Pak. J. Statist.
2009 Vol. 25(2), 205-220

IMPROVED ESTIMATION PROCEDURES FOR THE MEAN OF SENSITIVE VARIABLE USING RANDOMIZED RESPONSE MODEL

Zawar Hussain[1] and **Javid Shabbir**[2]
Department of Statistics, Quaid-i-Azam University, Islamabad, Pakistan.
Email: [1]zhlangah@yahoo.com; [2]jsqau@yahoo.com

ABSTRACT

In this article, we have considered the problem of estimation of mean of a rare sensitive quantitative variable μ_A, while surveying the human population, using the scrambled response technique suggested by Ryu et al. (2005). We have presented a family of estimators $\hat{\mu}_{A\lambda}$ to estimate the population mean μ_A and examined its properties. The obtained optimal estimator depends on the optimal value of the constant, say λ, which involves the different unknown population parameters, so proposed estimator has little practical application. To cover this deficiency, we have suggested two other estimators based on the estimated optimum value of λ and discussed their properties. We have compared the proposed estimators with Ryu et al. (2005) estimator using simple random sampling and stratified random sampling protocols.

KEY WORDS

Estimation of mean; simple random sampling with replacement; stratified random sampling; scrambled response; and mean squared error.

1. INTRODUCTION

Warner (1965) proposed a method to collect the data from the individuals in a randomized way for estimating the population proportion of individuals with sensitive attribute. This idea was further enhanced by Greenberg et al. (1971) to the estimation of mean of sensitive quantitative variables. Till now, there has been a rich growth in the realm of randomized response models. In fact, randomized response (RR) models are used as a tool to decrease the evasive answer bias and, of course, to provide privacy protection to the respondents in order to get them ready to divulge their response honestly. Some recent randomized response models, which allow the scrambling of true responses, are Eichhorn and Hayre (1983), Gupta et al. (2002), Singh and Mathur (2002a, b), Espejo and Singh (2003), Singh and Mathur (2003a, b), Singh and Mathur (2004a, b, c), Gupta and Shabbir (2004), Bar-Lev et al. (2004), Ryu et al. (2005), Singh and Mathur (2005), Arnab and Dorffner (2006), Singh and Mathur (2006), Hussain and Shabbir (2007), Hussain et al. (2007), Singh and Mathur (2007) and many others. In one sense or the other, these models have some demerits. For instance, in Eichhorn and Hayre (1983) model only one possibility of reporting the response is available. Gupta et al. (2002) method provided only one possibility of scrambling the true response. Ryu et al. (2005) proposed a simple extension of Gupta et al. (2002) model in two stages.

Using the concept of Gupta et al. (2002) and Mangat and Singh (1990), Ryu et al. (2005) have presented a two stage RR model. A sample of size n is taken using simple random sample with replacement. The i^{th} respondent selected in the sample is requested to use the randomization device R_1 which consists of two statements: (i) “Report the true response A of sensitive question” and (ii) “Go to randomization device R_2 in the second stage” represented with probabilities Q_1 and $1-Q_1$ respectively. The randomization device R_2 consists of two statements: (i) “Report the true response A of sensitive question” and (ii) “Report the scrambled response AB of sensitive question” represented by probabilities Q_2 and $1-Q_2$ respectively. Using the assumption of known distribution of scrambling variable B such that $\mu_B = 1$ and $\sigma_B^2 = \psi^2$, the response of i^{th} respondent can be written as

$$U_i = \alpha A_i + (1-\alpha)\left[\beta A_i + (1-\beta) A_i B_i\right], \tag{1.1}$$

where $\alpha = 1$, if a respondent is randomly assigned to the statement (i) in R_1, and $\alpha = 0$, if a respondent is randomly assigned to statement (ii) in R_1. Further, $\beta = 1$, if a respondent is randomly assigned to statement (i) in R_2, and $\beta = 0$, if a respondent is randomly assigned to statement (ii) in R_2. The expected value of the observed response is given by

$$E(U_i) = E\left[\alpha A_i + (1-\alpha)\left\{\beta A_i + (1-\beta) A_i B_i\right\}\right]$$

$$E(U_i) = E(\alpha)E(A_i) + E(1-\alpha)\left[E(\beta)E(A_i) + E(1-\beta)E(A_i B_i)\right]$$

$$= Q_1\mu_A + (1-Q_1)\left\{Q_2\mu_A + (1-Q_2)\mu_A\mu_B\right\} = \mu_A, \tag{1.2}$$

where α and β are Bernoulli random variables with means Q_1 and Q_2, and variances $Q_1(1-Q_1)$, $Q_2(1-Q_2)$ respectively. Ryu et al. (2005) have suggested an unbiased estimator of the mean μ_A as

$$\hat{\mu}_A = \frac{1}{n}\sum_{i=1}^{n} U_i . \tag{1.3}$$

The variance of $\hat{\mu}_A$ is given by

$$Var(\hat{\mu}_A) = \frac{1}{n}\left\{\sigma_A^2 + \left(\mu_A^2 + \sigma_A^2\right)(1-Q_1)(1-Q_2)\psi^2\right\}, \tag{1.4}$$

where σ_A^2 (may or may not be known) is the population variance of the sensitive variable under study.

2. PROPOSED ESTIMATION METHOD

Motivated by Searls (1964), we present a family of estimators to estimate the population mean μ_A of the sensitive quantitative variable as follows

$$\hat{\mu}_{A\lambda} = \lambda\hat{\mu}_A, \tag{2.1}$$

where $0 < \lambda \leq 1$ is some constant whose value is to be determined by the interviewer. It is shown by Searls (1964) that estimator given in (2.1) is biased and the amount of bias is given by

$$Bias\left(\hat{\mu}_{A\lambda}\right) = \left(\lambda - 1\right)\mu_A. \tag{2.2}$$

The mean squared error of the estimator $\hat{\mu}_{A\lambda}$ is given by

$$MSE\left(\hat{\mu}_{A\lambda}\right) = E\left(\hat{\mu}_{A\lambda} - \mu_A\right)^2 = \lambda^2\left\{\mu_A^2 + Var\left(\hat{\mu}_A\right)\right\} + \mu_A^2\left(1 - 2\lambda\right). \tag{2.3}$$

The proposed estimator $\hat{\mu}_{A\lambda}$ is more efficient than the estimator $\hat{\mu}_A$ if

$$MSE\left(\hat{\mu}_{A\lambda}\right) - Var\left(\hat{\mu}_A\right) \leq 0.$$

From (1.4) and (2.3), we can show easily that $MSE\left(\hat{\mu}_{A\lambda}\right) - Var\left(\hat{\mu}_A\right) \leq 0$ when

$$\frac{\mu_A^2 - Var\left(\hat{\mu}_A\right)}{\mu_A^2 + Var\left(\hat{\mu}_A\right)} < \lambda \leq 1. \tag{2.4}$$

Using different values of Q_1, Q_2, μ_A, ψ^2, σ_A^2 and n, we calculate the ranges of values of λ in which proposed estimator is more efficient than the Ryu et al. (2005) estimator. The ranges of λ for different values of other parameters are given in the Tables 1, 2, and 3 (see Appendix).

We compute the percent relative efficiency (*PRE*) of the proposed estimator $\hat{\mu}_{A\lambda}$ relative to Ryu et al. (2005) estimator $\hat{\mu}_A$ for different values of the parameters and selection probabilities as $PRE\left(\hat{\mu}_{A\lambda}\right) = \frac{Var\left(\hat{\mu}_A\right)}{MSE\left(\hat{\mu}_{A\lambda}\right)} \times 100$. The *PREs* are given in Tables 4, 5, and 6 (see Appendix).

From Tables 1, 2, and 3, we observed that:

i) For all values of λ, our proposed estimator $\hat{\mu}_{A\lambda}$ is more efficient than Ryu et al.'s (2005) estimator when sample size is moderately small or the true value of population mean is small.
ii) For larger samples, Ryu et al.'s (2005) estimator performs well for the smaller values of λ. This suggests us defining a stratified estimator, which we will discuss in the Section 4.

iii) The range of values of λ depends on the sample size and the true value of the population mean. As the population mean increases the range of λ squeezes towards 1 for the other fixed parametric values. Same is the case when sample size increases from moderate to large.

From Tables 4, 5, and 6, we observe that for the values of λ near to zero, the *PRE* decreases as Q_1 or Q_2 moves away from 0.5 but it becomes stable when λ is closer to one. Large reduction in the mean squared error of $\hat{\mu}_{A\lambda}$ can be gained when sample size is smaller, by setting Q_1 and Q_2 closer to zero or one. From the above discussion we can conclude that, when sample size is small or it is expensive to study a large sample, then proposed estimator has reasonable scope and is advantageous as compared to Ryu et al. (2005) estimator.

3. OPTIMUM ESTIMATORS AMONGST THE FAMILY OF ESTIMATORS $\hat{\mu}_{A\lambda}$

We can also find an optimum estimator in the family of estimators $\hat{\mu}_{A\lambda}$ by differentiating (2.3) with respect to λ and setting it equal to zero. By doing so, the optimum value of λ is given by

$$\lambda_{\boldsymbol{opt}} = \frac{\mu_A^2}{\mu_A^2 + Var\left(\hat{\mu}_A\right)} = \frac{\mu_A^2}{\mu_A^2 + \frac{1}{n}\left\{\sigma_A^2 + \left(\mu_A^2 + \sigma_A^2\right)\left(1-Q_1\right)\left(1-Q_2\right)\psi^2\right\}}. \tag{3.1}$$

Thus the optimum estimator is given by

$$\hat{\mu}_{A\lambda_{opt}} = \lambda_{\boldsymbol{opt}}\hat{\mu}_A = \frac{\mu_A^2\hat{\mu}_A}{\mu_A^2 + \frac{1}{n}\left\{\sigma_A^2 + \left(\mu_A^2 + \sigma_A^2\right)\left(1-Q_1\right)\left(1-Q_2\right)\psi^2\right\}}. \tag{3.2}$$

The *MSE* of $\hat{\mu}_{A\lambda_{opt}}$ is given by

$$MSE\left(\hat{\mu}_{A\lambda_{opt}}\right) = \frac{\mu_A^2 \frac{1}{n}\left\{\sigma_A^2 + \left(\mu_A^2 + \sigma_A^2\right)\left(1-Q_1\right)\left(1-Q_2\right)\psi^2\right\}}{\mu_A^2 + \frac{1}{n}\left\{\sigma_A^2 + \left(\mu_A^2 + \sigma_A^2\right)\left(1-Q_1\right)\left(1-Q_2\right)\psi^2\right\}}. \tag{3.3}$$

Now the relative efficiency (*RE*) of the optimum estimator relative to Ryu et al.'s (2005) estimator is given by

$$RE = \frac{Var\left(\hat{\mu}_A\right)}{MSE\left(\hat{\mu}_{A\lambda_{opt}}\right)}$$

$$RE = 1 + \frac{\frac{1}{n}\left\{\sigma_A^2 + \left(\mu_A^2 + \sigma_A^2\right)\left(1-Q_1\right)\left(1-Q_2\right)\psi^2\right\}}{\mu_A^2}. \tag{3.4}$$

From (3.4), it is obvious that the proposed optimum estimator is always more efficient than the Ryu et al. (2005) estimator in terms of the variability, but it has a limitation because optimum value of λ depends on the unknown population mean and variance. So its practicability is not easy and useful. Mehta and Srinivasan (1971) have discussed such problems by using the prior knowledge or experience about the unknown population parameters. Intelligent guesses about the unknown parameters may be made by using the past data or study and then about the constant λ. Let us suppose that $\tilde{\lambda}_{opt} = \theta\lambda_{opt}$, $(\theta > 0)$ be the guessed optimum value of the constant λ. When $\theta = 1$, the guessed value of optimum λ is exactly equal to the true optimum λ. As θ moves away from 1, our guessed value departs more from the true optimum value of λ. Using the approach of Singh and Shukla (2002), we define an estimator of the population mean μ_A based on the guessed optimum λ as

$$\hat{\mu}_{\tilde{\lambda}_{opt}} = \tilde{\lambda}_{opt}\hat{\mu}_A = \theta\lambda_{opt}\hat{\mu}_A \,. \tag{3.5}$$

The estimator in (3.5) is biased with the amount of bias

$$Bias\left(\hat{\mu}_{A\tilde{\lambda}_{opt}}\right) = \theta E\left(\lambda_{opt}\hat{\mu}_A\right) - \mu_A = \left\{\frac{\theta}{1 + C\left(\hat{\mu}_A\right)} - 1\right\}\mu_A \,. \tag{3.6}$$

The *MSE* of the estimator in (3.5) is given by

$$MSE\left(\hat{\mu}_{A\tilde{\lambda}_{opt}}\right) = \mu_A^2\left[1 + \frac{\theta^2}{\left\{1 + \left(C\left(\hat{\mu}_A\right)\right)^2\right\}} - \frac{2\theta}{\left\{1 + \left(C\left(\hat{\mu}_A\right)\right)^2\right\}}\right], \tag{3.7}$$

where $C\left(\hat{\mu}_A\right) = \dfrac{\sqrt{Var\left(\hat{\mu}_A\right)}}{\mu_A}$.

Using (1.4) and (3.7), we can show that the estimator given in (3.5) will be more efficient than Ryu et al.'s (2005) estimator if $MSE\left(\hat{\mu}_{A\tilde{\lambda}_{optimum}}\right) \le Var\left(\hat{\mu}_A\right)$. It is true only when

$$\left(\theta - 1\right)^2 \le \left(\frac{\sqrt{Var\left(\hat{\mu}_A\right)}}{\mu_A}\right)^4$$

or if

$$\left(\theta - 1\right)^2 \le \left(\frac{\sigma_A^2 + \left(\mu_A^2 + \sigma_A^2\right)\left(1 - Q_1\right)\left(1 - Q_2\right)\psi^2}{n\mu_A^2}\right)^2 . \tag{3.8}$$

For different values of the parameters and sample size, we calculate the range of values of θ, in which proposed estimator, based on guessed value of λ, is more efficient than the Ryu et al. (2005) estimator. These ranges are given in Tables 7, 8, and 9 (see Appendix). From these tables, we observed that the range of dominance θ squeezes to 1 when sample size or the true value of mean (μ_A) increases.

Singh and Mathur (2002, 2006) have discussed that if it is difficult to guess the value of λ or the unknown population parameters then these parameters can be replaced by their consistent estimates from sample. So (3.1) can be written as

$$\hat{\lambda}_{\boldsymbol{opt}} = \frac{\hat{\mu}_A^2}{\hat{\mu}_A^2 + \frac{1}{n} s_U^2}, \tag{3.9}$$

where $s_U^2 = \frac{\sum_{i=1}^{n}\left(U_i - \bar{U}\right)^2}{n-1}$ and U_i is defined earlier in (1.1).

Substituting (3.9) in (3.2), we have another estimator of μ_A as

$$\hat{\mu}_{A\hat{\lambda}_{opt}} = \frac{\hat{\mu}_A^3}{\hat{\mu}_A^2 + \frac{1}{n} s_U^2}. \tag{3.10}$$

To derive the bias and mean squared error of the $\hat{\mu}_{A\hat{\lambda}_{opt}}$, we proceed as follows. Define $e_0 = \frac{\hat{\mu}_A - \mu_A}{\mu_A}$ and $e_1 = \frac{s_U^2 - S_U^2}{S_U^2}$ such that $E(e_0) = E(e_1) = 0$. To first order of approximation, we have

$$E\left(e_0^2\right) = \frac{C_U^2}{n},\ E\left(e_1^2\right) = \frac{(\omega_4 - 1)}{n}, \text{ and } E(e_0 e_1) = \frac{C_U \omega_3}{n},$$

where $\omega_r = \frac{\tau_r}{\tau_2^{r/2}}$, $\tau_r = \frac{1}{N}\sum_{i=1}^{N}\left(U_i - \bar{U}\right)^r$ and $C_U = \frac{S_U}{\mu_A}$.

Now rewriting the estimator given in (3.10) in terms of e_0 and e_1, we have

$$\hat{\mu}_{A\hat{\lambda}_{opt}} = \frac{\mu_A(1+e_0)}{1 + \frac{1}{n}\frac{S_U^2}{\mu_A^2}(1+e_1)(1+e_0)^{-2}} = \frac{\mu_A(1+e_0)}{1 + \frac{1}{n}\frac{S_U^2}{\mu_A^2}(1+e_1)\left(1 - 2e_0 + 3e_0^2 + ...\right)}.$$

Neglecting the higher order terms, we get

$$\hat{\mu}_{A\hat{\lambda}_{opt}} = \mu_A(1+e_0)\{1+x\}^{-1}, \text{ where } x = \frac{1}{n} C_U^2 \left(1 - 2e_0 + 3e_0^2 + e_1 - 2e_0 e_1\right).$$

Assuming that $|x| < 1$, using the binomial expansion, we get

$$\hat{\mu}_{A\hat{\lambda}_{opt}} - \mu_A = \mu_A \left\{ e_0 - \frac{1}{n} C_U^2 \left(1 - e_0 + e_0^2 + e_1 - e_0 e_1\right) \right.$$
$$\left. + \frac{1}{n^2} C_U^4 \left(1 + e_1^2 - 3e_0 + 6e_0^2 + 2e_1 - 6e_0 e_1\right) \right\}. \qquad (3.11)$$

After applying the expectation on both sides of (3.11), we obtain

$$Bias\left(\hat{\mu}_{A\hat{\lambda}_{opt}}\right) = \mu_A \left[-\frac{1}{n^2} C_U^2 + \frac{C_U^4}{n^2} \{\omega_4 - 1\} + \frac{C_U^3 \omega_3}{n^2} - \frac{6C_U^5 \omega_3}{n^3} - \frac{6C_U^6}{n^3} \right]. \qquad (3.12)$$

Similarly, using (3.11), we have

$$MSE\left(\hat{\mu}_{A\hat{\lambda}'_{opt}}\right) = E\left(\hat{\mu}_{A\hat{\lambda}_{opt}} - \mu_A\right)^2 = \mu_A^2 \left[\frac{C_U^2}{n} - \frac{2C_U^3 \omega_3}{n^2} + \frac{1}{n} C_U^4 \left(\frac{\omega_4 - 1}{n} - 1 \right) \right.$$
$$\left. - \frac{2}{n^3} C_U^6 \left(\frac{3(\omega_4 - 1)}{n} - \frac{1}{2} \right) + \frac{24C_U^7 \omega_3}{n^4} + \frac{C_U^8}{n^3} \left(-\frac{19}{n} + \frac{21}{n^2} + \frac{6(\omega_4 - 1)}{n^2} - 6 \right) - \frac{24C_U^9 \omega_3}{n^5} \right].$$
$$(3.13)$$

Comparing (1.4) and (3.13), we see that it is very difficult to derive an exact expression for the efficiency condition. So we obtained the simulation results which are given in Table 10 (see Appendix) and are based on 5000 runs.

4. STRATIFIED PROPOSED ESTIMATOR

As we have mentioned in Section 2 that proposed family of estimators works well for smaller sample size and/or when it may be guessed that true value of the mean of the sensitive variable is smaller. This guides us to apply stratified sampling protocol when large samples could not be taken and it could be suspected that true mean of the sensitive study variable would be near to zero.

Let the population of size N can be divided into H strata of sizes $N_h, h = 1, 2, ..., H$. Let n_h be the size of the sample drawn from h^{th} stratum. Application of Ryu at al. (2005) model in stratified sampling with fixed total sample size n and optimum allocation of sample sizes in different strata yields the following mean estimator, as also suggested by Ryu at al. (2005)

$$\hat{\mu}_{RA} = \sum_{h=1}^{H} W_h \hat{\mu}_{R_h}, \qquad (4.1)$$

where $W_h = \dfrac{N_h}{N}$ is the h^{th} stratum weight, and $\hat{\mu}_{R_h}$ is the Ryu et al. (2005) estimator from stratum h.

The variance of $\hat{\mu}_{RA}$ is given by

$$Var\left(\hat{\mu}_{RA}\right)=\frac{1}{n}\left[\sum_{h=1}^{H}W_h\left\{\sigma_{A_h}^2+\left(\mu_{A_h}^2+\sigma_{A_h}^2\right)\left(1-Q_{1h}\right)\left(1-Q_{2h}\right)\psi_h^2\right\}^{1/2}\right]^2. \tag{4.2}$$

Now we define a proposed stratified estimator as

$$\left(\hat{\mu}_{A\lambda_{opt}}\right)_{St}=\sum_{h=1}^{H}W_h\hat{\mu}_{A(\lambda_{opt})_h}, \tag{4.3}$$

where $\hat{\mu}_{A(\hat{\lambda}_{opt})_h}=\dfrac{\hat{\mu}_{A_h}^3}{\hat{\mu}_{A_h}^2+\frac{1}{n}s_{U_h}^2}$ is the mean estimator based on estimated value of λ_h from the h^{th} stratum. Using the notations of above section for individual stratum and neglecting the higher order terms, the bias, to the first order of approximation, of the estimator in (4.3) is obtained as

$$\begin{aligned}Bias\left(\left(\hat{\mu}_{A\hat{\lambda}_{opt}}\right)_{St}\right)&=E\left(\left(\hat{\mu}_{A\hat{\lambda}_{opt}}\right)_{St}-\mu_A\right)\\&=\sum_{h=1}^{H}W_h\mu_{A_h}\left[-\frac{1}{n_h^2}C_{U_h}^2+\frac{C_{U_h}^4}{n_h^2}\left\{\left(\omega_4\right)_h-1\right\}+\frac{C_{U_h}^3\left(\omega_3\right)_h}{n_h^2}-\frac{6C_{U_h}^5\left(\omega_3\right)_h}{n_h^3}-\frac{6C_{U_h}^6}{n_h^3}\right]\end{aligned} \tag{4.4}$$

The mean squared error, to the first order of approximation, of the proposed stratified estimator is given by

$$\begin{aligned}MSE\left(\left(\hat{\mu}_{A\hat{\lambda}_{opt}}\right)_{St}\right)&=E\left(\left(\hat{\mu}_{A\hat{\lambda}_{opt}}\right)_{St}-\mu_A\right)^2\\&=\sum_{h=1}^{H}W_h\mu_{A_h}^2\left[\frac{C_{U_h}^2}{n_h}-\frac{2C_{U_h}^3\left(\omega_3\right)_h}{n_h^2}+\frac{1}{n_h}C_{U_h}^4\left(\frac{\left(\omega_4\right)_h-1}{n_h}-1\right)\right.\\&\quad-\frac{2}{n_h^3}C_{U_h}^6\left(\frac{3\left(\left(\omega_4\right)_h-1\right)}{n_h}-\frac{1}{2}\right)+\frac{24C_{U_h}^7\left(\omega_3\right)_h}{n_h^4}\\&\quad\left.+\frac{C_{U_h}^8}{n_h^3}\left(-\frac{19}{n_h}+\frac{21}{n_h^2}+\frac{6\left(\left(\omega_4\right)_h-1\right)}{n_h^2}-6\right)-\frac{24C_{U_h}^9\left(\omega_3\right)_h}{n_h^5}\right].\end{aligned} \tag{4.5}$$

On comparing (4.2) and (4.5) we see that it is difficult to obtain an exact expression for the efficiency condition of the proposed stratified estimator. Therefore, taking into consideration the case of two strata, we have worked out its *PRE* relative to Ryu et al. (2005) stratified estimator using the simulation approach and results are given in Tables 11 and 12 (see Appendix). We assume both the study and scrambling variables to be normally distributed with means 0.001, 1.0, and variances 1.0, 0.5, respectively. The results are based on 5000 iterations.

5. CONCLUSIONS

To sum up, we conclude that, in the situations where it is impossible/difficult to study the large samples and it is suspected that the actual mean of study variable may be small, proposed approach of estimation may be fruitfully used to get the precise estimates. To achieve the maximum gain in efficiency the selection probabilities Q_1 and Q_2 should preferably be chosen small. The mean and variance of scrambling variable should also be smaller.

APPENDIX

Table 1:
Range of values of λ for different values of selection probabilities when $n = 10, 11, ..., 100$, $\mu_A = 0.1$, $\sigma_A^2 = 1$, and $\psi^2 = 0.5$

	Q_2				
Q_1	0.1	0.3	0.5	0.7	0.9
0.1	0→1	0→1	0→1	0→1	0→1
0.3	0→1	0→1	0→1	0→1	0→1
0.5	0→1	0→1	0→1	0→1	0→1
0.7	0→1	0→1	0→1	0→1	0→1
0.9	0→1	0→1	0→1	0→1	0→1

Table 2:
Range of values of λ for different values of selection probabilities when $n = 10, 11, ..., 30$, $\mu_A = 0.2$, $\sigma_A^2 = 1$, and $\psi^2 = 0.5$

	Q_2				
Q_1	0.1	0.3	0.5	0.7	0.9
0.1	0→1	0→1	0→1	0→1	0→1
0.3	0→1	0→1	0→1	0→1	0→1
0.5	0→1	0→1	0→1	0→1	0→1
0.7	0→1	0→1	0→1	0→1	0→1
0.9	0→1	0→1	0→1	0→1	0→1

Table 3:
Range of values of λ for different values of selection probabilities when $n = 35$, $\mu_A = 0.2$, $\sigma_A^2 = 1$, and $\psi^2 = 0.5$

	Q_2				
Q_1	0.1	0.3	0.5	0.7	0.9
0.1	0→1	0.02→1	0.06→1	0.10→1	0.14→1
0.3	0.02→1	0.05→1	0.08→1	0.11→1	0.14→1
0.5	0.06→1	0.08→1	0.10→1	0.12→1	0.15→1
0.7	0.10→1	0.11→1	0.12→1	0.14→1	0.15→1
0.9	0.14→1	0.14→1	0.15→1	0.15→1	0.16→1

Table 4:
***PRE* of the $\hat{\mu}_{A\lambda}$ relative to $\hat{\mu}_A$ for different values λ and selection probabilities when $n = 10$, $\mu_A = 0.1$, $\sigma_A^2 = 1$, and $\psi^2 = 0.5$**

	$\lambda = 0.1$				
	Q_2				
Q_1	0.1	0.3	0.5	0.7	0.9
0.1	1103.84	1143.14	1133.34	1123.53	1113.70
0.3	1481.80	1399.58	1315.77	1230.30	1143.14
0.5	1399.58	1334.54	1268.49	1201.44	1133.34
0.7	1315.77	1268.49	1220.70	1172.38	1123.53
0.9	1230.30	1201.44	1172.38	1143.14	1113.70
	$\lambda = 0.3$				
0.1	1103.84	730.62	728.19	725.73	723.23
0.3	801.44	786.33	769.66	751.20	730.62
0.5	786.33	773.51	759.65	744.59	728.19
0.7	769.66	759.65	749.02	737.74	725.73
0.9	751.20	744.59	737.74	730.62	723.23
	$\lambda = 0.5$				
0.1	720.70	365.07	364.76	364.45	364.13
0.3	373.49	371.79	369.86	367.64	365.07
0.5	371.79	370.31	368.67	366.83	364.76
0.7	369.86	368.67	367.38	365.97	364.45
0.9	367.64	366.83	365.97	365.07	364.13
	$\lambda = 0.7$				
0.1	363.80	200.55	200.52	200.49	200.45
0.3	201.45	201.27	201.07	200.83	200.55
0.5	201.27	201.12	200.94	200.74	200.52
0.7	201.07	200.94	200.80	200.65	200.49
0.9	200.83	200.74	200.65	200.55	200.45
	$\lambda = 0.9$				
0.1	200.41	123.31	123.31	123.30	123.30
0.3	123.34	123.34	123.33	123.32	123.31
0.5	123.34	123.33	123.32	123.31	123.31
0.7	123.33	123.32	123.32	123.31	123.30
0.9	123.32	123.31	123.31	123.31	123.30

Table 5:
PRE of the $\hat{\mu}_{A\lambda}$ relative to $\hat{\mu}_A$ for different values λ and selection probabilities when $n=10$, $\mu_A = 0.2,\ \sigma_A^2 = 1,$ and $\psi^2 = 0.5$

$\lambda = 0.1$					
	Q_2				
Q_1	0.1	0.3	0.5	0.7	0.9
0.1	123.30	312.97	309.96	306.94	303.93
0.3	420.21	393.62	366.89	340.00	312.97
0.5	393.62	372.84	351.97	331.01	309.96
0.7	366.89	351.97	337.01	322.00	306.94
0.9	340.00	331.01	322.00	312.97	303.93
$\lambda = 0.3$					
0.1	300.91	360.70	358.27	355.83	353.37
0.3	438.76	420.81	401.87	381.87	360.70
0.5	420.81	406.17	390.89	374.94	358.27
0.7	401.87	390.89	379.57	367.89	355.83
0.9	381.87	374.94	367.89	360.70	353.37
$\lambda = 0.5$					
0.1	350.89	289.41	288.61	287.79	286.97
0.3	312.14	307.38	302.08	296.13	289.41
0.5	307.38	303.31	298.86	293.98	288.61
0.7	302.08	298.86	295.42	291.74	287.79
0.9	296.13	293.98	291.74	289.41	286.97
$\lambda = 0.7$					
0.1	286.13	190.69	190.57	190.44	190.31
0.3	194.05	193.38	192.61	191.73	190.69
0.5	193.38	192.79	192.13	191.40	190.57
0.7	192.61	192.13	191.62	191.06	190.44
0.9	191.73	191.40	191.06	190.69	190.31
$\lambda = 0.9$					
0.1	190.18	122.87	122.87	122.86	122.85
0.3	123.02	122.99	122.96	122.92	122.87
0.5	122.99	122.97	122.94	122.91	122.87
0.7	122.96	122.94	122.92	122.89	122.86
0.9	122.92	122.91	122.89	122.87	122.85

Table 6:
PRE of the $\hat{\mu}_{A\lambda}$ relative to $\hat{\mu}_A$ for different values λ and selection probabilities when $n = 10$, $\mu_A = 0.3$, $\sigma_A^2 = 1$, and $\psi^2 = 0.5$

$\lambda = 0.1$					
	Q_2				
Q_1	0.1	0.3	0.5	0.7	0.9
0.1	119.78	141.86	140.40	138.95	137.50
0.3	193.89	180.93	167.94	154.92	141.86
0.5	180.93	170.83	160.71	150.57	140.40
0.7	167.94	160.71	153.47	146.21	138.95
0.9	154.92	150.57	146.21	141.86	137.50
$\lambda = 0.3$					
0.1	136.04	195.93	194.25	192.56	190.87
0.3	252.56	239.07	225.15	210.77	195.93
0.5	239.07	228.28	217.22	205.88	194.25
0.7	225.15	217.22	209.15	200.93	192.56
0.9	210.77	205.88	200.93	195.93	190.87
$\lambda = 0.5$					
0.1	189.17	215.29	214.25	213.20	212.14
0.3	246.24	239.52	232.18	224.14	215.29
0.5	239.52	233.87	227.81	221.29	214.25
0.7	232.18	227.81	223.20	218.34	213.20
0.9	224.14	221.29	218.34	215.29	212.14
$\lambda = 0.7$					
0.1	211.06	176.30	176.04	175.79	175.53
0.3	183.08	181.72	180.16	178.37	176.30
0.5	181.72	180.52	179.20	177.72	176.04
0.7	180.16	179.20	178.16	177.02	175.79
0.9	178.37	177.72	177.02	176.30	175.53
$\lambda = 0.9$					
0.1	175.26	122.16	122.14	122.13	122.12
0.3	122.51	122.44	122.36	122.27	122.16
0.5	122.44	122.38	122.31	122.23	122.14
0.7	122.36	122.31	122.26	122.20	122.13
0.9	122.27	122.23	122.20	122.16	122.12

Table 7:
Range of suitable values of θ for different values of selection probabilities when $n = 10, 11, ..., 100$, $\mu_A = 0.1$, $\sigma_A^2 = 1$, and $\psi^2 = 0.5$

	Q_2				
Q_1	0.1	0.3	0.5	0.7	0.9
0.1	0→1	0→1	0→1	0→1	0→1
0.3	0→1	0→1	0→1	0→1	0→1
0.5	0→1	0→1	0→1	0→1	0→1
0.7	0→1	0→1	0→1	0→1	0→1
0.9	0→1	0→1	0→1	0→1	0→1

Table 8:
Range of suitable values of θ for different values of selection probabilities when $n = 10, 11, ..., 25$, $\mu_A = 0.2$, $\sigma_A^2 = 1$, and $\psi^2 = 0.5$

	Q_2				
Q_1	0.1	0.3	0.5	0.7	0.9
0.1	0→1	0→1	0→1	0→1	0→1
0.3	0→1	0→1	0→1	0→1	0→1
0.5	0→1	0→1	0→1	0→1	0→1
0.7	0→1	0→1	0→1	0→1	0→1
0.9	0→1	0→1	0→1	0→1	0→1

Table 9:
Range of suitable values of θ for different values of selection probabilities when $n = 15$, $\mu_A = 0.3$, $\sigma_A^2 = 1$, and $\psi^2 = 0.5$

	Q_2				
Q_1	0.1	0.3	0.5	0.7	0.9
0.1	0→1	0→1	0.08→1	0.11→1	0.16→1
0.3	0→1	0.08→1	0.12→1	0.12→1	0.18→1
0.5	0.08→1	0.12→1	0.16→1	0.13→1	0.20→1
0.7	0.16→1	0.18→1	0.20→1	0.15→1	0.23→1
0.9	0.23→1	0.24→1	0.24→1	0.16→1	0.25→1

Table 10:
Simulated *RE* of the estimator $\hat{\mu}_{A\hat{\lambda}_{opt}}$ relative to $\hat{\mu}_A$ for different values of selection probabilities when $\sigma_A^2 = 1$, $\psi^2 = 0.5$ and $n = 10$

$\mu_A = 0.1$					
	Q_2				
Q_1	0.1	0.3	0.5	0.7	0.9
0.1	194.20	194.60	189.17	189.95	191.42
0.3	191.61	192.42	189.85	189.80	191.36
0.5	190.97	192.50	187.69	191.33	190.85
0.7	190.13	190.38	189.08	190.57	189.59
0.9	189.19	189.79	190.98	190.65	188.99
$\mu_A = 0.2$					
0.1	173.19	160.38	162.27	160.93	160.22
0.3	167.91	160.79	162.35	160.90	158.61
0.5	164.88	161.04	162.06	160.87	158.61
0.7	162.52	162.65	161.48	160.01	159.75
0.9	161.93	163.15	160.59	159.77	161.28
$\mu_A = 0.3$					
0.1	124.20	133.52	131.97	132.58	129.17
0.3	129.39	133.63	131.92	130.94	129.59
0.5	132.99	133.05	132.13	130.66	129.13
0.7	134.45	132.99	131.63	130.79	129.44
0.9	134.96	132.56	131.74	130.37	129.98
$\mu_A = 0.4$					
0.1	103.15	112.14	111.48	110.03	110.45
0.3	107.17	111.58	111.35	110.56	109.78
0.5	111.05	110.98	111.39	110.37	109.34
0.7	112.55	111.24	111.10	110.92	108.26
0.9	113.93	110.99	110.91	111.39	107.76
$\mu_A = 0.5$					
0.1	86.82	97.46	96.81	96.38	94.08
0.3	91.10	96.91	95.95	96.83	94.33
0.5	94.55	96.75	95.80	96.53	93.76
0.7	96.95	97.00	96.77	95.09	94.23
0.9	97.71	97.75	96.78	94.57	95.42

Table 11:
Simulated bias and *PRE* of the estimator $\left(\hat{\mu}_{A\hat{\lambda}_{opt}}\right)_{St}$ relative to $\hat{\mu}_{RA}$ when
$\mu_{A1}=0.5, \mu_{A2}=0.1, \sigma^2_{A1}=\sigma^2_{A2}=1, n_1=n_2=10, W_1=0.3, W_2=0.7$ **, and** $\psi^2_2=\psi^2_1=0.5$

	$Q_{21}=Q_{22}$									
$Q_{11}=Q_{12}$	0.1		0.3		0.5		0.7		0.9	
	$\lvert Bias \rvert$	PRE	$\lvert Bias \rvert$	PRE	$\lvert Bias \rvert$	PRE	$\lvert Bias \rvert$	PRE	$\lvert Bias \rvert$	PRE
0.1	0.05	238.8	0.04	223.7	0.05	230.1	0.05	246.6	0.05	232.9
0.3	0.06	256.0	0.05	230.6	0.05	232.1	0.05	243.2	0.05	230.0
0.5	0.05	249.3	0.05	233.2	0.05	233.7	0.05	233.6	0.05	238.5
0.7	0.06	251.0	0.05	235.0	0.05	231.5	0.05	238.5	0.05	234.0
0.9	0.05	246.7	0.05	232.7	0.05	238.0	0.05	233.0	0.05	234.8

Table 12:
Simulated bias and *PRE* of proposed stratified estimator when
$\mu_{A1}=0.3, \mu_{A2}=0.2, \sigma^2_{A1}=2, \sigma^2_{A2}=3, n_1=n_2=10, W_1=0.6, W_2=0.4$ **, and** $\psi^2_2=\psi^2_1=1$

	$Q_{21}=Q_{22}$									
$Q_{11}=Q_{12}$	0.1		0.3		0.5		0.7		0.9	
	$\lvert Bias \rvert$	PRE	$\lvert Bias \rvert$	PRE	$\lvert Bias \rvert$	PRE	$\lvert Bias \rvert$	PRE	$\lvert Bias \rvert$	PRE
0.1	0.09	219.7	0.08	194.6	0.09	192.2	0.08	188.9	0.08	184.5
0.3	0.09	211.3	0.08	193.7	0.08	190.5	0.08	188.6	0.08	182.7
0.5	0.09	202.6	0.08	195.1	0.09	189.7	0.08	186.3	0.07	183.0
0.7	0.09	198.8	0.08	194.9	0.09	190.4	0.08	184.2	0.08	183.2
0.9	0.08	195.4	0.08	194.9	0.09	190.1	0.08	183.9	0.08	183.8

REFERENCES

1. Arnab, R. and Dorffner, G. (2006). Randomized response technique for complex survey designs. *Statist. Papers*, 48, 131-141.
2. Bar-Lev, S.K., Bobovitch. E. and Boukai, B. (2004). A note on randomized response models. *Metrika*, 60, 255-260.
3. Eichhorn, B.H. and Hayre, L.S. (1983). Scrambled randomized response methods for obtaining sensitive quantitative data. *J. Statist. Plann. and Infer.*, 7, 307-316.
4. Espejo Ruiz, M. and Singh, H. P. (2003). Protection of privacy with objective prior distribution in randomized response. *Statistica.*, Anno LXIII, 4, 697-701,
5. Greenberg, B.G., Kuebler, R.R. Jr., Abernathy, J.R. and Hovertiz, D.G. (1971). Application of the randomized response techniques in obtaining quantitative data. *J. Amer. Statist. Assoc.*, 66, 243-250.
6. Gupta S., Gupta, B. and Singh, S. (2002). Estimation of Sensitivity level of personal interview survey questions. *J. Statist. Plann. and Infer.*, 100, 239-247.
7. Gupta S. and Shabbir J. (2004). Sensitivity estimation for personal interview survey question. *Statistica.*, Anno LXIV, 4, 643-653.

8. Hussain, Z., Shabbir, J. and Gupta, S. (2007). An alternative to Ryu randomized response model. *J. Statist. and Manag. Systems*, 10(4), 511-517.
9. Hussain, Z. and Shabbir, J. (2007). On estimation of mean of a sensitive quantitative variable. *Inter. Stat.*, July # 6.
10. Mangat, N.S. and Singh, R. (1990). An alternative randomized response procedure. *Biometrika*, 77, 439-442.
11. Mehta, J.S. and Srinivasan, R. (1971). Estimation of mean by shrinkage to a point. *J. Amer. Statist. Assoc.*, 66, 86-90.
12. Ryu, J.-B., Kim, J.-M., Heo, T-Y., and Park, C.G. (2005). On stratified Randomized response sampling. *Model Assisted Statist. App.*, 1(1), 31-36.
13. Searls, D.T. (1964). The utilization of a known coefficient of variation in the estimation procedure. *J. Amer. Statist. Assoc.*, 59, 1225-1226.
14. Singh, H.P. and Shukla, S.K. (2002). A class of shrinkage estimators for the variance of exponential distribution with type-I censoring. *IAPQR Transitions*, 27, 119-141.
15. Singh, H.P. and Mathur, N. (2002a). A revisit to alternative estimator for randomized response technique. *J. Ind. Soc. Agri. Statist.*, 55(1), 79-87.
16. Singh, H.P. and Mathur, N. (2002b). An alternative to an improved randomized response strategy. *Statistics in Transition*, 5(5), 873-886.
17. Singh, H.P. and Mathur, N. (2003a). An optimally randomized response technique. *Aligarh J. Statist.*, 23, 1-5.
18. Singh, H.P. and Mathur, N. (2003b). Modified optimal randomized response sampling technique. *J. Ind. Soc. Agri. Statist.*, 56(2), 199-206.
19. Singh, H.P. and Mathur, N. (2004a). Unknown repeated trials in the unrelated question randomized response model. *Biometrical J.*, 46(3), 375-378.
20. Singh, H.P. and Mathur, N. (2004b). Estimation of population mean with known coefficient of variation under optimal randomized response model using scrambled response technique. *Statistics in Transition*, 6(7), 1079-1093.
21. Singh, H.P. and Mathur, N. (2004c). Improved estimation of population proportion possessing sensitive attribute with unknown repeated trial in randomized response sampling. *Statistica.*, Anno LXIV, 3, 537-544.
22. Singh, H.P. and Mathur, N. (2005). Estimation of population mean when coefficient of variation is known using scrambled randomized response technique. *J. Statist. Plann. and Infer.*, 131(1), 135-144.
23. Singh, H.P. and Mathur, N. (2006). An improved estimation procedure for estimating the proportion of a population possessing sensitive attribute in unrelated question randomized response technique. *Brazilian J. Prob. Statist.*, 20, 93-110.
24. Singh, H.P. and Mathur, N. (2007). An improved estimator for proportion of sensitive group of population using optional randomized response technique. *Assam Statist. Review*, 21(1), 64-72.
25. Warner, S.L. (1965). Randomized response: a survey technique for eliminating evasive answer bias. *J. Amer. Statist. Assoc.*, 60, 63-69.