security 17)*, pp. 729–745, Vancouver, BC, 2017. USENIX Association. ISBN 978-1-931971-40-9.

Stanley L. Warner. Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association*, 60(309):63–69, 1965. doi: 10.1080/01621459.1965.10480775.

Min Ye and Alexander Barg. Optimal schemes for discrete distribution estimation under locally differential privacy. *IEEE Transactions on Information Theory*, 64(8):5662–5676, 2018. doi: 10.1109/TIT.2018.2809790.

A Proofs

This appendix provides proofs of the main theoretical results. Throughout, the output alphabet is

$$\mathcal{Y} = [k] \cup \{\Omega\}.$$

For k-ary randomized response with privacy parameter α, recall that

$$p_\alpha = \frac{e^\alpha}{e^\alpha + k - 1}, \qquad r_\alpha = \frac{1}{e^\alpha + k - 1},$$

and, for $\alpha > 0$,

$$\gamma_\alpha = p_\alpha - r_\alpha = \frac{e^\alpha - 1}{e^\alpha + k - 1}.$$

A.1 Input Independent Nonresponse

A.1.1 Proof of the Erasure Representation

Proof. Fix an input category $i \in [k]$. Let Z be the output of K_ε and let T_δ independently replace Z by Ω with probability δ.

For any singleton output $y \in [k]$,

$$\Pr(Y = y \mid X = i) = (1 - \delta) K_\varepsilon(y \mid i).$$

Therefore,

$$\Pr(Y = i \mid X = i) = (1 - \delta) \frac{e^\varepsilon}{e^\varepsilon + k - 1},$$

while, for $y \neq i$,

$$\Pr(Y = y \mid X = i) = (1 - \delta)\frac{1}{e^{\varepsilon} + k - 1}.$$

The erasure output satisfies

$$\Pr(Y = \Omega \mid X = i) = \delta.$$

These are exactly the transition probabilities of $Q_{\varepsilon,\delta}$. Hence

$$Q_{\varepsilon,\delta} = T_\delta K_\varepsilon.$$

□

Exact privacy level. For $0 \leq \delta < 1$, the erasure does not change the exact privacy parameter of k-ary randomized response. Indeed, for singleton outputs, the common factor $1 - \delta$ cancels from every likelihood ratio, while the likelihood ratio for Ω is one. Thus

$$\max_{i,j,y} \log \frac{Q_{\varepsilon,\delta}(y \mid i)}{Q_{\varepsilon,\delta}(y \mid j)} = \varepsilon.$$

The case $\delta = 1$ is degenerate because the mechanism always outputs Ω and therefore has privacy parameter zero.

A.1.2 Proof of Blackwell Dominance

Proof. By the erasure representation,

$$Q_{\varepsilon,\delta} = T_\delta K_\varepsilon.$$

Thus, an observation generated by $Q_{\varepsilon,\delta}$ can be obtained by first observing the output of K_ε and then applying the stochastic postprocessing channel T_δ.

Thus $Q_{\varepsilon,\delta}$ is a Blackwell garbling of K_ε. Therefore, K_ε Blackwell dominates $Q_{\varepsilon,\delta}$.

More explicitly, consider any prior distribution on the private category, any action space $\mathcal{A}$, and any loss function $L : [k] \times \mathcal{A} \to \mathbb{R}$. Every decision rule based on an observation from $Q_{\varepsilon,\delta}$ can be reproduced from an observation generated by K_ε by first applying T_δ and then applying the original decision rule. Hence the set of achievable decision rules under $Q_{\varepsilon,\delta}$ is contained in the set achievable under K_ε. It follows that

$$\inf_{\phi_K} \mathbb{E}L(X, \phi_K(Z)) \leq \inf_{\phi_Q} \mathbb{E}L(X, \phi_Q(Y)).$$

The same argument applies to n independent observations because

$$Q_{\varepsilon,\delta}^{\otimes n} = T_\delta^{\otimes n} K_\varepsilon^{\otimes n}.$$

□

A.1.3 Proof of the Information Contraction Identities

Let K_i and Q_i denote the output distributions conditional on $X = i$.

Total variation. For two distinct categories i and j, the distributions K_i and K_j differ only at outputs i and j. Hence

$$D_{\mathrm{TV}}(K_i, K_j) = \frac{1}{2}\sum_{y=1}^{k} |K_\varepsilon(y \mid i) - K_\varepsilon(y \mid j)| \tag{26}$$

$$= p_\varepsilon - r_\varepsilon \tag{27}$$

$$= \gamma_\varepsilon. \tag{28}$$

Under independent nonresponse,

$$Q_i(y) = (1 - \delta)K_i(y), \qquad y \in [k],$$

and

$$Q_i(\Omega) = Q_j(\Omega) = \delta.$$

Consequently,

$$D_{\mathrm{TV}}(Q_i, Q_j) = \frac{1}{2}\sum_{y\in\mathcal{Y}} |Q_i(y) - Q_j(y)| \tag{29}$$

$$= (1-\delta) D_{\mathrm{TV}}(K_i, K_j) \tag{30}$$

$$= (1-\delta)\gamma_\varepsilon. \tag{31}$$

f divergences. Let f be a convex function with $f(1) = 0$. For discrete distributions,

$$D_f(P\|Q) = \sum_y Q(y) f\left(\frac{P(y)}{Q(y)}\right).$$

For $y \in [k]$,

$$\frac{Q_i(y)}{Q_j(y)} = \frac{K_i(y)}{K_j(y)}.$$

If $\delta = 1$, all inputs induce the point mass at Ω, so both sides of the desired identity are zero. Assume now $\delta < 1$. When $\delta = 0$, the ignorance symbol has zero mass and contributes nothing to the divergence. When $0 < \delta < 1$, its contribution is

$$\delta f(1) = 0.$$

Therefore,

$$D_f(Q_i\|Q_j) = \sum_{y=1}^{k} (1-\delta) K_j(y) f\left(\frac{K_i(y)}{K_j(y)}\right) \tag{32}$$

$$= (1-\delta) D_f(K_i\|K_j). \tag{33}$$

Kullback–Leibler divergence. For $i \neq j$, only outputs i and j contribute nonzero terms:

$$D_{\mathrm{KL}}(K_i\|K_j) = p_\varepsilon \log\frac{p_\varepsilon}{r_\varepsilon} + r_\varepsilon \log\frac{r_\varepsilon}{p_\varepsilon} \tag{34}$$

$$= (p_\varepsilon - r_\varepsilon)\log\frac{p_\varepsilon}{r_\varepsilon}. \tag{35}$$

Since

$$\frac{p_\varepsilon}{r_\varepsilon} = e^\varepsilon,$$

we obtain

$$D_{\mathrm{KL}}(K_i\|K_j) = \varepsilon\gamma_\varepsilon.$$

The f divergence result therefore gives

$$D_{\mathrm{KL}}(Q_i\|Q_j) = (1-\delta)\varepsilon\gamma_\varepsilon.$$

Mutual information. Let E denote the indicator that the randomized response output is not erased:

$$E = \begin{cases} 1, & Y \neq \Omega, \\ 0, & Y = \Omega. \end{cases}$$

The erasure event is independent of both X and the randomized response output Z. Moreover, E is determined by Y. By the chain rule,

$$I(X;Y) = I(X;E) + I(X;Y \mid E) \tag{36}$$

$$= (1-\delta) I(X;Z \mid E = 1), \tag{37}$$

because $I(X;E) = 0$ and Y is constant when $E = 0$. Independence of E and (X, Z) gives

$$I(X;Z \mid E = 1) = I(X;Z).$$

Hence

$$I(X;Y) = (1-\delta) I(X;Z).$$

Fisher information. Consider any regular parameterization $\boldsymbol{\theta}$ of the population distribution. Let

$$g_{\boldsymbol{\theta}}(z)$$

denote the probability of randomized response output z. Under independent nonresponse,

$$q_{\boldsymbol{\theta}}(y) = \begin{cases} (1-\delta)g_{\boldsymbol{\theta}}(y), & y \in [k], \\ \delta, & y = \Omega. \end{cases}$$

For $y \in [k]$,

$$\nabla_{\boldsymbol{\theta}} \log q_{\boldsymbol{\theta}}(y) = \nabla_{\boldsymbol{\theta}} \log g_{\boldsymbol{\theta}}(y),$$

whereas the score at Ω is zero because δ does not depend on $\boldsymbol{\theta}$. Therefore,

$$\mathcal{I}_{\mathrm{NR}} = \sum_{y=1}^{k} (1-\delta) g_{\boldsymbol{\theta}}(y) \nabla \log g_{\boldsymbol{\theta}}(y) \nabla \log g_{\boldsymbol{\theta}}(y)^{\top} \tag{38}$$

$$= (1-\delta)\mathcal{I}_{\mathrm{RR}}. \tag{39}$$

A.1.4 Proof of the Exact Finite Sample Risk

Let

$$\boldsymbol{u} = \left(\frac{1}{k}, \ldots, \frac{1}{k}\right).$$

Conditioned on receiving an informative report, the output follows ordinary k-ary randomized response. Its categorical distribution is therefore

$$\boldsymbol{\theta} = \boldsymbol{u} + \gamma_{\varepsilon}(\boldsymbol{\pi} - \boldsymbol{u}).$$

Since each report is retained independently with probability $1-\delta$,

$$M \sim \mathrm{Binomial}(n, 1-\delta).$$

Conditional on $M = m > 0$,

$$(N_1, \ldots, N_k) \mid M = m \sim \mathrm{Multinomial}(m, \boldsymbol{\theta}).$$

Let

$$\boldsymbol{F} = \frac{1}{m}(N_1, \ldots, N_k).$$

Then

$$\mathbb{E}[\boldsymbol{F} \mid M = m] = \boldsymbol{\theta}.$$

The raw inverse estimator is

$$\widetilde{\boldsymbol{\pi}} = \boldsymbol{u} + \frac{\boldsymbol{F} - \boldsymbol{u}}{\gamma_{\varepsilon}}.$$

Using the expression for $\boldsymbol{\theta}$,

$$\mathbb{E}[\widetilde{\boldsymbol{\pi}} \mid M = m] = \boldsymbol{u} + \frac{\boldsymbol{\theta} - \boldsymbol{u}}{\gamma_{\varepsilon}} \tag{40}$$

$$= \boldsymbol{\pi}. \tag{41}$$

Thus the estimator is conditionally unbiased.

The multinomial covariance of $\boldsymbol{F}$ is

$$\mathrm{Cov}(\boldsymbol{F} \mid M = m) = \frac{1}{m}\left[\mathrm{Diag}(\boldsymbol{\theta}) - \boldsymbol{\theta}\boldsymbol{\theta}^{\top}\right].$$

Hence

$$\mathbb{E}\left[\|\widetilde{\boldsymbol{\pi}} - \boldsymbol{\pi}\|_2^2 \mid M = m\right] = \mathrm{tr}\left(\mathrm{Cov}(\widetilde{\boldsymbol{\pi}} \mid M = m)\right) \tag{42}$$

$$= \frac{1 - \|\boldsymbol{\theta}\|_2^2}{m\gamma_{\varepsilon}^2}. \tag{43}$$

Because

$$\boldsymbol{u}^\top(\boldsymbol{\pi}-\boldsymbol{u})=0,$$

we have

$$\|\boldsymbol{\theta}\|_2^2=\|\boldsymbol{u}+\gamma_\varepsilon(\boldsymbol{\pi}-\boldsymbol{u})\|_2^2 \tag{44}$$

$$=\frac{1}{k}+\gamma_\varepsilon^2\|\boldsymbol{\pi}-\boldsymbol{u}\|_2^2. \tag{45}$$

Therefore,

$$\mathbb{E}\left[\|\widetilde{\boldsymbol{\pi}}-\boldsymbol{\pi}\|_2^2 \mid M=m\right]=\frac{1}{m}\left[\frac{1-1/k}{\gamma_\varepsilon^2}-\|\boldsymbol{\pi}-\boldsymbol{u}\|_2^2\right] \tag{46}$$

$$=\frac{C_{\varepsilon,k}(\boldsymbol{\pi})}{m}. \tag{47}$$

Conditioning on $M>0$ and applying the law of total expectation gives

$$\mathbb{E}\left[\|\widetilde{\boldsymbol{\pi}}-\boldsymbol{\pi}\|_2^2 \mid M>0\right] \tag{48}$$

$$=C_{\varepsilon,k}(\boldsymbol{\pi})\sum_{m=1}^{n}\frac{1}{m}\Pr(M=m\mid M>0) \tag{49}$$

$$=C_{\varepsilon,k}(\boldsymbol{\pi})\frac{\displaystyle\sum_{m=1}^{n}\frac{1}{m}\binom{n}{m}(1-\delta)^m\delta^{n-m}}{1-\delta^n}. \tag{50}$$

For the unconditional risk used in the experiments, define the estimator on the remaining event $M=0$ by

$$\widetilde{\boldsymbol{\pi}}=\boldsymbol{u}.$$

Because $\Pr(M=0)=\delta^n$, the law of total expectation gives

$$R_{n,\delta}:=\mathbb{E}\left[\|\widetilde{\boldsymbol{\pi}}-\boldsymbol{\pi}\|_2^2\right] \tag{51}$$

$$=\delta^n\|\boldsymbol{\pi}-\boldsymbol{u}\|_2^2+C_{\varepsilon,k}(\boldsymbol{\pi})\sum_{m=1}^{n}\frac{1}{m}\binom{n}{m}(1-\delta)^m\delta^{n-m}. \tag{52}$$

This is exactly Eq. (15).

To obtain the asymptotic expression, write

$$q=1-\delta.$$

For fixed $q>0$, the $M=0$ contribution is exponentially small. Moreover, a Taylor expansion around nq, together with standard binomial concentration, gives

$$\mathbb{E}\left[\frac{\mathbf{1}\{M>0\}}{M}\right]=\frac{1}{nq}+O(n^{-2}).$$

Consequently,

$$R_{n,\delta}=\frac{C_{\varepsilon,k}(\boldsymbol{\pi})}{n(1-\delta)}+O(n^{-2}).$$

For $\delta=0$,

$$R_{n,0}=\frac{C_{\varepsilon,k}(\boldsymbol{\pi})}{n}.$$

Therefore,

$$\frac{R_{n,\delta}}{R_{n,0}}\longrightarrow\frac{1}{1-\delta}.$$

Finally, since the probability simplex Δ_k is closed and convex, Euclidean projection satisfies

$$\|\Pi_{\Delta_k}(\boldsymbol{x})-\boldsymbol{z}\|_2\le\|\boldsymbol{x}-\boldsymbol{z}\|_2$$

for every $\boldsymbol{z}\in\Delta_k$. Taking $\boldsymbol{x}=\widetilde{\boldsymbol{\pi}}$ and $\boldsymbol{z}=\boldsymbol{\pi}$ gives

$$\|\widehat{\boldsymbol{\pi}}-\boldsymbol{\pi}\|_2\le\|\widetilde{\boldsymbol{\pi}}-\boldsymbol{\pi}\|_2.$$

□

A.2 Category Dependent Nonresponse

A.2.1 Proof of the Exact Privacy Cost

Proof. Because the output alphabet is finite, the exact LDP parameter is the largest pointwise log likelihood ratio over all inputs and outputs.

First consider the nonresponse output Ω. For inputs i and j,

$$\frac{Q_{\alpha,\boldsymbol{\delta}}(\Omega \mid i)}{Q_{\alpha,\boldsymbol{\delta}}(\Omega \mid j)} = \frac{\delta_i}{\delta_j}.$$

The largest such ratio is

$$\frac{\delta_{\max}}{\delta_{\min}},$$

which contributes

$$\log \frac{\delta_{\max}}{\delta_{\min}}.$$

Now consider an informative output $y \in [k]$. Since $a_i = 1 - \delta_i$,

$$\frac{Q_{\alpha,\boldsymbol{\delta}}(y \mid i)}{Q_{\alpha,\boldsymbol{\delta}}(y \mid j)} = \frac{a_i}{a_j} \frac{K_\alpha(y \mid i)}{K_\alpha(y \mid j)}.$$

For k-ary randomized response,

$$\max_{i,j,y} \frac{K_\alpha(y \mid i)}{K_\alpha(y \mid j)} = e^\alpha.$$

Furthermore,

$$\max_{i,j} \frac{a_i}{a_j} = \frac{1 - \delta_{\min}}{1 - \delta_{\max}}.$$

These two maxima can be achieved simultaneously. If $\delta_{\min} < \delta_{\max}$, choose i attaining $\delta_{\min}$ and j attaining $\delta_{\max}$; then $i \neq j$. If all δ_i are equal, choose any two distinct categories $i \neq j$, which is possible because $k \geq 2$. In either case, choosing the informative output $y = i$ gives

$$\frac{K_\alpha(i \mid i)}{K_\alpha(i \mid j)} = \frac{p_\alpha}{r_\alpha} = e^\alpha.$$

Hence the largest informative-output privacy loss is

$$\alpha + \log \frac{1 - \delta_{\min}}{1 - \delta_{\max}}.$$

Taking the larger of the nonresponse and informative-output terms gives

$$\varepsilon_Q = \max \left\{ \log \frac{\delta_{\max}}{\delta_{\min}},\ \alpha + \log \frac{1 - \delta_{\min}}{1 - \delta_{\max}} \right\}.$$

□

Boundary cases. The theorem assumes $0 < \delta_i < 1$ for every category. If $\delta_i = 0$ for one category and $\delta_j > 0$ for another, then Ω has positive probability under one input and zero probability under another, so no finite LDP parameter exists. Similarly, if $\delta_i = 1$ for one category while another category has $\delta_j < 1$, singleton outputs create a support mismatch and the privacy loss is infinite. The two homogeneous extremes are finite and should be handled separately: if $\delta_i = 0$ for every i, the channel is ordinary k-RR and has exact privacy parameter α; if $\delta_i = 1$ for every i, the channel always outputs Ω and has privacy parameter zero.

A.2.2 Proof of the No Free Heterogeneous Nonresponse Corollary

Proof. Set $\alpha = \varepsilon$. The exact privacy formula gives

$$\varepsilon_Q \geq \varepsilon + \log \frac{1-\delta_{\min}}{1-\delta_{\max}}.$$

Since

$$\delta_{\min} \leq \delta_{\max},$$

we have

$$\frac{1-\delta_{\min}}{1-\delta_{\max}} \geq 1.$$

Thus the logarithmic term is nonnegative. The mechanism can satisfy $\varepsilon_Q \leq \varepsilon$ only if

$$\log \frac{1-\delta_{\min}}{1-\delta_{\max}} = 0,$$

which implies

$$\delta_{\min} = \delta_{\max}.$$

Hence all category specific nonresponse probabilities must be equal. □

A.2.3 Derivation of the Privacy Budget Correction

To satisfy a target privacy level ε, the informative-output term must obey

$$\alpha + \log \frac{1-\delta_{\min}}{1-\delta_{\max}} \leq \varepsilon.$$

Therefore,

$$\alpha \leq \varepsilon - \log \frac{1-\delta_{\min}}{1-\delta_{\max}}.$$

The largest feasible informative privacy parameter is consequently

$$\alpha^\star = \varepsilon - \log \frac{1-\delta_{\min}}{1-\delta_{\max}},$$

provided that

$$\log \frac{\delta_{\max}}{\delta_{\min}} \leq \varepsilon$$

and $\alpha^\star \geq 0$.

A.2.4 Proof of the KL Decomposition

Proof. Let R denote the response indicator. Under private category i,

$$R \sim \mathrm{Ber}(a_i).$$

Conditional on $R = 1$, the informative output follows $K_\alpha(\cdot \mid i)$, while conditional on $R = 0$ the output is deterministically Ω.

Since R is determined by the released output, the chain rule for KL divergence gives

$$\begin{aligned} D_{\mathrm{KL}}(Q_i \| Q_j) = {} & D_{\mathrm{KL}}\left(\mathrm{Ber}(a_i) \| \mathrm{Ber}(a_j)\right) && (53)\\ & + a_i D_{\mathrm{KL}}\left(K_\alpha(\cdot \mid i) \| K_\alpha(\cdot \mid j)\right) && (54)\\ & + (1-a_i) D_{\mathrm{KL}}(\delta_\Omega \| \delta_\Omega). && (55) \end{aligned}$$

The last term is zero. Therefore,

$$D_{\mathrm{KL}}(Q_i \| Q_j) = D_{\mathrm{KL}}\left(\mathrm{Ber}(a_i) \| \mathrm{Ber}(a_j)\right) + a_i D_{\mathrm{KL}}\left(K_\alpha(\cdot \mid i) \| K_\alpha(\cdot \mid j)\right).$$

□

A.2.5 Proof of the Mutual Information Decomposition

Proof. Let

$$\overline{a} = \Pr(R = 1) = \sum_{i=1}^{k} \pi_i a_i.$$

Bayes' rule gives

$$\Pr(X = i \mid R = 1) = \frac{\pi_i a_i}{\overline{a}} = \pi_i^{(R)}.$$

Since R is a deterministic function of the released output Y, the chain rule gives

$$I(X;Y) = I(X;R) + I(X;Y \mid R).$$

When $R = 0$, $Y = \Omega$ deterministically, so

$$I(X;Y \mid R = 0) = 0.$$

When $R = 1$, the released content is precisely the randomized response output Z. Hence

$$I(X;Y \mid R = 1) = I_{\boldsymbol{\pi}^{(R)}}(X;Z).$$

Averaging over R yields

$$I(X;Y) = I(X;R) + \overline{a} I_{\boldsymbol{\pi}^{(R)}}(X;Z).$$

□

A.2.6 Proof of the Selection Bias Formula

Proof. By Bayes' rule,

$$\pi_i^{(R)} = \frac{\pi_i a_i}{\overline{a}}.$$

Subtracting π_i gives

$$\pi_i^{(R)} - \pi_i = \frac{\pi_i a_i}{\overline{a}} - \pi_i \tag{56}$$

$$= \frac{\pi_i (a_i - \overline{a})}{\overline{a}}. \tag{57}$$

Thus conditioning on response leaves the population distribution unchanged if and only if the response propensities do not alter the relative category weights. □

A.2.7 Proof of the Known-Propensity Estimator

Proof. Let

$$D_a = \operatorname{diag}(a_1, \ldots, a_k).$$

The unconditional probability of singleton output j is

$$s_j = \sum_{i=1}^{k} \pi_i a_i K_\alpha(j \mid i).$$

In vector form,

$$\boldsymbol{s} = K_\alpha^\top D_a \boldsymbol{\pi}.$$

For $\alpha > 0$, the k-ary randomized response matrix is invertible. Indeed,

$$K_\alpha = r_\alpha \mathbf{1}\mathbf{1}^\top + \gamma_\alpha I.$$

Its eigenvalue in the direction of $\mathbf{1}$ is one, while each eigenvalue orthogonal to $\mathbf{1}$ is $\gamma_\alpha > 0$. Therefore K_α is nonsingular.

Since $a_i > 0$ for every i, D_a is also nonsingular, and

$$\boldsymbol{\pi} = D_a^{-1} (K_\alpha^\top)^{-1} \boldsymbol{s}.$$

For each user, define the vector $\boldsymbol{V} \in \mathbb{R}^k$ by

$$V_j = \mathbf{1}\{Y = j\}.$$

If $Y = \Omega$, then $\boldsymbol{V}$ is the zero vector. We have

$$\mathbb{E}[\boldsymbol{V}] = \boldsymbol{s}.$$

The empirical singleton-frequency vector is

$$\widehat{\boldsymbol{s}} = \frac{1}{n} \sum_{\ell=1}^{n} \boldsymbol{V}_\ell.$$

Hence

$$\mathbb{E}[\widehat{\boldsymbol{s}}] = \boldsymbol{s}.$$

Therefore,

$$\mathbb{E}[\widetilde{\boldsymbol{\pi}}_{\text{known}}] = D_a^{-1}(K_\alpha^\top)^{-1}\mathbb{E}[\widehat{\boldsymbol{s}}] \tag{58}$$
$$= D_a^{-1}(K_\alpha^\top)^{-1}\boldsymbol{s} \tag{59}$$
$$= \boldsymbol{\pi}. \tag{60}$$

Thus the estimator is unbiased.

Because $\boldsymbol{V}$ is a categorical indicator vector with an additional zero outcome,

$$\mathrm{Cov}(\boldsymbol{V}) = \mathrm{Diag}(\boldsymbol{s}) - \boldsymbol{s}\boldsymbol{s}^\top.$$

Independence across users gives

$$\mathrm{Cov}(\widehat{\boldsymbol{s}}) = \frac{1}{n}\left[\mathrm{Diag}(\boldsymbol{s}) - \boldsymbol{s}\boldsymbol{s}^\top\right].$$

Applying the linear transformation

$$D_a^{-1}(K_\alpha^\top)^{-1}$$

therefore yields

$$\mathrm{Cov}(\widetilde{\boldsymbol{\pi}}_{\text{known}}) = \frac{1}{n} D_a^{-1}(K_\alpha^\top)^{-1}\left[\mathrm{Diag}(\boldsymbol{s}) - \boldsymbol{s}\boldsymbol{s}^\top\right] \tag{61}$$
$$\cdot\, K_\alpha^{-1} D_a^{-1}. \tag{62}$$

Since the estimator is unbiased,

$$\mathbb{E}\left[\|\widetilde{\boldsymbol{\pi}}_{\text{known}} - \boldsymbol{\pi}\|_2^2\right] = \mathrm{tr}\left[\mathrm{Cov}(\widetilde{\boldsymbol{\pi}}_{\text{known}})\right].$$

The covariance contains the factor $1/n$ exactly, so the squared risk also scales exactly as $1/n$. □

A.2.8 Proof of Nonidentifiability

As assumed in Theorem 4, we require $\alpha > 0$ so that K_α is invertible.

Proof. Define

$$w_i = \pi_i a_i$$

and

$$\boldsymbol{w} = (w_1, \ldots, w_k).$$

The singleton-output probability vector is

$$\boldsymbol{s} = K_\alpha^\top \boldsymbol{w}.$$

When $\alpha > 0$, K_α is invertible, as shown above. Hence the output law determines

$$\boldsymbol{w} = (K_\alpha^\top)^{-1}\boldsymbol{s}.$$

The nonresponse probability is

$$\Pr(Y = \Omega) = 1 - \sum_{i=1}^{k} w_i,$$

so it contains no additional information for separating $\boldsymbol{\pi}$ and $\boldsymbol{a}$.

Now suppose

$$\sum_{i=1}^{k} w_i < 1.$$

Let $\boldsymbol{\pi}'$ be any probability vector satisfying

$$\pi'_i \geq w_i \qquad \text{for every } i.$$

For coordinates with $\pi'_i > 0$, define

$$a'_i = \frac{w_i}{\pi'_i}.$$

If $\pi'_i = 0$, the constraint $\pi'_i \geq w_i \geq 0$ forces $w_i = 0$; for such a coordinate set $a'_i = 0$. Thus $0 \leq a'_i \leq 1$ and, in all cases,

$$\pi'_i a'_i = w_i.$$

Therefore,

$$K_\alpha^\top D_{a'} \boldsymbol{\pi}' = K_\alpha^\top \boldsymbol{w} = \boldsymbol{s}.$$

Moreover,

$$1 - \sum_{i=1}^{k} \pi'_i a'_i = 1 - \sum_{i=1}^{k} w_i,$$

so the probability of Ω is also unchanged.

Consequently, $(\boldsymbol{\pi}', \boldsymbol{a}')$ produces exactly the same complete output distribution as $(\boldsymbol{\pi}, \boldsymbol{a})$.

When $k \geq 2$ and

$$1 - \sum_{i} w_i > 0,$$

the remaining probability mass can generally be distributed among the coordinates of $\boldsymbol{\pi}'$ in more than one way. Hence there are multiple, and under interior conditions infinitely many, distinct parameter pairs that induce the same privatized output law. The original population distribution and category specific response propensities are therefore not separately identifiable without additional assumptions. □

Why $\alpha > 0$ is necessary. If $\alpha = 0$, then

$$K_0(j \mid i) = \frac{1}{k}$$

for all i and j. The matrix K_0 has rank one and is not invertible. In that case, even the vector $\boldsymbol{w}$ is not fully identifiable from the singleton outputs. Thus the statement that the output law identifies $\boldsymbol{w}$ requires $\alpha > 0$.

B Additional Experimental Details

B.1 Implementation and Repetition Counts

The final evidence package uses seed 20260828. For category-independent nonresponse, the privacy grid is $\varepsilon \in \{0.5, 1.0, 1.5, 2.0, 3.0\}$ and the principal nonresponse levels are $\delta \in \{0.1, 0.2\}$. The main synthetic benchmark uses 200 independent privatization repetitions per configuration, the real-data benchmark uses 150, the sample-size experiments use 200, estimator ablations use 120, and exact risk validation uses 20,000 Monte Carlo repetitions per configuration. Paired confidence intervals use 2,000 bootstrap resamples.

The category-dependent suite uses 30,000 repetitions for selection-bias validation and 50,000 repetitions for each known-propensity risk configuration. Its exported run manifest records Python 3.13.15, NumPy 2.1.3, pandas 2.2.3, SciPy 1.16.3, and Matplotlib 3.10.0 for the final run. These version numbers describe the reported execution environment rather than requirements of the theory.

B.2 Metrics and Estimators

Benchmark comparisons use projected ℓ_1 error $\|\widehat{\boldsymbol{\pi}} - \boldsymbol{\pi}\|_1$. Exact risk validations use raw squared ℓ_2 error because the analytical formulas are derived for the raw linear inverses. In the independent case, the raw estimator is completed with the uniform fallback $\boldsymbol{u}$ on $M = 0$, exactly as in Eq. (15). Benchmark estimates are projected onto the simplex after inversion.

For $k = 2$, k-RR is exactly Warner randomized response. We therefore do not report a separately seeded Warner implementation as a distinct mechanism. OUE, OLH, and HR-C1C2 are included as contextual frequency-estimation baselines; no global optimality claim is made for k-RR.

For OUE, we use $p = 1/2$ and $q = 1/(e^{\varepsilon} + 1)$, followed by the standard coordinate-wise inverse. For OLH, we use

$$g = \max\{2, \mathrm{round}(e^{\varepsilon} + 1)\}, \qquad p = \frac{e^{\varepsilon}}{e^{\varepsilon} + g - 1},$$

with an independently sampled hash mapping for each user and collision probability $1/g$. For Hadamard Response, we use the balanced C1/C2 construction with the smallest Hadamard order $K \geq k + 1$ and $p_{\mathrm{in}} = e^{\varepsilon}/(e^{\varepsilon} + 1)$.

B.3 Synthetic Distributions

The main synthetic scenarios are: binary skew $(0.7, 0.3)$ with $n = 5000$; Zipf $k = 5$ with $n = 5000$; a sparse $k = 10$ distribution with $n = 5000$; and Zipf $k = 20$ with $n = 10000$. Independent exact-risk validation also uses

$$\delta \in \{0, 0.05, 0.10, 0.20, 0.30, 0.40\},$$

together with multiple values of k, ε, and n.

B.4 Construction of Category-Dependent Nonresponse

For the controlled heterogeneous trajectory, let $\boldsymbol{t}$ be the equally spaced vector from -1 to 1 in k coordinates. Define

$$\widetilde{\boldsymbol{v}} = \boldsymbol{t} - (\boldsymbol{\pi}^{\top}\boldsymbol{t})\mathbf{1},$$
$$\boldsymbol{v} = \widetilde{\boldsymbol{v}}/\|\widetilde{\boldsymbol{v}}\|_{\infty}, \qquad \boldsymbol{\delta}(h) = \bar{\delta}\mathbf{1} + h\boldsymbol{v}. \tag{63}$$

Then $\boldsymbol{\pi}^{\top}\boldsymbol{v} = 0$, so $\sum_i \pi_i \delta_i(h) = \bar{\delta}$ exactly. The reported sweep uses Zipf $k = 5$, $\bar{\delta} = 0.2$, $h \in [0, 0.14]$, and target $\varepsilon = 1$. At each h, $\alpha^{\star}$ is recomputed from Eq. (19). This construction is a controlled illustration, not a claim that all forms of heterogeneity follow the same path.

B.5 Real-Data Preprocessing

The four UCI tasks estimate the empirical distribution of one selected categorical variable from Adult, Car Evaluation, Mushroom, and Nursery (Becker & Kohavi, 1996; Bohanec, 1988; UCI Machine Learning Repository, 1981; Rajkovic, 1989). The source datasets do not contain the response mechanisms studied here; after extracting the categorical distribution, we simulate the LDP and nonresponse channels under the stated experimental protocol:

Dataset	n	k	Variable	UCI ID
Adult	48,177	5	marital status	2
Car Evaluation	1,728	4	class	19
Mushroom	8,124	10	cap color	73
Nursery	12,960	5	class	76

For Adult, we intentionally retain only `Married-civ-spouse`, `Never-married`, `Divorced`, `Separated`, and `Widowed`; rows outside these five values are excluded before encoding. This produces $n = 48{,}177$ and $k = 5$. For the other three datasets, missing categorical labels are removed and all remaining observed categories of the selected variable are encoded. These are frequency estimation tasks, not supervised classification experiments.

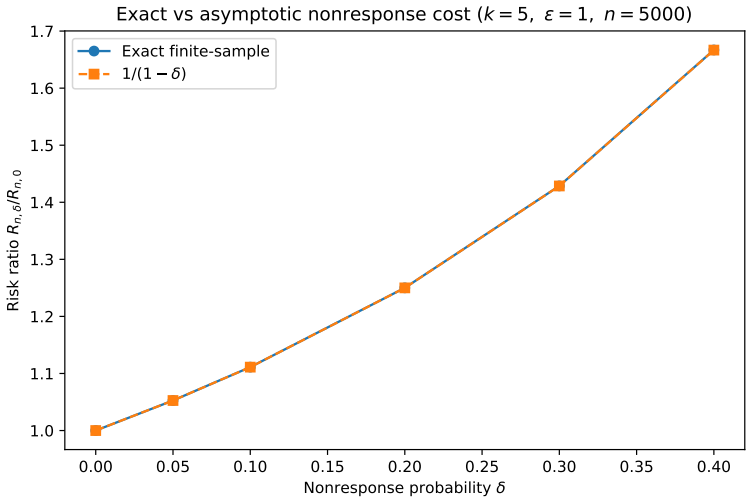

Figure 2: Exact finite sample risk inflation under independent nonresponse, compared with the asymptotic factor $1/(1-\delta)$ for Zipf $k = 5$, $n = 5000$, and $\varepsilon = 1$.

C Additional Tables and Figures

C.1 Independent Nonresponse

Table 2: Exact unconditional finite-sample squared ℓ_2 risk validation for Zipf $k = 5, \varepsilon = 1, n = 5000$, with the uniform fallback at $M = 0$.

δ	Empirical	Exact	Exact ratio	$1/(1-\delta)$
0.00	0.002414	0.002430	1.000000	1.000000
0.05	0.002571	0.002558	1.052643	1.052632
0.10	0.002716	0.002700	1.111136	1.111111
0.20	0.003018	0.003037	1.250063	1.250000
0.30	0.003465	0.003471	1.428694	1.428571
0.40	0.004041	0.004050	1.666889	1.666667

Table 3: Paired projected-ℓ_1 difference $\mathrm{NR}(\delta = .1) - \mathrm{kRR}$ on synthetic data at $\varepsilon = 1$. Intervals use paired bootstrap resampling.

Scenario	Mean difference	95% interval
Binary skew $(0.7, 0.3)$	-0.001633	$[-0.005227, 0.001806]$
Zipf $k = 5$	-0.000987	$[-0.007596, 0.005485]$
Sparse $k = 10$	0.010766	$[0.000356, 0.020938]$
Zipf $k = 20$	0.023020	$[0.010986, 0.035473]$

Figure 2 compares the exact unconditional risk in Eq. (15) with Monte Carlo simulation for Zipf $k = 5$, $n = 5000$, and $\varepsilon = 1$. At $\delta = 0.1$, the empirical raw risk is 0.002716 and the exact value is 0.002700; the exact inflation ratio is 1.111136, essentially the asymptotic value 1.111111. Across the full exact-risk validation suite, empirical risks remain close to the analytical values; Appendix Table 2 gives the representative sweep.

C.2 Category-Dependent Privacy, Selection, and Risk

This subsection collects the category-dependent results referenced in Sections 6.3–6.5. Table 4 and Figure 3 report the privacy-budget trajectory. Table 5 and Figure 4 show the category-level selection effect at $h = 0.10$, while Table 6 and Figure 5 validate the exact finite-sample risk of the

known-propensity estimator across sample sizes. Table 7 summarizes the controlled heterogeneity sweep at fixed average nonresponse and fixed total privacy.

Table 4: Privacy-budget trajectory at population-average nonresponse 0.2 and target $\varepsilon = 1$.

h	$\delta_{\min}$	$\delta_{\max}$	τ_{resp}	$\alpha^\star$
0.00	0.200000	0.200000	0.000000	1.000000
0.02	0.191532	0.220000	0.035847	0.964153
0.04	0.183065	0.240000	0.072241	0.927759
0.06	0.174597	0.260000	0.109221	0.890779
0.08	0.166130	0.280000	0.146826	0.853174
0.10	0.157662	0.300000	0.185101	0.814899
0.12	0.149195	0.320000	0.224090	0.775910
0.14	0.140727	0.340000	0.263847	0.736153

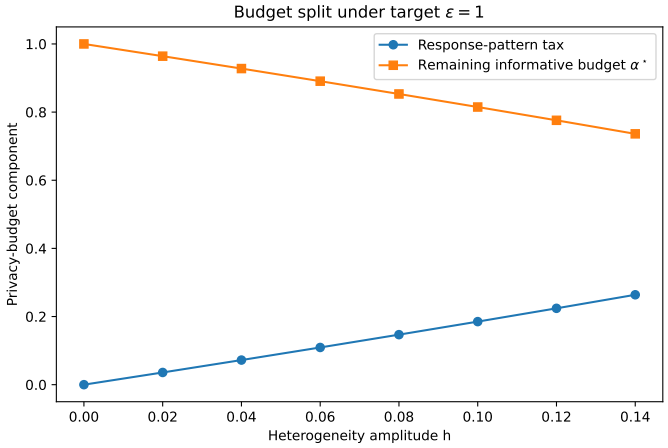

Figure 3: Privacy budget split along the controlled heterogeneity trajectory. The population average nonresponse rate is fixed at 0.2 and the target privacy level is $\varepsilon = 1$.

Table 5: Selection bias under category-dependent nonresponse at $h = 0.10$. The naive estimator tracks the responder distribution; the raw known-propensity estimator tracks the population.

Category	π_i	$\pi_i^{(R)}$	Naive mean	Known-prop. mean
1	0.437956	0.461134	0.461324	0.438172
2	0.218978	0.220827	0.220560	0.218733
3	0.145985	0.140724	0.140649	0.145915
4	0.109489	0.100673	0.100558	0.109373
5	0.087591	0.076642	0.076910	0.087906

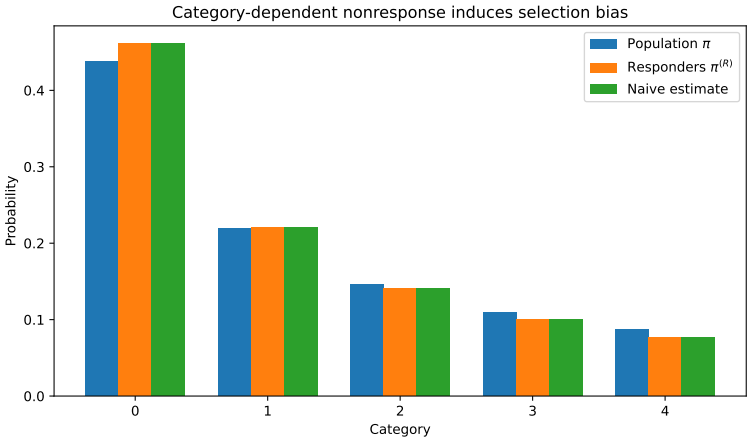

Figure 4: Category-level selection effect. The responder-only inverse follows $\boldsymbol{\pi}^{(R)}$, whereas the known-propensity estimator returns to the population target.

Table 6: Exact finite-sample risk validation for the known-propensity raw estimator.

n	Empirical L_2^2	95% MC interval	Exact L_2^2	nR_n
500	0.053137	$[0.052806, 0.053469]$	0.053139	26.569669
2000	0.013282	$[0.013200, 0.013364]$	0.013285	26.569669
5000	0.005327	$[0.005294, 0.005361]$	0.005314	26.569669
20000	0.001323	$[0.001315, 0.001332]$	0.001328	26.569669

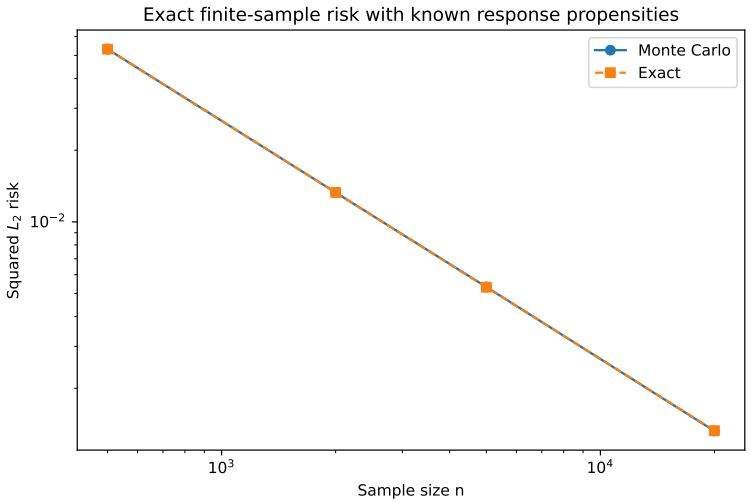

Figure 5: Monte Carlo and exact squared ℓ_2 risk for the known-propensity estimator over a forty-fold range of sample sizes.

Table 7: Fixed-average, fixed-total-privacy heterogeneity sweep. Risk ratios are relative to homogeneous nonresponse.

h	$\alpha^\star$	Privacy tax	Selection L_1	Risk ratio
0.00	1.000000	0.000000	0.000000	1.000000
0.02	0.964153	0.035847	0.010010	1.103124
0.05	0.909344	0.090656	0.025026	1.292328
0.08	0.853174	0.146826	0.040042	1.536874
0.10	0.814899	0.185101	0.050052	1.741536
0.12	0.775910	0.224090	0.060063	1.990411
0.14	0.736153	0.263847	0.070073	2.296712

C.3 Nonidentifiability Construction

The numerical construction uses

$$\boldsymbol{\pi}^{(1)} = (0.40, 0.30, 0.20, 0.10), \qquad \boldsymbol{a}^{(1)} = (0.90, 0.80, 0.70, 0.60),$$

which produce $\boldsymbol{w} = (0.36, 0.24, 0.14, 0.06)$. A second population uses

$$\boldsymbol{\pi}^{(2)} = (0.46, 0.26, 0.17, 0.11),$$

with

$$\boldsymbol{a}^{(2)} = (0.782609, 0.923077, 0.823529, 0.545455).$$

It has the same product vector. Both parameterizations yield the released law

$$(0.24807827, 0.21201957, 0.18197065, 0.15793151, 0.2)$$

up to floating-point precision, even though $\|\boldsymbol{\pi}^{(1)} - \boldsymbol{\pi}^{(2)}\|_1 = 0.14$. The maximum absolute difference between the two output laws is 1.11×10^{-16}.

C.4 Paired Real-Data Differences

Table 8 reports paired projected-ℓ_1 differences between independent nonresponse at $\delta = 0.1$ and plain k-RR on the four UCI tasks. The 95% intervals are obtained by paired bootstrap resampling over the shared privatization trials. Intervals containing zero indicate that no penalty is detectable at the reported sample size; they should not be interpreted as an information gain.

Table 8: Paired projected-ℓ_1 difference $\mathrm{NR}(\delta = .1) - \mathrm{kRR}$ on real data at $\varepsilon = 1$.

Dataset	Mean difference	95% interval
Adult	-0.000687	$[-0.002489, 0.001096]$
Car Evaluation	0.007949	$[-0.001013, 0.016698]$
Mushroom	0.011932	$[0.002117, 0.021785]$
Nursery	0.006388	$[0.001870, 0.011145]$

C.5 Estimator Ablation

Table 9 compares the projected linear inverse used in the benchmarks with a constrained maximum-likelihood estimator applied to the same released data. The two agree to the reported precision in all tested configurations, so the main text reports only the projected inverse.

Table 9: Projected inverse versus constrained maximum likelihood at $\varepsilon = 1$. The means are numerically indistinguishable in these tested settings.

Data	Mechanism	Projected inverse	Constrained MLE
Synthetic Zipf $k = 5$	k-RR	0.135275	0.135275
Synthetic Zipf $k = 5$	NR, $\delta = .1$	0.145580	0.145580
Adult	k-RR	0.026509	0.026509
Adult	NR, $\delta = .1$	0.028717	0.028717

C.6 Numerical Audit Summary

Table 10 summarizes representative numerical checks from the final run. The first four entries correspond to deterministic channel identities and are therefore expected to hold at machine precision. The final entry is the largest relative difference between Monte Carlo and exact risk across the known-propensity configurations and reflects sampling error rather than an analytical discrepancy.

Table 10: Representative final-run numerical audit. Machine-precision values refer to deterministic channel identities, not Monte Carlo risk estimates.

Check	Final-run discrepancy
Corrected heterogeneous privacy	2.22×10^{-16}
KL decomposition	1.67×10^{-16}
Mutual-information decomposition	6.94×10^{-18}
Nonidentifiability output-law equality	1.11×10^{-16}
Largest known-propensity MC relative risk error	3.896×10^{-3}

C.7 Full Privacy-Grid Benchmarks

Tables 11 and 12 extend the synthetic and real-data benchmarks to the full privacy grid

$$\varepsilon \in \{0.5, 1.0, 1.5, 2.0, 3.0\}.$$

Across the grid, the nonresponse extension broadly tracks k-RR, with larger δ generally increasing estimation error. The effect is more pronounced in several larger-alphabet settings, where OUE and OLH are also more competitive. Small finite-sample reversals can still occur after projection, consistent with the paired results reported above.

Table 11: Full synthetic privacy-grid benchmark: mean projected ℓ_1 error.

Scenario	ε	k-RR	OUE	OLH	HR-C1C2	NR .1	NR .2
Binary skew $(0.7, 0.3)$	0.5	0.041605	0.065946	0.066877	0.065758	0.044723	0.048899
Binary skew $(0.7, 0.3)$	1.0	0.025162	0.031934	0.034979	0.037904	0.023529	0.027746
Binary skew $(0.7, 0.3)$	1.5	0.017486	0.024752	0.024106	0.022301	0.017525	0.019222
Binary skew $(0.7, 0.3)$	2.0	0.013802	0.020105	0.019918	0.020780	0.015101	0.015555
Binary skew $(0.7, 0.3)$	3.0	0.011726	0.018150	0.017719	0.017558	0.011937	0.012515
Zipf $k = 5$	0.5	0.196486	0.201766	0.205114	0.210540	0.204758	0.218027
Zipf $k = 5$	1.0	0.092063	0.100788	0.103609	0.104706	0.091075	0.095736
Zipf $k = 5$	1.5	0.055474	0.066567	0.067206	0.078340	0.058667	0.060343
Zipf $k = 5$	2.0	0.039237	0.052413	0.051594	0.063424	0.042312	0.044523
Zipf $k = 5$	3.0	0.026679	0.040390	0.039301	0.051402	0.027736	0.031619
Sparse $k = 10$	0.5	0.420252	0.342065	0.350014	0.342941	0.414462	0.465018
Sparse $k = 10$	1.0	0.211500	0.191884	0.192346	0.204724	0.222266	0.230053
Sparse $k = 10$	1.5	0.131093	0.139170	0.136695	0.154687	0.132098	0.141406
Sparse $k = 10$	2.0	0.084509	0.100300	0.099120	0.133029	0.087277	0.096357
Sparse $k = 10$	3.0	0.049030	0.067345	0.065091	0.111775	0.052761	0.055334
Zipf $k = 20$	0.5	0.720246	0.491170	0.498192	0.497912	0.757279	0.768779
Zipf $k = 20$	1.0	0.379524	0.280030	0.283734	0.291678	0.402545	0.422936
Zipf $k = 20$	1.5	0.227272	0.191151	0.188691	0.233075	0.236406	0.252891
Zipf $k = 20$	2.0	0.140059	0.138131	0.140528	0.196725	0.150047	0.157718
Zipf $k = 20$	3.0	0.068829	0.085897	0.087819	0.167100	0.072032	0.075109

Table 12: Full real-data privacy-grid benchmark: mean projected ℓ_1 error.

Dataset	ε	k-RR	OUE	OLH	HR-C1C2	NR .1	NR .2
Adult marital status	0.5	0.064435	0.065082	0.064006	0.063412	0.065427	0.070939
Adult marital status	1.0	0.028180	0.032602	0.031248	0.032700	0.027493	0.030296
Adult marital status	1.5	0.016067	0.020384	0.020809	0.022997	0.016379	0.017786
Adult marital status	2.0	0.010500	0.015563	0.014977	0.019493	0.011493	0.011797
Adult marital status	3.0	0.005499	0.010288	0.010427	0.015026	0.006176	0.006918
Car Evaluation class	0.5	0.199034	0.230798	0.213772	0.197960	0.216214	0.221642
Car Evaluation class	1.0	0.098754	0.117920	0.119807	0.124071	0.106704	0.110513
Car Evaluation class	1.5	0.063308	0.088824	0.081279	0.080686	0.066282	0.066139
Car Evaluation class	2.0	0.040641	0.066128	0.060894	0.071451	0.043251	0.049827
Car Evaluation class	3.0	0.022871	0.045360	0.043981	0.057322	0.024427	0.027972
Mushroom cap color	0.5	0.323682	0.264306	0.251077	0.264984	0.349493	0.369722
Mushroom cap color	1.0	0.144730	0.137308	0.129757	0.147788	0.156662	0.166145
Mushroom cap color	1.5	0.088939	0.090431	0.090953	0.105583	0.095678	0.097405
Mushroom cap color	2.0	0.057606	0.068925	0.066893	0.086137	0.060147	0.063655
Mushroom cap color	3.0	0.028652	0.044858	0.044082	0.072521	0.031133	0.035507
Nursery class	0.5	0.099107	0.101222	0.105769	0.107300	0.108634	0.115269
Nursery class	1.0	0.046625	0.054402	0.052336	0.057785	0.053013	0.055228
Nursery class	1.5	0.028979	0.035283	0.036896	0.041438	0.030399	0.031858
Nursery class	2.0	0.019094	0.026687	0.026937	0.031313	0.019807	0.021743
Nursery class	3.0	0.009968	0.017716	0.018917	0.025454	0.011080	0.012375