Communications for Statistical Applications and Methods
2019, Vol. 26, No. 1, 11–25
https://doi.org/10.29220/CSAM.2019.26.1.011
Print ISSN 2287-7843 / Online ISSN 2383-4757

On efficient estimation of population mean under non-response

Shashi Bhushan[a], Abhay Pratap Pandey[1,b]

[a]Department of Mathematics and Statistics, D. S. M. N. R. University, Lucknow, India;
[b]Ramanujan College, University of Delhi, India

Abstract

The present paper utilizes auxiliary information to neutralize the effect of non-response for estimating the population mean. Improved ratio type estimators for population mean have been proposed and their properties are studied. These estimators are suggested for both single phase sampling and two phase sampling in presence of non-response. Empirical studies are conducted to validate the theoretical results and demonstrate the performance of the proposed estimators. The proposed estimators are shown to perform better than those used by Cochran (*Sampling Techniques* (3rd ed), John Wiley & Sons, 1977), Khare and Srivastava (In *Proceedings-National Academy Science, India, Section A*, **65**, 195–203, 1995), Rao (*Randomization Approach in Incomplete Data in Sample Surveys*, Academic Press, 1983; *Survey Methodology* **12**, 217–230, 1986), and Singh and Kumar (*Australian & New Zealand Journal of Statistics*, **50**, 395–408, 2008; *Statistical Papers*, **51**, 559–582, 2010) under the derived optimality condition. Suitable recommendations are put forward for survey practitioners.

Keywords: auxiliary variable, non-response, ratio type estimator, mean square error, efficiency

1. Introduction

The problem of non-response is inevitable in most sample surveys because the information cannot be obtained from all units selected in the survey due to various reasons. An estimator based on such incomplete information is biased and the final outcome may be misleading, when the respondents differ from non-respondents. In their seminal paper Hansen and Hurwitz (1946) considered a technique of sub-sampling the non-respondents in order to adjust for the non-response bias in a mail survey.

In sampling theory, it is well known that the efficiency of the estimators of unknown population parameters of the study variable can be increased by suitably using known information on an auxiliary variable. The ratio, product and regression methods of estimation are good examples in this context. Non-response adversely affects the estimate of population mean and population variance; in addition, many authors have suggested a number of estimators to estimate population parameter and their variance under the non-response for various situations. Cochran (1977) and Rao (1986) suggested a ratio method to estimate the population mean $\bar{Y}$ of the study variate y with sub-sampling from non-respondents.

Khare and Srivastava (1995) suggested an estimation procedure of the population mean using an auxiliary character in the presence of non-response, Khare and Srivastava (1995) proposed the studying of a conventional and alternative two phase sampling ratio, product and regression estimators in

[1] Corresponding author: Abhay Pratap Pandey, Ramanujan College, University of Delhi, Kalkaji New Delhi-110019, India. E-mail: abhaypratap.pandey@gmail.com

Published 31 January 2019 / journal homepage: http://csam.or.kr

the presence of non-response. Khare and Srivastava (1997) proposed transformed ratio type estimators for the population mean in the presence of non-response. Okafor and Lee (2000) proposed a double sampling scheme for ratio and regression estimation with sub sampling; in addition, the non-respondent also deal with the non-response problem. Khare and Srivastava (1993). Singh and Kumar (2008) proposed a general class of population mean estimators in survey sampling using auxiliary information with sub sampling for the non-respondent. Singh *et al.* (2010) also suggested a number of estimators to estimate population mean under non-response. Khare and Kumar (2011) proposed a method to estimate the population mean using known coefficient of variation of the study character in the presence of non-response. Singh and Bhushan (2012) proposed a generalized classes of two phase sampling estimators of population mean in presence of non-response. Shabbir and Khan (2013) also suggested a number of estimators to estimate the finite population mean using two auxiliary variables in two phase sampling in the presence of non-response. Sunil Kumar (Kumar, 2015) suggested an efficient use of auxiliary information in estimating the population ratio, product and mean in the presence of non-response.

An interesting finding of all these papers was that the regression (difference) estimators were found to be best in terms of mean squared error (MSE); in addition, any ratio type estimator can at best attain the MSE of these regression (difference) estimators. In this paper, we have proposed some improvement over regression as well as ratio estimators proposed by various authors in their earlier works.

2. Notations and existing results

Hansen and Hurwitz (1946) considered mail surveys in the first attempt-, and personal interviews in the second attempt. In the Hansen & Hurwitz method, the population of size of N is supposed to be composed of two strata, namely respondents and non-respondents; having sizes $N_1$and N_2 $(= N - N_1)$. Thus we label the data as $y_1, \ldots, y_{N_1}$ for the response group, and as $y_{N_1+1}, \ldots, y_{N_1+N_2}$ for the non-response group. Let $\bar{Y} = \sum_{i=1}^{N} y_i/N$ and $S_y^2 = \sum_{i=1}^{N}(y_i - \bar{Y})^2/(N-1)$ denote the population mean and variance, respectively. Let $\bar{Y}_1 = \sum_{i=1}^{N_1} y_i/N_1$ and $S_{y_1}^2 = \sum_{i=1}^{N_1}(y_i - \bar{Y}_1)^2/(N_1-1)$ denote the mean and variance of the response group, respectively. Similarly, let $\bar{Y}_2 = \sum_{i=N_1}^{N_1+N_2} y_i/N_2$ and $S_{y_2}^2 = \sum_{i=N_1+1}^{N_1+N_2}(y_i - \bar{Y}_2)^2/(N_2-1)$ denote the mean and variance of the non-response group, respectively. The population mean can be written as $\bar{Y} = W_1\bar{Y}_1 + W_2\bar{Y}_2$, $W_1 = N_1/N$, and $W_2 = N_2/N$. Let us consider a random sample of size n are drawn by using simple random sampling without replacement (SRSWOR). The random sample n should be made of two strata, namely n_1 respondents and $(n - n_1)$ non-respondents. The sample mean $\bar{y}_1 = \sum_{i=1}^{n_1} y_i/n_1$ is unbiased for $\bar{Y}_1$, but has a bias equal to $W_1(\bar{Y}_1 - \bar{Y}_2)$ in estimating the population mean $\bar{Y}$. The sample mean $\bar{y}_{2r} = \sum_{i=1}^{r} y_i/r$ is unbiased for the mean $\bar{y}_2$ of the n_2 units. An unbiased estimator for the population mean $\bar{Y}$ is

$$\bar{y}^* = w_1\bar{y}_1 + w_2\bar{y}_2,$$

where $w_1 = n_1/n$ and $w_2 = n_2/n$. The variance of $\bar{y}^*$ is given by

$$\mathrm{Var}(\bar{y}^*) = PS_y^2 + QS_{y_2}^2,$$

where $P = (1-f)/n$, $Q = W_2(k-1)/n$, and $f = n/N$.

Let $x_i\ (i = 1, 2, \ldots, N)$ denote an auxiliary variate correlated with study variate $y_i\ (i = 1, 2, \ldots, N)$. The population mean of the auxiliary variable x is $\bar{X} = \sum_{i=1}^{N} x_i/N$. Let $\bar{X}_1$ and $\bar{X}_2$ denote the means of the response and non-response groups. Let $\bar{x}$ denote the mean of all the n units. Let $\bar{x}_1$ and $\bar{x}_2$ denote the means of the n_1 responding units and the n_2 non-responding units. Further let $\bar{x}_{2r} = \sum_{i=1}^{r} x_i/r$ denote the mean of the subsampled units. The population variances of x and y are denoted by S_x^2 and S_y^2, and the population covariance by S_{xy}. The population correlation coefficient is $\rho = S_{xy}/S_x^2\, S_y^2$. The unbiased estimator of the population mean $\bar{X}$ of the auxiliary variable x is

$$\bar{x}^* = w_1\bar{x}_1 + w_2\bar{x}_2.$$

The variance of $\bar{x}^*$ is given by

$$\mathrm{Var}\,(\bar{x}^*) = PS_x^2 + QS_{x_2}^2,$$

where $S_{x_2}^2 = \sum_{i=N_1+1}^{N_1+N_2}(x_i - \bar{X}_2)^2/(N_2 - 1)$.

Similarly, In two phase sampling, we have n' observations on x from the first-phase sample, n_1 observations on y from the responding units of the n second-phase sample units, and r observations on y from the subsample units selected from the n_2 non response units of the second-phase sample. Let $\bar{x}'$ be the sample mean of auxiliary variable x based on the first-phase sample. Using the information on the auxiliary variable x collected from the first-phase sample, Cochran (1977), Khare and Srivastava (1995), Rao (1986), Okafor and Lee (2000), and Singh and Kumar (2008, 2010) suggested ratio and regression type estimators under non-response. We have classified these estimators into seven different strategies depending upon the available auxiliary information under both single phase sampling and two phase sampling.

Strategy I: When $\bar{y}^*$, $\bar{x}^*$, and $\bar{X}$ are used. If the auxiliary variable is known, the non-response occurs on the study variable y and information on the auxiliary variable x is not available from all the sample units along the population mean $\bar{X}$. The ratio and regression type estimators are

$$\begin{aligned} t_{r_1} &= \bar{y}^*\left(\frac{\bar{X}}{\bar{x}^*}\right), \\ t_{R_1} &= \bar{y}^*\left(\frac{\bar{X}}{\bar{x}^*}\right)^{\beta_1}, \\ t_1 &= \bar{y}^* + b^*\left(\bar{X} - \bar{x}^*\right), \end{aligned}$$

where $b^* = s_{xy}^*/s_x^{*^2}$ and the estimates s_{xy}^* and $s_x^{*^2}$ are based on the available data. To the first order of approximation, the MSEs of estimators t_{r_1}, t_{R_1}, and t_1 are given by

$$\begin{aligned} \mathrm{MSE}\,(t_{r_1}) &= P\left(S_y^2 + R^2S_x^2 - 2RS_{xy}\right) + Q\left(S_{y_2}^2 + R^2S_{x_2}^2 - 2RS_{xy_2}\right), \\ \mathrm{min.MSE}\,(t_{R_1}) &= PS_y^2 + QS_{y_2}^2 - \left\{\frac{\left(PS_{xy} + QS_{xy_2}\right)^2}{PS_y^2 + QS_{y_2}^2}\right\}, \\ \mathrm{min.MSE}\,(t_1) &= PS_y^2\left(1-\rho^2\right) + Q\left(S_{y_2}^2 + \beta^2S_{x_2}^2 - 2\beta S_{xy_2}\right), \end{aligned}$$

where $R = \bar{Y}/\bar{X}, \beta = S_{xy}/S_x^2$, and $\beta_{1\mathrm{opt}} = (PS_{xy} + QS_{xy_2})/(PS_y^2 + QS_{y_2}^2)$.

Strategy II: When $\bar{y}^*$, $\bar{x}$, and $\bar{X}$ are used. The non-response occurs on the study variable y, and information on the auxiliary variable x is available from all the sample units along the population mean $\bar{X}$ of the auxiliary variable is known. The ratio and regression type estimators are

$$t_{r_2} = \bar{y}^* \left(\frac{\bar{X}}{\bar{x}}\right),$$
$$t_{R_2} = \bar{y}^* \left(\frac{\bar{X}}{\bar{x}}\right)^{\beta_2},$$
$$t_2 = \bar{y}^* + b\left(\bar{X} - \bar{x}\right),$$

where $b = s^*_{xy}/s^2_x$. The first order approximate MSEs of the estimators t_{r_2}, t_{R_2}, and t_2 are given by

$$\text{MSE}\left(t_{r_3}\right) = P\left(S^2_y + R^2 S^2_x - 2RS_{xy}\right) + QS^2_{y_2}, \tag{2.1}$$
$$\text{min.MSE}\left(t_{R_2}\right) = \text{min.MSE}\left(t_2\right) = PS^2_y\left(1 - \rho^2\right) + QS^2_{y_2}, \tag{2.2}$$

where $\beta_2 = S_{xy_2}/S^2_{x_2}$ and $\beta_{2\text{opt}} = S_{xy}/S^2_x$.

Strategy III: When $\bar{y}^*$, $\bar{x}^*$, and $\bar{x}$ are used. The non-response occurs on the study variable y, and the information on the auxiliary variable x is obtained from all the sample units, but the population mean $\bar{X}$ of the auxiliary variable is not known. The ratio and regression type estimators are

$$t_{r_3} = \bar{y}^* \left(\frac{\bar{x}}{\bar{x}^*}\right),$$
$$t_{R_3} = \bar{y}^* \left(\frac{\bar{x}}{\bar{x}^*}\right)^{\beta_3},$$
$$t_3 = \bar{y}^* + b_{(2r)}\left(\bar{x} - \bar{x}^*\right),$$

where $b_{(2r)} = s_{xy_{(2r)}}/s^2_{x_{(2r)}}$. The MSEs to the first order of approximation, of the estimators t_{r_3}, t_{R_3}, and t_3 are given by

$$\text{MSE}\left(t_{r_3}\right) = PS^2_y + Q\left(S^2_{y_2} + R^2 S^2_{x_2} - 2RS_{xy_2}\right), \tag{2.3}$$
$$\text{min.MSE}\left(t_{R_3}\right) = \text{min.MSE}\left(t_3\right) = PS^2_y + QS^2_{y_2}\left(1 - \rho^2_2\right), \tag{2.4}$$

where $\beta_{3\text{opt}} = S_{xy_2}/S^2_{x_2}$.

Strategy IV: When $\bar{y}^*$, $\bar{x}$, $\bar{x}^*$, and $\bar{X}$ are used. Singh and Kumar (2008) suggested the estimators given below

$$t_{r_4} = \bar{y}^* \left(\frac{\bar{X}}{\bar{x}^*}\right)\left(\frac{\bar{X}}{\bar{x}}\right),$$
$$t_{R_4} = \bar{y}^* \left(\frac{\bar{X}}{\bar{x}^*}\right)^{\beta_4}\left(\frac{\bar{X}}{\bar{x}}\right)^{\beta_5},$$
$$t_4 = \bar{y}^* + d_1\left(\bar{x} - \bar{x}^*\right) + d_2\left(\bar{X} - \bar{x}\right).$$

To the first order of approximation, the MSEs of the estimators t_{r_4}, t_{R_4}, and t_4 are given by

$$\text{MSE}\left(t_{r_4}\right) = P\left\{S^2_y + 4RS^2_x\left(R - \beta\right)\right\} + Q\left\{S^2_{y_2} + RS^2_{x_2}\left(R - 2\beta_{(2)}\right)\right\},$$
$$\text{min.MSE}\left(t_4\right) = \text{min.MSE}\left(t_{R_4}\right) = PS^2_y\left(1 - \rho^2\right) + QS^2_{y_2}\left(1 - \rho^2_2\right),$$

where optimum values of β_4, β_5, d_1, and d_2 are given by $\beta_{4\text{opt}} = \beta_{(2)}/R, \beta_{5\text{opt}} = \beta - \beta_{(2)}/R, d_{1\text{opt}} = \beta_{(2)}$, and $d_{2\text{opt}} = \beta$. Now, we shall consider three more strategies under the two-phase sampling scheme proposed by Okafor and Lee (2000) and Singh and Kumar (2010).

Strategy V: When $\bar{y}^*$, $\bar{x}^*$, and $\bar{x}'$ are used. Okafor and Lee (2000) proposed a double sampling scheme for ratio estimation for sub sampling the non-respondent that also deals with the non-response problem.

$$t_{r_5} = \bar{y}^* \left(\frac{\bar{x}'}{\bar{x}^*}\right)$$
$$t_{R_5} = \bar{y}^* \left(\frac{\bar{x}'}{\bar{x}^*}\right)^{\beta_6}$$
$$t_5 = \bar{y}^* + b^* \left(\bar{x}' - \bar{x}^*\right)$$

To the first order of approximation, the MSEs of the estimators t_{r_5}, t_{R_5}, and t_5 are given by

$$\text{MSE}\left(t_{r_5}\right) = TS_y^2 + S\left(S_y^2 + R^2 S_x^2 - 2RS_{xy}\right) + Q\left(S_{y_2}^2 + R^2 S_{x_2}^2 - 2RS_{xy_2}\right),$$
$$\text{min.MSE}\left(t_{R_5}\right) = \text{min.MSE}\left(t_5\right) = TS_y^2 + SS_y^2\left(1 - \rho^2\right) + Q\left\{S_{y_2}^2 + \beta S_{x_2}^2\left(\beta - 2\beta_{(2)}\right)\right\},$$

where $\beta_{6\text{opt}} = \{(P - T)S_{xy} + QS_{xy_2}\}/(PS_x^2 + QS_{x_2}^2)$, $S = (1/n - 1/n')$, and $T = (1/n' - 1/N)$.

Strategy VI: When $\bar{y}^*$, $\bar{x}$, and $\bar{x}'$ are used. Khare and Srivastava (1995) and Okafor and Lee (2000) proposed a double sampling scheme for ratio estimation with sub sampling the non-respondent that also deals with the non-response problem.

$$t_{r_6} = \bar{y}^* \left(\frac{\bar{x}'}{\bar{x}}\right),$$
$$t_{R_6} = \bar{y}^* \left(\frac{\bar{x}'}{\bar{x}}\right)^{\beta_7},$$
$$t_6 = \bar{y}^* + b^{**}\left(\bar{x}' - \bar{x}\right).$$

To the first order of approximation, the MSE's of the estimators t_{r_6}, t_{R_6}, and t_6 are given by

$$\text{MSE}\left(t_{r_6}\right) = TS_y^2 + S\left(S_y^2 + R^2 S_x^2 - 2RS_{xy}\right) + QS_{y_2}^2,$$
$$\text{min.MSE}\left(t_{R_6}\right) = \text{min.MSE}\left(t_6\right) = TS_y^2 + SS_y^2\left(1 - \rho^2\right) + QS_{y_2}^2,$$

where $\beta_{7\text{opt}} = S_{xy}/S_x^2$.

Strategy VII: When $\bar{y}^*$, $\bar{x}^*$, $\bar{x}$, and $\bar{x}'$ are used. Singh and Kumar (2008) gave estimators for the population mean $\bar{Y}$ by using a double sampling scheme under non-response, which are mentioned as:

$$t_{r_7} = \bar{y}^* \left(\frac{\bar{x}'}{\bar{x}^*}\right)\left(\frac{\bar{x}'}{\bar{x}}\right),$$
$$t_{R_7} = \bar{y}^* \left(\frac{\bar{x}'}{\bar{x}^*}\right)^{\beta_8}\left(\frac{\bar{x}'}{\bar{x}}\right)^{\beta_9},$$
$$t_7 = \bar{y}^* + d_3\left(\bar{x} - \bar{x}^*\right) + d_4\left(\bar{x}' - \bar{x}\right).$$

To the first order of approximation, the MSEs of the above estimators are given by

$$\mathrm{MSE}\left(t_{r_7}\right) = S\left\{S_y^2 + 4RS_x^2\left(R-\beta\right)\right\} + Q\left\{S_{y_2}^2 + RS_{x_2}^2\left(R-2\beta_{(2)}\right)\right\} + TS_y^2,$$
$$\mathrm{min.MSE}\left(t_{R_7}\right) = \mathrm{min.MSE}\left(t_7\right) = TS_y^2 + SS_y^2\left(1-\rho^2\right) + QS_{y_2}^2\left(1-\rho_2^2\right),$$

where optimum values of β_8, β_9, d_3, and d_4 are given by $\beta_{8\mathrm{opt}} = \beta_{(2)}/R$, $\beta_{9\mathrm{opt}} = \beta/R$, $d_{3\mathrm{opt}} = \beta_{(2)}$, and $d_{4\mathrm{opt}} = \beta$.

Bhushan and Pandey (2017) proposed some improved regression type estimators under non-response in seven different strategies using Searls methodology (Searls, 1964). These estimators were an improvement over the corresponding regression estimators, which are BLUE, under non-response in seven different strategies stated as follow.

$$T_1 = \gamma_1\bar{y}^* + \delta_1\left(\bar{X}-\bar{x}^*\right),$$
$$T_2 = \gamma_2\bar{y}^* + \delta_2\left(\bar{X}-\bar{x}\right),$$
$$T_3 = \gamma_3\bar{y}^* + \delta_3\left(\bar{x}-\bar{x}^*\right),$$
$$T_4 = \gamma_4\bar{y}^* + \delta_4\left(\bar{x}-\bar{x}^*\right) + \eta_1\left(\bar{X}-\bar{x}\right),$$
$$T_5 = \gamma_5\bar{y}^* + \delta_5\left(\bar{x}^{'}-\bar{x}^*\right),$$
$$T_6 = \gamma_6\bar{y}^* + \delta_6\left(\bar{x}^{'}-\bar{x}\right),$$
$$T_7 = \gamma_7\bar{y}^* + \delta_7\left(\bar{x}-\bar{x}^*\right) + \eta_2\left(\bar{x}^{'}-\bar{x}\right),$$

where γ_i, δ_j, and η_k $\{i = j = 1, 2, \ldots, 7;\ k = 1, 2\}$ are suitably (optimally) chosen scalars to optimize MSE. The MSE of the estimators are given by

$$\mathrm{min.MSE}\left(T_i\right) = \frac{\bar{Y}^2\mathrm{MSE}\left(t_i\right)}{\bar{Y}^2 + \mathrm{MSE}\left(t_i\right)}, \quad i = 1, 2, \ldots, 7 \tag{2.5}$$

obviously,

$$\mathrm{min.MSE}(T_i) < \mathrm{MSE}(t_i).$$

The optimal values of γ_i, δ_j, and η_k $\{i = j = 1, 2, \ldots, 7;\ k = 1, 2\}$ are given by

$$\gamma_1 = \frac{\bar{Y}^2}{\left\{\bar{Y}^2 + PS_y^2 + QS_{y_2}^2 - \frac{\left(PS_{xy}+QS_{xy_2}\right)^2}{\left(QS_x^2+QS_{x_2}^2\right)}\right\}}, \quad \gamma_2 = \frac{\bar{Y}^2}{\left\{\bar{Y}^2 + PS_y^2(1-\rho^2) + QS_{y_2}^2\right\}},$$
$$\gamma_3 = \frac{\bar{Y}^2}{\left\{\bar{Y}^2 + PS_y^2 + QS_{y_2}^2(1-\rho_2^2)\right\}}, \quad \gamma_4 = \frac{\bar{Y}^2}{\left\{\bar{Y}^2 + PS_y^2(1-\rho^2) + QS_{y_2}^2(1-\rho_2^2)\right\}},$$
$$\gamma_5 = \frac{\bar{Y}^2}{\left\{\bar{Y}^2 + PS_y^2 + QS_{y_2}^2 - \frac{\left(SS_{xy}+QS_{xy_2}\right)^2}{\left(SS_x^2+QS_{x_2}^2\right)}\right\}}, \quad \gamma_6 = \frac{\bar{Y}^2}{\left\{\bar{Y}^2 + TS_y^2 + SS_y^2(1-\rho^2) + QS_{y_2}^2\right\}},$$
$$\gamma_7 = \frac{\bar{Y}^2}{\left\{\bar{Y}^2 + TS_y^2 + SS_y^2(1-\rho^2) + QS_{y_2}^2(1-\rho_2^2)\right\}},$$

$$\delta_1 = \frac{\bar{Y}^2\left(PS_{xy} + QS_{xy_2}\right)}{\left(PS_x^2 + QS_{x_2}^2\right)\left\{\bar{Y}^2 + PS_y^2 + QS_{y_2}^2 - \frac{\left(PS_{xy}+QS_{xy_2}\right)^2}{\left(PS_x^2+QS_{x_2}^2\right)}\right\}}, \quad \delta_2 = \frac{\bar{Y}^2 S_{xy}}{S_x^2\left\{\bar{Y}^2 + PS_y^2(1-\rho^2) + QS_{y_2}^2\right\}},$$

$$\delta_3 = \frac{\bar{Y}^2 S_{xy_2}}{S_{x_2}^2\left\{\bar{Y}^2 + PS_y^2 + QS_{y_2}^2(1-\rho_2^2)\right\}}, \quad \delta_4 = \frac{\bar{Y}^2 S_{xy_2}}{S_{x_2}^2\left\{\bar{Y}^2 + PS_y^2(1-\rho^2) + QS_{y_2}^2\left(1-\rho_2^2\right)\right\}},$$

$$\delta_5 = \frac{\bar{Y}^2\left(SS_{xy} + QS_{xy_2}\right)}{\left(SS_x^2+QS_{x_2}^2\right)\left\{\bar{Y}^2+PS_y^2+QS_{y_2}^2 - \frac{\left(SS_{xy}+BS_{xy_2}\right)^2}{\left(SS_x^2+QS_{x_2}^2\right)}\right\}}, \quad \delta_6 = \frac{\bar{Y}^2 S_{xy}}{S_x^2\left\{\bar{Y}^2 + TS_y^2 + SS_y^2(1-\rho^2) + QS_{y_2}^2\right\}},$$

$$\delta_7 = \frac{\bar{Y}^2 S_{xy_2}}{S_{x_2}^2\left\{\bar{Y}^2 + TS_y^2 + SS_y^2(1-\rho^2) + QS_{y_2}^2(1-\rho_2^2)\right\}},$$

$$\eta_1 = \frac{\bar{Y}^2 S_{xy}}{S_x^2\left\{\bar{Y}^2 + PS_y^2(1-\rho^2) + QS_{y_2}^2(1-\rho_2^2)\right\}}, \quad \eta_2 = \frac{\bar{Y}^2 S_{xy}}{S_x^2\left\{\bar{Y}^2 + TS_y^2 + SS_y^2(1-\rho^2) + QS_{y_2}^2(1-\rho_2^2)\right\}},$$

where $P = (1/n - 1/N)$, $Q = W_2(k-1)/n$, $S = (1/n - 1/n')$, and $T = (1/n' - 1/N)$.

Bhushan and Pandey (2017) showed that the these estimators were better than conventional regression estimators. In this article, we propose some new ratio type estimators and compared these with the corresponding regression estimators given earlier. The proposed estimators are motivated by Cochran (1977), Khare and Srivastava (1995), Rao (1986), Okafor and Lee (2000), and Singh and Kumar (2008) under the one phase and two phase sampling using seven different strategies under non-response.

3. Proposed improved ratio type estimators under a non-response

We propose improved ratio type estimators using Searls methodology (Searls, 1964), Searls (1964) proposed a technique to improve the conventional estimators by multiplying a tuning constant term α whose optimum value depends on the coefficient of variation, which is a fairly stable quantity, and we refer this technique of multiplication by a tuning constant α as Searls type transformation (STT), under seven different strategies in single phase sampling and two phase sampling, as follows.

$$T_s = \alpha\bar{y}^*$$

optimum value of α is given by

$$\alpha = \frac{\bar{Y}^2}{\bar{Y}^2 + PS_y^2 + QS_{y_2}^2}.$$

The proposed estimator under Strategy I, when $\bar{y}^*$, $\bar{x}^*$, and $\bar{X}$ are used, is given by

$$T_{s_1} = \alpha_1\bar{y}^*\left(\frac{\bar{X}}{\bar{x}^*}\right)^{\beta_1}.$$

The proposed estimator under Strategy II, when $\bar{y}^*$, $\bar{x}$, and $\bar{X}$ are used, is given by

$$T_{s_2} = \alpha_2\bar{y}^*\left(\frac{\bar{X}}{\bar{x}}\right)^{\beta_2}.$$

The proposed estimator under Strategy III, when $\bar{y}^*$, $\bar{x}^*$, and $\bar{x}$ are used, is given by

$$T_{s_3} = \alpha_3 \bar{y}^* \left(\frac{\bar{x}}{\bar{x}^*}\right)^{\beta_3}.$$

The proposed estimator under Strategy IV, when $\bar{y}^*$, $\bar{x}$, $\bar{x}^*$, and $\bar{X}$ are used, is given by

$$T_{s_4} = \alpha_4 \bar{y}^* \left(\frac{\bar{X}}{\bar{x}^*}\right)^{\beta_4} \left(\frac{\bar{X}}{\bar{x}}\right)^{\beta_5}.$$

The proposed estimator under Strategy V, when $\bar{y}^*$, $\bar{x}^*$, and $\bar{x}'$ are used, is given by

$$T_{s_5} = \alpha_5 \bar{y}^* \left(\frac{\bar{x}'}{\bar{x}^*}\right)^{\beta_6}.$$

The proposed estimator under Strategy VI, when $\bar{y}^*$, $\bar{x}$, and $\bar{x}'$ are used, is given by

$$T_{s_6} = \alpha_6 \bar{y}^* \left(\frac{\bar{x}'}{\bar{x}}\right)^{\beta_7}.$$

The proposed estimator under Strategy VII, when $\bar{y}^*$, $\bar{x}^*$, $\bar{x}$, and $\bar{x}'$ are used, is given by

$$T_{s_7} = \alpha_7 \bar{y}^* \left(\frac{\bar{x}'}{\bar{x}^*}\right)^{\beta_8} \left(\frac{\bar{x}'}{\bar{x}}\right)^{\beta_9},$$

where α_j $(j = 1, 2, \ldots 7)$ is a suitable chosen scalars. And the optimum values of α_j are defined in appendix.

Theorem 1. *The bias and minimum MSE of the new ratio type estimator T_{s_j} $(j = 1, 2, \ldots 7)$ is given by*

$$Bias\left(T_{s_j}\right) = \bar{Y}\left(\alpha_j - 1\right) \tag{3.1}$$

and

$$min.MSE_{\alpha_j}\left(T_{s_j}\right) = \bar{Y}^2\left(1 - \frac{B_j^2}{A_j}\right), \tag{3.2}$$

$j = 1, 2, \ldots, 7.$

Proof: *See Appendix.* □

It is interesting to note that simultaneous optimization with respect to the characterizing scalars γ_i and δ_i of the expression (2.5) of MSE is possible for regression (difference) type estimators. But simultaneous optimization with respect to the characterizing scalars α_i and β_i of the expression (3.2) not possible for ratio type estimators.

Theorem 2. *The proposed ratio type estimators T_{s_j}, $(j = 1, 2, \ldots, 7)$ are better than difference type estimators T_k, $(k = 1, 2, \ldots, 7)$ iff*

$$\frac{B_j^2}{A_j} > \gamma_{k_{opt}} \tag{3.3}$$

and vice versa. Otherwise both are equally efficient in case of equality in (3.3).

Proof: *It may be easily observed from (2.5) that the MSE of the difference type estimators T_i, ($i = 1, 2, \ldots, 7$) are given by*

$$min.MSE\,(T_i) = \bar{Y}^2\left(1 - \gamma_{i_{opt}}\right) \tag{3.4}$$

Comparing (3.4) with (3.2), we have the theorem. □

The only way ascertain (3.3) if this holds in practice is through a computational study.

Theorem 3. *The proposed ratio type estimators T_{s_j}, $j = 1, 2, \ldots, 7$ are better than the conventional ratio type estimators t_{R_j}, $j = 1, 2, \ldots, 7$ iff*

$$\frac{B_j^2}{A_j} > \left\{1 - \frac{min.MSE(t_{R_j})}{\bar{Y}^2}\right\} \tag{3.5}$$

and vice versa. Otherwise both are equally efficient in case of equality in (3.5).

Proof: *It may be easily observed from (3.2) and (2.2), we have the theorem.* □

Theorem 4. *The proposed ratio type estimators T_{s_j}, $j = 1, 2, \ldots, 7$ are better than the conventional regression type estimators t_i, $i = 1, 2, \ldots, 7$ iff*

$$\frac{B_j^2}{A_j} > \left\{1 - \frac{min.MSE(t_i)}{\bar{Y}^2}\right\} \tag{3.6}$$

and vice versa. Otherwise both are equally efficient in case of equality in (3.6).

Proof: *It may be easily observed from (3.2) and (2.4), we have the theorem.* □

4. Empirical study

In order to have a better understanding about the efficiency of the proposed estimators we have conducted a comprehensive empirical study on three populations and compared the proposed estimators with the existing estimators. The percentage relative efficiency (PRE) is calculated as

$$\text{PRE} = \frac{\text{Var}(\bar{y}^*)}{\text{min.MSE}(T_{s_j})} \times 100, \quad j = 1, 2, \ldots, 7.$$

1. The first population considered by Srivastava (1993, p.50) consists of a list of 70 villages in the administrative devision of Tehsil India that includes population and cultivated area (in acres) data from 1981. Here the cultivated area (in acres) is taken as the main study character and the population of village is taken as the auxiliary character.The parameters of the population are as follows: $N = 70$, $n' = 40$, $n = 25$, $\bar{Y} = 981.29$, $\bar{X} = 1755.53$, $S_y = 613.66$, $S_x = 1406.13$, $\bar{Y}_2 = 597.29$, $\bar{X}_2 = 1100.24$, $S_{y_2} = 244.11$, $S_{x_2} = 631.51$, $\rho = 0.778$, $\rho_2 = 0.445$, $R = 0.5589$, $\beta = 0.3395$, $\beta_2 = 0.1720$, $W_2 = 0.20$.

2. The second population considered by Khare and Kumar (2011). For the population of 96 villages of rural areas under Police Station, Singur, District Hooghly from district census Handbook (1981),

Table 1: Mean squared error and percentage relative efficiency of the estimators with respect to $\bar{y}^*$ for population 1

Estimator		k = 1	k = 2	k = 3	k = 4
	$\bar{y}^*$	10160.17 (100.00)	10636.88 (100.00)	11113.60 (100.00)	11590.32 (100.00)
	T_s	10054.08 (101.05)	10520.67 (101.10)	10986.79 (101.15)	11452.47 (101.20)
Strategy I	t_{R_1}	4288.77 (236.90)	4745.94 (224.13)	5195.20 (213.92)	5637.74 (205.58)
	T_1	4269.75 (237.95)	4722.66 (225.23)	5167.32 (215.07)	5604.92 (206.78)
	T_{s_1}	4247.51 (**239.20**)	4696.51 (**226.48**)	5137.05 (**216.34**)	5570.35 (**208.07**)
Strategy II	t_{R_2}	4298.93 (236.34)	4775.65 (222.73)	5252.36 (211.59)	5729.08 (202.31)
	T_2	4279.82 (237.39)	4752.08 (223.83)	5223.87 (212.74)	5695.19 (203.51)
	T_{s_2}	4259.43 (**238.53**)	4729.86 (**224.89**)	5199.82 (**213.73**)	5669.32 (**204.44**)
Strategy III	t_{R_3}	10065.76 (100.94)	10448.08 (101.81)	10830.39 (102.61)	11212.71 (103.37)
	T_3	9961.63 (101.99)	10335.93 (102.91)	10709.93 (103.77)	11083.65 (104.57)
	T_{s_3}	9959.45 (**102.01**)	10331.38 (**102.96**)	10702.83 (**103.84**)	11073.81 (**104.66**)
Strategy IV	t_{R_4}	4204.53 (241.65)	4586.84 (231.90)	4969.16 (223.65)	5351.47 (216.58)
	T_4	4186.25 (242.70)	4565.09 (233.00)	4961.16 (224.80)	5321.89 (217.78)
	T_{s_4}	4164.88 (**243.95**)	4540.74 (**234.25**)	4916.11 (**226.06**)	5291.00 (**219.06**)
Strategy V	t_{R_5}	6727.54 (151.02)	7182.84 (148.09)	7638.14 (145.50)	8093.44 (143.21)
	T_5	6680.87 (152.07)	7123.29 (149.32)	7554.66 (147.11)	7977.40 (145.29)
	T_{s_5}	6662.06 (**152.51**)	7104.86 (**149.71**)	7546.58 (**147.26**)	7987.23 (**145.11**)
Strategy VI	t_{R_6}	6741.11 (150.72)	7217.83 (147.37)	7694.55 (144.43)	8171.26 (141.84)
	T_6	6694.24 (151.77)	7164.13 (148.47)	7633.55 (145.58)	8102.50 (143.04)
	T_{s_6}	6677.77 (**152.15**)	7146.59 (**148.84**)	7614.95 (**145.94**)	8082.86 (**143.39**)
Strategy VII	t_{R_7}	6646.71 (152.86)	7029.02 (151.33)	7411.34 (149.95)	7793.66 (148.71)
	T_7	6601.14 (153.91)	6978.08 (152.43)	7354.73 (151.10)	7731.08 (149.91)
	T_{s_7}	6583.19 (**154.33**)	6957.40 (**152.88**)	7331.14 (**151.59**)	7704.39 (**150.44**)

published by the government of India, the data on the number of cultivators y, as a study character and the population of villages, as an auxiliary character x have been taken. The non-response rate in the population is considered to be 25%. The values of the parameters of the population are given as follows: $N = 96$, $n' = 65$, $n = 25$, $\bar{Y} = 185.22$, $\bar{X} = 1807.23$, $S_y = 195.03$, $S_x = 1921.77$, $S_{y_2} = 97.82$, $S_{x_2} = 1068.44$, $\rho = 0.904$, $\rho_2 = 0.895$, $R = 0.1025$, $W_2 = 0.25$.

3. Third population considered from Srivastava (1993, p.50). The data belongs to the data on physical growth of upper-socio-economic group of 95 school children of Varanasi under an ICMR study. The first 25% (i.e., 24 children) units have been considered as non-response units. The values of the parameters related to the study variate y (weight in kg) and the auxiliary variate x (chest circumference in cm) have been given below: $N = 95$, $n' = 70$, $n = 35$, $\bar{Y} = 19.497$, $\bar{X} = 55.8611$, $S_y = 3.0435$, $S_x = 3.2735$, $S_{y_2} = 2.3552$, $S_{x_2} = 2.5137$, $\rho = 0.8460$, $\rho_2 = 0.7290$, $R = 0.3490$, $\beta = 0.7865$, $\beta_2 = 0.6829$, $W_2 = 0.25$, $N_2 = 24$, $N_1 = 71$.

From perusal of above results it is observed that the new ratio type estimators proposed T_{s_i} are always better than the conventional ratio type counterparts t_{R_i}. Hence, we conclude that all proposed new ratio type estimators have higher efficiency in comparison to the conventional ratio type estimators. A comparison of STD (or regression) estimators T_i $(i = 1, 2, \ldots, 7)$ with new ratio type estimators T_{s_i} $(i = 1, 2, \ldots, 7)$, we observe that the new ratio type estimators are always better than the estimators proposed by Bhushan and Pandey regression type estimators under optimality condition (3.3). Therefore, in population I and II proposed ratio type estimators T_{s_j} are better than Bhushan and Pandey regression type estimators T_i, as (3.3) is satisfied. However, the case is reversed for population III.

Table 2: Mean squared error and percentage relative efficiency of the estimators with respect to $\bar{y}^*$ for population 2

Estimator		k: 1	2	3	4
	$\bar{y}^*$	1220.94 (100.00)	1316.63 (100.00)	1412.31 (100.00)	1508.00 (100.00)
	T_s	1178.98 (103.56)	1267.96 (103.84)	1356.47 (104.12)	1444.51 (104.40)
Strategy I	t_{R_1}	225.71 (540.93)	245.57 (536.15)	265.30 (532.34)	284.93 (529.25)
	T_1	224.23 (544.49)	243.83 (539.99)	263.26 (536.46)	282.58 (533.65)
	T_{s_1}	223.29 (**546.80**)	242.63 (**542.65**)	261.79 (**539.48**)	280.80 (**537.03**)
Strategy II	t_{R_2}	301.36 (405.14)	397.05(331.61)	492.74 (286.62)	588.43 (256.28)
	T_2	298.74 (408.69)	392.51(335.44)	485.76 (290.74)	578.50 (260.67)
	T_{s_2}	297.73 (**410.08**)	391.21 (**336.55**)	484.18 (**291.69**)	576.64 (**261.51**)
Strategy III	t_{R_3}	1144.29 (106.70)	1163.33 (113.18)	1182.37 (119.45)	1201.41 (125.52)
	T_3	1107.36 (110.26)	1125.18 (117.01)	1142.98 (123.56)	1160.76 (129.91)
	T_{s_3}	1106.75 (**110.32**)	1123.95 (**117.14**)	1141.10 (**123.77**)	1158.21 (**130.20**)
Strategy IV	t_{R_4}	224.72 (312.90)	243.76 (291.08)	262.79 (277.11)	281.84 (267.40)
	T_4	223.25 (546.88)	242.04 (543.98)	260.80 (541.53)	279.54 (539.46)
	T_{s_4}	222.32 (**549.18**)	240.86 (**546.63**)	259.36 (**544.54**)	277.81 (**542.82**)
Strategy V	t_{R_5}	380.12 (226.28)	400.02 (223.52)	419.93 (221.57)	439.83 (220.11)
	T_5	375.95 (324.76)	395.33 (333.05)	414.54 (340.69)	433.63 (347.76)
	T_{s_5}	374.68 (**325.86**)	393.68 (**334.44**)	412.61 (**342.28**)	431.48 (**349.49**)
Strategy VI	t_{R_6}	455.79 (184.51)	551.48 (162.44)	647.16 (149.51)	742.85 (141.01)
	T_6	449.81 (271.43)	542.75 (242.58)	635.18 (222.35)	727.11 (207.40)
	T_{s_6}	448.60 (**272.17**)	541.30 (**243.23**)	633.50 (**222.94**)	725.19 (**207.94**)
Strategy VII	t_{R_7}	379.14 (227.20)	398.18 (224.87)	417.22 (223.21)	436.26 (221.97)
	T_7	374.99 (325.59)	393.61 (334.50)	412.21 (342.62)	430.78 (350.06)
	T_{s_7}	373.74 (**326.68**)	392.05 (**335.83**)	410.31 (**344.20**)	428.53 (**351.90**)

5. Simulation study

In this section, simulation is conducted to evaluate the performance of the proposed class of estimators with respect to traditional estimators. For this study we have generated a population size $N = 1{,}000$ from standard normal distribution using MVRNORM package in software R, where study and auxiliary variable are correlated with correlation $\rho = 0.7$, draw sample of size $n = 200$ with 35% non-response. The whole simulation process starting from the drawing sample from variable Y and auxiliary variable X from normal population and calculating the estimates was repeated 50,000 times.

6. Conclusions

From the above computational results as shown in Tables 1–3, it may be concluded that the proposed estimators T_{s_j} dominate the over conventional ratio type estimators t_{R_i}. The most interesting and noticeable finding of this paper is the vitiation of conventional thought, that the ratio type estimator can at most match upto its regression estimator. Consequently in such a manner, we proved that the proposed ratio type estimators T_{s_j} $(j = 1, 2, \ldots, 7)$ provides an improvement over traditional ratio type estimators counterpart t_{R_i} $(i = 1, 2, \ldots, 7)$ while satisfying (3.5); in addition, we also proved that the proposed ratio estimator can provide an improvement over both traditional regression estimators t_i $(i = 1, 2, \ldots, 7)$ and proposed regression estimators T_i $(i = 1, 2, \ldots, 7)$ under the optimality conditions (3.3). Thus the proposed estimators are highly rewarding in terms of the increased precession of the estimates and negative impact of the non-response. Therefore, the proposed estimators may be recommended to survey practitioners for real-life applications.

Table 3: Mean squared error and percentage relative efficiency of the estimators with respect to $\bar{y}^*$ for population 3

Estimator		k = 1	k = 2	k = 3	k = 4
	$\bar{y}^*$	0.2067 (100.000)	0.2464 (100.000)	0.2860 (100.000)	0.3256 (100.000)
	T_s	0.2066 (100.054)	0.2462 (100.064)	0.2857 (100.075)	0.3253 (100.085)
Strategy I	t_{R_1}	0.0664 (311.050)	0.0853 (288.830)	0.1040 (274.840)	0.1227 (265.220)
	T_1	0.0664 (**311.104**)	0.0852 (**288.893**)	0.1040 (**274.913**)	0.1227 (**265.307**)
	T_{s_1}	0.0664(311.060)	0.0853 (288.840)	0.1040 (274.860)	0.1227 (265.250)
Strategy II	t_{R_2}	0.0871 (237.290)	0.1267 (194.380)	0.1663 (171.900)	0.2060 (158.070)
	T_2	0.0871 (**237.342**)	0.1267 (**194.441**)	0.1663 (**171.977**)	0.2058 (**158.158**)
	T_{s_2}	0.0871 (237.310)	0.1267 (194.410)	0.1663 (171.950)	0.2059 (158.130)
Strategy III	t_{R_3}	0.1857 (111.340)	0.2043 (120.610)	0.2228 (128.350)	0.2414 (134.890)
	T_3	0.1856 (**111.392**)	0.2041 (**120.680**)	0.2227 (**128.421**)	0.2412 (**134.974**)
	T_{s_3}	0.1856 (111.390)	0.2042 (120.670)	0.2227 (128.410)	0.2412 (134.960)
Strategy IV	t_{R_4}	0.0660 (312.890)	0.0846 (291.080)	0.1032 (277.110)	0.1217 (267.400)
	T_4	0.0660 (**312.950**)	0.0846 (**291.144**)	0.1031 (**277.186**)	0.1217 (**267.486**)
	T_{s_4}	0.0660 (312.910)	0.0846 (291.090)	0.1032 (277.130)	0.1217 (267.430)
Strategy V	t_{R_5}	0.0913 (226.280)	0.1102 (223.520)	0.1290 (221.570)	0.1479 (220.110)
	T_5	0.0913 (**226.331**)	0.1101 (**223.692**)	0.1288 (**221.952**)	0.1475 (**220.722**)
	T_{s_5}	0.0913 (226.300)	0.1102 (223.550)	0.1290 (221.600)	0.1479 (220.150)
Strategy VI	t_{R_6}	0.1120 (184.510)	0.1516 (162.440)	0.1913 (149.510)	0.2309 (141.010)
	T_6	0.1120 (184.568)	0.1516 (162.503)	0.1912 (**149.581**)	0.2307 (**141.098**)
	T_{s_6}	0.1120 184.540)	0.1516 (162.480)	0.1912 (149.560)	0.2308 (141.080)
Strategy VII	t_{R_7}	0.0910 (227.200)	0.1095 (224.870)	0.1281 (223.210)	0.1467 (221.970)
	T_7	0.0909 (**227.259**)	0.1095 (**224.934**)	0.1280 (**223.286**)	0.1466 (**222.057**)
	T_{s_7}	0.0909 (227.230)	0.1095 (224.890)	0.1281 (223.240)	0.1467 (222.010)

Table 4: Percentage relative efficiency (PRE) of the proposed estimators with respect to $\bar{y}^*$ using simulation

Estimator		PRE
	$\bar{y}^*$	100
	T_s	100.086
Strategy I	t_{R_1}	165.163
	T_1	166.775
	T_{s_1}	**167.016**
Strategy II	t_{R_2}	142.838
	T_2	143.0174
	T_{s_2}	**143.637**
Strategy III	t_{R_3}	120.771
	T_3	**121.753**
	T_{s_3}	120.795
Strategy IV	t_{R_4}	168.449
	T_4	168.503
	T_{s_4}	**168.727**
Strategy V	t_{R_5}	148.608
	T_5	149.331
	T_{s_5}	**149.486**
Strategy VI	t_{R_6}	130.484
	T_6	130.495
	T_{s_6}	**131.025**
Strategy VII	t_{R_7}	153.491
	T_7	153.893
	T_{s_7}	**154.491**

Appendix: Outline of derivation of Theorem 1

The MSE of T_{s_j} $(j = 1, 2, \ldots, 7)$ is given by The MSE of T_{s_1} is given by

$$\mathrm{MSE}(T_{s_1}) = \bar{Y}^2\Big[1+\alpha_1^2\left\{1+P\Big(C_y^2+\beta_1 C_x^2(2\beta_1+1)-4\beta_1\rho C_y C_x\Big)+Q\Big(C_{y_2}^2+\beta_1 C_{x_2}^2(2\beta_1+1)-4\beta_1\rho C_{y_2}C_{x_2}\Big)\right\}$$
$$-2\alpha_1\left\{1+P\left(\frac{\beta_1}{2}(\beta_1+1)C_x^2-\beta_1\rho C_y C_x\right)+Q\left(\frac{\beta_1}{2}(\beta_1+1)C_{x_2}^2-2\beta_1\rho C_{y_2}C_{x_2}\right)\right\}\Big]$$

which can be expressed as

$$\mathrm{MSE}\,(T_{s_1}) = \bar{Y}^2\left[1+\alpha_1^2 A_1 - 2\alpha_1 B_1\right]$$

For optimum value of α_1 differentiating the MSE(T_{s_1}) with respect to α_1 and equating to zero we get,

$$\alpha_{1\mathrm{opt}} = \frac{B_1}{A_1}$$

substituting the optimum value of α_1 in MSE (T_{s_1}) we get minimum MSE

$$\mathrm{min.MSE}\,(T_{s_1}) = \bar{Y}^2\left(1-\frac{B_1^2}{A_1}\right)$$

The derivations for other estimators T_{s_j} $(j = 1, 2, \ldots, 7)$ can be done on similar lines. We then have

$$\mathrm{MSE}\left(T_{s_j}\right) = \bar{Y}^2\left[1+\alpha_j^2 A_j - 2\alpha_j B_j\right].$$

It is important to mention here that simultaneous optimization w.r.t, α_j and β_j of the expression of MSE is not possible and we use optimum value of $\beta_j = \beta_{jopt}$ when $\alpha_j = 1$ and use this within $\alpha_j = \alpha_{jopt}$ to obtain (3.2) as used recently by various authors including Singh and Solanki (2013).

$$\mathrm{min.MSE}_{\alpha_j}\left(T_{s_j}\right) = \bar{Y}^2\left(1-\frac{B_j^2}{A_j}\right)$$

The optimum values of scalars for different estimators involved are given as

$$A_1 = \bar{Y}^2 + P\left\{S_y^2+\beta_1 R^2 S_x^2(2\beta_1+1)-4\beta_1 RS_{xy}\right\}+Q\left\{S_{y_2}^2+\beta_1 R^2 S_{x_2}^2(2\beta_1+1)-4\beta_1 RS_{xy_2}\right\},$$
$$B_1 = \bar{Y}^2 + P\left\{\frac{\beta_1}{2}R^2 S_x^2(\beta_1+1)-\beta_1 RS_{xy}\right\}+Q\left\{\frac{\beta_1}{2}R^2 S_{x_2}^2(\beta_1+1)-2\beta_1 RS_{xy_2}\right\},$$
$$A_2 = \bar{Y}^2 + P\left\{S_y^2+\beta_2 R^2 S_x^2(2\beta_2+1)-4\beta_2 RS_{xy}\right\}+QS_{y_2}^2,$$
$$B_2 = \bar{Y}^2 + P\left\{\frac{\beta_2}{2}R^2 S_x^2(\beta_2+1)-\beta_2 RS_{xy}\right\},$$
$$A_3 = \bar{Y}^2 + PS_y^2 + Q\left\{S_{y_2}^2+\beta_3 R^2 S_{x_2}^2(2\beta_3+1)-4\beta_3 RS_{xy_2}\right\},$$
$$B_3 = \bar{Y}^2 + Q\left\{\frac{\beta_3}{2}R^2 S_{x_2}^2(\beta_3+1)-\beta_3 RS_{xy_2}\right\},$$
$$A_4 = \bar{Y}^2 + P\left\{S_y^2+\beta_4 R^2 S_x^2(2\beta_4+1)+\beta_5 R^2 S_x^2(2\beta_5+1)-4\beta_4 RS_{xy}-4\beta_5 RS_{xy}+4\beta_4\beta_5 R^2 S_x^2\right\}$$
$$+\,Q\left\{S_{y_2}^2+\beta_4 R^2 S_{x_2}^2(2\beta_4+1)-4\beta_4 RS_{xy_2}\right\},$$

$$
\begin{aligned}
B_4 &= \bar{Y}^2 + P\left\{\frac{\beta_4}{2}R^2S_x^2(\beta_4+1) + \frac{\beta_5}{2}R^2S_x^2(\beta_5+1) - \beta_4RS_{xy} - \beta_5RS_{xy} + \beta_4\beta_5R^2S_x^2\right\} \\
&\quad + Q\left\{\frac{\beta_4}{2}R^2S_{x_2}^2(\beta_4+1) - \beta_4RS_{xy_2}\right\}, \\
A_5 &= \bar{Y}^2 + P\left(S_y^2 + \beta_6R^2S_x^2(2\beta_6+1) - 4\beta_6RS_{xy}\right) + Q\left(S_{y_2}^2 + \beta_6R^2S_{x_2}^2(2\beta_6+1) - 4\beta_6RS_{xy_2}\right) \\
&\quad - T\left(\beta_6R^2S_x^2(2\beta_6+1) - 4\beta_6RS_{xy}\right), \\
B_5 &= \bar{Y}^2 + P\left\{\frac{\beta_6}{2}R^2S_x^2(\beta_6+1) - \beta_6RS_{xy}\right\} + Q\left\{\frac{\beta_6}{2}R^2S_{x_2}^2(\beta_6+1) - \beta_6RS_{xy_2}\right\} \\
&\quad - T\left\{\frac{\beta_6}{2}R^2S_x^2(\beta_6+1) - \beta_6RS_{xy}\right\}, \\
A_6 &= \bar{Y}^2 + P\left(S_y^2 + \beta_7R^2S_x^2(2\beta_7+1) - 4\beta_7RS_{xy}\right) + QS_{y_2}^2 - T\left(\beta_7R^2S_x^2(2\beta_7+1) - 4\beta_7RS_{xy}\right), \\
B_6 &= \bar{Y}^2 + P\left\{\frac{\beta_7}{2}R^2S_x^2(\beta_7+1) - \beta_7RS_{xy}\right\} - T\left\{\frac{\beta_7}{2}R^2S_x^2(\beta_7+1) - \beta_7RS_{xy}\right\}, \\
A_7 &= \bar{Y}^2 + P\left(S_y^2 + \beta_9R^2S_x^2(2\beta_9+1) - 4\beta_9RS_{xy}\right) + Q\left(S_{y_2}^2 + \beta_8R^2S_{x_2}^2(2\beta_8+1) - 4\beta_8RS_{xy_2}\right) \\
&\quad - T\left(\beta_9R^2S_x^2(2\beta_9+1) - 4\beta_9RS_{xy}\right), \\
B_7 &= \bar{Y}^2 + P\left\{\frac{\beta_9}{2}R^2S_x^2(\beta_9+1) - \beta_9RS_{xy}\right\} + Q\left\{\frac{\beta_8}{2}R^2S_{x_2}^2(\beta_8+1) - \beta_8RS_{xy_2}\right\} \\
&\quad - T\left\{\frac{\beta_9}{2}R^2S_x^2(\beta_9+1) - \beta_9RS_{xy}\right\},
\end{aligned}
$$

where $P = (1/n - 1/N)$, $Q = W_2(k-1)/n$, $S = (1/n - 1/n')$, and $T = (1/n' - 1/N)$.

References

Bhushan S and Pandey AP (2017). An efficient estimation procedure for population mean under non-response, (Submitted).

Cochran WG (1977). *Sampling Techniques* (3rd ed), John Wiley & Sons, New York.

Hansen MH and Hurwitz WN (1946). The problem of non-response in sample surveys, *Journal of the American Statistical Association*, **41**, 517–529.

Khare BB and Kumar S (2011). Estimation of population mean using known coefficient of variation of the study character in the presence of non-response, *Communications in Statistics - Theory and Methods*, **40**, 2044–2058.

Khare BB and Srivastava S (1993). Estimation of population mean using auxiliary character in presence of non-response, *National Academy Science Letters-India*, **16**, 111–114.

Khare BB and Srivastava S (1995). Study of conventional and alternative two-phase sampling ratio, product and regression estimators in presence of non-response. In *Proceedings-National Academy Science, India, Section A*, **65**, 195–203.

Khare BB and Srivastava S (1997). Transformed ratio type estimators for the population mean in the presence of non-response, *Communications in Statistics - Theory and Methods*, **26**, 1779–1791.

Kumar S (2015). Efficient use of auxiliary information in estimating the population ratio, product and mean in presence of non-response, *Journal of Advanced Computing*, **4**, 68–87.

Okafor FC and Lee H (2000). Double sampling for ratio and regression estimation with sub sampling the non-respondents, *Survey Methodology*, **26**, 183–188.

Rao PSRS (1983). *Randomization Approach in Incomplete Data in Sample Surveys*, Vol. 2, WG

Madow, I Olkin, and DB Rubin (Eds), pp. 33–44, Academic Press, New York.

Rao PSRS (1986). Ratio estimation with subsampling the nonrespondents, *Survey Methodology* **12**, 217–230.

Searls DT (1964). The Utilization of a Known Coefficient of Variation in the Estimation Procedure, *Journal of the American Statistical Association*, **59**, 1225–1226.

Shabbir J and Khan NS (2013). On estimating the finite population mean using two auxiliary variables in two phase sampling in the presence of non response, *Communications in Statistics - Theory and Methods*, **42**, 4127–4145.

Singh HP and Kumar S (2008). A regression approach to the estimation of the finite population mean in the presence of non-response, *Australian* & *New Zealand Journal of Statistics*, **50**, 395–408.

Singh HP and Kumar S (2010). Estimation of mean in presence of non-response using two phase sampling scheme, *Statistical Papers*, **51**, 559–582.

Singh HP, Kumar S, and Kozak M (2010). Improved estimation of finite population mean using subsampling to deal with non response in two-phase sampling scheme, *Communications in Statistics - Theory and Methods*, **39**, 791–802.

Singh HP and Solanki RS (2013). An efficient class of estimators for the population mean using auxiliary information, *Communications in Statistics - Theory and Methods*, **42**, 145–163.

Singh RK and Bhushan S (2012). Generalized classes of two phase sampling estimators of population mean in presence of non-response. In *Proceeding of VII ISOS*, Aligarh Muslim University, Aligarh.

Srivastava S (1993). *Some problems on the estimation of population mean using auxiliary character in presence of non-response in sample surveys* (Ph.D thesis), BHU, Varanasi, U.P., India.

Received June 5, 2018; Revised October 12, 2018; Accepted November 25, 2018