Brazilian Journal of Biometrics, **43**,1-15 (2025)
doi:10.28951/bjb.v43i4.798

BRAZILIAN JOURNAL OF BIOMΣTRICS
ISSN:2764-5290

ARTICLE

Enhancing the efficiency of randomized response techniques using direct responses

Muhammad Azeem*[1]

[1]Department of Statistics, University of Malakand, Pakistan
*Corresponding author. Email: azeemstats2017@gmail.com

(Received: November 20, 2024; Revised: March 25, 2025; Accepted: April 1, 2025; Published: September 17, 2025)

Abstract
The quality of the findings of a research study largely depends on data it is based on. Many research studies require data collection from a sample of human participants. Respondents' refusals and untruthful responses are common issues in surveys based on human participants. Randomized response survey methods ensure the respondents' privacy protection, motivating them for participation in the survey. The existing scrambling methods do not offer the respondents to report the direct response. In practice, survey statisticians face situations where some of the survey participants may be willing to report their direct responses, thus avoiding the complex process of scrambling the responses. Moreover, some recent models involve a complicated scrambling process which may put a burden on the respondents which makes it difficult to practically implement such models in sample surveys. We propose two novel randomized response techniques using both the direct and scrambling response options. The proposed techniques are found to achieve a significant boost in efficiency over the competitor techniques. The efficiency conditions have been derived and are found mathematically strong. The results suggest that the proposed techniques are suitable for implementation in practical sample surveys.

Keywords: Direct response, Efficiency, Sampling theory, Scrambled response, Variance.

1. Introduction

For a survey researcher, refusals and/or untruthful responses due to fear of punishment is a common issue. Most of the respondents do not want to expose their true response to the researcher in sample surveys related to sensitive characteristics. The sensitive issues may be abortion, alcoholism, sexual abuse, number of failures in exam, criminal activities, consumption on luxury items, the amount payable for income tax, number of violation of rules by citizens and so forth. To overcome such refusals and untruthful responses from respondents, Warner (1965) proposed a unique technique commonly known as randomized response technique. This technique is easy to apply for practical situations and uses scrambled responses to enhance efficiency and to care the privacy of the respondents.

Over the past few decades, many researchers have presented efficient, and privacy protected scrambling models. Pollock & Bek (1976) proposed a quantitative variant of scrambling techniques to get reliable responses. For estimation of sensitive characteristics, Himmelfarb and Edgell (1980) revised the Pollock and Bek (1976) additive model. Eichhorn and Hayre (1983) used multiplicative scrambling to estimate the mean of a population. Franklin (1989) proposed a unique scrambling device which features some interesting qualities as compared with competitor scrambling methods. The main benefit of the Franklin (1989) technique is flexibility which means that other randomized techniques may be obtained from this model. The Mangat and Singh (1990), Singh and Sedory (2013), and Singh and Grewal (2013) models are all special forms of the Franklin (1989) technique.

Gupta *et al.* (2002) developed the concept of randomized response techniques. In this method, the respondents were free to report the true response or provide a random response. The decision to provide the true or scrambled response is based on whether the question is sensitive or not. When the question is sensitive, the respondent reports his/her true response otherwise a scrambled response is reported. Gupta *et al.* (2002) showed that optional techniques are more efficient than the conventional techniques, and they also proved it empirically as well as theoretically. The Gupta *et al.* (2002) technique was based on additive scrambling procedure. Bar-Lev *et al.* (2004) introduced a novel multiplicative optional model. Yan *et al.* (2008) introduced a technique to measure the privacy level of a scrambled response model. Gjestvang and Singh (2009) studied an improved scrambling technique for sensitive surveys. Hussain *et al.* (2016) used additive – subtractive variables to develop a new scrambling method.

Gupta *et al.* (2018) proposed a novel procedure to measure privacy as well as efficiency in each model simultaneously. The studies of Zamanzade and Wang (2018) and Santiago *et al.* (2019) analyzed estimation of population proportion under ranked set sampling design. Murtaza *et al.* (2020) presented an additive optional scramble technique by using correlated scrambling variable. Sanaullah *et al.* (2022) proposed a general scrambling method by using a two-phase sampling technique. Kumar and Kour (2022) conducted a detailed study on optional randomized procedures and measured the combined influence of different types of non-sampling errors. Mahdizadeh and Zamanzade (2022) discussed estimation of the area under the ROC curve in ranked set sampling. Narjis and Shabbir (2023) proposed an efficient scrambling technique by using an additional variable in the Gjestvang and Singh (2009) model.

Abbasi *et al.* (2022) and Abbasi and Asghar (2024) analyzed randomized response techniques under ranked set sampling design. Azeem (2023) presented a unified weighted measure for assessment of randomized techniques. Azeem and Ali (2023) compared different scrambling models under simple random sampling. Another recent study of Azeem (2024) presented an optional model which is better than the Gupta *et al.* (2024) technique in terms of efficiency. The studies of the Tarray and Singh (2015), Bouza (2016), Tiwari and Mehta (2017), Mehta and Aggarwal (2018), Kumar *et al.* (2023), and Lovig *et al.* (2023) provide further insights about scrambling models. This research paper proposes two novel optional randomized response models which improve the efficiency of the available models.

2. Selected Existing Models

Suppose the target population contains N units $\Omega = \{1, 2, 3, ...N\}$ from which n units are selected with replacement by simple random sampling. Suppose that the main sensitive variable is denoted by Y with mean $E(Y) = \mu_y$. Consider another scrambling variable W with mean $E(W) = \theta$. Moreover, let T be another scrambling variable with mean $E(T) = 1$. which is uncorrelated with variable Y, the variances of the variables are denoted by $V(Y) = \sigma_y^2$, $V(W) = \sigma_w^2$ and $V(T) = \sigma_T^2$.

2.1 Narjis and Shabbir (2023) Model

In this scrambled response technique, the data is collected using a randomization device, such as a deck of cards, or a spinner. The Narjis and Shabbir (2023) model gives three options to the respondents. Mathematically, every respondent provides his/her scrambled response (Z_i) as:

$$Z_i = \begin{cases} Y_i - \beta S, & \text{with probability } \dfrac{\alpha}{\alpha+\beta+\gamma}, \\ Y_i + \alpha S, & \text{with probability } \dfrac{\beta}{\alpha+\beta+\gamma}, \\ Y_i, & \text{with probability } \dfrac{\gamma}{\alpha+\beta+\gamma}, \end{cases} \tag{1}$$

where α, β and γ are some positive constants decided by researchers. An unbiased mean estimator using the n observed responses, is expressed as:

$$\hat{\mu}_{NS} = \frac{1}{n}\sum_{i=1}^{n} Z_i . \tag{2}$$

The estimator $\hat{\mu}_{NS}$ has variance:

$$Var\left(\hat{\mu}_{NS}\right) = \frac{1}{n}\left[\frac{\alpha\beta\left(\alpha+\beta\right)\left(\sigma_S^2+\theta^2\right)}{\alpha+\beta+\gamma} + \sigma_Y^2\right]. \tag{3}$$

2.2 Murtaza *et al.* (2020) Technique

In the Murtaza *et al.* (2020) scrambling technique, the observed responses Z_i are given as:

$$Z_i = \begin{cases} Y_i, & \text{with probability } 1\text{-}W, \\ TY + \alpha S, & \text{with probability } W. \end{cases} \tag{4}$$

An unbiased mean estimator on the basis of n observed responses, is given by:

$$\hat{\mu}_M = \frac{1}{n}\sum_{i=1}^{n} Z_i . \tag{5}$$

The estimator $\hat{\mu}_M$ has been shown to be unbiased with variance:

$$Var\left(\hat{\mu}_M\right) = \frac{1}{n}\left[W\sigma_T^2\left(\sigma_Y^2+\mu_Y^2\right)+\sigma_Y^2+\alpha^2\sigma_S^2\right]. \tag{6}$$

3. Proposed Models

In this section, two novel scrambling techniques are presented by using direct responses.

3.1 Proposed Model-I

Motivated by Narjis and Shabbir (2023), a new modified optional quantitative randomized response procedure is presented. In this technique, every respondent is requested to choose either the direct response or protected response, depending on the respondent's perception of the sensitive question. The researcher notes the choices of each respondent and hence knows the actual number of respondents who selected the true or protected responses.

Let n_1 be the number of participants reporting the direct response Y_i, and let n_2 be the number of respondents reporting the protected response Z_i. If a respondent opts for the protected scrambling response, he is provided three options of questions. The respondent selects one question at random, unobserved by the researcher, and reports his/her response accordingly.

The three options of questions are:

i. Provide the scrambled response $Y_i - \beta S$, with probability $\dfrac{\alpha}{\alpha+\beta+\gamma}$.

ii. Provide the scrambled response $Y_i + \alpha S$, with probability $\dfrac{\beta}{\alpha+\beta+\gamma}$.

iii. Provide the true response Y_i, with probability $\dfrac{\gamma}{\alpha+\beta+\gamma}$.

The mean of the respondents opting for the direct response is:

$$\bar{Y} = \frac{1}{n_1}\sum_{i=1}^{n1} Y_i . \tag{7}$$

The mean of the respondents opting for the protected response is:

$$\bar{Z} = \frac{1}{n_2}\sum_{i=1}^{n2} Z_i . \tag{8}$$

The overall sample mean can be calculated as:

$$\hat{\mu}_{P1} = \frac{n_1\bar{Y} + n_2\bar{Z}}{n_1 + n_2}, \tag{9}$$

where $n_1 + n_2 = n$.

3.2 Proposed Model-II

Our second prosed model is a motivation from the study of Murtaza *et al.* (2020). In this technique, like our first proposed model, every respondent is requested to choose either the direct response or protected response, depending on the respondent's perception of the sensitive question. The researcher notes the choices of each respondent and hence knows the actual number of respondents who selected the true or protected responses. If a respondent chooses to report the protected response, he is provided with two options of sensitive questions. The respondent selects one question at random, unobserved by the researcher, and reports his/her response accordingly.

The two options of questions are:

i. Report true responses Y_i, with probability $1-W$,

ii. Report protected responses $TY+\alpha S$, with probability W.

The overall sample mean can be calculated as:

$$\hat{\mu}_{P2}=\frac{n_1\bar{Y}+n_2\bar{Z}}{n_1+n_2}, \tag{10}$$

where $n_1+n_2=n$.

4. Mean and Variance

In this section, the mathematical proofs are shown that the mean estimator is unbiased along with the derivation of sampling variance under the proposed models.

Theorem 1: The mean estimator $\hat{\mu}_{P1}$ is unbiased for μ_Y.

Proof: Using expected value on equation (9) leads to:

$$E\left(\hat{\mu}_{P1}\right)=E\left(\frac{n_1\bar{Y}+n_2\bar{Z}}{n_1+n_2}\right)=\frac{n_1E\left(\bar{Y}\right)+n_2E\left(\bar{Z}\right)}{n_1+n_2}. \tag{11}$$

Taking expectation of equation (7), we get:

$$E\left(\bar{Y}\right)=E\left(\frac{1}{n_1}\sum_{i=1}^{n1}Y_i\right)=\mu_Y. \tag{12}$$

Now taking expectation both sides of the equation (8), we get:

$$E\left(\bar{Z}\right)=E\left(\frac{1}{n_2}\sum_{i=1}^{n2}Z_i\right). \tag{13}$$

Now,

$$E\left(Z_i\right)=\frac{\alpha}{\alpha+\beta+\gamma}\,E\left(Y-\beta S\right)\frac{\beta}{\alpha+\beta+\gamma}\,E\left(Y+\alpha S\right)+\frac{\gamma}{\alpha+\beta+\gamma}E\left(Y\right),$$

or

$$E\left(Z_i\right)=\frac{\alpha}{\alpha+\beta+\gamma}\,E\left(Y\right)-\beta E\left(S\right)\ \frac{\beta}{\alpha+\beta+\gamma}\,E\left(Y\right)+\alpha E\left(S\right)\frac{\gamma}{\alpha+\beta+\gamma}\,E\left(Y\right),$$

or

$$E\left(Z_i\right)=\frac{\alpha}{\alpha+\beta+\gamma}\left(\mu_Y-\beta\theta\right)\frac{\beta}{\alpha+\beta+\gamma}\left(\mu_Y+\alpha\theta\right)\frac{\gamma}{\alpha+\beta+\gamma}\,\mu_Y,$$

or

$$E\left(Z_i\right)=\frac{\alpha\mu_Y+\beta\mu_Y+\gamma\mu_Y}{\alpha+\beta+\gamma}=\mu_Y. \tag{14}$$

Using equation (14) in equation (13) yields:

$$E\left(\bar{Z}\right)=\frac{1}{n_2}\sum_{i=1}^{n2}\mu_y=\mu_Y\,. \tag{15}$$

Using equation (12) and equation (15) in equation (11), we get:

$$E\left(\hat{\mu}_{P1}\right)=\frac{n_1\mu_Y+n_2\mu_Y}{n_1+n_2}=\mu_Y\,. \tag{16}$$

In a similar way, the unbiasedness of $\hat{\mu}_{P2}$ can be easily proved.

Theorem 2: The variance of the mean estimators $\hat{\mu}_{P1}$ and $\hat{\mu}_{P2}$ are given by:

$$Var\left(\hat{\mu}_{P1}\right)=\frac{n_1\sigma_Y^2+n_2\left[\sigma_Y^2+\dfrac{\alpha\beta\left(\alpha+\beta\right)\left(\sigma_S^2+\theta^2\right)}{\left(\alpha+\beta+\gamma\right)}\right]}{\left(n_1+n_2\right)^2}. \tag{17}$$

and

$$Var\left(\hat{\mu}_{P2}\right)=\frac{n_1\sigma_Y^2+n_2\left[W\sigma_T^2\left(\sigma_Y^2+\mu_Y^2\right)+\sigma_Y^2+\alpha^2\sigma_S^2\right]}{\left(n_1+n_2\right)^2}. \tag{18}$$

Proof: Applying variance on both sides of equation (9) yields:

$$Var\left(\hat{\mu}_{P1}\right)=Var\left(\frac{n_1\bar{Y}+n_2\bar{Z}}{n_1+n_2}\right)=\frac{n_1^2Var\left(\bar{Y}\right)+n_2^2Var\left(\bar{Z}\right)}{\left(n_1+n_2\right)^2}. \tag{19}$$

Now

$$Var\left(\bar{Y}\right)=Var\left(\frac{1}{n_1}\sum_{i=1}^{n1}Y_i\right)=\frac{\sigma_Y^2}{n_1}, \tag{20}$$

and

$$Var\left(\bar{Z}\right)=Var\left(\frac{1}{n_2}\sum_{i=1}^{n2}Z_i\right)=\frac{1}{n_2}Var\left(Z_i\right). \tag{21}$$

The variance may be obtained as follows:

$$Var\left(Z_i\right)=E\left(Zi\right)^2-\left(E\left(Z_i\right)\right)^2. \tag{22}$$

$E\left(Z_i^2\right)$ can be expressed as:

$$E\left(Z_i^2\right)=\frac{\alpha}{\alpha+\beta+\gamma}E\left(Y-\beta S\right)^2+\frac{\beta}{\alpha+\beta+\gamma}E\left(Y+\alpha S\right)^2+\frac{\gamma}{\alpha+\beta+\gamma}E\left(Y\right)^2,$$

or

$$E\left(Z_i^2\right)=\frac{\alpha}{\alpha+\beta+\gamma}E\left(Y^2+\beta^2S^2-2\beta YS\right)+\frac{\beta}{\alpha+\beta+\gamma}E\left(Y^2+\alpha^2S^2+2\alpha YS\right)+\frac{\gamma}{\alpha+\beta+\gamma}E\left(Y\right)^2,$$

or

$$E(Z_i)^2 = \frac{\alpha}{\alpha+\beta+\gamma}\left(E(Y^2)+\beta^2E(S^2)-2\beta E(Y)E(S)\right)+\frac{\beta}{\alpha+\beta+\gamma}\left\{E(Y^2)+\alpha^2E(S^2)+2\alpha E(Y)E(S)\right\}+\frac{\gamma}{\alpha+\beta+\gamma}E(Y)^2$$

Using the assumptions given in Section 2, and the independence of variables, the above equation simplifies to:

$$E(Z_i)^2 = \frac{\alpha}{\alpha+\beta+\gamma}\left\{\left(\sigma_Y^2+\mu_Y^2\right)+\beta^2\left(\sigma_S^2+\theta^2\right)-2\beta\theta\mu_Y\right\}$$

$$+\frac{\beta}{\alpha+\beta+\gamma}\left\{\left(\sigma_Y^2+\mu_Y^2\right)+\alpha^2\left(\sigma_S^2+\theta^2\right)+2\alpha\theta\mu_Y\right\}+\frac{\gamma}{\alpha+\beta+\gamma}\left(\sigma_Y^2+\mu_Y^2\right),$$

or

$$E(Z_i)^2 = \frac{\alpha}{\alpha+\beta+\gamma}\left\{\left(\sigma_Y^2+\mu_Y^2\right)+\beta^2\left(\sigma_S^2+\theta^2\right)-2\beta\theta\mu_Y\right\}+\frac{\beta}{\alpha+\beta+\gamma}\left\{\left(\sigma_Y^2+\mu_Y^2\right)\right.$$

$$\left.+\alpha^2\left(\sigma_S^2+\theta^2\right)+2\alpha\theta\mu_Y\right\}+\frac{\gamma}{\alpha+\beta+\gamma}\left(\sigma_Y^2+\mu_Y^2\right),$$

or

$$E(Z_i)^2 = \sigma_Y^2+\mu_Y^2+\frac{\alpha\beta(\alpha+\beta)\left(\sigma_S^2+\theta^2\right)}{(\alpha+\beta+\gamma)}. \tag{23}$$

Using equation (16) and equation (23) in equation (22) yields:

$$Var(Z_i) = \sigma_Y^2+\mu_Y^2+\frac{\alpha\beta(\alpha+\beta)\left(\sigma_S^2+\theta^2\right)}{(\alpha+\beta+\gamma)}-(\mu_Y)^2.$$

or

$$Var(Z_i) = \sigma_Y^2+\frac{\alpha\beta(\alpha+\beta)\left(\sigma_S^2+\theta^2\right)}{(\alpha+\beta+\gamma)}. \tag{24}$$

Using equation (24) in equation (21) yields:

$$Var(\bar{Z}) = \frac{1}{n_2}\left[\sigma_Y^2+\frac{\alpha\beta(\alpha+\beta)\left(\sigma_S^2+\theta^2\right)}{(\alpha+\beta+\gamma)}\right]. \tag{25}$$

Using equation (20) and equation (25) in equation (19) yields:

$$Var(\hat{\mu}_{P1}) = \frac{n_1^2\left[\frac{\sigma_Y^2}{n_1}\right]+n_2^2\left[\frac{1}{n_2}\left(\sigma_Y^2+\frac{\alpha\beta(\alpha+\beta)\left(\sigma_S^2+\theta^2\right)}{(\alpha+\beta+\gamma)}\right)\right]}{(n_1+n_2)^2}.$$

$Var\left(\hat{\mu}_{P1}\right)$ can be simplified as:

$$Var\left(\hat{\mu}_{P1}\right)=\frac{n_1\sigma_Y^2+n_2\left[\sigma_Y^2+\frac{\alpha\beta\left(\alpha+\beta\right)\left(\sigma_S^2+\theta^2\right)}{\left(\alpha+\beta+\gamma\right)}\right]}{\left(n_1+n_2\right)^2}. \quad (26)$$

5. Model-Quality Evaluation

A general method for comparison of the performance of quality of two models in terms of efficiency is the relative efficiency. Mathematically, the relative efficiency $\left(RE\right)$ may be expressed as follows:

$$RE=\frac{Var\left(\hat{\mu}_{NS}\right)}{Var\left(\hat{\mu}_{P1}\right)}. \quad (27)$$

It may be observed that $RE=1$ indicates that the suggested technique and existing technique are equally efficient. Likewise, $RE>1$ indicates that the suggested technique is more efficient than the existing technique. In relative efficiency measures there have no idea about the privacy of the two models, However, for parameter estimation based on randomized response models, respondent privacy is also very important. To overcome this problem, Yan *et al.* (2008) developed a procedure to measure the privacy level offered by the scrambling technique.

The metric of privacy may be calculated using the formula:

$$\nabla \ =E\left[Z-Y\right]^2. \quad (28)$$

Large values of ∇ suggest a higher level of the respondent-privacy offered by in the model.
Another performance measure was introduced by Gupta *et al.* (2018) to measure the privacy and efficiency of the given technique simultaneously.

The unified metric of privacy and efficiency can be calculated using the formula:

$$\delta \ =\frac{Var\left(\hat{\mu}\ \right)}{\nabla}. \quad (29)$$

A model with a smaller value of δ is interpreted as the better model with respect to the simultaneous consideration of the efficiency and privacy level of the respondents.
Under the Narjis and Shabbir (2023) scrambling technique, the privacy metric can be obtained in the form:

$$\nabla_{NS}=\frac{\alpha\beta\left(\alpha+\beta\right)\left(\sigma_S^2+\theta^2\right)}{\alpha+\beta+\gamma}. \quad (30)$$

The measure of privacy of the suggested and available models are same because the privacy metric is independent of the number of respondents.

The unified measure of efficiency and privacy based on Narjis and Shabbir (2023) technique can be calculated as:

$$\delta_{NS}=\frac{Var\left(\hat{\mu}_{NS}\right)}{\nabla_{NS}}=\frac{1}{n}\left[\frac{\frac{\alpha\beta(\alpha+\beta)(\sigma_S^2+\theta^2)}{\alpha+\beta+\gamma}+\sigma_Y^2}{\frac{\alpha\beta(\alpha+\beta)(\sigma_S^2+\theta^2)}{\alpha+\beta+\gamma}}\right]. \quad (31)$$

The unified measure of efficiency and privacy based on suggested Model-I can be calculated as:

$$\delta_{P1}=\frac{Var\left(\hat{\mu}_{P1}\right)}{\nabla_{P1}}=\frac{\dfrac{n_1\sigma_Y^2+n_2\left[\sigma_Y^2+\dfrac{\alpha\beta\left(\alpha+\beta\right)\left(\sigma_S^2+\theta^2\right)}{\alpha+\beta+\gamma}\right]}{\left(n_1+n_2\right)^2}}{\dfrac{\alpha\beta\left(\alpha+\beta\right)\left(\sigma_S^2+\theta^2\right)}{\alpha+\beta+\gamma}}. \tag{32}$$

Under the Murtaza *et al.* (2020), the measure of privacy can be obtained as:

$$\nabla_M=W\sigma_T^2\left(\sigma_Y^2+\mu_Y^2\right)+\alpha^2\sigma_S^2. \tag{33}$$

The unified measure of efficiency and privacy based on Murtaza *et al.* (2020) can be calculated as:

$$\delta_M=\frac{Var\left(\hat{\mu}_M\right)}{\nabla_M}=\frac{\frac{1}{n}\left[W\sigma_T^2\left(\sigma_Y^2+\mu_Y^2\right)+\sigma_Y^2+\alpha^2\sigma_S^2\right]}{W\sigma_T^2\left(\sigma_Y^2+\mu_Y^2\right)+\alpha^2\sigma_S^2}. \tag{34}$$

The unified measure of efficiency and privacy based on suggested Model-II can be calculated as:

$$\delta_{P2}=\frac{Var\left(\hat{\mu}_{P2}\right)}{\nabla_{P2}}=\frac{n_1\sigma_Y^2+n_2\left[W\sigma_T^2\left(\sigma_Y^2+\mu_Y^2\right)+\sigma_Y^2+\alpha^2\sigma_S^2\right]}{W\sigma_T^2\left(\sigma_Y^2+\mu_Y^2\right)+\alpha^2\sigma_S^2}. \tag{35}$$

6. Empirical Comparison

In this section, we find the conditions by comparing the variances of the proposed models $Var\left(\hat{\mu}_{P1}\right)$ and $Var\left(\hat{\mu}_{P2}\right)$ with the variances of other existing models i.e. $Var\left(\hat{\mu}_{NS}\right)$ and $Var\left(\hat{\mu}_M\right)$.

The suggested scrambling Model-I will be more precise than the Narjis and Shabbir (2023) technique if:

$$Var\left(\hat{\mu}_{p1}\right)<Var\left(\hat{\mu}_{NS}\right),$$

or

$$\frac{n_1\sigma_Y^2+n_2\left[\sigma_Y^2+\dfrac{\alpha\beta\left(\alpha+\beta\right)\left(\sigma_S^2+\theta^2\right)}{\left(\alpha+\beta+\gamma\right)}\right]}{\left(n_1+n_2\right)^2}<\frac{1}{n}\left[\frac{\left(\alpha+\beta\right)\alpha\beta\left(\theta^2+\sigma_S^2\right)}{\alpha+\beta+\gamma}+\sigma_Y^2\right].$$

The proposed scrambling Model-II is more precise than the Murtaza *et al.* (2020) randomized technique if:

$$Var\left(\hat{\mu}_{p2}\right)<Var\left(\hat{\mu}_M\right),$$

or

$$\frac{n_1\sigma_Y^2+n_2\left[W\sigma_T^2\left(\sigma_Y^2+\mu_Y^2\right)+\sigma_Y^2+\alpha^2\sigma_S^2\right]}{\left(n_1+n_2\right)^2}<\frac{1}{n}\left[W\sigma_T^2\left(\sigma_Y^2+\mu_Y^2\right)+\sigma_Y^2+\alpha^2\sigma_S^2\right].$$

The values of variances and privacy are calculated for proposed randomized response models using the various values of constants α, β, γ and θ. The following tables show the different metrics of suggested optional scrambling techniques with respect to existing models.

Table 1. Comparisons of variances of the mean based on the suggested and available techniques

Parameters	β	σ_S^2	α	$Var(\hat{\mu}_{NS})$	$Var(\hat{\mu}_M)$	$Var(\hat{\mu}_{P1})$	$Var(\hat{\mu}_{P2})$
$W=0.1$, $\sigma_Y^2=2$	2	0.1	10	0.159	0.104	0.0195	0.014
			15	0.240	0.129	0.028	0.017
			20	0.322	0.164	0.036	0.020
			25	0.404	0.209	0.044	0.025
	4	0.3	10	0.332	0.144	0.037	0.018
			15	0.502	0.219	0.054	0.026
			20	0.673	0.324	0.071	0.036
			25	0.844	0.459	0.088	0.050
$W=0.5$, $\sigma_Y^2=5$	6	0.5	10	0.527	0.515	0.062	0.061
			15	0.794	0.640	0.088	0.073
			20	1.062	0.815	0.115	0.091
			25	1.330	1.040	0.142	0.113
	8	0.7	10	0.734	0.555	0.082	0.065
			15	1.105	0.730	0.119	0.082
			20	1.477	0.975	0.157	0.107
			25	1.851	1.290	0.194	0.138
$W=0.9$, $\sigma_Y^2=8$	10	0.9	10	0.484	0.795	0.063	0.094
			15	0.963	0.930	0.111	0.107
			20	1.446	1.155	0.159	0.130
			25	1.931	1.470	0.208	0.161
	12	1.1	10	0.604	0.805	0.075	0.095
			15	1.202	0.970	0.135	0.111
			20	1.806	1.245	0.195	0.139
			25	2.412	1.630	0.256	0.177

Table 1 displays the sampling variances using the suggested techniques with respect to the Narjis and Shabbir (2023) model and Murtaza *et al.* (2020) model for different values of the parameters. It is clearly observed that the suggested models have a significant reduction in the variances of the means, making them preferable over the existing models.

Table 2. Comparisons of the percent relative efficiency of the suggested and existing techniques

Parameters	β	σ_S^2	α	PRE_{P1NS}	PRE_{P2NS}	PRE_{P1M}	PRE_{P2M}	PRE_{P1P2}
$W = 0.1$ $\sigma_Y^2 = 2$	2	0.1	10	815.347	1132.198	535.493	743.590	138.861
			15	869.693	1452.667	468.380	782.346	167.032
			20	899.409	1606.215	459.365	820.359	178.586
			25	918.115	1644.821	476.300	853.301	179.152
	4	0.3	10	902.073	1838.243	392.798	800.444	203.780
			15	933.044	1964.232	408.061	859.045	210.519
			20	949.189	1865.988	457.867	900.111	196.588
			25	959.078	1703.130	522.205	927.331	177.580
$W = 0.5$, $\sigma_Y^2 = 5$	6	0.5	10	854.219	871.678	834.190	851.240	102.044
			15	898.175	1087.494	724.088	876.712	121.078
			20	921.854	1173.133	707.659	900.553	127.258
			25	936.628	1177.159	732.297	920.354	125.681
	8	0.7	10	890.757	1137.752	673.666	860.465	127.729
			15	924.666	1347.175	611.041	890.244	145.693
			20	942.577	1387.152	622.082	915.493	147.166
			25	953.631	1341.268	664.623	934.783	140.649
$W = 0.9$, $\sigma_Y^2 = 8$	10	0.9	10	770.756	515.385	1266.252	846.711	66.867
			15	869.902	896.184	840.576	865.972	103.021
			20	909.432	1112.791	726.685	889.180	122.361
			25	930.613	1196.302	708.521	910.803	128.550
	12	1.1	10	807.434	635.977	1077.031	848.325	78.765
			15	893.037	1078.837	720.814	870.783	120.805
			20	926.139	1299.556	638.797	896.358	140.320
			25	943.653	1359.105	637.972	918.846	144.026

Table 2 displays the percent relative efficiency of the existing and suggested techniques. Observing Table 2, all the values are greater than 100 with same level of privacy which means that the suggested techniques are more efficient than the existing techniques.

Table 3. Comparisons of the δ values of the proposed and existing models

Parameters	β	σ_S^2	α	δ_{NS}	δ_M	δ_{P1}	δ_{P2}
$W=0.1$, $\sigma_Y^2=2$	2	0.1	10	0.00205	0.00208	0.00025	0.00028
			15	0.00203	0.00206	0.00023	0.00026
			20	0.00203	0.00205	0.00023	0.00025
			25	0.00202	0.00204	0.00022	0.00024
	4	0.3	10	0.00202	0.00206	0.00022	0.00026
			15	0.00202	0.00204	0.00022	0.00024
			20	0.00201	0.00203	0.00021	0.00023
			25	0.00201	0.00202	0.00021	0.00022
$W=0.5$, $\sigma_Y^2=5$	6	0.5	10	0.00204	0.00204	0.00024	0.00024
			15	0.00203	0.00203	0.00023	0.00023
			20	0.00202	0.00203	0.00022	0.00023
			25	0.00202	0.00202	0.00022	0.00022
	8	0.7	10	0.00203	0.00204	0.00023	0.00024
			15	0.00202	0.00203	0.00022	0.00023
			20	0.00201	0.00202	0.00021	0.00022
			25	0.00201	0.00202	0.00021	0.00022
$W=0.9$, $\sigma_Y^2=8$	10	0.9	10	0.00207	0.00204	0.00027	0.00024
			15	0.00203	0.00204	0.00023	0.00024
			20	0.00202	0.00203	0.00022	0.00023
			25	0.00202	0.00202	0.00022	0.00022
	12	1.1	10	0.00205	0.00204	0.00025	0.00024
			15	0.00203	0.00203	0.00023	0.00023
			20	0.00202	0.00203	0.00022	0.00023
			25	0.00201	0.00202	0.00021	0.00022

Table 3 displays the unified metric of efficiency and privacy of the suggested and available techniques. It is observed that the values of δ based on both of the suggested techniques are smaller than the Narjis and Shabbir (2023) model and the Murtaza *et al.* (2020) model. These findings suggest that the proposed optional models are useful for application in real-world surveys.

7. Discussion and Conclusion

Two novel scrambling techniques have been introduced in this paper. Section 3 presented our proposed optional randomized response models utilizing direct responses. The proposed models are the updated version of the Narjis and Shabbir (2023) and Murtaza *et al.* (2020) quantitative models. The findings of this study suggest that the suggested techniques are more precise than existing techniques in terms of both efficiency and overall performance when privacy protection and efficiency are jointly considered.

Table 1 presents the variances of the developed and existing techniques. We observe that the suggested techniques provide smaller variances than the existing models. Table 2 presents the percentage relative efficiency of the suggested and the already available techniques. Table 2 shows the improvement of the proposed models in term of efficiency over the available techniques. It is observed that the suggested Model-I is superior to the Narjis and Shabbir (2023) model, and model-II is better than the Murtaza *et al.* (2020) quantitative technique. Moreover, one may observe that among the two suggested models, Model-II is the better model in terms of efficiency.

Table 3 displays the unified measure of efficiency and privacy of the suggested and available techniques. All the values of the δ are smaller under the proposed models than the Narjis and Shabbir (2023) model and Murtaza *et al.* (2020) technique.

It is also clearly observed from the tables that as the value of σ_S^2 decreases, the variances of estimators under the proposed models also decrease. Further, we also observe that the values of σ_S^2 are greatly affect the variances of the models. The findings of the efficiency comparison and unified measure clearly show that the proposed models are practically suitable for use with sensitive surveys.

8. Future Research

I suggest future researchers to investigate estimation of population median and population variance under the suggested models. The proposed method can also be utilized in some other sampling plans. Further, the effect of non-sampling errors on different estimators of parameters can also be investigated using the proposed models.

Acknowledgments

I would like to thank reviewers and editors for their comments and suggestions.

Funding Source

The author received no funding for this study.

Conflicts of Interest

The author has no conflict of interest to declare.

Data Availability Statement

The data associated with my study has not been deposited into a publicly available repository. All relevant data is available within the article and its references.

Author's contribution

Conceptualization: AZEEM, M.; **Data curation:** AZEEM, M.; **Formal analysis:** AZEEM, M.; **Funding acquisition:-; Investigation:** AZEEM, M.; **Methodology:** AZEEM, M.; **Project administration:** AZEEM, M. **Software:** AZEEM, M.; **Resources:** AZEEM, M.; **Supervision:** AZEEM, M.; **Validation:** AZEEM, M.; **Visualization:** AZEEM, M..; **Writing - original draft:** AZEEM, M.; **Writing - review and editing:** AZEEM, M.

References

1. Abbasi, A.M., & Asghar, A. An improved quantitative randomized response model under ranked set sampling. *Pakistan Journal of Statistics* **40**, 225-242 (2024).
2. Abbasi, A.M., Shad, M.Y., & Ahmad, A. On partial randomized response model using ranked set sampling. *PLOS ONE* **17**, e0277497. https://doi.org/10.1371/journal.pone.0277497
3. Azeem, M., & Ali, S. A neutral comparative analysis of additive, multiplicative, and mixed quantitative randomized response models. *PLOS ONE* **18**, e0284995 (2023). https://doi.org/10.1371/journal.pone.0284995

4. Azeem, M. Incorporating direct responses into optional randomized response models without compromising respondents' privacy: Accepted-March 2024. *REVSTAT – Statistical Journal* (2024). https://revstat.ine.pt/index.php/REVSTAT/article/view/663

5. Azeem, M. Introducing a weighted measure of privacy and efficiency for comparison of quantitative randomized response models. *Pakistan Journal of Statistics* **39**, 377-385 (2023).

6. Bar-Lev, S. K., Bobovitch, E., & Boukai, B. A note on randomized response models for quantitative data. *Metrika* **60**, 255-260 (2004). https://doi.org/10.1007/s001840300308

7. Bouza-Herrera, C. N. Behavior of some scrambled randomized response models under simple random sampling, ranked set sampling and Rao–Hartley–Cochran designs. In *Handbook of Statistics* **34**, 209-220 (2016). https://doi.org/10.1016/bs.host.2016.01.011

8. Eichhorn, B. H., & Hayre, L. S. Scrambled randomized response methods for obtaining sensitive quantitative data. *Journal of Statistical Planning and Inference* **7**, 307-316 (1983). https://doi.org/10.1016/0378-3758(83)90002-2

9. Franklin, L. R. A. A comparison of estimators for randomized response sampling with continuous distributions from a dichotomous population. *Communications in Statistics-Theory and Methods* **18**, 489-505 (1989). https://doi.org/10.1080/03610928908829913

10. Gupta, S., Gupta, B., & Singh, S. Estimation of sensitivity level of personal interview survey questions. *Journal of Statistical Planning and Inference* **100**, 239-247 (2002). https://doi.org/10.1016/S0378-3758(01)00137-9

11. Gjestvang, C. R., & Singh, S. An improved randomized response model: Estimation of mean. *Journal of Applied Statistics* **36**, 1361-1367 (2009).https://doi.org/10.1080/02664760802684151

12. Gupta, S., Mehta, S., Shabbir, J., & Khalil, S. A unified measure of respondent privacy and model efficiency in quantitative RRT models. *Journal of Statistical Theory and Practice* **12**, 506-511 (2018). https://doi.org/10.1080/15598608.2017.1415175

13. Gupta, S., Zhang, J., Khalil, S., & Sapra, P. Mitigating lack of trust in quantitative randomized response technique models. *Communications in Statistics - Simulation and Computation* **53**, 2624-2632 (2024). https://doi.org/10.1080/03610918.2022.2082477

14. Himmelfarb, S., & Edgell, S. E. Additive constants model: A randomized response technique for eliminating evasiveness to quantitative response questions. *Psychological Bulletin* **87**, 525 (1980). https://doi.org/10.1037/0033-2909.87.3.525

15. Hussain, Z., Al-Sobhi, M. M., Al-Zahrani, B., Singh, H. P., & Tarray, T. A. Improved randomized response in additive scrambling models. *Mathematical Population Studies* **23**, 205-221 (2016). https://doi.org/10.1080/08898480.2015.1087773

16. Kumar, S., & Kour, S. P. The joint influence of estimation of sensitive variable under measurement error and non-response using ORRT models. *Journal of Statistical Computation and Simulation* **92**, 3583-3604 (2022). https://doi.org/10.1080/00949655.2022.2075362

17. Kumar, S., Kour, S. P., & Singh, H. P. Applying ORRT for the estimation of population variance of sensitive variable. *Communications in Statistics-Simulation and Computation* 1-11 (2023). https://doi.org/10.1080/03610918.2023.2292966

18. Lovig, M., Khalil, S., Rahman, S., Sapra, P., & Gupta, S. A mixture binary RRT model with a unified measure of privacy and efficiency. *Communications in statistics-simulation and computation* **52**, 2727-2737 (2023). https://doi.org/10.1080/03610918.2021.1914092

19. Mahzizadeh, M., & Zamanzade, E. On estimating the area under ROC curve in ranked set sampling. *Statistical Methods in Medical Research* **31**, 1500-1514 (2022). https://doi.org/10.1177/09622802221097211

20. Mangat, N. S., & Singh, R. An alternative randomized response procedure. *Biometrika* **77**, 439-442 (1990). https://doi.org/10.1093/biomet/77.2.439

21. Mehta, S., & Aggarwal, P. Bayesian estimation of sensitivity level and population proportion of a sensitive characteristic in a binary optional unrelated question RRT model. *Communications in Statistics-Theory and Methods* **47**, 4021-4028 (2018). https://doi.org/10.1080/03610926.2017.1367812

22. Murtaza, M., Singh, S., & Hussain, Z. An innovative optimal randomized response model using correlated scrambling variables. *Journal of Statistical Computation and Simulation* **90**, 2823-2839 (2020). https://doi.org/10.1080/00949655.2020.1791118

23. Narjis, G., & Shabbir, J. An efficient new scrambled response model for estimating sensitive population mean in successive sampling. *Communications in Statistics-Simulation and Computation* **52**, 5327-5344 (2023). https://doi.org/10.1080/03610918.2021.1986528

24. Pollock, K. H., & Bek, Y. A comparison of three randomized response models for quantitative data. *Journal of the American Statistical Association* **71**, 884-886 (1976). https://www.tandfonline.com/doi/abs/10.1080/01621459.1976.10480963

25. Santiago, A., Sautto, J.M., & Bouza, C.N. Randomized estimation a proportion using ranked set sampling and Warner's procedure. *Investigacion Operacional* **40**, 356-361 (2019). http://dx.doi.org/10.13140/RG.2.2.11102.95043

26. Singh, S., & Grewal, I. S. Geometric distribution as a randomization device: implemented to the Kuk's model. *International Journal of Contemporary Mathematical Sciences* **8**, 243-248 (2013).

27. Singh, S., & Sedory, S. A. A new randomized response device for sensitive characteristics: an application of the negative hypergeometric distribution. *Metron* **71**, 3-8 (2013). https://doi.org/10.1007/s40300-013-0002-3

28. Sanaullah, A., Saleem, I., Gupta, S., & Hanif, M. Mean estimation with generalized scrambling using two-phase sampling. *Communications in Statistics – Simulation and Computation* **51**, 5643-5657 (2022). https://doi.org/10.1080/03610918.2020.1778032

29. Tarray, T. A., & Singh, H. P. A general procedure for estimating the mean of a sensitive variable using auxiliary information. *Investigación Operacional* **36** (2015). https://revistas.uh.cu/invoperacional/article/view/4612

30. Tiwari, N., & Mehta, P. Additive randomized response model with known sensitivity level. *International Journal of Computational and Theoretical Statistics* **4**, 83-93 (2017). http://dx.doi.org/10.12785/IJCTS/040201

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

32. Yan, Z., Wang, J., & Lai, J. An efficiency and protection degree-based comparison among the quantitative randomized response strategies. *Communications in Statistics – Theory and Methods* **38**, 400-408 (2008). https://doi.org/10.1080/03610920802220785

33. Zamanzade, E., & Wang, X. Proportion estimation in ranked set sampling in the presence of tie information. *Computational Statistics* **33**, 1349-1366 (2018). https://doi.org/10.1007/s00180-018-0807-x