Risk Management and Healthcare Policy

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Study to Alter the Nuisance Effect of Non-Response Using Scrambled Mechanism

Chandraketu Singh[1]
Mustafa Kamal[2]
Garib Nath Singh[1]
Jong-Min Kim[3]

[1]Department of Mathematics & Computing, Indian Institute of Technology (Indian School of Mines), Dhanbad, 826004, India; [2]Department of Basic Sciences, College of Science and Theoretical Studies, Saudi Electronic University, Dammam, Kingdom of Saudi Arabia; [3]Statistics Discipline, Division of Sciences and Mathematics, University of Minnesota-Morris, Morris, MN, 56267-2134, USA

Introduction: In biometric sample surveys, our objective is to get ready-made information for future planning and policy implementations related to the subject matters of highly sensitive issues. In such situations, we apply randomized response/scrambled response techniques. There are many highly sensitive issues which need to be examined over time as they may have a tendency to change. To get rid of these types of practical cases we need a scrambled response technique on successive occasions.

Methods: Using an additive and multiplicative technique, we proposed new effective scrambled response models to estimate the population mean of quantitative sensitive characteristics. Degree of privacy protection and unified measure approaches are used to examine the efficacy of the proposed models. Efficiency of the proposed models has been checked using MATLAB software. The utility of the proposed models under two occasions of successive sampling has been also explored using exponential-type estimators. Empirical and simulation studies are carried out to justify the proposition of the proposed estimators using MATLAB software.

Results: The percent relative efficiencies of the proposed models are always greater than 100 with respect to the well-known Bar-Lev et al model. In terms of degree of privacy protection, most of the values are greater than 0.5 and closer to 1. Similarly, the values of the proposed models are smaller with respect to the Bar-Lev et al model in terms of a unified measure approach. When the proposed scrambled response models are used on successive occasions, the percent relative efficiency is always found greater than 100 for all cases over its competitors.

Discussion: In this study, after deeply examining the properties of the proposed models, we found that the proposed models performed better over the well-known existing model. The proposed models may be used in human survey when we deal with highly sensitive issues. The proposed models also performed better when we utilized them in successive sampling. Hence, if sensitive characteristics change with time, the proposed estimators may be the best alternative to deal with these types of situations.

Mathematics Subject Classification: 62D05.

Keywords: scrambled response model, privacy protection, successive sampling, mean square error, Monte Carlo simulation

Background

In social surveys, obtaining reliable data through direct questioning may be difficult when potentially sensitive questions like sexual indulgence during teenage years, voluntary prostitution, negligence of government rules, drug intake etc. are included in the survey. To avoid these situations, Warner[24] introduced the data collection technique that protects anonymity of the respondent known as the randomized

Correspondence: Chandraketu Singh
Department of Mathematics & Computing, Indian Institute of Technology (Indian School of Mines), Dhanbad, 826004, India
Email Chandraketu.lko@gmail.com

response technique (RRT). Whereas, estimation of the mean of quantitative sensitive variables utilizing randomized response models was initiated by Greenberg et al.[7] Later, Pollock and Bek[15] introduced an additive technique for mean estimation of quantitative sensitive variables. Eichhorn and Hayre[6] investigated the Pollock and Bek[15] models in depth and introduced the scrambled response method for estimating the population parameters of quantitative sensitive characteristics. Some other developments for different practical situations to estimate the population mean of positive quantitative sensitive characteristics are by Bar-Lev et al.,[3] Singh and Tarray,[23] Kim and Elam,[10] Diana and Perri,[5] Diana and Perri,[4] Son and Kim,[17] among others.

In real-life problems, there are well-documented situations where sensitive studies need to be monitored over time to understand the problem in a better way. To handle such situations, Jessen,[9] Priyanka et al.[14] and Singh and Sharma[18] among others used successive sampling. Arnab and Singh[1] and Priyanka and Trisandhya[12] among others used a scrambled response technique to handle sensitive issues on successive occasions.

Motivated by the above cited works, in the present article an attempt has been made to propose efficient scrambled response models. The degree of privacy protection and unified measures for the proposed models have been discussed. Efficacy of the proposed strategy has been justified.

Application of the proposed models under successive sampling has also been discussed. To estimate the population mean of sensitive variables, exponential types of estimators have been proposed. Properties of the proposed estimators have been discussed under the proposed scrambled response models. Empirical studies using a real data set and Monte Carlo simulation studies have been performed under the proposed scrambled response models. Empirical and simulation results indicate the dominance of the proposed estimators over some well-known estimators.

Proposed Technique

The scrambling technique has been used in many randomization devices, with the goal of increasing the respondent's cooperation. Pollock and Bek[15] have considered the additive model in which the respondent is asked to sum his sensitive attribute by a random value from a known distribution.

Eichhorn and Hayre[6] have considered another survey model involving a quantitative response variable and proposed an RR technique for it. Such models are very useful in studies involving a measured response variable which is highly sensitive in its nature. The model considered by Eichhorn and Hayre[6] is known as a multiplicative model in which the respondent multiplies his answers to the sensitive question by a random number from a known distribution. Motivated by Pollock and Bek[15] and Eichhorn and Hayre[6] we have proposed these models. The proposed scrambled response models are very useful to deal with highly sensitive issues. For example, studies addressing issues such as: i) not only whether or not a woman had an abortion, but in addition, how many abortions she underwent; ii) not only whether the subject used illicit drugs, but also the number of occasions on which drugs were taken; iii) not only if an individual cheated on his income tax report, but also the amount of under-reporting etc. The proposed models will encourage researchers to think more on these lines. The key issue when choosing a model is to find the right trade-off between privacy protection and efficiency in the estimates. In Measure of Privacy Protection we have tested the privacy level of our models, in Efficiency Comparison the efficiency of the models and in Unified Measure of Models Quality we use a unified measure of model quality. The proposed models protect the privacy of the respondents and perform better in terms of unified measure of model quality. Any researcher may think on this line and produce another model by using additive and multiplicative models but they have to test their efficiency level, privacy level and unified measure of model quality. In many cases, it may be possible that a model may be efficient but it is not necessary that the model also protects the respondent's privacy and performs better in terms of unified measure.

Suppose $\Omega=(1,2,\ldots,N)$ be the finite population and $Y \geq 0$ be a quantitative sensitive variable of interest with unknown mean and variance (μ_Y, σ_Y^2) respectively. Let W_1, W_2, S, and U be the scrambling variables independent of Y with means $\mu_{W_1}, \mu_{W_2}, \mu_S, \mu_U$ and variances $\sigma_{W_1}^2, \sigma_{W_2}^2, \sigma_S^2, \sigma_U^2$ respectively. A random sample of size n respondents was drawn from the population under a simple random sampling with replacement scheme. Each selected respondent in the sample was provided with a random device having three types of cards bearing statements: i) green cards with the statement: report sensitive variable Y; ii) red cards with the statement: report the scrambled response YW_1+W_2; iii) yellow card: report scramble value Z=YS (for model 1) and $Z=W_3(Y+U)$ (for model 2) with probabilities P_1, P_2 and P_3 respectively, such that $\sum_{i=1}^{3} P_i = 1$.

Model 1

In our models, the hypothesis is to develop whether our models are effective or not for estimating the population mean of quantitative sensitive characteristics and in protecting the privacy of respondents.

Let the response be α_1 whose distribution is as follows:

$$\alpha_1 = \begin{cases} Y & \text{with prob. } P_1 \\ Y\,W_1 + W_2 & \text{with prob. } P_2 \\ Y\,S & \text{with prob. } P_3 \end{cases}$$

Hence, for the observed reported response $\alpha_1 = Y\,P_1 + (Y\,W_1 + W_2)P_2 + Y\,S\,P_3$, the mean and variance of quantitative sensitive variable Y are as follows:

$$\hat{\mu}_{Y_1} = \frac{\bar{\alpha}_1 - P_2\mu_{W_2}}{\{P_1 + P_2\mu_{W_1} + P_3\mu_S\}} \text{ and } \quad V(\hat{\mu}_{Y_1}) = \frac{\sigma^2_{\alpha_1}}{n\{P_1 + P_2\mu_{W_1} + P_3\mu_S\}^2} \tag{1}$$

where

$$\sigma^2_{\alpha_1} = [P_1(\sigma^2_Y + \mu^2_Y) + P_2\{(\sigma^2_Y + \mu^2_Y)(\sigma^2_{W_1} + \mu^2_{W_1}) + (\sigma^2_{W_2} + \mu^2_{W_2})\} + P_3(\sigma^2_Y + \mu^2_Y)(\sigma^2_S + \mu^2_S)] - [P_1\mu_Y + P_2\{\mu_Y\mu_{W_1} + \mu_{W_2}\} + P_3\mu_Y\mu_S]^2$$

Model 2

Let the response be α_2 whose distribution is as follows:

$$\alpha_2 = \begin{cases} Y & \text{with prob. } P_1 \\ Y\,W_1 + W_2 & \text{with prob. } P_2 \\ W_3(Y + U) & \text{with prob. } P_3 \end{cases}$$

Hence, for the observed reported response $\alpha_2 = Y\,P_1 + (Y\,W_1 + W_2)P_2 + \{W_3(Y + U)\}\,P_3$, the mean and variance of quantitative sensitive variable Y are given by

$$\hat{\mu}_{Y_2} = \frac{\bar{\alpha}_2 - P_2\mu_{W_2} - P_3\mu_{W_3}\mu_U}{\{P_1 + P_2\mu_{W_1} + P_3\mu_{W_3}\}} \text{ and } \quad V(\hat{\mu}_{Y_2}) = \frac{\sigma^2_{\alpha_2}}{n\{P_1 + P_2\mu_{W_1} + P_3\mu_{W_3}\}^2} \tag{2}$$

where

$$\sigma^2_{\alpha_2} = \Big[P_1(\sigma^2_Y + \mu^2_Y) + P_2\Big\{(\sigma^2_Y + \mu^2_Y)(\sigma^2_{W_1} + \mu^2_{W_1}) + (\sigma^2_{W_2} + \mu^2_{W_2})\Big\} + P_3\Big\{(\sigma^2_Y + \mu^2_Y)(\sigma^2_{W_3} + \mu^2_{W_3}) + (\sigma^2_U + \mu^2_U)(\sigma^2_{W_3} + \mu^2_{W_3})\Big\}\Big] - [P_1\mu_Y + P_2\{\mu_Y\mu_{W_1} + \mu_{W_2}\} + P_3\{\mu_Y\mu_{W_3} + \mu_S\mu_{W_3}\}]^2$$

Measure of Privacy Protection

Following the work of Diana and Perri[5] the square of correlation coefficient between observed response and quantitative sensitive variables from models 1 and 2 are denoted by $\rho^2_{Y\theta_{D_i}}$ $(i = 1, 2)$ and given as

$$\rho^2_{Y\theta_{D_1}} = \frac{\{P_1 + P_2\mu_{W_1} + P_3\,\mu_S\}^2}{\sigma^2_{\theta_{D_1}}} \tag{3}$$

and

$$\rho^2_{Y\theta_{D_2}} = \frac{\{P_1 + P_2\mu_{W_1} + P_3\,\mu_{W_3}\}^2}{\sigma^2_{\theta_{D_2}}} \tag{4}$$

Efficiency Comparison

To show the performances of the proposed models, we have compared with the Bar-Lev et al[3] model in terms of percent relative efficiencies (PREs) using the formula:

$$PRE = \frac{V(\hat{\mu}_{YBL})}{V(\hat{\mu}_{Y_i})} x100 \quad \text{for } i = 1, 2$$

We have considered that scrambling variables follow a Poisson distribution. The data sets used for empirical comparison are given in Table 1.

From Tables 2 and 3, it may be seen that when we decrease the value of P_1 (probability of reporting the sensitive variable), the values of percent relative efficiencies increase. The values of percent relative efficiencies are high when the probability of selecting the third statement is also high (probability of reporting scrambled value). The level of privacy protection is closer to 1 in maximum cases. Hence, our proposed models perform better to deal with highly sensitive issues than the competitor when the probability of reporting scrambling variables is high.

Table 1 Data Set Used for Efficiency Comparison

Set	Y	W_1	W_2	W_3	S	U
A	4	3	5	4	2	3
B	7	3	5	4	2	5
C	7	4	8	5	2	1
D	5	2	8	5	3	1

http://doi.org/10.2147/RMHP.S294731

Table 2 PRE of the Suggested Models 1, 2 for Data Set A with Respect to the Bar Lev et al[3] Model Along with PP and (ψ_D)

	P_1	P_2	Model 1			Model 2		
			PRE	PP	ψ_D	PRE	PP	ψ_D
Set A	0.6	0.1	621.0964	0.9313	0.0791	734.0638	0.9188	0.0679
		0.2	507.3141	0.9274	0.0973	587.3111	0.9159	0.0851
		0.3	438.5585	0.9200	0.1134	480.0347	0.9124	0.1045
		0.4	399.4805	0.9084	0.1261	399.4805	0.9084	0.1261
	0.5	0.1	718.2527	0.9315	0.1028	983.3566	0.9063	0.0771
		0.2	607.4803	0.9262	0.1222	808.0478	0.9018	0.0943
		0.3	539.7898	0.9176	0.1388	675.9839	0.8968	0.1134
		0.4	503.2854	0.9046	0.1510	574.3161	0.8911	0.1343
	0.4	0.1	797.1256	0.9307	0.1295	1334.7000	0.8840	0.0814
		0.2	693.5569	0.9242	0.1499	1121.7000	0.8775	0.0976
		0.3	630.8819	0.9144	0.1666	957.7873	0.8701	0.1153
		0.4	600.5440	0.9000	0.1778	829.4807	0.8619	0.1344
	0.3	0.1	858.0844	0.9291	0.1594	1912.4000	0.8420	0.0789
		0.2	764.6166	0.9216	0.1803	1641.9000	0.8316	0.0931
		0.3	709.9589	0.9106	0.1966	1431.8000	0.8197	0.1083
		0.4	688.7410	0.8946	0.2063	1266.1000	0.8062	0.1245
	0.2	0.1	901.4702	0.9269	0.1924	3155.5000	0.7442	0.0684
		0.2	820.0522	0.9184	0.2134	2782.7000	0.7230	0.0799
		0.3	775.6281	0.9062	0.2287	2498.1000	0.6977	0.0922
		0.4	765.8839	0.8884	0.2362	2280.8000	0.6676	0.1056

Table 3 PRE of the Suggested Models 1, 2 for Data Set C with Respect to the Bar Lev et al[3] Model Along with PP and (ψ_D)

	P_1	P_1	Model 1			Model 2		
			PRE	PP	ψ_D	PRE	PP	ψ_D
Set C	0.6	0.1	901.9565	0.9766	0.2873	1004.2000	0.9740	0.2587
		0.2	567.8047	0.9764	0.4564	608.1222	0.9748	0.4269
		0.3	405.9730	0.9743	0.6398	420.5453	0.9734	0.6182
		0.4	317.9679	0.9706	0.8200	317.9679	0.9706	0.8200
	0.5	0.1	1105.3000	0.9760	0.3563	1276.8000	0.9723	0.3096
		0.2	733.2070	0.9751	0.5375	815.4293	0.9723	0.4847
		0.3	543.3190	0.9725	0.7273	585.0477	0.9704	0.6769
		0.4	437.8670	0.9681	0.9066	455.8927	0.9668	0.8719
	0.4	0.1	1273.0000	0.9751	0.4321	1529.3000	0.9701	0.3615
		0.2	883.0640	0.9737	0.6238	1023.2000	0.9695	0.5407
		0.3	675.8570	0.9705	0.8177	759.8299	0.9668	0.7301
		0.4	559.5810	0.9654	0.9928	610.0249	0.9623	0.9137
	0.3	0.1	1398.5000	0.9740	0.5144	1753.2000	0.9674	0.4131
		0.2	1008.7000	0.9720	0.7146	1223.0000	0.9661	0.5930
		0.3	795.2860	0.9682	0.9100	938.6643	0.9625	0.7755
		0.4	675.7270	0.9623	1.0775	776.8682	0.9567	0.9428
	0.2	0.1	1476.9000	0.9727	0.6027	1940.3000	0.9641	0.4628
		0.2	1102.9000	0.9702	0.8092	1406.8000	0.9619	0.6398
		0.3	893.9090	0.9657	1.0030	1115.5000	0.9573	0.8109
		0.4	778.9050	0.9589	1.1593	953.2830	0.9497	0.9564

In Table 4, for model 1, the values of percent relative efficiencies increase for decreasing values of P_1 (probability of reporting sensitive variable). The values of percent relative efficiencies are high when the probability of selecting the third statement is also high (probability of reporting scrambled value). For model 2, the values of PREs follow the same trend except for the point P_1=0.2. We get the highest value of percent relative efficiency when the probability of selecting the first statement is 0.3 and the probability of selecting the third statement is 0.6. In almost all cases the privacy level is high for both models. Hence, both of our scrambled response models may be useful when dealing with sensitive issues.

In Table 5, for model 2, the values of percent relative efficiencies increase for decreasing values of P_1 (probability of reporting sensitive variables). The values of percent relative efficiencies are high when the probability of selecting the third statement is also high (probability reporting scrambled value). For model 1, the values of PREs follow the same trend as in model 2 except for the probabilities P_1=0.3 and 0.2. We get the highest value of percent relative efficiency when the probability of selecting the first statement is 0.4 and the probability of selecting third statement is 0.5. In almost all cases the privacy level is high for both models.

Unified Measure of Models Quality

Following the work of Gupta et al[8] unified measures are calculated for the proposed models and Bar-Lev et al[3] models using the formula

$$\psi_{BL}=\frac{Var\left(\hat{\mu}_Y\right)}{(PL_{BL})},\ \psi_D=\frac{Var\left(\hat{\mu}_{Y_i}\right)}{(PL_D)_i}\ \text{for i=1, 2.}$$

where $Var\left(\hat{\mu}_{Y_i}\right)$ and (PL_B) are the variance and privacy level of the Bar-Lev et al[3] model and $Var\left(\hat{\mu}_{Y_i}\right)$ and $(PL_D)_i$ are the variance and privacy level of the proposed models respectively.

The graphical representations of the proposed models with respect to the Bar–Lev et al[3] model are shown in Figures 1–8.

Successive Sampling Scheme

Let $\Omega = \{1, 2, 3, \ldots N\}$ be a finite population of size N, which has been sampled over two occasions. The character under study is a sensitive variable denoted by x(y) on the first (second) occasion and z is a non-sensitive auxiliary variable available at both occasions. On the first (second)

Table 4 PRE of the Suggested Models 1, 2 for Data Set B with Respect to the Bar Lev et al[3] Model Along with PP and (ψ_D)

	P_1	P_2	Model 1			Model 2		
			PRE	PP	ψ_D	PRE	PP	ψ_D
Set B	0.6	0.1	628.9684	0.9731	0.1933	621.8278	0.9734	0.1955
		0.2	500.3729	0.9723	0.2432	510.9955	0.9717	0.2383
		0.3	416.7076	0.9706	0.2925	426.4308	0.9699	0.2861
		0.4	360.6787	0.9680	0.3389	360.6787	0.9680	0.3389
	0.5	0.1	718.6984	0.9733	0.2518	884.2605	0.9671	0.2059
		0.2	592.7823	0.9719	0.3057	712.5727	0.9662	0.2558
		0.3	507.9907	0.9697	0.3575	584.9739	0.9651	0.3119
		0.4	450.3687	0.9667	0.4045	487.8431	0.9639	0.3745
	0.4	0.1	787.7386	0.9730	0.3174	1392.5000	0.9522	0.1834
		0.2	668.8384	0.9711	0.3745	1072.4000	0.9537	0.2378
		0.3	587.1055	0.9685	0.4277	850.2804	0.9544	0.2997
		0.4	531.5301	0.9650	0.4742	689.7350	0.9546	0.3694
	0.3	0.1	836.4214	0.9723	0.3898	3009.9000	0.9003	0.1170
		0.2	727.5593	0.9701	0.4492	2000.1000	0.9178	0.1727
		0.3	652.0846	0.9671	0.5027	1441.4000	0.9273	0.2372
		0.4	601.5560	0.9631	0.5472	1093.4000	0.9329	0.3108
	0.2	0.1	864.9232	0.9714	0.4689	1095.0000	0.6376	0.0564
		0.2	768.1719	0.9688	0.5293	4875.2000	0.8020	0.1008
		0.3	701.3500	0.9654	0.5818	2942.2000	0.8547	0.1566
		0.4	658.2670	0.9609	0.6228	2012.6000	0.8803	0.2223

Table 5 PRE of the Suggested Models 1, 2 for Data Set D with Respect to the Bar Lev et al[3] Model Along with PP and (ψ_D)

	P_1	P_2	Model 1			Model 2		
			PRE	PP	ψ_D	PRE	PP	ψ_D
Set D	0.6	0.1	755.4813	0.9553	0.1958	899.3230	0.9468	0.1659
		0.2	1144.1000	0.9469	0.1304	1340.8000	0.9378	0.1124
		0.3	1987.7000	0.9288	0.0765	2245.8000	0.9195	0.0684
		0.4	4883.7000	0.8672	0.0334	4883.7000	0.8672	0.0334
	0.5	0.1	766.0197	0.9538	0.2956	972.1170	0.9413	0.2360
		0.2	1092.2000	0.9469	0.2089	1378.8000	0.9330	0.1679
		0.3	1706.8000	0.9340	0.1355	2126.3000	0.9178	0.1107
		0.4	3200.9000	0.9029	0.0747	3844.0000	0.8834	0.0636
	0.4	0.1	770.1999	0.9507	0.4152	1054.8000	0.9325	0.3091
		0.2	1058.0000	0.9443	0.3043	1457.2000	0.9233	0.2260
		0.3	1559.9000	0.9331	0.2089	2161.7000	0.9073	0.1550
		0.4	2609.4000	0.9099	0.1280	3642.7000	0.8742	0.0955
	0.3	0.1	767.2133	0.9465	0.5528	1153.4000	0.9196	0.3785
		0.2	1028.5000	0.9400	0.4152	1577.1000	0.9080	0.2803
		0.3	1463.7000	0.9291	0.2952	2310.9000	0.8880	0.1956
		0.4	2303.9000	0.9082	0.1919	3831.9000	0.8473	0.1236
	0.2	0.1	756.6400	0.9412	0.7059	1280.1000	0.9005	0.4361
		0.2	998.1845	0.9341	0.5392	1758.3000	0.8840	0.3235
		0.3	1389.6000	0.9226	0.3921	2608.0000	0.8548	0.2255
		0.4	2112.9000	0.9015	0.2639	4469.5000	0.7917	0.1421

occasion the sensitive variables x(y) are coded to g (h) with the aid of scrambling variables. The scrambling variable may follow a probability distribution.

The following notations have been considered for further use:

$\bar{X}, \bar{Y}, \bar{Z}, \bar{W}_1, \bar{W}_2, \bar{S}, \bar{U}$: population means of the variables x, y, z, W_1, W_2, S, U.

$\bar{g}_u, \bar{h}_m, \bar{g}_m, \bar{h}_n$: sample means of the respective variables based on sample sizes shown in suffices.

$\bar{z}_u, \bar{z}_m, \bar{z}_n$: sample means of the non-sensitive auxiliary variable based on the sample sizes shown in suffices.

$\rho_{yx}, \rho_{xz}, \rho_{yz}, \rho_{gh}, \rho_{hz}, \rho_{gz}$: correlation coefficients between the variables shown in suffices.

$S_x^2, S_y^2, S_z^2, S_{W_1}^2, S_{W_2}^2, S_S^2, S_U^2$: population variance of the variables x, y, z, W_1, W_2, S, U.

Proposed Scrambled Response Models Under Successive Sampling

In this section our hypothesis is to determine whether our models are effective or not for estimating the quantitative sensitive characteristics and for estimating the sensitive population mean when it changes frequently with the passage of time according to its nature.

The proposed randomized devices under the first and second occasions for the ith respondent scrambled response is given as

Model 1:

$$G_1=\begin{cases} X_i & \text{with prob. } P_1 \\ X_i\, W_1+W_2 & \text{with prob. } P_2 \\ X_i\, S & \text{with prob. } P_3 \end{cases}$$

$$H_1=\begin{cases} Y_i & \text{with prob. } P_1 \\ Y_i\, W_1+W_2 & \text{with prob. } P_2 \\ Y_i\, S & \text{with prob. } P_3 \end{cases}$$

For estimating the sensitive population mean on two occasions, successive sampling for the sensitive variables x(y) which are coded as g (h) and given as

$$(\bar{Y})_1 = \frac{\bar{H}_1 - W_2 P_2}{P_1 + W_1 P_2 + S P_3} \tag{5}$$

$$\rho_{h_1 g_1} = \frac{\begin{array}{l} P_1^2 \rho_{yx} S_y S_x + P_2^2\left(\rho_{yx} S_y S_x S_{w_1}^2 + \rho_{yx} S_y S_x \bar{w}_1^2 + \bar{x}\,\bar{y}\, S_{w_1}^2 + S_{w_2}^2\right) \\ +P_3^2\left(\rho_{yx} S_y S_x S_s^2 + \rho_{yx} S_y S_x \bar{S}^2 + \bar{x}\,\bar{y} S_s^2\right) \\ +2\rho_{yx} S_y S_x\left(\bar{w}_1 P_1 P_2 + \bar{S}\,\bar{w}_1 P_2 P_3 + \bar{S} P_1 P_3\right) \end{array}}{\sqrt{D_1}\,\sqrt{D_2}}$$

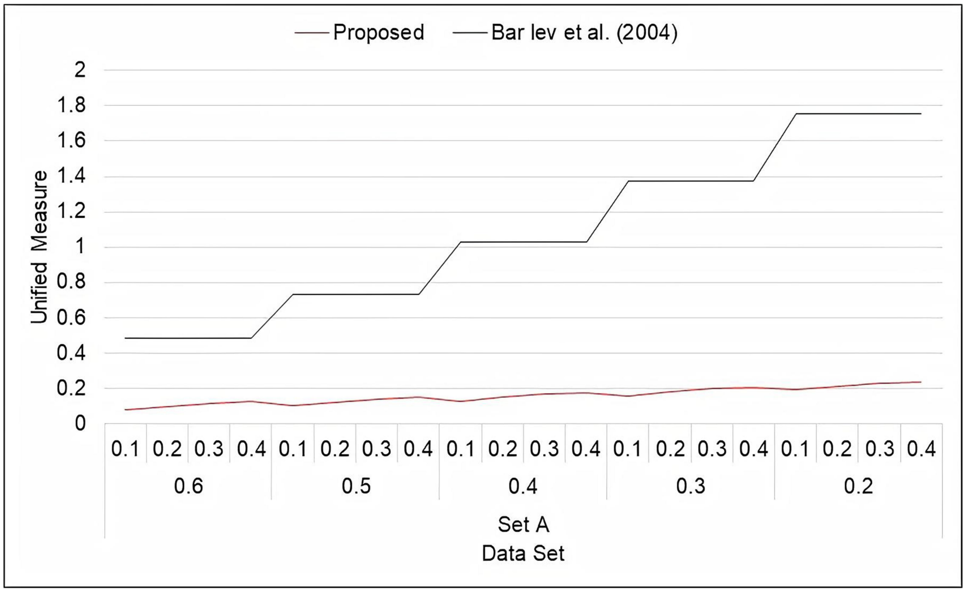

Figure 1 Unified measure for Model 1 (Set A).

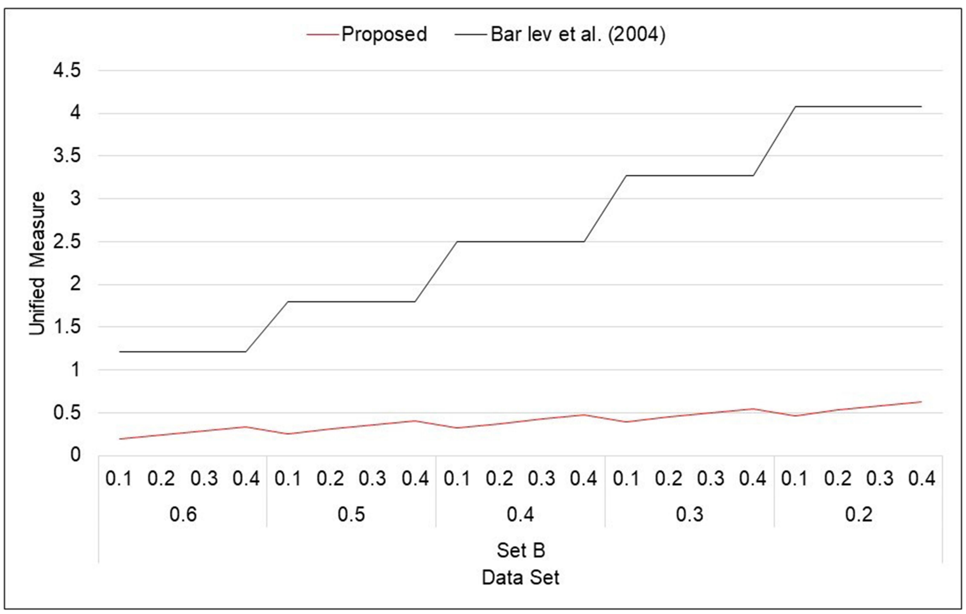

Figure 2 Unified measure for Model 1 (Set B).

$$\begin{aligned}D_1 = {} & S_y^2P_1^2+P_2^2\left(S_y^2S_{w_1}^2+S_y^2\bar{w}_1^2+\bar{y}^2S_{w_1}^2+S_{w_2}^2\right)\\ & +P_3^2\left(S_y^2S_s^2+S_y^2\bar{S}^2+\bar{y}^2S_y^2\right)\\ & +2S_y^2(\bar{w}_1P_1P_2+\bar{S}P_1P_3+\bar{S}\bar{w}_1P_2P_3)\end{aligned}$$

$$\begin{aligned}D_2 = {} & S_x^2P_1^2+P_2^2\left(S_x^2S_{w_1}^2+S_x^2\bar{w}_1^2+\bar{x}^2S_{w_1}^2+S_{w_2}^2\right)\\ & +P_3^2\left(S_x^2S_s^2+S_x^2\bar{S}^2+\bar{x}^2S_x^2\right)\\ & +2S_x^2(\bar{w}_1P_1P_2+\bar{S}P_1P_3+\bar{S}\bar{w}_1P_2P_3)\end{aligned}$$

$$\rho_{h_1z}=\frac{\rho_{yz}S_y(P_1+\bar{w}_1P_2+\bar{S}P_3)}{\sqrt{D_1}},\ \rho_{g_1z}=\frac{\rho_{xz}S_x(P_1+\bar{w}_1P_2+\bar{S}P_3)}{\sqrt{D_2}}$$

Model 2:

$$G_2=\begin{cases} X_i & \text{with prob. } P_1\\ X_i\,W_1+W_2 & \text{with prob. } P_2\\ W_3(X_i+U) & \text{with prob. } P_3\end{cases}$$

$$H_2=\begin{cases} Y_i & \text{with prob. } P_1\\ Y_i\,W_1+W_2 & \text{with prob. } P_2\\ W_3(Y_i+U) & \text{with prob. } P_3\end{cases}$$

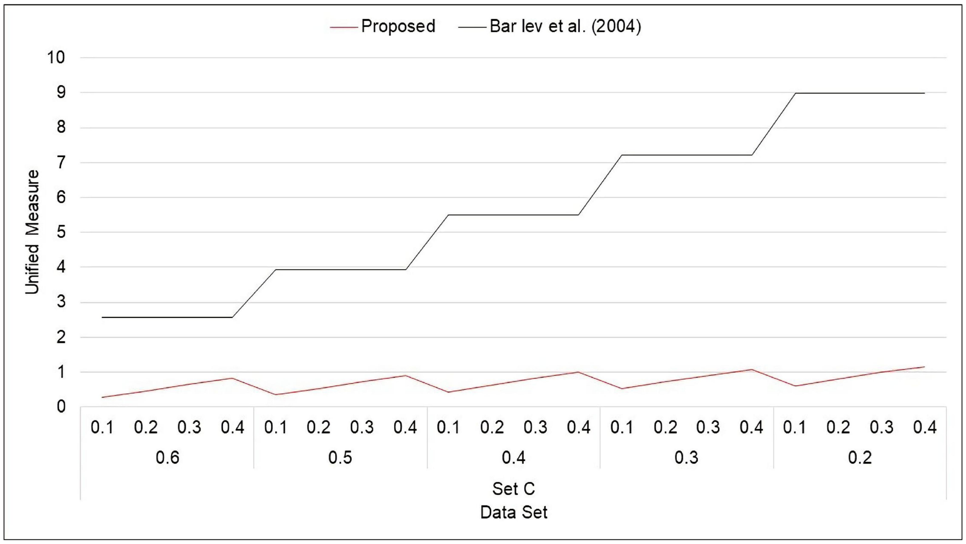

Figure 3 Unified measure for Model 1 (Set C).

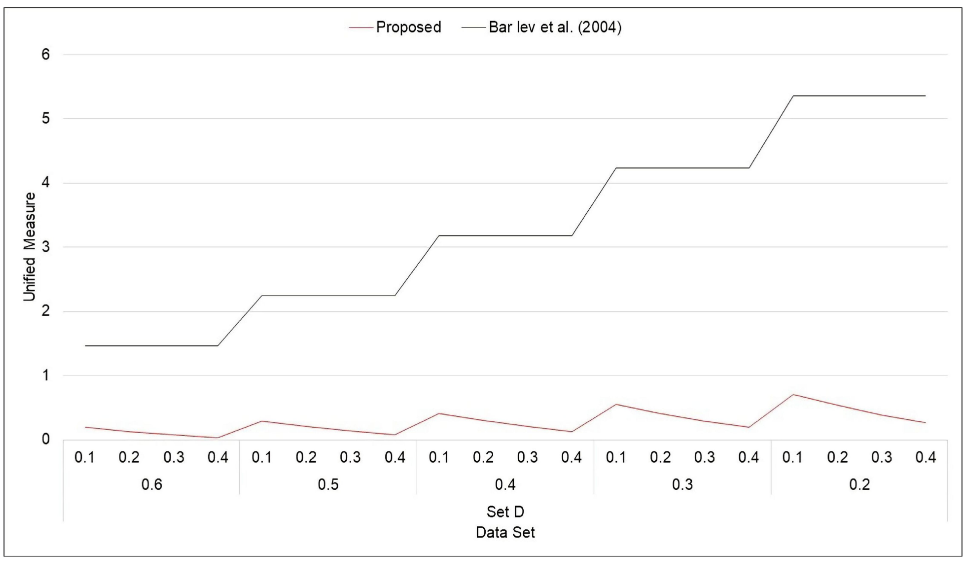

Figure 4 Unified measure for Model 1 (Set D).

The sensitive variables x(y) are coded as g (h) and are given by

$$(\bar{Y})_2 = \frac{\bar{H}_2 - P_2 W_2 - P_3 W_3 U}{P_1 + P_2 W_1 + P_3 W_3} \tag{6}$$

$$\rho_{h_2 g_2} = \frac{\begin{array}{l} P_1^2 \rho_{yx} S_y S_x \\ + P_2^2 \left(\rho_{yx} S_y S_x S_{w_1}^2 + \rho_{yx} S_y S_x \bar{w}_1^2 + \bar{x}\,\bar{y}\, S_{w_1}^2 + S_{w_2}^2 \right) \\ + P_3^2 \left(\begin{array}{l} \rho_{yx} S_y S_x S_{w_3}^2 + \rho_{yx} S_y S_x \bar{w}_3^2 + \bar{x}\,\bar{y}\, S_{w_3}^2 + \bar{x}\,\bar{u}\, S_{w_3}^2 \\ + \bar{y}\,\bar{u}\, S_{w_3}^2 + S_u^2 S_{w_3}^2 + S_u^2 \bar{w}_3^2 + \bar{u}\, S_{w_3}^2 \end{array} \right) \\ + 2\rho_{yx} S_y S_x (\bar{w}_1 P_1 P_2 + \bar{w}_3 P_1 P_3 + \bar{w}_1 \bar{w}_3 P_1 P_3) \end{array}}{\sqrt{D_3}\,\sqrt{D_4}}$$

$$D_3 = S_y^2 P_1^2 + P_2^2 \left(S_y^2 S_{w_1}^2 + S_y^2 \bar{w}_1^2 + \bar{y}^2 S_{w_1}^2 + S_{w_2}^2 \right) + P_3^2 \left(S_y^2 S_{w_3}^2 + S_{w_3}^2 \bar{y}^2 + \bar{w}_3^2 S_y^2 + S_{w_3}^2 S_u^2 + S_{w_3}^2 \bar{u}^2 + \bar{w}_3^2 S_u^2 + 2 S_{w_3}^2 \bar{y}\,\bar{u} \right) + 2 S_y^2 (\bar{w}_1 P_1 P_2 + \bar{w}_3 P_1 P_3 + \bar{w}_1 \bar{w}_3 P_2 P_3)$$

$$D_4 = S_x^2 P_1^2 + P_2^2 \left(S_x^2 S_{w_1}^2 + S_x^2 \bar{w}_1^2 + \bar{x}^2 S_{w_1}^2 + S_{w_2}^2 \right) + P_3^2 \left(S_x^2 S_{w_3}^2 + S_{w_3}^2 \bar{x}^2 + \bar{w}_3^2 S_x^2 + S_{w_3}^2 S_u^2 + S_{w_3}^2 \bar{u}^2 + \bar{w}_3^2 S_u^2 + 2 S_{w_3}^2 \bar{x}\,\bar{u} \right) + 2 S_y^2 (\bar{w}_1 P_1 P_2 + \bar{w}_3 P_1 P_3 + \bar{w}_1 \bar{w}_3 P_2 P_3)$$

$$\rho_{h_2 z} = \frac{\rho_{yz} S_y (P_1 + \bar{w}_1 P_2 + \bar{w}_3 P_3)}{\sqrt{D_3}}, \quad \rho_{g_2 z} = \frac{\rho_{xz} S_x (P_1 + \bar{w}_1 P_2 + \bar{w}_3 P_3)}{\sqrt{D_4}}$$

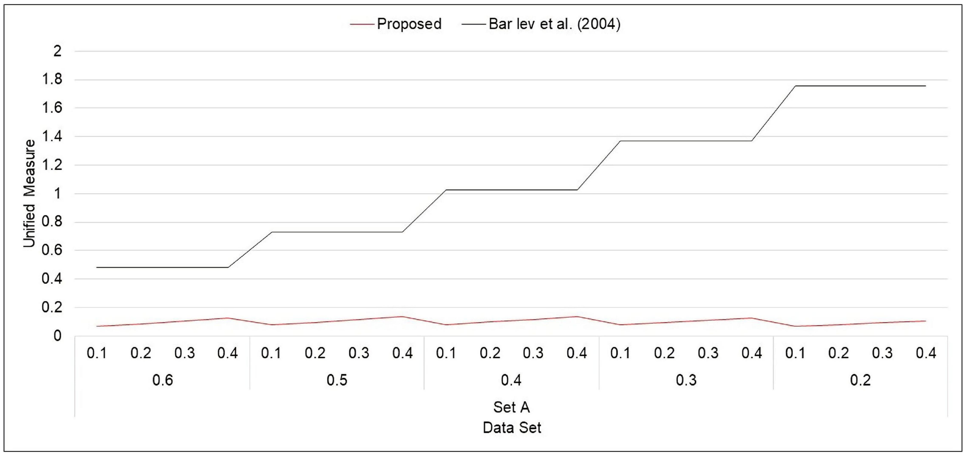

Figure 5 Unified measure for Model 2 (Set A).

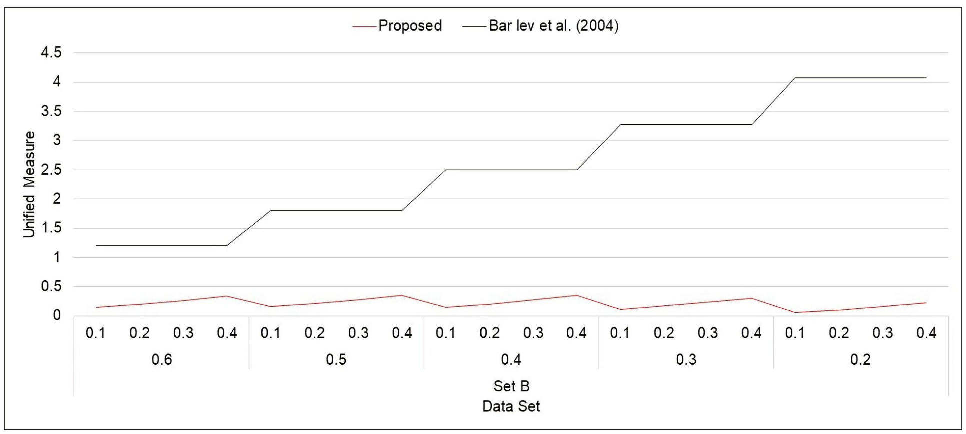

Figure 6 Unified measure for Model 2 (Set B).

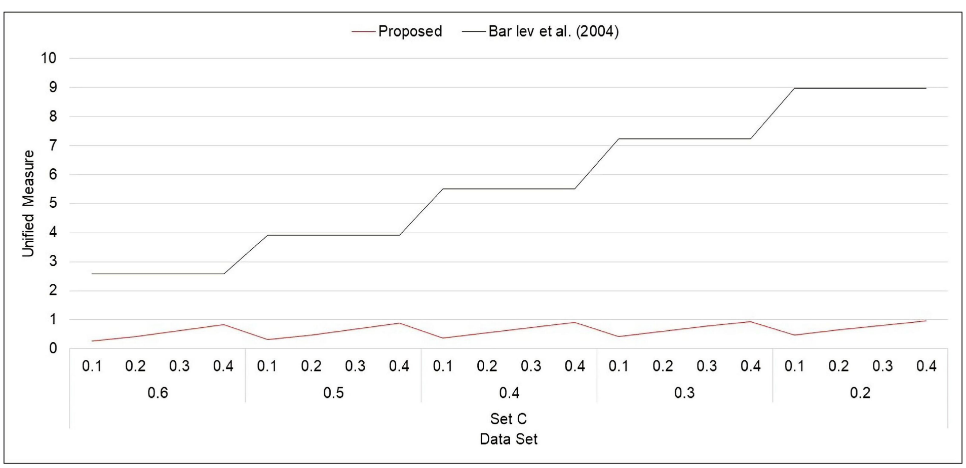

Figure 7 Unified measure for Model 2 (Set C).

Proposed Estimators

Exponential-type estimators play a vital role in increasing the precision of the estimates and are more stable over the sampling fluctuations. Bahl and Tuteja[2] were the first who proposed the exponential estimators and discussed certain regularity conditions under which their proposed

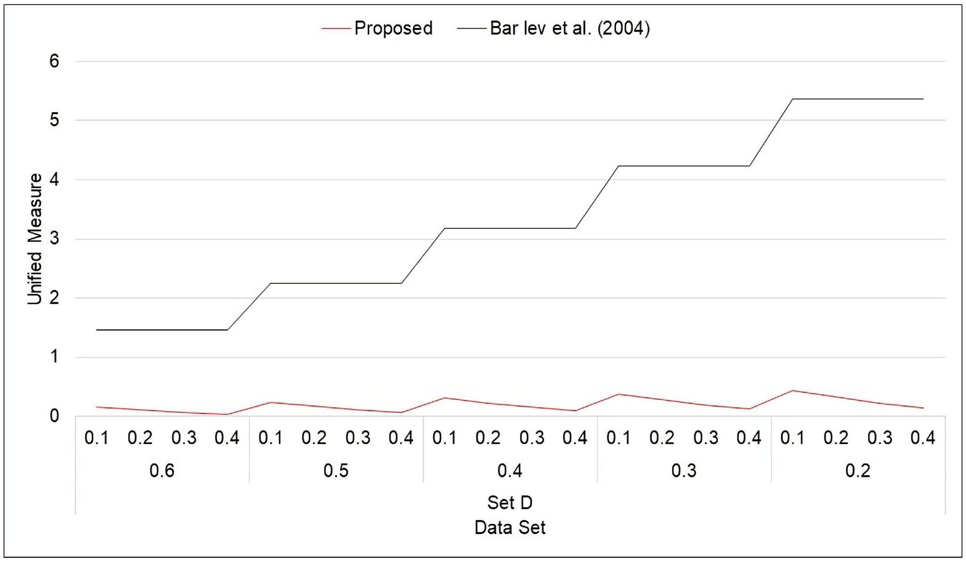

Figure 8 Unified measure for Model 2 (Set D).

estimators were better in comparison with mean per unit and ratio and product estimators. Further, they have reduced the sampling variance/MSE of their proposed estimators to the level of regression estimators. Thus, the exponential estimator given by Bahl and Tuteja[2] may be treated as an alternative of regression estimators. Thereafter, following the lines of Bahl and Tuteja,[2] various notable authors such as Singh and Homa,[19] Priyanka and Mittal,[13] Singh et al.,[20] Singh and Pal,[21] Singh et al[22] among others have proposed efficient estimators using an exponential function with effective results which are published in reputed journals. Since, exponential-type estimators are the better alternative for increasing the precision of the estimates. Encouraged by these works and motivated to put forward our idea to sample survey audiences, we have proposed exponential-type estimators and used the proposed scrambled response models in this under successive sampling. The proposed estimators found better in terms of percent relative efficiency and may be used when we deal with quantitative sensitive characteristics and for estimating the sensitive population mean when it changes frequently with the passage of time according to its nature.

For estimating the population mean of sensitive variables on the current (second) occasion, two independent estimators are proposed as follows:

$$T_u = \bar{h}_u \exp\left(\frac{\bar{Z}-\bar{z}_u}{\bar{Z}+\bar{z}_u}\right) \quad (7)$$

$$T_m = \bar{h}^*_m \exp\left(\frac{\bar{g}_n-\bar{g}_m}{\bar{g}_n+\bar{g}_m}\right)\exp\left(\frac{\bar{Z}-\bar{z}_n}{\bar{Z}+\bar{z}_n}\right) \textit{where}\ \bar{h}^*_m = \frac{\bar{h}_m}{\bar{g}_m}\bar{g}_n \quad (8)$$

The estimators defined in Equations (7 and 8) are further reproduced in the functional form as

$$T_u = g\left(\bar{h}_u, \bar{z}_u\right) \text{and}\ T_m = f\left(\bar{h}_m, \bar{g}_n, \bar{g}_m, \bar{z}_n\right) \quad (9)$$

The final estimator is a convex linear combination of two estimators T_u and T_m shown as

$$T_{prop} = \phi\, T_u + (1-\phi)\, T_m \quad (10)$$

where $\phi\,(0 \le \phi \le 1)$ is a constant to be determined under some criteria such that T_{prop} is more precise.

Properties of the Proposed Estimator

To obtain the bias and mean square error of the proposed estimators (T_u, T_m), we consider the following transformation.

$$\bar{h}_u = \bar{H}(1+e_0), \bar{h}_m = \bar{H}(1+e_1),$$
$$\bar{g}_m = \bar{G}(1+e_2), \bar{g}_n = \bar{G}(1+e_3),$$
$$\bar{z}_u = \bar{Z}(1+e_4), \bar{z}_n = \bar{Z}(1+e_5)$$

Such that $E(e_i) = 0$; $|e_i| < 1$, where i= 0, 1, 2, 3, 4, 5.

Thus, we have the following expressions

$$E\left(e_0^2\right) = \left(\frac{1}{u} - \frac{1}{N}\right)\left(\frac{S_h^2}{\bar{H}^2}\right), E\left(e_1^2\right) = \left(\frac{1}{m} - \frac{1}{N}\right)\left(\frac{S_h^2}{\bar{H}^2}\right),$$

$$E\left(e_2^2\right) = \left(\frac{1}{m} - \frac{1}{N}\right)\left(\frac{S_g^2}{\bar{G}^2}\right), E\left(e_3^2\right) = \left(\frac{1}{n} - \frac{1}{N}\right)\left(\frac{S_g^2}{\bar{G}^2}\right)$$

$$E\left(e_4^2\right) = \left(\frac{1}{u} - \frac{1}{N}\right)\left(\frac{S_z^2}{\bar{Z}^2}\right), E\left(e_5^2\right) = \left(\frac{1}{n} - \frac{1}{N}\right)\left(\frac{S_z^2}{\bar{Z}^2}\right),$$

$$E(e_0e_1) = -\frac{1}{N}\left(\frac{S_h^2}{\bar{H}^2}\right),$$

$$E(e_0e_2) = E(e_0e_3) = -\frac{1}{N}\left(\frac{S_h}{\bar{H}}\right)\left(\frac{S_g}{\bar{G}}\right)$$

$$E(e_0e_4) = \left(\frac{1}{u} - \frac{1}{N}\right)\rho_{hz}\left(\frac{S_h}{\bar{H}}\right)\left(\frac{S_z}{\bar{Z}}\right),$$

$$E(e_0e_5) = -\frac{1}{N}\rho_{hz}\left(\frac{S_h}{\bar{H}}\right)\left(\frac{S_z}{\bar{Z}}\right) = E(e_1e_4),$$

$$E(e_1e_2) = \left(\frac{1}{m} - \frac{1}{N}\right)\rho_{hg}\left(\frac{S_h}{\bar{H}}\right)\left(\frac{S_g}{\bar{G}}\right)$$

$$E(e_1e_3) = \left(\frac{1}{n} - \frac{1}{N}\right)\rho_{hg}\left(\frac{S_h}{\bar{H}}\right)\left(\frac{S_g}{\bar{G}}\right),$$

$$E(e_1e_5) = \left(\frac{1}{n} - \frac{1}{N}\right)\rho_{hz}\left(\frac{S_h}{\bar{H}}\right)\left(\frac{S_z}{\bar{Z}}\right),$$

$$E(e_2e_3) = \left(\frac{1}{n} - \frac{1}{N}\right)\left(\frac{S_g^2}{\bar{G}^2}\right),$$

$$E(e_2e_4) = -\frac{1}{N}\rho_{hz}\left(\frac{S_h}{\bar{H}}\right)\left(\frac{S_z}{\bar{Z}}\right),$$

$$E(e_2e_5) = E(e_3e_5) = \left(\frac{1}{n} - \frac{1}{N}\right)\rho_{gz}\left(\frac{S_g}{\bar{G}}\right)\left(\frac{S_z}{\bar{Z}}\right),$$

$$E(e_3e_4) = -\frac{1}{N}\rho_{gz}\left(\frac{S_g}{\bar{G}}\right)\left(\frac{S_z}{\bar{Z}}\right),$$

$$E(e_4e_5) = -\frac{1}{N}\left(\frac{S_z^2}{\bar{Z}^2}\right)$$

Using Taylor series expansions up to the first order, we expand the functional form of the estimator

$$T_u = g\left(\bar{h}_u, \bar{z}_u\right) \tag{11}$$

Thus, we have

$$\begin{aligned} T_u &= \bar{H} + \frac{\partial T_u}{\partial \bar{h}_u}\left(\bar{h}_u - \bar{H}\right) + \frac{\partial T_u}{\partial \bar{z}_u}(\bar{z}_u - \bar{Z}) \\ &+ \frac{1}{2}\{\left(\bar{h}_u - \bar{H}\right)^2\frac{\partial^2 T_u}{\partial \bar{h}_u^2} + (\bar{z}_u - \bar{Z})^2\frac{\partial^2 T_u}{\partial \bar{z}_u^2} \\ &+ 2\left(\bar{h}_u - \bar{H}\right)(\bar{z}_u - \bar{Z})\frac{\partial^2 T_u}{\partial \bar{h}_u \partial \bar{z}_u}\} \\ T_u &= \bar{H} + g_1\left(\bar{h}_u - \bar{H}\right) + g_2(\bar{z}_u - \bar{Z}) \\ &+ \frac{1}{2}\{\left(\bar{h}_u - \bar{H}\right)^2 g_{11} + (\bar{z}_u - \bar{Z})^2 g_{22} \\ &+ 2\left(\bar{h}_u - \bar{H}\right)(\bar{z}_u - \bar{Z})g_{12}\} \end{aligned} \tag{12}$$

where $g_1 = \left.\frac{\partial T_u}{\partial \bar{h}_u}\right|_{\left(\bar{h}_u = \bar{H}, \bar{z}_u = \bar{Z}\right)} = 1,$

$$g_2 = \left.\frac{\partial T_u}{\partial \bar{z}_u}\right|_{\left(\bar{h}_u = \bar{H}, \bar{z}_u = \bar{Z}\right)} = -\frac{1}{2}\frac{\bar{H}}{\bar{Z}},$$

$$g_{11} = \left.\frac{\partial^2 T_u}{\partial \bar{h}_u^2}\right|_{\left(\bar{h}_u = \bar{H}, \bar{z}_u = \bar{Z}\right)} = 0$$

$$g_{22} = \left.\frac{\partial^2 T_u}{\partial \bar{z}_u^2}\right|_{\left(\bar{h}_u = \bar{H}, \bar{z}_u = \bar{Z}\right)} = \frac{3\bar{H}}{4\bar{Z}^2},$$

$$g_{12} = \left.\frac{\partial^2 T_u}{\partial \bar{h}_u \partial \bar{z}_u}\right|_{\left(\bar{h}_u = \bar{H}, \bar{z}_u = \bar{Z}\right)} = -\frac{1}{2\bar{Z}}$$

The bias and mean square error of the estimator T_u are-given as

$$\begin{aligned} B(T_u) &= E(T_u - \bar{H}) = E[g_1\bar{H}e_0 + g_2\bar{Z}e_4 \\ &+ \frac{1}{2}\{g_{11}\bar{H}^2e_0^2 + g_{22}\bar{Z}^2e_4^2 + 2g_{12}e_0e_4\}] \end{aligned} \tag{13}$$

$$MSE(T_u) = E(T_u - \bar{H})^2 = E[g_1\bar{H}e_0 + g_2\bar{Z}e_4]^2 \tag{14}$$

Remark: s g and h are two coded response variables over two successive moves and z is a stable non-sensitive auxiliary variable. Hence, following Murthy[11] and Reddy,[16] we assume that the coefficient of variations g, h and z are almost equal $\left(i.e.\, C_h \cong C_g \cong C_z\right)$.

Substituting the values of g_1, g_2, g_{11}, g_{22} and g_{12} in Equations (13 and 14), we get the bias and mean square of the estimator T_u as

$$B(T_u) = \bar{H}\left[\left(\frac{1}{u} - \frac{1}{N}\right)\left(\frac{3}{8} - \frac{1}{2}\rho_{hz_1}\right)\right]C_h^2 \tag{15}$$

$$M(T_u) = \left[\left(\frac{1}{u} - \frac{1}{N}\right)\left(\frac{5}{4} - \rho_{hz_1}\right)\right]S_h^2 \tag{16}$$

The function $T_u = g\left(\bar{h}_u, \bar{z}_u\right)$ is based on statistics $g\left(\bar{h}_u, \bar{z}_u\right)$ and satisfies the following regularity conditions:

The point $T_u = g\left(\bar{h}_u, \bar{z}_u\right)$ assumes the value in the closed convex subset R^2 of two-dimensional space containing the point $(\bar{H}, \bar{Z})$.

The function $g\left(\bar{h}_u, \bar{z}_u\right)$ is continuous and bounded in R^2.

$g(\bar{H}, \bar{Z}) = \bar{H}$ and $g_1(\bar{H}, \bar{Z}) = 1$ where $g_1(\bar{H}, \bar{Z}) = 1$ is the first-order derivative of g with respect to $\bar{h}_u$.

The first-order partial derivatives of $g\left(\bar{h}_u, \bar{z}_u\right)$ exists and is continuous and bounded in R^2

Similarly, the bias and mean square error of the estimator T_m are derived using the following steps

$$T_m = f\left(\bar{h}_m, \bar{g}_n, \bar{g}_m, \bar{z}_n\right) \tag{17}$$

Following Taylor series expressions, we have

$$\begin{aligned} T_m &= \bar{H} + \frac{\partial T_m}{\partial \bar{h}_m}\left(\bar{h}_m - \bar{H}\right) + \frac{\partial T_m}{\partial \bar{g}_n}(\bar{g}_n - \bar{G}) \\ &+ \frac{\partial T_m}{\partial \bar{g}_m}(\bar{g}_m - \bar{G}) + \frac{\partial T_m}{\partial \bar{z}_n}(\bar{z}_n - \bar{Z}) \\ &+ \frac{1}{2}\left\{\frac{\partial^2 T_m}{\partial \bar{h}_m^2}\left(\bar{h}_m - \bar{H}\right) + \frac{\partial^2 T_m}{\partial \bar{g}_n^2}(\bar{g}_n - \bar{G})^2\right. \\ &+ \frac{\partial^2 T_m}{\partial \bar{g}_m^2}(\bar{g}_m - \bar{G})^2 + \frac{\partial^2 T_m}{\partial \bar{z}_n^2}(\bar{z}_n - \bar{Z})^2 \\ &+ 2\frac{\partial^2 T_m}{\partial \bar{h}_m \partial \bar{g}_n}\left(\bar{h}_m - \bar{H}\right)(\bar{g}_n - \bar{G}) \\ &+ 2\frac{\partial^2 T_m}{\partial \bar{h}_m \bar{g}_m}\left(\bar{h}_m - \bar{H}\right)(\bar{g}_m - \bar{G}) \\ &+ 2\frac{\partial^2 T_m}{\partial \bar{h}_m \partial \bar{z}_n}\left(\bar{h}_m - \bar{H}\right)(\bar{z}_n - \bar{Z}) \\ &+ 2\frac{\partial^2 T_m}{\partial \bar{g}_n \partial \bar{g}_m}(\bar{g}_n - \bar{G})(\bar{g}_m - \bar{G}) \\ &+ 2\frac{\partial^2 T_m}{\partial \bar{g}_n \partial \bar{z}_n}(\bar{g}_n - \bar{G})(\bar{z}_n - \bar{Z}) \\ &+ 2\left.\frac{\partial^2 T_m}{\partial \bar{g}_m \partial \bar{z}_n}(\bar{g}_m - \bar{G})(\bar{z}_n - \bar{Z})\right\} \end{aligned}$$

$$\begin{aligned} T_m &= \bar{H} + f_1\left(\bar{h}_m - \bar{H}\right) + f_2(\bar{g}_n - \bar{G}) + f_3(\bar{g}_m - \bar{G}) \\ &+ f_4(\bar{z}_n - \bar{Z}) + \frac{1}{2}\{f_{11}\left(\bar{h}_m - \bar{H}\right) + f_{22}(\bar{g}_n - \bar{G})^2 \\ &+ f_{33}(\bar{g}_m - \bar{G})^2 + f_{44}(\bar{z}_n - \bar{Z})^2 \\ &+ 2f_{12}\left(\bar{h}_m - \bar{H}\right)(\bar{g}_n - \bar{G}) + 2f_{13}\left(\bar{h}_m - \bar{H}\right)(\bar{g}_m - \bar{G}) \\ &+ 2f_{14}\left(\bar{h}_m - \bar{H}\right)(\bar{z}_n - \bar{Z}) + 2f_{23}(\bar{g}_n - \bar{G})(\bar{g}_m - \bar{G}) \\ &+ 2f_{24}(\bar{g}_n - \bar{G})(\bar{z}_n - \bar{Z}) + 2f_{34}(\bar{g}_m - \bar{G})(\bar{z}_n - \bar{Z})\} \end{aligned} \tag{18}$$

where

$$f_1 = \left.\frac{\partial T_m}{\partial \bar{h}_m}\right|_{\left(\bar{h}_m=\bar{H},\bar{g}_n=\bar{G},\bar{g}_m=\bar{G},\bar{z}_n=\bar{Z}\right)} = 1,$$

$$f_2 = \left.\frac{\partial T_m}{\partial \bar{g}_n}\right|_{\left(\bar{h}_m=\bar{H},\bar{g}_n=\bar{G},\bar{g}_m=\bar{G},\bar{z}_n=\bar{Z}\right)} = \frac{3\bar{H}}{2\bar{G}},$$

$$f_2 = \left.\frac{\partial T_m}{\partial \bar{g}_n}\right|_{\left(\bar{h}_m=\bar{H},\bar{g}_n=\bar{G},\bar{g}_m=\bar{G},\bar{z}_n=\bar{Z}\right)} = \frac{\bar{H}}{2\bar{G}}$$

$$f_2 = \left.\frac{\partial T_m}{\partial \bar{z}_n}\right|_{\left(\bar{h}_m=\bar{H},\bar{g}_n=\bar{G},\bar{g}_m=\bar{G},\bar{z}_n=\bar{Z}\right)} = -\frac{\bar{H}}{2\bar{Z}},$$

$$f_{11} = \left.\frac{\partial^2 T_m}{\partial \bar{h}_m^2}\right|_{\left(\bar{h}_m=\bar{H},\bar{g}_n=\bar{G},\bar{g}_m=\bar{G},\bar{z}_n=\bar{Z}\right)} = 0,$$

$$f_{22} = \left.\frac{\partial^2 T_m}{\partial \bar{g}_n^2}\right|_{\left(\bar{h}_m=\bar{H},\bar{g}_n=\bar{G},\bar{g}_m=\bar{G},\bar{z}_n=\bar{Z}\right)} = \frac{3\bar{H}}{4\bar{G}^2}$$

$$f_{33} = \left.\frac{\partial^2 T_m}{\partial \bar{g}_m^2}\right|_{\left(\bar{h}_m=\bar{H},\bar{g}_n=\bar{G},\bar{g}_m=\bar{G},\bar{z}_n=\bar{Z}\right)} = \frac{7\bar{H}}{4\bar{G}^2},$$

$$f_{44} = \left.\frac{\partial^2 T_m}{\partial \bar{z}_n^2}\right|_{\left(\bar{h}_m=\bar{H},\bar{g}_n=\bar{G},\bar{g}_m=\bar{G},\bar{z}_n=\bar{Z}\right)} = \frac{3\bar{H}}{4\bar{Z}^2},$$

$$f_{12} = \left.\frac{\partial^2 T_m}{\partial \bar{h}_m \partial \bar{g}_n}\right|_{\left(\bar{h}_m=\bar{H},\bar{g}_n=\bar{G},\bar{g}_m=\bar{G},\bar{z}_n=\bar{Z}\right)} = \frac{3}{2\bar{G}},$$

$$f_{13} = \left.\frac{\partial^2 T_m}{\partial \bar{h}_m \bar{g}_m}\right|_{\left(\bar{h}_m=\bar{H},\bar{g}_n=\bar{G},\bar{g}_m=\bar{G},\bar{z}_n=\bar{Z}\right)} = -\frac{3}{2\bar{G}}$$

$$f_{14} = \left.\frac{\partial^2 T_m}{\partial \bar{h}_m \partial \bar{z}_n}\right|_{\left(\bar{h}_m=\bar{H},\bar{g}_n=\bar{G},\bar{g}_m=\bar{G},\bar{z}_n=\bar{Z}\right)} = -\frac{1}{2\bar{Z}},$$

$$f_{23} = \left.\frac{\partial^2 T_m}{\partial \bar{g}_n \partial \bar{g}_m}\right|_{\left(\bar{h}_m=\bar{H},\bar{g}_n=\bar{G},\bar{g}_m=\bar{G},\bar{z}_n=\bar{Z}\right)} = -\frac{9\bar{H}}{4\bar{G}^2}$$

$$f_{24} = \left.\frac{\partial^2 T_m}{\partial \bar{g}_n \partial \bar{z}_n}\right|_{\left(\bar{h}_m=\bar{H},\bar{g}_n=\bar{G},\bar{g}_m=\bar{G},\bar{z}_n=\bar{Z}\right)} = -\frac{3\bar{H}}{4\bar{G}\bar{Z}},$$

$$f_{34} = \left.\frac{\partial^2 T_m}{\partial \bar{g}_m \partial \bar{z}_n}\right|_{\left(\bar{h}_m=\bar{H},\bar{g}_n=\bar{G},\bar{g}_m=\bar{G},\bar{z}_n=\bar{Z}\right)} = \frac{3\bar{H}}{4\bar{G}\bar{Z}}$$

The bias and mean square error of the estimator T_m are given as

$$B(T_m) = B(T_m - \bar{H}) = \begin{bmatrix} f_1\bar{H}e_1 + f_2\bar{G}e_3 + f_3\bar{G}e_2 + f_4\bar{Z}e_5 \\ +\frac{1}{2}\left\{f_{11}\bar{H}^2e_1^2 + f_{22}\bar{G}^2e_3^2 + f_{33}\bar{G}^2e_2^2 + f_{44}\bar{Z}^2e_5^2\right. \\ +2f_{12}\bar{H}\bar{G}e_1e_3 + 2f_{13}\bar{H}\bar{G}e_1e_2 + 2f_{14}\bar{H}\bar{Z}e_1e_5 \\ \left. +2f_{23}\bar{G}^2e_2e_3 + 2f_{24}\bar{G}\bar{Z}e_3e_5 + 2f_{34}\bar{G}\bar{Z}e_2e_5 \right\} \end{bmatrix} \tag{19}$$

$$\begin{aligned} MSE(T_m) &= E(T_m - \bar{H})^2 \\ &= [f_1\bar{H}e_1 + f_2\bar{G}e_3 + f_3\bar{G}e_2 + f_4\bar{Z}e_5]^2 \end{aligned} \tag{20}$$

Substituting the values of $f_1, f_2, f_3, f_4, f_{11}, f_{22}, f_{33}, f_{44}, f_{12}, f_{13}, f_{14}, f_{23}, f_{24}, f_{34}$ in Equations (19 and 20), we get the bias and mean square of the estimator T_m as

$$B(T_m) = \bar{H}\left[\left(\frac{1}{m} - \frac{1}{n}\right)\left(\frac{15}{8} - \frac{3}{2}\rho_{hg}\right) + \left(\frac{1}{n} - \frac{1}{N}\right)\left(\frac{3}{8} - \frac{1}{2}\rho_{hz_1}\right)\right] \tag{21}$$

$$\mathrm{M}(T_{\mathrm{m}})=\left[\frac{1}{\mathrm{m}}\left(\frac{13}{4}-3\rho_{\mathrm{hg}}\right)+\frac{1}{\mathrm{n}}\left(3\rho_{\mathrm{hg}}-\rho_{\mathrm{hz}_1}-2\right)\right.$$
$$\left.-\frac{1}{\mathrm{N}}\left(\frac{5}{4}-\rho_{\mathrm{hz}_1}\right)\right]\mathrm{S}_{\mathrm{h}}^2 \tag{22}$$

The function $\mathrm{T_m}=\mathrm{f}(\bar{h}_{\mathrm{m}},\bar{g}_{\mathrm{n}},\bar{g}_{\mathrm{m}},\bar{z}_{\mathrm{n}})$ is based on statistics $\mathrm{f}(\bar{h}_{\mathrm{m}},\bar{g}_{\mathrm{n}},\bar{g}_{\mathrm{m}},\bar{z}_{\mathrm{n}})$ and satisfies the similar regularity conditions as Equation (11).

From Equations (12 and 18) we get the covariance between T_u and T_m

$$C(T_u,T_{\mathrm{m}})=-\frac{1}{\mathrm{N}}\left(\frac{5}{4}-\rho_{\mathrm{hz}}\right)\mathrm{S}_{\mathrm{h}}^2 \tag{23}$$

Theorem 1: The bias of the estimator T_{prop} to be the first order of approximations is obtained as

$$\begin{aligned}\mathrm{B}(T_{prop}) &= E(T_{prop}-\bar{H})\\ &= E[\phi(T_u-\bar{H})+(1-\phi)(T_m-\bar{H})]\\ &= \phi B(T_{\mathrm{u}})+(1-\phi)B(T_m)\end{aligned} \tag{24}$$

where $B(T_{\mathrm{u}})$ and $B(T_m)$ are given in Equations (15 and 21) respectively.

Theorem 2: Mean square error of the estimator T_{prop} to the first order of approximations is derived as

$$\begin{aligned}M(T_{ptop}) &= E(T_{prop}-\bar{H})^2\\ &= E[\phi(T_u-\bar{H})+(1-\phi)(T_m-\bar{H})]^2\\ &= \phi^2M(T_{\mathrm{u}})+(1-\phi)^2M(T_m)\\ &\quad+2\phi(1-\phi)C(T_{\mathrm{u}},T_m)\end{aligned} \tag{25}$$

where $M(T_{\mathrm{u}})$, $M(T_m)$ and $C(T_u,T_m)$ are given in Equations (16, 22 and 23) respectively.

Minimum Mean Square Error of the Estimator T_{prop}

The MSE of the estimator T_{prop} given in Equation (25) is a function of unknown constant ϕ, therefore, to obtain the optimum choice of ϕ, we minimize Equation (25) with respect to ϕ as

$$\phi_{opt}=\frac{M(T_m)-C(T_u,T_m)}{M(T_u)+M(T_m)-2C(T_u,T_m)} \tag{26}$$

Putting the value of ϕ_{opt} in Equation (25), we have optimum MSE of the estimator T_{prop} as

$$M(T_{prop})_{opt}=\frac{M(T_u)\,.\,M(T_m)-[C(T_u,T_m)]^2}{M(T_u)+M(T_m)-2C(T_u,T_m)} \tag{27}$$

Further, substituting the expressions of $\mathrm{M}(T_{\mathrm{u}}),\ \mathrm{M}(T_m)\ and\ C(T_{\mathrm{u}},T_m)$ in Equations (26 and 27), the simplified values of ϕ_{opt} and $\mathrm{M}(T_{prop})_{opt}$ are obtained as

$$\phi_{opt}=\left[\frac{\mu(A_3+A_2\mu)}{A_3-A_2\mu^2}\right]\ and\ M(\xi)_{opt}=\left[\frac{(1-f)\,A_4-A_5\,\mu+A_6\,\mu^2}{A_3-A_2\,\mu^2}\right]\frac{S_h^2}{n} \tag{28}$$

where

$$\mathrm{A}_1=\left(\frac{13}{4}-3\rho_{\mathrm{hg}}\right),\ \mathrm{A}_2=\left(3\rho_{\mathrm{hg}}-\rho_{\mathrm{hz}_1}-2\right),$$
$$\mathrm{A}_3=\left(\frac{5}{4}-\rho_{\mathrm{hz}_1}\right),\ \mathrm{A}_3=(\mathrm{A}_1+\mathrm{A}_2),$$
$$\mathrm{A}_3^2=\mathrm{A}_4,\ \mathrm{A}_2\,\mathrm{A}_3=\mathrm{A}_5$$
$$\mathrm{A}_2\,\mathrm{A}_3\,\mathrm{f}=\mathrm{A}_6,\ \mathrm{f}=(\mathrm{n}/\mathrm{N})\ \text{and}\ \mu=(\mathrm{u}/\mathrm{n})$$

Optimum Replacement Strategy

To determine the optimum value of μ (fraction of sample to be drawn afresh on the current occasion) so that $\bar{H}$ be estimated with maximum precision and minimum cost, we minimize $M(T_{prop})_{opt}$ μ which results in a quadratic equation in μ given as

$$\mu^2 A_1-2\mu A_2+A_3=0 \tag{29}$$

Solving Equation (29), the solutions of $\mu\,(say\,\hat{\mu})$ are given as

$$\hat{\mu}=\frac{A_2\pm\sqrt{A_1A_3}}{A_1} \tag{30}$$

The real value of $\hat{\mu}$ lies only if $(A_1A_3)\geq 0$. So, using Equation (30) and putting the admissible value of $\hat{\mu}\,(say\,\mu_0)$ in Equation (28), the optimum value for the mean square error of the estimator $\mathrm{T_{prop}}$ is written as

$$M(T_{prop})^{(0)}_{opt}=\left[\frac{(1-f)\,A_4-A_5\,\mu_0+A_6\,\mu_0^2}{A_3-A_2\,\mu_0^2}\right]\frac{S_h^2}{n} \tag{31}$$

Estimators for Sensitive Population Mean at Current Occasion Under Proposed Scrambled Response Models

To obtain the estimator of the sensitive population mean at current occasion estimators $T_{prop},\ t_s,\ t_n\ and\ T_{ss}$ the coded response variable $\bar{H}$ on the current move in Equations 1 and 2 is replaced by its estimators respectively and presented in Table 6.

Table 6 Sensitive Population Mean Estimators and Their Mean Square Errors Under the Proposed Scrambled Response Models

Model	Sensitive Population Mean Estimator	MSE of Sensitive Population Mean Estimator
Model 1	$\bar{Y}_{prop} = \frac{T_{prop} - W_2 P_2}{P_1 + W_1 P_2 + S P_3}$	$MSE(\bar{Y}_{prop}) = \frac{MSE(T_{prop})_{opt}}{[P_1 + W_1 P_2 + S P_3]^2}$
	$\bar{Y}_s = \frac{\bar{h}_n - W_2 P_2}{P_1 + W_1 P_2 + S P_3}$	$MSE(\bar{Y}_s) = \frac{MSE(\bar{h}_n)}{[P_1 + W_1 P_2 + S P_3]^2}$
	$\bar{Y}_n = \frac{\hat{\hat{H}} - W_2 P_2}{P_1 + W_1 P_2 + S P_3}$	$MSE(\bar{Y}_n) = \frac{MSE(\hat{\hat{H}})}{[P_1 + W_1 P_2 + S P_3]^2}$
	$\bar{Y}_{ss} = \frac{T_{ss} - W_2 P_2}{P_1 + W_1 P_2 + S P_3}$	$MSE(\bar{Y}_{ss}) = \frac{MSE(T_{ss})_{opt}}{[P_1 + W_1 P_2 + S P_3]^2}$
Model 2	$\bar{Y}_{prop} = \frac{T_{prop} - W_2 P_2 - P_3 W_3 U}{P_1 + W_1 P_2 + W_3 P_3}$	$MSE(\bar{Y}_{prop}) = \frac{MSE(T_{prop})}{[P_1 + W_1 P_2 + W_3 P_3]^2}$
	$\bar{Y}_s = \frac{\bar{h}_n - W_2 P_2 - P_3 W_3 U}{P_1 + W_1 P_2 + W_3 P_3}$	$MSE(\bar{Y}_s) = \frac{MSE(\bar{h}_n)}{[P_1 + W_1 P_2 + W_3 P_3]^2}$
	$\bar{Y}_n = \frac{\hat{\hat{H}} - W_2 P_2 - P_3 W_3 U}{P_1 + W_1 P_2 + W_3 P_3}$	$MSE(\bar{Y}_n) = \frac{MSE(\hat{\hat{H}})}{[P_1 + W_1 P_2 + W_3 P_3]^2}$
	$\bar{Y}_{ss} = \frac{T_{ss} - W_2 P_2 - P_3 W_3 U}{P_1 + W_1 P_2 + W_3 P_3}$	$MSE(\bar{Y}_{sl}) = \frac{MSE(T_{ss})_{opt}}{[P_1 + W_1 P_2 + W_3 P_3]^2}$

Monte Carlo Simulation Study

The simulation study has been carried out by considering 5000 different samples using Monte Carlo simulation for Population-I. The simulated PRE of the proposed estimator with respect to the sample mean estimator, natural successive sampling estimator and Singh and Sharma[18] estimators have been computed using the data set $n = 20$, $u = 4$, $m = 16$.

The following steps summarize the simulation study as follows:

Step 1. Consider a real population of size N=51 (Population-I), from which 5000 SRSWOR different samples of size n=20 have been selected.

Step 2. From each of the selected samples, m=16 units were retained as a matched portion from sample n.

Step 3. To draw a SRSWOR of size u=4 (fresh) from the remaining part of the population of size N-n=31 as an unmatched portion.

Step 4. Calculate the value of T_u from new unmatched units, u on the current occasion and T_m from the m units retained units.

Step 5. Calculate the value of the estimator T_{prop} based on the value of T_u and T_m.

Step 6. Repeat step (2), (3), (4) and (5), 5000 times. Thus, we obtain 5000 values for the suggested estimator T_{prop}.

Step 7. The MSE of T_{prop} is obtained by $M(T_{prop}) = \frac{1}{5000}\sum_{i=1}^{5000}(T_{prop} - \bar{H})^2$.

The efficiency of the estimator T_{prop} with respect to the considered estimator is defined by:

$$PRE = \frac{M(.)}{M(T_{prop})} \times 100$$

From Table 7, we deal with quantitative sensitive characteristics and for estimating the population mean when it changes frequently with the passing of time according to its nature. The values of percent relative efficiencies are always greater than 100 for all cases but do not follow any specific pattern for decreasing values of P_1 and increasing values of P_2. Therefore, the proposed estimators may be used when we want to monitor sensitive issues which change according to time.

Empirical Study

To show the performances of the proposed estimator T_{prop}, we compare with sample mean estimator $\bar{h}_n$ (when there is no matching), natural successive sampling estimator $\hat{\hat{H}} = \phi\bar{h}_u + (1-\phi)\bar{h}'_m$ where $\bar{h}'_m = \bar{h}_m + \beta_{hg}(\bar{G}_n - \bar{G}_m)$ and the Singh and Sharma[18] estimator T_{SS}. The estimator suggested by Singh and Sharma[18] is as follows

$$T_{SS} = \phi\, T_{su} + (1-\phi)T_{sm} \tag{32}$$

where $T_{su} = \bar{h}_u\left(\frac{\bar{Z}_1 + A}{\bar{z}_{1u} + A}\right)\exp\left(\frac{\bar{Z}_1 - \bar{z}_{1u}}{\bar{Z}_1 + \bar{z}_{1u}}\right)$ and $T_{sm} = \frac{\bar{h}_m}{\bar{g}_m}\,\bar{x}_n\left(\frac{\bar{Z}_1 + A}{\bar{z}_{1u} + A}\right)\exp\left(\frac{\bar{Z}_1 - \bar{z}_{1m}}{\bar{Z}_1 + \bar{z}_{1m}}\right)$

The variance of $\bar{h}_n$, optimum variance of $\hat{\hat{H}}$ and the optimum mean square error of Singh and Sharma[18] estimators are as follows

$$V(\bar{h}_n) = \left(\frac{1}{n} - \frac{1}{N}\right)S_h^2,\ V\left(\hat{\hat{H}}\right) = \left[1 + \sqrt{1 - \rho_{hg}^2}\right]\frac{S_h^2}{2n} - \frac{S_h^2}{N}$$

$$M(T_{ss})_{opt} = \left[\frac{A_3 - A_4\mu_1^{(0)} - A_5\mu_1^{(0)^2}}{A_1 - (A_1 - A_2)\mu_1^{(0)^2}}\right]\frac{S_h^2}{n}$$

where

$$A_1 = \left(k^2 + k + 5/4\right) - (2k+1)\rho_{yz_1},$$
$$A_2 = (9/4) - 2\rho_{yx},\ A_3 = (1-f)A_1^2,$$
$$A_4 = A_1^2 - A_1A_2,\ A_2 = fA_4,\ f = n/N$$

Population-I: source Statistical Abstracts of United States.

Let y, x and z be the number of abortions reported in the states of the US during the year 2007, 2005 and 2004 respectively. We consider a real data of N=51 units with the following parameters:

Table 7 Percent Relative Efficiencies of T_{prop} with Respect to $\bar{h}_n$, $\hat{\bar{H}}$ and Singh and Sharma.[18]

P_1	P_2	Model 1			Model 2		
		E_1	E_2	E_3	E_1	E_2	E_3
0.6	0.1	531.8465	329.8350	398.4332	135.8645	251.9519	522.9993
	0.2	322.0406	305.6881	376.0890	635.7316	145.0556	334.3851
	0.3	345.4841	272.1959	399.2807	202.8125	201.6089	407.7525
	0.4	210.0684	247.0109	293.6007	126.5340	166.3455	258.0254
0.5	0.1	580.2067	342.7783	466.8862	226.7746	293.9675	578.2768
	0.2	285.6870	361.0272	502.0864	113.1070	213.5440	278.0264
	0.3	291.0438	367.5730	502.0864	203.8290	338.3325	395.0837
	0.4	205.2069	135.5575	362.7108	593.4900	188.9976	377.6263
0.4	0.1	209.5301	229.6845	521.6015	154.7853	219.4040	341.8411
	0.2	273.6734	243.8209	415.6563	154.7853	196.2041	409.3476
	0.3	186.4002	164.2299	180.1534	156.4836	286.9223	109.5825
	0.4	229.6534	452.6183	344.1958	122.8926	286.9223	259.2876
0.3	0.1	177.3793	348.7761	243.7971	207.1104	335.4644	475.7623
	0.2	161.7847	209.6734	196.5551	172.5561	583.4740	332.0107
	0.3	177.6283	321.8474	268.6266	284.1577	234.7380	233.3851
	0.4	118.5369	321.8474	180.0502	228.0400	212.2720	495.1331
0.2	0.1	173.7929	318.0804	485.0828	123.9482	320.5842	842.0112
	0.2	302.2107	366.5077	163.6724	118.1140	250.1179	300.6180
	0.3	195.3811	253.6954	247.4197	128.1928	388.1533	449.2047
	0.4	395.1311	422.3936	548.8376	475.0541	362.5715	186.8346

$$N = 51,\ S_x^2 = 2.3416\,e+09,\ S_y^2 = 1.5465\,e+09,$$
$$S_z^2 = 1.5128\,e+09,\ \bar{X} = 2.3416\,e+04$$
$$\bar{Y} = 2.4286\,e+04,\ \bar{Z} = 2.3963\,e+04,$$
$$\rho_{yx} = 0.9904,\ \rho_{xz} = 0.9987,\ \rho_{yz} = 0.9906.$$

Population–II: source Priyanka and Trisandhya.[12]

$$N = 315,\ S_x^2 = 1.2463 * 10^6,\ S_y^2 = 2.1926 * 10^6,$$
$$S_z^2 = 1.4670 * 10^7,\ \bar{X} = 370.5238$$
$$\bar{Y} = 504.8095,\ \bar{Z} = 4.0233 * 10^3,$$
$$\rho_{yx} = 0.8937,\ \rho_{xz} = 0.6491,\ \rho_{yz} = 0.7012.$$

Utilizing the above data, the percent relative efficiencies of the proposed estimator T_{prop} with respect to $\bar{h}_n$, $\hat{\bar{H}}$ and the Singh and Sharma[18] estimators are computed using the formula given below and presented in Tables 7 and 8.

$$E_1 = \frac{V(\bar{h}_n)}{M(T_{prop})^{(0)}_{opt}} \text{x}100,\ E_2 = \frac{V(\hat{\bar{H}})}{M(T_{prop})^{(0)}_{opt}} \text{x}100,$$
$$E_3 = \frac{M(T_{ss})_{opt}}{M(T_{prop})^{(0)}_{opt}} \text{x}100.$$

From Table 8, for models 1 and 2, it may be read that the percent relative efficiencies of the proposed estimator are decreasing with decreasing values of P_1 with respect to the sample mean estimator and natural successive estimator. The minimum value of μ_0 is 0.5352 which indicates that the fraction of fresh sample to be replaced at the current occasion is as low as about 53% of the total sample size, which reduces the cost of the survey. The values of percent relative efficiencies increase with the decreasing values of P_1 with respect to the Singh and Sharma[18] estimator.

From Table 9, for models 1 and 2, it may be read that the percent relative efficiencies of the proposed estimator decrease with the decreasing values of P_1 with respect to the sample mean estimator and natural successive estimator. The minimum value of μ_0 is 0.5112, which indicates that the fraction of fresh sample to be replaced at the current occasion is as low as about 51% of the total sample size, which reduces the cost of the survey. The values of percent relative efficiencies increase with decreasing value of P_1 with respect to the Singh and Sharma[18] estimator.

Table 8 Optimum Values μ_0 and PRE of T_{prop} with Respect to $\bar{h}_n$, $\hat{\bar{H}}$ and Singh and Sharma[18] for Population-I.

P1	P2	Model 1					Model 2				
		μ_0	E1	E2	μ_{0s}	E3	μ_0	E1	E2	μ_{0s}	E3
0.6	0.1	0.5352	289.1200	151.6900	0.5597	108.1200	0.5538	257.4100	135.4400	0.5538	122.0100
	0.2	0.5357	288.2600	151.2500	0.6220	117.2900	0.5424	276.5600	145.2600	0.5424	112.5800
	0.3	0.5479	267.1300	140.4200	0.7519	117.3400	0.5448	272.3900	143.1200	0.5448	114.6500
	0.4	0.5630	242.8700	127.9600	0.7404	120.4200	0.5630	242.8700	127.9600	0.5630	129.0600
0.5	0.1	0.5469	268.9400	141.3500	0.5661	117.9900	0.5644	240.8600	126.9300	0.5644	130.0200
	0.2	0.5423	276.7600	145.3600	0.6231	120.7600	0.5531	258.5600	136.0300	0.5531	121.4500
	0.3	0.5486	266.0900	139.8900	0.7182	176.5100	0.5486	266.0500	139.8700	0.5486	117.7800
	0.4	0.5600	247.5900	130.3900	0.7779	119.4200	0.5555	254.7800	134.0800	0.5555	123.2900
0.4	0.1	0.5569	252.4900	132.9100	0.5723	125.8200	0.5719	229.5100	121.0800	0.5719	135.4600
	0.2	0.5496	264.3000	138.9700	0.6214	125.3600	0.5620	244.5300	128.8200	0.5620	128.2600
	0.3	0.5514	261.4300	137.5000	0.6966	150.8600	0.5552	255.2300	134.3200	0.5552	123.0700
	0.4	0.5591	249.0300	131.1300	0.8290	121.2200	0.5553	255.0500	134.2300	0.5553	123.1600
0.3	0.1	0.5653	239.3500	126.1500	0.5778	131.9600	0.5775	221.3600	116.8700	0.5775	139.3400
	0.2	0.5569	252.5600	132.9400	0.6141	129.9300	0.5689	233.9800	123.3900	0.5689	133.3200
	0.3	0.5555	254.8300	134.1100	0.6836	142.8300	0.5618	244.8200	128.9600	0.5618	128.1200
	0.4	0.5598	247.8500	130.5200	0.7878	126.2600	0.5586	249.7600	131.5100	0.5586	125.7300
0.2	0.1	0.5724	228.8300	120.7300	0.5826	136.8200	0.5818	215.2700	113.7300	0.5818	142.2200
	0.2	0.5635	242.1500	127.5900	0.6100	134.0600	0.5743	225.9800	119.2600	0.5743	137.1400
	0.3	0.5601	247.5200	130.3600	0.6762	141.2200	0.5676	236.0000	124.4200	0.5676	132.3600
	0.4	0.5618	244.8300	128.9700	0.7533	139.6300	0.5631	242.8700	127.9600	0.5631	129.0600
0.1	0.1	0.5783	220.3300	116.3400	0.5868	140.7200	0.5852	210.5700	111.3000	0.5852	144.4400
	0.2	0.5695	233.1500	122.9600	0.6078	137.6700	0.5787	219.7700	116.0500	0.5787	140.0900
	0.3	0.5648	240.2600	126.6200	0.6684	141.6400	0.5724	228.7800	120.7000	0.5724	135.8000
	0.4	0.5645	240.7100	126.8500	0.8600	180.7100	0.5675	236.1100	124.4800	0.5675	132.3000

Scrambling Mechanism versus Direct Questioning

If no scrambled response technique has been used, then the estimator under the direct method is given as

$$T_{du}=\bar{y}_u \exp\left(\frac{\bar{Z}_1-\bar{z}_{1u}}{\bar{Z}_1+\bar{z}_{1u}}\right) \tag{33}$$

Another estimator based on a sample of size m common to both occasions is a modified exponential-type estimator and structured as

$$T_{dm}= \bar{y}_m^* \exp\left(\frac{\bar{x}_n-\bar{x}_m}{\bar{x}_n+\bar{x}_m}\right)\exp\left(\frac{\bar{Z}_1-\bar{z}_{1n}}{\bar{Z}_1+\bar{z}_{1n}}\right) where\ \bar{y}_m^*=\frac{\bar{y}_m}{\bar{x}_m}\,\bar{x}_n \tag{34}$$

The final estimator is a convex linear combination of two estimators ξ_u and ξ_m shown as

$$T_{propd} = \phi\, T_{du} + (1-\phi)\, T_{dm} \tag{35}$$

The minimum mean square error of the estimator up to first-order approximations is given as

$$M\left(T_{propd}\right)_{opt} = \left[\frac{(1-f)\,A_4 - A_5\,\mu + A_6\,\mu^2}{A_3 - A_2\,\mu^2}\right]\frac{S_y^2}{n} \tag{36}$$

where

$$A_1=\left(\frac{13}{4}-3\rho_{yx}\right), A_2=\left(3\rho_{yx}-\rho_{yz_1}-2\right), A_3=\left(\frac{5}{4}-\rho_{yz_1}\right),$$

$$A_3=(A_1+A_2),\ A_3^2=A_4,\ A_2\,A_3=A_5$$

$$A_2\,A_3\,f=A_6, f=(n/N)\ \text{and}\ \mu=(u/n)$$

The percent relative efficiencies have been computed using data in Estimators for Sensitive Population Mean at Current Occasion Under Proposed Scrambled Response Models for different choices of probabilities and are graphically presented in Figures 9 and 10.

Table 9 Optimum Values μ_0 and PRE of T_{prop} with Respect to $\bar{h}_n$, $\hat{\hat{H}}$ and Singh and Sharma[18] for Population-II.

P_1	P_2	Model 1					Model 2				
		μ_0	E_1	E_2	μ_{0s}	E_3	μ_0	E_1	E_2	μ_{0s}	E_3
0.6	0.1	0.5112	166.4000	115.0700	0.5157	142.2300	0.5196	159.4100	110.0400	0.5196	145.9100
	0.2	0.5114	166.2300	114.9500	0.5278	143.5400	0.5143	163.7900	113.1900	0.5143	143.3200
	0.3	0.5168	161.7000	111.6900	0.5488	147.9000	0.5154	162.8800	112.5400	0.5154	143.8600
	0.4	0.5240	155.7100	107.4000	0.5720	153.2800	0.5240	155.7100	107.4000	0.5240	148.0700
0.5	0.1	0.5163	162.1100	111.9900	0.5199	144.6800	0.5248	155.1700	107.0000	0.5248	148.4000
	0.2	0.5143	163.8400	113.2300	0.5281	144.7100	0.5192	159.6800	110.2400	0.5192	145.7400
	0.3	0.5171	161.4600	111.5200	0.5451	147.6200	0.5171	161.4500	111.5100	0.5171	144.7000
	0.4	0.5225	156.9500	108.2800	0.5660	152.0200	0.5203	158.7600	109.5900	0.5203	146.2800
0.4	0.1	0.5210	158.2000	109.1800	0.5240	146.9200	0.5287	151.9800	104.7200	0.5287	150.2700
	0.2	0.5176	161.0500	111.2200	0.5240	146.9200	0.5235	156.1500	107.7000	0.5235	147.8200
	0.3	0.5184	160.3700	110.7400	0.5427	147.8800	0.5202	158.8700	109.6600	0.5202	146.2200
	0.4	0.5221	157.3200	108.5500	0.5609	151.2900	0.5202	158.8300	109.6300	0.5202	146.2400
0.3	0.1	0.5252	154.7600	106.7200	0.5277	148.9000	0.5318	149.5400	102.9700	0.5318	151.7000
	0.2	0.5210	158.2100	109.1900	0.5308	147.6300	0.5271	153.2700	105.6400	0.5271	149.5200
	0.3	0.5203	158.7800	109.6000	0.5414	148.4900	0.5234	156.2200	107.7600	0.5234	147.7800
	0.4	0.5224	157.0200	108.3300	0.5569	151.0000	0.5219	157.5000	108.6800	0.5219	147.0200
0.2	0.1	0.5289	151.7900	104.5900	0.5310	150.6000	0.5342	147.6300	101.6100	0.5342	152.8200
	0.2	0.5243	155.5200	107.2600	0.5326	149.0700	0.5300	150.9400	103.9700	0.5300	150.8800
	0.3	0.5225	156.9300	108.2700	0.5408	149.2900	0.5264	153.8300	106.0400	0.5264	149.1800
	0.4	0.5234	156.2300	107.7700	0.5539	151.0300	0.5241	155.7100	107.3900	0.5241	148.0800
0.1	0.1	0.5322	149.2300	102.7600	0.5339	152.0700	0.5362	146.1000	100.5100	0.5362	153.7200
	0.2	0.5274	153.0400	105.4800	0.5345	150.4100	0.5324	149.0500	102.6200	0.5324	151.9900
	0.3	0.5249	155.0100	106.8900	0.5407	150.1800	0.5290	151.7700	104.5700	0.5290	150.3900
	0.4	0.5248	155.1300	106.9800	0.5516	151.3000	0.5264	153.8600	106.0700	0.5264	149.1600

Interpretation of Results

The following interpretations may be read from Tables 2–9 and Figures 1–10.

(I) From Tables 2–5, it may be seen that the PREs of the proposed randomized response models with respect to the Bar-Lev et al[3] model are always greater than 100, which indicates the dominating behaviors of the proposed models over the Bar–Lev et al[3] model.

(II) From Tables 2–5, it is clear that all the values of τ are greater than 0.5 and closer to 1. Hence, more privacy is protected and greater cooperation may be expected for the proposed models.

(III) From Figures 1–8, it is visible that the values of unified measure of model quality for the proposed models are smaller than for the Bar-Lev et al[3] model. Hence, the proposed models are better than the contemporary randomized response models given by Bar-Lev et al.[3]

(IV) From Tables 7–9, it may be seen that the suggested estimator under the proposed scrambled response models performs better than sample mean estimator, natural successive sampling estimator and Singh and Sharma[18] estimators.

(V) From Figures 9 and 10, it is visible that the proposed estimator under the scrambled response models when compared with the direct method performs better in terms of percent relative efficiency.

Conclusions and Recommendations

From the above results and interpretations we may conclude that the proposed models are uniformly dominating over the Bar-Lev et al[3] model in terms of enhanced precision of estimates. It has been found that to deal with extremely

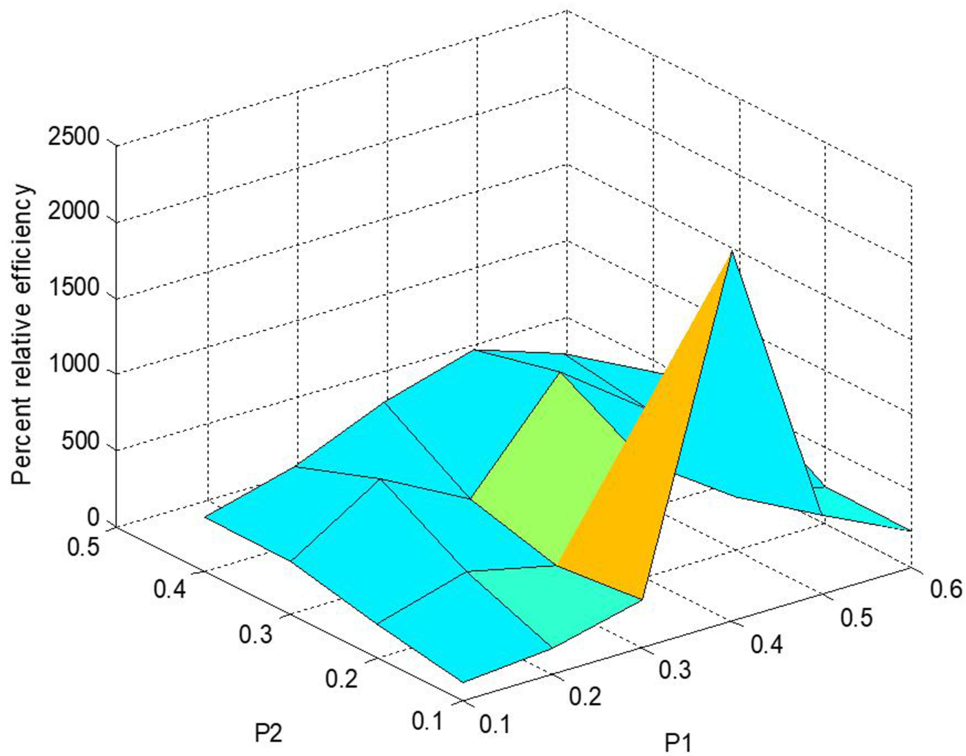

Figure 9 Percent relative efficiency E_{1d} for Model 1.

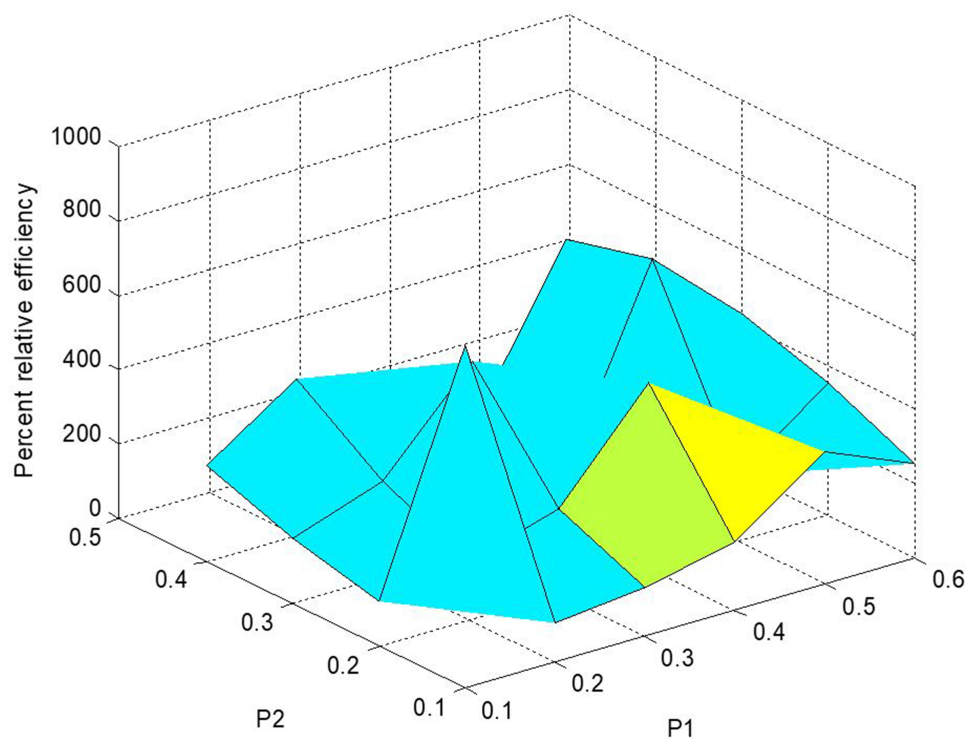

Figure 10 Percent relative efficiency E_{1d} for Model 2.

sensitive questions, the proposed models are suitable with a high degree of privacy protection. The suggested estimator accomplishes good percent relative efficiency. Thus, the proposed randomized response models and proposed estimators may be recommended to survey practitioners encouragingly for use in the real life problems, whenever they intend to deal with the quantitative sensitive characteristics and for estimating the sensitive population mean when it changes frequently with the passage of time according to its nature.

Acknowledgments

The authors are thankful to the Indian Institute of Technology (Indian School of Mines), Dhanbad and College of Science and Theoretical Studies, Saudi Electronic University, KSA for providing financial and necessary infrastructural support to carry out the present research work. Authors are also thankful to the honorable reviewers, honorable editor and honorable editorial board for their valuable suggestions which improved the quality of the manuscript.

Disclosure

The authors reported no conflicts of interest for this work.

References

1. Arnab R, Singh S. Estimation of mean of sensitive characteristics for successive sampling. *Commun Stat Theory Methods*. 2013;42(14):2499–2524. doi:10.1080/03610926.2011.605233
2. Bahl S, Tuteja RK. Ratio and product type exponential estimators. *J Inf Optim Sci*. 1991;12(1):159–164. doi:10.1080/02522667.1991.10699058
3. Bar-Lev SK, Bobovitch E, Boukai B. A note on randomized response models for quantitative data. *Metrika*. 2004;60(3):255–260. doi:10.1007/s001840300308
4. Diana G, Perri PF. New scrambled response models for estimating the mean of a sensitive quantitative character. *J Appl Stat*. 2010;37(11):1875–1890. doi:10.1080/02664760903186031
5. Diana G, Perri PF. A class of estimators for quantitative sensitive data. *Stat Pap*. 2011;52(3):633–650. doi:10.1007/s00362-009-0273-1
6. Eichhorn BH, Hayre LS. Scrambled randomized response methods for obtaining sensitive quantitative data. *J Stat Plan Inference*. 1983;7(4):307–316. doi:10.1016/0378-3758(83)90002-2
7. Greenberg BG, Kuebler RR, Abernathy JR, Horvitz DG. Application of the randomized response technique in obtaining quantitative data. *J Am Stat Assoc*. 1971;66(334):243–250. doi:10.1080/01621459.1971.10482248
8. Gupta S, Mehta S, Shabbir J, Khalil S. A unified measure of respondent privacy and model efficiency in quantitative RRT models. *J Stat Theory Pract*. 2018;12(3):506–511. doi:10.1080/15598608.2017.1415175
9. Jessen RJ. Statistical investigation of a sample survey for obtaining farm facts. *Iowa Agric Home Econ Exp Station Res Bull*. 1942;26(304):1–104.
10. Kim JM, Elam ME. A stratified unrelated question randomized response model. *Stat Pap*. 2007;48(2):215–233. doi:10.1007/s00362-006-0327-6
11. Murthy MN. *Sampling Theory and Methods*. Calcutta: Statistical Publication Society; 1967.
12. Priyanka K, Trisandhya P. Item sum techniques for quantitative sensitive estimation on successive occasions. *Commun Stat Appl Methods*. 2019;26(2):175–189. doi:10.29220/CSAM.2019.26.2.175
13. Priyanka K, Mittal R. New approaches using exponential type estimator with cost modelling for population mean on successive waves. *Stat Trans New Series*. 2017;18(4):569–587. doi:10.21307/stattrans-2017-001
14. Priyanka K, Mittal R, Kim JM. Multivariate rotation design for population mean in sampling on successive occasions. *Commun Stat Appl Methods*. 2015;22(5):445–462. doi:10.5351/CSAM.2015.22.5.445
15. Pollock KH, Bek Y. A comparison of three randomized response models for quantitative data. *J Am Stat Assoc*. 1976;71(356):884–886. doi:10.1080/01621459.1976.10480963
16. Reddy VN. A study on the use of prior knowledge on certain population parameters in estimation. *Sankhya C*. 1978;40:29–37.
17. Son CK, Kim JM. Calibration estimation of adjusted Kuk's randomized response model for sensitive attribute. *Braz J Prob Stat*. 2017;31(1):160–178. doi:10.1214/16-BJPS307
18. Singh GN, Sharma AK. An alternative rotation patterns in two occasion successive sampling. *Int J Math Stat*. 2014;15(3):9–22.
19. Singh GN, Homa F. Effective rotation patterns in successive sampling over two occasions. *J Stat Theory Pract*. 2013;7(1):146–155. doi:10.1080/15598608.2012.755484
20. Singh GN, Singh C, Pandey AK, Suman S. An improved exponential method of estimation for current population mean in two-occasion successive sampling. *J Stat Appl Prob*. 2017;6(3):587–595. doi:10.18576/jsap/060314
21. Singh HP, Pal SK. Improved estimation of current population mean over two occasions. *Sri Lankan J Appl Stat*. 2015;16(1):1–19. doi:10.4038/sljastats.v16i1.7804
22. Singh GN, Homa F, Maurya S. Exponential method of estimation in two-occasion successive sampling. *Int J Stat Econ*. 2013;12(3):26–39.
23. Singh HP, Tarray T. An improved randomized response additive model. *Sri Lankan J Appl Stat*. 2014;15(2):131–138. doi:10.4038/sljastats.v15i2.7412
24. Warner SL. Randomized response: a survey technique for eliminating evasive answer bias. *J Am Stat Assoc*. 1965;60(309):63–69. doi:10.1080/01621459.1965.10480775