Simultaneous Estimation of Mean of Sensitive Variable and Sensitivity level by using Generalized Optional Scrambling

Muhammad Noor-ul-Amin
COMSATS Institute of Information and Technology, Lahore, Pakistan
nooramin.stats@gmail.com

Nadia Mushtaq
National College of Business Administration and Economics, Lahore, Pakistan

Muhammad Hanif
National College of Business Administration and Economics, Lahore, Pakistan

Abstract

Randomized response technique introduces anonymity into subjects' responses hence encouraging more honest responses. In quantitative randomized response model, additive and multiplicative models have been developed to reduce bias. However, additive and multiplicative models may not be sufficient to reduce this bias so the generalized optional scrambling randomized response model proposed is able to reduce these problems. We also improved mean estimation utilizing information from a non-sensitive auxiliary variable by way of ratio and regression estimators in the proposed model.

1. Introduction

Randomized response technique (RRT), pioneered by (Warner, 1965) which helps interviewers extract reliable data corresponding to sensitive questions while maintaining respondent anonymity. The quantitative optional randomized response model was introduced by Gupta et al. (2002). In this model, the respondents decide themselves whether they want to tell the truth (or scramble their true response) depending upon whether the question being asked is perceived by them as non-sensitive (or sensitive). The sensitivity level of the question is the proportion of respondents who consider the question sensitive and is usually denoted by W. Gupta et al. (2002) and the Gupta and Shabbir (2004) models were based on multiplicative scrambling whereas the Gupta et al. (2010) model is based on additive scrambling which works better than the multiplicative scrambling, as shown in Gupta et al. (2012a). Mushtaq et al. (2016) proposed estimation of a population mean of a sensitive variable in stratified two-phase sampling. Mushtaq et al. (2017) presented a family of estimators of a sensitive variable using auxiliary information in stratified random sampling.

As we know that in a survey, different questions may have different sensitivity levels and it may be useful to quantify this sensitivity. In this paper we consider generalized optional scrambling that allows simultaneous estimation of mean of sensitive variable and the sensitivity level of a sensitive question. In this model we draw two subsamples and obtain two responses from each respondent using two different generalized scrambling variables. A theoretical comparison and simulation study is conducted to analyze the performance of the suggested estimators for proposed model.

2. Proposed Methods

2.1 Mean Estimator and Sensitivity Level

Let the population size N and sample size n and sample size split into two subsamples of sizes n_1 and $n_2 (n_1 + n_2 = n)$. Let Y be the sensitive study variable, X be a non-sensitive auxiliary variable that is correlated with Y and S_1, S_2 be scrambling variables. Let the mean and variance of Y be μ_Y and σ_Y^2, respectively, for the auxiliary variable X be μ_X and σ_X^2 for $S_i (i = 1,2)$ be θ_i and σ_{Si}^2 respectively. In each subsample, we will observe X directly and the study variable Y will observe by using scramble response. In each subsample, respondents provide a scrambled response if they consider the question sensitive and a true response otherwise. Let W be the sensitivity level of the underlying sensitive question. So k_1 and k_2 are suitably chosen scalars.

According to the model, the reported response Z_1 and Z_2 are given by

$$Z_1 = \begin{cases} Y & \text{with probability } (1\text{-W}) \\ Y + k_1 S_1 & \text{with probability W} \end{cases}, \tag{1}$$

$$Z_2 = \begin{cases} Y & \text{with probability } (1\text{-W}) \\ Y - k_2 S_2 & \text{with probability W} \end{cases}, \tag{2}$$

where $0 \le k_1 \le 1$ and $0 \le k_2 \le 1$.

The mean and variance of Z_1 and Z_2 are given by

$$\mathrm{E}(Z_1) = \mu_Y + W k_1 \theta_1 \tag{3}$$

$$\mathrm{E}(Z_2) = \mu_Y - W k_2 \theta_2 \tag{4}$$

and

$$\sigma_{Z_i}^2 = \sigma_Y^2 + W(1-W) k_i^2 \theta_i^2 + W k_i^2 \sigma_{S_i}^2 . \tag{5}$$

The proposed mean and sensitivity estimators are given by respectively

$$\hat{\mu}_Y = \frac{\theta_2 k_2 \bar{Z}_1 + \theta_1 k_1 \bar{Z}_2}{\theta_2 k_2 + \theta_1 k_1}, \tag{6}$$

and

$$\hat{W} = \frac{\bar{Z}_1 - \bar{Z}_2}{\theta_2 k_2 + \theta_1 k_1}. \tag{7}$$

The variances of $\hat{\mu}_Y$ and $\hat{W}$ are given below:

$$V\left(\hat{\mu}_Y\right)=\frac{1}{\left(\theta_2 k_2+\theta_1 k_1\right)^2}\left(\theta_2^2 k_2^2\frac{\sigma_{Z1}^2}{n_1}+\theta_1^2 k_1^2\frac{\sigma_{Z2}^2}{n_2}\right), \tag{8}$$

$$V\left(\hat{W}\right)=\frac{1}{\left(\theta_2 k_2+\theta_1 k_1\right)^2}\left(\frac{\sigma_{Z1}^2}{n_1}+\frac{\sigma_{Z2}^2}{n_2}\right), \tag{9}$$

Theorem 1: $\hat{\mu}_Y$ is unbiased estimator of μ_Y

Proof: From (6) we have given as:

$$\mathrm{E}\left(\hat{\mu}_Y\right)=\mathrm{E}\left(\frac{\theta_2 k_2\bar{Z}_1+\theta_1 k_1\bar{Z}_2}{\theta_2 k_2+\theta_1 k_1}\right), \tag{10}$$

$$=\frac{1}{\left(\theta_2 k_2+\theta_1 k_1\right)}\left[\theta_2 k_2 E\left(\bar{Z}\right)_1+\theta_1 k_1 E\left(\bar{Z}_2\right)\right]$$

$$=\frac{1}{\left(\theta_2 k_2+\theta_1 k_1\right)}\left(\theta_2 k_2+\theta_1 k_1\right)\mu_Y$$

$$\mathrm{E}\left(\hat{\mu}_Y\right)=\mu_Y\,. \tag{11}$$

Theorem 2: $\hat{W}$ is unbiased estimator of W

Proof: From (7) we have given as:

$$E\left(\hat{W}\right)=E\left(\frac{\bar{Z}_1-\bar{Z}_2}{\theta_2 k_2+\theta_1 k_1}\right), \tag{12}$$

$$=\frac{1}{\left(\theta_2 k_2+\theta_1 k_1\right)}\left(E\left(\bar{Z}_1\right)-E\left(\bar{Z}_2\right)\right),$$

$$=\frac{1}{\left(\theta_2 k_2+\theta_1 k_1\right)}W\left(\theta_2 k_2+\theta_1 k_1\right),$$

$$E\left(\hat{W}\right)=W\,. \tag{13}$$

2.2 Sample Size Optimization for Model-II

We find optimum sub-sample sizes which help in minimizing variance which is helps to improve the efficiency of the model and estimation of the mean prevalence of the sensitive characteristic. Thus, taking both variances into account, one can try to find n_1 and n_2 that minimize $\left[Var\left(\hat{\mu}_Y\right)+Var\left(\hat{W}\right)\right]$. We do this by taking partial derivatives with respect to n_1 and n_2, respectively, setting the derivatives to zero, then solving for

n_1 and n_2 to find specific optimal sample sizes. The optimal sample sizes subject to $n_1 + n_2 = n$, are

$$n_1 = \frac{n\sigma_{Z_1}\sqrt{\theta_2^2+1}}{\sigma_{Z_1}\sqrt{\theta_2^2+1}+\sigma_{Z_2}\sqrt{\theta_1^2+1}}, \tag{14}$$

$$n_2 = \frac{n\sigma_{Z_2}\sqrt{\theta_1^2+1}}{\sigma_{Z_1}\sqrt{\theta_2^2+1}+\sigma_{Z_2}\sqrt{\theta_1^2+1}}. \tag{15}$$

2.3 Ratio Estimator following Proposed Model

We propose the following ratio estimator based on a two sample approach using two different generalized scrambling variables in optional randomized response technique for proposed model, is given as:

$$\hat{\mu}_{RP1} = \left(\frac{\theta_2 k_2 \bar{Z}_1 + \theta_1 k_1 \bar{Z}_2}{\theta_1 k_1 + \theta_2 k_2}\right)\left(\frac{\mu_X}{\bar{x}_1}+\frac{\mu_X}{\bar{x}_2}\right)\frac{1}{2}. \tag{16}$$

Theorem 3: The MSE of $\hat{\mu}_{RP1}$ is given by

$$MSE\left(\hat{\mu}_{RP1}\right) \cong \left(\frac{1-f_1}{n_1}\right)\left[\left(\frac{\theta_2 k_2}{\left(\theta_1 k_1+\theta_2 k_2\right)}\right)^2 \sigma_{Z_1}^2 + \frac{\mu_Y^2}{4}C_x^2 - \frac{\theta_2 k_2}{\left(\theta_1 k_1+\theta_2 k_2\right)}\mu_Y \rho_{yx}\sigma_y C_x\right]$$
$$+\left(\frac{1-f_2}{n_2}\right)\left[\left(\frac{\theta_1 k_1}{\left(\theta_1 k_1+\theta_2 k_2\right)}\right)^2 \sigma_{Z_2}^2 + \frac{\mu_Y^2}{4}C_x^2 - \frac{\theta_2 k_2}{\left(\theta_1 k_1+\theta_2 k_2\right)}\mu_Y \rho_{yx}\sigma_y C_x\right].$$

Proof:

$$e_{z_1} = \frac{\bar{z}_1-\mu_{Z_1}}{\mu_{Z_1}},\ e_{z_2} = \frac{\bar{z}_2-\mu_{Z_2}}{\mu_{Z_2}},\ e_{x_1} = \frac{\bar{x}_1-\mu_X}{\mu_X},\ e_{x_2} = \frac{\bar{x}_2-\mu_X}{\mu_X}. \tag{17}$$

Substituting for $\bar{z}_1, \bar{z}_2, \bar{x}_1$ and $\bar{x}_2$ in (16) and we have the following:

$$\hat{\mu}_{RP1} = \frac{1}{\left(\theta_1 k_1+\theta_2 k_2\right)}\left[\theta_2 k_2\left(\mu_{Z_1}+e_{z_1}\mu_{Z_1}\right)+\theta_1 k_1\left(\mu_{Z_2}+e_{z_2}\mu_{Z_2}\right)\right]\frac{1}{2}\left[\left(1+e_{x_1}\right)^{-1}+\left(1+e_{x_2}\right)^{-1}\right] \tag{18}$$

By solving (18), we have the following results:

$$\hat{\mu}_{RP1} = \left[\frac{1}{2}\mu_Y + \frac{\theta_2 k_2}{2\left(\theta_1 k_1+\theta_2 k_2\right)}\mu_{Z_1}e_{z_1} + \frac{\theta_1 k_1}{2\left(\theta_1 k_1+\theta_2 k_2\right)}\mu_{Z_2}e_{z_2}\right]\left[\left(1-e_{x1}+e_{x1}^2\right)+\left(1-e_{x2}+e_{x2}^2\right)\right]$$

$$\hat{\mu}_{RP1}-\mu_Y = \frac{\mu_Y}{2}\left(-e_{x1}+e_{x1}^2-e_{x2}+e_{x2}^2\right)+\frac{\theta_2 k_2}{\left(\theta_1 k_1+\theta_2 k_2\right)}\mu_{Z_1}\left(2e_{z_1}-e_{z_1}e_{x1}-e_{z_1}e_{x2}\right)$$
$$+\frac{\theta_1 k_1}{\left(\theta_1 k_1+\theta_2 k_2\right)}\mu_{Z_2}\left(2e_{z_2}-e_{z_2}e_{x1}-e_{z_2}e_{x2}\right) \tag{20}$$

Let us define the following terms:

$$E\left(e_{z_1}\right)=E\left(e_{z_2}\right)=E\left(e_{x_1}\right)=E\left(e_{x_2}\right)=0\ E\left(e_{z_i}e_{x_i}\right)=\rho_{z_i x_i}C_{z_i}C_{x_i}\qquad (i=1,2),$$

Independent sample so we have the following:

$$E\left(\bar{z}_1,\bar{x}_2\right)=0=E\left(\bar{z}_2,\bar{x}_1\right),C_{z_i}^2=\frac{\sigma_Y^2+W(1-W)k_i^2\theta_i^2+Wk_i^2\sigma_{Si}^2}{\left(\mu_Y\pm Wk_i\theta_i\right)^2},$$

$$\rho_{z_i x_i}=\frac{\sigma_{yx_i}}{\sigma_{Z_i}\sigma_{x_i}}=\frac{\rho_{YX}}{\sqrt{1+\frac{W(1-W)k_i^2\theta_i^2}{\sigma_Y^2}+\frac{Wk_i^2\sigma_{Si}^2}{\sigma_Y^2}}}.$$

By squaring and solving (20), given as:

$$\begin{aligned}MSE\left(\hat{\mu}_{RP1}\right)\cong&\left(\frac{1-f_1}{n_1}\right)\left[\left(\frac{\theta_2k_2}{\left(\theta_1k_1+\theta_2k_2\right)}\right)^2\sigma_{Z_1}^2+\frac{\mu_Y^2}{4}C_{x_1}^2-\frac{\theta_2k_2}{\left(\theta_1k_1+\theta_2k_2\right)}\mu_{Z_1}\mu_Y\rho_{z_1x_1}C_{z_1}C_{x_1}\right]\\&+\left(\frac{1-f_2}{n_2}\right)\left[\left(\frac{\theta_1k_1}{\left(\theta_1k_1+\theta_2k_2\right)}\right)^2\sigma_{Z_2}^2+\frac{\mu_Y^2}{4}C_{x_2}^2-\frac{\theta_2k_2}{\left(\theta_1k_1+\theta_2k_2\right)}\mu_{Z_2}\mu_Y\rho_{z_2x_2}C_{z_2}C_{x_2}\right]\end{aligned}\qquad(21)$$

And we noted that

$$C_{x_1}=C_{x_2}=C_x,\ \rho_{z_1x_1}=\rho_{z_1x},\ \rho_{z_2x_2}=\rho_{z_2x},$$

$$\rho_{zx}=\frac{\rho_{yx}\sigma_y}{\mu_Z C_Z}\ \text{and}\ \rho_{yx}\sigma_y=\mu_Z C_Z\rho_{zx}.$$

By using the above values in the (22), so we have:

$$\begin{aligned}MSE\left(\hat{\mu}_{RP1}\right)\cong&\left(\frac{1-f_1}{n_1}\right)\left[\left(\frac{\theta_2k_2}{\left(\theta_1k_1+\theta_2k_2\right)}\right)^2\sigma_{Z_1}^2+\frac{\mu_Y^2}{4}C_x^2-\frac{\theta_2k_2}{\left(\theta_1k_1+\theta_2k_2\right)}\mu_Y\rho_{yx}\sigma_yC_x\right]\\&+\left(\frac{1-f_2}{n_2}\right)\left[\left(\frac{\theta_1k_1}{\left(\theta_1k_1+\theta_2k_2\right)}\right)^2\sigma_{Z_2}^2+\frac{\mu_Y^2}{4}C_x^2-\frac{\theta_2k_2}{\left(\theta_1k_1+\theta_2k_2\right)}\mu_Y\rho_{yx}\sigma_yC_x\right].\end{aligned}\qquad(22)$$

2.4 Regression Estimator following Proposed Model

We propose the following regression estimator based on two sample approach using two different generalized scrambling variables in optional randomized response technique for proposed model given as:

$$\hat{\mu}_{\mathrm{Re}\,gP1}=\left(\frac{\theta_2k_2\bar{z}_1+\theta_1k_1\bar{z}_2}{k_1\theta_1+k_2\theta_2}\right)+\left\{\beta_{z_1x_1}\left(\mu_x-\bar{x}_1\right)+\beta_{z_2x_2}\left(\mu_x-\bar{x}_2\right)\right\}\left(\frac{1}{2}\right),\qquad(23)$$

where $\hat{\beta}_{z_i x_i}\,(i=1,2)$ are the sample regression coefficients between Z_i and X_i respectively and $\bar{z}_i$, $\bar{x}_i$ $(i=1,2)$ are the two sub-sample means.

Theorem 4: The MSE of $\hat{\mu}_{\mathrm{Re}\,gP1}$ is given by

$$MSE\left(\hat{\mu}_{\mathrm{Re}\,gP1}\right) \cong \left(\frac{1-f_1}{n_1}\right)\left[\left(\frac{\theta_2 k_2}{(\theta_1 k_1+\theta_2 k_2)}\right)^2 \sigma_{Z_1}^2 + \frac{\beta_{z_1x_1}^2 \mu_{x_1}^2}{4} C_{x_1}^2 - \frac{\theta_2 k_2}{(\theta_1 k_1+\theta_2 k_2)} \beta_{z_1x_1}\mu_{z_1}\mu_{x_1}\rho_{z_1x_1}C_{z_1}C_{x_1}\right]$$

$$+\left(\frac{1-f_2}{n_2}\right)\left[\left(\frac{\theta_1 k_1}{(\theta_1 k_1+\theta_2 k_2)}\right)^2 \sigma_{Z_2}^2 + \frac{\beta_{z_2x_2}^2 \mu_{x2}^2}{4} C_{x_2}^2 - \frac{\theta_1 k_1}{(\theta_1 k_1+\theta_2 k_2)} \beta_{z_2x_2}\mu_{z_2}\mu_{x_2}\rho_{z_2x_2}C_{z_2}C_{x_2}\right].$$

Proof:

Now by expanding the (23), we have:

$$\hat{\mu}_{\mathrm{Re}\,gP1} = \frac{1}{(\theta_1 k_1+\theta_2 k_2)}\left[\theta_2 k_2\left(\mu_{Z_1} + e_{z_1}\mu_{Z_1}\right) + \theta_1 k_1\left(\mu_{Z_2} + e_{z_2}\mu_{Z_2}\right)\right]$$
$$-\left[\beta_{z_1x_1}\mu_x e_{x_1} + \beta_{z_2x_2}\mu_x e_{x_2}\right]\left(\frac{1}{2}\right),$$

$$\hat{\mu}_{\mathrm{Re}\,gP1} = \frac{1}{(\theta_1 k_1+\theta_2 k_2)}\left[\theta_2 k_2\mu_{Z_1} + \theta_2 k_2 e_{z_1}\mu_{Z_1} + \theta_1 k_1\mu_{Z_2} + \theta_1 k_1 e_{z_2}\mu_{Z_2}\right]$$
$$-\left[\beta_{z_1x_1}\mu_x e_{x_1} + \beta_{z_2x_2}\mu_x e_{x_2}\right]\left(\frac{1}{2}\right) \tag{24}$$

By substituting the values of μ_{Z_1} and μ_{Z_2} in (24). And by solving we have:

$$\hat{\mu}_{\mathrm{Re}\,gP1} = \frac{1}{(\theta_1 k_1+\theta_2 k_2)}\left[\theta_2 k_2\left(\mu_Y + wk_1\theta_1\right) + \theta_2 k_2 e_{z_1}\mu_{Z_1} + \theta_1 k_1\left(\mu_Y - wk_2\theta_2\right) + \theta_1 k_1 e_{z_2}\mu_{Z_2}\right]$$
$$-\left(\frac{1}{2}\right)\left[\beta_{z_1x_1}\mu_x e_{x_1} + \beta_{z_2x_2}\mu_x e_{x_2}\right]$$

$$E\left(\hat{\mu}_{\mathrm{Re}\,gP1} - \hat{\mu}_Y\right)^2 = E\left[\frac{\theta_2 k_2}{(\theta_1 k_1+\theta_2 k_2)}\mu_{Z_1}e_{z_1} + \frac{\theta_1 k_1}{(\theta_1 k_1+\theta_2 k_2)}\mu_{Z_2}e_{z_2}\right.$$
$$\left.-\left(\frac{1}{2}\right)\mu_x\left[\beta_{z_1x_1}e_{x_1} + \beta_{z_2x_2}e_{x_2}\right]\right]^2 \tag{25}$$

$$MSE\left(\hat{\mu}_{\mathrm{Re}\,gP1}\right) \cong \left(\frac{1-f_1}{n_1}\right)\left[\left(\frac{\theta_2 k_2}{(\theta_1 k_1+\theta_2 k_2)}\right)^2 \sigma_{Z_1}^2 + \frac{\beta_{z_1x_1}^2 \mu_{x_1}^2}{4} C_{x_1}^2 - \frac{\theta_2 k_2}{(\theta_1 k_1+\theta_2 k_2)} \beta_{z_1x_1}\mu_{z_1}\mu_{x_1}\rho_{z_1x_1}C_{z_1}C_{x_1}\right] \tag{26}$$
$$+\left(\frac{1-f_2}{n_2}\right)\left[\left(\frac{\theta_1 k_1}{(\theta_1 k_1+\theta_2 k_2)}\right)^2 \sigma_{Z_2}^2 + \frac{\beta_{z_2x_2}^2 \mu_{x2}^2}{4} C_{x_2}^2 - \frac{\theta_1 k_1}{(\theta_1 k_1+\theta_2 k_2)} \beta_{z_2x_2}\mu_{z_2}\mu_{x_2}\rho_{z_2x_2}C_{z_2}C_{x_2}\right].$$

By solving (26), we have given as

$$MSE\left(\hat{\mu}_{\mathrm{Re}\,gP1}\right) \cong \frac{1}{\left(\theta_1 k_1 + \theta_2 k_2\right)^2}\left[\theta_2^2 k_2^2\left(\frac{1-f_1}{n_1}\right)\sigma_{Z_1}^2 + \theta_1^2 k_1^2\left(\frac{1-f_2}{n_2}\right)\sigma_{Z_2}^2\right] + \frac{\rho_{yx}^2\sigma_Y^2}{4}\alpha - \rho_{yx}^2\sigma_Y^2\beta. \qquad (27)$$

Where $\alpha = \left(\frac{1-f_1}{n_1}\right) + \left(\frac{1-f_2}{n_2}\right),$

$$\beta = \left(\frac{1-f_1}{n_1}\right)\left(\frac{\theta_2 k_2}{\left(\theta_1 k_1 + \theta_2 k_2\right)}\right) + \left(\frac{1-f_2}{n_2}\right)\left(\frac{\theta_1 k_1}{\left(\theta_1 k_1 + \theta_2 k_2\right)}\right).$$

3. Simulation Study

In this simulation study is conducted to analyze the performance of the suggested estimators for proposed model. The comparison has been made by taking proposed mean estimator and Gupta et al. (2010). And the ratio and regression estimators compared with proposed mean estimator. For numerical comparison, we consider the following populations given as:

Population 1

$$\mu = \begin{bmatrix} 6 & 4 \end{bmatrix},\ N = 5000,\ \Sigma = \begin{bmatrix} 9 & 4.8 \\ 4.8 & 4 \end{bmatrix}, \qquad \rho_{XY} = 0.8.$$

Population 2

$$\mu = \begin{bmatrix} 6 & 4 \end{bmatrix},\ N = 5000,\ \Sigma = \begin{bmatrix} 9 & 1.84 \\ 1.84 & 4 \end{bmatrix}, \qquad \rho_{XY} = 0.3.$$

In Tables 1 to 3, the empirical and theoretical MSE's of the estimators based on the first-order approximation. And following expression is use to obtain percent relative efficiency (PRE) of different estimators with respect to $\hat{\mu}_{YG}$:

$$PRE\left(\hat{\mu}_{YP1}\right) = \frac{MSE\left(\hat{\mu}_{YG}\right)}{MSE\left(\hat{\mu}_{YP1}\right)} \times 100.$$

And the percent relative efficiency (PRE) for ratio and regression estimators for proposed model is given as:

$$PRE\left(\hat{\mu}_{RP1}\right) = \frac{MSE\left(\hat{\mu}_{YP1}\right)}{MSE\left(\hat{\mu}_{RP1}\right)} \times 100 \text{ and}$$

$$PRE\left(\hat{\mu}_{\mathrm{Re}\,gP1}\right) = \frac{MSE\left(\hat{\mu}_{YP1}\right)}{MSE\left(\hat{\mu}_{\mathrm{Re}\,gP1}\right)} \times 100.$$

Table 1: Empirical and Theoretical MSE, PRE for the Proposed Mean Estimator in proposed Model with respect to $\hat{\mu}_{YG}$ for Population 1 and 2

ρ_{YX}	n	n_1	n_2	W	Estimation	MSE Estimation		PRE
						Theoretical	Empirical	
0.8	500	250	250	0.3	$\hat{\mu}_{YG}$	0.0445	0.0437	100
					$\hat{\mu}_{YP1}$	0.0287	0.0288	155.13
				0.5	$\hat{\mu}_{YG}$	0.0456	0.0431	100
					$\hat{\mu}_{YP1}$	0.0287	0.0277	158.55
				0.7	$\hat{\mu}_{YG}$	0.0464	0.0456	100
					$\hat{\mu}_{YP1}$	0.0287	0.0289	161.65
	1000	500	500	0.3	$\hat{\mu}_{YG}$	0.0211	0.0196	100
					$\hat{\mu}_{YP1}$	0.0136	0.0131	155.13
				0.5	$\hat{\mu}_{YG}$	0.0216	0.0210	100
					$\hat{\mu}_{YP1}$	0.0136	0.0136	158.55
				0.7	$\hat{\mu}_{YG}$	0.0220	0.0207	100
					$\hat{\mu}_{YP1}$	0.0136	0.0132	161.65
0.3	500	250	250	0.3	$\hat{\mu}_{YG}$	0.0445	0.0433	100
					$\hat{\mu}_{YP1}$	0.0287	0.0284	155.13
				0.5	$\hat{\mu}_{YG}$	0.0456	0.0444	100
					$\hat{\mu}_{YP1}$	0.0287	0.0291	158.55
				0.7	$\hat{\mu}_{YG}$	0.0464	0.0443	100
					$\hat{\mu}_{YP1}$	0.0287	0.0275	161.65
	1000	500	500	0.3	$\hat{\mu}_{YG}$	0.0211	0.0195	100
					$\hat{\mu}_{YP1}$	0.0136	0.0131	155.138
				0.5	$\hat{\mu}_{YG}$	0.0216	0.0204	100
					$\hat{\mu}_{YP1}$	0.0136	0.0134	158.55
				0.7	$\hat{\mu}_{YG}$	0.0220	0.0201	100
					$\hat{\mu}_{YP1}$	0.0136	0.0130	161.65

Table 2: Empirical and Theoretical MSE, PRE for the Ratio Estimator In proposed Model with respect to $\hat{\mu}_{YP1}$ for Population 1 and 2

ρ_{YX}	n	n_1	n_2	W	Estimation	MSE Estimation		PRE
						Theoretical	Empirical	
0.8	500	250	250	0.3	$\hat{\mu}_{YP1}$	0.0287	0.0286	100
					$\hat{\mu}_{RP1}$	0.0183	0.0181	156.99
				0.5	$\hat{\mu}_{YP1}$	0.0287	0.0275	100
					$\hat{\mu}_{RP1}$	0.0183	0.0175	157.17
				0.7	$\hat{\mu}_{YP1}$	0.0287	0.0288	100
					$\hat{\mu}_{RP1}$	0.0183	0.0182	157.0276
	1000	500	500	0.3	$\hat{\mu}_{YP1}$	0.0136	0.0138	100
					$\hat{\mu}_{RP1}$	0.0086	0.0083	156.99
				0.5	$\hat{\mu}_{YP1}$	0.0136	0.0139	100
					$\hat{\mu}_{RP1}$	0.0086	0.0090	157.17
				0.7	$\hat{\mu}_{YP1}$	0.0136	0.0137	100
					$\hat{\mu}_{RP1}$	0.0086	0.0089	157.02
0.3	500	250	250	0.3	$\hat{\mu}_{YP1}$	0.0287	0.0281	100
					$\hat{\mu}_{RP1}$	0.0354	0.0345	81.16
				0.5	$\hat{\mu}_{YP1}$	0.0287	0.0269	100
					$\hat{\mu}_{RP1}$	0.0354	0.0326	81.26
				0.7	$\hat{\mu}_{YP1}$	0.0287	0.0269	100
					$\hat{\mu}_{RP1}$	0.0354	0.0358	81.20
	1000	500	500	0.3	$\hat{\mu}_{YP1}$	0.0136	0.0136	100
					$\hat{\mu}_{RP1}$	0.0167	0.0170	81.16
				0.5	$\hat{\mu}_{YP1}$	0.0136	0.0134	100
					$\hat{\mu}_{RP1}$	0.0167	0.0169	81.26
				0.7	$\hat{\mu}_{YP1}$	0.0136	0.0139	100
					$\hat{\mu}_{RP1}$	0.0167	0.0175	81.20

Table 3: Empirical and Theoretical MSE, PRE for the Proposed Regression Estimator in proposed Model with respect to $\hat{\mu}_{YP1}$ for Population 1 and 2

ρ_{YX}	n	n_1	n_2	W	Estimation	MSE Estimation		PRE
						Theoretical	Empirical	
0.8	500	250	250	0.3	$\hat{\mu}_{YP1}$	0.0287	0.0286	100
					$\hat{\mu}_{\mathrm{Re}\,gP1}$	0.0168	0.0168	169.77
				0.5	$\hat{\mu}_{YP1}$	0.0287	0.0275	100
					$\hat{\mu}_{\mathrm{Re}\,gP1}$	0.0176	0.0175	163.19
				0.7	$\hat{\mu}_{YP1}$	0.0287	0.0288	100
					$\hat{\mu}_{\mathrm{Re}\,gP1}$	0.0176	0.0177	161.28
	1000	500	500	0.3	$\hat{\mu}_{YP1}$	0.0136	0.0138	100
					$\hat{\mu}_{\mathrm{Re}\,gP1}$	0.0083	0.0083	161.31
				0.5	$\hat{\mu}_{YP1}$	0.0136	0.0139	100
					$\hat{\mu}_{\mathrm{Re}\,gP1}$	0.0083	0.0080	168.79
				0.7	$\hat{\mu}_{YP1}$	0.0136	0.0137	100
					$\hat{\mu}_{\mathrm{Re}\,gP1}$	0.0083	0.0082	164.19
0.3	500	250	250	0.3	$\hat{\mu}_{YP1}$	0.0287	0.0281	100
					$\hat{\mu}_{\mathrm{Re}\,gP1}$	0.0270	0.0280	101.84
				0.5	$\hat{\mu}_{YP1}$	0.0287	0.0269	100
					$\hat{\mu}_{\mathrm{Re}\,gP1}$	0.0270	0.0271	105.38
				0.7	$\hat{\mu}_{YP1}$	0.0287	0.0269	100
					$\hat{\mu}_{\mathrm{Re}\,gP1}$	0.0270	0.0278	102.631
	1000	500	500	0.3	$\hat{\mu}_{YP1}$	0.0136	0.0136	100
					$\hat{\mu}_{\mathrm{Re}\,gP1}$	0.0127	0.0125	105.69
				0.5	$\hat{\mu}_{YP1}$	0.0136	0.0134	100
					$\hat{\mu}_{\mathrm{Re}\,gP1}$	0.0128	0.0129	105.69
				0.7	$\hat{\mu}_{YP1}$	0.0136	0.0139	100
					$\hat{\mu}_{\mathrm{Re}\,gP1}$	0.0128	0.0122	105.69

4. Concluding Remarks

From the Table 1, we can conclude that, the proposed mean estimator in will always perform better than the proposed mean estimator of Gupta et al. (2010) estimator. In the Tables 2 and 3, it is observed that, the proposed ratio and regression estimators will always perform better than the proposed mean estimator in suggested model. We consider the data set similar to data used in Gupta et al. (2014). The percent relative efficiency of the proposed estimator is greater than the mean estimator in high correlation and also perform better in low correlation coefficient. It is noted that the proposed estimator also give maximum efficiency when the sensitivity level equal to 0.5. We also compare for two different sample sizes such as for 500 and 1000. So by increasing sample size there is no effect on efficiency of increasing sample size.

References

1. Gupta, S., Gupta, B. and Singh, S. (2002). Estimation of sensitivity level of personal interview survey questions. *Journal of Statistical Planning and Inference*, 100, 239-247.
2. Gupta, S., Shabbir, J. and Sehra, S. (2010). Mean and sensitivity estimation in optional randomized response models. Journal of Statistical Planning and Inference, 140(10), 2870-2874.
3. Gupta, S.N., Mehta, S., Shabbir, J. and Dass, B. K. (2012). Generalized scrambling in quantitative optional randomized response models. Communications in Statistics - Theory and Methods, in press.
4. Gupta, S.N and Shabbir, J. (2004). Sensitivity estimation for personal interview survey questions. Statistica, 64, 643-653.
5. Gupta, S., Kalucha, G., Shabbir, J. and Dass, B.K. (2014). Estimation of finite population mean using optional models in the presence of non-sensitive auxiliary information. American Journal of Mathematical and Management Sciences, 33(2), 147-159.
6. Mushtaq, N., Noor-ul-Amin, M. and Hanif, M. (2016). Estimation of a Population Mean of a Sensitive Variable in Stratified Two-Phase Sampling, Pakistan Journal of Statistics 32(5), 393-404.
7. Mushtaq, N., Noor-ul-Amin, M. and Hanif, M. (2017). A Family of Estimators of a Sensitive Variable Using Auxiliary Information in Stratified Random Sampling, Pakistan Journal of Statistics and operation research. XIII (1), 141-155.
8. Warner, S.L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. Journal of the American Statistical Association, 60, 63-69.