

Partial optional randomized response technique with calibration weighting to adjust non-response in successive sampling

Kumari Priyanka^{1,a}, Pidugu Trisandhya^b, Ajay Kumar^b

^aDepartment of Mathematics, Shivaji College (University of Delhi), India

^bDepartment of Mathematics, University of Delhi, India

Abstract

The present article endeavours to develop partial optional randomized - response technique (PORT) to deal with sensitive issues in presence of non-response in successive sampling. Calibration techniques have been embedded with PORT to estimate sensitive population mean at current move in two move successive sampling in presence of non-response. Optimum calibration weights are computed at each move with the aid of constraints based on auxiliary information. Detailed properties of the proposed estimators have been discussed. Possible cases in which non-response may creep at two moves has been explored. The proposed technique has been compared with the modified existing technique. Simulation results indicate that the proposed technique is more efficient than existing, modified one. Suitable recommendations are forwarded.

Keywords: partial optional randomized - response technique, sensitive variable, successive moves, calibration estimators, non-response, population mean, auxiliary information

1. Introduction

In many social surveys, people often do not respond genuinely when socially sensitive questions are asked. To account for this behaviour, different techniques were introduced by many statisticians to reduce no reporting, under-reporting or over reporting.

The sensitive issues may be matters of medical/legal malpractices, addiction of drugs, criminal conviction, induced abortions, acid attacks, domestic violence, etc. Two widely practised indirect questioning techniques dealing with these issues are randomized response technique (RRT) and scrambled response technique (SRT) that protect the privacy of respondents and mask the sensitive response, thereby motivating the respondents to give accurate response.

Warner (1965) was the first to provide such a randomizing model and followed by sizeable literature was added by Horvitz *et al.* (1967), Greenberg *et al.* (1971), Christofides (2003), Kim and Elam (2007), Wu *et al.* (2008), Yan *et al.* (2009), Arnab (2011), Diana and Perri (2011), Arnab *et al.* (2012). The SRT was initiated by Pollock and Bek (1976). Further, Eichhorn and Hayre (1983), Saha (2007), Diana and Perri (2010, 2011), Perri and Diana (2013) and Hussain and Al-Zahrani (2016) added substantive literature in this area. In RRT or SRT the respondent need to provide randomized or scrambled response. However, depending on the thought of respondents it might be possible that some issue may be sensitive to one respondent but the same issue may not be sensitive to other.

¹ Corresponding author: Department of Mathematics, Shivaji College (University of Delhi), New Delhi 110 027, India.
priyanka.ism@gmail.com

Hence, to address these problems optional randomized response technique (ORT) may be used. ORT can be classified into two categories: full optional randomized response technique (FORT) and partial optional randomized response technique (PORT). The ORT is more efficient than compulsory randomized response technique (CRT) because the probability of obtaining true responses in ORT is much higher than that in the CRT (Arnab, 2004). PORT was discussed by many authors such as Mangat and Singh (1994), Gupta *et al.* (2002, 2013), Pal (2008), Chaudhuri and Dihidar (2009), and Sanaullah *et al.* (2020).

It is common that the sensitive variable may be dynamic over time, in such a situation one point survey may not be sufficient. Continuous monitoring of sensitive variable may be required. The dynamics of such sensitive variables may be studied using successive sampling. Addressing the sensitive variable, Arnab and Singh (2013), Yu *et al.* (2015), Naeem and Shabbir (2016), Singh *et al.* (2017), Priyanka *et al.* (2018), Priyanka and Trisandhya (2019a, 2019b), Priyanka *et al.* (2019), Singh *et al.* (2018) contributed rich literature. These researchers used simple random sampling design in successive sampling and used either RRT or SRT to deal with sensitive issues. However, no attempt has been made to estimate sensitive population mean using PORT in successive sampling.

Generally, all sample surveys are affected by the problem of non-response. When the issues are sensitive, then they are more prone to occur and can severely affect the validity and generalizability of the results. Non-response are generally of two types namely unit non-response and item non-response. In unit non-response, sampled unit fails to respond completely, however in item non-response, the sampled unit responds to the survey but fails to respond to a particular question.

Hence, before proceeding with any method the kind of non-response creeping in the survey must be identified and suitable measure must be devices.

Therefore, in this paper an attempt has been made to estimate sensitive population mean at current move using PORT with calibration weighting to adjust unit non-response in two move successive sampling. A new model for PORT has been proposed and the existing model by Sanaullah *et al.* (2020) have been modified to work as PORT model for successive sampling. The paper is structured as follows, in Section 2, the proposed PORT models on two successive moves have been presented along with the calibration estimators in presence of unit non-response at both the moves for the coded response variable at current move. In Section 3, the asymptotic behaviour of the estimators developed in previous section have been studied. Section 4 is devoted to discuss the study under simple random sampling without replacement. In Section 5, the possible cases in which the non-response may creep has been discussed. Section 6 presents the corresponding estimators for sensitive population mean along with the expression for their variances. Section 7 is dedicated to a simulation study to compare the precision of the proposed estimators with the usual modified estimator and the comparison of the proposed PORT models. Finally, Section 8 is devoted to the discussion of result and conclusion.

2. Survey design and development of estimator

2.1. Formulation and notation

A finite population $U = (U_1, U_2, \dots, U_N)$ of N identifiable units is considered for sampling over two moves. The size of population is constant while the value of units do change over the moves. Let the sensitive study character be denoted by y_1 the first move and y_2 the second move. It is assumed that the information on non-sensitive auxiliary variable x , whose population mean is known and stable over moves, is readily available on both the moves. In two moves successive sampling, we intend to focus on the probability design to draw relevant samples on different moves under unit non-response for the estimation of population mean of sensitive variable. At the first move, a sample s_n of size n is

Table 1: Inclusion probabilities

| Size | First order inclusion probability | Second order inclusion probability |
|-------|--|---|
| r_1 | $\pi_{1i} = \sum_{i \in s_{r_1}} P_{d_1}(s_{r_1})$ | $\pi_{1ij} = \sum_{i, j \in s_{r_1}} P_{d_1}(s_{r_1})$ |
| m | $\pi_i s_{r_1} = \sum_{i \in s_m} P_{d_2}(s_m)$ | $\pi_{ij} s_{r_1} = \sum_{i, j \in s_m} P_{d_2}(s_m)$ |
| r_2 | $\pi_i = \sum_{i \in s_{r_2}} P_{d_3}(s_{r_2})$ | $\pi_{ij} = \sum_{i, j \in s_{r_2}} P_{d_3}(s_{r_2})$ |

Table 2: Sample design basic weights

| Response Set | Size | Sampling design basic weight for selecting i^{th} unit in corresponding sample |
|--------------|-------|--|
| s_{r_1} | r_1 | $\alpha_{1i} = \frac{1}{\pi_{1i}}$ |
| s_m | m | $\alpha_i^* = \frac{1}{\pi_{1i} \pi_{i s_{r_1}}}$ |
| s_{r_2} | r_2 | $\beta_i^* = \frac{1}{(\pi_{1i})^c \pi_i s_m^c}$ |

drawn with the design d_1 having probability $p_{d_1}(s_n)$. We assume that s_{r_1} of size r_1 be the response set from s_n ; $s_{r_1} \subset s_n \subset U$, having probability $p_{d_1}(s_{r_1})$. A random sub-sample s_m of m units from s_{r_1} (the set of responding units at first move) are preserved with design d_2 having the probability $p_{d_2}(s_m)$ to be used at current move. Next, an unmatched sample s_u of size u is drawn afresh (without replacement) at second move with the design d_3 having probability $p_{d_3}(s_u)$. Let the response set r_2 of size s_{r_2} is obtained from s_u ; $s_{r_2} \subset s_u \subset (U - s_n)$ having the probability $p_{d_3}(s_{r_2})$. Let S_1 and S_2 be two scrambling variables that will be used to code the response for the sensitive variables. It is assumed that two scrambling variables are independent of each other and their mean and variance are known. Further, let $z_1(z_2)$ be the coded response variable corresponding to sensitive variables $y_1(y_2)$ on two moves. The first order and second order positive inclusion probabilities for different samples on two moves are shown in Table 1 and sampling design basic weights are presented in Table 2.

Let the known mean and variance of scrambling variables be assumed as $E(S_1) = \bar{S}_1, E(S_2) = \bar{S}_2, V(S_1) = \sigma_{S_1}^2, V(S_2) = \sigma_{S_2}^2$.

Based on the considered sampling design, we intend to apply partial optional randomization technique (PORT) on successive moves to handle sensitivity of study variable.

2.2. Partial optional randomization technique on successive moves

Motivated by recent work of Sanaullah *et al.* (2020), we intend to modify their generalized randomised response model to be applicable for successive sampling under PORT. The model is modified as,

• **PORT-I: Modified model** (Sanaullah *et al.*, 2020)

The response obtained from the j^{th} respondent on first and second move are respectively given as,

$$z_{1j} = \begin{cases} y_{1j}, & \text{with probability } p, \\ y_{1j}S_{1j} + aS_{2j}, & \text{with probability } (1 - p), \end{cases} \tag{2.1}$$

$$z_{2j} = \begin{cases} y_{2j}, & \text{with probability } p, \\ y_{2j}S_{1j} + aS_{2j}, & \text{with probability } (1 - p), \end{cases} \tag{2.2}$$

where $a \in (-\infty, \infty)$ is a suitable constant chosen by the investigator (Himmelfarb and Edgell, 1980).

Taking expectation on both sides of equation (2.1) and (2.2) respectively, we get

$$\bar{Z}_1 = p\bar{Y}_1 + (1-p)\left[\bar{Y}_1\bar{S}_1 + a\bar{S}_2\right],$$

$$\bar{Z}_2 = p\bar{Y}_2 + (1-p)\left[\bar{Y}_2\bar{S}_1 + a\bar{S}_2\right].$$

Therefore, the population mean of sensitive variable at current move is given as

$$\bar{Y}_2 = \frac{\bar{Z}_2 - (1-p)a\bar{S}_2}{p + (1-p)\bar{S}_1}, \quad (2.3)$$

such that

$$\rho_{z_1 z_2} = \frac{p^2(\rho_{y_1 y_2} \sigma_{y_1} \sigma_{y_2}) + 2p(1-p)(\rho_{y_1 y_2} \sigma_{y_1} \sigma_{y_2} \bar{S}_1) + (1-p)^2 \left[(\rho_{y_1 y_2} \sigma_{y_1} \sigma_{y_2}) (\sigma_{s_1}^2 + \bar{S}_1^2) + \bar{Y}_1 \bar{Y}_2 \sigma_{s_1}^2 \right]}{\sqrt{I_1} \sqrt{I_2}},$$

$$\rho_{z_1 y_1} = \frac{(p + (1-p)\bar{S}_1) \sigma_{y_1}}{\sqrt{I_1}},$$

$$\rho_{z_2 y_2} = \frac{(p + (1-p)\bar{S}_1) \sigma_{y_2}}{\sqrt{I_2}},$$

$$I_1 = p^2 \sigma_{y_1}^2 + (1-p)^2 \left[\sigma_{y_1}^2 (\sigma_{s_1}^2 + \bar{S}_1^2) + \sigma_{s_1}^2 \bar{Y}_1^2 + a^2 \sigma_{s_2}^2 \right],$$

$$I_2 = p^2 \sigma_{y_2}^2 + (1-p)^2 \left[\sigma_{y_2}^2 (\sigma_{s_1}^2 + \bar{S}_1^2) + \sigma_{s_1}^2 \bar{Y}_2^2 + a^2 \sigma_{s_2}^2 \right].$$

• PORT-II: Proposed model

Since the modified Sanaullah *et al.* (2020) model involves an unknown constant to be chosen by the investigator. This selection of constant may involve an added source of randomness in the model. Hence, we intend to propose a model for PORT on successive moves which is independent of arbitrary chosen constant. The response received from the j^{th} respondent under the proposed PORT model on first and second move are respectively given as

$$z_{1j} = \begin{cases} y_{1j}, & \text{with probability } p, \\ \frac{y_{1j}S_{1j} + S_{2j} - \bar{S}_2}{\bar{S}_1} & \text{with probability } (1-p), \end{cases} \quad (2.4)$$

$$z_{2j} = \begin{cases} y_{2j}, & \text{with probability } p, \\ \frac{y_{2j}S_{1j} + S_{2j} - \bar{S}_2}{\bar{S}_1}, & \text{with probability } (1-p). \end{cases} \quad (2.5)$$

On taking expectation on both sides of equation (2.4), we observe that

$$\begin{aligned} E(z_{1j}) &= pE[y_{1j}] + (1-p)E\left[\frac{y_{1j}S_{1j} + S_{2j} - \bar{S}_2}{\bar{S}_1}\right], \\ &= p\bar{Y}_1 + (1-p)\left[\bar{Y}_1 + \frac{\bar{S}_2}{\bar{S}_1} - \frac{\bar{S}_2}{\bar{S}_1}\right], \\ \implies \bar{Z}_1 &= \bar{Y}_1. \end{aligned}$$

Similarly, taking expectations on both the sides of equation (2.5) we get the population mean of sensitive variable at current move as

$$\bar{Y}_2 = \bar{Z}_2, \quad (2.6)$$

such that

$$\rho_{z_1 z_2} = \frac{p^2(\rho_{y_1 y_2} \sigma_{y_1} \sigma_{y_2}) + p(1-p) [2\rho_{y_1 y_2} \sigma_{y_1} \sigma_{y_2} + \bar{Y}_1 \bar{Y}_2] + (1-p)^2 [I_3]}{\sqrt{I_4} \sqrt{I_5}},$$

$$I_3 = \frac{(\rho_{y_1 y_2} \sigma_{y_1} \sigma_{y_2} + \bar{Y}_1 \bar{Y}_2)(\sigma_{s_1}^2 + \bar{S}_1^2) + \sigma_{s_2}^2 + \bar{S}_2^2 + \sigma_{s_2}^2/n}{\sigma_{s_1}^2/n + \bar{S}_1^2} + \bar{S}_2 \bar{Y}_2 - \bar{Y}_1 \bar{Y}_2,$$

$$\rho_{z_1 y_1} = \frac{p\sigma_{y_1}^2 + (1-p)\sigma_{y_1}^2}{\sigma_{y_1} \sqrt{I_4}},$$

$$\rho_{z_2 y_2} = \frac{p\sigma_{y_2}^2 + (1-p)\sigma_{y_2}^2}{\sigma_{y_2} \sqrt{I_5}},$$

$$I_4 = p^2\sigma_{y_1}^2 + (1-p)^2 \left[\frac{(\sigma_{s_1}^2 + \bar{S}_1^2)(\sigma_{y_1}^2 + \bar{Y}_1^2)}{\sigma_{s_1}^2/n + \bar{S}_1^2} - \bar{Y}_1^2 + \frac{\sigma_{s_2}^2 + \bar{S}_2^2}{\sigma_{s_1}^2/n + \bar{S}_1^2} - \frac{\sigma_{s_2}^2/n + \bar{S}_2^2}{\sigma_{s_1}^2/n + \bar{S}_1^2} \right] + 2p(1-p)\sigma_{y_1}^2,$$

$$I_5 = p^2\sigma_{y_2}^2 + (1-p)^2 \left[\frac{(\sigma_{s_1}^2 + \bar{S}_1^2)(\sigma_{y_2}^2 + \bar{Y}_2^2)}{\sigma_{s_1}^2/n + \bar{S}_1^2} - \bar{Y}_2^2 + \frac{\sigma_{s_2}^2 + \bar{S}_2^2}{\sigma_{s_1}^2/n + \bar{S}_1^2} - \frac{\sigma_{s_2}^2/n + \bar{S}_2^2}{\sigma_{s_1}^2/n + \bar{S}_1^2} \right] + 2p(1-p)\sigma_{y_2}^2.$$

Remark 1. The mean and variance of sensitive variable at current move in two move successive sampling are obtained in terms of mean of coded response variable. Hence, efficient estimators need to be investigated to estimate coded response variable so that the estimate of sensitive variable get improved and became more effective. Hence, in next section we investigate the suitable estimators for coded response variable in presence of non-response at both the moves.

Remark 2. Since the study character is sensitive in nature. Even though the investigator try so hard, there will always be scope for some non-response. Hence, in order to deal with non-response, calibration technique applied over successive moves may be a good alternative. The calibration technique becomes more effective if auxiliary information is available and in successive sampling the information from previous move may also be used as an auxiliary information at current move together with the availability of additional auxiliary variable. Hence, in the next section a weighted calibration estimator for coded response variable have been proposed to adjust the effect due to presence of non-response.

2.3. Calibration estimator for coded response variable

Devil and Särndal (1992) invoked calibration technique in survey sampling, which is proved to be an efficient technique to adjust non-response by Lundström and Särndal (1999). Therefore, calibration technique has been used to adjust non-response with the aid of an additional auxiliary variable to estimate coded response variable which will be further used to estimate population mean of sensitive variable.

2.3.1. Calibration estimator based on fresh sample

Let the basic design weight β_i^* be replaced by the new weight w_{r_2} . Hence, for the estimation of sensitive population mean at current move, the proposed calibration estimator \hat{T}_{cu}^{NR} based on fresh

sample of size u is given as

$$\hat{T}_{cu}^{NR} = \frac{1}{N} \sum_{i \in s_{r_2}} w_{r_{2i}} z_{2i}, \quad (2.7)$$

In order to obtain the calibrated weight $w_{r_{2i}}$, we minimize the chi-square type function

$$\sum_{i \in s_{r_2}} \frac{(w_{r_{2i}} - \beta_i^*)^2}{q_{ui} \beta_i^*}, \quad (2.8)$$

subject to calibration constraint

$$\frac{1}{N} \sum_{i \in s_{r_2}} w_{r_{2i}} x_i = \bar{X}, \quad (2.9)$$

with q_{ui} being known positive constant unrelated to β_i^* , x_i and \bar{X} . This results in to calibrated weights given by

$$w_{r_{2i}} = \beta_i^* + \beta_i^* q_{ui} N \left[\frac{(\bar{X} - \frac{1}{N} \sum_{i \in s_{r_2}} \beta_i^* x_i) x_i}{\sum_{i \in s_{r_2}} q_{ui} \beta_i^* x_i^2} \right]. \quad (2.10)$$

After substituting the calibrated weights $w_{r_{2i}}$ in equation (2.7) we obtain the final calibrated estimator \hat{T}_{cu}^{NR} as,

$$\hat{T}_{cu}^{NR} = \left[\frac{1}{N} \sum_{i \in s_{r_2}} \beta_i^* z_{2i} + \hat{b}_u \left(\bar{X} - \frac{1}{N} \sum_{i \in s_{r_2}} \beta_i^* x_i \right) \right], \quad (2.11)$$

with

$$\hat{b}_u = \left(\sum_{i \in s_{r_2}} \beta_i^* q_{ui} x_i^2 \right)^{-1} \left(\sum_{i \in s_{r_2}} \beta_i^* q_{ui} x_i z_{2i} \right),$$

where q_{ui} being known positive constant unrelated to β_i^* and \bar{X} .

2.3.2. Calibration estimator based on matched sample

In successive sampling, to ameliorate the performance of the estimators on the current move, it is quotidian practice to use the information collected on the first move as auxiliary information in addition to availability of additional non-sensitive auxiliary variable. The calibration estimator in presence of non-response is proposed based on sample of size m at current (second) move with the new calibrated weights as

$$\hat{T}_{cm}^{NR} = \frac{1}{N} \sum_{i \in s_m} w_{mi} z_{2i}. \quad (2.12)$$

To find the calibrated weight w_{mi} , we minimize the chi-square function

$$\sum_{i \in s_m} \frac{(w_{mi} - \alpha_i^*)^2}{q_{mi} \alpha_i^*}, \tag{2.13}$$

subject to calibration constraints

$$\frac{1}{N} \sum_{i \in s_m} w_{mi} z_{1i} = \bar{z}_{1n}^{*NR}, \tag{2.14}$$

$$\frac{1}{N} \sum_{i \in s_m} w_{mi} x_i = \bar{X}, \tag{2.15}$$

with q_{mi} as known positive constant unrelated to α_i^* and \bar{X} . Following similar procedure, we obtain the calibrated estimator for \bar{z}_{1n}^{*NR} based on sample s_{r_1} , the response set obtained from s_n drawn at first move as

$$\bar{z}_{1n}^{*NR} = \left[\frac{1}{N} \sum_{i \in s_{r_1}} \alpha_{1i} z_{1i} + \left(\bar{X} - \frac{1}{N} \sum_{i \in s_{r_1}} \alpha_{1i} x_i \right) \hat{b}_n \right], \tag{2.16}$$

with

$$\hat{b}_n = \left(\sum_{i \in s_{r_1}} \alpha_{1i} q_{ni} x_i^2 \right)^{-1} \left(\sum_{i \in s_{r_1}} \alpha_{1i} q_{ni} x_i z_{1i} \right),$$

where q_{ni} is known positive constant unrelated to α_{1i} and \bar{X} .

Now, minimizing the chi-square function in equation (2.13) subject to constraints in equation (2.14) and (2.15) lead to the calibrated weights given by

$$w_{mi} = \alpha_i^* + \left[\frac{1}{N} x_i \alpha_i^* q_{mi} \left\{ \left(\bar{X} - \frac{1}{N} \sum_{i \in s_m} \alpha_i^* x_i \right) \left(\sum_{i \in s_m} \alpha_i^* q_{mi} z_{1i}^2 \right) - \left(\bar{z}_{1n}^{*NR} - \frac{1}{N} \sum_{i \in s_m} \alpha_i^* z_{1i} \right) \left(\sum_{i \in s_m} \alpha_i^* q_{mi} z_{1i} x_i \right) \right\} + \frac{1}{N} z_{1i} q_{mi} \alpha_i^* \left\{ \left(\bar{z}_{1n}^{*NR} - \frac{1}{N} \sum_{i \in s_m} \alpha_i^* z_{1i} \right) \left(\sum_{i \in s_m} \alpha_i^* q_{mi} x_i^2 \right) - \left(\bar{X} - \frac{1}{N} \sum_{i \in s_m} \alpha_i^* x_i \right) \left(\sum_{i \in s_m} \alpha_i^* q_{mi} z_{1i} x_i \right) \right\} \right] \hat{b}_m, \tag{2.17}$$

where

$$\hat{b}_m = \left[\left(\sum_{i \in s_m} \alpha_i^* q_{mi} (z_{1i})^2 \right) \left(\frac{1}{N^2} \sum_{i \in s_m} \alpha_i^* q_{mi} x_i^2 \right) - \frac{1}{N^2} \left(\sum_{i \in s_m} \alpha_i^* q_{mi} z_{1i} x_i \right)^2 \right]^{-1}.$$

Substituting the calibrated weight w_{mi} in equation (2.12), the final proposed calibrated estimator \hat{T}_{cm}^{NR} based on sample size m at current move becomes

$$\hat{T}_{cm}^{NR} = \left[\frac{1}{N} \sum_{i \in s_m} \alpha_i^* z_{2i} + \hat{b}_{m1} \left(\bar{z}_{1n}^{*NR} - \frac{1}{N} \sum_{i \in s_m} \alpha_i^* z_{1i} \right) + \hat{b}_{m2} \left(\bar{X} - \frac{1}{N} \sum_{i \in s_m} \alpha_i^* x_i \right) \right], \tag{2.18}$$

with

$$\hat{b}_{m1} = \left(\frac{1}{N^2} \sum_{i \in s_m} \alpha_i^* q_{mi} z_{1i} z_{2i} \right) \left[\left(\sum_{i \in s_m} \alpha_i^* q_{mi} x_i^2 \right) - \left(\sum_{i \in s_m} \alpha_i^* q_{mi} z_{1i} x_i \right) \right] \hat{b}_m,$$

$$\hat{b}_{m2} = \left(\frac{1}{N^2} \sum_{i \in s_m} \alpha_i^* q_{mi} x_i z_{2i} \right) \left[\left(\sum_{i \in s_m} \alpha_i^* q_{mi} (z_{1i})^2 \right) - \left(\sum_{i \in s_m} \alpha_i^* q_{mi} z_{1i} x_i \right) \right] \hat{b}_m,$$

q_{mi} being known positive constant unrelated to α_i^* and \bar{z}_{1n}^{*NR} .

2.3.3. Combined calibrated estimator

The final calibrated estimator in presence of non-response at both moves is considered as convex linear combination of the two calibrated estimators \hat{T}_{cu}^{NR} and \hat{T}_{cm}^{NR} respectively and is given as,

$$\hat{T}_c^{NR} = \phi \hat{T}_{cu}^{NR} + (1 - \phi) \hat{T}_{cm}^{NR}, \quad (2.19)$$

where \hat{T}_{cu}^{NR} and \hat{T}_{cm}^{NR} are given in equation (2.11) and (2.18) respectively and $\phi \in [0, 1]$ is a scalar quantity to be chosen suitably.

3. Asymptotic variance

This section is dedicated to elaboration of asymptotic properties of proposed calibration estimator \hat{T}_c^{NR} . Since, the estimator \hat{T}_c^{NR} depends on the estimators \hat{T}_{cu}^{NR} and \hat{T}_{cm}^{NR} given in equation (2.11) and (2.18) respectively, so we first discuss the asymptotic properties of \hat{T}_{cu}^{NR} and \hat{T}_{cm}^{NR} . In addition, the results suggested by Randles (1982), may be used to discuss the asymptotic variance of estimators.

Proposition 1. *The asymptotic behaviour of the calibrated estimator \hat{T}_{cu}^{NR} is same as that of*

$$\hat{T}_{cu|b}^{NR} = \frac{1}{N} \sum_{i \in s_{r2}} \beta_i^* z_{2i} + \left(\bar{X} - \frac{1}{N} \sum_{i \in s_{r2}} \beta_i^* x_i \right) b, \quad (3.1)$$

with

$$b = \left(\sum_{i \in s_{r2}} q_{ui} x_i^2 \right)^{-1} \left(\sum_{i \in s_{r2}} q_{ui} x_i z_{2i} \right). \quad (3.2)$$

Proof: Assuming,

$$\hat{T}_{cu}^{NR}(\gamma) = \frac{1}{N} \sum_{i \in s_{r2}} \beta_i^* z_{2i} + \left(\bar{X} - \frac{1}{N} \sum_{i \in s_{r2}} \beta_i^* x_i \right) \gamma, \quad (3.3)$$

for any variable γ , equation (3.3) shows $\hat{T}_{cu}^{NR}(\gamma)$ is the calibration estimator \hat{T}_{cu}^{NR} when \hat{b}_u in equation (2.11) is replaced by γ . Therefore, the limiting mean of $\hat{T}_{cu}^{NR}(\gamma)$ when the actual parameter value is b , given in equation (3.2), can be written as

$$\mu(\gamma) = \lim_{u \rightarrow +\infty} E_b \left[\hat{T}_{cu}^{NR}(\gamma) \right] = \tilde{Z}_2, \quad (3.4)$$

where \tilde{Z}_2 is the limiting value of \bar{Z}_2 as $N \rightarrow \infty$. Using Randles (1982), the estimator \hat{T}_{cu}^{NR} has the same asymptotic behaviour as that of the estimator in equation (3.1). \square

Proposition 2. *The variance of the estimator \hat{T}_{culb}^{NR} in equation (3.1) is given by*

$$V(\hat{T}_{culb}^{NR}) = \left[\frac{1}{N^2} \sum_{i \in U} \sum_{j \in U} \Delta_{ij} \frac{z_{2i}}{(\pi_i)^c} \frac{z_{2j}}{(\pi_j)^c} \right] + E_1 \left[\frac{1}{N^2} \sum_{i \in (s_n)^c} \sum_{j \in (s_n)^c} \frac{\Delta_{ij}(s_n)^c}{(\pi_{1i})^c (\pi_{1j})^c} \frac{e_i}{\pi_i |s_n^c} \frac{e_j}{\pi_j |s_n^c} \right], \quad (3.5)$$

where $e_i = z_{2i} - x_i b$, $\Delta_{ij} = \pi_{ij} - \pi_i \pi_j$, $\Delta_{ij}(s_n)^c = \pi_{ij}(s_n)^c - \pi_i(s_n)^c \pi_j(s_n)^c$ and E_1 is the expectation under the design d_1 .

Proof: Since the estimator \hat{T}_{culb}^{NR} is unbiased, so its variance is given by

$$V(\hat{T}_{culb}^{NR}) = V_1(E_3[\hat{T}_{culb}^{NR}]) + E_1[V_3(\hat{T}_{culb}^{NR})], \quad (3.6)$$

where E_1 and V_1 are the expectation and variance under the design d_1 respectively, and E_3 and V_3 represent the conditional expectation and conditional variance under design d_3 respectively

$$V_1(E_3[\hat{T}_{culb}^{NR}]) = V_1\left(\frac{1}{N} \sum_{i \in s_{r_2}} \beta_i^* z_{2i}\right) = \frac{1}{N^2} \sum_{i \in U} \sum_{j \in U} \Delta_{ij} \frac{z_{2i}}{(\pi_i)^c} \frac{z_{2j}}{(\pi_j)^c}. \quad (3.7)$$

Now,

$$\begin{aligned} E_1[V_3(\hat{T}_{culb}^{NR})] &= E_1 \left[V_3 \left\{ \left(\frac{1}{N} \sum_{i \in s_{r_2}} \beta_i^* z_{2i} \right) + \left(\bar{X} - \frac{1}{N} \sum_{i \in s_{r_2}} \beta_i^* x_i \right) b \right\} \right] \\ &= E_1 \left[V_3 \left(\frac{1}{N} \sum_{i \in s_{r_2}} \beta_i^* z_{2i} - \frac{1}{N} \sum_{i \in s_{r_2}} \beta_i^* x_i b \right) \right] \\ &= E_1 \left[V_3 \left(\frac{1}{N} \sum_{i \in s_{r_2}} \beta_i^* e_i \right) \right] \\ &= E_1 \left[\frac{1}{N^2} \sum_{i \in (s_n)^c} \sum_{j \in (s_n)^c} \frac{\Delta_{ij}(s_n)^c}{(\pi_{1i})^c (\pi_{1j})^c} \frac{e_i}{\pi_i |s_n^c} \frac{e_j}{\pi_j |s_n^c} \right]. \end{aligned} \quad (3.8)$$

Using equation (3.7) and (3.8) in equation (3.6), we get the expression for variance as in equation (3.5). \square

Remark 3. From Proposition 1 and Proposition 2, the estimator \hat{T}_{cu}^{NR} and \hat{T}_{cm}^{NR} are asymptotically unbiased and their asymptotic variances are given as

$$V(\hat{T}_{cu}^{NR}) = \frac{1}{N^2} \left[\sum_{i \in U} \sum_{j \in U} \Delta_{ij} \frac{z_{2i}}{(\pi_i)^c} \frac{z_{2j}}{(\pi_j)^c} \right] + \frac{1}{N^2} E_1 \left[\sum_{i \in (s_n)^c} \sum_{j \in (s_n)^c} \frac{\Delta_{ij}(s_n)^c}{(\pi_{1i})^c (\pi_{1j})^c} \frac{e_i}{\pi_i |s_n^c} \frac{e_j}{\pi_j |s_n^c} \right]. \quad (3.9)$$

Similarly,

$$V(\hat{T}_{cm}^{NR}) = \frac{1}{N^2} \left[\sum_{i \in U} \sum_{j \in U} \Delta'_{ij} \frac{z_{2i}}{(\pi_i)^c} \frac{z_{2j}}{(\pi_j)^c} \right] + \frac{1}{N^2} E_2 \left[\sum_{i \in s_{r_1}} \sum_{j \in s_{r_1}} \frac{\Delta_{ij|s_{r_1}}}{\pi_{1i}\pi_{1j}} \frac{e_i}{\pi_i|s_{r_1}} \frac{e_j}{\pi_j|s_{r_1}} \right], \tag{3.10}$$

where $\Delta'_{ij} = \pi_{1ij} - \pi_{1i}\pi_{1j}$ and E_2 is the expectation under the design d_2 .

Proposition 3. *The asymptotic variance of proposed calibration estimator is obtained as*

$$V(\hat{T}_c^{NR}) = \phi^2 V(\hat{T}_{cu}^{NR}) + (1 - \phi)^2 V(\hat{T}_{cm}^{NR}), \tag{3.11}$$

where $V(\hat{T}_{cu}^{NR})$ and $V(\hat{T}_{cm}^{NR})$ are given in equation (3.9) and (3.10) respectively.

Proof: The asymptotic variance of calibration estimator \hat{T}_c^{NR} is given by

$$\begin{aligned} V(\hat{T}_c^{NR}) &= E \left[\hat{T}_c^{NR} - \bar{Z}_2 \right]^2, \\ &= E \left[\phi \hat{T}_{cu}^{NR} + (1 - \phi) \hat{T}_{cm}^{NR} - \bar{Z}_2 \right]^2, \\ &= \phi^2 V(\hat{T}_{cu}^{NR}) + (1 - \phi)^2 V(\hat{T}_{cm}^{NR}) + 2\phi(1 - \phi) \text{cov}(\hat{T}_{cu}^{NR}, \hat{T}_{cm}^{NR}). \end{aligned} \tag{3.12}$$

The values of \hat{T}_{cu}^{NR} and \hat{T}_{cm}^{NR} have been computed in equation (3.9) and (3.10) respectively and as the estimators $V(\hat{T}_{cu}^{NR})$ and $V(\hat{T}_{cm}^{NR})$ are based on two non-overlapping samples of sizes u and m respectively. So, $\text{cov}(\hat{T}_{cu}^{NR}, \hat{T}_{cm}^{NR}) = 0$. By using these values in equation (3.12), we have the expression for the asymptotic variance of the calibrated estimator in presence of non-response as in equation (3.11). □

Remark 4. From the equation (3.11), it can be concluded that, $V(\hat{T}_c^{NR})$ is a function of unknown constant ϕ . Therefore, it is optimized with respect to ϕ and subsequently the optimum value of ϕ is obtained as

$$\phi_{opt.} = \frac{V(\hat{T}_{cm}^{NR})}{V(\hat{T}_{cu}^{NR}) + V(\hat{T}_{cm}^{NR})}. \tag{3.13}$$

Substituting the value of $\phi_{(opt.)}$ from equation (3.13) in (3.11), we get the optimum variance of the proposed estimator \hat{T}_c^{NR} as

$$V(\hat{T}_c^{NR})_{opt.} = \frac{V(\hat{T}_{cu}^{NR}) \times V(\hat{T}_{cm}^{NR})}{V(\hat{T}_{cu}^{NR}) + V(\hat{T}_{cm}^{NR})}. \tag{3.14}$$

4. Study under simple random sampling without replacement sampling design

In this section, calibration estimator in presence of non-response has been considered for simple random sampling without replacement (SRSWOR) design on both the moves. For that the relevant suppositions are given as

$$\pi_{1i} = \frac{r_1}{N}, \quad \pi_{1ij} = \frac{r_1(r_1 - 1)}{N(N - 1)}.$$

Because the sample s_n is drawn from U with SRSWOR of size n , it implies that the complement, $s_n^c = U - s_n$ is a simple random sample without replacement of size $N - n$, therefore we have

$$\pi_{1i}^c = \frac{N - n}{N}, \quad \pi_{1ij}^c = \frac{(N - n)(N - n - 1)}{N(N - 1)}.$$

Also, we suppose that the matched sample s_m is drawn from s_{r_1} , with SRSWOR of size m so

$$\pi_{i|s_{r_1}} = \frac{m}{r_1}, \quad \pi_{ij|s_{r_1}} = \frac{m(m - 1)}{r_1(r_1 - 1)}.$$

Finally, the unmatched sample s_u is drawn from s_n^c with SRSWOR of size u . Thus, we have

$$\pi_{i|s_n^c} = \frac{r_2}{N - r_2}, \quad \pi_{ij|s_n^c} = \frac{r_2(r_2 - 1)}{(N - n)(N - n - 1)}$$

Now, based on sample of size r_2 on current move, the proposed calibration estimator \hat{T}_{cu}^{NR} under SRSWOR sampling design becomes

$$\hat{T}_{cu}^{NR}(s) = \left[\bar{z}_{2u} + \hat{b}_u(s)(\bar{X} - \bar{x}_u) \right], \tag{4.1}$$

with

$$\hat{b}_u(s) = \left(\sum_{i \in s_{r_2}} q_{ui} x_i^2 \right)^{-1} \left(\sum_{i \in s_{r_2}} q_{ui} x_i z_{2i} \right).$$

Similarly, based on sample of size m on current move, the proposed calibration estimator \hat{T}_{cm}^{NR} under SRSWOR scheme is obtained as

$$\hat{T}_{cm}^{NR}(s) = \left[\bar{z}_{2m} + \hat{b}_{m1}(s)(\bar{z}_{1n}^{*NR}(s) - \bar{z}_{1m}) + \hat{b}_{m2}(s)(\bar{X} - \bar{x}_m) \right]. \tag{4.2}$$

with

$$\begin{aligned} \hat{b}_{m1}(s) &= \left(\sum_{i \in s_m} q_{mi} z_{1i} z_{2i} \right) \left[\left(\sum_{i \in s_m} q_{mi} x_i^2 \right) - \left(\sum_{i \in s_m} q_{mi} z_{1i} x_i \right) \right] \hat{b}_m(s), \\ \hat{b}_{m2}(s) &= \left(\sum_{i \in s_m} q_{mi} x_i z_{2i} \right) \left[\left(\sum_{i \in s_m} q_{mi} (z_{1i})^2 \right) - \left(\sum_{i \in s_m} q_{mi} z_{1i} x_i \right) \right] \hat{b}_m(s), \\ \hat{b}_m(s) &= \left[\left(\sum_{i \in s_m} q_{mi} (z_{1i})^2 \right) \left(\sum_{i \in s_m} q_{mi} x_i^2 \right) - \left(\sum_{i \in s_m} q_{mi} z_{1i} x_i \right)^2 \right]^{-1}, \\ \bar{z}_{1n}^{*NR}(s) &= \left[\bar{z}_{1n} + \hat{b}_n(s)(\bar{X} - \bar{x}_n) \right], \\ \hat{b}_n(s) &= \left(\sum_{i \in s_{r_1}} q_{ni} x_i^2 \right)^{-1} \left(\sum_{i \in s_{r_1}} q_{ni} x_i z_{1i} \right). \end{aligned}$$

Now, the estimator $\hat{T}_c^{NR}(s)$ becomes

$$\hat{T}_c^{NR}(s) = \phi^s \hat{T}_{cu}^{NR}(s) + (1 - \phi^s) \hat{T}_{cm}^{NR}(s), \quad (4.3)$$

where $\hat{T}_{cu}^{NR}(s)$ and $\hat{T}_{cm}^{NR}(s)$ are given in equation (4.1) and (4.2) respectively and $\phi^s \in [0, 1]$ is a scalar quantity to be chosen suitably.

Remark 5. Further if we assume, $q_{ui} = q_{ni} = q_{mi} = 1$ in \hat{T}_c^{NR} , then the calibration estimator \hat{T}_c^{NR} be denoted as \hat{T}_c^{*NR} .

Remark 6. The proposed calibration estimator have been compared with general successive sampling estimator in presence of non-response at both the moves, so the general successive sampling estimator have been modified for estimation of coded response variable and is given as

$$\hat{T}_g^{NR} = \Psi_g \hat{T}_{gu}^{NR} + (1 - \Psi_g) \hat{T}_{gm}^{NR}, \text{ where } \Psi_g \in [0, 1], \quad (4.4)$$

with $\hat{T}_{gu}^{NR} = \bar{z}_{2r_2}$, $\hat{T}_{gm}^{NR} = \bar{z}_{2m} + \beta_{z_2 z_1} (\bar{z}_{1r_1} - \bar{z}_{1m}) \beta_{z_2 z_1} = \frac{S_{z_2 z_1}}{S_{z_1}^2}$.

5. Possible cases

There might be a possibility that non-response may occur only at current move or only at previous move or there may be no non-response at any move. Therefore, in order to retain similar estimators in all possible situations, the calibration technique have been retained and possible modifications has been done in the constraints as per the situation and calibration estimators in different possible situations have been obtained, which are described in following cases.

5.1. Case-I: when there is non-response only at current move

In this situation the proposed estimator \hat{T}_c^{NR} of the coded response variable \bar{Z}_2 changes to

$$\hat{T}_{c_1}^{NR} = \Psi_1 \hat{T}_{cu}^{NR} + (1 - \Psi_1) \hat{T}_{cm}^{NR}; \Psi_1 \in [0, 1], \quad (5.1)$$

where the estimator \hat{T}_{cm}^{NR} can be obtained by replacing r_1 by n in equation (2.16) and other equations that depends on it and \hat{T}_{cu}^{NR} is defined in equation (2.11).

5.2. Case-II: when there is non-response only at first (previous) move

In the presence of non-response only at first (previous) move, the estimator \hat{T}_c^{NR} of the coded response variable \bar{Z}_2 changes to

$$\hat{T}_{c_2}^{NR} = \Psi_2 \hat{T}_{cu}^{NR} + (1 - \Psi_2) \hat{T}_{cm}^{NR}; \Psi_2 \in [0, 1], \quad (5.2)$$

where the estimator \hat{T}_{cu}^{NR} can be obtained by replacing r_2 by u in equation (2.11) and other equations that depends on it and \hat{T}_{cm}^{NR} is defined in equation (2.18).

5.3. Case-III: when there is no non-response at any move

In the presence of no non-response at any move, the estimator \hat{T}_c^{NR} of the coded response variable \bar{Z}_2 changes to

$$\hat{T}_{c_3} = \Psi_3 \hat{T}_{cu}^{NR} + (1 - \Psi_3) \hat{T}_{cm}^{NR}; \Psi_3 \in [0, 1], \quad (5.3)$$

Table 3: Estimators of sensitive population mean and their variances

| | Estimators | Variance |
|---------|--|---|
| PORT-I | $\hat{Y}_{2c} = \frac{\hat{T}_c^{*NR} - (1-p)a\bar{S}_2}{p+(1-p)\bar{S}_1}$ | $V[\hat{Y}_{2c}] = \frac{V(\hat{T}_c^{*NR})_{opt.}}{[p+(1-p)\bar{S}_1]^2}$ |
| | $\hat{Y}_{2g} = \frac{\hat{T}_g^{NR} - (1-p)a\bar{S}_2}{p+(1-p)\bar{S}_1}$ | $V[\hat{Y}_{2g}] = \frac{V(\hat{T}_g^{NR})_{opt.}}{[p+(1-p)\bar{S}_1]^2}$ |
| | $\hat{Y}_{2c_1} = \frac{\hat{T}_{c_1}^{NR} - (1-p)a\bar{S}_2}{p+(1-p)\bar{S}_1}$ | $V[\hat{Y}_{2c_1}] = \frac{V(\hat{T}_{c_1}^{NR})_{opt.}}{[p+(1-p)\bar{S}_1]^2}$ |
| | $\hat{Y}_{2c_2} = \frac{\hat{T}_{c_2}^{NR} - (1-p)a\bar{S}_2}{p+(1-p)\bar{S}_1}$ | $V[\hat{Y}_{2c_2}] = \frac{V(\hat{T}_{c_2}^{NR})_{opt.}}{[p+(1-p)\bar{S}_1]^2}$ |
| | $\hat{Y}_{2c_3} = \frac{\hat{T}_{c_3}^{NR} - (1-p)a\bar{S}_2}{p+(1-p)\bar{S}_1}$ | $V[\hat{Y}_{2c_3}] = \frac{V(\hat{T}_{c_3}^{NR})_{opt.}}{[p+(1-p)\bar{S}_1]^2}$ |
| PORT-II | $\hat{Y}_{2c}^* = \hat{T}_c^{*NR}$ | $V[\hat{Y}_{2c}^*] = V(\hat{T}_c^{*NR})_{opt.}$ |
| | $\hat{Y}_{2g}^* = \hat{T}_g^{NR}$ | $V[\hat{Y}_{2g}^*] = V(\hat{T}_g^{NR})_{opt.}$ |
| | $\hat{Y}_{2c_1}^* = \hat{T}_{c_1}^{NR}$ | $V[\hat{Y}_{2c_1}^*] = V(\hat{T}_{c_1}^{NR})_{opt.}$ |
| | $\hat{Y}_{2c_2}^* = \hat{T}_{c_2}^{NR}$ | $V[\hat{Y}_{2c_2}^*] = V(\hat{T}_{c_2}^{NR})_{opt.}$ |
| | $\hat{Y}_{2c_3}^* = \hat{T}_{c_3}^{NR}$ | $V[\hat{Y}_{2c_3}^*] = V(\hat{T}_{c_3}^{NR})_{opt.}$ |

where the estimator \hat{T}_{cu} can be obtained by replacing r_2 by u in equation (2.11) and \hat{T}_{cm} can be obtained by replacing r_1 by n in equation (2.16) and other equations that depends on it and \hat{T}_{cm}^{NR} is defined in equation (2.18).

6. Estimators for sensitive population mean under PORT

Replacing the population mean of coded response variable \bar{Z}_2 in equation (2.3) and (2.6) by its estimators $\hat{T}_c^{*NR}, \hat{T}_g^{NR}, \hat{T}_{c_1}^{NR}, \hat{T}_{c_2}^{NR}$ and \hat{T}_{c_3} , the respective estimators for sensitive population mean at current move becomes $\hat{Y}_{2c}, \hat{Y}_{2g}, \hat{Y}_{2c_1}, \hat{Y}_{2c_2}$ and \hat{Y}_{2c_3} under PORT-I model and $\hat{Y}_{2c}^*, \hat{Y}_{2g}^*, \hat{Y}_{2c_1}^*, \hat{Y}_{2c_2}^*$ and $\hat{Y}_{2c_3}^*$ under PORT-II model respectively, which are presented in Table 3.

Remark 7. For the considered model in equation (2.3) and (2.6), the two scrambling variables \bar{S}_1 and \bar{S}_2 used to perturb the true response through the PORT models may follow any distribution in two move successive sampling. Hence, following Pollock and Bek (1976) and Eichhorn and Hayre (1983), we consider scrambling variable S_1 to follow normal distribution with mean 0 and variance 1. However, the scrambling variable S_2 has been assumed to follow normal distribution with mean 1 and variance 1.

7. Simulation study

In this section, a simulation study has been carried out to reveal the behaviour of the proposed estimators. For this purpose, a natural population has been considered from statistical abstracts of United States. The considered population comprise of $N = 51$ states, which is described as: y_1 : Rate of abortion in 2007 y_2 : Rate of abortion in 2008 x : Rate of abortion in 2005. As discussed in Remark 6, the scrambling variables $S_1 \sim N(0, 1)$ and $S_2 \sim N(1, 1)$. The data for S_1 and S_2 have been generated by MATLAB software.

To judge the performance of both the PORT models under the proposed calibration estimators in presence of non-response to estimate the sensitive population mean in two move successive sampling,

we have studied the behaviour of the estimators by considering different choices for rate of non-response at both moves. For simulation, 10,000 independent replications of considered sampling design in two move successive sampling via MATLAB have been considered. All the samples are obtained under simple random sampling without replacement. An environment through simulation process has been created for non-response by assuming non-response rates as 10%, 20%, and 30% at both the moves.

The entire simulation has been replicated for different values of n, m and u which are considered as different sets given as

$$Set - I : n = 20; \quad m = 12; u = 8, \quad Set - II : n = 20; \quad m = 10; u = 10.$$

Calibration estimators have been compared with general successive sampling estimator under both the considered PORT models in terms of percent relative efficiency (PRE), which are given as

$$PRE_1 = \frac{MSE(\hat{Y}_{2g})}{MSE(\hat{Y}_{2c})} \times 100,$$

$$PRE_2 = \frac{MSE(\hat{Y}_{2g}^*)}{MSE(\hat{Y}_{2c}^*)} \times 100,$$

where $MSE(\hat{Y}_{2g}) = 1/10,000 \sum_{i=1}^{10,000} [\hat{Y}_{2gi} - \bar{Y}_2]^2$. Similarly, $MSE(\hat{Y}_{2g}^*)$, $MSE(\hat{Y}_{2c})$ and $MSE(\hat{Y}_{2c}^*)$ can be computed. The simulation results have been represented in Tables 4–5 respectively. Further, in order to identify the better PORT model, the percent relative efficiency (PRE) of calibration estimator under PORT-I with respect to calibration estimator under PORT-II have been computed as

$$PRE_3 = \frac{MSE(\hat{Y}_{2c})}{MSE(\hat{Y}_{2c}^*)} \times 100,$$

where $MSE(\hat{Y}_{2c}) = 1/10,000 \sum_{i=1}^{10,000} [\hat{Y}_{2ci} - \bar{Y}_2]^2$ and similarly, $MSE(\hat{Y}_{2c}^*)$ can be computed. The results are presented in Figures 1–2.

8. Discussion of results and conclusion

Following interpretation can be drawn from the simulation results presented in Tables 4–5 and also in Figures 1–2.

- (i) From Tables 4–5, it is observed that the proposed calibration estimator is performing better than general successive sampling estimator in the presence of non-response at both the moves under PORT-I as well as PORT-II models. This shows that the use of calibration technique to adjust the effect due to non-response is fruitful.
- (ii) Figures 1–2 show that the calibration estimator under proposed PORT-II model is better than the same calibration estimator under PORT-I model for all considered choices of constants and non-response rates.
- (iii) It is also observed from Figures 1–2 that for the fixed value of p , the PRE_3 first increases up to $\Psi = 0.5$, then decreases for $\Psi > 0.5$. This shows that if more weight is attached to

Table 4: Percent relative efficiency of \hat{Y}_{2g} with respect to \hat{Y}_{2c} for different sets and different choices of non-response (NR) rates under PORT-I where $\Psi \in \{\phi, \Psi_g\}$

| Ψ | p | SET-I | | | | | | SET-II | | | | | |
|--------|-----|------------------|-----------|------------------|-----------|------------------|-----------|------------------|-----------|------------------|-----------|------------------|-----------|
| | | NR = 10% | | NR = 20% | | NR = 30% | | NR = 10% | | NR = 20% | | NR = 30% | |
| | | PRE ₁ | | PRE ₁ | | PRE ₁ | | PRE ₁ | | PRE ₁ | | PRE ₁ | |
| | | $a = -10$ | $a = -19$ | $a = -10$ | $a = -19$ | $a = -10$ | $a = -19$ | $a = -10$ | $a = -19$ | $a = -10$ | $a = -19$ | $a = -10$ | $a = -19$ |
| 0.1 | 0.3 | 110.7 | 137.9 | 178.6 | 137.2 | 175.8 | 140.4 | 100.5 | 138.1 | 109.1 | 136.7 | 136.1 | 135.6 |
| | 0.5 | 152.7 | 173.3 | 152.4 | 173.0 | 251.9 | 172.12 | 156.3 | 174.2 | 155.0 | 173.4 | 151.6 | 172.6 |
| | 0.7 | 362.6 | 196.9 | 349.1 | 195.1 | 336.4 | 198.6 | 342.6 | 191.0 | 315.7 | 192.9 | 315.1 | 189.1 |
| | 0.9 | 319.1 | 387.3 | 372.9 | 419.4 | 406.2 | 436.2 | 158.0 | 234.2 | 201.7 | 227.8 | 181.4 | 255.1 |
| 0.3 | 0.3 | 120.4 | 118.4 | 103.5 | 107.8 | 120.0 | 146.4 | 115.6 | 118.5 | 115.0 | 117.5 | 132.3 | 106.4 |
| | 0.5 | 140.9 | 148.9 | 138.8 | 148.1 | 137.1 | 146.5 | 141.1 | 178.6 | 141.7 | 138.2 | 140.1 | 137.2 |
| | 0.7 | 302.5 | 173.8 | 292.7 | 173.1 | 286.7 | 174.1 | 295.9 | 171.6 | 289.8 | 172.0 | 281.7 | 172.1 |
| | 0.9 | 309.9 | 355.2 | 332.3 | 413.2 | 428.3 | 435.9 | 185.2 | 246.1 | 207.1 | 224.7 | 239.0 | 277.9 |
| 0.5 | 0.3 | 110.4 | 106.5 | 116.5 | 105.2 | 113.8 | 100.0 | 100.4 | 100.1 | 108.3 | 109.6 | 116.5 | 118.6 |
| | 0.5 | 128.6 | 163.6 | 126.4 | 122.4 | 124.7 | 181.3 | 129.6 | 123.6 | 128.6 | 142.8 | 127.4 | 134.5 |
| | 0.7 | 244.7 | 150.3 | 234.5 | 149.6 | 231.0 | 152.0 | 243.0 | 150.2 | 237.2 | 151.9 | 234.9 | 152.2 |
| | 0.9 | 405.7 | 378.5 | 366.0 | 410.2 | 454.6 | 448.2 | 219.4 | 298.1 | 231.9 | 277.7 | 249.5 | 232.5 |
| 0.7 | 0.3 | 112.3 | 113.9 | 109.1 | 171.8 | 108.5 | 70.4 | 103.9 | 183.8 | 102.5 | 112.7 | 100.6 | 101.7 |
| | 0.5 | 115.8 | 119.0 | 114.7 | 186.9 | 113.4 | 85.5 | 117.6 | 118.6 | 118.0 | 108.3 | 116.0 | 117.1 |
| | 0.7 | 184.9 | 129.1 | 179.3 | 228.9 | 174.8 | 127.3 | 187.0 | 130.2 | 185.0 | 129.4 | 182.3 | 230.2 |
| | 0.9 | 424.1 | 325.7 | 413.4 | 339.3 | 423.6 | 353.0 | 261.6 | 285.2 | 293.1 | 287.6 | 303.4 | 287.4 |
| 0.9 | 0.3 | 100.0 | 97.2 | 94.0 | 184.9 | 102.0 | 183.2 | 100.1 | 100.1 | 102.0 | 186.9 | 100.0 | 135.7 |
| | 0.5 | 105.3 | 102.7 | 104.3 | 191.6 | 102.5 | 189.8 | 107.3 | 103.7 | 105.9 | 193.1 | 105.0 | 192.2 |
| | 0.7 | 134.7 | 108.4 | 132.0 | 208.4 | 131.0 | 207.3 | 233.4 | 109.4 | 135.1 | 209.1 | 133.7 | 209.3 |
| | 0.9 | 355.8 | 220.9 | 349.9 | 223.7 | 341.7 | 231.6 | 334.2 | 204.1 | 327.2 | 211.8 | 328.7 | 217.0 |

Table 5: Percent relative efficiency of \hat{Y}_{2g}^* with respect to \hat{Y}_{2c}^* for different sets and different choices of non-response (NR) rates under PORT-II where $\Psi \in \{\phi, \Psi_g\}$

| Ψ | p | SET-I | | | SET-II | | |
|--------|-----|------------------|------------------|------------------|------------------|------------------|------------------|
| | | NR = 10% | NR = 20% | NR = 30% | NR = 10% | NR = 20% | NR = 30% |
| | | PRE ₂ | PRE ₂ | PRE ₂ | PRE ₂ | PRE ₂ | PRE ₂ |
| 0.1 | 0.3 | 142.0 | 144.7 | 160.1 | 130.6 | 131.2 | 140.6 |
| | 0.5 | 177.0 | 208.1 | 235.0 | 239.0 | 244.4 | 257.5 |
| | 0.7 | 277.7 | 339.0 | 456.8 | 900.6 | 154.3 | 396.4 |
| | 0.9 | 3809.6 | 4150.1 | 5548.5 | 1578.5 | 1523.0 | 2549.4 |
| 0.3 | 0.3 | 152.6 | 161.8 | 183.6 | 126.1 | 126.6 | 129.9 |
| | 0.5 | 214.2 | 209.0 | 252.0 | 254.2 | 356.8 | 462.4 |
| | 0.7 | 389.2 | 363.9 | 453.9 | 460.0 | 488.0 | 568.5 |
| | 0.9 | 3420.2 | 4141.8 | 4400.8 | 1504.1 | 1849.2 | 2485.5 |
| 0.5 | 0.3 | 185.2 | 197.5 | 117.7 | 136.7 | 249.1 | 163.4 |
| | 0.5 | 253.0 | 259.3 | 259.7 | 148.7 | 396.0 | 188.4 |
| | 0.7 | 304.9 | 326.2 | 338.3 | 225.1 | 582.8 | 228.0 |
| | 0.9 | 2408.2 | 2407.6 | 2181.4 | 1582.8 | 1622.9 | 1960.6 |
| 0.7 | 0.3 | 135.9 | 137.9 | 132.8 | 107.9 | 113.2 | 106.6 |
| | 0.5 | 173.1 | 161.4 | 159.0 | 137.5 | 134.5 | 164.0 |
| | 0.7 | 247.0 | 229.0 | 207.2 | 213.4 | 224.5 | 212.8 |
| | 0.9 | 1015.5 | 931.2 | 784.2 | 1204.8 | 1035.5 | 1012.1 |
| 0.9 | 0.3 | 148.5 | 134.4 | 122.9 | 167.8 | 162.8 | 150.0 |
| | 0.5 | 155.3 | 142.5 | 131.0 | 193.9 | 180.3 | 155.7 |
| | 0.7 | 173.7 | 165.0 | 144.8 | 224.6 | 193.5 | 176.4 |
| | 0.9 | 388.4 | 350.2 | 300.6 | 496.8 | 428.5 | 387.6 |

matched sample than PRE is higher as compared to the situation when more weight is attached to unmatched/fresh sample at current move. These results are in accordance with the theory of

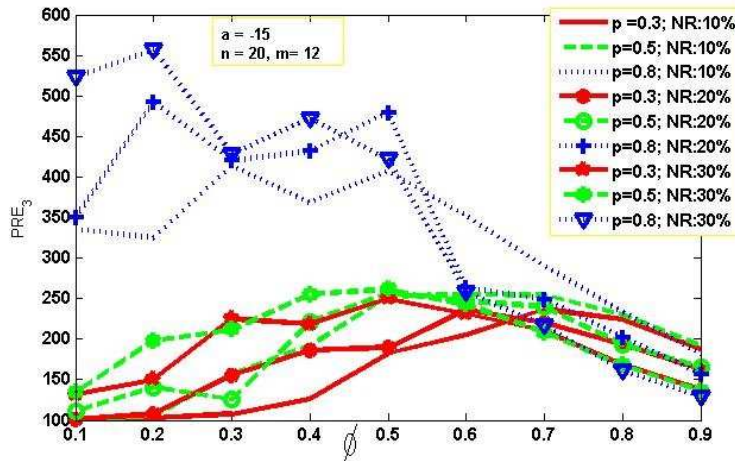


Figure 1: PRE of the calibration estimator under PORT-I with respect to the calibration estimator under PORT-II for different choices of non-response rates and varying ϕ for $a = -15$.

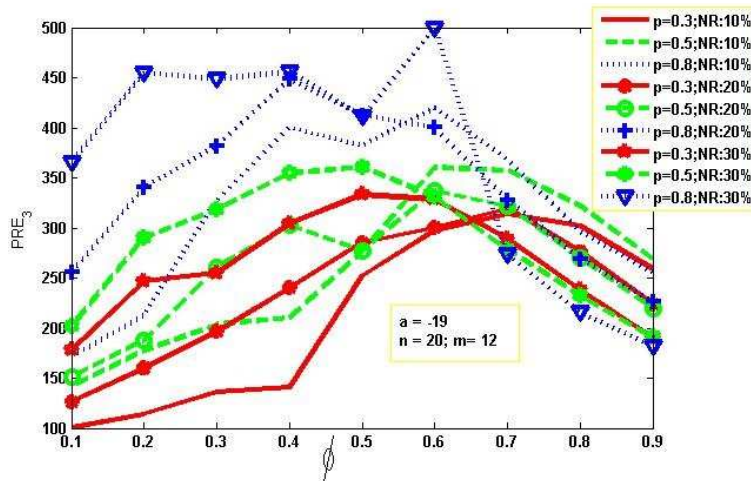


Figure 2: PRE of the calibration estimator under PORT-I with respect to the calibration estimator under PORT-II for different choices of non-response rates and varying ϕ for $a = -19$.

successive sampling.

- (iv) Figures 1–2 also shows that higher percent relative efficiency is observed for larger value of p i.e., PRE_3 in general increased as p increases.

8.1. Conclusion

The estimation of sensitive population mean at current move in two move successive sampling is feasible using PORT. The calibration technique applied to adjust the effect due to non-response is proved to be fruitful under both the considered models. The proposed model PORT-II is coming out to be more efficient than the modified Sanullah *et al.* (2020) model (PORT-I) on successive moves. Therefore, it is concluded that the proposed PORT model may be used for the estimation of sensitive

population mean at current move in two move successive sampling.

Acknowledgements

Authors are thankful to the reviewers for their valuable suggestions, which led to improvement over the earlier version of the paper. Authors are also thankful to SERB, New Delhi, India for providing the financial assistance to carry out the present work. Authors sincerely acknowledged the free access to data by statistical abstracts of United States.

References

- Arnab R (2004). Optional randomized response techniques for complex survey designs, *Biometrical Journal*, **46**, 114–124.
- Arnab R (2011). Alternative estimators for randomized response techniques in multi-character surveys, *Communications in Statistics-Theory and Methods*, **40**, 1839–1848.
- Arnab R, Singh S, and North D (2012). Use of two decks of cards in randomized response techniques for complex survey designs, *Communications in Statistics-Theory and Methods*, **41**, 3198–3210.
- Arnab R and Singh S (2013). Estimation of mean of sensitive characteristics for successive sampling, *Communications in Statistics-Theory and Methods*, **42**, 2499–2524.
- Christofides TC (2003). A generalized randomized response technique, *Metrika*, **57**, 195–200.
- Chaudhuri A and Dihidar K (2009). Estimating means of stigmatizing qualitative and quantitative variables from discretionary responses randomized or direct, *Sankhya*, **71**, 123–136.
- Diana G and Perri PF (2010). New scrambled response models for estimating the mean of a sensitive quantitative character, *Journal of Applied Statistics*, **37**, 1875–1890.
- Diana G and Perri PF (2011). A class of estimators for quantitative sensitive data, *Stat Papers*, **52**, 633–650.
- Deville JC and Särndal CE (1992). Calibration estimators in survey sampling, *Journal of the American Statistical Association*, **87**, 376–382.
- Eichhorn BH and Hayre LS (1983). Scrambled randomized response method for obtaining sensitive quantitative data, *Journal of Statistical Planning and Inference*, **7**, 307–316.
- Gupta S, Gupta B, and Singh S (2002). Estimation of sensitivity level of personal interview survey question, *Journal of Statistical Planning and Inference*, **100**, 239–247.
- Gupta S, Mehta S, Shabbir J, and Dass BK (2013). Generalized scrambling quantitative optional randomized response models, *Communication in Statistics-Theory and Methods*, **42**, 4034–4042.
- Greenberg BG, Kubler RR, and Horvitz DG (1971). Application of RR technique in obtaining quantitative data, *Journal of the American Statistical Association*, **66**, 243–250.
- Horvitz DG, Shah BV, and Simmons WR (1967). The unrelated question randomized response model, *Journal of the American Statistical Association*, 65–72.
- Himmelfarb S and Edgell SE (1980). Additive constants model: A randomized response technique for eliminating evasiveness to quantitative response questions, *Psychological Bulletin*, **87**, 525–530.
- Hussain Z and Al-Zahrani B (2016). Mean and sensitivity estimation of a sensitive variable through additive scrambling, *Communications in Statistics-Theory and Methods*, **45**, 182–193.
- Kim JM and Elam ME (2007). A stratified unrelated question randomized response model, *Statistical Papers*, **48**, 215–233.
- Lundström, S and Särndal CE (1999). Calibration as a standard method for treatment of non-response, *Journal of Official Statistics*, **15**, 305–327.
- Mangat NS and Singh S (1994). An optional randomised response sampling technique, *Journal of*

- Indian Statistical Association*, **32**, 71–75.
- Naeem N and Shabbir J (2016). Use of scrambled responses on two occasions successive sampling under non-response, Hacettepe University Bulletin of Natural Sciences and Engineering Series B: Mathematics and Statistics, **46**.
- Pal S (2008). Unbiasedly estimating the total of a stigmatizing variable from a complex survey on permitting options for direct or randomized responses, *Statistical Papers*, **49**, 157–164.
- Pollock KH and Bek Y (1976). A comparison of three randomized response models for quantitative data, *Journal of the American Statistical Association*, **71**, 884–886.
- Perri PF and Diana G (2013). Scrambled response models Based on auxiliary variables, *Advances in Theoretical and Applied Statistics*(pp281-291), Springer, Berlin.
- Priyanka K and Trisandhya P (2019a). A Composite Class of Estimators using Scrambled Response Mechanism for Sensitive Population mean in Successive Sampling, *Communications in Statistics-Theory and Methods*, **48**, 1009–1032.
- Priyanka K and Trisandhya P (2019b). A Item sum techniques for quantitative sensitive estimation on successive occasions, *Communications for Statistical Applications and Methods*, **26**, 175–189.
- Priyanka K, Trisandhya P, and Mittal R (2018). Dealing sensitive characters on successive occasions through a general class of estimators using scrambled response techniques, *Metron*, **76**, 203–230.
- Priyanka K, Trisandhya P, and Mittal R (2019). Scrambled response Techniques in Two Wave Rotation Sampling for Estimating Population Mean of Sensitive Characteristics with Case Study, *Journal of Indian Society of Agricultural Statistics*, **73**, 41–52.
- Randles R (1982). On the asymptotic normality of statistics with estimated parameters, *Annals of Statistics*, **10**, 462–474.
- Saha A (2007). A simple randomized response technique in complex surveys, *Metron*, **LXV**, 59–66.
- Singh GN, Suman S, Khetan M, and Paul C (2017). Some estimation procedures of sensitive character using scrambled response techniques in successive sampling, *Communications in Statistics-Theory and Methods*.
- Singh GN, Khetan M, and Suman S (2018). Assessment of Scrambled Response on Second Call in Two-Occasion Successive Sampling under Non-Response, *Journal of Indian Society of Agricultural Statistics*, **72**, 147–156.
- Sanaullah A, Saleem I, Gupta S, and Hanif M (2020). Mean estimation with generalized scrambling using two-phase sampling, *Communications in Statistics-Simulation and Computations*.
- Wu JW, Tian GL, and Tang ML (2008). Two new models for survey sampling with sensitive characteristics: Design and Analysis, *Metrika*, **67**, 251–263.
- Warner SL (1965). Randomized response: a survey technique for eliminating evasive answer bias, *Journal of the American Statistical Association*, **60**, 63–69.
- Yan Z, Wang J, and Lai J (2009). An efficiency and protection degree-based comparison among the quantitative randomized response strategies, *Communications in Statistics-Theory and Methods*, **38**, 400–408.
- Yu B, Jin Z, Tian J, and Gao G (2015). Estimation of sensitive proportion by randomized response data in successive sampling, *Computational and Mathematical Methods in Medicine*, **2015**, 1–6.