

A STRATIFIED MANGAT AND SINGH'S OPTIONAL RANDOMIZED RESPONSE MODEL USING PROPORTIONAL AND OPTIMAL ALLOCATION

Housila P. Singh

School of Studies in Statistics, Vikram University, Ujjain – 456010 – India

Tanveer A. Tarray

School of Studies in Statistics, Vikram University, Ujjain – 456010 – India

1. INTRODUCTION

The randomized response (RR) data - gathering device to procure thrust worthy data on sensitive issues by protecting privacy of the respondents was first developed by Warner (1965). Subsequently, several other workers have proposed different RR strategies for instance, see the review oriented references like Hedayat and Sinha (1991), Chaudhuri and Mukerjee (1988), Tracy and Mangat (1996) and the papers by Mangat et al. (1997), Singh *et al.* (2000), Chang and Huang (2001), Chang *et al.* (2004 a b), Gupta *et al.* (2006), Nazuk and Shabbir (2010), Lee *et al.* (2011) and Singh and Tarray (2012).

Hong *et al.* (1994) suggested a stratified RR technique under the proportional sampling assumption. Kim and Warde (2004) and Kim and Elam (2005) have presented a stratified RR techniques using an optimal allocation which are more efficient than a stratified RR technique using a proportional allocation. Kim and Elam (2007) have mentioned that the extension of the randomized response technique to stratified random sampling may be useful if the investigator is interested in estimating the proportion of HIV / AIDS positively affected persons at different levels such as by rural areas or urban areas, age groups, or income groups.

In the stratified random sampling, the population to be used to conduct the survey is partitioned into strata. A sample is then selected by simple random sampling with replacement (SRSWR) from each stratum is known. To get the full benefit from stratification, it is assumed that the number of units in each stratum is known. In the stratified Warner's randomized response model, an individual respondent in the sample from stratum ' i ' is instructed to use the randomization device R_i which consists of a sensitive question (S) card with probability P_i and its negative question (\bar{S}) card with probability $(1 - P_i)$. The respondent answers the question with "Yes" or "No" without reporting which question card he or she has. A respondent belonging to the sample in different strata will perform different randomization device, each having different pre assigned probabilities. Under the assumption that these "Yes" or "No" reports are made truthfully and P_i is set by the researcher, the probability of "Yes" answers in stratum ' i ' for the stratified Warner's RR model is:

$$Z_i = P_i \pi_{S_i} + (1 - P_i)(1 - \pi_{S_i}), \text{ for } (i = 1, 2, \dots, k); \quad (1)$$

where Z_i is the proportion of "Yes" answers in a stratum i and π_{S_i} is the proportion of respondents with the sensitive trait in a stratum i . Let n_i denote the number of units in the sample from stratum i and n denote the total number of units in sample from all stratum so that

$n = \sum_{i=1}^k n_i$. The maximum likelihood estimate $\hat{\pi}_S$ (which is unbiased) of sensitive proportion $\pi_S = \sum_{i=1}^k w_i \pi_{S_i}$ is given by

$$\hat{\pi}_S = \sum_{i=1}^k w_i \hat{\pi}_{S_i} = \sum_{i=1}^k w_i \left[\frac{\hat{Z}_i - (1 - P_i)}{2P_i - 1} \right] \quad (2)$$

where $w_i = (N_i / N)$ for $(i = 1, 2, \dots, k)$ so that $w = \sum_{i=1}^k w_i = 1$, N is the number of units in

the whole population and N_i is the total number of units in the stratum i and \hat{Z}_i is a point estimate of Z_i .

The variance of $\hat{\pi}_S$ in (2) is given by

$$\begin{aligned} V(\hat{\pi}_S) &= \sum_{i=1}^k w_i^2 V(\hat{\pi}_{S_i}) \\ &= \sum_{i=1}^k \frac{w_i^2}{n_i} \left[\pi_{S_i}(1 - \pi_{S_i}) + \frac{P_i(1 - P_i)}{(2P_i - 1)^2} \right] = MSE(\hat{\pi}_S), \end{aligned} \quad (3)$$

where $MSE(\cdot)$ stands for the mean square error of (\cdot) .

Under proportional allocation (i.e. $n_i = n(N_i / N)$), the variance /MSE of $\hat{\pi}_S$ is given by

$$\begin{aligned} V(\hat{\pi}_S)_P &= \frac{1}{n} \sum_{i=1}^k w_i \left[\pi_{S_i}(1 - \pi_{S_i}) + \frac{P_i(1 - P_i)}{(2P_i - 1)^2} \right] \\ &= MSE(\hat{\pi}_S)_P \end{aligned} \quad (4)$$

which is due to Hong *et al.* (1994).

If the prior information on π_{S_i} is available from the past experience, then under optimum allocation:

$$\frac{n_i}{n} = \frac{\omega_i \left[\pi_{S_i}(1 - \pi_{S_i}) + \frac{P_i(1 - P_i)}{(2P_i - 1)^2} \right]^{1/2}}{\sum_{i=1}^k \omega_i \left[\pi_{S_i}(1 - \pi_{S_i}) + \frac{P_i(1 - P_i)}{(2P_i - 1)^2} \right]^{1/2}}, \quad (5)$$

Kim and Warde (2004) obtained the minimal variance/ MSE of the estimator $\hat{\pi}_S$ as

$$\begin{aligned} V(\hat{\pi}_S)_O &= \frac{1}{n} \left[\sum_{i=1}^k \omega_i \left\{ \pi_{S_i}(1 - \pi_{S_i}) + \frac{P_i(1 - P_i)}{(2P_i - 1)^2} \right\}^{1/2} \right]^2 \\ &= \text{MSE}(\hat{\pi}_S)_O \quad . \end{aligned} \quad (6)$$

In this paper we have suggested a stratified optional randomized response technique based on Mangat and Singh's (1994) optional randomized response model and studied its properties are studied under proportional allocation and optimum allocation. Numerically we have shown that the proposed stratified optional randomized response technique is better than Mangat and Singh (1994), Hong et al. (1994) and Kim and Warde (2004) estimators.

2. PROPOSED MODEL

In this proposed model, the population is partitioned into strata, and a sample is selected by simple random sampling with replacement (SRSWR) from each stratum. To get the full benefit from stratification, it is assumed that the number of units in each stratum is known. In this procedure, the randomized response device R_i and method for sampling the respondents in each stratum ' i ' remains same as in Kim and Warde (2004) model. However, it differs in the sense that the respondent is free to give answer in terms of "Yes" and "No" either by using RR device or without using it. It is not revealing to the interviewer which mode has been followed for giving answer.

Let n_i denote the number of units in the sample from stratum i and n denote the total number of units in samples from all stratum so that $n = \sum_{i=1}^k n_i$ if T_i is the probability that a respondent gives answer without using RR device then assuming completely truthful reporting, the probability Y_i of a "Yes" answer in stratum i is given by

$$\begin{aligned} Y_i &= T_i \pi_{S_i} + (1 - T_i) Z_i, \quad \text{for } i = 1, 2, \dots, k \\ &= T_i \pi_{S_i} + (1 - T_i) \{ P_i \pi_{S_i} + (1 - P_i)(1 - \pi_{S_i}) \}, \quad \text{for } i = 1, 2, \dots, k; \end{aligned} \quad (7)$$

where Z_i is given by (1).

For this procedure, we consider the following estimator of π_{S_i} :

$$\hat{\pi}_{mi} = \frac{n'_i}{n_i}, \quad (8)$$

where n'_i is the observed number of "Yes" answers obtained from the n_i respondents including in the sample from stratum i .

Since the selection in different strata are made independently, the estimators for individual strata can be added together to obtain an estimator for the whole population. Thus the estimator $\hat{\pi}_{ST}$ of

$\pi_S = \sum_{i=1}^k w_i \pi_{S_i}$ is given by

$$\hat{\pi}_{ST} = \sum_{i=1}^k w_i \hat{\pi}_{mi} = \sum_{i=1}^k w_i (n'_i / n_i) \quad (9)$$

Since n'_i / n_i is distributed as a binomial variate $B(n_i, Y_i)$, we, therefore, have the theorem below.

THEOREM 1. *The estimator $\hat{\pi}_{ST}$ is biased and the expression for bias is given by*

$$B(\hat{\pi}_{ST}) = \sum_{i=1}^k w_i (1 - T_i)(1 - P_i)(1 - 2\pi_{S_i})$$

PROOF. We have

$$\begin{aligned} B(\hat{\pi}_{ST}) &= E(\hat{\pi}_{ST}) - \pi_S \\ &= E\left(\sum_{i=1}^k w_i \hat{\pi}_{mi}\right) - \sum_{i=1}^k w_i \pi_{S_i} \\ &= \sum_{i=1}^k w_i E(\hat{\pi}_{mi} - \pi_{S_i}) \\ &= \sum_{i=1}^k w_i (Y_i - \pi_{S_i}) \\ &= \sum_{i=1}^k w_i [T_i \pi_{S_i} + (1 - T_i)\{P_i \pi_{S_i} + (1 - P_i)(1 - \pi_{S_i})\} - \pi_{S_i}] \\ &= \sum_{i=1}^k w_i (1 - T_i)(1 - P_i)(1 - 2\pi_{S_i}) \end{aligned}$$

which proves the theorem.

In order to study the performance of the estimator $\hat{\pi}_{ST}$, we need its mean square error (MSE) which is given by

$$\sum_{i=1}^k MSE(\hat{\pi}_{ST}) = \sum_{i=1}^k w_i^2 \frac{Y_i(1 - Y_i)}{n_i} + \left[\sum_{i=1}^k w_i (1 - T_i)(1 - P_i)(1 - 2\pi_{S_i}) \right]^2, \quad (10)$$

where Y_i is given by (7).

Now we will obtain the MSE ($\hat{\pi}_{ST}$) under (i) Proportional allocation, and (ii) Optimum allocation.

(i) Proportional Allocation

The MSE ($\hat{\pi}_{ST}$) under the proportional allocation is given in the following theorem.

THEOREM 2. Under the proportional allocation (i.e. $n_i = n(N_i / N)$), the MSE ($\hat{\pi}_{ST}$) is given by

$$MSE(\hat{\pi}_{ST})_P = \frac{1}{n} \left[\sum_{i=1}^k w_i Y_i (1 - Y_i) \right] + \left[\sum_{i=1}^k w_i (1 - T_i) (1 - P_i) (1 - 2\pi_{Si}) \right]^2 \quad (11)$$

Proof is simple so omitted

(ii) Optimum Allocation

Information on π_{Si} and T_i are usually unavailable. But if prior information on π_{Si} and T_i are available from past experience then it helps to derive the following optimal allocation formula.

THEOREM 3. The optimal allocation of n to n_1, n_2, \dots, n_{k-1} and n_k to derive the minimum mean square error subject to $n = \sum_{i=1}^k n_i$ is approximately given by

$$\frac{n_i}{n} = \frac{w_i \sqrt{Y_i (1 - Y_i)}}{\sum_{i=1}^k w_i \sqrt{Y_i (1 - Y_i)}} \quad (12)$$

Proof is simple so omitted.

THEOREM 4. Under the optimal allocation (12), the minimum mean square error of $\hat{\pi}_{ST}$ is given by

$$MSE(\hat{\pi}_{ST})_O = \frac{1}{n} \left[\sum_{i=1}^k w_i \sqrt{Y_i (1 - Y_i)} \right]^2 + \left[\sum_{i=1}^k w_i (1 - T_i) (1 - P_i) (1 - 2\pi_{Si}) \right]^2 \quad (13)$$

PROOF. Inserting (12) in (10) one can easily get (13).

3. RELATIVE EFFICIENCY

Suppose that there are two strata (i.e. $k = 2$) in the population, $\pi_{S1} \neq \pi_{S2}$, $P = P_1 = P_2$ and $T = T_1 = T_2$, the mean square error of Mangat and Singh (1994) estimator $\hat{\pi}_{mS}$ is given by

$$\begin{aligned} \tilde{MSE}(\hat{\pi}_{mS}) = & \frac{[\pi_S(1 - \pi_S)\{2P - 1 + 2T(1 - P)\}^2 + (1 - T)(1 - P)\{1 - (1 - T)(1 - P)\}]}{n} \\ & + [(1 - T)(1 - P)(1 - 2\pi_S)]^2 \end{aligned} \quad (14)$$

where $\pi_S = w_1\pi_{S1} + w_2\pi_{S2}$.

3.1. Proportional Allocation

For two strata (i.e. $k = 2$) in the population and $P = P_1 = P_2$, $MSE(\hat{\pi}_S)$ in (4) reduces to:

$$\tilde{MSE}(\hat{\pi}_S)_P = \frac{1}{n} \left[\sum_{i=1}^k w_i \left\{ \pi_{Si}(1 - \pi_{Si}) + \frac{P(1-P)}{(2P-1)^2} \right\} \right] \quad (15)$$

Under the assumptions $k = 2$, $P = P_1 = P_2$ and $T = T_1 = T_2$, the $MSE(\hat{\pi}_{ST})_P$ in (5) reduces to:

$$\tilde{MSE}(\hat{\pi}_{ST})_P = \frac{1}{n} \left[\sum_{i=1}^k w_i Y_i^* (1 - Y_i^*) \right] + [(1-T)(1-P)(1-2\pi_S)]^2 \quad (16)$$

where $\pi_S = (w_1\pi_{S1} + w_2\pi_{S2})$ and $Y_i^* = T\pi_{Si} + (1-T)\{P\pi_{Si} + (1-P)(1-\pi_{Si})\}$, for $i = 1, 2, \dots, k$.

From (14), (15) and (16), the percent relative efficiency (PRE) of the proposed estimator $\hat{\pi}_{ST}$ (under proportional allocation) with respect to Mangat and Singh's (1994) estimator $\hat{\pi}_{mS}$ and the estimator $\hat{\pi}_S$ (under proportional allocation i.e. Hong *et al.*'s estimator) are respectively given by

$$PRE((\hat{\pi}_{ST})_P, \hat{\pi}_{mS}) = \frac{\tilde{MSE}(\hat{\pi}_{mS})}{\tilde{MSE}(\hat{\pi}_{ST})_P} \times 100 \quad (17)$$

and

$$PRE((\hat{\pi}_{ST})_P, (\hat{\pi}_S)_P) = \frac{\tilde{MSE}(\hat{\pi}_S)_P}{\tilde{MSE}(\hat{\pi}_{ST})_P} \quad (18)$$

We have computed the percent relative efficiencies $PRE(\hat{\pi}_{ST}, \hat{\pi}_{mS})$ and $PRE(\hat{\pi}_{ST}, \hat{\pi}_S)$ for different values of $n, P, w_1, w_2, \pi_{S1}, \pi_{S2}$ and T . Findings are shown in Table 1 and Table 2 respectively. Diagrammatic representations are also given in Fig. 1 and Fig. 2.

It is observed from Table 1 to 2 and Fig. 1 to 2 that:

the values of percent relative efficiencies $PRE((\hat{\pi}_{ST})_P, \hat{\pi}_{mS})$ and $PRE((\hat{\pi}_{ST})_P, \hat{\pi}_S)$ are larger than 100. We can say that the envisaged estimator $\hat{\pi}_{ST}$ (under proportional allocation) is more efficient than the proposed estimator $\hat{\pi}_S$ (under proportional allocation i.e. Hong *et al.*'s estimator) and that of the Mangat and Singh's (1994) estimator $\hat{\pi}_{mS}$. Fig. 1 and Fig. 2 show results for $P=0.72, 0.81, T=0.5, 0.7, 0.9$ and different values of $n, \pi_{S1}, \pi_{S2}, w_1, w_2$.

It is observed from Table 1 that the values of the relative efficiency $PRE((\hat{\pi}_{ST})_P, \hat{\pi}_{mS})$ increase as the value of P increases. Table 2 exhibits that the values of the percent relative efficiency $PRE(\hat{\pi}_{ST}, \hat{\pi}_S)$ decrease as the value of P increases.

We further note from the results of Fig. 1 and 2 that there is large gain in efficiency by using the suggested estimator $\hat{\pi}_{ST}$ under proportional allocation over the proposed estimator $\hat{\pi}_S$ (Hong *et al.*'s (1994) estimator) as well as Mangat and Singh's (1994) estimator $\hat{\pi}_{mS}$. Thus the proposed

estimator $\hat{\pi}_Y$ under proportional allocation is to be proposed over Hong *et al.*'s (1994) estimator $\hat{\pi}_S$ and the Mangat and Singh's (1994) estimator $\hat{\pi}_{mS}$.

TABLE 1;
 Percent relative efficiency of the proposed estimator $\hat{\pi}_{ST}$ (under proportional allocation) with respect to
 Mangat and Singh (1994) estimator $\hat{\pi}_{mS}$.

n	π_{s1}	π_{s2}	w_1	w_2	T	π_s	P							
							0.6	0.63	0.66	0.69	0.72	0.75	0.78	0.81
10	0.08	0.13	0.90	0.10	0.50	0.09	188.43	206.37	228.24	255.05	288.07	328.92	379.72	443.16
10	0.08	0.13	0.80	0.20	0.50	0.09	190.02	208.03	229.95	256.71	289.53	329.97	379.98	442.04
10	0.08	0.13	0.70	0.30	0.50	0.10	191.62	209.71	231.64	258.34	290.96	330.97	380.18	440.87
10	0.08	0.13	0.60	0.40	0.50	0.10	193.22	211.38	233.33	259.96	292.36	331.92	380.33	439.66
10	0.08	0.13	0.50	0.50	0.50	0.11	194.84	213.05	235.01	261.55	293.73	332.83	380.43	438.41
10	0.08	0.13	0.40	0.60	0.50	0.11	196.46	214.73	236.68	263.13	295.07	333.70	380.48	437.12
10	0.08	0.13	0.30	0.70	0.50	0.12	198.09	216.40	238.34	264.69	296.37	334.52	380.48	435.79
10	0.08	0.13	0.20	0.80	0.50	0.12	199.72	218.08	240.00	266.22	297.63	335.29	380.42	434.43
10	0.08	0.13	0.10	0.90	0.50	0.13	201.36	219.75	241.63	267.73	298.86	336.01	380.32	433.02
20	0.18	0.23	0.90	0.10	0.70	0.19	309.02	331.19	355.59	382.33	411.50	443.14	477.19	513.50
20	0.18	0.23	0.80	0.20	0.70	0.19	310.79	332.71	356.75	383.02	411.59	442.47	475.59	510.79
20	0.18	0.23	0.70	0.30	0.70	0.20	312.52	334.18	357.86	383.66	411.63	441.75	473.95	508.06
20	0.18	0.23	0.60	0.40	0.70	0.20	314.21	335.60	358.92	384.24	411.60	440.97	472.27	505.33
20	0.18	0.23	0.50	0.50	0.70	0.21	315.87	336.98	359.92	384.76	411.52	440.15	470.57	502.58
20	0.18	0.23	0.40	0.60	0.70	0.21	317.49	338.30	360.87	385.23	411.38	439.28	468.82	499.83
20	0.18	0.23	0.30	0.70	0.70	0.22	319.06	339.58	361.76	385.63	411.18	438.36	467.05	497.07
20	0.18	0.23	0.20	0.80	0.70	0.22	320.59	340.81	362.60	385.98	410.93	437.38	465.24	494.30
20	0.18	0.23	0.10	0.90	0.70	0.23	322.08	341.99	363.38	386.26	410.62	436.37	463.40	491.53
30	0.28	0.33	0.90	0.10	0.90	0.29	499.84	507.19	514.50	521.77	528.97	536.11	543.17	550.13
30	0.28	0.33	0.80	0.20	0.90	0.29	497.81	504.94	512.03	519.08	526.08	533.01	539.86	546.63
30	0.28	0.33	0.70	0.30	0.90	0.30	495.75	502.67	509.56	516.40	523.19	529.92	536.57	543.15
30	0.28	0.33	0.60	0.40	0.90	0.30	493.69	500.40	507.08	513.72	520.31	526.85	533.31	539.70
30	0.28	0.33	0.50	0.50	0.90	0.31	491.60	498.12	504.60	511.05	517.45	523.79	530.07	536.28
30	0.28	0.33	0.40	0.60	0.90	0.31	489.50	495.83	502.12	508.38	514.59	520.75	526.85	532.88
30	0.28	0.33	0.30	0.70	0.90	0.32	487.39	493.53	499.63	505.71	511.74	517.72	523.64	529.51
30	0.28	0.33	0.20	0.80	0.90	0.32	485.26	491.22	497.14	503.04	508.89	514.70	520.46	526.17
30	0.28	0.33	0.10	0.90	0.90	0.33	483.12	488.90	494.65	500.37	506.06	511.71	517.30	522.85

TABLE 2;
 Percent relative efficiency of the proposed estimator $\hat{\pi}_{ST}$ (under proportional allocation) with respect to the
 Hong et al.'s estimator $\hat{\pi}_S$ (under proportional allocation).

n	π_{S1}	π_{S2}	w_1	w_2	T	π_S	P							
							0.6	0.63	0.66	0.69	0.72	0.75	0.78	0.81
10	0.08	0.13	0.90	0.10	0.50	0.09	1311.29	844.84	606.72	467.56	378.48	317.51	273.48	240.12
10	0.08	0.13	0.80	0.20	0.50	0.09	1326.97	854.14	612.77	471.69	381.34	319.46	274.72	240.77
10	0.08	0.13	0.70	0.30	0.50	0.10	1342.77	863.47	618.82	475.80	384.18	321.38	275.93	241.40
10	0.08	0.13	0.60	0.40	0.50	0.10	1358.68	872.85	624.87	479.90	386.99	323.27	277.10	241.99
10	0.08	0.13	0.50	0.50	0.50	0.11	1374.71	882.25	630.92	483.97	389.77	325.12	278.24	242.56
10	0.08	0.13	0.40	0.60	0.50	0.11	1390.85	891.69	636.97	488.02	392.51	326.93	279.35	243.09
10	0.08	0.13	0.30	0.70	0.50	0.12	1407.08	901.15	643.01	492.05	395.22	328.71	280.42	243.60
10	0.08	0.13	0.20	0.80	0.50	0.12	1423.42	910.64	649.03	496.05	397.90	330.45	281.46	244.09
10	0.08	0.13	0.10	0.90	0.50	0.13	1439.84	920.14	655.04	500.01	400.54	332.15	282.46	244.54
20	0.18	0.23	0.90	0.10	0.70	0.19	2004.32	1251.28	868.76	645.61	502.54	404.20	332.81	278.63
20	0.18	0.23	0.80	0.20	0.70	0.19	2017.75	1257.96	872.22	647.31	503.20	404.21	332.42	278.01
20	0.18	0.23	0.70	0.30	0.70	0.20	2030.96	1264.49	875.57	648.92	503.79	404.18	332.01	277.38
20	0.18	0.23	0.60	0.40	0.70	0.20	2043.94	1270.86	878.80	650.45	504.32	404.10	331.57	276.74
20	0.18	0.23	0.50	0.50	0.70	0.21	2056.68	1277.06	881.90	651.89	504.79	403.99	331.12	276.09
20	0.18	0.23	0.40	0.60	0.70	0.21	2069.17	1283.08	884.88	653.24	505.20	403.84	330.64	275.44
20	0.18	0.23	0.30	0.70	0.70	0.22	2081.39	1288.93	887.73	654.50	505.55	403.65	330.13	274.77
20	0.18	0.23	0.20	0.80	0.70	0.22	2093.34	1294.59	890.46	655.67	505.83	403.42	329.61	274.11
20	0.18	0.23	0.10	0.90	0.70	0.23	2105.00	1300.06	893.05	656.75	506.06	403.15	329.07	273.43
30	0.28	0.33	0.90	0.10	0.90	0.29	2825.45	1676.94	1108.44	785.62	584.57	450.75	357.10	288.93
30	0.28	0.33	0.80	0.20	0.90	0.29	2810.43	1667.70	1102.21	781.17	581.29	448.28	355.23	287.51
30	0.28	0.33	0.70	0.30	0.90	0.30	2795.46	1658.52	1096.02	776.76	578.04	445.84	353.38	286.11
30	0.28	0.33	0.60	0.40	0.90	0.30	2780.54	1649.38	1089.88	772.39	574.83	443.43	351.55	284.72
30	0.28	0.33	0.50	0.50	0.90	0.31	2765.67	1640.29	1083.77	768.06	571.64	441.05	349.74	283.36
30	0.28	0.33	0.40	0.60	0.90	0.31	2750.85	1631.25	1077.72	763.77	568.50	438.69	347.96	282.01
30	0.28	0.33	0.30	0.70	0.90	0.32	2736.08	1622.26	1071.71	759.52	565.38	436.36	346.21	280.69
30	0.28	0.33	0.20	0.80	0.90	0.32	2721.37	1613.32	1065.74	755.30	562.29	434.06	344.47	279.38
30	0.28	0.33	0.10	0.90	0.90	0.33	2706.70	1604.43	1059.81	751.12	559.24	431.78	342.75	278.09

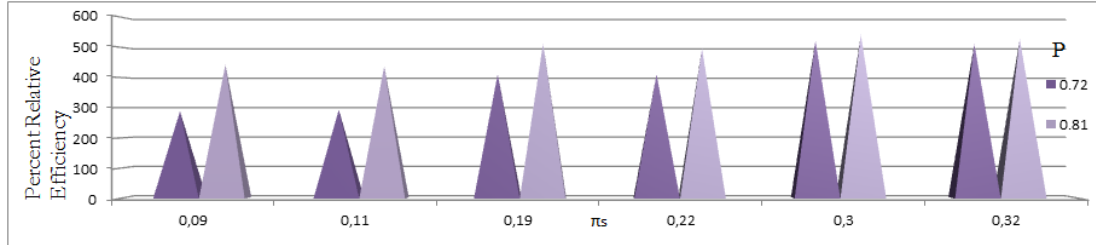


Figure 1 – Percent relative efficiency of the proposed estimator $\hat{\pi}_{ST}$ (under proportional allocation) with respect to Mangat and Singh (1994) estimator $\hat{\pi}_{mS}$.

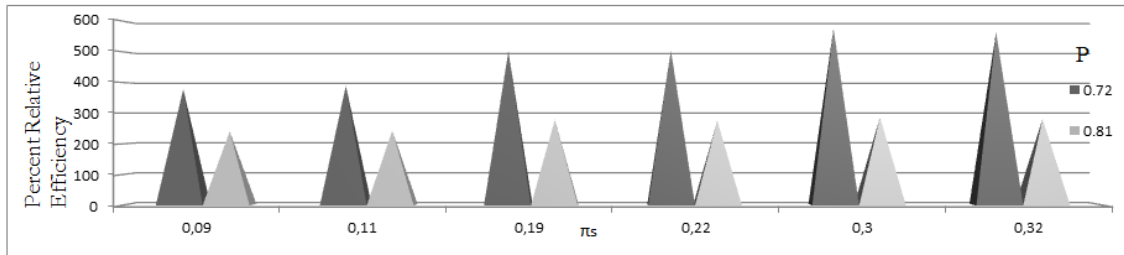


Figure 2 – Percent relative efficiency of the proposed estimator (under proportional allocation) with respect to the estimator (under proportional allocation i.e. Hong et al.'s (1994) estimator).

3.2 Optimum Allocation

Under the assumption $k = 2$ (i.e. two strata in the population), $P = P_1 = P_2$, $T = T_1 = T_2$,

$MSE(\hat{\pi}_S)$ in (6) and $MSE(\hat{\pi}_{ST})$ in (13) respectively reduce to:

$$\tilde{MSE}(\hat{\pi}_S)_O = \frac{1}{n} \left[\sum_{i=1}^2 w_i \left\{ \pi_{Si}(1 - \pi_{Si}) + \frac{P(1-P)}{(2P-1)^2} \right\}^{1/2} \right]^2 \tag{19}$$

and

$$\tilde{MSE}(\hat{\pi}_{ST})_O = \frac{1}{n} \left[\sum_{i=1}^2 w_i \sqrt{Y_i^*(1 - Y_i^*)} \right]^2 + [(1-T)(1-P)(1-2\pi_S)]^2 \tag{20}$$

From (14), (19) and (20), the percent relative efficiency (PRE) of the proposed estimator $\hat{\pi}_{ST}$ (under optimum allocation) with respect to Mangat and Singh's (1994) estimator $\hat{\pi}_{mS}$ and the estimator $\hat{\pi}_S$ (under optimum allocation i.e. Kim and Warde's (2004) estimator) are respectively given by

$$PRE((\hat{\pi}_{ST})_O, \hat{\pi}_{mS}) = \frac{\tilde{MSE}(\hat{\pi}_{mS})}{\tilde{MSE}(\hat{\pi}_{ST})_O} \times 100 \quad (21)$$

and

$$PRE((\hat{\pi}_{ST})_O, (\hat{\pi}_S)_O) = \frac{\tilde{MSE}(\hat{\pi}_S)_O}{\tilde{MSE}(\hat{\pi}_{ST})_O} \times 100 \quad (22)$$

We have computed the percent relative efficiencies $PRE(\hat{\pi}_{ST}, \hat{\pi}_{mS})$ and $PRE((\hat{\pi}_{ST})_O, (\hat{\pi}_S)_O)$ for various values of n , P , w_1 , w_2 , π_{S1} , π_{S2} and T . Results are compiled in Table 3 and Table 4 respectively. Diagrammatic representations are also given in Fig. 3 and Fig. 4. The values of $PRE((\hat{\pi}_{ST})_O, \hat{\pi}_{mS})$ and $PRE((\hat{\pi}_{ST})_O, (\hat{\pi}_S)_O)$ are greater than 100 for all values of n , P , w_1 , w_2 , π_{S1} , π_{S2} and T considered here. So we can say that the envisaged estimator $\hat{\pi}_{ST}$ (under optimum allocation) is more efficient than the Kim and Warde's (2004) estimator $\hat{\pi}_S$ (under optimum allocation) and that of the Mangat and Singh's (1994) estimator $\hat{\pi}_{mS}$. Fig. 3 and Fig. 4 show results for $P = 0.72, 0.81$, $T = 0.5, 0.7, 0.9$ and different values of n , P , w_1 , w_2 , π_{S1} , π_{S2} . We note from Table 3 that the values of the percent relative efficiencies $PRE(\hat{\pi}_{ST}, \hat{\pi}_{mS})$ increase as the value of P increases. Table 4 demonstrates that the values of the percent relative efficiency $PRE(\hat{\pi}_{ST}, \hat{\pi}_S)$ decrease as the value of P increases. It is further observed from the results of Fig. 3 and 4 that there is large gain in efficiency by using the proposed estimator $\hat{\pi}_{ST}$ under optimum allocation to that of the Kim and Warde's (2004) estimator $\hat{\pi}_S$ as well as Mangat and Singh's (1994) estimator $\hat{\pi}_{mS}$. Thus our recommendation is to use the proposed estimator $\hat{\pi}_{ST}$ (under optimum allocation) over Kim and Warde's (2004) estimator $\hat{\pi}_S$ and Mangat and Singh's (1994) estimator $\hat{\pi}_{mS}$ in presence of prior information of π_{S1} and π_{S2} .

TABLE 3
 Percent relative efficiency of the proposed estimator $\hat{\pi}_{ST}$ (under optimum allocation) with respect to Mangat
 and Singh's (1994) estimator $\hat{\pi}_{mS}$.

n	π_{s1}	π_{s2}	w_1	w_2	T	π_s	P							
							0.6	0.63	0.66	0.69	0.72	0.75	0.78	0.81
10	0.08	0.13	0.90	0.10	0.50	0.09	188.44	206.38	228.27	255.09	288.12	329.01	379.86	443.39
10	0.08	0.13	0.80	0.20	0.50	0.09	190.04	208.06	229.99	256.77	289.63	330.12	380.23	442.44
10	0.08	0.13	0.70	0.30	0.50	0.10	191.64	209.74	231.69	258.42	291.09	331.17	380.51	441.40
10	0.08	0.13	0.60	0.40	0.50	0.10	193.25	211.42	233.39	260.05	292.51	332.16	380.70	440.26
10	0.08	0.13	0.50	0.50	0.50	0.11	194.87	213.10	235.08	261.66	293.89	333.08	380.82	439.03
10	0.08	0.13	0.40	0.60	0.50	0.11	196.49	214.77	236.75	263.23	295.22	333.93	380.85	437.71
10	0.08	0.13	0.30	0.70	0.50	0.12	198.11	216.44	238.40	264.77	296.50	334.72	380.80	436.30
10	0.08	0.13	0.20	0.80	0.50	0.12	199.74	218.10	240.04	266.29	297.73	335.45	380.67	434.81
10	0.08	0.13	0.10	0.90	0.50	0.13	201.38	219.76	241.66	267.77	298.92	336.10	380.45	433.23
20	0.18	0.23	0.90	0.10	0.70	0.19	309.06	331.24	355.64	382.40	411.59	443.25	477.32	513.68
20	0.18	0.23	0.80	0.20	0.70	0.19	310.85	332.78	356.85	383.15	411.75	442.66	475.83	511.09
20	0.18	0.23	0.70	0.30	0.70	0.20	312.60	334.28	357.99	383.82	411.83	442.00	474.26	508.46
20	0.18	0.23	0.60	0.40	0.70	0.20	314.30	335.71	359.06	384.42	411.83	441.26	472.63	505.77
20	0.18	0.23	0.50	0.50	0.70	0.21	315.97	337.10	360.07	384.95	411.76	440.45	470.93	503.04
20	0.18	0.23	0.40	0.60	0.70	0.21	317.58	338.42	361.02	385.41	411.61	439.56	469.17	500.27
20	0.18	0.23	0.30	0.70	0.70	0.22	319.14	339.69	361.89	385.79	411.38	438.60	467.35	497.45
20	0.18	0.23	0.20	0.80	0.70	0.22	320.66	340.89	362.70	386.10	411.08	437.57	465.47	494.59
20	0.18	0.23	0.10	0.90	0.70	0.23	322.12	342.03	363.43	386.33	410.70	436.47	463.52	491.69
30	0.28	0.33	0.90	0.10	0.90	0.29	499.91	507.26	514.57	521.84	529.05	536.19	543.25	550.22
30	0.28	0.33	0.80	0.20	0.90	0.29	497.92	505.06	512.16	519.21	526.21	533.15	540.01	546.78
30	0.28	0.33	0.70	0.30	0.90	0.30	495.90	502.83	509.72	516.57	523.37	530.10	536.77	543.35
30	0.28	0.33	0.60	0.40	0.90	0.30	493.85	500.58	507.27	513.92	520.52	527.06	533.53	539.93
30	0.28	0.33	0.50	0.50	0.90	0.31	491.77	498.30	504.79	511.25	517.65	524.00	530.29	536.51
30	0.28	0.33	0.40	0.60	0.90	0.31	489.67	496.00	502.30	508.56	514.78	520.95	527.06	533.10
30	0.28	0.33	0.30	0.70	0.90	0.32	487.53	493.67	499.79	505.87	511.90	517.89	523.83	529.70
30	0.28	0.33	0.20	0.80	0.90	0.32	485.37	491.33	497.26	503.16	509.02	514.84	520.60	526.31
30	0.28	0.33	0.10	0.90	0.90	0.33	483.18	488.96	494.72	500.44	506.13	511.78	517.38	522.93

TABLE 4
 Percent relative efficiency of the proposed estimator $\hat{\pi}_{ST}$ (under optimum allocation) with respect to the Kim and Warde's (2004) estimator $\hat{\pi}_S$.

n	π_{S1}	π_{S2}	w_1	w_2	T	π_S	P							
							0.6	0.63	0.66	0.69	0.72	0.75	0.78	0.81
10	0.08	0.13	0.90	0.10	0.50	0.09	1311.36	844.90	606.77	467.61	378.54	317.58	273.56	240.21
10	0.08	0.13	0.80	0.20	0.50	0.09	1327.09	854.24	612.87	471.79	381.45	319.58	274.86	240.93
10	0.08	0.13	0.70	0.30	0.50	0.10	1342.93	863.61	618.95	475.94	384.33	321.54	276.11	241.60
10	0.08	0.13	0.60	0.40	0.50	0.10	1358.87	873.01	625.03	480.06	387.16	323.45	277.31	242.22
10	0.08	0.13	0.50	0.50	0.50	0.11	1374.91	882.43	631.09	484.14	389.94	325.31	278.46	242.80
10	0.08	0.13	0.40	0.60	0.50	0.11	1391.04	891.86	637.13	488.19	392.68	327.12	279.56	243.33
10	0.08	0.13	0.30	0.70	0.50	0.12	1407.26	901.31	643.15	492.20	395.38	328.87	280.60	243.81
10	0.08	0.13	0.20	0.80	0.50	0.12	1423.55	910.76	649.14	496.16	398.02	330.58	281.59	244.24
10	0.08	0.13	0.10	0.90	0.50	0.13	1439.92	920.21	655.10	500.08	400.60	332.22	282.54	244.63
20	0.18	0.23	0.90	0.10	0.70	0.19	2004.53	1251.43	868.89	645.72	502.64	404.29	332.90	278.71
20	0.18	0.23	0.80	0.20	0.70	0.19	2018.13	1258.24	872.45	647.51	503.37	404.37	332.57	278.14
20	0.18	0.23	0.70	0.30	0.70	0.20	2031.47	1264.87	875.87	649.18	504.02	404.38	332.20	277.55
20	0.18	0.23	0.60	0.40	0.70	0.20	2044.53	1271.29	879.14	650.74	504.58	404.34	331.79	276.94
20	0.18	0.23	0.50	0.50	0.70	0.21	2057.30	1277.51	882.26	652.20	505.06	404.24	331.34	276.29
20	0.18	0.23	0.40	0.60	0.70	0.21	2069.77	1283.52	885.23	653.53	505.46	404.08	330.85	275.63
20	0.18	0.23	0.30	0.70	0.70	0.22	2081.93	1289.31	888.04	654.76	505.77	403.85	330.32	274.94
20	0.18	0.23	0.20	0.80	0.70	0.22	2093.75	1294.89	890.69	655.87	506.01	403.58	329.75	274.23
20	0.18	0.23	0.10	0.90	0.70	0.23	2105.23	1300.23	893.19	656.86	506.16	403.24	329.15	273.50
30	0.28	0.33	0.90	0.10	0.90	0.29	2825.81	1677.16	1108.60	785.73	584.65	450.82	357.15	288.97
30	0.28	0.33	0.80	0.20	0.90	0.29	2811.07	1668.10	1102.48	781.36	581.43	448.40	355.32	287.58
30	0.28	0.33	0.70	0.30	0.90	0.30	2796.29	1659.02	1096.36	777.01	578.23	445.99	353.49	286.20
30	0.28	0.33	0.60	0.40	0.90	0.30	2781.48	1649.95	1090.26	772.68	575.04	443.60	351.68	284.83
30	0.28	0.33	0.50	0.50	0.90	0.31	2766.64	1640.88	1084.18	768.35	571.86	441.22	349.88	283.47
30	0.28	0.33	0.40	0.60	0.90	0.31	2751.77	1631.81	1078.10	764.05	568.70	438.85	348.09	282.12
30	0.28	0.33	0.30	0.70	0.90	0.32	2736.87	1622.75	1072.04	759.76	565.56	436.50	346.32	280.78
30	0.28	0.33	0.20	0.80	0.90	0.32	2721.96	1613.69	1065.99	755.48	562.43	434.16	344.55	279.44
30	0.28	0.33	0.10	0.90	0.90	0.33	2707.03	1604.63	1059.95	751.22	559.32	431.84	342.80	278.12

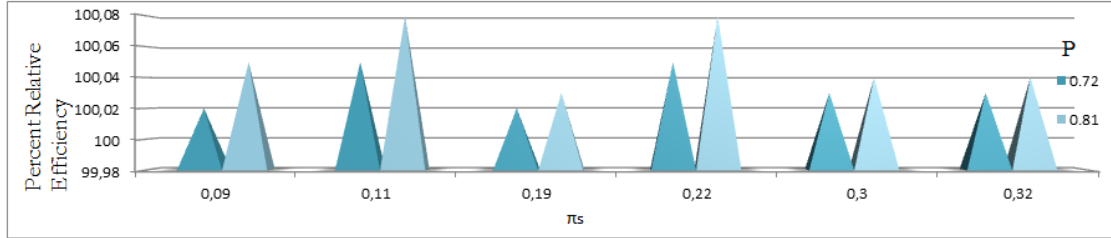


Figure 3 - Percent relative efficiency of the proposed estimator (under optimum allocation) with respect to Manga and Singh (1994) estimator

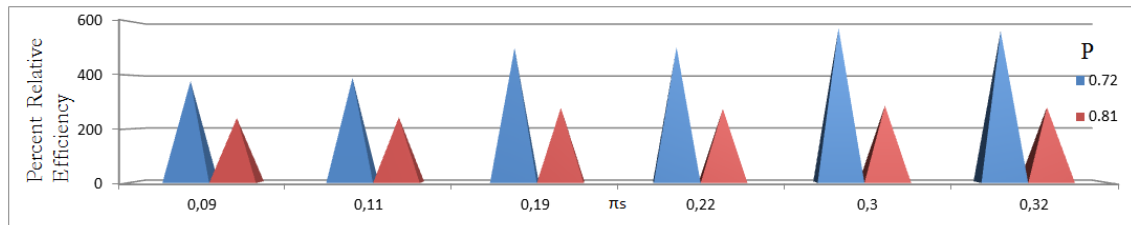


Figure 4 - Percent relative efficiency of the proposed estimator $\hat{\pi}_{ST}$ (under optimum allocation) with respect to th Kim and Warde's

To have the tangible idea about the performance of the proposed estimator $\hat{\pi}_{ST}$ under optimum allocation to that of under proportional allocation, we have computed the percent relative efficiency of the proposed estimator $\hat{\pi}_{ST}$ under optimum allocation with respect to $\hat{\pi}_{ST}$ under proportional allocation by using the formula:

$$PRE ((\hat{\pi}_{ST})_O, (\hat{\pi}_{ST})_P) = \frac{\tilde{MSE}(\hat{\pi}_{ST})_P}{\tilde{MSE}(\hat{\pi}_{ST})_O} \times 100 \tag{23}$$

For different values of $n, P, w_1, w_2, \pi_{S1}, \pi_{S2}$ and T .

Findings are shown in Table 5 and its diagrammatic representation is given by the figure 5.

We note from Table 5 that the values of the percent relative efficiencies $PRE(\hat{\pi}_{ST}, \hat{\pi}_{ST})$ increase as the value of P increases.

Fig. 5 show results for $P = 0.72, 0.81, T = 0.5, 0.7, 0.9$ and different values of $\pi_{S1}, \pi_{S2}, w_1, w_2$.

We further note from the results of Fig. 5 that there is a marginal gain in efficiency by using the suggested estimator $\hat{\pi}_{ST}$ under optimum allocation to that of under proportional allocation.

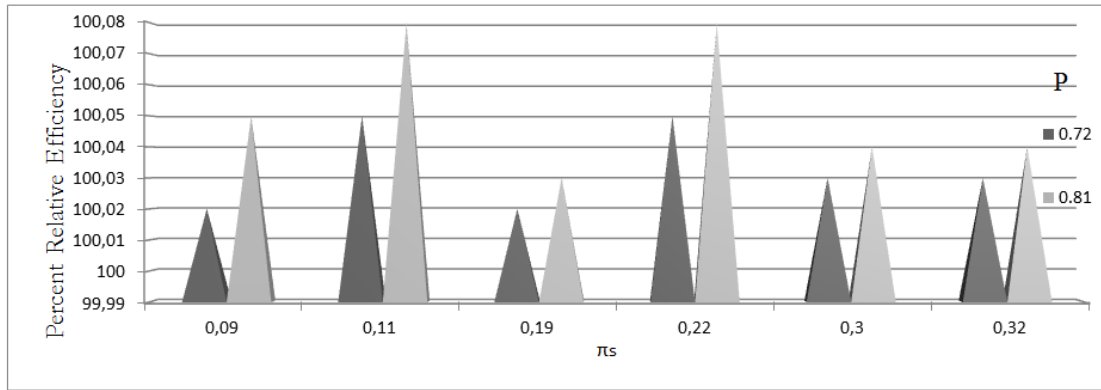


Figure 5 - Percent relative efficiency of the proposed estimator $\hat{\pi}_{ST}$ under optimum allocation with respect to $\hat{\pi}_{ST}$ under proportional allocation

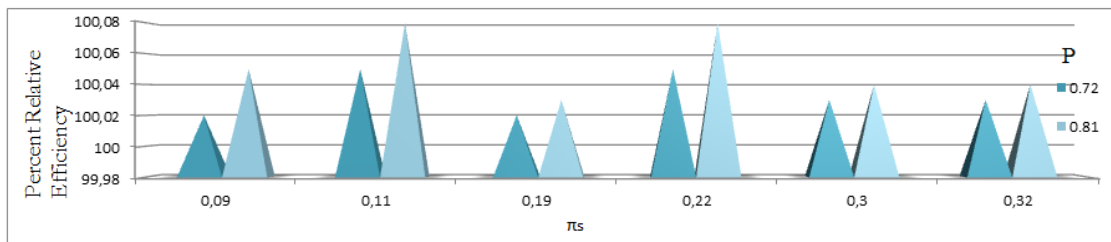


Figure 3 - Percent relative efficiency of the proposed estimator $\hat{\pi}_{ST}$ (under optimum allocation) with respect to Mangat and Singh (1994) estimator $\hat{\pi}_{mS}$.

4. DISCUSSION

This paper addresses the problem of estimating the proportion π_S of the population belonging to a sensitive group using optional randomized response technique in stratified sampling. A stratified Mangat and Singh's optional randomized response model using Mangat and Singh's (1994) model has been proposed. It has been shown that the proposed randomized response model is more efficient than the Hong *et al.*'s (1994), Mangat and Singh's (1994) and Kim and Warde's (2004) stratified randomize response models.

In addition to the gain in efficiency, the proposed methods are more beneficial than the previous methods.

Two more advantages exist with stratified RR models using optimal allocation. The first is that they solve a limitation of RR technique which is the loss of individual characteristics of the respondents. Also, using optimal allocation helps to overcome the high cost incurred because of the difficulty in obtaining a proportional sample from a stratum (as mentioned in the Kim and Warde's (2004) model).

ACKNOWLEDGEMENTS

The authors are grateful to the Editor – in- Chief and to the learned referee for annotating the original manuscript in the present form.

REFERENCES

- A. CHAURHURI, R. MUKERJEE (1985), *Optionally randomized response techniques*. "Calcutta Statistical Association Bulletin, 34, pp.225 – 229.
- A. CHAURHURI, R. MUKERJEE (1988). *Randomized Response: Theory and Techniques*. Marcel-Dekker, New York, USA.
- A. NAZUK, J. SHABIR (2010). *A new mixed randomized response model*. "International Journal of Business Social Sciences", 1 No. 1.
- A.S. HEDAYAY, B.K. SINGHA (1991). *Design and Inference in Finite Population Sampling*. New York, Wiley.
- B. GREENBERG, A. ABUL-ELA, W.R. SIMMONS, D.G. HORVITZ (1969). *The unreleased question randomized response: Theoretical framework*. "Journal of American Statistical Association, 64, pp. 529-539.
- D. S. TRACY, N.S. MANGAT (1995). *A partial randomized response strategy*. "Test", 4 (2), pp. 315–321.
- D.S. TRACY, N.S. MANGAT (1996). *Some developments in randomized response sampling during the last decade – A follow up of review by Chaudhuri and Mukherjee*. "Journal of Applied Statistical Sciences, 4 (2/3), pp.147–158.
- D.S. TRACY, S.S. OSAHAN (1999). *An improved randomized response technique*. "Pakistan Journal of Statistics, 15, (1), pp.1–6.
- G.S. LEE, D. UHM, J.M. KIM (2011). *Estimation of a rare sensitive attribute in a stratified sample using Poisson distribution*. "Statistics, iFirst", pp. 1-15.
- H.J. CHANG, C.L. WANG, K.C. HUANG (2004 a). *On estimating the proportion of a qualitative sensitive character using randomized response sampling*. "Quality and Quantity", 38, pp. 675-680.
- H.J. CHANG, C.L. WANG, K.C. HUANG (2004 b). *Using randomized response to estimate the proportion and truthful reporting probability in a dichotomous finite population*. "Journal of Applied Statistics", 31, pp. 565-573.
- H.J. CHANG, K.C. HUANG (2001). *Estimation of proportion and sensitivity of a qualitative character*. "Metrika", 53, pp. 269-280.

- H.P. SINGH, T. A. TARRAY (2012). *A stratified unknown repeated trials in randomized response sampling*. "Communication of the Korean Statistical Society, 19(6), pp. 751-759.
- J.A. FOX, P.E. TRACY (1986). *Randomized Response: A method of Sensitive Surveys*. Newbury Park, CA. SEGE Publications.
- J.B. RYU, K.H. HONG, G.S. LEE (1993). *Randomized response model*, Freedom Academy, Seoul, Korea.
- J.M. KIM, M.E. ELAM (2003). *A stratified unrelated question randomized response model*, "Journal of Statistical Planning and Inference", inreview.
- J.M. KIM, M.E. ELAM (2005). *A two-stage stratified Warner's randomized response model using optimal allocation*, "Metrika", 61, pp. 1-7.
- J.M. KIM, M.E. ELAM (2007). *A stratified unrelated randomized response model*, "Statistical Papers", 48, pp. 215-233.
- J.M. KIM, J.M. TEBBS, S.W. AN (2006). *Extensions of Mangat's randomized response model*. "Journal of Statistical Planning and Inference", 136, pp.1554-1567.
- J.M. KIM, W.D. WARDE (2005). *A mixed randomized response model*. "Journal of Statistical Planning and Inference", 133, pp. 211-221.
- J.M. KIM, W.D. WARDE (2004). *A stratified Warner randomized response model*. "Journal of Statistical Planning and Inference", 120, pp. 155-165.
- K. HONG, J.YUM AND H. LEE (1994). *A stratified randomized response technique*, "Korean Journal of Applied Statistics", 7, pp. 141-147.
- M. LAND, S. SINGH, S.A. SEDORY (2011). *Estimation of a rare attribute using Poisson distribution*. "Statistics, iFirst", pp.1-10.
- M. MOHAMMOD, S. SINGH, S. HORN (1998). *On the confidentiality guaranteed under randomized response sampling: a comparison with several new techniques*. "Biometrical Journal", 40:2, pp. 237-242.
- N.S. MANGAT (1994). *An improved randomized response strategy*. "Journal of Royal Statistical Society, B, 56 (1), pp. 93-95.
- N.S. MANGAT, S. SINGH (1994). *An optional randomized response sampling techniques*, "Journal of Indian Statistical Association, 32, pp. 71-75
- N.S. MANGAT, R. SINGH, S. SINGH (1997). *Violation of respondents privacy in Moor's - itsrectification through a random group strategy*. "Communication in Statistics theory and Methods, 26:3, pp. 743-754.
- N.S. MANGAT, R. SINGH (1990). *An alternative randomized procedure*. "Biometrika", 77, pp. 439-442.
- R. SINGH, N.S. MANGAT (1996). *Elements of Survey Sampling*, Kluwer Academic Publishers, Dordrecht, The Netherlands.
- S. GUPTA, J. SHABIR, R. LEMBO (2006). *Modification to Warner's model using blank cards*, "American

Journal of Mathematical Management Sciences”, 26, pp. 185-196.

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

S. SINGH, R. SINGH, N.S. MANGAT (2000). *Some alternative strategies to Moor's model in randomized response sampling*. “Journal of Statistical Planning and Inference”, 83, pp. 243–255.

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

S.L. WARNER (1965). *Randomized response: A survey technique for eliminating evasive answer bias*. “Journal of American Statistical Association”, 60, pp. 63-69.

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

SUMMARY

A stratified Mangat and Singh's optional randomized response model using proportional and optimal allocation

This paper suggests a stratified optional randomized response model based on Mangat and Singh (1994) model that has proportional and optimal allocation and larger gain in efficiency. Numerically it is found that the suggested model is more efficient than Kim and Warde (2004) stratified randomized response model and Mangat and Singh (1994) model. Graphical representations are also given in support of the present study.

Keywords: Randomized response technique; Stratified random sampling; Simple random sampling with replacement; Estimation of proportion; Mean square error.