*Journal of Interdisciplinary Mathematics*
ISSN: 0972-0502 (Print), ISSN: 2169-012X (Online)
Vol. 28 (2025), No. 5, pp. 1735–1758
DOI : 10.47974/JIM-1743

# Improved mean estimation of sensitive study variable under two phase sampling with sub-sampling the non-respondents

Hafiza Fakarul Nisa [#]
*Department of Statistics*
*Lahore College for Women University*
*Lahore*
*Pakistan*

Sadia Khalil *
*Department of Statistics*
*Lahore College for Women University*
*Lahore*
*Pakistan*

*and*

*Department of Mathematics and Statistics*
*University of North Carolina*
*Greensboro*
*U.S.A.*

Sat Gupta
*Department of Mathematics and Statistics*
*University of North Carolina*
*Greensboro*
*U.S.A.*

**Abstract**

In this study, we present generalized mean estimators aimed at tackling the challenges associated with non-sampling errors in sensitive survey data. Our approach utilizes both full and optional Randomized Response Technique (RRT) models within a two-phase simple random sampling framework, incorporating information from auxiliary variables. To evaluate the performance of our proposed estimators, we employ a unified measure of

# *E-mail:* **Fakharunnisa153@gmail.com**
* *E-mail:* **s_khali2@uncg.edu; (Corresponding author)**
**Sadia.Khalil@icwu.edu.pk**

efficiency and privacy protection. The validity of these estimators is supported by extensive simulations.

***Subject Classification (2010):*** *62D05.*

***Keywords:*** *Randomized response models (RRT), Measurement errors (ME), Non-response, Efficiency, Respondent privacy.*

## 1. Introduction

Auxiliary information plays a crucial role in enhancing the precision of estimators, as demonstrated in ratio, product, and regression estimation methods. When the information of the auxiliary variable at population level is unknown, two-phase sampling can be employed. This method was first introduced by Neyman [1].

In real life, obtaining data for the situations where variables are sensitive such as tax evasions, money laundering, abortions, and any kind of crime etc. Getting truthful response from the respondents is sometimes even impossible. To address sensitive variables, researchers employ Randomized Response Technique (RRT) models, which were first introduced by Warner in 1965 [2]. Then Eichhorn and Hayre [3], Gupta, Gupta, and Singh [4], Gupta and Shabbir [5], Sousa et al. [6], and Gupta et al. [7, 8], are some of the authors who have made significant contributions in this area. In collecting data on such variables, one may face two common problems, i.e., errors in measurement and low response rate.

Non-response is a common issue in surveys, often leading to biased results and reduced data quality. Hansen and Hurwitz [9] were the first researcher who gave the idea of a procedure based on taking a sub-sample of those who do not respond and then obtain the information from this group by contacting them personally. Another type of issue in surveys is the Measurement error. Most of the time, researchers assume ME to be non- existent or very small. But, if these errors are not small enough, there is a chance we may get estimates which are not reliable. Which is a more serious scenario.

The impact of measurement errors on the estimation of population parameters has been addressed by Kumar et al. [10], Khalil, Noor-ul-Amin, and Hanif [11], and Singh, Vishwakarma, and Kim [12]. Also, Singh and Sharma [13], Singh and Vishwakarma [14] and Audu et al. [15] have addressed how the estimation procedure can be affected with errors in measurement and low response rate. Khalil et al. [11, 16] proposed estimators for the unknown population mean using optional RRT models, specifically addressing the issue of measurement error in sensitive study variables. Using the parameters of auxiliary variables, Khalil, Noor-ul-Amin, and Hanif [17] derived a more generalized estimator of unknown mean for the population. Zhang et al. [18] studied the ratio estimation of the mean under RRT models. Following this, Zhang, Khalil, and Gupta, [19] introduced efficient mean estimators using optional RRT models that account for both errors. Subsequently, Shuja, Khalil, and Gupta [20] examined mean estimation while considering the effects of errors in measurement using RRT models in a two-phase sampling framework.

This paper is organized as follows: Section 1 describes various sampling strategies. Section 2 presents a generalized mean estimator that accounts for errors in measurement and low response rate. Section 3 is based on the simulation results, while Section 4 provides the conclusion.

1.1 *Sampling strategy:*

Let $B=(B_1,B_2,B_3,...,B_N)$ constitute a bounded population of $N$ identifiable units. Take a sample having size $n'$ using SRS without replacement at the first phase and obtain the required data on the auxiliary variable *(X)*. In second phase, data on the sensitive variable*(Y)* is collected from a smaller sample of size $n(n<n')$ taken from $n'$. Suppose that at the first phase, there is no measurement error and there is complete response, so $x_i'=X_i'$, that is, the observed value is same as the correct value. While in the second phase, measurement error and non-response effect the study and auxiliary variable. Out of $n$ units, $n_1$ units give a response and $n_2$ units refuse to answer. Now, we apply Hansen Hurwitz's [9] approach to sub-sample the non-respondents, as follows:

- $n_1$ of the sampled units answered the questionnaire by mail,
- $(n_2-n_s)$ units did not respond to the postal questionnaire and were not contacted during personal interviews, and
- $n_s\left(n_s=\frac{n_2}{h}, h>1\right)$ units who were non-respondents to the mailed questionnaire but they gave response later when they were approached by personal interview method.

1.2 *Optional RRT (ORRT) Models*

In full RRT models, everyone has to give a scramble response whatever the question is. However, we can consider that a survey question may be personal or touchy for some individuals, but not for everyone. Gupta et al.

[4] gave the idea that the freedom should be given to the respondent, he can give the direct answer or provide a scramble response if he feels the question is touchy. With this in mind, they proposed an Optional RRT (ORRT) model, which enables researchers to calculate the mean of the sensitive study variable and sensitivity level $W$ (the proportion of individuals who perceive the research question as sensitive).

The respondent can use a very simple RRT model with only one scrambling variable $S$ to answer the sensitive question as $Y + S$ (Gupta et al.) [7], or using two scrambling variables $S$ and $T$, as in Diana and Perri [21], where the reported response is $TY + S$. By taking Var($T$)=0 and E($T$)=1, the Diana and Perri [21] model reduces to the simple additive model. Khalil, Noor-ul-Amin, and Hanif [11] have shown that while the general model offers higher privacy, but the additive model performs better in terms of efficiency. However, Diana and Perri [21] model is better when we use a combined measure $\delta = \frac{Var(\hat{\mu})}{\Delta_{DP}}$ proposed by Gupta et al. [22] for efficiency and privacy protection, where $\Delta_{DP}$ is the measure showing the privacy level proposed by Yan, Wang, and Lai [23] and given by $E$ $E(Z-Y)^2$. $\hat{\mu}$ is the mean estimator based on the same model. The model having smaller δ value is favorable over others as it shows a higher level of privacy, a smaller value of $Var(\hat{\mu})$, or both. One can notice that:

$$\delta_{additive\,RRT} = 1 + \frac{\sigma_y^2}{\sigma_s^2} > 1 + \frac{\sigma_y^2}{\sigma_s^2 + \sigma_T^2(\mu_y^2 + \sigma_y^2)} = \delta_{general\ RRT} \tag{1}$$

As a result, when employing the Diana and Perri [21] model, the multiplicative scrambling variable $T$ serves to decrease model efficiency while enhancing privacy. Overall, this general model is regarded as the most effective due to its unified measure.

Hence, we decided to use the this model. The reported response will be of this form

$$Z = \begin{cases} Y & \text{for probability } 1 - W \\ TY + S & \text{for probability } W \end{cases}$$

where, the two scrambling variables $T$, $S$ with their known means and variances as $\mu_T$, $\mu_S$, and $\sigma_T^2$, $\sigma_S^2$ respectively. The reported response Z has mean and variance as,

$$E(Z) = E(Y)(1-W) + E(TY+S)W = E(Y) \tag{2}$$

$$Var(Z) = E(Z^2) - (E(Z))^2 = \sigma_y^2 + \sigma_s^2 W + \sigma_T^2(\sigma_y^2 + \mu_y^2)W \tag{3}$$

Var($Z$) will increase with the increased values of $W$, also ORRT model serves as the full RRT model at the highest value of W which is one.

### 1.3 *Existing Mean Estimators*

Below are the existing mean estimators, adapted for Optional RRT (ORRT) under two-phase sampling to examine the effects of errors in measurement and low response rate on the estimation procedure.

1) **Ordinary Mean Estimator**

$$\hat{\mu}_{ord} = \hat{\bar{y}}^* = w_1\bar{y}_1 + w_2\bar{y}_2^*, \text{ where } \bar{y}_2^* = \frac{1}{n_s}\sum_{i=1}^{n_s} z_i \tag{4}$$

The expression of $MSE$ of $\hat{\mu}_{ord}$ is:

$$MSE(\hat{\mu}_{ord}) = \lambda(S_Z^2 + S_U^2) + \theta(S_{Z(2)}^2 + S_{U(2)}^2) + H \tag{5}$$

MSE without measurement error and non-response is,

$$MSE^*(\hat{\mu}_{ord}) = \lambda * S_Z^2 \tag{6}$$

$$\text{where } \lambda = \frac{N-n}{Nn}, \theta = \frac{W_2(h-1)}{n} \; and \; H = \frac{W_2 h}{n}\left[\frac{\{(S_{y(2)}^2 + \mu_{y(2)}^2)S_T^2 + S_s^2\}W}{(\mu_T W + 1 - W)^2}\right].$$

2) **Traditional Ratio Estimator**

$$\hat{\mu}_{rt} = \frac{\hat{\bar{z}}^*}{\hat{\bar{x}}^*}\bar{x}' \tag{7}$$

The expression of $MSE$ of $\hat{\mu}_{rt}$ is given by

$$MSE^*(\hat{\mu}_{rt}) \approx \begin{bmatrix} \lambda S_Z^2 + (\lambda-\lambda')\{R^2S_X^2 - 2R\rho_{ZX}S_ZS_X\} + \lambda\{S_U^2 + R^2S_V^2\} + \theta\{S_{U(2)}^2 + R^2S_{V(2)}^2\} + \\ \theta\{S_{Z(2)}^2 + R^2S_{X(2)}^2 - 2R\rho_{ZX(2)}S_{Z(2)}S_{X(2)}\} + H \end{bmatrix} \tag{8}$$

$$\text{where } R = \frac{\bar{Z}}{\bar{X}}, \; \rho_{ZX} = \frac{\rho_{YX}}{\sqrt{1 + \frac{S_s^2 W}{S_Y^2} + \frac{S_T^2(S_Y^2 + \mu_Y^2)W}{S_Y^2}}} \text{ and } H = \frac{W_2 h}{n}\left[\frac{\{(S_{y(2)}^2 + \mu_{y(2)}^2)S_T^2 + S_s^2\}W}{(\mu_T W + 1 - W)^2}\right]$$

The expression of $MSE$ without ME and Non-response is given by,

$$MSE^*(\hat{\mu}_{rt}) \approx \lambda S_Z^2 + (\lambda - \lambda')\{R^2S_X^2 - 2R\rho_{ZX}S_ZS_X\} \tag{9}$$

3) **Regression Estimator Diana, Riaz, and Shabbir [24]**

$$\hat{\mu}_{reg} = \hat{\bar{z}}^{*} - b_{ZX}(\bar{x} - \bar{x}_1) \tag{10}$$

The expression of $MSE$ of $\hat{\mu}_{reg}$ is,

$$MSE^{*}(\hat{\mu}_{reg}) = S_z^2[\lambda - (\lambda - \lambda')\rho_{ZX}^2] + \lambda S_U^2 + \theta(S_{z(2)}^2 + S_{u(2)}^2) \tag{11}$$

MSE without errors is given by,

$$MSE^{*}(\hat{\mu}_{reg}) = S_Z^2[\lambda(1-\rho_{ZX}^2) + \lambda'\rho_{ZX}^2] \tag{12}$$

## 2. Proposed generalized RRT mean estimator

We have developed a generalized RRT mean estimator for the sensitive study variable using Optional RRT (ORRT) to estimate the unknown population mean under two-phase simple random sampling without replacement. This is applicable when both $Y$ and $X$ are affected by errors, as outlined in the following situations:

i. ME in both Y and X with complete response.
ii. Non-response on Y only, along with ME in both Y and X
iii. Non-response and ME in Y only, and
iv. Non-response and ME in both Y and X.

Following Shuja, Khalil, and Gupta [20], the proposed estimator is given by,

$$\hat{\mu}_G = [\hat{\bar{z}} - k(\bar{x} - \bar{x}')]\left[\frac{(\alpha_i \bar{x}' + \beta_i)}{\pi(\alpha_i \bar{x} + \beta_i) + (1-\pi)(\alpha_i \bar{x}' + \beta_i)}\right]^{v}, \tag{13}$$

where $v$ and $k$ are suitably chosen constants and $\pi$ is intended to be a mixing constant determined through optimality consideration. We can generate a variety of estimators by changing the values of known parameters of auxiliary variable $\alpha_i(\alpha_i \neq 0)$ and $\beta_i$. We also get various RRT ratio mean estimators and various RRT product mean estimators by putting $v=1$ and $v=-1$ respectively.

The following notations are employed to derive the bias and MSE for all the proposed RRT mean estimators.

The following notation is used for first phase variable.

$$\bar{x}' = \bar{X} + \bar{X}' \tag{14}$$

When variable is free of errors

$$\tilde{\bar{x}} = \bar{X} + \tilde{\bar{X}} \tag{15}$$

When variables are contaminated with ME

$$U_i = z_i - Z_i;\ \ V_i = x_i - X_i;\hat{\bar{x}} = \bar{X} + \hat{\bar{X}};\ \hat{\bar{z}} = \bar{Z} + \hat{\bar{Z}} \tag{16}$$

When variables are contaminated with ME and non-response

$$U_i^* = z_i^* - Z_i^*;\ V_i^* = x_i^* - X_i^*;\hat{\bar{z}}^* = \bar{Z} + \hat{\bar{Z}}^*;\ \hat{\bar{x}}^* = \bar{X} + \hat{\bar{X}}^* \tag{17}$$

where,

$$\bar{X}' = \frac{1}{n'}\Omega_{X'};\ \hat{\bar{X}} = \frac{1}{n}(\Omega_X + \Omega_V);\ \hat{\bar{Z}} = \frac{1}{n}(\Omega_Z + \Omega_U);\tilde{\bar{X}} = \frac{1}{n}\Omega_X;\hat{\bar{X}}^* = \frac{1}{n}(\Omega_X^* + \Omega_V^*);$$
$$\hat{\bar{Z}}^* = \frac{1}{n}(\Omega_Z^* + \Omega_U^*)$$

$$\begin{aligned}
&E(\bar{X}')^2 = \lambda' S_X^2, E(\tilde{\bar{X}})^2 = \lambda S_X^2, E(\hat{\bar{Z}})^2 = \lambda(S_Z^2 + S_U^2), E(\hat{\bar{X}})^2 = \lambda(S_X^2 + S_V^2),\\
&E(\hat{\bar{Z}}^*)^2 = \lambda(S_Z^2 + S_U^2) + \theta(S_{Z(2)}^2 + S_{U(2)}^2) + H,\\
&E(\hat{\bar{X}}^*)^2 = \lambda(S_X^2 + S_V^2) + \theta(S_{X(2)}^2 + S_{V(2)}^2), E(\hat{\bar{Z}}\bar{X}') = \lambda' \rho_{ZX} S_Z S_X,\\
&E(\hat{\bar{Z}}\hat{\bar{X}}) = \lambda \rho_{ZX} S_Z S_X,\ E(\hat{\bar{X}}\bar{X}') = \lambda' S_X^2, E(\bar{X}'\tilde{\bar{X}}) = \lambda' S_X^2,\\
&E(\hat{\bar{Z}}^*\hat{\bar{X}}) = \lambda \rho_{ZX} S_Z S_X, E(\hat{\bar{Z}}^*\bar{X}') = \lambda' \rho_{ZX} S_Z S_X,\\
&E(\hat{\bar{Z}}^*\hat{\bar{X}}^*) = \lambda \rho_{ZX} S_Z S_X + \theta \rho_{ZX(2)} S_{X(2)} S_{Z(2)},\\
&E(\hat{\bar{X}}^*\bar{X}') = \lambda' S_X^2, E(\hat{\bar{Z}}^*\tilde{\bar{X}}) = \lambda \rho_{ZX} S_Z S_X,
\end{aligned} \tag{18}$$

$$\lambda' = \frac{(1-f')}{n'}, \lambda = \frac{(1-f)}{n}, \theta = \frac{W_2(h-1)}{n}, S_Z^2 = S_Y^2 + S_S^2, f' = \frac{n'}{N}, f = \frac{n}{N},$$
$$H = \frac{W_2 h}{n}\left[\frac{\{(S_{y(2)}^2 + \mu_{y(2)}^2)S_T^2 + S_s^2\}W}{(\mu_T W + 1 - W)^2}\right], \rho_{ZX} = \frac{\rho_{YX}}{\left(\sqrt{1 + \frac{S_S^2 W}{S_Y^2} + \frac{S_T^2(S_Y^2 + \mu_Y^2)W}{S_Y^2}}\right)},$$

**Case-1:** When ($Y$) and ($X$) are contaminated with measurement error

The estimator is expressed as:

$$\hat{\mu}_{G_1} = [\hat{\bar{z}} - k(\hat{\bar{x}} - \bar{x}')]\left[\frac{(\alpha_i \bar{x}' + \beta_i)}{\pi_1(\alpha_i \hat{\bar{x}} + \beta_i) + (1-\pi_1)(\alpha_i \bar{x}' + \beta_i)}\right]^{\nu}, \tag{19}$$

By using eq- (16) and (18) in eq- (19), the estimator will take a form as:

$$\hat{\mu}_{G_1} = [(\bar{Z}+\hat{\bar{Z}}) - k(\bar{X}+\hat{\bar{X}}-\bar{X}-\bar{X}')]\left[\frac{\alpha_i(\bar{X}+\bar{X}')+\beta_i}{\pi_1(\alpha_i\bar{X}+\alpha_i\hat{\bar{X}}+\beta_i)+(1-\pi_1)(\alpha_i\bar{X}+\alpha_i\bar{X}'+\beta_i)}\right]^v, \quad (20)$$

After simplification, we get the bias (after getting second order approximation) and *MSE* (after getting first order approximation) as:

$$Bias(\hat{\mu}_{G_1}) \approx \left[\begin{array}{l} \lambda' S_X^2\left(\frac{v(v+1)}{2!}\omega_i^2\bar{Z} - v^2\omega_i^2\bar{Z} + \frac{v(v-1)}{2!}\omega_i^2\bar{Z}\right) + \lambda S_V^2\left(v\pi_1\omega_i k + \frac{v(v+1)}{2!}\omega_i^2\pi_1^2\bar{Z}\right) \\ +(\lambda-\lambda')S_X^2\left(v\pi_1 k\omega_i + \frac{v(v+1)}{2!}\omega_i^2\pi_1^2\bar{Z}\right) - v\pi_1\omega_i\rho_{ZX}S_Z S_X(\lambda-\lambda') \end{array}\right] \quad (21)$$

where $\omega_i = \frac{\alpha_i}{\alpha_i\bar{X}+\beta_i}$.

$$MSE(\hat{\mu}_{G_1}) \approx \left[\begin{array}{l} \{\lambda S_Z^2 + (\lambda-\lambda')(k^2S_X^2 + v^2\omega_i^2\pi_1^2\bar{Z}^2S_X^2 + 2v\omega_i\pi_1 k\bar{Z}S_X^2 - 2\rho_{ZX}S_Z S_X(k+v\omega_i\pi_1\bar{Z}))\} \\ +\lambda\{S_U^2 + k^2S_V^2 + v^2\pi_1^2\omega_i^2\bar{Z}^2S_V^2 + 2kv\pi_1\omega_i\bar{Z}S_V^2\} \end{array}\right] \quad (22)$$

$$\pi_1 = \pi_{opt_1} = \frac{(\lambda-\lambda')\rho_{ZX}S_Z S_X - k[(\lambda-\lambda')S_X^2 + \lambda S_V^2]}{[S_X^2(\lambda-\lambda') + \lambda S_V^2]\bar{Z}\omega_i v}$$

Substituting the value of $\pi_{opt_1}$ in eq (22), the $MSE^*(\hat{\mu}_{G_1})$ has the following minimum value:

$$MSE_{\min}(\hat{\mu}_{G_1}) \approx [\lambda S_Z^2 + P_1^2\rho_{ZX}^2S_Z^2S_X^2(S_X^2(\lambda-\lambda') + \lambda S_V^2) - 2P_1\rho_{ZX}^2S_Z^2S_X^2(\lambda-\lambda') + \lambda S_U^2]$$

where $P_1 = \frac{(\lambda-\lambda')}{(\lambda-\lambda')S_X^2 + \lambda S_V^2}$

By putting $S_U^2 = S_V^2 = 0$ in eq- (23), the *MSE* without measurement error can be obtained as:

$$MSE^*_{\min}(\hat{\mu}_{G_1}) = S_Z^2[\lambda(1-\rho_{ZX}^2) + \lambda'\rho_{ZX}^2], \quad (24)$$

The suggested RRT mean estimators $(\hat{\mu}_{G_1})$ will perform more efficient versus the ordinary $(\hat{\mu}_{ord})$ estimator when, $MSE(\hat{\mu}_{G_1}) < MSE^*(\hat{\mu}_{ord})$,
or if

$$\lambda S_Z^2 + P_1^2\rho_{ZX}^2S_Z^2S_X^2(S_X^2(\lambda-\lambda') + \lambda S_V^2) - 2P_1\rho_{ZX}^2S_Z^2S_X^2(\lambda-\lambda') + \lambda S_U^2 < \lambda(S_Z^2 + S_U^2) \quad (25)$$

or
$$-\frac{((\lambda-\lambda')\rho_{ZX}S_X S_Z)^2}{(\lambda-\lambda')S_X^2 + \lambda_2 S_V^2} < 0 \quad (26)$$

**Case-2: When both variables are having measurement errors, but only ($Y$) is affected by non-response error.**

The estimator is defined as:

$$\hat{\mu}_{G_2} = [\hat{\bar{z}}^* - k(\hat{\bar{x}} - \bar{x}')]\left[\frac{(\alpha_i \bar{x}' + \beta_i)}{\pi_2(\alpha_i \hat{\bar{x}} + \beta_i) + (1-\pi_2)(\alpha_i \bar{x}' + \beta_i)}\right]^v, \tag{27}$$

By using eq (16), (17) and (18) in eq (27), the estimator can be written as:

$$\hat{\mu}_{G_2} = \left[(\bar{Z} + \hat{\bar{Z}}^*) - k(\bar{X} + \hat{\bar{X}} - \bar{X} - \bar{X}')\right]\left[\frac{\alpha_i(\bar{X} + \bar{X}') + \beta_i}{\pi_2(\alpha_i \bar{X} + \alpha_i \hat{\bar{X}} + \beta_i) + (1-\pi_2)(\alpha_i \bar{X} + \alpha_i \bar{X}' + \beta_i)}\right]^v \tag{28}$$

After simplification, the bias and $MSE$ are given by:

$$Bias(\hat{\mu}_{G_2}) \approx \left[\begin{array}{l} \lambda' S_X^2 \left\{\frac{v(v+1)}{2!}\omega_i^2 \bar{Z} - v^2 \omega_i^2 \bar{Z} + \frac{v(v-1)}{2!}\omega_i^2 \bar{Z}\right\} + \lambda S_V^2 \left\{v\pi_2 \omega_i k + \frac{v(v+1)}{2!}\omega_i^2 \pi_2^2 \bar{Z}\right\} + \\ (\lambda - \lambda') S_X^2 \left\{v\pi_2 k \omega_i + \frac{v(v+1)}{2!}\omega_i^2 \pi_2^2 \bar{Z}\right\} - v\pi_2 \omega_i \rho_{ZX} S_Z S_X (\lambda - \lambda') \end{array}\right] \tag{29}$$

$$MSE(\hat{\mu}_{G_2}) \approx \left[\begin{array}{l} \{\lambda S_Z^2 + (\lambda - \lambda')((k + v\omega_i \pi_3 \bar{Z})^2 S_X^2 - 2\rho_{ZX} S_Z S_X (k + v\omega_i \pi_3 \bar{Z}))\} + \\ \{\lambda (S_U^2 + (k + v\omega_i \pi_3 \bar{Z})^2 S_V^2)\} + \{\theta(S_{Z(2)}^2 + S_{U(2)}^2)\} + H \end{array}\right] \tag{30}$$

$$\pi_2 = \pi_{opt_2} = \frac{(\lambda - \lambda')\rho_{ZX} S_Z S_X - k[(\lambda - \lambda') S_X^2 + \lambda S_V^2]}{\bar{Z}\omega_i v[S_X^2(\lambda - \lambda') + \lambda S_V^2]}$$

The minimum MSE of $\hat{\mu}_{G_2}$ is obtained by putting the value of $\pi_{opt_2}$ in eq (30),

$$MSE_{\min}(\hat{\mu}_{G_2}) \approx \left[\begin{array}{l} \lambda S_Z^2 + P_2^2 \rho_{ZX}^2 S_Z^2 S_X^2 (S_X^2(\lambda - \lambda') + \lambda S_V^2) - 2P_2 \rho_{ZX}^2 S_Z^2 S_X^2 (\lambda - \lambda') + \lambda S_U^2 \\ + \theta(S_{Z(2)}^2 + S_{U(2)}^2) + H \end{array}\right] \tag{31}$$

Where, $P_2 = \frac{(\lambda - \lambda')}{(\lambda - \lambda') S_X^2 + \lambda S_V^2}$

We may obtain the $MSE$ without non-response and measurement error by setting $S_U^2 = S_V^2 = S_{U(2)}^2 = S_{Z(2)}^2 = 0$ in eq-(31) as:

$$MSE_{\min}^*(\hat{\mu}_{G_2}) = S_Z^2[\lambda(1 - \rho_{ZX}^2) + \lambda' \rho_{ZX}^2], \tag{32}$$

The suggested RRT mean estimators $(\hat{\mu}_{G_2})$ will perform more efficiently than the ordinary RRT mean estimator $(\hat{\mu}_{ord})$ when $MSE(\hat{\mu}_{G_2}) <$ $MSE^*(\hat{\mu}_{ord})$,
or if

$$\lambda S_Z^2 + P_2^2\rho_{ZX}^2 S_Z^2 S_X^2 (S_X^2(\lambda-\lambda') + \lambda S_V^2) + \lambda S_U^2 - 2P_2\rho_{ZX}^2 S_Z^2 S_X^2(\lambda-\lambda') + \theta(S_{Z(2)}^2 + S_{U(2)}^2) + G < \lambda(S_Z^2 + S_U^2) + \theta(S_{Z(2)}^2 + S_{U(2)}^2) + G \quad (33)$$

or

$$-\frac{((\lambda-\lambda')\rho_{ZX}S_X S_Z)^2}{(\lambda-\lambda')S_X^2 + \lambda S_V^2} < 0 \quad (34)$$

**Case-3:** When Non-response and ME on (Y) only:
The estimator is defined as:

$$\hat{\mu}_{G_3} = [\hat{\bar{z}}^* - k(\tilde{\bar{x}} - \bar{x}')]\left[\frac{(\alpha_i \bar{x}' + \beta_i)}{\pi_3(\alpha_i \tilde{\bar{x}} + \beta_i) + (1-\pi_3)(\alpha_i \bar{x}' + \beta_i)}\right]^{v}, \quad (35)$$

By putting eq- (16), (17) and (18) in eq- (35), the estimator can be written as follows,

$$\hat{\mu}_{G_3} = [(\bar{Z} + \hat{\bar{Z}}^*) - k(\bar{X} + \tilde{\bar{X}} - \bar{X} - \bar{X}')]\left[\frac{\alpha_i(\bar{X} + \bar{X}') + \beta_i}{\pi_3(\alpha_i \bar{X} + \alpha_i \tilde{\bar{X}} + \beta_i) + (1-\pi_3)(\alpha_i \bar{X} + \alpha_i \bar{X}' + \beta_i)}\right]^{v}, \quad (36)$$

After simplification, the bias and *MSE* will be:

$$Bias(\hat{\mu}_{G_3}) \approx \left[\begin{array}{l} \lambda' S_X^2\left\{\frac{v(v+1)}{2!}\omega_i^2\bar{Z} - v^2\omega_i^2\bar{Z} + \frac{v(v-1)}{2!}\omega_i^2\bar{Z}\right\} - v\pi_3\omega_i\rho_{ZX}S_Z S_X(\lambda-\lambda') + \\ (\lambda-\lambda')S_X^2\left\{v\pi_3 k\omega_i + \frac{v(v+1)}{2!}\omega_i^2\pi_3^2\bar{Z}\right\} \end{array}\right] \quad (37)$$

$$MSE(\hat{\mu}_{G_3}) \approx \left[\begin{array}{l} \lambda S_Z^2 + (\lambda-\lambda')\{(k + v\omega_i\pi_3\bar{Z})^2 S_X^2 - 2\rho_{ZX}S_Z S_X(k + v\omega_i\pi_3\bar{Z})\} + \lambda S_U^2 + \\ \theta(S_{Z(2)}^2 + S_{U(2)}^2) + H \end{array}\right] \quad (38)$$

$$\pi_3 = \pi_{opt_3} = \frac{(\lambda-\lambda')\rho_{ZX}S_Z S_X - k(\lambda-\lambda')S_X^2}{S_X^2(\lambda-\lambda')\bar{Z}\omega_i v}$$

The minimum MSE of $\hat{\mu}_{G_3}$ is obtained by putting the value of $\pi_{opt_3}$ in eq (38)

$$MSE_{\min}(\hat{\mu}_{G_3}) \approx [\lambda S_Z^2 - \rho_{ZX}^2 S_Z^2(\lambda-\lambda')] + [\lambda S_U^2] + [\theta(S_{Z(2)}^2 + S_{U(2)}^2)] + H \quad (39)$$

We may obtain the MSE without measurement error and non-response by setting. $S_U^2 = S_{U(2)}^2 = S_{Z(2)}^2 = 0.$ in eq-(39)

$$MSE_{\min}^{*}(\hat{\mu}_{G_3}) = S_Z^2[\lambda(1-\rho_{ZX}^2) + \lambda'\rho_{ZX}^2], \tag{40}$$

The suggested RRT mean estimators $(\hat{\mu}_{G_3})$ will perform more efficiently than the ordinary RRT mean estimator $(\hat{\mu}_{ord})$ when $MSE^{*}(\hat{\mu}_{G_3}) < MSE^{*}(\hat{\mu}_{ord})$
or if

$$\lambda S_Z^2 - \rho_{ZX}^2 S_Z^2(\lambda-\lambda') + \lambda S_U^2 + \theta(S_{Z(2)}^2 + S_{U(2)}^2) + G < \lambda(S_Z^2 + S_U^2) + \theta(S_{Z(2)}^2 + S_{U(2)}^2) + G \tag{41}$$

or

$$-\rho_{ZX}^2 S_Z^2(\lambda-\lambda') < 0 \tag{42}$$

**Case-4:** When measurement error and non-response on both variables,

The estimator is defined as:

$$\hat{\mu}_{G_4} = [\hat{\bar{z}}^{*} - k(\hat{\bar{x}}^{*} - \bar{x}')]\left[\frac{(\alpha_i\bar{x}'+\beta_i)}{\pi_4(\alpha_i\hat{\bar{x}}^{*}+\beta_i)+(1-\pi_4)(\alpha_i\bar{x}'+\beta_i)}\right]^{v}, \tag{43}$$

By putting eq- (17) and (18) in eq- (43) it becomes,

$$\hat{\mu}_{G_4} = [(\bar{Z}+\hat{\bar{Z}}^{*}) - k(\bar{X}+\hat{\bar{X}}^{*} - \bar{X} - \bar{X}')]\left[\frac{\alpha_i(\bar{X}+\bar{X}')+\beta_i}{\pi_4(\alpha_i\bar{X}+\alpha_i\hat{\bar{X}}^{*}+\beta_i)+(1-\pi_4)(\alpha_i\bar{X}+\alpha_i\bar{X}'+\beta_i)}\right]^{v} \tag{44}$$

The expressions of the bias and *MSE* are:

$$Bias(\hat{\mu}_{G_4}) \approx \begin{bmatrix} \lambda' S_X^2\left\{\frac{v(v+1)}{2!}\omega_i^2\bar{Z} - v^2\omega_i^2\bar{Z} + \frac{v(v-1)}{2!}\omega_i^2\bar{Z}\right\} + \lambda S_V^2\left\{v\pi_4\,\omega_i k + \frac{v(v+1)}{2!}\omega_i^2\pi_4^2\bar{Z}\right\} + \\ (\lambda-\lambda')S_X^2\left\{v\pi_4 k\,\omega_i + \frac{v(v+1)}{2!}\omega_i^2\pi_4^2\bar{Z}\right\} - v\pi_4\,\omega_i\rho_{ZX}S_Z S_X(\lambda-\lambda') + \\ \theta(S_{X(2)}^2 + S_{V(2)}^2)\left\{\pi_4 v\omega_i k + \frac{v(v+1)}{2!}\pi_4^2\omega_i^2\bar{Z}\right\} - \theta\pi_4 v\omega_i\rho_{ZX(2)}S_{X(2)}S_{Z(2)} \end{bmatrix} \tag{45}$$

$$MSE(\hat{\mu}_{G_4}) \approx \begin{bmatrix} \lambda S_Z^2 + (\lambda-\lambda')\{(k+v\omega_i\pi_4\bar{Z})^2 S_X^2 - 2\rho_{ZX}S_Z S_X(k+v\,\omega_i\,\pi_4\bar{Z})\} + \\ \lambda\{S_U^2 + (k+v\omega_i\pi_4\bar{Z})^2 S_V^2\} + \theta\{S_{U(2)}^2 + (k+v\omega_i\pi_4\bar{Z})^2 S_{V(2)}^2\} + \\ \theta\{S_{Z(2)}^2 + (k+v\omega_i\pi_4\bar{Z})^2 S_{X(2)}^2 - 2\rho_{ZX(2)}S_{Z(2)}S_{X(2)}(k+v\,\omega_i\,\pi_4\bar{Z})\} + H \end{bmatrix} \tag{46}$$

$$\pi_4 = \pi_{opt_4} = \frac{(\lambda-\lambda')\rho_{ZX}S_Z S_X - k((\lambda-\lambda')S_X^2 + \lambda S_V^2) + \theta(\rho_{ZX(2)}S_{Z(2)}S_{X(2)} - k(S_{X(2)}^2 + S_{V(2)}^2))}{v\omega_i \bar{Z}[S_X^2(\lambda-\lambda') + \lambda S_V^2 + \theta(S_{X(2)}^2 + S_{V(2)}^2)]}.$$

The minimum MSE of $\hat{\mu}_{G_4}$ is obtained by putting the value of $\pi_{opt4}$ in eq (46)

$$MSE_{\min}(\hat{\mu}_{G_4}) \approx \begin{bmatrix} \lambda S_Z^2 + (\lambda-\lambda')(P_3^2 S_X^2 - 2P_3\rho_{ZX}S_X S_Z) + \theta(S_{Z(2)}^2 + P_3^2 S_{X(2)}^2 - \\ 2P_3\rho_{ZX(2)}S_{X(2)}S_{Z(2)}) + \lambda(S_U^2 + P_3^2 S_V^2) + \theta(S_{U(2)}^2 + P_3^2 S_{V(2)}^2) + H \end{bmatrix} \tag{47}$$

Where, $P_3 = \frac{(\lambda-\lambda')\rho_{ZX}S_X S_Z + \theta(\rho_{ZX(2)}S_{X(2)}S_{Z(2)})}{(\lambda-\lambda')S_X^2 + \lambda S_V^2 + \theta(S_{X(2)}^2 + S_{V(2)}^2)}$

We can get the expression of *MSE* without *ME* and non-response by putting $S_U^2 = S_V^2 = S_{U(2)}^2 = S_{V(2)}^2 = 0$ in eq (47)

$$MSE_{\min}^{*}(\hat{\mu}_{G_4}) = S_Z^2[\lambda(1-\rho_{ZX}^2) + \lambda'\rho_{ZX}^2], \tag{48}$$

The suggested RRT mean estimators $(\hat{\mu}_{G_4})$ will perform more efficiently than the ordinary RRT mean estimator $(\hat{\mu}_{ord})$ when $MSE_{\min}(\hat{\mu}_{G_4})$ $< MSE^{*}(\hat{\mu}_{ord})$,
or if

$$\lambda S_Z^2 + (\lambda-\lambda')(P_3^2 S_X^2 - 2P_3\rho_{ZX}S_X S_Z) + \theta(S_{Z(2)}^2 + P_3^2 S_{X(2)}^2 - 2P_3\rho_{ZX(2)}S_{X(2)}S_{Z(2)}) + \lambda(S_U^2 + P_3^2 S_V^2) + \theta(S_{U(2)}^2 + P_3^2 S_{V(2)}^2) + G < \lambda(S_Z^2 + S_U^2) + \theta(S_{Z(2)}^2 + S_{U(2)}^2) + G \tag{49}$$

or
$$-\frac{((\lambda-\lambda')\rho_{ZX}S_X S_Z + \theta\rho_{ZX(2)}S_{X(2)}S_{Z(2)})^2}{(\lambda-\lambda')S_X^2 + \lambda S_V^2 + \theta(S_{X(2)}^2 + S_{V(2)}^2)} < 0 \tag{50}$$

**Remark 1:** Note that the $MSE_{\min}^{*}(\hat{\mu}_{G_1})$, $(\hat{\mu}_{G_2})$, $(\hat{\mu}_{G_3})$ and $(\hat{\mu}_{G_4})$ expressions given in eq's- (24),(32),(40) and (48) are the same as given in eq- (12) which represents exactly the variance of the linear regression estimator.

**Remark 2:** From the generalized RRT ratio estimator given in eq-(13), by putting $v = 1$ and different choices of unknown constants $\alpha_i, \beta_i$, one can have a number of ratio estimators. Also taking $v = -1$ will produce several product estimators. The classes of ratio and product estimators are listed below in Table 1.

**Table 1**

**Ratio and Product RRT Mean Estimators.**

| **RRT Ratio Estimators** ($v$ = 1 and π = 1) | **RRT product Estimators** ($v$ = -1 and π = 1) |
|---|---|
| If we put $k = 0$ | |
| $\hat{\mu}_{GR}^{1} = \hat{\hat{z}}\left[\frac{(\alpha_i \bar{x}' + \beta_i)}{(\alpha_i \bar{x} + \beta_i)}\right].$ | $\hat{\mu}_{GP}^{1} = \hat{\hat{z}}\left[\frac{(\alpha_i \bar{x} + \beta_i)}{(\alpha_i \bar{x}' + \beta_i)}\right].$ |
| If we put $k = 1$ | |
| $\hat{\mu}_{GR}^{2} = [\hat{\hat{z}} - (\bar{x} - \bar{x}')]\left[\frac{(\alpha_i \bar{x}' + \beta_i)}{(\alpha_i \bar{x} + \beta_i)}\right].$ | $\hat{\mu}_{GP}^{2} = [\hat{\hat{z}} - (\bar{x} - \bar{x}')]\left[\frac{(\alpha_i \bar{x} + \beta_i)}{(\alpha_i \bar{x}' + \beta_i)}\right].$ |
| If we put $k = b_{ZX}$ | |
| $\hat{\mu}_{GR}^{3} = [\hat{\hat{z}} - b_{ZX}(\bar{x} - \bar{x}')]\left[\frac{(\alpha_i \bar{x}' + \beta_i)}{(\alpha_i \bar{x} + \beta_i)}\right].$ | $\hat{\mu}_{GP}^{3} = [\hat{\hat{z}} - b_{ZX}(\bar{x} - \bar{x}')]\left[\frac{(\alpha_i \bar{x} + \beta_i)}{(\alpha_i \bar{x}' + \beta_i)}\right].$ |

## 3. Analysis through simulation

The efficiency of the proposed RRT mean estimator using two-phase simple random sampling has been assessed through simulation. The values of $k$ and $v$ are chosen as one and $\pi$ will take its optimum value. The various parameters of auxiliary variable can be used for $\alpha, \beta$ such as coefficient of variation, kurtosis, median, quartiles etc. It is observed that *MSE* is not much affected with different parameters in any significant way with different choices of $\alpha, \beta$. For this reason, the only choice to be considered here is $\alpha = 1$ and $\beta = 0$. The additive scrambling variable $S$ is distributed as normal with mean = 0 and variance $(\sigma_s^2 = 0.2 * \sigma_x^2)$ and the multiplicative $T$ is also following normal distribution with $\mu_T = 1$, and varying $\sigma_T^2 = 0, 0.5, 1$. Following Zhang et al. [19], a population of size $N$=5000 is generated with the following parameters for $Y$ and $X$.

Populations

$$\mu = \begin{bmatrix} 10 \\ 6 \end{bmatrix}, \qquad \Sigma = \begin{bmatrix} 16 & 9.051 \\ 9.051 & 8 \end{bmatrix}, \rho_{XY} = 0.8$$

We consider sample sizes for first and second phase as $n' = 750$ and $n$ = 150 respectively and the response rate on the survey is assumed to be 60% in the second phase. At the second phase, out of 150 sample units, only $n_1$ = 90 respondents give response and $n_2$ = 60 do not give response. We take another sample from non-respondent group with $n_s = n_2 / h$ by taking $h$=2. The empirical *MSE* of $\hat{\mu}_{G_i}$ is obtained as,

$$MSE^*(\hat{\mu}_{G_i}) = \frac{1}{10000}\sum_{i=1}^{10000}(\hat{\mu}_{G_i} - \mu_y)^2, \qquad i = 1,2,3,4$$

where, $\hat{\mu}_{G_i} = \hat{\mu}_{G_1}, \hat{\mu}_{G_2}, \hat{\mu}_{G_3}, \hat{\mu}_{G_4}$ and $\mu_y$ is the population mean of $Y$.

The PREs of the estimator are calculated as,

$$PRE = \frac{MSE^*(\hat{\mu}_{ord})}{MSE^*(\hat{\mu}_{G_i})} * 100.$$

A unified measure introduced by Gupta et al. [22], which addresses both privacy and efficiency, is calculated using the following formula:

$$\delta = \frac{MSE^*(\hat{\mu}_{G_i})}{\Delta_{DP}}$$

For the biased estimators, the word *MSE* is used instead of $Var(.)$. The results of MSEs, PREs and Unified measure for all the cases are given in Tables (2, 3, 4, 5). The theoretical MSEs are given in bold.

**Table 2**

**Case-I: Theoretical/Empirical MSEs, PREs along unified measure when $S_s^2 = 0.2 * S_x^2$**

| W | Estimators | $S_T^2$ | MSE | | PRE | | $\delta$ |
|---|---|---|---|---|---|---|---|
| | | | Without ME | With ME | Without ME | With ME | |
| 0.5 | $\hat{\mu}_{ord1}$ | 0 | **0.1087** 0.1089 | **0.1179** 0.1187 | 100.0000 | 100.0000 | 0.0723 |
| | | 0.5 | **0.2861** 0.2993 | **0.2945** 0.3142 | 100.0000 | 100.0000 | 0.0050 |
| | | 1 | **0.4764** 0.4979 | **0.4869** 0.5062 | 100.0000 | 100.0000 | 0.0041 |
| | $\hat{\mu}_{rt1}$ | 0 | **0.0669** 0.0738 | **0.0916** 0.0865 | 162.4813 | 128.7118 | 0.0562 |
| | | 0.5 | **0.2466** 0.2709 | **0.2720** 0.2782 | 116.0178 | 108.2720 | 0.0045 |
| | | 1 | **0.4319** 0.4604 | **0.4633** 0.4749 | 110.3033 | 105.0939 | 0.0039 |

*Contd...*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | $\hat{\mu}_{G_1}$ | 0 | **0.0547**<br>0.0601 | **0.0681**<br>0.0661 | 198.7203 | 173.1277 | 0.0417 |
| | | 0.5 | **0.2341**<br>0.2519 | **0.2484**<br>0.2610 | 122.2127 | 118.5587 | 0.0041 |
| | | 1 | **0.4203**<br>0.4429 | **0.4406**<br>0.4553 | 113.3476 | 110.5084 | 0.0037 |
| 0.8 | $\hat{\mu}_{ord1}$ | 0 | **0.1118**<br>0.1104 | **0.1211**<br>0.1214 | 100.0000 | 100.0000 | 0.0743 |
| | | 0.5 | **0.3969**<br>0.4135 | **0.4064**<br>0.4258 | 100.0000 | 100.0000 | 0.0068 |
| | | 1 | **0.6993**<br>0.7251 | **0.7142**<br>0.7077 | 100.0000 | 100.0000 | 0.0060 |
| | $\hat{\mu}_{rt1}$ | 0 | **0.0699**<br>0.0751 | **0.0947**<br>0.0916 | 159.9427 | 127.8775 | 0.0580 |
| | | 0.5 | **0.3574**<br>0.4030 | **0.3838**<br>0.4088 | 111.0520 | 105.8884 | 0.0064 |
| | | 1 | **0.6548**<br>0.6882 | **0.6907**<br>0.6892 | 106.7960 | 103.4023 | 0.0059 |
| | $\hat{\mu}_{G_1}$ | 0 | **0.0577**<br>0.0611 | **0.0712**<br>0.0706 | 193.7608 | 170.0843 | 0.0437 |
| | | 0.5 | **0.3449**<br>0.3784 | **0.3603**<br>0.3834 | 115.0768 | 112.7949 | 0.0060 |
| | | 1 | **0.6433**<br>0.6699 | **0.6680**<br>0.6636 | 108.7051 | 106.9162 | 0.0057 |
| 1 (Full RRT) | $\hat{\mu}_{ord1}$ | 0 | **0.1138**<br>0.1121 | **0.1232**<br>0.1233 | 100.0000 | 100.0000 | 0.0756 |
| | | 0.5 | **0.4708**<br>0.4983 | **0.4810**<br>0.5181 | 100.0000 | 100.0000 | 0.0080 |
| | | 1 | **0.8480**<br>0.8834 | **0.8658**<br>0.8491 | 100.0000 | 100.0000 | 0.0074 |
| | $\hat{\mu}_{rt1}$ | 0 | **0.0720**<br>0.0826 | **0.0968**<br>0.0927 | 158.0555 | 127.2727 | 0.0594 |
| | | 0.5 | **0.4313**<br>0.4860 | **0.4584**<br>0.4914 | 109.1583 | 104.9302 | 0.0077 |
| | | 1 | **0.8034**<br>0.8432 | **0.8282**<br>0.8422 | 105.5514 | 104.5310 | 0.0070 |
| | $\hat{\mu}_{G_1}$ | 0 | **0.0597**<br>0.0646 | **0.0733**<br>0.0717 | 190.6197 | 168.0764 | 0.0450 |
| | | 0.5 | **0.4187**<br>0.4633 | **0.4349**<br>0.4716 | 112.4433 | 110.6001 | 0.0073 |
| | | 1 | **0.7919**<br>0.8250 | **0.7976**<br>0.8196 | 107.0842 | 108.5506 | 0.0068 |

**Table 3**

**Case-2: Theoretical/Empirical MSEs, PREs along unified measure when response rate is 60%, $h$=2 and $S_s^2 = 0.2 * S_x^2$**

| $W$ | Estimators | | MSE | | PRE | | $\delta$ |
|---|---|---|---|---|---|---|---|
| | | $S_T^2$ | Without ME | With ME | Without ME | With ME | |
| 0.5 | $\hat{\mu}_{ord2}$ | 0 | **0.1463**<br>0.1484 | **0.1584**<br>0.1586 | 100.0000 | 100.0000 | 0.0988 |
| | | 0.5 | **0.2992**<br>0.2964 | **0.3142**<br>0.3052 | 100.0000 | 100.0000 | 0.0051 |
| | | 1 | **0.4365**<br>0.4529 | **0.4559**<br>0.4719 | 100.0000 | 100.0000 | 0.0039 |
| | $\hat{\mu}_{rt2}$ | 0 | **0.1043**<br>0.1596 | **0.1346**<br>0.1854 | 140.2684 | 117.6820 | 0.0840 |
| | | 0.5 | **0.2578**<br>0.2980 | **0.2895**<br>0.3496 | 116.0589 | 108.5319 | 0.0047 |
| | | 1 | **0.3985**<br>0.4794 | **0.4329**<br>0.5086 | 109.5357 | 105.3130 | 0.0037 |
| | $\hat{\mu}_{G_2}$ | 0 | **0.0925**<br>0.1280 | **0.1121**<br>0.1480 | 158.1621 | 141.3024 | 0.0710 |
| | | 0.5 | **0.2451**<br>0.2710 | **0.2654**<br>0.2920 | 122.0726 | 118.3873 | 0.0043 |
| | | 1 | **0.3857**<br>0.4340 | **0.4089**<br>0.4536 | 113.1708 | 111.4942 | 0.0035 |
| 0.8 | $\hat{\mu}_{ord2}$ | 0 | **0.1488**<br>0.1520 | **0.1608**<br>0.1610 | 100.0000 | 100.0000 | 0.1003 |
| | | 0.5 | **0.3915**<br>0.3842 | **0.4065**<br>0.4123 | 100.0000 | 100.0000 | 0.0066 |
| | | 1 | **0.6116**<br>0.6447 | **0.6280**<br>0.6677 | 100.0000 | 100.0000 | 0.0054 |
| | $\hat{\mu}_{rt2}$ | 0 | **0.1068**<br>0.1655 | **0.1370**<br>0.2020 | 139.3258 | 117.3722 | 0.0855 |
| | | 0.5 | **0.3501**<br>0.3962 | **0.3819**<br>0.4458 | 111.8251 | 106.4414 | 0.0062 |
| | | 1 | **0.5737**<br>0.6700 | **0.6051**<br>0.6991 | 106.6062 | 103.7845 | 0.0052 |

*Contd...*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | $\hat{\mu}_{G_2}$ | 0 | **0.0950** 0.1324 | **0.1146** 0.1494 | 156.6315 | 140.3141 | 0.0715 |
| | | 0.5 | **0.3374** 0.3662 | **0.3578** 0.3916 | 116.0343 | 113.6109 | 0.0058 |
| | | 1 | **0.5609** 0.6226 | **0.5810** 0.6472 | 109.0390 | 108.0895 | 0.0050 |
| 1 (Full RRT) | $\hat{\mu}_{ord2}$ | 0 | **0.1505** 0.1563 | **0.1621** 0.1625 | 100.0000 | 100.0000 | 0.1011 |
| | | 0.5 | **0.4526** 0.4405 | **0.4676** 0.4713 | 100.0000 | 100.0000 | 0.0076 |
| | | 1 | **0.7270** 0.7592 | **0.7396** 0.7919 | 100.0000 | 100.0000 | 0.0064 |
| | $\hat{\mu}_{rt2}$ | 0 | **0.1085** 0.1665 | **0.1387** 0.2055 | 138.7096 | 116.8709 | 0.0865 |
| | | 0.5 | **0.4112** 0.4449 | **0.4429** 0.5281 | 110.0680 | 105.5769 | 0.0072 |
| | | 1 | **0.6891** 0.7900 | **0.7167** 0.8317 | 105.4999 | 103.1952 | 0.0062 |
| | $\hat{\mu}_{G_2}$ | 0 | **0.0967** 0.1352 | **0.1162** 0.1508 | 155.6360 | 139.5008 | 0.0725 |
| | | 0.5 | **0.3985** 0.4173 | **0.4189** 0.4628 | 113.5760 | 111.6257 | 0.0068 |
| | | 1 | **0.6763** 0.7436 | **0.6926** 0.7774 | 107.4967 | 106.7860 | 0.0060 |

**Table 4**

**Case-3: Theoretical/Empirical MSEs, PREs along unified measure when response rate is 60%, *h*=2 $S_s^2 = 0.2 * S_x^2$**

| W | Estimators | $S_T^2$ | MSE Without ME | MSE With ME | PRE Without ME | PRE With ME | $\delta$ |
|---|---|---|---|---|---|---|---|
| 0.5 | $\hat{\mu}_{ord3}$ | 0 | **0.1463** 0.1484 | **0.1588** 0.1611 | 100.0000 | 100.0000 | 0.0942 |
| | | 0.5 | **0.2992** 0.2964 | **0.3223** 0.3205 | 100.0000 | 100.0000 | 0.0054 |
| | | 1 | **0.4365** 0.4529 | **0.4998** 0.4891 | 100.0000 | 100.0000 | 0.0044 |

*Contd...*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | $\hat{\mu}_{rt3}$ | 0 | **0.1043**<br>0.1596 | **0.1166**<br>0.1659 | 140.2684 | 136.1921 | 0.0692 |
| | | 0.5 | **0.2578**<br>0.2980 | **0.2787**<br>0.3310 | 116.0589 | 115.6440 | 0.0046 |
| | | 1 | **0.3985**<br>0.4794 | **0.4580**<br>0.5038 | 109.5357 | 109.1266 | 0.0040 |
| | $\hat{\mu}_{G_3}$ | 0 | **0.0925**<br>0.1280 | **0.1043**<br>0.1360 | 158.1621 | 152.2531 | 0.0619 |
| | | 0.5 | **0.2451**<br>0.2710 | **0.2692**<br>0.3013 | 122.0726 | 119.7251 | 0.0045 |
| | | 1 | **0.3857**<br>0.4340 | **0.4465**<br>0.4745 | 113.1708 | 111.9373 | 0.0039 |
| 0.8 | $\hat{\mu}_{ord3}$ | 0 | **0.1488**<br>0.1520 | **0.1616**<br>0.1626 | 100.0000 | 100.0000 | 0.0959 |
| | | 0.5 | **0.3915**<br>0.3842 | **0.4243**<br>0.4283 | 100.0000 | 100.0000 | 0.0070 |
| | | 1 | **0.6116**<br>0.6447 | **0.7172**<br>0.6898 | 100.0000 | 100.0000 | 0.0063 |
| | $\hat{\mu}_{rt3}$ | 0 | **0.1068**<br>0.1655 | **0.1194**<br>0.1676 | 139.3258 | 135.3433 | 0.0708 |
| | | 0.5 | **0.3501**<br>0.3962 | **0.3808**<br>0.4375 | 111.8252 | 111.4233 | 0.0063 |
| | | 1 | **0.5737**<br>0.6700 | **0.6754**<br>0.7173 | 106.6062 | 106.1889 | 0.0059 |
| | $\hat{\mu}_{G_3}$ | 0 | **0.0950**<br>0.1324 | **0.1071**<br>0.1398 | 156.6316 | 150.8870 | 0.0635 |
| | | 0.5 | **0.3374**<br>0.3662 | **0.3712**<br>0.4089 | 116.0343 | 114.3050 | 0.0062 |
| | | 1 | **0.5609**<br>0.6226 | **0.6639**<br>0.6680 | 109.0390 | 108.0283 | 0.0058 |
| | $\hat{\mu}_{ord3}$ | 0 | **0.1505**<br>0.1563 | **0.1635**<br>0.1669 | 100.0000 | 100.0000 | 0.0970 |
| | | 0.5 | **0.4526**<br>0.4405 | **0.4929**<br>0.4814 | 100.0000 | 100.0000 | 0.0082 |
| | | 1 | **0.7270**<br>0.7592 | **0.8665**<br>0.8036 | 100.0000 | 100.0000 | 0.0076 |

*Contd...*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | $\hat{\mu}_{rt3}$ | 0 | **0.1085**<br>0.1665 | **0.1213**<br>0.1712 | 138.7097 | 134.7898 | 0.0720 |
| | | 0.5 | **0.4112**<br>0.4449 | **0.4494**<br>0.4868 | 110.0681 | 109.6795 | 0.0075 |
| | | 1 | **0.6891**<br>0.7900 | **0.8247**<br>0.8159 | 105.4999 | 105.0685 | 0.0072 |
| | $\hat{\mu}_{G_3}$ | 0 | **0.0967**<br>0.1352 | **0.1090**<br>0.1419 | 155.6360 | 150.0000 | 0.0647 |
| | | 0.5 | **0.3985**<br>0.4173 | **0.4398**<br>0.4586 | 113.5760 | 112.0737 | 0.0073 |
| | | 1 | **0.6763**<br>0.7436 | **0.8132**<br>0.7890 | 107.4967 | 106.5543 | 0.0071 |

**Table 5**

**Case-4: Theoretical/Empirical MSEs, PREs along unified measure when response rate is 60%, *h* =2 and $S_s^2 = 0.2 * S_x^2$**

| *W* | Estimators | $S_T^2$ | MSE | | PRE | | $\delta$ |
|---|---|---|---|---|---|---|---|
| | | | Without ME | With ME | Without ME | With ME | |
| 0.5 | $\hat{\mu}_{ord4}$ | 0 | **0.1343**<br>0.1513 | **0.1575**<br>0.1602 | 100.0000 | 100.0000 | 0.1005 |
| | | 0.5 | **0.2911**<br>0.2942 | **0.3082**<br>0.3202 | 100.0000 | 100.0000 | 0.0055 |
| | | 1 | **0.4729**<br>0.4750 | **0.4806**<br>0.4694 | 100.0000 | 100.0000 | 0.0040 |
| | $\hat{\mu}_{rt4}$ | 0 | **0.0851**<br>0.1219 | **0.1216**<br>0.1404 | 157.8143 | 129.5230 | 0.0776 |
| | | 0.5 | **0.2281**<br>0.2647 | **0.2675**<br>0.2906 | 127.6194 | 115.2149 | 0.0048 |
| | | 1 | **0.4075**<br>0.4514 | **0.4396**<br>0.4488 | 116.0490 | 109.3267 | 0.0036 |
| | $\hat{\mu}_{G_4}$ | 0 | **0.0657**<br>0.0820 | **0.0889**<br>0.0872 | 204.4140 | 177.1653 | 0.0567 |
| | | 0.5 | **0.2096**<br>0.2110 | **0.2327**<br>0.2359 | 138.8836 | 132.4452 | 0.0042 |
| | | 1 | **0.4021**<br>0.3905 | **0.4091**<br>0.3938 | 117.6075 | 117.4774 | 0.0034 |

*Contd...*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 0.8 | $\hat{\mu}_{ord4}$ | 0 | **0.1465**<br>0.1552 | **0.1600**<br>0.1628 | 100.0000 | 100.0000 | 0.1021 |
| | | 0.5 | **0.3675**<br>0.3812 | **0.3949**<br>0.4121 | 100.0000 | 100.0000 | 0.0071 |
| | | 1 | **0.6360**<br>0.6780 | **0.6737**<br>0.6870 | 100.0000 | 100.0000 | 0.0056 |
| | $\hat{\mu}_{rt4}$ | 0 | **0.0977**<br>0.1315 | **0.1242**<br>0.1411 | 149.9488 | 128.8244 | 0.0793 |
| | | 0.5 | **0.3134**<br>0.3428 | **0.3541**<br>0.3787 | 117.2622 | 111.5221 | 0.0064 |
| | | 1 | **0.6108**<br>0.6135 | **0.6526**<br>0.6414 | 104.1257 | 103.2332 | 0.0054 |
| | $\hat{\mu}_{G_4}$ | 0 | **0.0783**<br>0.0843 | **0.0915**<br>0.0885 | 187.1009 | 174.8634 | 0.0584 |
| | | 0.5 | **0.2949**<br>0.3003 | **0.3193**<br>0.3271 | 124.6185 | 123.6768 | 0.0057 |
| | | 1 | **0.5225**<br>0.5425 | **0.6002**<br>0.5945 | 121.7225 | 112.2460 | 0.0050 |
| 1<br>(Full RRT) | $\hat{\mu}_{ord4}$ | 0 | **0.1483**<br>0.1562 | **0.1645**<br>0.1674 | 100.0000 | 100.0000 | 0.1050 |
| | | 0.5 | **0.4324**<br>0.4560 | **0.4517**<br>0.4704 | 100.0000 | 100.0000 | 0.0081 |
| | | 1 | **0.7731**<br>0.8150 | **0.8023**<br>0.7881 | 100.0000 | 100.0000 | 0.0067 |
| | $\hat{\mu}_{rt4}$ | 0 | **0.0994**<br>0.1316 | **0.1259**<br>0.1428 | 149.1951 | 130.6592 | 0.0804 |
| | | 0.5 | **0.3693**<br>0.4228 | **0.4109**<br>0.4390 | 117.0864 | 109.9294 | 0.0074 |
| | | 1 | **0.7478**<br>0.7621 | **0.7813**<br>0.7497 | 103.3832 | 102.6878 | 0.0065 |
| | $\hat{\mu}_{G_4}$ | 0 | **0.0800**<br>0.0851 | **0.0953**<br>0.0914 | 185.3750 | 172.6128 | 0.0608 |
| | | 0.5 | **0.3477**<br>0.3508 | **0.3762**<br>0.3844 | 124.3600 | 120.0691 | 0.0068 |
| | | 1 | **0.7288**<br>0.7120 | **0.7299**<br>0.7003 | 109.9191 | 105.9186 | 0.0060 |

The MSEs and PREs, of the RRT mean estimators are presented in Tables [2, 3, 4, 5] when the variance of $T$ is taken as 0, 0.5, 1 and the response rate (RR) is 60% in the second phase with $S_s^2 = 0.2 * S_x^2$. When comparing all these tables, our proposed generalized RRT mean estimator demonstrates superior performance over the ordinary and ratio estimators, particularly in terms of high efficiency, whether there is error in measurement and low response rate. In all the tables, when variance of $T$ increases, the mean estimation becomes less efficient. As the $Var(T)$ increases from 0 to 1, the MSE increases while $\delta$ goes down. For example, when ($W$) is set to (0.8), the MSE of the propose estimator increases from (0.0915) to (0.6002), but the value of $\delta$ falls from (0.0584) to (0.0050) in Table 5 and this pattern is observed in all the cases. Another observation is, when non-response is present, mean estimators under the ORRT model are better than full RRT model when evaluation is done through the unified measure. We can also

**Table 6**

**Case-4: Theoretical/Empirical MSEs, PREs of the RRT Mean Estimator when $S_T^2 = 0.5$, $W = 0.5$ and $S_s^2 = 0.2 * S_x^2$**

| Estimator | $h$ | MSE with ME | | | PRE with ME | | |
|---|---|---|---|---|---|---|---|
| | | 40% | 60% | 75% | 40% | 60% | 75% |
| $\hat{\mu}_{ord4}$ | 2 | **0.4143**<br>0.4213 | **0.3082**<br>0.3202 | **0.2491**<br>0.2501 | 100.0000 | 100.0000 | 100.0000 |
| | 3 | **0.6140**<br>0.6326 | **0.4437**<br>0.4534 | **0.3233**<br>0.3272 | 100.0000 | 100.0000 | 100.0000 |
| | 4 | **0.7948**<br>0.8168 | **0.5766**<br>0.5908 | **0.3959**<br>0.4406 | 100.0000 | 100.0000 | 100.0000 |
| $\hat{\mu}_{rt4}$ | 2 | **0.3728**<br>0.4146 | **0.2675**<br>0.2906 | **0.2124**<br>0.2329 | 111.1320 | 115.2149 | 117.2787 |
| | 3 | **0.5526**<br>0.5563 | **0.3936**<br>0.4131 | **0.2802**<br>0.3055 | 111.1111 | 112.7286 | 115.3818 |
| | 4 | **0.7185**<br>0.7672 | **0.5131**<br>0.5599 | **0.3487**<br>0.3825 | 110.6193 | 112.3757 | 113.5360 |
| $\hat{\mu}_{G_4}$ | 2 | **0.3328**<br>0.3437 | **0.2327**<br>0.2359 | **0.1823**<br>0.1835 | 124.4892 | 132.4452 | 136.6429 |
| | 3 | **0.5031**<br>0.4791 | **0.3498**<br>0.3542 | **0.2469**<br>0.2477 | 122.0433 | 126.8440 | 130.9437 |
| | 4 | **0.6529**<br>0.6640 | **0.4592**<br>0.4805 | **0.3050**<br>0.3304 | 121.7338 | 125.5662 | 129.8033 |

see that by increasing the value of (W) from (0.5) to (1), the *MSE* also increases for all the cases. It shows that when respondents perceive the survey question as non-sensitive, the ORRT model achieves greater efficiency.

Table 6 indicates that as the response rate increases, the mean squared error (MSE) decreases while the percentage of relative efficiency (PRE) increases for each estimator. Furthermore, when the value of h increases from 2 to 4, the mean square error also increases, but PRE decreases for each estimator. This table also demonstrates that our proposed generalized RRT mean estimator consistently outperforms both the ordinary RRT and RRT ratio estimators across all situations.

## 4. Concluding Remarks

There are two major take-aways from the simulations results. We can see that the performance in terms of MSE and unified measure worsens with an increased sensitivity level $W$. This also shows that the Optional RRT models are better choice since the worst case is when $W = 1$, which corresponds to the non-optional model. We also note that we should not be afraid to use the extra multiplicative noise because it improves the privacy more than it hurts the efficiency, as shown by the unified measure.

## Acknowledgement

The authors wish to express their sincere gratitude to the reviewers and the editor for their invaluable suggestions and constructive feedback.

## References

[1] J. Neyman, "Contribution to the theory of sampling human populations," *Journal of the American Statistical Association*, vol. 33, no. 201, pp. 101–116 (1938).

[2] S. L. Warner, "Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias," *Journal of the American Statistical Association*, vol. 60, no. 309, pp. 63–69 (1965).

[3] B. H. Eichhorn and L. S. Hayre, "Scrambled randomized response methods for obtaining sensitive quantitative data," *Journal of Statistical Planning and Inference*, vol. 7, no. 4, pp. 307–316 (1983).

[4] S. Gupta, B. Gupta, and S. Singh, "Estimation of sensitivity level of personal interview survey questions," *Journal of Statistical Planning and Inference*, vol. 100, no. 2, pp. 239–247 (2002).

[5] S. Gupta and J. Shabbir, "Sensitivity Estimation for Personal Interview Survey Questions," *Statistica*, vol. 64, no. 4, pp. 643–653 (2004).

[6] R. Sousa, J. Shabbir, P. C. Real, and S. Gupta, "Ratio Estimation of the Mean of a Sensitive Variable in the Presence of Auxiliary Information," *Journal of Statistical Theory and Practice*, vol. 4, no. 3, pp. 495–507 (2010).

[7] S. Gupta, J. Shabbir, R. Sousa, and P. Corte-Real, "Estimation of the Mean of a Sensitive Variable in the Presence of Auxiliary Information," *Communications in Statistics-Theory and Methods*, vol. 41, no. 13-14, pp. 2394–2404 (2012).

[8] S. Gupta, J. Shabbir, R. Sousa, and P. Corte-Real, "Improved Exponential Type Estimators of the Mean of a Sensitive Variable in the Presence of Non-sensitive Auxiliary Information," *Communications in Statistics-Simulation and Computation*, vol. 45, no. 9, pp. 3317–3328 (2016).

[9] M. H. Hansen and W. N. Hurwitz, "The problem of non-response in sample surveys," *Journal of the American Statistical Association*, vol. 41, no. 236, pp. 517–529 (1946).

[10] M. Kumar, R. Singh, A. K. Singh, and F. Smarandache, "Some ratio type estimators under measurement errors," *World Applied Sciences Journal*, vol. 14, no. 2, pp. 272–276 (2011).

[11] S. Khalil, M. Noor-ul-Amin, and M. Hanif, "Estimation of population mean for a sensitive variable in the presence of measurement error," *Journal of Statistics and Management Systems,* vol. 21, no. 1, pp. 81–91 (2018).

[12] N. Singh, G. K. Vishwakarma, and J. M. Kim, "Computing the effect of measurement error on efficient variant of the product and ratio estimator using auxiliary variable," *Communications in Statistics-Simulation and Computation*, vol. 51, no. 2, pp. 604–625 (2019).

[13] S. R. Singh and P. Sharma, "Method of estimation in the presence of non-response and measurement errors simultaneously," *Journal of Modern Applied Statistical Methods*, vol. 14, no. 1, pp. 107–121 (2015).

[14] N. Singh and G. K. Vishwakarma, "A generalized class of estimator of population mean with the combined effect of measurement errors and non-response in sample survey," *Revista Investigacion Operacional*, vol. 40, no. 2, pp. 275–285 (2019).

[15] A. Audu, R. Singh, S. Khare, and N. S. Dauran, "Almost unbiased estimators for population mean in the presence of non-response and measurement error," *Journal of Statistics and Management Systems*, vol. 24, no. 3, pp. 573–589 (2021).

[16] S. Khalil, Q. Zhang, and S. Gupta, "Mean estimation of sensitive variables under measurement errors using optional RRT models," *Communications in Statistics-Simulation and Computation*, vol. 50, no. 5, pp. 1417–1426 (2021).

[17] S. Khalil, M. Noor-ul-Amin, and M. Hanif, "Generalized estimator of population mean by using conventional and non-conventional measures in the presence of measurement errors," *Communications in Statistics-Simulation and Computation*, vol. 48, no. 2, pp. 516–529 (2019).

[18] Q. Zhang, S. Gupta, G. Kalucha, and S. Khalil, "Ratio estimation of the mean under RRT models," *Journal of Statistics and Management Systems*, vol. 22, no. 1, pp. 97–113 (2019).

[19] Q. Zhang, S. Khalil, and S. Gupta, "Mean estimation of sensitive variables under non-response and measurement errors using optional RRT models," *Journal of Statistical Theory and Practice*, vol. 15, no. 3, pp. 1–15 (2021).

[20] U. Shuja, S. Khalil, and S. Gupta, "Mean estimation using scrambled responses with the impact of measurement errors in two-phase sampling," *Journal of Interdisciplinary Mathematics*, vol. 25, no. 6, pp. 1675–1695 (2022).

[21] G. Diana and P. F. Perri, "A class of estimators for quantitative sensitive data," *Stat Papers*, vol. 52, pp. 633–650 (2011).

[22] S. Gupta, S. Mehta, J. Shabbir, and S. Khalil, "A unified measure of respondent privacy and model efficiency in quantitative RRT models," *Journal of Statistical Theory and Practice*, vol. 12, no. 3, pp. 506–511 (2018).

[23] Z. Yan, J. Wang, and J. Lai, "An efficiency and protection degree-based comparison among the quantitative randomized response strategies," *Communications in Statistics: Theory and Methods*, vol. 38, no. 3, pp. 400–408 (2009).

[24] G. Diana, S. Riaz, and J. Shabbir, "Hansen and Hurwitz estimator with scrambled response on the second call," *Journal of Applied Statistics*, vol. 41, no. 3, pp. 596–611 (2014).

*Received December, 2022*