



Estimating the prevalence of sensitive attribute with optional unrelated question randomized response models under simple and stratified random sampling

G. Narjis* and J. Shabbir

Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan.

Received 24 October 2018; received in revised form 14 October 2019; accepted 23 December 2019

KEYWORDS

Multi-question approach;
 Randomized response;
 Proportional allocation;
 Relative efficiency.

Abstract. This study proposes optional Randomized Response Technique (RRT) models in a binary response situation. The application of the proposed optional RRT models under stratification is also explored. Gupta et al. [1] introduced an ingenious idea of optional RRT model in which a question may appear sensitive to one respondent while being not sensitive to another. This study focuses on estimating π , the prevalence of sensitive attribute, ω , the sensitivity level of the underlying sensitive question when the proportion of unrelated innocuous attribute π_x is unknown. A new multi-question approach is proposed for estimation of parameters (π, ω) . A comparison between the proposed optional RRT models and the corresponding full RRT models is carried out numerically under simple and stratified random sampling.

© 2021 Sharif University of Technology. All rights reserved.

1. Introduction

Respondents usually either distort the truth or refuse to answer when potentially sensitive questions such as illegal use of drugs, tax evasion, homosexuality, cheating in exams, extra marital affairs, illegal income, or domestic violence are included in surveys. Non-response or false response produces biased estimates; therefore, Randomized Response Technique (RRT) is useful for reducing the bias and procuring reliable data. Warner [2] suggested an ingenious method of RRT to estimate the unknown population prevalence of sensitive attribute (π). In Warner's model, a randomly selected proportion p of respondents is asked a sensitive attribute, say A , and the remaining proportion $(1 - p)$

of respondents are asked complement of the sensitive attribute A^c . The researcher does not know the respondents' answers, whether they are of sensitive question or its complement. Greenberg et al. [3] extended the Warner's model by introducing unrelated innocuous attribute say X as a replacement of A^c in their RRT model. Some other developments in RRT result from Chaudhuri and Mukerjee [4], Mahmood et al. [5], Perri [6], Hussain and Shabbir [7], Lee et al. [8], Abdelfatah and Mazloun [9], Tanveer and Singh [10,11], Blair et al. [12], Singh and Gorey [13], Bose [14], and Abid et al. [15].

RRT models can be categorized as full, partial, and optional. In full RRT model, a Randomization Device (RD) is provided to respondents who are requested to respond to the question (i.e., sensitive or non-sensitive) occurring on RD. The element of a truthful response is incorporated in partial RRT model by adding the additional stage or RD. For example, in the two-stage RRT model, the first RD has two options:

*. Corresponding author. Tel.: +92-051 9064-2184
 E-mail address: gnarjis@stat.qau.edu.pk (G. Narjis)

- (i) Do you belong to the sensitive group?
- (ii) Go to the second RD with known probabilities T and $(1 - T)$.

The second RD is exactly the same as that provided in the full RRT model. In the optional RRT model, respondents are requested to provide truthful responses if they consider the question non-sensitive and use the RD if they consider the sensitive question and provides the response after using it. Mangat and Singh [16] and Gupta et al. [1] introduced the partial RRT model and optional RRT model, respectively. The characteristics of optional RRT models were discussed further by Gupta and Shabbir [17], Gupta et al. [18–20], and Chhabra et al. [21]. The purpose of all RRT models is to protect privacy and promote cooperation among respondents.

Researchers remain mostly interested in investigating both the prevalence of sensitive attribute and sensitivity level of the question. To this end, Gupta et al. [22] described a two-step approach to estimating the mean of a sensitive variable and sensitivity level. In this approach, the respondents are asked two separate questions:

- Question 1 is the auxiliary question about whether or not the main research question is sensitive enough for the respondent to opt for a scrambled response;
- Question 2 is the main research question that the respondent answers using an additive optional RRT model.

Sihm et al. [23], Chhabra et al. [21], and Narjis and Shabbir [24] discussed some improvement on a two-step approach.

Motivated by Sihm et al. [23], we propose optional RRT models for the randomized devices that consist of three types of statements and three techniques proposed by Mahmood et al. [5]. We propose a multi-question approach to estimate the prevalence of sensitive attribute and sensitivity level when an unrelated innocuous attribute is unknown. The basic purpose of this study is to obtain the truthful responses from some proportion of people and promote the efficiency of the models. This paper is organized as follows: Section 2 presents RRT models given by Mahmood et al. [5]. Section 3 introduces a multi-question approach to estimate the parameters of the proposed optional RRT models under Simple Random Sampling With Replacement (SRSWR) and numerically compares it with the RRT models proposed by Mahmood et al. [5]. Section 4 proposes optional RRT models under stratification and numerically compares them with the mentioned RRT models. Finally, Section 5 concludes the paper.

2. Some existing RRT models

This section discusses RRT models from Mahmood et al. [5]. Mahmood et al. [5] introduced three binary unrelated question models to estimate the prevalence of sensitive attribute:

Technique-I: The RD consists of three types of statements:

- (i) I belong to group A ;
- (ii) I belong to group X^c ;
- (iii) I belong to group X with probabilities p_1, p_2 , and p_3 , such that $\sum_{i=1}^3 p_i = 1$.

The proportion of the unrelated question π_x is unknown; thus, to estimate π_x , another sample of size n_2 is required such that $(n_1 + n_2) = n$. The probability of a “Yes” response in the first sample of size n_1 respondents is given as follows:

$$\lambda_{M_1} = \pi p_1 + p_2(1 - \pi_x) + p_3 \pi_x. \quad (1)$$

In the second independent sample of n_2 respondents, the question is asked only on the unrelated question X to estimate the proportion π_x of the unrelated innocuous attribute. An unbiased estimator for population proportion π is given by:

$$\hat{\pi}_{M_1} = \frac{\hat{\lambda}_{M_1} - p_2(1 - \hat{\pi}_x) - p_3 \hat{\pi}_x}{p_1}, \quad (2)$$

taking $\hat{\lambda}_{M_1} \sim B(n_1, \lambda_{M_1})$ and $\hat{\pi}_x \sim B(n_2, \pi_x)$. The minimum variance of $\hat{\pi}_{M_1}$ is given by:

$$\begin{aligned} Var(\hat{\pi}_{M_1})_{\min} &= \frac{\left[\sqrt{\lambda_{M_1}(1 - \lambda_{M_1})} + (p_2 - p_3)\sqrt{\pi_x(1 - \pi_x)} \right]^2}{np_1^2}. \end{aligned} \quad (3)$$

The ratio of optimal sample sizes n_1 and n_2 for the two independent samples is given by:

$$\left(\frac{n_1}{n_2} \right)_{opt} = \sqrt{\frac{\lambda_{M_1}(1 - \lambda_{M_1})}{(p_2 - p_3)^2 \pi_x(1 - \pi_x)}}. \quad (4)$$

Technique-II: The statement (ii) from Technique-I is simply replaced with “Try once again”. The respondents are advised to repeat the process if the selected card shows “Try once again” statement. In the second draw, if statement “Try once again” reappears, then respondents are asked to say “Yes” irrespective of their actual status. The probability of a “Yes” response in the first sample of size n_1 respondents is given as follows:

$$\lambda_{M_2} = (1 + p_2)(p_1 \pi + p_3 \pi_x) + p_2^2. \quad (5)$$

In the second independent sample of n_2 respondents, the question is asked only on the unrelated question X to estimate the proportion π_x of the unrelated character. An unbiased estimator of population proportion π is given by:

$$\hat{\pi}_{M_2} = \frac{\hat{\lambda}_{M_2} - p_2^2 - p_3(1 + p_2)\hat{\pi}_x}{p_1(1 + p_2)}, \quad (6)$$

taking $\hat{\lambda}_{M_2} \sim B(n_1, \lambda_{M_2})$ and $\hat{\pi}_x \sim B(n_2, \pi_x)$. The minimum variance of $\hat{\pi}_{M_2}$ is given by:

$$\begin{aligned} Var(\hat{\pi}_{M_2})_{\min} \\ = \frac{\left[\sqrt{\lambda_{M_2}(1 - \lambda_{M_2})} + p_3(1 + p_2)\sqrt{\pi_x(1 - \pi_x)} \right]^2}{n\{p_1(1 + p_2)\}^2}. \end{aligned} \quad (7)$$

The ratio of optimal sample sizes n_1 and n_2 for the two independent samples is given by:

$$\left(\frac{n_1}{n_2} \right)_{opt} = \sqrt{\frac{\lambda_{M_2}(1 - \lambda_{M_2})}{\{p_3(1 + p_2)\}^2 \pi_x(1 - \pi_x)}}. \quad (8)$$

Technique-III: This Technique differs from Technique-I in the sense that the second statement “I belong to group X^c ” is simply replaced by the statement “I belong to group A^c ”. The probability of a “Yes” response in the first sample of size n_1 respondent is determined as follows:

$$\lambda_{M_3} = \pi p_1 + p_2(1 - \pi) + p_3 \pi_x. \quad (9)$$

In the second independent sample of n_2 respondents, the question is asked only on the unrelated question X to estimate the proportion π_x of the unrelated character. An unbiased estimator for population proportion π is given by:

$$\hat{\pi}_{M_3} = \frac{\hat{\lambda}_{M_3} - p_2 - p_3 \hat{\pi}_x}{p_1 - p_2}, \quad (10)$$

taking $\hat{\lambda}_{M_3} \sim B(n_1, \lambda_{M_3})$ and $\hat{\pi}_x \sim B(n_2, \pi_x)$. The minimum variance of $\hat{\pi}_{M_3}$ is given by:

$$\begin{aligned} Var(\hat{\pi}_{M_3})_{\min} \\ = \frac{\left[\sqrt{\lambda_{M_3}(1 - \lambda_{M_3})} + p_3 \sqrt{\pi_x(1 - \pi_x)} \right]^2}{n(p_1 - p_2)^2}. \end{aligned} \quad (11)$$

The ratio of optimal sample sizes n_1 and n_2 for the two independent samples is given by:

$$\left(\frac{n_1}{n_2} \right)_{opt} = \sqrt{\frac{\lambda_{M_3}(1 - \lambda_{M_3})}{p_3^2 \pi_x(1 - \pi_x)}}. \quad (12)$$

In the next section, a new multi-question approach to estimating the prevalence of sensitive attribute and sensitivity level of modified binary optional RRT models is proposed.

3. Multi-question approach

The main motivation of the multi-question approach is to avoid split sample approach to parameter estimation which requires a larger sample size. To estimate the parameters of RRT models, i.e., unrelated innocuous attributes (π_x, π_y) , sensitivity level (ω) , and prevalence of sensitive attribute (π) in the multi-question approach, each respondent is asked more than two separate questions. In the multi-question approach, the Direct Question (DQ) method is applied first to estimate the proportion of unrelated innocuous attributes $(\hat{\pi}_x, \hat{\pi}_y)$ used in RRT models and, then, to apply two RDs, denoted by R_1 and R_2 , to estimate the proportion of sensitivity level $(\hat{\omega})$ and sensitive attribute $(\hat{\pi})$, respectively.

3.1. Estimation of sensitivity level

The question about sensitivity is asked via R_1 . In this randomization process, the sensitive question is “Is the main research question sensitive?” It is asked along with an unrelated innocuous question. We use R_1 device proposed by Greenberg et al. [3]. The probability of getting a “Yes” response to R_1 is:

$$\lambda_\omega = \omega p + (1 - p)\pi_y. \quad (13)$$

An unbiased estimator of population proportion ω is given by:

$$\hat{\omega} = \frac{\hat{\lambda}_\omega - (1 - p)\hat{\pi}_y}{p}, \quad (14)$$

where $\hat{\lambda}_\omega$ and $\hat{\pi}_y$ are the proportion of “Yes” answer reported by n respondents through R_1 and DQ method in the sample, taking $\hat{\lambda}_\omega \sim B(n, \lambda_\omega)$ and $\hat{\pi}_y \sim B(n, \pi_y)$. It is easy to prove that $\hat{\omega}$ is unbiased for ω with the following variance:

$$Var(\hat{\omega}) = \frac{1}{p^2} \left\{ \frac{\lambda_\omega(1 - \lambda_\omega)}{n} + (1 - p)^2 \frac{\pi_y(1 - \pi_y)}{n} \right\}. \quad (15)$$

3.2. Estimation of prevalence of the sensitive attribute

In this section, we propose three binary optional RRT models to estimate the prevalence of the sensitive attribute. Each respondent is given the option to answer the sensitive question directly if they consider question insensitive or provide the response after using device R_2 .

Technique-I: Optional RRT model

The probability of getting “Yes” response from the proposed binary optional RRT model of Technique-I or the main research question is measured as follows:

$$\lambda_1 = (1 - \omega)\pi + \omega \{p_1 \pi + p_2(1 - \pi_x) + p_3 \pi_x\}. \quad (16)$$

Solving Eq. (16) for π , we have an estimator of π as follows:

$$\hat{\pi}_{JS_1} = \frac{\hat{\lambda}_1 - \hat{\omega} \{p_2 - (p_2 - p_3) \hat{\pi}_x\}}{\{1 - \hat{\omega}(1 - p_1)\}}, \quad (17)$$

where $\hat{\omega}$ is obtained from Eq. (14) and $\hat{\lambda}_1$ and $\hat{\pi}_x$ are the proportions of “Yes” answers reported by n respondents through R_2 and DQ method in the sample, taking $\hat{\lambda}_1 \sim B(n, \lambda_1)$ and $\hat{\pi}_x \sim B(n, \pi_x)$. After applying first-order Taylor’s expansion to Eq. (17), we have:

$$\hat{\pi}_{JS_1} \approx \left[\begin{aligned} & \frac{\lambda_1 - \omega \{p_2 - (p_2 - p_3) \pi_x\}}{\{1 - \omega(1 - p_1)\}} \\ & + \frac{\hat{\lambda}_1 - \lambda_1}{\{1 - \omega(1 - p_1)\}} \\ & + \frac{\omega(p_2 - p_3)(\hat{\pi}_x - \pi_x)}{\{1 - \omega(1 - p_1)\}} \\ & + \frac{\theta_1(\hat{\omega} - \omega)}{\{1 - \omega(1 - p_1)\}^2} \end{aligned} \right], \quad (18)$$

where:

$$\theta_1 = (p_2 - p_3) \pi_x + (1 - p_1) \lambda_1 - p_2.$$

Theorem 3.1. The estimator $\hat{\pi}_{JS_1}$ is unbiased for π with variance as given below:

$$\begin{aligned} Var(\hat{\pi}_{JS_1}) &= \frac{\lambda_1(1 - \lambda_1) + \omega^2(p_2 - p_3)^2 \pi_x(1 - \pi_x)}{n\{1 - \omega(1 - p_1)\}^2} \\ &+ \frac{\theta_1^2 \{\lambda_\omega(1 - \lambda_\omega) + (1 - p)^2 \pi_y(1 - \pi_y)\}}{np^2 \{1 - \omega(1 - p_1)\}^4}. \end{aligned} \quad (19)$$

Proof. Taking expected value on both sides of Eq. (18), we have:

$$E(\hat{\pi}_{JS_1}) \approx \left[\begin{aligned} & \frac{\lambda_1 - \omega \{p_2 - (p_2 - p_3) \pi_x\}}{\{1 - \omega(1 - p_1)\}} \\ & + \frac{E(\hat{\lambda}_1) - \lambda_1}{\{1 - \omega(1 - p_1)\}} \\ & + \frac{\omega(p_2 - p_3)(E(\hat{\pi}_x) - \pi_x)}{\{1 - \omega(1 - p_1)\}} \\ & + \frac{\theta_1(E(\hat{\omega}) - \omega)}{\{1 - \omega(1 - p_1)\}^2} \end{aligned} \right]. \quad (20)$$

Given that $E(\hat{\lambda}_1) = \lambda_1$, $E(\hat{\pi}_x) = \pi_x$, $E(\hat{\lambda}_\omega) = \lambda_\omega$, and $E(\hat{\pi}_y) = \pi_y$, Eq. (20) becomes:

$$E(\hat{\pi}_{JS_1}) \approx \frac{\lambda_1 - \omega \{p_2 - (p_2 - p_3) \pi_x\}}{\{1 - \omega(1 - p_1)\}} = \pi. \quad (21)$$

Thus, $\hat{\pi}_{JS_1}$ is approximately an unbiased estimator of π . Now, applying variance on both sides of Eq. (18), we have:

$$\begin{aligned} Var(\hat{\pi}_{JS_1}) &= \left[\frac{Var(\hat{\lambda}_1) + \omega^2(p_2 - p_3)^2 Var(\hat{\pi}_x)}{\{1 - \omega(1 - p_1)\}^2} \right. \\ &\quad \left. + \frac{\theta_1^2 Var(\hat{\omega})}{\{1 - \omega(1 - p_1)\}^4} \right]. \end{aligned} \quad (22)$$

Using the fact that $Var(\hat{\lambda}_1) = \frac{\lambda_1(1 - \lambda_1)}{n}$ and $Var(\hat{\pi}_x) = \frac{\pi_x(1 - \pi_x)}{n}$ and substituting the values of $Var(\hat{\omega})$ given in Eq. (15), we have Eq. (19) that completes the proof.

Technique-II optional RRT model

In R_2 , the respondent has an option to answer the sensitive question directly or using Technique-II of Mahmood et al. [5]. Under R_2 , the probability of getting “Yes” response is:

$$\lambda_2 = (1 - \omega)\pi + \omega \{(1 + p_2)(p_1\pi + p_3\pi_x) + p_2^2\}. \quad (23)$$

Solving Eq. (23) for π , we have an estimator of π as follows:

$$\hat{\pi}_{JS_2} = \frac{\hat{\lambda}_2 - \hat{\omega} \{p_2^2 + p_3(1 + p_2)\hat{\pi}_x\}}{[1 - \hat{\omega}\{1 - p_1(1 + p_2)\}]}, \quad (24)$$

where $\hat{\omega}$ is obtained from Eq. (14) and $\hat{\lambda}_2$ and $\hat{\pi}_x$ are the proportions of “Yes” answers reported by n respondents through R_2 and DQ method in the sample by taking $\hat{\lambda}_2 \sim B(n, \lambda_2)$ and $\hat{\pi}_x \sim B(n, \pi_x)$.

After applying first-order Taylor’s expansion to Eq. (24), we have:

$$\hat{\pi}_{JS_2} \approx \left[\begin{aligned} & \frac{\hat{\lambda}_2 - \hat{\omega} \{p_2^2 + p_3(1 + p_2)\hat{\pi}_x\}}{[1 - \hat{\omega}\{1 - p_1(1 + p_2)\}]} \\ & + \frac{(\hat{\lambda}_2 - \lambda_2)}{[1 - \omega\{1 - p_1(1 + p_2)\}]} \\ & - \frac{\omega p_3(1 + p_2)(\hat{\pi}_x - \pi_x)}{[1 - \omega\{1 - p_1(1 + p_2)\}]} \\ & + \frac{\theta_2(\hat{\omega} - \omega)}{[1 - \omega\{1 - (1 + p_2)p_1\}]^2} \end{aligned} \right], \quad (25)$$

where:

$$\theta_2 = \lambda_2 - \{(1 + p_2)(p_3\pi_x + \lambda_2 p_1) + p_2^2\}.$$

Theorem 3.2. The estimator $\hat{\pi}_{JS_2}$ is an unbiased one for π with variance, as given by:

$$\begin{aligned} Var(\hat{\pi}_{JS_2}) &= \frac{\lambda_2(1 - \lambda_2) + \omega^2 \{p_3(1 + p_2)\}^2 \pi_x(1 - \pi_x)}{n[1 - \omega\{1 - p_1(1 + p_2)\}]^2} \\ &+ \frac{\theta_2^2 \{\lambda_\omega(1 - \lambda_\omega) + (1 - p)^2 \pi_y(1 - \pi_y)\}}{np^2 [1 - \omega\{1 - p_1(1 + p_2)\}]^4}. \end{aligned} \quad (26)$$

Proof. The proof is analogous to that of Theorem 3.1.

Technique-III optional RRT model

The probability of getting “Yes” response from the proposed binary optional RRT model is:

$$\lambda_3 = (1 - \omega)\pi + \omega \{p_1\pi + p_2(1 - \pi) + p_3\pi_x\}. \quad (27)$$

Solving Eq. (27) for π , we have an estimator of π as:

$$\hat{\pi}_{JS_3} = \frac{\hat{\lambda}_3 - \hat{\omega}(p_2 + p_3\hat{\pi}_x)}{\{1 - \hat{\omega}(1 - p_1 + p_2)\}}, \quad (28)$$

where $\hat{\omega}$ is obtained from Eq. (14) and $\hat{\lambda}_3$ and $\hat{\pi}_x$ are the proportions of “Yes” answers reported by n respondents through R_2 and DQ method in the sample by taking $\lambda_3 \sim B(n, \lambda_3)$ and $\pi_x \sim B(n, \pi_x)$. After applying first-order Taylor’s expansion to Eq. (28), we have:

$$\hat{\pi}_{JS_3} \approx \left[\begin{aligned} & \frac{\lambda_3 - \omega(p_2 + p_3\pi_x)}{\{1 - \omega(1 - p_1 + p_2)\}} \\ & + \frac{\hat{\lambda}_3 - \lambda_3}{\{1 - \omega(1 - p_1 + p_2)\}} \\ & - \frac{\omega p_3(\hat{\pi}_x - \pi_x)}{\{1 - \omega(1 - p_1 + p_2)\}} \\ & + \frac{\theta_3(\hat{\omega} - \omega)}{\{1 - \omega(1 - p_1 + p_2)\}^2} \end{aligned} \right], \quad (29)$$

where:

$$\theta_3 = \lambda_3(1 - p_1 + p_2) - \pi_x p_3 - p_2.$$

Thus, we have the following theorem and its proof is similar to that of Theorem 3.1.

Theorem 3.3. The estimator $\hat{\pi}_{JS_3}$ is an unbiased one for π with variance as given by:

$$\begin{aligned} Var(\hat{\pi}_{JS_3}) = & \frac{\lambda_3(1 - \lambda_3) + \omega^2 p_3^2 \pi_x(1 - \pi_x)}{n\{1 - \omega(1 - p_1 + p_2)\}^2} \\ & + \frac{\theta_3^2 \{\lambda_\omega(1 - \lambda_\omega) + (1 - p)^2 \pi_y(1 - \pi_y)\}}{np^2 \{1 - \omega(1 - p_1 + p_2)\}^4}. \end{aligned} \quad (30)$$

3.3. Privacy protection

Lanke [25] proposed a measure of privacy protection based on the idea that a person belonging to the sensitive group A might be reluctant to reveal the actual group, whereas a member of the complementary group A^c is expected to be quite willing to announce so. Thus, membership in A may be embarrassing, except for that in A^c . Thus, the higher the conditional probability $P(A|R)$, $R \in \{\text{“Yes”}, \text{“No”}\}$ of belonging to A given a certain answer R , the greater the embarrassment

caused by giving that response. The Lanke measure (hereafter L -measure) is:

$$L = \max\{P(A|Yes), P(A|No)\}. \quad (31)$$

However, it takes $L = P(A|Yes)$ when:

$$P(A|Yes) - P(A|No) = \frac{\pi[P(Yes|A) - P(Yes)]}{P(Yes)[1 - P(Yes)]} > 0,$$

that is when:

$$\begin{aligned} P(Yes|A) - P(Yes) = & (1 - \pi)[P(Yes|A) \\ & - P(Yes|A^c)] > 0. \end{aligned}$$

This is true for all the considered RRT models. Therefore, for these randomization procedures, L -measure reduces to:

$$L = \frac{\pi P(Yes|A)}{P(Yes)}. \quad (32)$$

The L -measure for the RRT models is given below:

$$L_{M_1} = \frac{\pi[p_1 + p_2(1 - \pi_x) + p_3\pi_x]}{\pi p_1 + p_2(1 - \pi_x) + p_3\pi_x}, \quad (33)$$

$$L_{M_2} = \frac{\pi[(1 + p_2)(p_1 + p_3\pi_x) + p_2^2]}{(1 + p_2)(p_1\pi + p_3\pi_x) + p_2^2}, \quad (34)$$

and:

$$L_{M_3} = \frac{\pi[p_1 + p_3\pi_x]}{\pi p_1 + p_2(1 - \pi) + p_3\pi_x}. \quad (35)$$

The L -measure for the proposed optional RRT models is given by:

$$L_{JS_1} = \frac{\pi[(1 - \omega) + \omega\{p_1 + p_2(1 - \pi_x) + p_3\pi_x\}]}{(1 - \omega)\pi + \omega\{p_1\pi + p_2(1 - \pi_x) + p_3\pi_x\}}, \quad (36)$$

$$L_{JS_2} = \frac{\pi[(1 - \omega) + \omega\{(1 + p_2)(p_1 + p_3\pi_x) + p_2^2\}]}{(1 - \omega)\pi + \omega\{(1 + p_2)(p_1\pi + p_3\pi_x) + p_2^2\}}, \quad (37)$$

and:

$$L_{JS_3} = \frac{\pi[(1 - \omega) + \omega\{p_1 + p_3\pi_x\}]}{(1 - \omega)\pi + \omega\{p_1\pi + p_2(1 - \pi) + p_3\pi_x\}}. \quad (38)$$

Based on the two RRT models, the model with the smaller value of Lanke [25] privacy measure will be more protective than the other.

If the inequality given below holds true, then the proposed optional RRT models are more protected than competitor models, respectively.

$$L_{JS_i} - L_{M_i} < 0, \quad (39)$$

where $i = 1, 2, 3$.

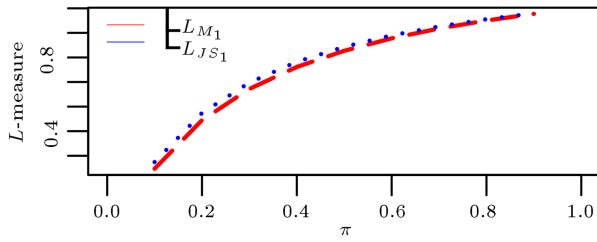


Figure 1. Measure of privacy protection of Technique-I optional RRT model with respect to Mahmood et al. [5] when $\omega = 0.9$, $\pi_x = 0.35$, $p_1 = 0.7$, $p_3 = 2(1 - p_1)/3$, and $p_2 = (1 - p_1 - p_3)$.

Inequality (39) under L_{JS_1} and L_{M_1} is:

$$(1 - \omega)(1 - \pi)\{p_2 - (p_2 - p_3)\pi_x\} < 0, \quad (40)$$

Inequality (39) under L_{JS_2} and L_{M_2} is:

$$(1 - \omega)(1 - \pi)\{p_2^2 + p_3(1 + p_2)\pi_x\} < 0, \quad (41)$$

Inequality (39) under L_{JS_3} and L_{M_3} is:

$$(1 - \omega)(1 - \pi)\{p_2 + p_3\pi_x\} < 0. \quad (42)$$

Privacy of the respondents and gain in efficiency are two alternating measures. The efficiency of the RRT model decreases when the degree of protection to the respondents increases, and vice versa. Theoretically, we observe that the proposed optional RRT models are less protective than competitor RRT models. However, as seen in Figure 1, we observe that for a highly sensitive question ($\omega = 0.9$), the proposed Technique-I optional RRT model and competitor model are equally protected at different values of design parameters. In the case of Technique-II and Technique-III, the same conclusion can be drawn. Chhabra et al. [21] also highlighted that when the asked question is highly sensitive, the proposed optional models achieved the same protection level as the corresponding RRT model. Our conclusion is consistent with that of Chhabra et al. [21].

3.4. Simulation study

In this section, the theoretical and empirical results of the proposed estimators $\hat{\pi}_{JS_i}$ ($i = 1, 2, 3$) and $\hat{\omega}$ are presented. The parameters ω , π_y , p_1 , $p_2 = (1 - p_1 - p_3)$, and $p_3 = 2(1 - p_1)/3$ were allowed to vary, while all others were fixed, $\pi = 0.45$, $\pi_x = 0.85$, $p = 0.7$, $n = 1000$, and 10,000 trials per simulation. For the simulation of $\hat{\pi}_{JS_i}$ and $\hat{\omega}$, the value of ω varies from 0.1 to 0.9 with the increment of 0.2. Similarly, the value of p_1 varies from 0.2 to 0.8 with the increment of 0.2. Moreover, π_y values are chosen as 0.35 and 0.65, respectively.

The simulated results of $\hat{\pi}_{JS_i}$ ($i = 1, 2, 3$) and $\hat{\omega}$ of optional RRT models are obtained using the multi-question approach which provides a strong support

for our earlier findings that $\hat{\pi}_{JS_i}$ and $\hat{\omega}$ are unbiased estimators of π and ω , respectively. As shown in Tables 1–3, the simulated values of $E(\hat{\pi})$ and $Var(\hat{\pi})$ are very close to the corresponding true parameter values of π and theoretical values of $Var(\hat{\pi})$, respectively. The first-order Taylor's approximation was used to calculate the theoretical values for $Var(\hat{\pi})$. Similarly, simulated values of $E(\hat{\omega})$ and $Var(\hat{\omega})$ are very close to the corresponding true parameter values of ω and theoretical value of $Var(\hat{\omega})$, respectively. The results presented in Tables 1–3 indicate that the proposed optional RRT models of the multi-question approach are more efficient than competitor RRT models. The results of each model with respect to their competitor RRT models are discussed below:

- (i) The findings in Table 1 indicate that Technique-I, i.e., proposed optional RRT model, is more efficient than that proposed by Mahmood et al. [5]. For example, the results in Table 1 for $\omega = 0.5$, $p_1 = 0.6$, and $\pi_y = 0.35$ show the empirical variance of $Var(\hat{\pi}) = 0.0004$ and theoretical variance of $Var(\hat{\pi}) = 0.0004$ for the proposed optional RRT model, respectively. However, the empirical and theoretical variances for the model of Mahmood et al. [5] include $Var(\hat{\pi}) = 0.0007$ and $Var(\hat{\pi}) = 0.0006$, respectively, which demonstrate greater variance than that for the proposed optional RRT model. Similar efficiency patterns are observed at different values of ω , p_1 , and π_y ;
- (ii) The results given in Table 2 show that for $\omega = 0.5$, $p_1 = 0.6$, and $\pi_y = 0.35$, empirical variance of $Var(\hat{\pi}) = 0.0003$ and theoretical variance of $Var(\hat{\pi}) = 0.0004$ for the Technique-II optional RRT model are achieved. However, the empirical and theoretical variances of Technique-II of Mahmood et al. [5] are $Var(\hat{\pi}) = 0.0005$ and $Var(\hat{\pi}) = 0.0008$, respectively, which demonstrate that the Technique-II optional RRT model is more efficient than the competitor model. The efficiency pattern remains similar for different values of ω , p_1 , and π_y ;
- (iii) Table 3 demonstrates that for $\omega = 0.5$, p_1 , and $\pi_y = 0.35$, the empirical and theoretical variances of Technique-III optional RRT model are $Var(\hat{\pi}) = 0.0004$ and $Var(\hat{\pi}) = 0.0004$, respectively. The result shows that Technique-III optional RRT model is more efficient than the competitor Technique-III of Mahmood et al. [5]. The empirical and theoretical variances of Technique-III of Mahmood et al. [5] are $Var(\hat{\pi}) = 0.0011$ and $Var(\hat{\pi}) = 0.0016$, respectively, indicating to its lower efficiency than that of the proposed optional RRT model. The same variance patterns are observed at different values of ω , p_1 , and π_y .

Table 1. Simulation results for $\pi = 0.45$ when $trials = 10,000$, $n = 1000$, $\pi_x = 0.85$, $\pi_y = 0.35, 0.65$, $p = 0.7$ and at various choices of ω , $p_1, p_3 = 2(1 - p_1)/3$, and $p_2 = (1 - p_1 - p_3)$.

p_1		Technique-I Mahmood et al. [5]	Technique-I optional RRT model									
			$\pi_y = 0.35$					$\pi_y = 0.65$				
			ω					ω				
			0.1	0.3	0.5	0.7	0.9	0.1	0.3	0.5	0.7	0.9
0.2	Empirical mean ($\hat{\pi}$)	0.4499	0.4382	0.4469	0.4594	0.4763	0.5037	0.4218	0.4289	0.4394	0.4540	0.4799
	Empirical var ($\hat{\pi}$)	0.0059	0.0003	0.0004	0.0006	0.0009	0.0014	0.0004	0.0005	0.0008	0.0012	0.0021
	Theoretical var ($\hat{\pi}$)	0.0040	0.0003	0.0004	0.0007	0.0014	0.0033	0.0003	0.0005	0.0007	0.0013	0.0033
	Empirical mean ($\hat{\omega}$)		0.1000	0.3002	0.5002	0.7002	0.9002	0.1002	0.3000	0.5000	0.7000	0.8999
	Empirical var ($\hat{\omega}$)		0.0003	0.0004	0.0005	0.0004	0.0004	0.0004	0.0004	0.0005	0.0004	0.0003
	Theoretical var ($\hat{\omega}$)		0.0003	0.0005	0.0005	0.0005	0.0004	0.0004	0.0005	0.0005	0.0005	0.0003
0.4	Empirical mean ($\hat{\pi}$)	0.4495	0.4417	0.4480	0.4563	0.4665	0.4788	0.4305	0.4359	0.4435	0.4528	0.4641
	Empirical var ($\hat{\pi}$)	0.0015	0.0003	0.0004	0.0005	0.0006	0.0008	0.0004	0.0004	0.0006	0.0007	0.0010
	Theoretical var ($\hat{\pi}$)	0.0011	0.0003	0.0004	0.0005	0.0008	0.0012	0.0003	0.0004	0.0005	0.0008	0.0012
	Empirical mean ($\hat{\omega}$)		0.1001	0.3003	0.4997	0.7002	0.8999	0.0999	0.3000	0.4998	0.7002	0.8999
	Empirical var ($\hat{\omega}$)		0.0003	0.0004	0.0005	0.0005	0.0004	0.0004	0.0005	0.0005	0.0004	0.0002
	Theoretical var ($\hat{\omega}$)		0.0003	0.0005	0.0005	0.0005	0.0004	0.0004	0.0005	0.0005	0.0005	0.0003
0.6	Empirical mean ($\hat{\pi}$)	0.4501	0.4447	0.4488	0.4539	0.4593	0.4655	0.4377	0.4416	0.4464	0.4515	0.4574
	Empirical var ($\hat{\pi}$)	0.0007	0.0003	0.0003	0.0004	0.0004	0.0005	0.0003	0.0004	0.0004	0.0005	0.0006
	Theoretical var ($\hat{\pi}$)	0.0006	0.0003	0.0003	0.0004	0.0005	0.0006	0.0003	0.0003	0.0004	0.0005	0.0006
	Empirical mean ($\hat{\omega}$)		0.1000	0.2997	0.5000	0.7002	0.9001	0.0999	0.3000	0.5003	0.7004	0.9001
	Empirical var ($\hat{\omega}$)		0.0003	0.0004	0.0005	0.0004	0.0004	0.0004	0.0005	0.0005	0.0004	0.0003
	Theoretical var ($\hat{\omega}$)		0.0003	0.0005	0.0005	0.0005	0.0004	0.0004	0.0005	0.0005	0.0005	0.0003
0.8	Empirical mean ($\hat{\pi}$)	0.4499	0.4475	0.4494	0.4515	0.4538	0.4566	0.4443	0.4462	0.4482	0.4504	0.4532
	Empirical var ($\hat{\pi}$)	0.0004	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0004
	Theoretical var ($\hat{\pi}$)	0.0004	0.0003	0.0003	0.0003	0.0003	0.0004	0.0003	0.0003	0.0003	0.0003	0.0004
	Empirical mean ($\hat{\omega}$)		0.1001	0.2998	0.4998	0.7000	0.9000	0.1002	0.2996	0.4998	0.7002	0.8998
	Empirical var ($\hat{\omega}$)		0.0002	0.0004	0.0005	0.0004	0.0004	0.0004	0.0005	0.0005	0.0004	0.0002
	Theoretical var ($\hat{\omega}$)		0.0003	0.0005	0.0005	0.0005	0.0004	0.0004	0.0005	0.0005	0.0005	0.0003

While comparing the results of the proposed optional RRT models with competitor RRT models, one should keep in mind that optional models simultaneously estimate the prevalence of sensitive attribute (π) and sensitivity level (ω). The proposed optional RRT models are characterized by their own merits and demerits; therefore, different models can be used under different scenarios.

4. Stratified random sampling scheme

In this section, the applications of RRT models proposed by Mahmood et al. [5] and the suggested optional RRT models under stratified random sampling are discussed; in this regard, the population consists of

L number of strata and the size of each stratum N_h is known. Let the population of size N be divided into disjoint L strata of size N_h ($h = 1, 2, \dots, L$) and $W_h = (\frac{N_h}{N})$ is the known proportion of population units falling in the h th stratum. Let n_h ($\sum_{h=1}^L n_h = n$) respondents be selected by SRSWR and asked to answer “Yes” or “No” according to the RDs.

4.1. The stratified Technique-I of Mahmood et al. [5]

Each respondent in stratum h is provided an RD that consists of three types of statements:

- (i) I belong to group A_h ;
- (ii) I belong to group X_h^c ;

Table 2. Simulation results for $\pi = 0.45$ when $trials = 10,000$, $n = 1000$, $\pi_x = 0.85$, $\pi_y = 0.35, 0.65$, $p = 0.7$ and at various choices of ω , $p_1, p_3 = 2(1 - p_1)/3$, and $p_2 = (1 - p_1 - p_3)$.

p_1		Technique-II Mahmood et al. [5]	Technique-II optional RRT model									
			$\pi_y = 0.35$					$\pi_y = 0.65$				
			ω					ω				
			0.1	0.3	0.5	0.7	0.9	0.1	0.3	0.5	0.7	0.9
0.2	Empirical mean ($\hat{\pi}$)	0.6501	0.4290	0.4634	0.5092	0.5721	0.6643	0.3924	0.4249	0.4690	0.5319	0.6294
	Empirical var ($\hat{\pi}$)	0.0019	0.0003	0.0004	0.0005	0.0006	0.0007	0.0004	0.0005	0.0006	0.0008	0.0010
	Theoretical var ($\hat{\pi}$)	0.0070	0.0003	0.0005	0.0007	0.0013	0.0026	0.0003	0.0005	0.0007	0.0012	0.0025
	Empirical mean ($\hat{\omega}$)		0.1000	0.3002	0.5002	0.7002	0.9002	0.1002	0.3000	0.5000	0.7000	0.8999
	Empirical var ($\hat{\omega}$)		0.0003	0.0004	0.0005	0.0004	0.0004	0.0004	0.0004	0.0005	0.0004	0.0003
	Theoretical var ($\hat{\omega}$)		0.0003	0.0005	0.0005	0.0005	0.0004	0.0004	0.0005	0.0005	0.0005	0.0003
0.4	Empirical mean ($\hat{\pi}$)	0.5660	0.4385	0.4662	0.4995	0.5394	0.5886	0.4156	0.4428	0.4759	0.5160	0.5663
	Empirical var ($\hat{\pi}$)	0.0008	0.0003	0.0003	0.0004	0.0004	0.0005	0.0003	0.0004	0.0004	0.0005	0.0006
	Theoretical var ($\hat{\pi}$)	0.0018	0.0003	0.0004	0.0005	0.0007	0.0009	0.0003	0.0004	0.0005	0.0007	0.0009
	Empirical mean ($\hat{\omega}$)		0.1001	0.3003	0.4997	0.7002	0.8999	0.0999	0.3000	0.4998	0.7002	0.8999
	Empirical var ($\hat{\omega}$)		0.0003	0.0004	0.0005	0.0005	0.0004	0.0004	0.0005	0.0005	0.0004	0.0002
	Theoretical var ($\hat{\omega}$)		0.0003	0.0005	0.0005	0.0005	0.0004	0.0004	0.0005	0.0005	0.0005	0.0003
0.6	Empirical mean ($\hat{\pi}$)	0.5227	0.4448	0.4642	0.4859	0.5096	0.5362	0.4318	0.4511	0.4729	0.4967	0.5236
	Empirical var ($\hat{\pi}$)	0.0005	0.0003	0.0003	0.0003	0.0004	0.0004	0.0003	0.0003	0.0004	0.0004	0.0004
	Theoretical var ($\hat{\pi}$)	0.0008	0.0003	0.0003	0.0004	0.0004	0.0005	0.0003	0.0003	0.0004	0.0004	0.0005
	Empirical mean ($\hat{\omega}$)		0.1000	0.2997	0.5000	0.7002	0.9001	0.0999	0.3000	0.5003	0.7004	0.9001
	Empirical var ($\hat{\omega}$)		0.0003	0.0004	0.0005	0.0004	0.0004	0.0004	0.0005	0.0005	0.0004	0.0003
	Theoretical var ($\hat{\omega}$)		0.0003	0.0005	0.0005	0.0005	0.0004	0.0004	0.0005	0.0005	0.0005	0.0003
0.8	Empirical mean ($\hat{\pi}$)	0.4857	0.4487	0.4587	0.4692	0.4802	0.4921	0.4431	0.4531	0.4637	0.4747	0.4867
	Empirical var ($\hat{\pi}$)	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
	Theoretical var ($\hat{\pi}$)	0.0004	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
	Empirical mean ($\hat{\omega}$)		0.1001	0.2998	0.4998	0.7000	0.9000	0.1002	0.2996	0.4998	0.7002	0.8998
	Empirical var ($\hat{\omega}$)		0.0002	0.0004	0.0005	0.0004	0.0004	0.0004	0.0005	0.0005	0.0004	0.0002
	Theoretical var ($\hat{\omega}$)		0.0003	0.0005	0.0005	0.0005	0.0004	0.0004	0.0005	0.0005	0.0005	0.0003

(iii) I belong to group X_h with probabilities p_{h_1} , p_{h_2} , and p_{h_3} , respectively, such that $\sum_{i=1}^3 p_{h_i} = 1$.

The proportion of the unrelated question π_{x_h} is unknown; therefore, estimating π_{x_h} needs another sample of size n_{h_2} such that $(n_{h_1} + n_{h_2} = n_h)$. The probability of a “Yes” response in the first sample of size n_{h_1} respondents is given by:

$$\lambda_{M_{h_1}} = \pi_h p_{h_1} + p_{h_2}(1 - \pi_{x_h}) + p_{h_3} \pi_{x_h}. \quad (43)$$

In the second independent samples of n_{h_2} respondents, the question is asked only on the unrelated question X_h to estimate the proportion π_{x_h} of the unrelated character. An unbiased estimator for the population

proportion $\hat{\pi}_{M_{1(st)}}$ is given below:

$$\hat{\pi}_{M_{1(st)}} = \sum_{h=1}^L W_h \left[\frac{\hat{\lambda}_{M_{h_1}} - p_{h_2}(1 - \hat{\pi}_{x_h}) - p_{h_3} \hat{\pi}_{x_h}}{p_{h_1}} \right]. \quad (44)$$

The variance of $\hat{\pi}_{M_{1(st)}}$ is measured as follows:

$$Var(\hat{\pi}_{M_{1(st)}}) = \sum_{h=1}^L \frac{W_h^2}{n_h} V_{M_1}, \quad (45)$$

where:

$$V_{M_1} = \frac{\left[\sqrt{\lambda_{M_{h_1}}(1 - \lambda_{M_{h_1}})} + (p_{h_2} - p_{h_3}) \sqrt{\pi_{x_h}(1 - \pi_{x_h})} \right]^2}{p_{h_1}^2}. \quad (46)$$

Table 3. Simulation results for $\pi = 0.45$ when $trials = 10,000$, $n = 1000$, $\pi_x = 0.85$, $\pi_y = 0.35, 0.65$, $p = 0.7$ and at various choices of ω , $p_1, p_3 = 2(1 - p_1)/3$, and $p_2 = (1 - p_1 - p_3)$.

		Technique-III optional RRT model										
p_1	Technique-III Mahmood et al. [5]	$\pi_y = 0.35$					$\pi_y = 0.65$					
		ω					ω					
		0.1	0.3	0.5	0.7	0.9	0.1	0.3	0.5	0.7	0.9	
0.2	Empirical mean ($\hat{\pi}$)	0.4493	0.4276	0.4440	0.4705	0.5187	0.6337	0.3944	0.4051	0.4237	0.4629	0.6006
	Empirical var ($\hat{\pi}$)	0.0401	0.0004	0.0006	0.0009	0.0017	0.0042	0.0005	0.0008	0.0013	0.0031	0.0140
	Theoretical var ($\hat{\pi}$)	0.0959	0.0003	0.0005	0.0008	0.0014	0.0032	0.0003	0.0005	0.0008	0.0014	0.0032
	Empirical mean ($\hat{\omega}$)		0.1000	0.3002	0.5002	0.7002	0.9002	0.1002	0.3000	0.5000	0.7000	0.8999
	Empirical var ($\hat{\omega}$)		0.0003	0.0004	0.0005	0.0004	0.0004	0.0004	0.0004	0.0005	0.0004	0.0003
	Theoretical var ($\hat{\omega}$)		0.0003	0.0005	0.0005	0.0005	0.0004	0.0004	0.0005	0.0005	0.0005	0.0003
0.4	Empirical mean ($\hat{\pi}$)	0.4484	0.4343	0.4462	0.4630	0.4865	0.5216	0.4123	0.4219	0.4359	0.4565	0.4892
	Empirical var ($\hat{\pi}$)	0.0052	0.0003	0.0004	0.0006	0.0009	0.0013	0.0004	0.0005	0.0008	0.0011	0.0019
	Theoretical var ($\hat{\pi}$)	0.0098	0.0003	0.0004	0.0005	0.0008	0.0012	0.0003	0.0004	0.0005	0.0008	0.0012
	Empirical mean ($\hat{\omega}$)		0.1001	0.3003	0.4997	0.7002	0.8999	0.0999	0.3000	0.4998	0.7002	0.8999
	Empirical var ($\hat{\omega}$)		0.0003	0.0004	0.0005	0.0005	0.0004	0.0004	0.0005	0.0005	0.0004	0.0002
	Theoretical var ($\hat{\omega}$)		0.0003	0.0005	0.0005	0.0005	0.0004	0.0004	0.0005	0.0005	0.0005	0.0003
0.6	Empirical mean ($\hat{\pi}$)	0.4501	0.4401	0.4477	0.4574	0.4685	0.4824	0.4270	0.4338	0.4426	0.4529	0.4659
	Empirical var ($\hat{\pi}$)	0.0011	0.0003	0.0004	0.0004	0.0005	0.0007	0.0003	0.0004	0.0005	0.0006	0.0008
	Theoretical var ($\hat{\pi}$)	0.0016	0.0003	0.0003	0.0004	0.0005	0.0006	0.0003	0.0003	0.0004	0.0005	0.0006
	Empirical mean ($\hat{\omega}$)		0.1000	0.2997	0.5000	0.7002	0.9001	0.0999	0.3000	0.5003	0.7004	0.9001
	Empirical var ($\hat{\omega}$)		0.0003	0.0004	0.0005	0.0004	0.0004	0.0004	0.0005	0.0005	0.0004	0.0003
	Theoretical var ($\hat{\omega}$)		0.0003	0.0005	0.0005	0.0005	0.0004	0.0004	0.0005	0.0005	0.0005	0.0003
0.8	Empirical mean ($\hat{\pi}$)	0.4498	0.4454	0.4490	0.4530	0.4574	0.4625	0.4395	0.4429	0.4467	0.4510	0.4560
	Empirical var ($\hat{\pi}$)	0.0005	0.0003	0.0003	0.0003	0.0003	0.0004	0.0003	0.0003	0.0003	0.0004	0.0004
	Theoretical var ($\hat{\pi}$)	0.0006	0.0003	0.0003	0.0003	0.0003	0.0004	0.0003	0.0003	0.0003	0.0003	0.0004
	Empirical mean ($\hat{\omega}$)		0.1001	0.2998	0.4998	0.7000	0.9000	0.1002	0.2996	0.4998	0.7002	0.8998
	Empirical var ($\hat{\omega}$)		0.0002	0.0004	0.0005	0.0004	0.0004	0.0004	0.0005	0.0005	0.0004	0.0002
	Theoretical var ($\hat{\omega}$)		0.0003	0.0005	0.0005	0.0005	0.0004	0.0004	0.0005	0.0005	0.0005	0.0003

4.2. The stratified Technique-II of Mahmood et al. [5]

The RD provided in each stratum consists of three types of statements:

- (i) I belong to group A_h ;
- (ii) Try once again;

- (iii) I belong to group X_h with probabilities p_{h_1}, p_{h_2} , and p_{h_3} , respectively, such that $\sum_{i=1}^3 p_{h_i} = 1$.

If the statement “Try once again” appears on the second trial, then the respondent is requested to report “Yes” irrespective of his/her actual status. The proportion of the unrelated question π_{x_h} is unknown;

thus, estimating π_{x_h} needs another sample of size n_{h_2} such that $(n_{h_1} + n_{h_2} = n_h)$. The probability of a “Yes” response in the first sample of size n_{h_1} respondents is assessed below:

$$\lambda_{M_{h_2}} = (1 + p_{h_2})(p_{h_1}\pi_h + p_{h_3}\pi_{x_h}) + p_{h_2}^2. \quad (47)$$

In the second independent sample of n_{h_2} respondents, the question is asked only on the unrelated question X_h to estimate the proportion π_{x_h} of the unrelated character. An unbiased estimator of population proportion $\hat{\pi}_{M_{2(st)}}$ is given by:

$$\hat{\pi}_{M_{2(st)}} = \sum_{h=1}^L W_h \left[\frac{\hat{\lambda}_{M_{h_2}} - p_{h_2}^2 - p_{h_3}(1 + p_{h_2})\hat{\pi}_{x_h}}{p_{h_1}(1 + p_{h_2})} \right]. \quad (48)$$

The variance of $\hat{\pi}_{M_{2(st)}}$ is estimated as follows:

$$\text{Var}(\hat{\pi}_{M_{2(st)}}) = \sum_{h=1}^L \frac{W_h^2}{n_h} V_{M_2}, \quad (49)$$

where:

$$V_{M_2} = \frac{\left[\sqrt{\lambda_{M_{h_2}}(1 - \lambda_{M_{h_2}})} + p_{h_3}(1 + p_{h_2})\sqrt{\pi_{x_h}(1 - \pi_{x_h})} \right]^2}{\{p_{h_1}(1 + p_{h_2})\}^2}. \quad (50)$$

4.3. The stratified Technique-III of Mahmood et al. [5]

Each respondent in stratum h provided with an RD that consists of three types of statements:

- (i) I belong to group A_h ;
- (ii) I belong to group A_h^c ;
- (iii) I belong to group X_h with probabilities p_{h_1} , p_{h_2} , and p_{h_3} , such that $\sum_{i=1}^3 p_{h_i} = 1$.

The proportion of the unrelated question π_{x_h} is unknown; thus, estimating π_{x_h} needs another sample of size n_{h_2} such that $(n_{h_1} + n_{h_2} = n_h)$. The probability of a “Yes” response in the first sample of size n_{h_1} respondents is given by:

$$\lambda_{M_{h_3}} = \pi_h p_{h_1} + p_{h_2}(1 - \pi_h) + p_{h_3}\pi_{x_h}. \quad (51)$$

In the second independent sample of n_{h_2} respondents, the question is asked only on the unrelated question X_h to estimate the proportion π_{x_h} of the unrelated character. An unbiased estimator of population proportion $\hat{\pi}_{M_{3(st)}}$ is given by:

$$\hat{\pi}_{M_{3(st)}} = \sum_{h=1}^L W_h \left[\frac{\hat{\lambda}_{M_{h_3}} - p_{h_2} - p_{h_3}\hat{\pi}_{x_h}}{p_{h_1} - p_{h_2}} \right]. \quad (52)$$

The variance of $\hat{\pi}_{M_{3(st)}}$ is assessed as follows:

$$\text{Var}(\hat{\pi}_{M_{3(st)}}) = \sum_{h=1}^L \frac{W_h^2}{n_h} V_{M_3}, \quad (53)$$

where:

$$V_{M_3} = \frac{\left[\sqrt{\lambda_{M_{h_3}}(1 - \lambda_{M_{h_3}})} + p_{h_3}\sqrt{\pi_{x_h}(1 - \pi_{x_h})} \right]^2}{(p_{h_1} - p_{h_2})^2}. \quad (54)$$

4.4. A stratified multi-question approach

To estimate the parameters of RRT models, i.e., unrelated innocuous characteristics (π_x, π_y), sensitivity level (ω), and prevalence of sensitive characteristic (π) in the multi-question approach, each respondent in stratum h is asked more than two separate questions. In the multi-question approach, DQ method is applied first to estimate the proportion of unrelated innocuous characteristics ($\hat{\pi}_{x_h}, \hat{\pi}_{y_h}$) used in RRT models and, then, apply two RDs, denoted by R_{h_1} and R_{h_2} , to estimates the proportion of sensitivity level ($\hat{\omega}_h$) and sensitive characteristic ($\hat{\pi}_h$) respectively.

4.4.1. Estimation of sensitivity level under stratification

The question about sensitivity is asked via R_{h_1} in stratum h . In this randomization process, the sensitive question is “Is the main research question sensitive?” It is asked along with an unrelated innocuous question. We use the device R_{h_1} suggested by Greenberg et al. [3]. The probability of getting “Yes” response to the R_{h_1} is:

$$\lambda_{\omega_h} = \omega_h p_h + (1 - p_h)\pi_{y_h}. \quad (55)$$

The estimator for ω_h is given by:

$$\hat{\omega}_h = \frac{\hat{\lambda}_{\omega_h} - (1 - p_h)\hat{\pi}_{y_h}}{p_h}, \quad (56)$$

where $\hat{\lambda}_{\omega_h}$ is the proportion of “yYs” answer reported by n respondents in the sample $\hat{\lambda}_{\omega_h} \sim B(n_h, \lambda_{\omega_h})$ and $\hat{\pi}_{y_h} \sim B(n_h, \pi_{y_h})$. An unbiased estimator for population proportion ω is given by:

$$\hat{\omega}_{(st)} = \sum_{h=1}^L W_h \left[\frac{\hat{\lambda}_{\omega_h} - (1 - p_h)\hat{\pi}_{y_h}}{p_h} \right]. \quad (57)$$

It is easy to prove that $\hat{\omega}_{(st)}$ is unbiased for ω with variance:

$$\text{Var}(\hat{\omega}_{(st)}) = \sum_{h=1}^L W_h^2 \text{Var}(\hat{\omega}_h), \quad (58)$$

where:

$$\text{Var}(\hat{\omega}_h) = \frac{1}{p_h^2} \left\{ \frac{\lambda_{\omega_h}(1 - \lambda_{\omega_h})}{n_h} + (1 - p_h)^2 \frac{\pi_{y_h}(1 - \pi_{y_h})}{n_h} \right\}. \quad (59)$$

Now, the respondents are given an option to answer the sensitive question directly if they consider question insensitive or provide the response after using device R_{h_2} .

4.4.2. Estimation of sensitive attribute

Each respondent in stratum h is given the option to answer the sensitive question directly if they consider the question insensitive or provide the response after using device R_{h_2} .

Technique-I optional RRT model under stratification

The probability of getting “Yes” response in stratum h of the proposed binary optional RRT model of Technique-I is:

$$\lambda_{h_1} = (1 - \omega_h)\pi_h + \omega_h\{\pi_h p_{h_1} + p_{h_2}(1 - \pi_{x_h}) + p_{h_3}\pi_{x_h}\}. \quad (60)$$

Solving Eq. (60) for π_h , we have an estimator of π_h as:

$$\hat{\pi}_{JS_{h_1}} = \frac{\hat{\lambda}_{h_1} - \hat{\omega}_h\{p_{h_2} - (p_{h_2} - p_{h_3})\hat{\pi}_{x_h}\}}{\{1 - \hat{\omega}_h(1 - p_{h_1})\}}, \quad (61)$$

where $\hat{\omega}_h$ is obtained from Eq. (56) and $\hat{\lambda}_{h_1}$ is the proportion of “Yes” answers reported by n_h respondents in the sample, $\hat{\lambda}_{h_1} \sim B(n_h, \lambda_{h_1})$ and $\hat{\pi}_{x_h} \sim B(n_h, \pi_{x_h})$. After applying the first-order Taylor’s expansion to Eq. (61), we have:

$$\hat{\pi}_{JS_{h_1}} \approx \left[\begin{aligned} &\frac{\lambda_{h_1} - \omega\{p_{h_2} - (p_{h_2} - p_{h_3})\pi_{x_h}\}}{\{1 - \omega_h(1 - p_{h_1})\}} \\ &+ \frac{\hat{\lambda}_{h_1} - \lambda_{h_1}}{\{1 - \omega_h(1 - p_{h_1})\}} \\ &+ \frac{\omega_h(p_{h_2} - p_{h_3})(\hat{\pi}_{x_h} - \pi_{x_h})}{\{1 - \omega_h(1 - p_{h_1})\}} \\ &+ \frac{\theta_{h_1}(\hat{\omega}_h - \omega_h)}{\{1 - \omega_h(1 - p_{h_1})\}^2} \end{aligned} \right], \quad (62)$$

where:

$$\theta_{h_1} = (p_{h_2} - p_{h_3})\pi_{x_h} + (1 - p_{h_1})\lambda_{h_1} - p_{h_2}.$$

The stratified estimator $\hat{\pi}_{JS_{1(st)}}$ is given by:

$$\hat{\pi}_{JS_{1(st)}} = \sum_{h=1}^L W_h \hat{\pi}_{JS_{h_1}}. \quad (63)$$

Theorem 4.1. The estimator $\hat{\pi}_{JS_{1(st)}}$ is unbiased for π with variance, as given by:

$$Var(\hat{\pi}_{JS_{1(st)}}) = \sum_{h=1}^L \frac{W_h^2}{n_h} V_{JS_1}, \quad (64)$$

where:

$$V_{JS_1} = \frac{\lambda_{h_1}(1 - \lambda_{h_1}) + \omega_h^2(p_{h_2} - p_{h_3})^2\pi_{x_h}(1 - \pi_{x_h})}{\{1 - \omega_h(1 - p_{h_1})\}^2} + \frac{\theta_{h_1}^2\{\lambda_{\omega_h}(1 - \lambda_{\omega_h}) + (1 - p_h)^2\pi_{y_h}(1 - \pi_{y_h})\}}{p_h^2\{1 - \omega_h(1 - p_{h_1})\}^4}. \quad (65)$$

Proof. Taking expectation on both sides of Eq. (63), we have:

$$E(\hat{\pi}_{JS_{1(st)}}) \approx \sum_{h=1}^L W_h \left[\begin{aligned} &\frac{\lambda_{h_1} - \omega\{p_{h_2} - (p_{h_2} - p_{h_3})\pi_{x_h}\}}{\{1 - \omega_h(1 - p_{h_1})\}} \\ &+ \frac{E(\hat{\lambda}_{h_1}) - \lambda_{h_1}}{\{1 - \omega_h(1 - p_{h_1})\}} \\ &+ \frac{\omega_h(p_{h_2} - p_{h_3})(E(\hat{\pi}_{x_h}) - \pi_{x_h})}{\{1 - \omega_h(1 - p_{h_1})\}} \\ &+ \frac{\theta_{h_1}(E(\hat{\omega}_h) - \omega_h)}{\{1 - \omega_h(1 - p_{h_1})\}^2} \end{aligned} \right]. \quad (66)$$

Using the fact that $E(\hat{\lambda}_{h_1}) = \lambda_{h_1}$, $E(\hat{\pi}_{x_h}) = \pi_{x_h}$, $E(\hat{\lambda}_{\omega_h}) = \lambda_{\omega_h}$, and $E(\hat{\pi}_{y_h}) = \pi_{y_h}$. Eq. (66) becomes:

$$E(\hat{\pi}_{JS_{1(st)}}) \approx \sum_{h=1}^L W_h \left[\frac{\lambda_{h_1} - \omega_h\{p_{h_2} - (p_{h_2} - p_{h_3})\pi_{x_h}\}}{\{1 - \omega_h(1 - p_{h_1})\}} \right] = \pi. \quad (67)$$

Thus, $\hat{\pi}_{JS_{1(st)}}$ is approximately an unbiased estimator of π . Now, applying variance on both sides of Eq. (63), we have:

$$V(\hat{\pi}_{JS_{1(st)}}) = \sum_{h=1}^L W_h^2 \left[\begin{aligned} &\frac{V(\hat{\lambda}_{h_1}) + \omega^2(p_{h_2} - p_{h_3})^2 V(\hat{\pi}_{x_h})}{\{1 - \omega_h(1 - p_{h_1})\}^2} \\ &+ \frac{\theta_{h_1}^2 V(\hat{\omega}_h)}{\{1 - \omega_h(1 - p_{h_1})\}^4} \end{aligned} \right]. \quad (68)$$

Using the fact that $Var(\hat{\lambda}_{h_1}) = \frac{\lambda_{h_1}(1 - \lambda_{h_1})}{n_h}$ and $Var(\hat{\pi}_{x_h}) = \frac{\pi_{x_h}(1 - \pi_{x_h})}{n_h}$ and substituting the values of $Var(\hat{\omega}_h)$ given in Eq. (59), we have Eq. (64) which completes the proof.

Technique-II optional RRT model under stratification
The probability of getting “Yes” response in stratum h of the proposed binary optional RRT model of

Technique-II is:

$$\lambda_{h_2} = (1 - \omega_h)\pi_h + \omega_h \{(1 + p_{h_2})(p_{h_1}\pi + p_{h_3}\pi_{x_h}) + p_{h_2}^2\}. \quad (69)$$

Solving Eq. (69) for π_h , we have an estimator of π_h as follows:

$$\hat{\pi}_{JS_{h_2}} = \frac{\hat{\lambda}_{h_2} - \hat{\omega}_h \{p_{h_2}^2 + p_{h_3}(1 + p_{h_2})\hat{\pi}_{x_h}\}}{[1 - \hat{\omega}_h \{1 - p_{h_1}(1 + p_{h_2})\}]}, \quad (70)$$

where $\hat{\omega}_h$ is obtained from Eq. (56) and $\hat{\lambda}_{h_2}$ is the proportion of “Yes” answers reported by n_h respondents in the sample, $\hat{\lambda}_{h_2} \sim B(n_h, \lambda_{h_2})$ and $\hat{\pi}_{x_h} \sim B(n_h, \pi_{x_h})$. After applying first-order Taylor’s expansion to Eq. (70), we have:

$$\hat{\pi}_{JS_{h_2}} \approx \left[\begin{aligned} & \frac{\hat{\lambda}_{h_2} - \hat{\omega}_h \{p_{h_2}^2 + p_{h_3}(1 + p_{h_2})\pi_{x_h}\}}{[1 - \hat{\omega}_h \{1 - p_{h_1}(1 + p_{h_2})\}]} \\ & + \frac{(\hat{\lambda}_{h_2} - \lambda_{h_2})}{[1 - \omega_h \{1 - p_{h_1}(1 + p_{h_2})\}]} \\ & - \frac{\omega_h p_{h_3}(1 + p_{h_2})(\hat{\pi}_{x_h} - \pi_{x_h})}{[1 - \omega_h \{1 - p_{h_1}(1 + p_{h_2})\}]} \\ & + \frac{\theta_{h_2}(\hat{\omega}_h - \omega_h)}{[1 - \omega_h \{1 - (1 + p_{h_2})p_{h_1}\}]^2} \end{aligned} \right], \quad (71)$$

where:

$$\theta_{h_2} = \lambda_{h_2} - \{(1 + p_{h_2})(p_{h_3}\pi_{x_h} + \lambda_{h_2}p_{h_1}) + p_{h_2}^2\}.$$

The stratified estimator $\hat{\pi}_{JS_{2(st)}}$ is obtained as follows:

$$\hat{\pi}_{JS_{2(st)}} = \sum_{h=1}^L W_h \hat{\pi}_{JS_{h_2}}. \quad (72)$$

Theorem 4.2. The estimator $\hat{\pi}_{JS_{2(st)}}$ is unbiased for π with variance as given by:

$$Var(\hat{\pi}_{JS_{2(st)}}) = \sum_{h=1}^L \frac{W_h^2}{n_h} V_{JS_2}, \quad (73)$$

where:

$$V_{JS_2} = \frac{\lambda_{h_2}(1 - \lambda_{h_2}) + \omega_h^2 \{p_{h_3}(1 + p_{h_2})\}^2 \pi_{x_h}(1 - \pi_{x_h})}{[1 - \omega_h \{1 - p_{h_1}(1 + p_{h_2})\}]^2} + \frac{\theta_{h_2}^2 \{\lambda_{\omega_h}(1 - \lambda_{\omega_h}) + (1 - p_h)^2 \pi_{y_h}(1 - \pi_{y_h})\}}{p_h^2 [1 - \omega_h \{1 - p_{h_1}(1 + p_{h_2})\}]^4}. \quad (74)$$

Proof. The proof is analogous to that of Theorem 4.1.

Technique-III optional RRT model under stratification
The probability of getting “Yes” response in stratum h of the proposed binary optional RRT model of Technique-III is:

$$\lambda_{h_3} = (1 - \omega_h)\pi_h + \omega \{p_{h_1}\pi_h + p_{h_2}(1 - \pi_h) + p_{h_3}\pi_{x_h}\}. \quad (75)$$

Solving Eq. (75) for π_h , we have an estimator for π_h as follows:

$$\hat{\pi}_{JS_{h_3}} = \frac{\hat{\lambda}_{h_3} - \hat{\omega}_h(p_{h_2} + p_{h_3}\hat{\pi}_{x_h})}{\{1 - \hat{\omega}_h(1 - p_{h_1} + p_{h_2})\}}, \quad (76)$$

where $\hat{\omega}_h$ is obtained from Eq. (56) and $\hat{\lambda}_{h_2}$ is the proportion of “Yes” answers reported by n_h respondents in the sample, $\hat{\lambda}_{h_3} \sim B(n_h, \lambda_{h_3})$ and $\hat{\pi}_{x_h} \sim B(n_h, \pi_{x_h})$. After applying first-order Taylor’s expansion to Eq. (76), we have:

$$\hat{\pi}_{JS_{h_3}} \approx \left[\begin{aligned} & \frac{\lambda_{h_3} - \omega(p_{h_2} + p_{h_3}\pi_{x_h})}{\{1 - \omega_h(1 - p_{h_1} + p_{h_2})\}} \\ & + \frac{\hat{\lambda}_{h_3} - \lambda_{h_3}}{\{1 - \omega_h(1 - p_{h_1} + p_{h_2})\}} \\ & - \frac{\omega_h p_{h_3}(\hat{\pi}_{x_h} - \pi_{x_h})}{\{1 - \omega_h(1 - p_{h_1} + p_{h_2})\}} \\ & + \frac{\theta_{h_3}(\hat{\omega}_h - \omega_h)}{\{1 - \omega_h(1 - p_{h_1} + p_{h_2})\}^2} \end{aligned} \right], \quad (77)$$

where:

$$\theta_{h_3} = \lambda_{h_3}(1 - p_{h_1} + p_{h_2}) - \pi_{x_h}p_{h_3} - p_{h_2}.$$

The stratified estimator $\hat{\pi}_{JS_{3(st)}}$ is given by:

$$\hat{\pi}_{JS_{3(st)}} = \sum_{h=1}^L W_h \hat{\pi}_{JS_{h_3}}. \quad (78)$$

Thus, we have the following theorem and its proof is similar to that of Theorem 4.1.

Theorem 4.3. The estimator $\hat{\pi}_{JS_{3(st)}}$ is unbiased for π with variance, as given by:

$$Var(\hat{\pi}_{JS_{3(st)}}) = \sum_{h=1}^L \frac{W_h^2}{n_h} V_{JS_3}, \quad (79)$$

where:

$$V_{JS_3} = \frac{\lambda_{h_3}(1 - \lambda_{h_3}) + \omega_h^2 p_{h_3}^2 \pi_{x_h}(1 - \pi_{x_h})}{\{1 - \omega_h(1 - p_{h_1} + p_{h_2})\}^2} + \frac{\theta_{h_3}^2 \{\lambda_{\omega_h}(1 - \lambda_{\omega_h}) + (1 - p_h)^2 \pi_{y_h}(1 - \pi_{y_h})\}}{p_h^2 \{1 - \omega_h(1 - p_{h_1} + p_{h_2})\}^4}. \quad (80)$$

Corollary 4.1. In proportional allocation, the size of each stratum is $n_h = n(\frac{N_h}{N})$ and the variances of $\hat{\pi}_{JS_{i(st)}}$ and $\hat{\pi}_{M_{i(st)}}$ are respectively given below:

$$Var(\hat{\pi}_{JS_{i(st)}}) = \frac{1}{n} \sum_{h=1}^L W_h V_{JS_i}, \quad (81)$$

and:

$$Var\left(\hat{\pi}_{M_{i(s_t)}}\right) = \frac{1}{n} \sum_{h=1}^L W_h V_{M_i}, \quad (82)$$

for $i = 1, 2, 3 \dots$

4.5. Relative efficiency

The Relative Efficiency (RE) of the proposed optional RRT models with respect to the RRT models of Mahmood et al. [5] under stratification using proportional allocation method is defined as:

$$RE(i) = \frac{Var\left(\hat{\pi}_{M_{i(s_t)}}\right)}{Var\left(\hat{\pi}_{JS_{i(s_t)}}\right)}, \quad i = 1, 2, 3. \quad (83)$$

To calculate it empirically, we have chosen $n = 1000$ with two strata having stratum weights ($W_1 = 0.3$, $W_2 = 0.7$) and ($W_1 = 0.7$, $W_2 = 0.3$) for different values of π_1 and π_2 . The design parameters are defined as $p_1 = p_{11} = p_{12}$, $p_2 = p_{21} = p_{22}$, $p_3 = p_{31} = p_{32}$, $\omega = \omega_1 = \omega_2$, $\pi_y = \pi_{y_1} = \pi_{y_2}$, $\pi_x = \pi_{x_1} = \pi_{x_2}$, and $p = p_1 = p_2$, respectively. The results of each model with respect to their competitor models are discussed as follows:

- (i) Table 4 shows that $RE(1) > 1$ decreases when p_1 and ω values increase from 0.2 to 0.8 and from 0.1 to 0.9 by increment of 0.2, respectively.

Table 4. Relative efficiency of the proposed Technique-I optional RRT over Technique-I RRT when $n = 1000$, $\pi_x = 0.85$, $p = 0.7$ and at various choices of ω , p_1 , $p_3 = 2(1 - p_1)/3$, and $p_2 = (1 - p_1 - p_3)$.

p_1	π_1	π_2	W_1	W_2	$\pi_y = 0.35$					$\pi_y = 0.65$				
					ω					ω				
					0.1	0.3	0.5	0.7	0.9	0.1	0.3	0.5	0.7	0.9
0.2	0.08	0.13	0.3	0.7	18.71	9.15	4.80	2.40	0.98	17.17	8.86	4.80	2.46	1.03
			0.7	0.3	19.64	9.22	4.77	2.37	0.96	17.82	8.91	4.77	2.43	1.02
	0.28	0.33	0.3	0.7	14.02	8.75	5.18	2.74	1.12	13.68	8.65	5.18	2.76	1.15
			0.7	0.3	14.25	8.76	5.14	2.70	1.11	13.86	8.65	5.14	2.74	1.14
	0.48	0.53	0.3	0.7	13.16	8.95	5.57	2.99	1.21	13.12	8.93	5.57	3.00	1.22
			0.7	0.3	13.13	8.90	5.53	2.97	1.21	13.08	8.88	5.53	2.98	1.21
	0.68	0.73	0.3	0.7	15.07	9.90	5.96	3.10	1.22	15.03	9.89	5.96	3.11	1.22
			0.7	0.3	14.70	9.76	5.92	3.10	1.22	14.67	9.75	5.92	3.10	1.23
	0.88	0.93	0.3	0.7	25.34	12.51	6.37	3.01	1.12	24.12	12.29	6.37	3.05	1.15
			0.7	0.3	23.23	12.11	6.33	3.03	1.14	22.33	11.93	6.33	3.06	1.16
0.4	0.08	0.13	0.3	0.7	6.36	3.60	2.21	1.38	0.84	6.00	3.53	2.21	1.41	0.86
			0.7	0.3	6.81	3.69	2.22	1.37	0.83	6.36	3.60	2.22	1.40	0.86
	0.28	0.33	0.3	0.7	4.36	3.10	2.16	1.45	0.91	4.29	3.08	2.16	1.46	0.92
			0.7	0.3	4.45	3.13	2.16	1.44	0.90	4.38	3.10	2.16	1.45	0.91
	0.48	0.53	0.3	0.7	3.94	2.99	2.18	1.50	0.94	3.93	2.99	2.18	1.50	0.94
			0.7	0.3	3.95	2.99	2.17	1.49	0.94	3.94	2.99	2.17	1.49	0.94
	0.68	0.73	0.3	0.7	4.34	3.21	2.28	1.53	0.94	4.33	3.21	2.28	1.53	0.94
			0.7	0.3	4.25	3.17	2.27	1.53	0.94	4.25	3.17	2.27	1.53	0.94
	0.88	0.93	0.3	0.7	7.33	4.24	2.57	1.54	0.89	7.10	4.19	2.57	1.56	0.90
			0.7	0.3	6.65	4.05	2.52	1.54	0.89	6.48	4.01	2.52	1.56	0.91
0.6	0.08	0.13	0.3	0.7	3.43	2.28	1.62	1.18	0.87	3.32	2.26	1.62	1.19	0.89
			0.7	0.3	3.73	2.38	1.64	1.18	0.86	3.59	2.34	1.64	1.19	0.88
	0.28	0.33	0.3	0.7	2.29	1.84	1.47	1.16	0.91	2.27	1.83	1.47	1.17	0.91
			0.7	0.3	2.34	1.86	1.48	1.16	0.90	2.32	1.86	1.48	1.17	0.91
	0.48	0.53	0.3	0.7	2.07	1.74	1.43	1.16	0.92	2.07	1.74	1.43	1.16	0.92
			0.7	0.3	2.07	1.74	1.43	1.16	0.92	2.07	1.74	1.43	1.16	0.92
	0.68	0.73	0.3	0.7	2.20	1.82	1.47	1.17	0.91	2.20	1.82	1.47	1.18	0.91
			0.7	0.3	2.17	1.80	1.47	1.17	0.91	2.17	1.80	1.47	1.17	0.92
	0.88	0.93	0.3	0.7	3.52	2.41	1.70	1.22	0.88	3.46	2.39	1.70	1.23	0.89
			0.7	0.3	3.19	2.28	1.66	1.21	0.89	3.15	2.27	1.66	1.22	0.89
0.8	0.08	0.13	0.3	0.7	2.00	1.61	1.32	1.11	0.94	1.98	1.60	1.32	1.11	0.94
			0.7	0.3	2.17	1.69	1.36	1.12	0.94	2.14	1.68	1.36	1.12	0.94
	0.28	0.33	0.3	0.7	1.45	1.31	1.18	1.06	0.95	1.45	1.31	1.18	1.06	0.95
			0.7	0.3	1.47	1.32	1.18	1.06	0.95	1.47	1.32	1.18	1.06	0.95
	0.48	0.53	0.3	0.7	1.36	1.25	1.14	1.04	0.95	1.36	1.25	1.14	1.04	0.95
			0.7	0.3	1.36	1.25	1.14	1.04	0.95	1.36	1.25	1.14	1.04	0.95
	0.68	0.73	0.3	0.7	1.40	1.27	1.16	1.05	0.95	1.40	1.27	1.16	1.05	0.95
			0.7	0.3	1.39	1.27	1.15	1.05	0.95	1.39	1.27	1.15	1.05	0.95
	0.88	0.93	0.3	0.7	1.93	1.58	1.31	1.10	0.93	1.92	1.58	1.31	1.10	0.93
			0.7	0.3	1.79	1.50	1.28	1.09	0.93	1.78	1.50	1.28	1.09	0.93

Table 5. Relative efficiency of the proposed Technique-II optional RRT over Technique-II RRT when $n = 1000$, $\pi_x = 0.85$, $p = 0.7$ and at various choices of ω , p_1 , $p_3 = 2(1 - p_1)/3$, and $p_2 = (1 - p_1 - p_3)$.

p_1	π_1	π_2	W_1	W_2	$\pi_y = 0.35$					$\pi_y = 0.65$				
					ω					ω				
					0.1	0.3	0.5	0.7	0.9	0.1	0.3	0.5	0.7	0.9
0.2	0.08	0.13	0.3	0.7	30.92	15.04	8.39	4.74	2.45	28.05	14.54	8.39	4.86	2.57
			0.7	0.3	32.29	15.16	8.36	4.72	2.45	29.03	14.63	8.36	4.84	2.58
	0.28	0.33	0.3	0.7	23.47	14.35	8.66	4.89	2.35	22.45	14.07	8.66	4.99	2.46
			0.7	0.3	23.89	14.38	8.63	4.88	2.36	22.76	14.08	8.63	4.98	2.47
	0.48	0.53	0.3	0.7	21.72	14.60	9.07	4.98	2.20	21.31	14.45	9.07	5.06	2.30
			0.7	0.3	21.72	14.53	9.02	4.98	2.22	21.26	14.36	9.02	5.06	2.32
	0.68	0.73	0.3	0.7	24.48	16.21	9.68	4.99	2.00	24.35	16.15	9.68	5.04	2.08
			0.7	0.3	23.89	15.95	9.61	4.99	2.02	23.74	15.88	9.61	5.05	2.11
	0.88	0.93	0.3	0.7	43.83	21.59	10.70	4.84	1.73	43.81	21.59	10.70	4.87	1.80
			0.7	0.3	39.34	20.67	10.57	4.86	1.76	39.33	20.66	10.57	4.89	1.83
0.4	0.08	0.13	0.3	0.7	10.40	5.86	3.78	2.59	1.80	9.78	5.74	3.78	2.62	1.85
			0.7	0.3	11.12	6.00	3.81	2.58	1.79	10.38	5.86	3.81	2.62	1.84
	0.28	0.33	0.3	0.7	6.98	5.01	3.63	2.61	1.84	6.82	4.96	3.63	2.64	1.87
			0.7	0.3	7.16	5.07	3.64	2.61	1.84	6.97	5.01	3.64	2.64	1.87
	0.48	0.53	0.3	0.7	6.05	4.71	3.56	2.60	1.80	6.00	4.69	3.56	2.61	1.82
			0.7	0.3	6.09	4.72	3.57	2.60	1.81	6.03	4.70	3.57	2.62	1.83
	0.68	0.73	0.3	0.7	6.30	4.84	3.56	2.50	1.65	6.29	4.83	3.56	2.51	1.67
			0.7	0.3	6.21	4.80	3.56	2.51	1.67	6.19	4.79	3.56	2.52	1.69
	0.88	0.93	0.3	0.7	9.93	5.89	3.63	2.23	1.33	9.93	5.89	3.63	2.23	1.34
			0.7	0.3	9.03	5.69	3.62	2.27	1.37	9.03	5.69	3.62	2.27	1.38
0.6	0.08	0.13	0.3	0.7	5.08	3.34	2.41	1.84	1.45	9.78	5.74	3.78	2.62	1.85
			0.7	0.3	5.54	3.47	2.45	1.84	1.44	10.38	5.86	3.81	2.62	1.84
	0.28	0.33	0.3	0.7	3.26	2.67	2.20	1.82	1.51	6.82	4.96	3.63	2.64	1.87
			0.7	0.3	3.35	2.71	2.22	1.83	1.51	6.97	5.01	3.64	2.64	1.87
	0.48	0.53	0.3	0.7	2.82	2.44	2.10	1.79	1.50	6.00	4.69	3.56	2.61	1.82
			0.7	0.3	2.84	2.46	2.11	1.79	1.51	6.03	4.70	3.57	2.62	1.83
	0.68	0.73	0.3	0.7	2.82	2.42	2.05	1.71	1.41	6.29	4.83	3.56	2.51	1.67
			0.7	0.3	2.80	2.42	2.06	1.72	1.42	6.19	4.79	3.56	2.52	1.69
	0.88	0.93	0.3	0.7	3.90	2.75	2.00	1.47	1.09	9.93	5.89	3.63	2.23	1.34
			0.7	0.3	3.62	2.68	2.01	1.52	1.14	9.03	5.69	3.62	2.27	1.38
0.8	0.08	0.13	0.3	0.7	2.52	1.98	1.62	1.37	1.17	2.50	1.98	1.62	1.37	1.18
			0.7	0.3	2.74	2.07	1.66	1.37	1.16	2.72	2.07	1.66	1.37	1.17
	0.28	0.33	0.3	0.7	1.76	1.60	1.47	1.34	1.23	1.76	1.60	1.47	1.34	1.23
			0.7	0.3	1.79	1.62	1.47	1.35	1.23	1.79	1.62	1.47	1.35	1.23
	0.48	0.53	0.3	0.7	1.59	1.50	1.41	1.32	1.24	1.59	1.50	1.41	1.32	1.24
			0.7	0.3	1.60	1.50	1.41	1.32	1.24	1.60	1.50	1.41	1.32	1.24
	0.68	0.73	0.3	0.7	1.58	1.48	1.38	1.28	1.19	1.58	1.48	1.38	1.28	1.19
			0.7	0.3	1.57	1.48	1.38	1.29	1.20	1.57	1.48	1.38	1.29	1.20
	0.88	0.93	0.3	0.7	1.88	1.57	1.33	1.13	0.97	1.88	1.57	1.33	1.13	0.97
			0.7	0.3	1.80	1.55	1.34	1.17	1.02	1.80	1.55	1.34	1.17	1.02

So, we conclude that the efficiency of Technique-I optional RRT model may either increase or decrease in the case of different combinations of design parameters;

- (ii) According to Table 5, when the value of p_1 increases from 0.2 to 0.8, $RE(2) > 1$ decreases. Similarly, an attenuating pattern is observed at

fixed values of p_1 , π_1 , and π_2 when ω values increase from 0.1 to 0.9. Overall, result shows that Technique-II optional RRT model is more efficient than the competitor Technique-II of Mahmood et al. [5];

- (iii) In Table 6, we observe that the proposed Technique-III optional RRT model is more effi-

Table 6. Relative efficiency of the proposed Technique-III optional RRT over Technique-III RRT when $n = 1000$, $\pi_x = 0.85$, $p = 0.7$ and at various choices of ω , p_1 , $p_3 = 2(1 - p_1)/3$, and $p_2 = (1 - p_1 - p_3)$.

p_1	π_1	π_2	W_1	W_2	$\pi_y = 0.35$					$\pi_y = 0.65$				
					ω					ω				
					0.1	0.3	0.5	0.7	0.9	0.1	0.3	0.5	0.7	0.9
0.2	0.08	0.13	0.3	0.7	284.72	104.30	34.06	6.01	0.01	247.11	98.45	34.06	6.47	0.02
			0.7	0.3	288.25	102.15	32.86	5.73	0.01	247.32	96.15	32.86	6.18	0.01
	0.28	0.33	0.3	0.7	271.81	129.41	49.25	9.91	0.02	256.50	125.51	49.25	10.48	0.03
			0.7	0.3	271.49	126.56	47.45	9.40	0.02	254.58	122.44	47.45	9.97	0.03
	0.48	0.53	0.3	0.7	295.26	163.00	70.69	17.01	0.06	291.40	161.60	70.69	17.48	0.07
			0.7	0.3	291.03	159.16	68.33	16.13	0.05	286.35	157.52	68.33	16.63	0.07
	0.68	0.73	0.3	0.7	376.31	208.71	93.33	25.64	0.28	375.92	208.64	93.33	25.67	0.32
			0.7	0.3	363.71	203.46	91.41	24.99	0.23	363.55	203.43	91.41	25.05	0.27
	0.88	0.93	0.3	0.7	663.16	271.75	101.25	24.76	0.35	626.90	266.05	101.25	25.25	0.40
			0.7	0.3	609.25	264.49	101.56	25.43	0.42	583.22	259.98	101.56	25.83	0.46
0.4	0.08	0.13	0.3	0.7	44.21	20.85	10.11	4.44	1.45	39.76	19.96	10.11	4.66	1.66
			0.7	0.3	46.30	20.99	10.00	4.34	1.40	41.14	20.03	10.00	4.57	1.62
	0.28	0.33	0.3	0.7	33.58	20.28	11.35	5.48	1.93	32.41	19.90	11.35	5.63	2.10
			0.7	0.3	34.11	20.26	11.21	5.37	1.88	32.77	19.85	11.21	5.53	2.05
	0.48	0.53	0.3	0.7	32.04	21.38	12.92	6.66	2.51	31.81	21.28	12.92	6.72	2.59
			0.7	0.3	31.91	21.18	12.75	6.54	2.45	31.62	21.06	12.75	6.61	2.55
	0.68	0.73	0.3	0.7	37.91	25.07	15.07	7.78	3.00	37.88	25.06	15.07	7.78	3.01
			0.7	0.3	36.80	24.51	14.82	7.68	2.97	36.79	24.51	14.82	7.69	2.98
	0.88	0.93	0.3	0.7	70.42	36.03	18.32	8.46	3.04	67.90	35.52	18.32	8.56	3.11
			0.7	0.3	63.30	34.20	17.92	8.42	3.06	61.60	33.82	17.92	8.50	3.12
0.6	0.08	0.13	0.3	0.7	9.63	5.78	3.67	2.37	1.52	9.04	5.63	3.67	2.43	1.61
			0.7	0.3	10.43	6.00	3.73	2.38	1.52	9.69	5.83	3.73	2.45	1.61
	0.28	0.33	0.3	0.7	6.26	4.61	3.33	2.33	1.57	6.16	4.57	3.33	2.36	1.61
			0.7	0.3	6.42	4.67	3.35	2.33	1.56	6.29	4.63	3.35	2.36	1.61
	0.48	0.53	0.3	0.7	5.59	4.36	3.29	2.38	1.62	5.57	4.36	3.29	2.39	1.63
			0.7	0.3	5.60	4.36	3.29	2.37	1.62	5.58	4.35	3.29	2.38	1.63
	0.68	0.73	0.3	0.7	6.20	4.79	3.57	2.54	1.71	6.20	4.78	3.57	2.54	1.71
			0.7	0.3	6.06	4.70	3.52	2.52	1.70	6.06	4.70	3.52	2.52	1.70
	0.88	0.93	0.3	0.7	11.29	7.18	4.65	2.98	1.85	11.09	7.13	4.65	3.00	1.87
			0.7	0.3	10.04	6.68	4.45	2.91	1.83	9.91	6.65	4.45	2.93	1.85
0.8	0.08	0.13	0.3	0.7	3.60	2.78	2.20	1.77	1.45	3.53	2.75	2.20	1.79	1.47
			0.7	0.3	3.98	2.97	2.30	1.83	1.48	3.88	2.93	2.30	1.84	1.50
	0.28	0.33	0.3	0.7	2.32	2.02	1.75	1.52	1.30	2.31	2.02	1.75	1.52	1.31
			0.7	0.3	2.37	2.05	1.77	1.53	1.31	2.36	2.05	1.77	1.53	1.32
	0.48	0.53	0.3	0.7	2.09	1.86	1.65	1.45	1.26	2.08	1.86	1.65	1.45	1.27
			0.7	0.3	2.09	1.87	1.65	1.45	1.27	2.09	1.87	1.65	1.45	1.27
	0.68	0.73	0.3	0.7	2.19	1.94	1.70	1.48	1.28	2.19	1.94	1.70	1.48	1.28
			0.7	0.3	2.16	1.92	1.69	1.47	1.28	2.16	1.92	1.69	1.47	1.28
	0.88	0.93	0.3	0.7	3.43	2.72	2.18	1.76	1.42	3.41	2.71	2.18	1.76	1.43
			0.7	0.3	3.10	2.53	2.07	1.70	1.39	3.08	2.52	2.07	1.70	1.40

cient than the competitor Technique-III of Mahmood et al. [5]. The value of $RE(3)$ is higher than 1 for all combinations of design parameters.

5. Concluding remarks

According to the numerical comparison, the proposed optional Randomized Response Technique (RRT) mod-

els estimated by using the multi-question approach were more efficient than the corresponding Techniques of Mahmood et al. [5]. It was demonstrated that for a higher sensitive question, the proposed optional RRT model and the competitor models provided equal protection for the respondents. Moreover, the proposed optional RRT models simultaneously estimated the prevalence of sensitive attribute and sensitivity

level, compared to competitor RRT models. We, therefore, preferred the proposed optional RRT models over competitor models. An extensive simulation study showed that the empirical mean and variance of $\hat{\pi}_{JS_i}$ ($i = 1, 2, 3$) and $\hat{\omega}$ were in good agreement with the corresponding theoretical values. It was also shown that $\hat{\pi}_{JS_i}$ in first-order approximation worked very well. The superiority of the suggested optional RRT under stratification was revealed through a numerical comparison and it was observed that the proposed optional RRT models under stratified random sampling outperformed stratified techniques of Mahmood et al. [5]. Moreover, we also observed that design parameters played an important role in increasing or decreasing the efficiency of the suggested optional RRT models.

Acknowledgments

The authors are thankful to the editor and anonymous learned referees for their valuable suggestions which improved the quality of this work.

References

1. Gupta, S.N., Gupta, B.C., and Singh, S. "Estimation of sensitivity level of personnel interview survey questions", *Journal of Statistical Planning and Inference*, **100**, pp. 239–247 (2002).
2. Warner, S.L. "Randomized response: A survey technique for eliminating evasive answer bias", *Journal of the American Statistical Association*, **60**(309), pp. 63–69 (1965).
3. Greenberg, B., Abul-El, A., Simmons, W., and Horvitz, D. "The unrelated question randomized response model: theoretical framework", *Journal of the American Statistical Association*, **64**(326), pp. 520–539 (1969).
4. Chaudhuri, A. and Mukerjee, R., *Randomized Response: Theory and Techniques*, Marcel Dekker, Inc (1998).
5. Mahmood, M., Singh, S., and Horn, S. "On the confidentiality guaranteed under randomized response sampling: a comparisons with several techniques", *Biometrical Journal*, **40**, pp. 237–242 (1998).
6. Perri, P.F. "Modified randomized devices for Simmons' model", *Model Assisted Statistics and Applications*, **3**(3), pp. 233–239 (2008).
7. Hussain, Z. and Shabbir, J. "Estimation of the mean of a socially undesirable characteristic", *Scientia Iranica*, **20**(3), pp. 839–845 (2013).
8. Lee, G.S., Hong, K-H., Kim, J-M., and Son, C-K. "An estimation of a sensitive attribute based on a two stage stratified randomized response model with stratified unequal probability sampling", *Brazilian Journal of Probability and Statistics*, **28**(3), pp. 381–408 (2014).
9. Abdelfatah, S. and Mazloun, R. "Improved randomized response model using three decks of cards", *Model Assisted Statistical Application*, **9**, pp. 63–72 (2014).
10. Tanveer, T.A. and Singh, H.P. "Some improved additive randomized response models utilizing higher order moments ratios of scrambling variable", *Models Assisted Statistics and Applications*, **10**(4), pp. 361–383 (2015).
11. Singh, H.P. and Tanveer, T.A. "An improvement over Kim and Elam stratified unrelated question randomized response model using Neyman allocation", *Sankhya: The Indian Journal of Statistics*, **77**(B), pp. 91–107 (2015).
12. Blair, G., Imai, K., and Zhou, Y-Y. "Design and analysis of the randomized response technique", *Journal of the American Statistical Association*, **110**(511), pp. 1304–1319 (2015).
13. Singh, H.P. and Gorey, S.M. "An efficient new randomized response model", *Communications in Statistics-Theory and Methods*, **46**(19), pp. 9629–9635 (2016).
14. Bose, M. "Respondent privacy and estimation efficiency in randomized response surveys for discrete-valued sensitive variables", *Statistical Papers*, **56**(4), pp. 1055–1069 (2015).
15. Abid, M., Naeem, A., Hussain, Z., Riaz, M., and Tahir, M. "Investigating the impact of simple and mixture priors on estimating sensitive proportion through a general class of randomized response models", *Scientia Iranica*, **26**(2), pp. 1009–1022 (2019).
16. Mangat, N.S. and Singh, R. "An alternative randomized response procedure", *Biometrika*, **77**(2), pp. 439–442 (1990).
17. Gupta, S. and Shabbir, J. "Sensitivity estimation for personal interview survey questions", *Statistica*, **64**(4), pp. 643–653 (2004).
18. Gupta, S., Shabbir, J., and Sehra, S. "Mean and sensitivity estimation in optional randomized response models", *Journal of Statistical Planning and Inference*, **140**(10), pp. 2870–2874 (2010).
19. Gupta, S., Thornton, B., Shabbir, J., and Singhal, S. "A comparison of multiplicative and additive optional RRT models", *Journal of Statistical Theory and Applications*, **5**, pp. 226–239 (2006).
20. Gupta, S.N., Tuck, A., Gill, T.S., and Crowe, M. "Optional unrelated question randomized response models", *Involve: A Journal of Mathematics*, **6**(4), pp. 483–492 (2013).
21. Chhabra, A., Dass, B., and Mehta, S. "Multistage optional unrelated question RRT model", *Journal of Statistical Theory and Applications*, **15**(1), pp. 80–95 (2016).
22. Gupta, S., Shabbir, J., and Kalucha, G. "A two-step approach to ratio and regression estimation of finite population mean using optional randomized response models", *Hacettepe Journal of Mathematics and Statistics*, **45**(06), pp. 1819–1830 (2016).

23. Sihm, J.S., Chhabra, A., and Gupta, S.N. “An optional unrelated question RRT model”, *Involve: A Journal of Mathematics*, **9**(2), pp. 195–209 (2016).
24. Narjis, G. and Shabbir, J. “Estimation of population proportion and sensitivity level using optional unrelated question randomized response techniques”, *Communications in Statistics-Simulation and Computation*, **49**(12), pp. 3212–3226 (2020). DOI: 10.1080/03610918.2018.1538453
25. Lanke, J. “On the degree of protection in randomized interviews”, *International Statistical Review*, **44**, pp. 197–203 (1976).

Biographies

Ghulam Narjis completed her PhD degree in Statistics from Quaid-i-Azam University Islamabad, Pakistan in November 2019. She has published five article

papers in internationally reputed journals in the field of survey methodologies, randomized response, and data analysis. Her areas of interest include sampling theory and Bayesian analysis. The current article is a part of her PhD research.

Javid Shabbir is working as a Tenured Professor of Statistics at Department of Statistics Quaid-I-Azam University Islamabad, Pakistan. He has completed his PhD in Statistics from Kent University at Canterbury, UK in 1997. He availed Post-Doctoral positions at University of Southern Maine 2003, USA in 2003 and University of North Carolina at Greensboro USA in 2005. He has published about 300 articles in different internationally reputed journals. His area of research includes survey sampling and randomized response techniques. He has supervised many MPhil and PhD students at the department.