STATISTICS IN TRANSITION new series, Winter 2014
Vol. 15, No. 1, pp. 67–82

A MODIFIED MIXED RANDOMIZED RESPONSE MODEL

Housila P. Singh, Tanveer A. Tarray[1]

ABSTRACT

Socio-economic investigations often relate to certain personal features that people wish to hide from others in comprehensive inquiries, detailed questionnaires include numerous items. Data on most of them are frequently easy to procure merely by asking, but a few others can be on sensitive issues for which people are not inclined to state honest responses. For example, most people prefer to conceal the truth regarding their savings, the extent of their accumulated wealth, their history of intentional tax evasion and other illegal and or unethical practices leading to earnings from clandestine sources, crimes, trade in contraband goods, susceptibility to intoxication, expenditures on addictions of various forms, homosexuality, and similar issues which are customarily disapproved of by society. Open or direct queries often fail to yield reliable data on such confidential aspects of human life. Warner (1965) developed an alternative survey technique that is known as randomized response (RR) technique. Greenberg et al. (1971) presented a revised version of Warner's (1965) technique for qualitative variables. Later various modifications were given by several researchers [see Chaudhuri (2011)]. Kim and Warde (2005) and Nazuk and Shabir (2010) presented mixed randomized response models using simple random sampling with replacement sampling scheme which improves the privacy of respondents. In this paper we have suggested a modified mixed randomized response model to estimate the proportion of a qualitative sensitive variable. Properties of the proposed randomized response model have been studied along with recommendations. It has been shown that the suggested randomized response model is always better than Kim and Warde's (2005) model while it is better than Nazuk and Shabbir's (2010) model under some realistic conditions. Numerical illustrations and graphs are also given in support of the present study.

Key words: randomized response technique, simple random sampling, dichotomous population, estimation of proportion, privacy of respondents, sensitive characteristics.

[1] School of Studies in Statistics, Vikram University, Ujjain – 456010 – India.
E-mail: tanveerstat@gmail.com.

1. Introduction

In situations where potentially embarrassing or incriminating responses are sought, the randomized response (RR) technique is effective in reducing non-sampling errors in sample surveys.

Refusal to respond and lying in surveys are two main sources of such non-sampling errors, as the stigma attached to certain practices (e.g. sexual behaviours and the use of illegal drugs) often leads to discrimination. Warner (1965) was first to introduce a randomized response (RR) model to estimate the proportion for sensitive attributes including homosexuality, drug addiction or abortion. Greenberg et al. (1969) proposed the unrelated question RR model that is a variation of Warner's (1965) RR model. Since the work by Warner (1965) a huge literature has emerged on the use and formulation of different randomization devices to estimate the population proportion of a sensitive attribute in survey sampling. Mention may be made of the work of Tracy and Mangat (1996), Cochran (1977), Singh and Mangat (1996), Chaudhuri and Mukherjee (1988), Ryu et al. (1993), Fox and Tracy (1986), Singh (2003), Singh and Tarray (2012, 2013 a,b,c,d) and the references cited therein.

Mangat et al. (1997) and Singh et al. (2000) pointed out the privacy problem with Moors' (1971) model. To implement the privacy problem with the Moors' (1971) model, Mangat et al. (1997) and Singh et al. (2000) presented several strategies as an alternative to Moors' model, but their models can lose a large portion of data information and require a high cost to obtain confidentiality of the respondents. Kim and Warde (2005) suggested a mixed randomized response model using simple random sampling which rectifies the privacy problem. Amitava (2005) and Hussain and Shabbir (2007) suggested improvements over Kim and Warde's (2005) mixed randomized response technique in complex surveys situations and illustrated the superiority of their models over Kim and Warde's (2005) procedure. Later, Nazuk and Shabbir (2010) presented a modification of Kim and Warde's (2005) model to estimate the proportion of a qualitative sensitive variable using simple random sampling with replacement (SRSWR), which reduces the variance of the estimator and improves the privacy protection of respondents.

In this paper we have suggested a modified mixed randomized response model and its properties are studied. We have shown that the suggested mixed randomized response model is always better than Kim and Warde's (2005) model and it is more efficient than the one recently proposed by Nazuk and Shabbir's (2010) estimator under some realistic conditions.

2. Kim and Warde (2005) and Nazuk and Shabbir's (2010) models

2.1. Kim and Warde's (2005) mixed randomized response model

Kim and Warde (2005) introduced a mixed randomized response procedure for estimating the proportion π_S of a population possessing a sensitive attribute using simple random sampling with replacement (SRSWR) which rectifies the privacy problem. Following them, a single sample with the size n is selected by SRSWR from the population. Each respondent selected in the sample is instructed to answer the direct question "I am a member of the innocuous trait group". If a respondent answers "Yes" to the direction question, then she or he is instructed to go to the randomization device R_1 consisting of the statements (i) "I am a member of the sensitive trait group" and (ii) "I am a member of the innocuous trait group" with pre-assigned probability of selection P_1 and $1-P_1$, respectively. If a respondent answers "No" to the direct question, then the respondent is instructed to use the randomization device R_2 consisting of the statement (i) "I am a member of the sensitive trait group" and (ii) "I am not a member of the sensitive trait group" with pre-assigned probability P and 1-P, respectively. The survey procedures are performed under the assumption that both the sensitive and innocuous questions are unrelated and independent in the randomization device R_1. To protect the respondents' privacy, the respondents should not disclose to the interviewer the question they answered from either R_1 or R_2. Let n be the sample size confronted with a direct question and n_1 and n_2 (= n- n_1) denote the number of "Yes" and "No" answers from the sample. Since all respondents using the randomization device R_1 already responded "Yes" from the initial direct innocuous question, the proportion Y of getting "Yes" answers from the respondents using the randomization device R_1 should be

$$Y = P_1\pi_S + (1-P_1)\pi_1 = P_1\pi_S + (1-P_1), \tag{2.1}$$

where π_S is the proportion of "Yes" answers from the sensitive trait and π_1 is the proportion of "Yes" answers from the innocuous question [see Kim and Warde (2005,p.212)].

An unbiased estimator of π_S is given by

$$\hat{\pi}_a = \frac{\hat{Y} - (1-P_1)}{P_1} \tag{2.2}$$

where $\hat{Y}$ is the sample proportion of "Yes" responses.

The proportion of "Yes" answers from the respondents using the randomization device R_2 is given by

$$X = [P\pi_S + (1-P)(1-\pi_S)] = [(2P-1)\pi_S + (1-P)] \tag{2.3}$$

Thus, an unbiased estimator of π_S, in terms of the sample proportion of "Yes" responses $\hat{X}$, is

$$\hat{\pi}_b = \frac{\hat{X}-(1-P)}{(2P-1)} \tag{2.4}$$

Pooling the two unbiased estimators $\hat{\pi}_a$ and $\hat{\pi}_b$ using weights, Kim and Warde (2005) suggested an unbiased estimator $\hat{\pi}_a$ and $\hat{\pi}_b$ for π_S as

$$\hat{\pi}_{kw} = \frac{n_1}{n}\hat{\pi}_a + \frac{(n-n_1)}{n}\hat{\pi}_b \text{ , for } 0 < \frac{n_1}{n} < 1 \tag{2.5}$$

Applying Lanke's (1976) arguments, Kim and Warde (2005) derived

$$P = \frac{1}{2-P_1} \tag{2.6}$$

and hence obtained the variance of the estimator $\hat{\pi}_{kw}$ as

$$V(\hat{\pi}_{kw}) = \frac{\pi_S(1-\pi_S)}{n} + \frac{(1-P_1)\,[\,\lambda P_1(1-\pi_S)+(1-\lambda)]}{nP_1^2} \tag{2.7}$$

for $n = n_1 + n_2$ and $\lambda = \frac{n_1}{n}$.

2.2. Nazuk and Shabbir's (2010) model

Nazuk and Shabbir (2010) presented a modified version of Kim and Warde's (2005) model which differs from Kim and Warde's (2005) procedure only in the formation of the randomization device R_2. The description of Nazuk and Shabbir's (2010) model is given below.

Let a random sample of size n be selected using SRSWR. Each respondent in the sample is instructed to answer an innocuous question "I possess the innocuous character Y". If the answer to the initial direct question is "Yes" then the respondent is instructed to go the randomization device R_1, otherwise R_2, where R_1 consists of two statements (i) "I belong to the sensitive group" and (ii) "I belong to the innocuous group", with respective probability P_1 and $(1-P_1)$, while R_2 consists of the same pair of statements as in R_1 but with respective probability P_2 and $(1-P_2)$. In order to offer privacy to the respondents they are not required to say that which randomization device they have used. Let n_1 and n_2 be the number of respondents using R_1 and R_2 respectively such that $(n_1 + n_2) = n$. Note that the respondents coming to R_1 have reported "Yes" to the initial direct question, therefore $\pi_1 = 1$ in R_1 [see Nazuk and Shabbir (2010, pp.186-187)].

The probability of “Yes” answers is (the same as given in (2.1))

$$Y = [P_1\pi_S + (1-P_1)\pi_1] = P_1\pi_S + (1-P_1), \tag{2.8}$$

An unbiased estimator of π_S is (the same as given in (2.2))

$$\hat{\pi}_a = \frac{\hat{Y} - (1-P_1)}{P_1} \tag{2.9}$$

where $\hat{Y}$ is the same as defined earlier.

Note that the respondents using R_2 have reported a “No” to the initial direct question, therefore $\pi_1 = 0$ in R_2. Denote by X_2 the probability of “Yes” answers from the respondents using R_2 which is given by

$$X_2 = [P_2\pi_S + (1-P_2)\pi_1] = P_2\pi_S \tag{2.10}$$

Let $\hat{X}_2$ be the sample proportion of “Yes” response from the randomization device R_2, then an unbiased estimator of π_S is given by

$$\hat{\pi}_c = \frac{\hat{X}_2 - (1-P_2)}{P_2} \tag{2.11}$$

Pooling the two unbiased estimators $\hat{\pi}_a$ and $\hat{\pi}_c$, Nazuk and Shabbir (2010) suggested an unbiased estimator for π_S as

$$\hat{\pi}_{ns} = \frac{n_1}{n}\hat{\pi}_a + \frac{(n-n_1)}{n}\hat{\pi}_c \text{ , for } 0 < \frac{n_1}{n} < 1 \tag{2.12}$$

With $P_2 = \frac{1}{2-P_1}$ Nazuk and Shabbir (2010) obtained the variance of $\hat{\pi}_{ns}$ as

$$V(\hat{\pi}_{ns}) = \frac{\pi_S(1-\pi_S)}{n} + \frac{(1-P_1)[\lambda(1-\pi_S) + (1-\lambda)\pi_S P_1]}{nP_1} \tag{2.13}$$

3. The suggested model

The suggested procedure differs from Kim and Warde (2005) and Nazuk and Shabbir’s (2010) procedures only in the contribution of the randomization device R_2. Let a random sample of size n be selected using simple random sampling with replacement (SRSWR). Each respondent from the sample is instructed to answer the direct question “I am a member of the innocuous group”. If a respondent answers “Yes” to the direct question, then she or he is instructed to go to the randomization device R_1 consisting of the statements (i) “I am the member of the sensitive trait group” and (ii) “I am a member of the innocuous trait group” with respective probabilities P_1 and $(1-P_1)$. If a respondent answers “No” to the direct question, then the respondent is instructed to use the randomization device R_2 using three statements: (i) “I possess the sensitive attribute “A” ”, (ii) “Yes” and (iii) “No” with known probabilities P, (1-P)w and (1-P)w respectively, where $w \in [0,1]$. It is to be mentioned that the randomization device R_2 is due to Singh et al. (1995). The survey procedures are performed under the assumption that both the sensitive and innocuous questions are unrelated and independent in the randomization device R_1. To protect the respondent’s privacy, the respondents should not disclose to the interviewer the question they answered from either R_1 or R_2.

We explain the suggested procedure with the help of an example earlier considered by Hussain and Shabbir (2007). Consider that we are interested in the estimation of the proportion π_S of carriers of HIV in a particular county/ locality/ town or district. Each survey respondent is asked a direct innocuous (non-sensitive) question “Were you born in the first three months of a calendar year?”. On receiving a “Yes” response he/she is requested to use the randomization device R_1 consisting of the two statements, (i) “I do carry HIV” and (ii) “My birthday falls in the first three months of a calendar year” presented with predetermined probabilities P_1 and $(1-P_1)$. If the respondent says “No” to the direct question he/she is requested to use the randomization device R_2. Now, from this random device, if the statement (i) is chosen, the respondent will reply according to his actual status with respect to carriers of HIV. In the case the statement (ii) or (iii) is selected, one will report “Yes” or “No” as observed on the outcome of the random device R_2 presented with predetermined probabilities P, (1-P)w and (1-P)w respectively, where $w \in [0,1]$.

Let n be the sample size confronted with a direct question and n_1 and n_2 $(= n - n_1)$ denote the number of “Yes” and “No” answers from the sample. Note that the respondents coming to R_1 have reported a “Yes” to the initial direct question, therefore $\pi_1 = 1$ in R_1, where π_1 is the proportion of “Yes” answers from the innocuous question.

Denote by ‘Y’ the probability of “Yes” from the respondents using R_1. Then

$$Y = P_1\pi_S + (1-P_1)\pi_1 = P_1\pi_S + (1-P_1), \tag{3.1}$$

where π_S is the proportion of “Yes” answers from the sensitive trait.

An unbiased estimator of π_S, in terms of the sample proportion of “Yes” responses $\hat{Y}$, becomes

$$\hat{\pi}_a = \frac{\hat{Y} - (1 - P_1)}{P_1} . \tag{3.2}$$

The variance of $\hat{\pi}_a$ is

$$V(\hat{\pi}_a) = \frac{Y(1-Y)}{n_1 P_1^2} = \frac{(1-\pi_S)\,[\,P_1\pi_S + (1-P_1)]}{n_1 P_1}$$
$$= \frac{1}{n_1}\left[\pi_S(1-\pi_S) + \frac{(1-\pi_S)(1-P_1)}{P_1}\right]. \tag{3.3}$$

The proportion of "Yes" answers from the respondents using randomization device R_2 follows:

$$X_3 = P\pi_S + (1-P)w \tag{3.4}$$

An unbiased estimator of π_S, in terms of the sample proportion of "Yes" responses $\hat{X}_3$, becomes

$$\hat{\pi}_d = \frac{\hat{X}_3 - (1-P)w}{P} . \tag{3.5}$$

The variance of $\hat{\pi}_d$ is given by

$$V(\hat{\pi}_d) = \frac{X_3(1-X_3)}{n_2 P^2} = \left[\frac{\pi_S(1-\pi_S)}{n_2} + \frac{(1-\pi_S)(1-P)w}{n_2 P^2}\right] \tag{3.6}$$

The estimator of π_S, in terms of the sample proportions of "Yes" responses $\hat{Y}$ and $\hat{X}_3$, is

$$\hat{\pi}_t = \frac{n_1}{n}\hat{\pi}_a + \frac{n_2}{n}\hat{\pi}_d$$
$$= \frac{n_1}{n}\hat{\pi}_a + \frac{(n-n_1)}{n}\hat{\pi}_d \text{ , for } 0 < \frac{n_1}{n} < 1 \tag{3.7}$$

As both $\hat{\pi}_a$ and $\hat{\pi}_d$ are unbiased estimators, the expected value of $\hat{\pi}_t$ is

$$E(\hat{\pi}_t) = E\left[\frac{n_1}{n}\hat{\pi}_a + \frac{n_2}{n}\hat{\pi}_d\right]$$
$$= \frac{n_1}{n}\pi_S + \frac{(n-n_1)}{n}\pi_S = \pi_S$$

Thus, the proposed estimator $\hat{\pi}_t$ is an unbiased estimator π_S.

Now, the variance of $\hat{\pi}_t$ is given by

$$V(\hat{\pi}_t) = \left(\frac{n_1}{n}\right)^2 V(\hat{\pi}_a) + \left(\frac{n_2}{n}\right)^2 V(\hat{\pi}_d)$$

$$= \left(\frac{n_1}{n}\right)^2 \frac{1}{n_1}\left[\pi_S(1-\pi_S) + \frac{(1-\pi_S)(1-P_1)}{P_1}\right]$$

$$+ \left(\frac{n_2}{n}\right)^2 \frac{1}{n_2}\left[\pi_S(1-\pi_S) + \frac{(1-\pi_S)(1-P)w}{P^2}\right]$$

$$= \frac{n_1}{n^2}\left[\pi_S(1-\pi_S) + \frac{(1-\pi_S)(1-P_1)}{P_1}\right] + \frac{n_2}{n^2}\left[\pi_S(1-\pi_S) + \frac{(1-\pi_S)(1-P)w}{P^2}\right] \tag{3.8}$$

Since our mixed RR model also uses Horvitz's et al. (1967) method when $\pi_1 = 1$, we can apply Lanke's (1976) idea to our suggested model. Thus, using Lanke's (1976) result for P with $\pi_1 = 1$, we get

$$P = \frac{1}{2 - P_1}. \tag{3.9}$$

Putting $P = (2 - P_1)^{-1}$ in (3.6), we get

$$V(\hat{\pi}_d) = \frac{\pi_S(1-\pi_S)}{(n-n_1)} + \frac{(1-\pi_S)\,(1-P_1)w}{(n-n_1)}$$

$$= \frac{1}{(n-n_1)}\left[\pi_S(1-\pi_S) + (1-\pi_S)(1-P_1)w\right] \tag{3.10}$$

Thus, we have established the following theorem.

Theorem 3.1. The variance of $\hat{\pi}_t$ is given by

$$V(\hat{\pi}_t) = \frac{\pi_S(1-\pi_S)}{n} + \frac{(1-P_1)\,[\,\lambda(1-\pi_S) + (1-\lambda)P_1(1-\pi_S)w]}{nP_1} \tag{3.11}$$

for $n = n_1 + n_2$ and $\lambda = \frac{n_1}{n}$.

Remark 3.1. Following Chaudhuri (2001, 2004), Amitava (2005) and Hussain and Shabbir (2007), the present study can be extended for complex surveys.

4. Efficiency comparisons

In this section we have made a comparison of the suggested model under a completely truthful reporting case with Kim and Warde (2005) and Nazuk and Shabbir's (2010) models.

From (2.7) and (3.11) we have

$$V(\hat{\pi}_t) < V(\hat{\pi}_{kw}) \text{ if}$$

$$[\lambda(1-\pi_S)+(1-\lambda)P_1(1-\pi_S)w] < \frac{[\lambda P_1(1-\pi_S)+(1-\lambda)]}{P_1}$$

$$\text{i.e. if } P_1^2(1-\pi_S)w < 1$$

which is always true.

Thus, the proposed model is always better than Kim and Warde's (2005) model.

Further, from (2.13) and (3.11) we have

$$V(\hat{\pi}_{ns}) - V(\hat{\pi}_t) = \frac{(1-P_1)(1-\lambda)}{n}\{\pi_S(1+w)-w\}$$

which is positive if

$$\pi_S > \frac{w}{(1+w)}. \tag{4.1}$$

It follows from (4.1) that f for $\pi_S \geq 1/2$, the proposed randomized response model is always superior to Nazuk and Shabbir's (2010) model. Further, for $\pi_S = 2/5, 3/5, 1/5, 1/10$, the proposed model is better than Nazuk and Shabbir's (2010) model in the respective ranges of w:

$$w \in (0,\ 2/3),\ w \in (0,\ 3/7),\ w \in (0, 1/4) \text{ and } w \in (0, 1/9)$$

It is observed from the above that when the value of $\pi(< 1/2)$ decreases the ranges of w decrease.

To have a tangible idea about the performance of the proposed estimator $\hat{\pi}_t$ over Kim and Warde's (2005) estimator $\hat{\pi}_{kw}$ and Nazuk and Shabbir's (2010) estimator $\hat{\pi}_{ns}$, we have computed the percent relative efficiency of the proposed estimator $\hat{\pi}_t$ with respect to Kim and Warde's (2005) estimator $\hat{\pi}_{kw}$ and Nazuk and Shabbir's (2010) estimator $\hat{\pi}_{ns}$ by using the formulae:

$$\text{PRE}\ (\hat{\pi}_t, \hat{\pi}_{kw}) = \frac{V(\hat{\pi}_{kw})}{V(\hat{\pi}_t)} \times 100$$

$$= \frac{\left[\pi_S(1-\pi_S)P_1^2 + (1-P_1)\{\lambda P_1(1-\pi_S)+(1-\lambda)\}\right]}{P_1\left[\pi_S(1-\pi_S)P_1 + (1-P_1)\{\lambda(1-\pi_S)+(1-\lambda)P_1(1-\pi_S)w\}\right]} \times 100 \tag{4.2}$$

and

$$\text{PRE }(\hat{\pi}_t, \hat{\pi}_{ns}) = \frac{V(\hat{\pi}_{ns})}{V(\hat{\pi}_t)} \times 100$$

$$= \frac{\left[\pi_S(1-\pi_S)P_1 + (1-P_1)\{\lambda(1-\pi_S)+(1-\lambda)\pi_S P_1\}\right]}{\left[\pi_S(1-\pi_S)P_1 + (1-P_1)\{\lambda(1-\pi_S)+(1-\lambda)P_1(1-\pi_S)w\}\right]} \times 100 \tag{4.3}$$

for different values of , P_1,w, n and n_1.

We have obtained the values of the percent relative efficiencies $\text{PRE }(\hat{\pi}_t, \hat{\pi}_{kw})$ for λ = (0.7, 0.5, 0.3), n = 1000 and for different cases of of π_S , w, n_1 and P_1. Findings are shown in Table 1. Diagrammatic representation is also given in Fig. 1.

It is observed from Fig. 1 and Table 1that:

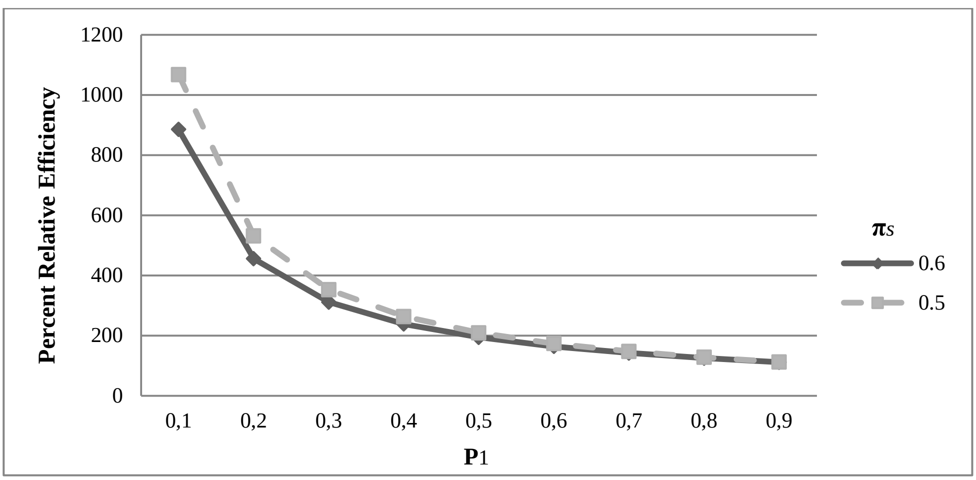

Figure 1. Percent relative efficiency of the proposed estimator $\hat{\pi}_t$ with respect to Kim and Warde's (2005) estimator $\hat{\pi}_{kw}$ when $\lambda = 0.7$ and $w = 0.25$

Table 1. Percent relative efficiency of the proposed estimator $\hat{\pi}_t$ with respect to Kim and Warde's (2005) estimator $\hat{\pi}_{kw}$

π_S	n = 1000		λ	w	P_1								
	n_1	n_2			0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.50	700.00	300.00	0.70	0.25	885.33	455.36	311.09	238.24	193.94	163.88	141.93	125.00	111.37
0.50	700.00	300.00	0.70	0.50	876.71	447.37	303.80	231.71	188.24	159.09	138.15	122.34	109.97
0.50	700.00	300.00	0.70	0.75	868.26	439.66	296.84	225.52	182.86	154.57	134.56	119.79	108.59
0.50	500.00	500.00	0.50	0.25	1858.19	865.38	538.40	377.36	282.35	220.13	176.45	144.23	119.53
0.50	500.00	500.00	0.50	0.50	1818.18	833.33	512.82	357.14	266.67	208.33	168.07	138.89	116.96
0.50	500.00	500.00	0.50	0.75	1779.86	803.57	489.56	338.98	252.63	197.74	160.44	133.93	114.49
0.50	300.00	700.00	0.30	0.25	3848.10	1614.13	914.09	587.68	405.63	292.93	218.14	165.98	128.20
0.50	300.00	700.00	0.30	0.50	3675.68	1500.00	836.60	534.48	369.23	268.52	202.53	156.98	124.25
0.50	300.00	700.00	0.30	0.75	3518.04	1400.94	771.22	490.12	338.82	247.86	189.00	148.90	120.54
0.60	700.00	300.00	0.70	0.25	1067.81	531.79	352.90	263.27	209.35	173.25	147.34	127.77	112.42
0.60	700.00	300.00	0.70	0.50	1057.57	522.73	344.98	256.47	203.64	168.64	143.83	125.39	111.20
0.60	700.00	300.00	0.70	0.75	1047.52	513.97	337.40	250.00	198.23	164.27	140.48	123.09	110.01
0.60	500.00	500.00	0.50	0.25	2256.12	1022.22	619.63	423.68	309.68	236.16	185.44	148.75	121.23
0.60	500.00	500.00	0.50	0.50	2208.45	985.71	591.70	402.50	293.88	224.73	177.62	143.95	119.00
0.60	500.00	500.00	0.50	0.75	2162.76	951.72	566.19	383.33	279.61	214.36	170.44	139.45	116.86
0.60	300.00	700.00	0.30	0.25	4650.76	1896.91	1048.23	659.09	445.57	315.45	230.45	172.10	130.51
0.60	300.00	700.00	0.30	0.50	4448.13	1769.23	965.12	604.17	409.30	291.96	215.94	164.01	127.09
0.60	300.00	700.00	0.30	0.75	4262.43	1657.66	894.22	557.69	378.49	271.73	203.14	156.65	123.85
0.70	700.00	300.00	0.70	0.25	1372.48	660.11	423.80	306.41	236.52	190.34	157.68	133.43	114.77
0.70	700.00	300.00	0.70	0.50	1359.50	649.17	414.67	298.91	230.51	185.69	154.28	131.22	113.68
0.70	700.00	300.00	0.70	0.75	1346.75	638.59	405.93	291.78	224.79	181.26	151.03	129.08	112.62
0.70	500.00	500.00	0.50	0.25	2921.41	1286.90	758.87	504.92	359.08	266.32	203.27	158.33	125.16
0.70	500.00	500.00	0.50	0.50	2860.83	1242.53	726.35	481.25	342.14	254.58	195.56	153.81	123.16
0.70	500.00	500.00	0.50	0.75	2802.71	1201.11	696.50	459.70	326.73	243.82	188.42	149.54	121.22
0.70	300.00	700.00	0.30	0.25	5998.59	2380.72	1283.95	788.84	521.07	360.06	256.18	185.70	136.04
0.70	300.00	700.00	0.30	0.50	5744.28	2227.83	1188.34	727.94	482.27	335.83	241.78	178.01	132.94
0.70	300.00	700.00	0.30	0.75	5510.65	2093.39	1105.98	675.77	448.84	314.65	228.90	170.93	129.98
0.90	700.00	300.00	0.70	0.25	3814.24	1694.15	1000.32	661.65	464.12	336.52	248.40	184.63	136.83
0.90	700.00	300.00	0.70	0.50	3779.14	1667.54	980.42	647.06	453.73	329.44	243.91	182.13	135.80
0.90	700.00	300.00	0.70	0.75	3744.68	1641.75	961.29	633.09	443.80	322.65	239.58	179.69	134.78
0.90	500.00	500.00	0.50	0.25	8261.22	3430.00	1901.23	1182.61	779.49	529.00	362.52	246.43	162.47
0.90	500.00	500.00	0.50	0.50	8096.00	3319.35	1827.01	1133.33	747.54	509.17	351.09	240.70	160.40
0.90	500.00	500.00	0.50	0.75	7937.25	3215.63	1758.38	1088.00	718.11	490.76	340.36	235.23	158.38
0.90	300.00	700.00	0.30	0.25	16862.28	6343.75	3253.67	1896.91	1180.58	758.78	491.58	313.12	189.05
0.90	300.00	700.00	0.30	0.50	16183.91	5970.59	3037.64	1769.23	1105.45	715.95	468.81	302.63	185.63
0.90	300.00	700.00	0.30	0.75	15558.01	5638.89	2848.51	1657.66	1039.32	677.69	448.06	292.82	182.34

The values of percent relative efficiencies $PRE(\hat{\pi}_t, \hat{\pi}_{kw})$ is more than 100. We can say that the envisaged estimator $\hat{\pi}_t$ is more efficient than Kim and Warde's (2005) estimator $\hat{\pi}_{kw}$. Fig. 1 shows the results for $\lambda = 0.7$, w = 0.25 and different values of P_1 and π_S.

We note from Table 1 that the values of the percent relative efficiencies $PRE(\hat{\pi}_t, \hat{\pi}_{kw})$ decrease as the value of P_1 increases. Also, the values of the percent relative efficiencies $PRE(\hat{\pi}_t, \hat{\pi}_{kw})$ increase as the value of λ decreases for fixed values of P_1.

We further note from the results of Fig. 1 that there is a large gain in efficiency by using the suggested estimator $\hat{\pi}_t$ over Kim and Warde's (2005) estimator $\hat{\pi}_{kw}$ when the proportion of the stigmatizing attribute is moderately large.

Fig. 2 and Table 2exhibit that:

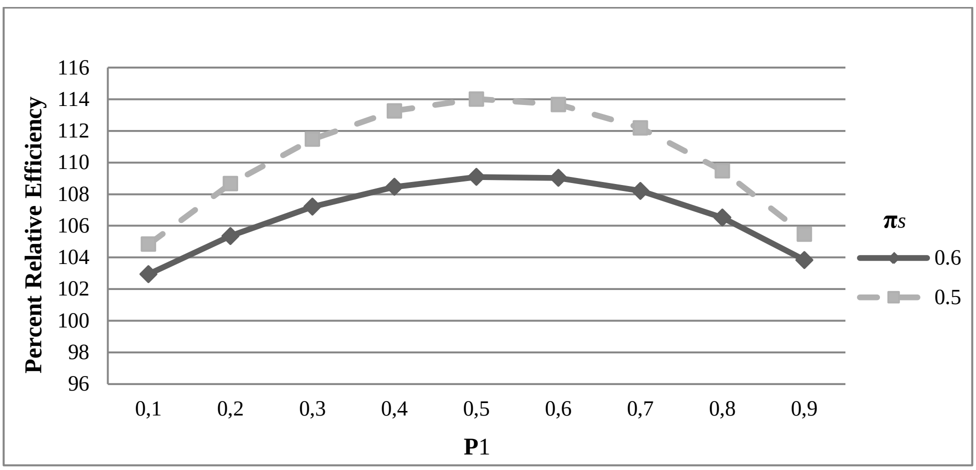

Figure 2. Percent relative efficiency of the proposed estimator $\hat{\pi}_t$ with respect to Nazuk and Shabbir's (2010) estimator $\hat{\pi}_{ns}$ when $\lambda = 0.7$ and $w = 0.25$

Table 2. Percent relative efficiency of the proposed estimator $\hat{\pi}_t$ with respect to Nazuk and Shabbir's (2010) estimator $\hat{\pi}_{ns}$

π_S	n = 1000		λ	w	P_1								
	n_1	n_2			0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.50	700.00	300.00	0.70	0.25	102.95	105.36	107.21	108.46	109.09	109.03	108.21	106.52	103.84
0.50	700.00	300.00	0.70	0.50	101.95	103.51	104.69	105.49	105.88	105.84	105.33	104.26	102.53
0.50	700.00	300.00	0.70	0.75	100.96	101.72	102.29	102.67	102.86	102.84	102.59	102.08	101.25
0.50	500.00	500.00	0.50	0.25	106.60	111.54	114.96	116.98	117.65	116.98	114.96	111.54	106.60
0.50	500.00	500.00	0.50	0.50	104.31	107.41	109.50	110.71	111.11	110.71	109.50	107.41	104.31
0.50	500.00	500.00	0.50	0.75	102.11	103.57	104.54	105.08	105.26	105.08	104.54	103.57	102.11
0.50	300.00	700.00	0.30	0.25	114.07	122.83	127.79	129.86	129.58	127.27	123.13	117.21	109.53
0.50	300.00	700.00	0.30	0.50	108.96	114.14	116.96	118.10	117.95	116.67	114.31	110.85	106.16
0.50	300.00	700.00	0.30	0.75	104.29	106.60	107.81	108.30	108.24	107.69	106.68	105.15	102.99
0.60	700.00	300.00	0.70	0.25	104.84	108.67	111.48	113.27	114.02	113.68	112.20	109.49	105.47
0.60	700.00	300.00	0.70	0.50	103.84	106.82	108.98	110.34	110.91	110.65	109.52	107.45	104.33
0.60	700.00	300.00	0.70	0.75	102.85	105.03	106.59	107.56	107.96	107.78	106.98	105.49	103.21
0.60	500.00	500.00	0.50	0.25	110.79	118.52	123.60	126.32	126.88	125.42	122.01	116.67	109.36
0.60	500.00	500.00	0.50	0.50	108.45	114.29	118.03	120.00	120.41	119.35	116.87	112.90	107.35
0.60	500.00	500.00	0.50	0.75	106.21	110.34	112.94	114.29	114.56	113.85	112.14	109.38	105.41
0.60	300.00	700.00	0.30	0.25	122.78	136.08	143.06	145.45	144.30	140.23	133.61	124.65	113.44
0.60	300.00	700.00	0.30	0.50	117.43	126.92	131.72	133.33	132.56	129.79	125.19	118.79	110.47
0.60	300.00	700.00	0.30	0.75	112.52	118.92	122.04	123.08	122.58	120.79	117.78	113.46	107.65
0.70	700.00	300.00	0.70	0.25	107.96	114.04	118.34	120.89	121.74	120.89	118.34	114.04	107.96
0.70	700.00	300.00	0.70	0.50	106.94	112.15	115.79	117.93	118.64	117.93	115.79	112.15	106.94
0.70	700.00	300.00	0.70	0.75	105.94	110.33	113.35	115.12	115.70	115.12	113.35	110.33	105.94
0.70	500.00	500.00	0.50	0.25	117.65	129.76	137.31	140.98	141.25	138.46	132.83	124.51	113.56
0.70	500.00	500.00	0.50	0.50	115.21	125.29	131.43	134.38	134.59	132.35	127.80	120.95	111.74
0.70	500.00	500.00	0.50	0.75	112.87	121.11	126.03	128.36	128.53	126.76	123.13	117.59	109.98
0.70	300.00	700.00	0.30	0.25	136.89	157.19	167.05	169.72	167.05	160.14	149.66	136.01	119.42
0.70	300.00	700.00	0.30	0.50	131.09	147.09	154.61	156.62	154.61	149.36	141.24	130.37	116.70
0.70	300.00	700.00	0.30	0.75	125.76	138.22	143.89	145.39	143.89	139.94	133.72	125.19	114.10
0.90	700.00	300.00	0.70	0.25	132.51	155.85	171.06	178.95	180.15	175.18	164.42	148.17	126.64
0.90	700.00	300.00	0.70	0.50	131.29	153.40	167.66	175.00	176.12	171.50	161.45	146.15	125.69
0.90	700.00	300.00	0.70	0.75	130.09	151.03	164.39	171.22	172.26	167.96	158.58	144.20	124.74
0.90	500.00	500.00	0.50	0.25	171.43	216.67	242.17	252.17	249.57	236.36	213.95	183.33	145.19
0.90	500.00	500.00	0.50	0.50	168.00	209.68	232.71	241.67	239.34	227.50	207.21	179.07	143.34
0.90	500.00	500.00	0.50	0.75	164.71	203.13	223.97	232.00	229.92	219.28	200.87	175.00	141.54
0.90	300.00	700.00	0.30	0.25	246.71	318.75	348.91	352.58	337.86	309.40	269.97	221.29	164.42
0.90	300.00	700.00	0.30	0.50	236.78	300.00	325.75	328.85	316.36	291.94	257.47	213.88	161.45
0.90	300.00	700.00	0.30	0.75	227.62	283.33	305.46	308.11	297.44	276.34	246.07	206.94	158.58

The values of percent relative efficiencies $PRE(\hat{\pi}_t, \hat{\pi}_{ns})$ is more than 100. We can say that the envisaged estimator $\hat{\pi}_t$ is more efficient than Nazuk and Shabbir's (2010) estimator$_c$. Fig. 2 shows the results for $\lambda = 0.7$, w = 0.25 and different values of P_1 and π_S.

We note from Table 2 that the values of the percent relative efficiencies $PRE(\hat{\pi}_t, \hat{\pi}_{ns})$ increase as the value of P_1 increases up to $P_1 \leq 0.5$ and decreases $P_1 > 0.5$ onwards. Also, the values of the percent relative efficiencies $PRE(\hat{\pi}_t, \hat{\pi}_{ns})$ increase as the value of λ decreases for fixed value of P_1.

We further note from the results of Fig. 2 that there is a large gain in efficiency by using the suggested estimator $\hat{\pi}_t$ over Nazuk and Shabbir's (2010) estimator $\hat{\pi}_{ns}$ when the proportion π_S of the stigmatizing attribute and P_1 are moderately large.

It is observed from Table 1 and Table 2 that a larger gain in efficiency is obtained by using the proposed estimator $\hat{\pi}_t$ over Kim and Warde's (2005) estimator $\hat{\pi}_{kw}$ as compared to Nazuk and Shabbir's (2010) estimator $\hat{\pi}_{ns}$.

5. Conclusions

In this paper we have proposed a mixed randomized response model to estimate the proportion of qualitative sensitive character. It has been shown that the proposed mixed randomized response model is more efficient than Kim and Warde (2005) and Nazuk and Shabbir's (2010) mixed randomize response models with a larger gain in efficiency. Thus, this paper attempts to extend the methodology of the mixed randomized response techniques.

Acknowledgements

The authors are grateful to the referee for fruitful comments to bring the original manuscript in the present form.

REFERENCES

AMITAVA, S., (2005). Kim and Warde's mixed randomized response technique for complex surveys. Jour Mod. Appl. Statist. Meth. , 4(2), 538–544.

CHAUDHURI, A., MUKERJEE, R., (1988): Randomized Response: Theory and Techniques. Marcel-Dekker, New York, USA.

CHAUDHURI, A., (2011). Randomized response and indirect questioning techniques in surveys. CRC Press, Taylor and Frances group, USA.

CHAUDHURI, A., (2004). Christofides' randomized response technique in complex sample surveys. Metrika, 60(3), 23–228.

CHAUDHURI, A., (2002). Estimating sensitive proportions from randomized responses in unequal probability sampling. Cal. Statist. Assoc. Bull., 52, 315–322.

COCHRAN, W. G., (1977). Sampling Technique, 3rd Edition. New York: John Wiley and Sons, USA.

FOX, J. A., TRACY, P. E., (1986). Randomized Response: A method of Sensitive Surveys. Newbury Park, CA: SEGE Publications.

GREENBERG, B., ABUL-ELA, A., SIMMONS, W. R., HORVITZ, D. G., (1969). The unreleased question randomized response: Theoretical framework. Jour. Amer. Statist. Assoc., 64, 529–539.

HORVITZ, D. G., SHAH, B. V., SIMMONS, W. R., (1967). The unrelated question randomized response model. Proc. Soc. Statist. Sec. Amer. Statistical Assoc. 65–72.

HUSSAIN, Z., SHABBIR, J., (2007). Improvement of Kim and Warde's mixed randomized response technique for complex surveys. InterStat, July # 003.

KIM, J. M., TEBBS, J. M., AN, S. W., (2006). Extensions of Mangat's randomized response model. Jour. Statist. Plan. Inference, 136, 1554–1567.

KIM, J. M., WARDE, W. D., (2005). A mixed randomized response model. Jour. Statist. Plan. Inference, 133, 211–221.

LANKE, J., (1976). On the degree of protection in randomized interview Internet. Statist. Rev. 44, 80–83.

MANGAT, N. S., SINGH, R., (1990). An alternative randomized procedure. Biometrika, 77, 439–442.

MANGAT, N. S., SINGH, R., SINGH, S., (1997). Violation of respondent's privacy in Moors model – its rectification through a random group strategy response model. Comm. Statist. Theo. Meth., (3), 243–255.

MOORS, J. A., (1971). Optimization of the unrelated question randomized response model. Jour. Amer. Statist. Assoc., 66, 627–629.

NAZUK, A., SHABBIR, J., (2010). A new mixed randomized response model. Inter. Jour Buss. Soc. Sci. 1, 186–190.

RYU, J. B., HONG, K. H., LEE, G. S., (1993). Randomized response model, Freedom Academy, Seoul, Korea.

SINGH, H. P., TARRAY, T. A., (2012). A Stratified Unknown repeated trials in randomized response sampling. Comm. Korean Statist. Soc., 19, (6), 751–759.

SINGH, H. P., TARRAY, T. A., (2013a). An alternative to Kim and Warde's mixed randomized response model. Statist. Oper. Res. Trans., 37 (2), 189–210.

SINGH, H. P., TARRAY, T. A., (2013b). An alternative to stratified Kim and Warde's randomized response model using optimal (Neyman) allocation, Model Assist. Statist. Appl., 9, 37–62.

SINGH, H. P., TARRAY, T. A., (2013c). An improved mixed randomized response model. Model Assist. Statist. Appl., 9, 73–87.

SINGH, H. P., TARRAY, T. A., (2013d). An alternative to Kim and Warde's mixed randomized response technique. Accepted in Statistica.

SINGH, R., MANGAT, N. S., (1996). Elements of Survey Sampling, Kluwer Academic Publishers, Dordrecht, The Netherlands.

SINGH, S., SINGH, R., MANGAT, N. S., TRACY, D. S., (1995). An improved two-stage randomized response strategy. Statistical Papers, 36, 265–271.

SINGH, S., (2003). Advanced sampling theory with applications. Kluwer Academic Publishers, Dordrecht.

SING, S., SINGH, R., MANGAT, N. S., (2000). Some alternative strategies to Moor's model in randomized response model. Jour. Statist. Plan. Inference, 83, 243–255.

TRACY, D. S., MANGAT, N. S., (1996). Some developments in randomized response sampling during the last decade – A follow up of review by Chaudhuri and Mukherjee. Jour. Applied. Statist. Sci., 4 (2/3), 147–158.

WARNER, S. L., (1965). Randomized response: A survey technique for eliminating evasive answer bias. Jour. Amer. Statist. Assoc., 60, 63–69.