Research Article GU J Sci 32(3): 1021-1043 (2019) DOI: 10.35378/gujs.418681

Gazi University

Journal of Science

http://dergipark.gov.tr/gujs

Estimation of Mean of a Sensitive Quantitative Variable in Complex Survey: Improved Estimator and Scrambled Randomized Response Model

Iram SALEEM[1], Aamir SANAULLAH[2], Nursel KOYUNCU[3,*], Muhammad HANIF[1]

[1]National College of Business Administration and Economics Lahore, Pakistan

[2]COMSATS Institute Of Information Technology, Lahore, Pakistan

[3]Department of Statistics, Hacettepe University, Beytepe, Ankara, Turkey

Highlights

- This paper introduces a new Scrambled Randomized Response model.
- A simulation study is done to compare the efficiency of the proposed model.
- The results showed that the proposed model performs better than existing model.

Article Info

Received: 26/04/2018
Accepted: 17/12/2019

Keywords

Scrambled Randomized Response
Sensitive Variable
Exponential Estimator
Simulation

Abstract

With the intention to control a true swapping between the efficiency and the privacy protection this paper introduces a scrambled randomized response (SRR) model to be alternative of Saha's scrambling mechanism. The basic initiative is to provide an assortment of the additive, the subtractive and the multiplicative models. The simulation and the empirical studies are provided for various sample sizes to compare the efficiency of the proposed model. The results obtained from simulation showed that the proposed model performs better than Pollock and Bek's additive model. Also, the proposed generalized estimator of mean has been studied using a new SRR model presented in this article and shown that the proposed estimator and its class of estimators are more efficient than existing estimators. It is also shown that gain in efficiency is more when the proposed SRR model is used. The efficiency of the proposed class of estimators over existing estimators using both models is also provided using real data and with a simulation study.

1. INTRODUCTION

In surveys having sensitive issues such as tax evasion, gambling, child abuse and drug misuse, we do not obtain trustworthy information easily from the respondents. To get reliable and valid information in such surveys, Warner [1] introduced a randomized response technique (RRT) to protect the privacy of respondents. This method is easy to use for gathering information about sensitive qualitative response variable, for example, to estimate the proportion of women in a community enduring induced abortions. Unfortunately, this proposed method is not applicable for sensitive quantitative response variable, for instance, to find out the average abortion number in a community. Greenberg et al. [2] and Pollock and Bek [3] extended the work of Warner [1] to obtain a reliable response for quantitative variables. Himmelfarb and Edgell [4] revisited the Pollock and Bek's [3] additive model and studied it in more detail for estimating the mean of a quantitative sensitive variable.

Further, Eichhorn and Hayre [5] introduced a multiplicative model for quantitative surveys. Eichhorn and Hayre [5] presented the concept of scrambled randomized response technique (SRR) in which the respondent provides a quantitative reply to the sensitive question asked and then response is multiplied to

*Corresponding author, e-mail: nkoyuncu@hacettepe.edu.tr

the random number generated by the respondents themselves from some randomized device. Further contribution and modifications in SRR technique by developing some new models, have been made by several authors such as, Saha [6], Diana and Perri [7,8]. Hussain [9] considered a subtractive SRR model which involves the respondent subtracting the scrambling variable from the answer to sensitive questions. Many different modifications in SRR technique to develop some new models have been made by several authors such as, Chen and Singh [10], Singh and Tarray [11] and Hussain and Al-Zahrani [12]. We can list recent important studies about scrambled response models under ranked set sampling (RSS) as follows: Bouza [13] considered some important scrambled response models under the use of simple random sampling, RSS and Rao–Hartley–Cochran designs. Bouza [14] examined scrambled responses in RSS. Bouza et al. [15] studied the an Optional Scrambling Randomized Response Modeling procedure under RSS design.

Several authors such as Sousa et al. [16], Gupta et al. [17] , Tarray and Singh [18] and Koyuncu et al. [19] presented different estimators to estimate the population mean of a sensitive study variable using a non-sensitive auxiliary variable.

Saha [6, 20]'s model can be considered when survey questions are extremely sensitive. In such a situation scrambling device provides a greater protection to the respondent but on the other hand efficiency of the estimators cannot be expected to be good. Therefore, this study aims to look into a possible alternative to Saha's approach that is more flexible in terms of correctness and secrecy protection. With intention to hold a right trade-off between the efficiency and secrecy protection, an attempt is made successfully to provide a blend of the three models, Pollock and Bek's [3] additive, Hussain's [9] subtractive and Eichhron and Hayre's [5] multiplicative models are provided and show that proposed model performs better. Further aim of the study is extended to show that the proposed estimator performs better than some existing estimators when either our proposed SRR model or the Pollock and Bek's model is used. The main objective is to provide the evidence that under the proposed SRR model the proposed estimators as well as the existing estimators, perform more efficiently. In Section2, a general SRR model is proposed with its properties. In Section3, an improved class of generalized estimators is discussed. Further, the proposed estimator is also shown to be an almost unbiased and the most efficient estimator. Also the optimal conditions in order to attain the minimum mean square errors are discussed. In Section4, a simulation study is used to show the dominance of the proposed procedure over existing Pollock and Bek's additive model and the efficiency of the proposed generalized estimator. The concluding remarks are given in Section5.

Pollock and Bek [3] introduced an additive model to obtain reliable answers for a quantitative sensitive variable under study. In this approach, the respondent is asked to sum of the sensitive value say Y and the random scrambling value say S from the known distribution, independent to Y. Suppose a random sample of size n is drawn without replacement from a finite population $U=(U_1,U_2,\ldots,U_N)$. For the ith unit $(\mathrm{i}=1,2,\ldots,\mathrm{N})$ let Y_i and S_i be the values of study and scrambling variable. Let $\bar{y}=\sum_{i=1}^{n}Y_i \Big/ n$ and $\bar{Y}=\sum_{i=1}^{N}Y_i \Big/ N$ be the sample mean and population mean of Y respectively and $\bar{S}=\sum_{i=1}^{N}S_i \Big/ N$ denotes the population mean of S. We can define following expectations $E(S)=\bar{S}=0$ and $E(Y)=\bar{Y}$

The reported response following Pollock and Bek's [3] additive model is given by,

$$Z_{PB}=Y+S \tag{1}$$

for the ith unit $(\mathrm{i}=1,2,\ldots,\mathrm{N})$, Z_i denotes the sum of the sensitive value and the random scrambling value the estimated mean for Z_{PB} can be estimated by

$$\hat{\mu}_{Z_{PB}} = \bar{z} = \frac{\sum_{i=1}^{n} Z_i}{n}, \tag{2}$$

with variance

$$Var\left(\hat{\mu}_{Z_{PB}}\right) = \theta\left(S_Y^2 + S_S^2\right), \tag{3}$$

where $\theta = \frac{1}{n} - \frac{1}{N}$, $S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(Y_i - \bar{Y}^2\right)$ and $S_s^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(S_i - \bar{S}^2\right)$ are the variances of Y and S, respectively.

Eichhorn and Hayre [5] presented a multiplicative approach adapting the initial idea of Pollock and Bek [3] in depth and assuming another scrambling random variable R with the mean and variance respectively as $E\left(R\right) = \bar{R}$ and $Var(R) = S_R^2$. Then the observed scrambled response is

$$Z_{EH} = YR \tag{4}$$

and the estimated mean of $\bar{Y}$ is

$$\hat{\mu}_{Z_{EH}} = \frac{\bar{z}}{\bar{R}}, \tag{5}$$

where $\bar{z}_{EH} = \frac{\sum_{i=1}^{n} z_{EH(i)}}{n}$ is the sample mean of *n* coded responses. The variance of $\hat{\mu}_{EH}$ is given by,

$$Var\left(\hat{\mu}_{Z_{EH}}\right) = \frac{1}{n}\left[S_Y^2 + \bar{Y}^2\left(1 + C_Y^2\right)C_R^2\right] \tag{6}$$

where $C_Y^2 = \frac{S_y^2}{\bar{Y}^2}$ and $C_R^2 = \frac{S_R^2}{\bar{R}^2}$.

The auxiliary information can be used at both design and estimation stages to gain precision of the estimators of population parameters. Diana and Perri [7, 8] proposed some classes of estimators using a sensitive auxiliary variable. Likewise, different estimators to estimate the population mean of a sensitive study variable using a non-sensitive auxiliary variable were presented by several authors. Sousa et al. [16] presented a ratio estimator for sensitive study variable (Y) using a non-sensitive auxiliary variable (X_1). The estimator t_S presented by Sousa et al. [16] is given by,

$$t_S = \bar{z}\,\frac{\bar{X}_1}{\bar{x}_1}, \tag{7}$$

where $\bar{x}_1$ and $\bar{X}_1$ are the sample mean and population mean of non-sensitive auxiliary variable respectively. The MSE of t_S is given by,

$$MSE\left(t_s\right)=\theta\bar{Y}^2\left(C_z^2+C_{x_1}^2-2C_zC_{x_1}\rho_{zx_1}\right), \tag{8}$$

where C_z and C_{x_1} are the coefficient of variation of Z and X_1 respectively. The ratio estimators are widely used when the correlation coefficient between the randomized response variable and auxiliary variable is strong and positive. The estimator t_S provides better performance than usual unbiased mean estimator $\bar{z}$ when $\rho_{ZX_1}>\frac{C_{X_1}}{2C_Z}$ and usual unbiased estimator is preferred over t_S if $-\frac{C_{X_1}}{2C_Z}<\rho_{zx}<\frac{C_{X_1}}{2C_Z}$.
Gupta et al. [17] proposed a regression-cum-ratio estimator is written as,

$$t_G=\left[k_1\bar{z}+k_2\left(\bar{X}_1-\bar{x}_1\right)\right]\left(\frac{\bar{X}_1}{\bar{x}_1}\right), \tag{9}$$

where k_1 and k_2 are constants.

$$\begin{aligned}MSE\left(t_s\right)=&\left(k_1-1\right)\bar{Y}^2+k_1^2\bar{Y}^2\theta\left(C_z^2+3C_{x_1}^2-4C_zC_{x_1}\rho_{zx_1}\right)+k_2^2\bar{X}^2\theta C_{x_1}^2\\&-2k_1\bar{Y}^2\theta\left(C_{x_1}^2-C_zC_{x_1}\rho_{zx_1}\right)-2k_2\overline{YX}\theta C_{x_1}^2-2k_1k_2\overline{YX}\theta\left(C_zC_{x_1}\rho_{zx_1}-2C_{x_1}^2\right).\end{aligned} \tag{10}$$

Koyuncu et al. [19] suggested two generalized regression-cum-exponential estimator using non-sensitive auxiliary variables. The first proposed estimator and its MSE are given by,

$$t_{K_1}=\left[w_1\bar{z}+w_2\left(\bar{X}_1-\bar{x}_1\right)\right]\exp\left(\frac{\bar{X}_1-\bar{x}_1}{\bar{X}_1+\bar{x}_1}\right), \tag{11}$$

$$\begin{aligned}\mathrm{MSE}\left(\mathrm{t}_{K_1}\right)=&\left\{\bar{\mathrm{Z}}^2+\mathrm{w}_1{}^2\bar{\mathrm{Z}}^2\left(1\text{-}1\left(\mathrm{C}_z^2+\mathrm{C}_{x_1}^2\text{ - }2\mathrm{C}_{zx_1}\right)\right)+\mathrm{w}_2^2\bar{\mathrm{X}}_1^2\,\mathrm{C}_{x1}^2\right.\\&\left.+\mathrm{w}_1\bar{\mathrm{Z}}^2\left(1\left(\mathrm{C}_{zx_1}+\frac{3}{4}\mathrm{C}_{x_1}^2\right)\text{- }2\right)\text{- }\mathrm{w}_2\bar{\mathrm{Z}}\bar{\mathrm{X}}_1\mathrm{C}_{x1}^2+2\mathrm{w}_1\mathrm{w}_2\bar{\mathrm{Z}}\bar{\mathrm{X}}_1\left(\mathrm{C}_{x_1}^2\text{ - }\mathrm{C}_{zx_1}\right)\right\}.\end{aligned} \tag{12}$$

Their second estimator and its MSE are given by,

$$t_{K_2}=\left[d_1\bar{z}+d_2\left(\bar{X}_1-\bar{x}_1\right)+d_3\left(\bar{X}_2-\bar{x}_2\right)\right]\exp\left(\frac{\left(\bar{X}_1-\bar{x}_1\right)+\left(\bar{X}_2-\bar{x}_2\right)}{\left(\bar{X}_1+\bar{x}_1\right)+\left(\bar{X}_2+\bar{x}_2\right)}\right), \tag{13}$$

$$\begin{aligned}MSE\left(t_{K_2}\right)=&\left\{\bar{Z}^2+d_1A-d_2B-d_3C+d_1^2D+d_2^2\bar{X}_1^2\theta C_{x_1}^2+d_3^2\bar{X}_2^2\theta C_{x2}^2\right.\\&\left.+2d_1d_2F+2d_1d_3G+2d_1d_3\bar{X}_1\bar{X}_2\theta C_{x_1x_2}\right\},\end{aligned} \tag{14}$$

Where

$$A=\bar{Z}^2\left(-2+\theta\left\{\frac{\bar{X}_1C_{zx_1}}{\left(\bar{X}_1+\bar{X}_2\right)}+\frac{\bar{X}_2C_{zx_2}}{\left(\bar{X}_1+\bar{X}_2\right)}-\frac{3\bar{X}_1^2C^2{}_{x_1}}{4\left(\bar{X}_1+\bar{X}_2\right)^2}-\frac{6\bar{X}_1\bar{X}_2C_{x_1x_2}}{4\left(\bar{X}_1+\bar{X}_2\right)^4}-\frac{3\bar{X}_2^2C^2{}_{x_2}}{4\left(\bar{X}_1+\bar{X}_2\right)^2}\right\}\right),$$

$$B=\theta\frac{\bar{Z}}{\left(\bar{X}_1+\bar{X}_2\right)}\left(\bar{X}_1^2C^2{}_{x_1}+\bar{X}_1\bar{X}_2C_{x_1x_2}\right),\ C=\theta\frac{\bar{Z}}{\left(\bar{X}_1+\bar{X}_2\right)}\left(\bar{X}_2^2C^2{}_{x_2}+\bar{X}_1\bar{X}_2C_{x_1x_2}\right),$$

$$D=\bar{Z}^2\left(1+\theta\left\{C_z^2+\frac{\bar{X}_1^2C^2{}_{x_1}}{\left(\bar{X}_1+\bar{X}_2\right)^2}+\frac{\bar{X}_2^2C^2{}_{x_2}}{\left(\bar{X}_1+\bar{X}_2\right)^2}-2\frac{\bar{X}_1C_{zx_1}}{\left(\bar{X}_1+\bar{X}_2\right)}-2\frac{\bar{X}_2C_{zx_2}}{\left(\bar{X}_1+\bar{X}_2\right)}+2\frac{\bar{X}_1\bar{X}_2C_{x_1x_2}}{\left(\bar{X}_1+\bar{X}_2\right)^4}\right\}\right),$$

$$F=\theta\bar{Z}\left(\frac{\bar{X}_1^2C^2{}_{x_1}}{\left(\bar{X}_1+\bar{X}_2\right)}-\bar{X}_1C_{zx_1}+\frac{\bar{X}_1\bar{X}_2C_{x_1x_2}}{\left(\bar{X}_1+\bar{X}_2\right)}\right),\ G=\theta\bar{Z}\left(\frac{\bar{X}_2^2C^2{}_{x_2}}{\left(\bar{X}_1+\bar{X}_2\right)}-\bar{X}_2C_{zx_2}+\frac{\bar{X}_1\bar{X}_2C_{x_1x_2}}{\left(\bar{X}_1+\bar{X}_2\right)}\right).$$

2. THE PROPOSED SRR MODEL

In this section, we have introduced a new SRR model by combining the three models, namely Pollock and Bek [3] additive, Hussain [9] subtractive and Eichhorn and Hayre [5] multiplicative model. The reported response Z_{new} is given as,

$$Z_{new}=g\left(Y+aS\right)+\left(1-g\right)YS, \tag{15}$$

where $g\in[0,1]$ and $\alpha\in[-1,1]$ are suitable constants controlled by the researcher. Considering $\bar{S}=E\left(S\right)=0$ and S_S^2 be the mean and variance of scrambling random variable S. The choice of $(g=1,\alpha=1)$ provides Pollock and Bek [3] additive model whereas $(g=1,\alpha=-1)$, generates the subtractive model of Hussain [9]. Also, we notice that (g=0) will result in the multiplicative model of Eichhorn and Hayre [5], whereas for (g=1) we may get Tarray and Singh [18] SRR model. The resulting mean, variance and covariance for the proposed model are established as respectively

$$E\left(Z_{new}\right)=g\bar{Y} \tag{16}$$

$$Var\left(Z_{new}\right)=g^2\left(S_y^2+a^2S_s^2\right)+\left(1-g\right)^2S_s^2\left(S_y^2+\bar{Y}^2\right)+2ag\left(1-g\right)\bar{Y}S_s^2 \tag{17}$$

and

$$\sigma_{X_iZ_{new}}=g\sigma_{X_iY}. \tag{18}$$

Therefore, the coefficient of variation $\left(C_Z\right)$ and the correlation $\left(\rho_{ZX_1},\rho_{ZX2}\right)$ will become,

$$C^2_{z_{new}}=\frac{1}{\bar{Z}^2_{new}}\left[g^2\left(S_y^2+a^2S_s^2\right)+\left(1-g\right)^2S_s^2\left(S_y^2+\bar{Y}^2\right)+2ag\left(1-g\right)\bar{Y}S_s^2\right], \tag{19}$$

$$\rho_{X_iZ_{new}}=\rho_{YX_i}\frac{gS_Y}{\sqrt{\left[g^2\left(S_y^2+a^2S_s^2\right)+\left(1-g\right)^2S_s^2\left(S_y^2+\bar{Y}^2\right)+2ag\left(1-g\right)\bar{Y}S_s^2\right]}},\ (i=1,2). \tag{20}$$

Following Gupta et al. [21] the unified measure is used to evaluate the privacy level of the proposed model and some existing models. The unified measure is given as,

$$\delta=\frac{Var\left(\hat{\mu}_Y\right)}{Privacy\,level}, \tag{21}$$

where $Var\left(\hat{\mu}_Y\right)$ is the variance of the mean estimator of the model and privacy level is the privacy measure for same model using Yan et al. [22] i-e. $E\left(Z-Y\right)^2$. In Table 1, the unified measures of different models are presented with the $Var\left(\hat{\mu}_Y\right)$ and privacy level.

3. THE PROPOSED GENERALIZED REGRESSION-CUM-EXPONENTIAL ESTIMATOR

In this section, we have proposed an improved exponential type estimator of the finite population mean of sensitive study variable using two non-sensitive auxiliary variables. The expressions of the bias and MSE of the proposed estimator are obtained up to first order of approximation. Let Y be a sensitive study variable, which cannot be directly obtained. Let X_1 and X_2 be the non-sensitive auxiliary variables, positively correlated with Y. Consider a scrambling variable S, be independent of Y, X_1 and X_2. The respondents are asked to provide true response for X_1 and X_2 and to report a scrambled response for Y given as $Z=Y+S$. Suppose a random sample n is drawn without replacement from a finite population $U=(U_1,U_2,\ldots,U_N)$. Let Y_i, X_{1i} and X_{2i} be the *i*th unit values of study variable Y and auxiliary variables X_1 and X_2, respectively. Let the sample means be $\bar{y}=\frac{\sum_{i=1}^{n}Y_i}{n}$, $\bar{x}_1=\frac{\sum_{i=1}^{n}X_{1i}}{n}$, $\bar{x}_2=\frac{\sum_{i=1}^{n}X_{2i}}{n}$ and $\bar{z}=\frac{\sum_{i=1}^{n}Z_i}{n}$ corresponding to the $\bar{Y}=E\left(Y\right)$, $\bar{X}_1=E\left(X_1\right)$, $\bar{X}_2=E\left(X_2\right)$ and $\bar{Z}=E\left(Z\right)$ be the population means of Y, X_1, X_2 and Z, respectively. The auxiliary variables X_1 and X_2 are assumed to be known and $\bar{S}=E\left(S\right)=0$. Therefore, $E\left(Z\right)=E\left(Y\right)$ and $C_Z^2=C_Y^2+\frac{S_s^2}{\bar{Y}^2}$, where C_Z^2 and C_Y^2 are the coefficient of variations of Z and Y, respectively.

Table 1. *Unified Measure of the Different Models*

Models	$Var(\hat{\mu}_Y)$	**Privacy Level**	δ
Pollock and Bek [3]	$Var(\hat{\mu}_{PB})=\frac{S_Y^2+S_S^2}{n}$	S_S^2	$\frac{S_Y^2+S_S^2}{nS_S^2}$
Eichhorn and Hayre [5]	$Var(\hat{\mu}_{EH})=\frac{1}{n}\left[S_Y^2+\bar{Y}^2+C_R^2\left(1+C_Y^2\right)\right]$	$\left(\bar{Y}^2+S_Y^2\right)\left(1-\bar{R}\left(2-\bar{R}\right)+S_R^2\right)$	$\frac{\left[1+\frac{C_R^2\left(1+C_Y^2\right)}{\left(S_Y^2+\bar{Y}^2\right)}\right]}{n\left[1-\bar{R}\left(2-\bar{R}\right)+S_R^2\right]}$
Saha [20]	$Var(\hat{\mu}_{Saha})$ $=\frac{1}{n}\left[\left(\bar{R}^2+S_R^2\right)\left(\bar{Y}^2+S_Y^2+S_S^2\right)-\bar{Y}^2\right]$	$\left(\bar{R}^2+S_R^2\right)\left[\bar{Y}^2+S_Y^2+S_S^2\right]$ $+\bar{Y}^2-2\bar{R}\left(\bar{Y}^2+S_Y^2\right)$	$\frac{\left\{1-\frac{\bar{Y}^2}{\left[\left(\bar{Y}^2+S_Y^2+S_S^2\right)\left(\bar{R}^2+S_R^2\right)\right]}\right\}}{n\left[1+\frac{\bar{Y}^2-2\bar{R}\left(\bar{Y}^2+S_Y^2\right)}{\left[\left(\bar{Y}^2+S_Y^2+S_S^2\right)\left(\bar{R}^2+S_R^2\right)\right]}\right]}$
Hussian [9]	$Var(\hat{\mu}_H)=\frac{S_Y^2+S_S^2}{n}$	S_S^2	$\frac{S_Y^2+S_S^2}{nS_S^2}$
Singh and Tarray [11]	$Var(\hat{\mu}_{ST})=\frac{S_Y^2+a^2S_S^2}{n}$	$a^2S_S^2$	$\frac{S_Y^2+S_S^2}{na^2S_S^2}$
Proposed SRR model	$Var(\hat{\mu}_{Z_{new}})=\frac{1}{n}\left[g^2\left(S_Y^2+a^2S_S^2\right)\right.$ $\left.+(1-g)^2S_S^2\left(\bar{Y}^2+S_Y^2\right)+2ag(1-g)\bar{Y}S_s^2\right]$	$\left(\bar{Y}^2+S_Y^2\right)\left((1-g)^2S_S^2+g^2+1\right)$ $+g\left(2\bar{Y}\left(a(1-g)S_S^2-\bar{Y}\right)+ga^2S_S^2\right)$	$\frac{\left\{1+\frac{g^2\left(S_Y^2+a^2S_S^2\right)+2ag(1-g)\bar{Y}S_s^2}{\left(\bar{Y}^2+S_Y^2\right)}\right\}}{n\left[\left(1+g^2+(1-g)^2S_S^2\right)+\frac{g\left\{2\bar{Y}\left(a(1-g)S_S^2-\bar{Y}\right)+a^2gS_S^2\right\}}{\left(\bar{Y}^2+S_Y^2\right)}\right]}$

Following Rao [23] and considering a linear relation between the auxiliary variable X_1 and the sensitive study variable Y, we may consider a regression-type estimator t_{reg} using single non-sensitive auxiliary variable under randomized response model $Z=Y+S$ as,

$$t_{reg}=\alpha_0\bar{z}+\alpha_1\left(\bar{X}_1-\bar{x}_1\right), \tag{22}$$

where α_0 and α_1 are the suitable constants.

Similarly following Bahl and Tuteja [24] and considering an exponential-type relation between the auxiliary variable X_2 and the sensitive study variable Y, we may reproduce the exponential-type ratio estimator t_{er} and exponential-type product estimator t_{ep} respectively under randomized response model $Z=Y+S$ as,

$$t_{er}=\bar{z}\exp\left(\frac{\bar{X}_2-\bar{x}_2}{\bar{X}_2+\bar{x}_2}\right) \text{ and } t_{ep}=\bar{z}\exp\left(\frac{\bar{x}_2-\bar{X}_2}{\bar{X}_2+\bar{x}_2}\right). \tag{23}$$

Then the estimators in (22)-(23) may lead us to propose a generalized regression-cum-exponential type estimator as,

$$t_g=\left(\alpha_0\bar{z}+\alpha_1\left(\bar{X}_1-\bar{x}_1\right)\right)\left[\exp\left(\frac{b\left(\bar{X}_2-\bar{x}_2\right)}{\bar{X}_2+(\alpha_2-1)\bar{x}_2}\right)\right], \tag{24}$$

where $b(-1,0,1)$ is a generalization constant its values yield regression estimator, class of regression-cum-exponential ratio and regression-cum-exponential product estimator respectively. α_0,α_1 and α_2 are assumed to be unknown constants and need to be estimated such that the MSE of t_g is minimum.

It is observed that regression estimator t_{reg} in (22) may be obtained for $b=0$. Also, we get exponential estimator based on single auxiliary variable. Also for different choices of these constants, we may get different estimators, as see in Table 2. It provides some example of different estimators based on two auxiliary variables.

Table 2. *Class of Estimators for Two Auxiliary Variables*

Class of estimators	α_0	α_1	b	α_2
$t_{g1}=\left(\alpha_0\bar{z}+\left(\bar{X}_1-\bar{x}_1\right)\right)\left[\exp\left(\frac{\left(\bar{X}_2-\bar{x}_2\right)}{\bar{X}_2}\right)\right]$	α_0	1	1	1
$t_{g2}=\left(\alpha_0\bar{z}+b_{zx1}\left(\bar{X}_1-\bar{x}_1\right)\right)\left[\exp\left(\frac{\left(\bar{X}_2-\bar{x}_2\right)}{\bar{X}_2}\right)\right]$	α_0	b_{zx1}	1	1
$t_{g3}=\left(\alpha_0\bar{z}+\rho_{zx1}\left(\bar{X}_1-\bar{x}_1\right)\right)\left[\exp\left(\frac{\left(\bar{X}_2-\bar{x}_2\right)}{\bar{X}_2}\right)\right]$	α_0	ρ_{zx1}	1	1

$t_{g4}=\left(\alpha_0\bar{z}+\left(\bar{X}_1-\bar{x}_1\right)\right)\left[\exp\left(\frac{\bar{X}_2-\bar{x}_2}{2\left(\bar{X}_2+\bar{x}_2\right)}\right)\right]$	α_0	1	½	2
$t_{g5}=\left(\alpha_0\bar{z}+b_{zx1}\left(\bar{X}_1-\bar{x}_1\right)\right)\left[\exp\left(\frac{\bar{X}_2-\bar{x}_2}{2\bar{X}_2}\right)\right]$	α_0	b_{zx1}	½	1
$t_{g6}=\left(\alpha_0\bar{z}+C_{x1}\left(\bar{X}_1-\bar{x}_1\right)\right)\left[\exp\left(\frac{2\left(\bar{X}_2-\bar{x}_2\right)}{\bar{X}_2}\right)\right]$	α_0	C_{x1}	2	1
$t_{g7}=\left(\alpha_0\bar{z}+\rho_{zx1}\left(\bar{X}_1-\bar{x}_1\right)\right)\left[\exp\left(\frac{\rho_{zx2}\left(\bar{X}_2-\bar{x}_2\right)}{\bar{X}_2}\right)\right]$	α_0	ρ_{zx1}	ρ_{zx2}	1
$t_{g8}=\left(\bar{z}+b_{zx1}\left(\bar{X}_1-\bar{x}_1\right)\right)\left[\exp\left(\frac{\left(\bar{X}_2-\bar{x}_2\right)}{\bar{X}_2}\right)\right]$	1	b_{zx1}	1	1
$t_{g9}=\left(\bar{z}+\rho_{zx1}\left(\bar{X}_1-\bar{x}_1\right)\right)\left[\exp\left(\frac{\left(\bar{X}_2-\bar{x}_2\right)}{\bar{X}_2}\right)\right]$	1	ρ_{zx1}	1	1

To obtain the bias and MSE, let us define the following notations

$$e_z=\left(\bar{z}-\bar{Z}\right)/\bar{Z},\quad e_{X_1}=\left(\bar{x}_1-\bar{X}_1\right)/\bar{X}_1,\ e_{X_2}=\left(\bar{x}_2-\bar{X}_2\right)/\bar{X}_2$$

such that $E\left(e_i\right)=0 \quad i=z,x_1,x_2,$

and expectations are as,

$$E\left(e_z^2\right)=\theta C_z^2,\ E\left(e_{x_1}^2\right)=\theta C_{x_1}^2, E\left(e_{x_2}^2\right)=\theta C_{x_2}^2, E\left(e_z e_{x_1}\right)=\theta C_z C_{x_1}\rho_{zx_1},$$

$$E\left(e_z e_{x_2}\right)=\theta C_z C_{x_2}\rho_{zx_2}, E\left(e_{x_1}e_{x_2}\right)=\theta C_{x_1}C_{x2}\rho_{x_1x_2}, \tag{25}$$

where $\theta=\left(1-f\right)/n,\ f=n/N$. We may express (24) in terms of e's as,

$$\begin{aligned} t_g &\cong \left(\alpha_0\bar{Z}\left(1+e_Z\right)-\alpha_1\bar{X}_1e_{X_1}\right)\exp\left(\frac{-be_{X_2}}{\alpha_2}\left(1+\left(1-\frac{1}{\alpha_2}\right)e_{X_2}\right)^{-1}\right) \\ &=\left(\alpha_0\bar{Z}\left(1+e_Z\right)-\alpha_1\bar{X}_1e_{X_1}\right)\exp\left(\frac{-be_{X_2}}{\alpha_2}\left(1-\left(1-\frac{1}{\alpha_2}\right)e_{X_2}+...\right)\right). \end{aligned} \tag{26}$$

The bias and MSE of t_g to the first order approximation $O(n^{-1})$ are given by,

$$Bias\left(t_g\right)\cong\bar{Z}\left[\alpha_0\left(1+\frac{b^2}{\alpha^2_2}\theta C^2{}_{X_2}-\frac{b}{\alpha_2}\theta C_Z C_{X_2}\rho_{ZX_2}\right)+\alpha_1\frac{\bar{X}_1}{\bar{Z}}\frac{b}{\alpha_2}\theta C_{X_1}C_{X_2}\rho_{X_1X_2}-1\right]. \tag{27}$$

Equating (27) with zero, we may obtain

$$\alpha_2=\left(\frac{2\alpha_0 b}{A_1-\sqrt{A_1^2-4A_2}},\frac{2\alpha_0 b}{A_1+\sqrt{A_1^2-4A_2}}\right), \tag{28}$$

where

$$A_1=\left(\alpha_0 K_{ZX_2}-\alpha_1\frac{K_{X_1X_2}}{R_{ZX_1}}\right), A_2=\frac{\alpha_0\left(\alpha_0-1\right)}{\theta C^2_{X_2}},\ K_{ZX_2}=\frac{C_Z}{C_{X_2}}\rho_{ZX_2}, K_{X_1X_2}=\frac{C_{X_1}}{C_{X_2}}\rho_{X_1X_2}, R_{ZX_1}=\frac{\bar{Z}}{\bar{X}_1}. \tag{29}$$

By using the values of α_2, one may get an almost unbiased class of estimators for the population mean of sensitive study variable. Providing different values of the constants to (27), we can get expressions for the bias of different estimators.

A general MSE expression of the t_g can be obtained as,

$$MSE\left(t_{gi}\right)\cong\theta\bar{Z}^2\Bigg[\Bigg\{\alpha_{0i}^2\left(1+C^2{}_Z\right)+\alpha^2{}_{1i}\frac{\bar{X}_1^2}{\bar{Z}^2}C^2{}_{X_1}+\alpha_{0i}^2v_i^2C^2{}_{X_2}-2\alpha_{0i}\alpha_{1i}\frac{\bar{X}_1}{\bar{Z}}C_ZC_{X_2}\rho_{ZX_1}$$
$$-2\alpha_{0i}^2v_iC_ZC_{X_2}\rho_{ZX_2}+2\alpha_{0i}^2\alpha_{1i}\frac{\bar{X}_1}{\bar{Z}}v_iC_{X_1}C_{X_2}\rho_{X_1X_2}\Bigg\}+1-2\alpha_{0i}\Bigg], \tag{30}$$

where $v=\dfrac{b}{\alpha_2}$ and α_0, α_1 and v attain their optimum values as,

$$\alpha_0=\frac{1}{\left(1+C_ZR_3\right)},\ \alpha_1=R_1\frac{\bar{Z}C_Z}{\bar{X}C_{x_1}\left(1+C_ZR_3\right)},\ v=R_2\frac{C_Z}{C_{X_2}} \tag{31}$$

where

$$R_1=\frac{\left(\rho_{x1z}-\rho_{x2z}\rho_{x1x2}\right)}{\left(1-\rho^2{}_{x1x2}\right)},\ R_2=\frac{\left(\rho_{zx2}-\rho_{zx1}\rho_{x1x2}\right)}{\left(1-\rho^2{}_{x1x2}\right)}$$

and $R_3=R_1^2+R_2^2-2R_1\rho_{zx1}-2R_2\rho_{zx2}+2R_1R_2\rho_{x1x2}$.

Substituting the optimum values in (30), we may get the minimum value of the $MSE\left(t_g\right)$ as,

$$\min MSE\left(t_g\right)=\theta\bar{Z}^2C^2{}_Z\left[1-\frac{1}{\left(1+R_3C_z^2\right)}\right]. \tag{32}$$

We may get the minimum MSEs for t_{gi} $\left(i=1,2,\ldots,9\right)$ using different values of $\alpha_0,\alpha_1,\alpha_2$ and b in (32), (e.g. as given in Table 2).

Observe that optimum choices of the constants α_0,α_1 and α_2 are based on some parameters whose values may be from the prior surveys or may be presumed through the experience drawn in due course of time, for case in point see, Horvitz and Thompson [25], Sanaullah et al. [26], Asghar et al. [27], and Sanaullah et al. [28]. In many real life situations, it may not be possible to guess these values through experience or prior surveys, so it is better to replace these values with the estimates drawn from a pilot surveys for case in point see Koyuncu and Kadilar [29] and Jabeen et al. [30]. Also, we can verify that the generalized regression-cum-exponential estimator t_g in (24) is more efficient than the estimators t_m, t_S and t_G when

$$\left[S_S^2+\frac{S_Y^2}{\left(1+R_3C_Z^2\right)}\right]>0,\left[C_{X1}^2-2C_{ZX1}+\frac{C_Z^2}{\left(1+R_3C_Z^2\right)}\right]>0 \text{ and}$$

$$\left[\frac{\left(1-\rho_{ZX1}^2\right)\left(1+\theta C_X^2\right)}{C_Z^2\left(1-\rho_{ZX1}^2\right)\theta+\left(1+\theta C_X^2\right)}-\frac{R_3C_Z^2}{\left(1+R_3C_Z^2\right)}\right]>0 \text{ , respectively.}$$

Eq. (27)-(32) can be reproduced under the proposed model following the results in Eq. (15)-(20) for the proposed estimator as well as for the existing estimators.

4. SIMULATION STUDY

In this section the efficiency of the proposed class of estimators over existing estimators using proposed model and Pollock and Bek [3] is provided using real data and with a simulation study. This study is involved to evaluate the MSEs of the estimators both empirically and theoretically. This simulation study considers two finite populations of size N=1000 each generated from multivariate normal distribution with theoretical mean vector μ=[5,5,5] for [Y,X_1,X_2] with covariance matrices for two populations respectively as,

Population I

$$\sigma^2=\begin{bmatrix}10 & 3 & 2.9\\ 3 & 2 & 1.1\\ 2.9 & 1.1 & 2\end{bmatrix} \quad \rho_{x1y}=0.6817,\ \ \rho_{x2y}=0.6705$$

Population II

$$\sigma^2=\begin{bmatrix}6 & 3 & 2.9\\ 3 & 2 & 1.1\\ 2.9 & 1.1 & 2\end{bmatrix} \quad \rho_{x1y}=0.8706,\ \ \rho_{x2y}=0.8428$$

The scrambling variable S is taken from the normal distribution with mean zero and standard deviation equal to 10% of the standard deviation of X_1.

The reported response for the two models are taken as,

i) Pollock and Bek [3] → *Z=Y+S*
ii) New model
a) $Z=0.3\left(Y+0.5S\right)+\left(1-0.3\right)YS$ iff g=0.3 & a=0.5
b) $Z=0.6\left(Y+0.5S\right)+\left(1-0.6\right)YS$ iff g=0.6 & a=0.5
c) $Z=0.9\left(Y+0.5S\right)+\left(1-0.9\right)YS$ iff g=0.9 & a=0.5

The sample size for each population considered as n=50, 100, 200 and 300. The percent relative efficiency (PRE)s for Himmelfarb and Edgel [4] model and new model are calculated with the following equations:

$$PRE=\frac{MSE\left(t_m\right)}{MSE\left(t_j\right)}\times 100,$$

where $j=m,S,G,K_1,K_2,g\ and\ g_i\ (\mathrm{i}=1,2,\ldots,9)$.

In Table 3, the simulation results of the unified measure for different models are presented under both populations.

Table 3. *Simulation results for unified measure*

n	**SRR Model**	**Unified Measure**	
		Population I	**Population II**
50	Pollock and Bek [3]	9.85535	7.47312
	Eichhorn and Hayre [5]	0.02042	0.02042
	Saha [20]	0.00342	0.00235
	Hussian [9]	9.85535	7.47312
	Singh and Tarray [11]	39.4214	29.8925
	Proposed SRR model	**0.00308**	**0.00227**
100	Pollock and Bek [3]	4.92767	3.09094
	Eichhorn and Hayre [5]	0.01010	0.01010
	Saha [20]	0.00171	0.00119
	Hussian [9]	4.92767	3.09094
	Singh and Tarray [11]	19.7107	12.3637
	Proposed SRR model	**0.00154**	**0.00112**
200	Pollock and Bek [3]	2.61486	1.54547
	Eichhorn and Hayre [5]	0.00502	0.00502
	Saha [20]	0.00091	0.00059
	Hussian [9]	2.61486	1.54547
	Singh and Tarray [11]	10.4594	6.18188
	Proposed SRR model	**0.00085**	**0.00056**
300	Pollock and Bek [3]	1.74324	1.03031
	Eichhorn and Hayre [5]	0.00334	0.00334
	Saha [20]	0.00061	0.00039
	Hussian [9]	1.74324	1.03031
	Singh and Tarray [11]	6.97297	4.12125
	Proposed SRR model	**0.00057**	**0.00037**

Using the two models, the proposed SRR model and Pollock and Bek's model the PRE values for each estimator are shown in Tables 4-11.

5. CONCLUSION

This paper has proposed a new SRR model and a generalized regression-cum-exponential estimator using two non-sensitive auxiliary variables to obtain efficient results for estimating the population mean of sensitive variable and some class of estimators have also been shown in Table2. The unified measure is calculated for different models presented in Table 1.The unified measure of the proposed SRR model has been obtained using g=0.6. From Table 3, it has been noticed that the proposed model has minimum value unified measure as compared to existing models. As the unified measure decreases the privacy of the respondent increases.

Also, from the simulation and empirical study using two population datasets provided in Table (4-11), it has been observed that the MSE values for the proposed class of estimators are less than the MSEs of the existing estimators t_m, t_S, t_G, t_{K_1}, t_{K_2}, t_{reg}, and t_{er} when both models are used. It shows that the class of generalized estimators are more efficient than existing estimators. From both populations, we have noticed that as the sample size increases, there is a reduction in the *MSE* of the estimators discussed.

From Table (4-11) it is observed that using the proposed SRR model the generalized class of estimators including the existing estimators attains less MSE values than the MSE value using the Pollock and Bek's model therefore it can be concluded that the proposed SRR model improves the efficiency of the proposed class of estimators and the efficiency of the existing estimators too. Furthermore, we observed what can be noted from Table (4-11) is that as the value of *g* increases in the proposed SRR model, the MSE's of estimators also increases. From both populations, we have noticed that as the sample size increases, there is a reduction in the *MSE* of the estimators discussed. The same type of results is noted in Diana and Perri [8] model.

CONFLICTS OF INTEREST

No conflict of interest was declared by the authors.

REFERENCES

[1] Warner, S. L., "Randomized response: A survey technique for eliminating evasive answer bias", Journal of the American Statistical Association, 60(309): 63-69, (1965).

[2] Greenberg, B. G., Kuebler, R. R., Jr., Abernathy, J. R., Horvitz, D. G., "Application of the randomized response technique in obtaining quantitative data", Journal of the American Statistical Association, 66(334): 243-250, (1971).

[3] Pollock, K. H. and Bek, Y., "A comparison of three randomized response models for quantitative data", Journal of the American Statistical Association, 71(356):884-886, (1976).

[4] Himmelfarb, S., Edgell, S. E., "Additive constants model: A randomized response technique for eliminating evasiveness to quantitative response questions", Psychological Bulletin, 87(3): 525, (1980).

[5] Eichhorn, B.H., Hayre, L. S., "Scrambled randomized response models for obtaining sensitive quantitative data", Journal of Statistical Planning and Inference, 7(4): 307-316, (1983).

[6] Saha, A., "Optional randomized response in stratified unequal probability sampling-A simulation based numerical study with Kuk's method", Test, 16(2): 346-354, (2007).

[7] Diana, G., Perri, P. F., "New scrambled response models for estimating the mean of a sensitive quantitative character", Journal of Applied Statistics, 37(11): 1875-1890, (2010).

[8] Diana, G., Perri, P. F., "A class of estimators for quantitative sensitive data", Statistical Papers, 52(3): 633-650, (2011).

[9] Hussain, Z., “Improvement of the Gupta and Thornton scrambling model through double use of randomization device”, International Journal Academic Research Business and Social Science, 2:91-97, (2012).

[10] Chen, C. C., Singh, S., “The Franklin's randomized response model for two sensitive attributes”, Section Survey Research Methods, 4171-4185, (2009).

[11] Singh, H., Tarray, T., “An improved randomized response additive model”, Sri Lankan Journal of Applied Statistics, 15(2): 131-138, (2014).

[12] Hussain, Z., Al-Zahrani, B., “Mean and sensitivity estimation of a sensitive variable through additive scrambling”, Communications in Statistics-Theory and Methods, 45(1): 182-193, (2016).

[13] Bouza-Herrera, C. N., “Behavior of some scrambled randomized response models under simple random sampling, ranked set sampling and Rao–Hartley–Cochran designs”, Handbook of Statistics, 34, 209-220, Elsevier, (2016).

[14] Bouza-Herrera, C. N., “Data gathering, analysis and protection of privacy through randomized response techniques: qualitative and quantitative human traits”, Investigación Operacional, 39(1): 151-152, (2018).

[15] Bouza, C. N., Singh, P., Singh, R., “Ranked set sampling and optional scrambling randomized response modeling”, Investigación Operacional, 39(1): 100-107, (2018).

[16] Sousa, R., Shabbir, J., Real, P. C., Gupta, S., “Ratio estimation of the mean of a sensitive variable in the presence of auxiliary information”, Journal of Statistical Theory and Practice, 4(3): 495-507, (2010).

[17] Gupta, S., Shabbir, J., Sousa, R., Corte-real, P., “Estimation of the mean of a sensitive variable in the presence of auxiliary information”, Communications in Statistics-Theory and Methods, 41(13-14): 2394–2404, (2012).

[18] Tarray, T. A., Singh, H. P., “A general procedure for estimating the mean of a sensitive variable using auxiliary information”, Investigacion Operacionel, 36(3): 268-279, (2015).

[19] Koyuncu, N., Gupta, S., Sousa, R., “Exponential-type estimators of the mean of a sensitive variable in the presence of non-sensitive auxiliary information”, Communications in Statistics-Simulation and Computation, 43(7): 1583-1594, (2014).

[20] Saha, A., “A randomized response technique for quantitative data under unequal probability sampling”, Journal of Statistical Theory and Practice, 2(4): 589-596, (2008).

[21] 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(3): 506-511, (2018).

[22] 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(3): 400-408, (2008).

[23] Rao, T. J., “On certail methods of improving ration and regression estimators”, Communications in Statistics-Theory and Methods, 20(10): 3325-3340, (1991).

[24] Bahl, S., Tuteja, R. K., “Ratio and product type exponential estimators”, Journal of Information and Optimization Sciences”, 12(1): 159-164, (1991).

[25] Horvitz, D.G., Thompson, D.J., “A generalization of sampling without replacement from a finite universe”, J. Amer. Statist. Assoc., 47(260): 663-685, (1952).

[26] Sanaullah, A., Ali, H.A., Noor ul Amin, M., Hanif, M., “Generalized exponential chain ratio estimators under stratified two-phase random sampling”, Applied Mathematics and Computation, 226: 541-547, (2014).

[27] Asghar, A., Sanaullah, A., Hanif, M., “Generalized exponential type estimator for population variance in survey sampling”, Revista Colombiana de Estadística, 37(1): 213-224, (2014).

[28] Sanaullah, A., Hanif, M., Asghar. A., “Generalized exponential estimators for population variance under two-phase sampling”, Int. J. Appl. Comput. Math., 2(1): 75-84, (2016).

[29] Koyuncu, N., Kadilar, C., “Family of estimators of population mean using two auxiliary variables in stratified random sampling”, Communications in Statistics-Theory and Methods, 38(14): 2398–2417, (2009).

[30] Jabeen, R., Sanaullah, A., Hanif, M., “Generalized estimator for estimating population mean under two stage sampling”, Pak. J. Statist., 30(4): 465-486, (2014).

Table 4. *Simulation and empirical results at sample size n=50 for the MSEs and PREs of the estimators using population I [N=1000,* $\rho_{x1y} = 0.6817$ & $\rho_{x2y} = 0.6705$ *]*

	MSE Estimation using Population–I											
n=50	**Himmelfarb and Edgel Model**			**Proposed Model**								
				g=0.3 & a=0.5			g=0.6 & a=0.5			g=0.9 & a=0.5		
Estimators	Empirical	Theoretical	**PRE**	Empirical	Theoretical	**PRE**	Empirical	Theoretical	**PRE**	Empirical	Theoretical	**PRE**
t_m	0.2148	0.1904	100.00	0.0180	0.0199	100.00	0.0208	0.0693	100.00	0.0233	0.1540	100.00
t_S	0.1168	0.1125	169.18	0.0174	0.0166	119.74	0.0228	0.0425	163.00	0.0339	0.0911	169.17
t_G	0.0924	0.1017	187.25	0.0157	0.0156	127.81	0.0167	0.0386	179.59	0.0202	0.0823	187.27
t_{K_1}	0.0924	0.1016	187.33	0.0156	0.0156	127.89	0.0166	0.0386	179.68	0.0202	0.0822	187.34
t_{K_2}	3.4328	0.4666	40.80	0.1195	0.0486	40.91	0.1909	0.1713	40.48	0.3578	0.3809	40.44
t_1	0.0924	0.1017	187.27	0.0157	0.0156	127.81	0.0167	0.0386	179.59	0.0202	0.0823	187.27
t_2	0.1619	0.1429	133.22	0.0175	0.0194	102.84	0.0178	0.0535	129.68	0.0264	0.1157	133.17
Proposed												
t_g	**0.0729**	**0.0780**	**243.95***	**0.0157**	**0.0135**	**147.85***	**0.0167**	**0.0301**	**230.37***	**0.0202**	**0.0631**	**244.03***
t_{g1}	0.0749	0.0782	243.61	0.0587	0.0326	61.12	0.0983	0.0367	188.73	0. 1113	0.0638	241.50
t_{g2}	0.0862	0.0897	212.19	0.0233	0.0144	137.91	0.0511	0.0340	204.18	0.0545	0.0718	214.62
t_{g3}	0.0809	0.0813	234.14	0.0223	0.0166	119.59	0.0465	0.0304	227.87	0.0533	0.0646	238.47
t_{g4}	0.0984	0.0964	197.53	0.0454	0.0276	72.00	0.0617	0.0358	193.52	0.0638	0.0757	203.37
t_{g5}	0.0796	0.0863	220.68	0.0193	0.0167	119.23	0.0308	0.0431	160.99	0.0341	0.0923	166.88
t_{g6}	0.1097	0.0962	198.00	0.0363	0.0170	117.13	0.1161	0.0406	170.66	0.1381	0.0871	176.88
t_{g7}	0.0890	0.0979	194.56	0.0172	0.0315	63.15	0.0210	0.0663	104.60	0.0267	0.1232	125.05
t_{g8}	0.0862	0.0897	212.19	0.0233	0.0144	137.91	0.0511	0.0340	204.18	0.0545	0.0718	214.62
t_{g9}	0.0809	0.0813	234.14	0.0223	0.0166	119.59	0.0465	0.0304	227.87	0.0533	0.0646	238.47

***shows that most efficient estimator under each model**

Table 5. *Simulation and empirical results at sample size n=100 for the MSEs and PREs of the estimators using population I [N=1000,* $\rho_{x1y} = 0.6817$ & $\rho_{x2y} = 0.6705$ *]*

	MSE Estimation using Population–I											
n=100	**Himmelfarb and Edgel Model**			**Proposed Model**								
				g=0.3 & a=0.5			g=0.6 & a=0.5			g=0.9 & a=0.5		
Estimators	Empirical	Theoretical	**PRE**	Empirical	Theoretical	**PRE**	Empirical	Theoretical	**PRE**	Empirical	Theoretical	PRE
t_m	0.1010	0.0902	100.00	0.0056	0.0094	100.00	0.0065	0.0329	100.00	0.0107	0.0730	100.00
t_S	0.0574	0.0533	169.19	0.0054	0.0079	119.82	0.0089	0.0202	163.03	0.0186	0.0431	169.16
t_G	0.0504	0.0483	186.86	0.0051	0.0074	127.43	0.0059	0.0183	179.21	0.0102	0.0391	186.84
t_{K_1}	0.0504	0.0483	186.90	0.0051	0.0073	128.65	0.0059	0.0183	179.61	0.0102	0.0390	186.89
t_{K_2}	2.4119	0.2223	40.56	0.0488	0.0390	24.19	0.2220	0.1449	22.68	0.7442	0.3238	22.54
t_1	0.0504	0.0483	186.86	0.0051	0.0074	127.43	0.0059	0.0183	179.21	0.0102	0.0391	186.84
t_2	0.0775	0.0677	133.22	0.0056	0.0092	102.84	0.0071	0.0253	129.69	0.0124	0.0548	133.19
Proposed												
t_g	**0.0389**	**0.0370**	**243.53***	**0.0051**	**0.0064**	**147.57***	**0.0059**	**0.0143**	**230.04***	**0.0102**	**0.0300**	**243.61***
t_{g1}	0.0390	0.0371	243.20	0.0252	0.0154	61.11	0.0305	0.0174	188.58	0.0424	0.0303	241.19
t_{g2}	0.0441	0.0425	212.04	0.0087	0.0069	137.66	0.0117	0.0161	204.04	0.0254	0.0340	214.46
t_{g3}	0.0418	0.0386	233.63	0.0082	0.0079	119.52	0.0127	0.0144	227.49	0.0225	0.0307	238.04
t_{g4}	0.0494	0.0458	197.03	0.0193	0.0131	71.98	0.0195	0.0170	193.35	0.0239	0.0360	202.84
t_{g5}	0.0429	0.0409	220.44	0.0068	0.0079	118.92	0.0074	0.0204	160.71	0.0150	0.0438	166.58
t_{g6}	0.0544	0.0457	197.55	0.0138	0.0081	116.71	0.0299	0.0193	170.30	0.0621	0.0414	176.44
t_{g7}	0.0459	0.0465	194.10	0.0056	0.0150	62.99	0.0064	0.0315	104.39	0.0104	0.0585	124.76
t_{g8}	0.0441	0.0425	212.04	0.0087	0.0069	137.66	0.0117	0.0161	204.04	0.0254	0.0340	214.46
t_{g9}	0.0418	0.0386	233.63	0.0082	0.0079	119.52	0.0127	0.0144	227.49	0.0225	0.0307	238.04

Table 6. *Simulation and empirical results at sample size n=200 for the MSEs and PREs of the estimators using population I [N=1000,* $\rho_{x1y} = 0.6817$ & $\rho_{x2y} = 0.6705$ *]*

	MSE Estimation using Population–I											
n=200	**Himmelfarb and Edgel Model**			**Proposed Model**								
				g=0.3 & a=0.5			g=0.6 & a=0.5			g=0.9 & a=0.5		
Estimators	Empirical	Theoretical	**PRE**	Empirical	Theoretical	**PRE**	Empirical	Theoretical	**PRE**	Empirical	Theoretical	**PRE**
t_m	0.0399	0.0401	100.00	0.0030	0.0042	100.00	0.0033	0.0146	100.00	0.0039	0.0324	100.00
t_S	0.0241	0.0237	169.19	0.0029	0.0035	119.71	0.0045	0.0090	162.95	0.0075	0.0192	169.17
t_G	0.0227	0.0215	186.68	0.0028	0.0033	126.97	0.0031	0.0082	178.92	0.0037	0.0174	186.70
t_{K_1}	0.0226	0.0214	187.46	0.0027	0.0033	128.92	0.0030	0.0081	180.47	0.0036	0.0174	186.81
t_{K_2}	1.7998	0.0991	40.45	0.0965	0.0341	12.28	0.2769	0.1315	11.10	0.8645	0.2950	10.99
t_1	0.0227	0.0215	186.68	0.0028	0.0033	126.97	0.0031	0.0082	178.92	0.0037	0.0174	186.70
t_2	0.0301	0.0301	133.20	0.0030	0.0041	102.95	0.0036	0.0113	129.66	0.0050	0.0244	133.18
Proposed												
t_g	**0.0178**	**0.0165**	**243.35***	**0.0028**	**0.0029**	**147.02***	**0.0031**	**0.0064**	**229.56***	**0.0037**	**0.0133**	**243.47***
t_{g1}	0.0179	0.0165	243.06	0.0114	0.0069	61.08	0.0149	0.0077	188.63	0.0211	0.0135	240.94
t_{g2}	0.0203	0.0189	211.95	0.0040	0.0031	137.38	0.0063	0.0072	203.91	0.0119	0.0151	214.34
t_{g3}	0.0181	0.0172	233.43	0.0042	0.0035	119.37	0.0062	0.0064	227.41	0.0111	0.0136	237.76
t_{g4}	0.0212	0.0204	196.76	0.0086	0.0058	71.99	0.0093	0.0076	193.12	0.0104	0.0160	202.56
t_{g5}	0.0196	0.0182	220.22	0.0032	0.0035	118.70	0.0040	0.0091	160.62	0.0061	0.0195	166.39
t_{g6}	0.0223	0.0203	197.34	0.0071	0.0036	116.39	0.0153	0.0086	170.16	0.0331	0.0184	176.25
t_{g7}	0.0196	0.0207	193.90	0.0030	0.0067	62.91	0.0032	0.0140	104.29	0.0038	0.0260	124.63
t_{g8}	0.0203	0.0189	211.95	0.0040	0.0031	137.38	0.0063	0.0072	203.91	0.0119	0.0151	214.34
t_{g9}	0.0181	0.0172	233.43	0.0042	0.0035	119.37	0.0062	0.0064	227.41	0.0111	0.0136	237.76

Table 7. *Simulation and empirical results at sample size n=300 for the MSEs and PREs of the estimators using population I [N=1000,* $\rho_{x1y} = 0.6817 \, \& \, \rho_{x2y} = 0.6705$ *]*

	MSE Estimation using Population–I											
n=300	**Himmelfarb and Edgel Model**			**Proposed Model**								
				g=0.3 & a=0.5			g=0.6 & a=0.5			g=0.9 & a=0.5		
Estimators	Empirical	Theoretical	**PRE**	Empirical	Theoretical	**PRE**	Empirical	Theoretical	**PRE**	Empirical	Theoretical	**PRE**
t_m	0.0247	0.0234	100.00	0.0016	0.0024	100.00	0.0022	0.0085	100.00	0.0030	0.0189	100.00
t_S	0.0149	0.0138	169.18	0.0015	0.0020	119.61	0.0030	0.0052	163.22	0.0053	0.0112	169.23
t_G	0.0131	0.0125	186.59	0.0015	0.0019	127.08	0.0020	0.0048	178.99	0.0028	0.0101	186.59
t_{K_1}	0.0130	0.0125	187.04	0.0014	0.0019	129.79	0.0020	0.0047	182.83	0.0028	0.0100	189.20
t_{K_2}	2.0667	0.0579	40.41	0.0583	0.0325	7.51	0.1720	0.1271	6.70	0.4629	0.2854	6.63
t_1	0.0130	0.0125	186.59	0.0015	0.0019	127.08	0.0020	0.0048	178.99	0.0028	0.0101	186.59
t_2	0.0189	0.0176	133.22	0.0016	0.0024	102.52	0.0024	0.0066	129.68	0.0035	0.0142	133.24
Proposed												
t_g	**0.0102**	**0.0096**	**243.29***	**0.0015**	**0.0017**	**146.99***	**0.0020**	**0.0037**	**229.65**	**0.0028**	**0.0078**	**243.50***
t_{g1}	0.0102	0.0096	243.04	0.0067	0.0040	61.00	0.0102	0.0045	188.50	0.0139	0.0079	241.02
t_{g2}	0.0116	0.0110	211.97	0.0023	0.0018	137.08	0.0044	0.0042	203.83	0.0073	0.0088	214.27
t_{g3}	0.0109	0.0100	233.33	0.0022	0.0021	119.02	0.0041	0.0038	227.20	0.0068	0.0080	237.69
t_{g4}	0.0129	0.0119	196.64	0.0052	0.0034	71.76	0.0065	0.0044	193.20	0.0073	0.0093	202.57
t_{g5}	0.0112	0.0106	220.15	0.0018	0.0021	118.45	0.0029	0.0053	160.45	0.0041	0.0114	166.40
t_{g6}	0.0137	0.0119	197.30	0.0037	0.0021	116.19	0.0099	0.0050	170.06	0.0192	0.0107	176.16
t_{g7}	0.0119	0.0121	193.70	0.0016	0.0039	62.72	0.0022	0.0082	104.28	0.0029	0.0152	124.56
t_{g8}	0.0116	0.0110	211.97	0.0023	0.0018	137.08	0.0044	0.0042	203.83	0.0073	0.0088	214.27
t_{g9}	0.0109	0.0100	233.33	0.0022	0.0021	119.02	0.0041	0.0038	227.20	0.0068	0.0080	237.69

Table 8. *Simulation and empirical results at sample size n=50 for the MSEs and PREs of the estimators using population II [N=1000,* $\rho_{x1y} = 0.8706 \& \rho_{x2y} = 0.8428$*]*

	MSE Estimation using Population–II											
n=50	**Himmelfarb and Edgel Model**			**Proposed Model**								
				g=0.3 & a=0.5			g=0.6 & a=0.5			g=0.9 & a=0.5		
Estimators	Empirical	Theoretical	**PRE**	Empirical	Theoretical	**PRE**	Empirical	Theoretical	**PRE**	Empirical	Theoretical	**PRE**
t_m	0.0734	0.1144	100.00	0.0097	0.0123	100.00	0.0119	0.0417	100.00	0.0135	0.0925	100.00
t_S	0.0328	0.0378	302.75	0.0083	0.0092	134.64	0.0166	0.0154	271.39	0.0284	0.0305	303.25
t_G	0.0292	0.0279	409.38	0.0080	0.0082	149.70	0.0102	0.0118	352.53	0.0122	0.0225	410.57
t_{K_1}	0.0292	0.0279	409.52	0.0079	0.0082	150.43	0.0102	0.0118	354.63	0.0122	0.0225	410.75
t_{K_2}	1.7912	0.4621	24.75	0.0858	2.7010	0.46	0.1479	10.8033	0.39	0.3636	24.3072	0.38
t_1	0.0292	0.0279	409.38	0.0080	0.0082	149.70	0.0102	0.0118	352.53	0.0122	0.0225	410.57
t_2	0.0445	0.0684	167.20	0.0095	0.0119	103.53	0.0130	0.0264	158.05	0.0168	0.0553	167.20
Proposed												
t_g	**0.0088**	**0.0063**	**1809.81***	**0.0076**	**0.0063**	**195.56***	**0.0102**	**0.0041**	**1028.08***	**0.0122**	**0.0050**	**1845.51***
t_{g1}	0.0089	0.0066	1725.19	0.0503	0.0256	48.07	0.0723	0.0110	378.77	0.0841	0.0059	1559.19
t_{g2}	0.0134	0.0181	633.33	0.0143	0.0629	19.58	0.0345	0.0372	112.17	0.0408	0.0210	439.45
t_{g3}	0.0093	0.0069	1667.35	0.0189	0.0127	97.08	0.0304	0.0070	592.90	0.0414	0.0052	1774.66
t_{g4}	0.0220	0.0228	501.45	0.0342	0.0205	60.07	0.0423	0.0093	446.90	0.0414	0.0162	570.04
t_{g5}	0.0150	0.0135	846.01	0.0101	0.0584	21.10	0.0205	0.0380	109.78	0.0205	0.0370	249.96
t_{g6}	0.0304	0.0269	425.20	0.0302	0.0101	121.62	0.0700	0.0156	268.42	0.1156	0.0310	297.97
t_{g7}	0.0108	0.0134	856.78	0.0105	0.0292	42.15	0.0127	0.0468	89.15	0.0139	0.0674	137.26
t_{g8}	0.0134	0.0181	633.33	0.0143	0.0629	19.58	0.0345	0.0372	112.17	0.0408	0.0210	439.45
t_{g9}	0.0093	0.0069	1667.35	0.0189	0.0127	97.08	0.0304	0.0070	592.90	0.0414	0.0052	1774.66

Table 9. *Simulation and empirical results at sample size n=100 for the MSEs and PREs of the estimators using population II [N=1000,* $\rho_{x1y}=0.8706 \& \rho_{x2y}=0.8428$ *]*

	MSE Estimation using Population–II											
n=100	**Himmelfarb and Edgel Model**			**Proposed Model**								
				g=0.3 & a=0.5			g=0.6 & a=0.5			g=0.9 & a=0.5		
Estimators	Empirical	Theoretical	**PRE**	Empirical	Theoretical	**PRE**	Empirical	Theoretical	**PRE**	Empirical	Theoretical	**PRE**
t_m	0.0734	0.1144	100.00	0.0097	0.0123	100.00	0.0119	0.0417	100.00	0.0135	0.0925	100.00
t_S	0.0328	0.0378	302.75	0.0083	0.0092	134.64	0.0166	0.0154	271.39	0.0284	0.0305	303.25
t_G	0.0292	0.0279	409.38	0.0080	0.0082	149.70	0.0102	0.0118	352.53	0.0122	0.0225	410.57
t_{K_1}	0.0292	0.0279	409.52	0.0079	0.0082	150.43	0.0102	0.0118	354.63	0.0122	0.0225	410.75
t_{K_2}	1.7912	0.4621	24.75	0.0858	2.7010	0.46	0.1479	10.8033	0.39	0.3636	24.3072	0.38
t_1	0.0292	0.0279	409.38	0.0080	0.0082	149.70	0.0102	0.0118	352.53	0.0122	0.0225	410.57
t_2	0.0445	0.0684	167.20	0.0095	0.0119	103.53	0.0130	0.0264	158.05	0.0168	0.0553	167.20
Proposed												
t_g	**0.0088**	**0.0063**	**1809.81***	**0.0076**	**0.0063**	**195.56***	**0.0102**	**0.0041**	**1028.08***	**0.0122**	**0.0050**	**1845.51***
t_{g1}	0.0089	0.0066	1725.19	0.0503	0.0256	48.07	0.0723	0.0110	378.77	0.0841	0.0059	1559.19
t_{g2}	0.0134	0.0181	633.33	0.0143	0.0629	19.58	0.0345	0.0372	112.17	0.0408	0.0210	439.45
t_{g3}	0.0093	0.0069	1667.35	0.0189	0.0127	97.08	0.0304	0.0070	592.90	0.0414	0.0052	1774.66
t_{g4}	0.0220	0.0228	501.45	0.0342	0.0205	60.07	0.0423	0.0093	446.90	0.0414	0.0162	570.04
t_{g5}	0.0150	0.0135	846.01	0.0101	0.0584	21.10	0.0205	0.0380	109.78	0.0205	0.0370	249.96
t_{g6}	0.0304	0.0269	425.20	0.0302	0.0101	121.62	0.0700	0.0156	268.42	0.1156	0.0310	297.97
t_{g7}	0.0108	0.0134	856.78	0.0105	0.0292	42.15	0.0127	0.0468	89.15	0.0139	0.0674	137.26
t_{g8}	0.0134	0.0181	633.33	0.0143	0.0629	19.58	0.0345	0.0372	112.17	0.0408	0.0210	439.45
t_{g9}	0.0093	0.0069	1667.35	0.0189	0.0127	97.08	0.0304	0.0070	592.90	0.0414	0.0052	1774.66

Table 10. *Simulation and empirical results at sample size n=200 for the MSEs and PREs of the estimators using population II* *[N=1000,* $\rho_{x1y} = 0.8706 \& \rho_{x2y} = 0.8428$ *]*

	MSE Estimation using Population–II											
n=200	**Himmelfarb and Edgel Model**			**Proposed Model**								
				g=0.3 & a=0.5			g=0.6 & a=0.5			g=0.9 & a=0.5		
Estimators	Empirical	Theoretical	**PRE**	Empirical	Theoretical	**PRE**	Empirical	Theoretical	**PRE**	Empirical	Theoretical	**PRE**
t_m	0.0305	0.0241	100.00	0.0016	0.0026	100.00	0.0023	0.0088	100.00	0.0035	0.0195	100.00
t_S	0.0104	0.0080	302.89	0.0014	0.0019	134.20	0.0033	0.0032	271.30	0.0076	0.0064	303.27
t_G	0.0075	0.0059	408.83	0.0013	0.0017	148.85	0.0022	0.0025	353.01	0.0034	0.0048	409.89
t_{K_1}	0.0074	0.0058	415.17	0.0013	0.0017	151.46	0.0020	0.0025	353.01	0.0031	0.0048	409.89
t_{K_2}	2.7479	0.0984	24.47	0.0861	2.3316	0.11	0.0997	9.3264	0.09	0.4852	20.9843	0.09
t_1	0.0075	0.0059	408.83	0.0013	0.0017	148.85	0.0020	0.0025	353.01	0.0031	0.0048	409.89
t_2	0.0174	0.0144	167.22	0.0016	0.0025	103.19	0.0025	0.0056	158.09	0.0044	0.0116	167.27
Proposed												
t_g	**0.0017**	**0.0013**	**1810.53***	**0.0013**	**0.0013**	**194.74***	**0.0020**	**0.0009**	**1034.12***	**0.0031**	**0.0011**	**1854.29***
t_{g1}	0.0018	0.0014	1720.00	0.0086	0.0054	47.96	0.0141	0.0023	378.88	0.0242	0.0013	1557.60
t_{g2}	0.0059	0.0038	633.68	0.0021	0.0133	19.55	0.0054	0.0078	112.12	0.0115	0.0044	439.50
t_{g3}	0.0017	0.0015	1660.69	0.0028	0.0027	97.00	0.0057	0.0015	593.92	0.0118	0.0011	1770.00
t_{g4}	0.0063	0.0048	500.62	0.0067	0.0043	59.95	0.0086	0.0020	446.19	0.0114	0.0034	569.30
t_{g5}	0.0041	0.0029	844.91	0.0016	0.0123	21.06	0.0032	0.0080	109.88	0.0055	0.0078	249.94
t_{g6}	0.0092	0.0057	424.69	0.0045	0.0021	121.03	0.0130	0.0033	267.99	0.0349	0.0065	297.71
t_{g7}	0.0018	0.0028	856.94	0.0017	0.0062	42.05	0.0025	0.0099	89.15	0.0037	0.0142	137.21
t_{g8}	0.0059	0.0038	633.68	0.0021	0.0133	19.55	0.0054	0.0078	112.12	0.0115	0.0044	439.50
t_{g9}	0.0017	0.0015	1660.69	0.0028	0.0027	97.00	0.0057	0.0015	593.92	0.0118	0.0011	1770.00

Table 11. *Simulation and empirical results at sample size n=300 for the MSEs and PREs of the estimators using population II [N=1000,* $\rho_{x1y} = 0.8706 \& \rho_{x2y} = 0.8428$ *]*

	MSE Estimation using Population–II											
n=300	**Himmelfarb and Edgel Model**			**Proposed Model**								
				g=0.3 & a=0.5			g=0.6 & a=0.5			g=0.9 & a=0.5		
Estimators	Empirical	Theoretical	PRE	Empirical	Theoretical	PRE	Empirical	Theoretical	PRE	Empirical	Theoretical	PRE
t_m	0.0161	0.0141	100.00	0.0011	0.0015	100.00	0.0015	0.0051	100.00	0.0022	0.0114	100.00
t_S	0.0052	0.0046	302.80	0.0010	0.0011	134.82	0.0022	0.0019	271.43	0.0047	0.0037	303.48
t_G	0.0037	0.0034	409.62	0.0009	0.0010	149.50	0.0014	0.0015	351.37	0.0020	0.0028	409.75
t_{K_1}	0.0036	0.0034	413.24	0.0009	0.0010	149.50	0.0013	0.0014	358.74	0.0019	0.0028	412.73
t_{K_2}	2.1126	0.0575	24.44	0.0760	2.2969	0.07	0.2156	9.1875	0.06	0.4797	20.6718	0.05
t_1	0.0037	0.0034	409.62	0.0009	0.0010	149.50	0.0014	0.0015	351.37	0.0020	0.0028	409.75
t_2	0.0096	0.0084	167.26	0.0011	0.0015	103.42	0.0017	0.0032	158.33	0.0027	0.0068	167.16
Proposed												
t_g	**0.0010**	**0.0008**	**1801.28***	**0.0009**	**0.0008**	**193.59***	**0.0014**	**0.0005**	**1026.00***	**0.0020**	**0.0006**	**1830.65***
t_{g1}	0.0010	0.0008	1734.57	0.0060	0.0066	22.78	0.0093	0.0014	380.00	0.0141	0.0007	1554.79
t_{g2}	0.0028	0.0022	632.88	0.0017	0.0181	8.36	0.0037	0.0046	112.25	0.0064	0.0026	438.22
t_{g3}	0.0010	0.0008	1672.62	0.0021	0.0069	22.01	0.0038	0.0009	589.66	0.0068	0.0006	1773.44
t_{g4}	0.0032	0.0028	500.00	0.0045	0.0228	6.62	0.0055	0.0012	446.09	0.0069	0.0020	567.50
t_{g5}	0.0020	0.0017	846.39	0.0012	0.0135	11.17	0.0022	0.0047	109.85	0.0032	0.0045	250.00
t_{g6}	0.0048	0.0033	424.47	0.0033	0.0269	5.61	0.0094	0.0019	268.59	0.0200	0.0038	297.12
t_{g7}	0.0011	0.0016	856.71	0.0012	0.0134	11.31	0.0016	0.0058	89.22	0.0023	0.0083	137.08
t_{g8}	0.0028	0.0022	632.88	0.0017	0.0181	8.36	0.0037	0.0046	112.25	0.0064	0.0026	438.22
t_{g9}	0.0010	0.0008	1672.62	0.0021	0.0069	22.01	0.0038	0.0009	589.66	0.0068	0.0006	1773.44