STATISTICS IN TRANSITION-new series, December 2012
Vol. 13, No. 3, pp. 495—518

CHAIN RATIO ESTIMATOR FOR THE POPULATION MEAN IN THE PRESENCE OF NON-RESPONSE

B. B. Khare[1], U. Srivastava[2], K. Kumar[1]

ABSTRACT

In this paper we have proposed two chain ratio type estimators for population mean using two auxiliary variables in the presence of non-response. The proposed estimators have been found to be more efficient than the relevant estimators for the fixed values of preliminary sample of size n' and subsample of size $n(< n')$ taken from the preliminary sample of size n' under the specified conditions. The proposed estimators are more efficient than the corresponding estimators for population mean ($\bar{Y}$) of study variable y in the case of the fixed cost and have less total cost in comparison to the cost incurred by the corresponding relevant estimators for specified variance. The results have been supported by empirical study as well as Monte Carlo simulation study.

Key words: chain ratio type estimator, preliminary sample, bias, mean square error, non-response, additional auxiliary variable.

1. Introduction

The research work using auxiliary information in proposing selection procedure and the estimators for the population parameters was initiated by Bowely (1926), Neyman (1934, 1938), Watson (1937), Cochran (1940, 1942), Hansen et al. (1953) and Robson (1957). The population mean of the auxiliary variable (x) is not known but the population mean of an additional auxiliary variable (z) is known which is cheaper and less correlated to the study variable (y) in comparison to the main auxiliary variable ($i.e. \rho_{yx} > \rho_{yz}, \rho_{yx}, \rho_{yz} > 0$). In such case Chand (1975), Kiregyera (1980, 1984) and Srivastava et al. (1990) proposed chain ratio type estimators using additional variable with known population mean.

Sometimes, it may not be possible to collect the complete information for all the units selected in the sample due to non-response. The missing observations

[1] Department of Statistics, Banaras Hindu University, Varanasi-221005, India.
E-mail: bbkhare56@yahoo.com.
[2] Statistics Sections M.M.V., Banaras Hindu University, Varanasi-221005, India.

due to non-response may occur during the investigation, which may be at random, and the ignorance of such missing observations may lead to biased estimator, though the amount of the bias may be very negligible. If the missing observation due to non-response is not at random then the amount of bias in the estimator will be large and may increase the error in the estimation, and the sampling error will also increase. Little and Rubin (2002) suggested to ignore the missing data completely if the percentage of incomplete cases are very low. This practice will reduce the sample size and may increase the bias and the variance of the estimator when the incomplete cases are large. However, some imputation techniques to replace the missing observation are considered by Rao and Toutenburge (1995), and Toutenburge and Srivastava (1998, 2003). Estimation of the population mean in sample surveys when some observations are missing due to non-response not at random was considered by Hansen and Hurwitz (1946), Rao (1986, 1987), Khare and Srivastava (1993, 1995).

In this paper, we have proposed two chain type estimators for the population mean of the study variable in the presence of non-response. The expressions for the bias and mean square error of the proposed estimator are obtained and a comparison of the proposed estimators has been made with the relevant estimators.

The optimum values of the preliminary sample (n'), subsample (n) and the sub-sampling fraction $(k^{-1}, k>1)$ have been obtained for fixed cost $C \le C_0$ and for the specified variance $V = V_0'$. A comparative study of the proposed estimators with the relevant estimators has been made with the help of an empirical study as well as Monte Carlo simulation study.

2. The estimators

Let Y_i, X_i and Z_i be the non-negative values for the ith unit of the population $U = U_1, U_2, ...U_N$ on the study variable y, the auxiliary variable x and the additional auxiliary variable z with their population means $\bar{Y}$, $\bar{X}$ and $\bar{Z}$. Here $\bar{X}$ is unknown, but $\bar{Z}$, the population mean of additional auxiliary variable (z) (closely related to x) is known, which may be cheaper and less correlated to the study variable (y) in comparison to the main auxiliary variable (x). Now a preliminary sample of size $n'(n' < N)$ is selected from the population of size N using simple random sampling without replacement (SRSWOR) scheme, and we estimate the population mean $\bar{X}$ using additional auxiliary variable with known population mean $\bar{Z}$ and n' observations on x and z. Again, subsample of size $n(< n')$ is selected from the preliminary sample of size n' by using simple random sampling without replacement (SRSWOR) scheme, and it has been observed that n_1 units respond and n_2 unit do not respond in the sample of size

n for the study variable y. It is also assumed that the population of size N is composed of N_1 responding and N_2 non-responding units, though they are unknown. Further, a sub sample of size r $(r = n_2 k^{-1}, k > 1)$ from n_2 non-responding units has been drawn by using SRSWOR method of sampling by making extra effort. Hence, we have $(n_1 + r)$ observations on the y variable.

Using Hansen and Hurwitz (1946) technique, the estimator for population mean using $(n_1 + r)$ observations on y variable is given by

$$\bar{y}^* = \frac{n_1}{n}\bar{y}_1 + \frac{n_2}{n}\bar{y}_2', \tag{2.1}$$

where $\bar{y}_1$ and $\bar{y}_2'$ are the sample means of variable y based on n_1 and r units respectively. The estimator $\bar{y}^*$ is unbiased and has variance

$$V(\bar{y}^*) = \frac{f}{n}S_y^2 + \frac{W_2(k-1)}{n}S_{y(2)}^2, \tag{2.2}$$

where $f = 1 - \frac{n}{N}, W_2 = \frac{N_2}{N}$, $S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})^2$ and $S_{y(2)}^2 = \frac{1}{N_2 - 1}\sum_{i=1}^{N_2}(Y_{i(2)} - \bar{Y}_{(2)})^2$ are the population mean square of y for the entire population and for the non-responding part of the population.

Similarly, the estimator $\bar{x}^*$ for the population mean $\bar{X}$ in the presence of non-response based on corresponding $(n_1 + r)$ observations is given by

$$\bar{x}^* = \frac{n_1}{n}\bar{x}_1 + \frac{n_2}{n}\bar{x}_2', \tag{2.3}$$

where $\bar{x}_1$ and $\bar{x}_2'$ are the sample means of variable x based on n_1 and r units respectively, we also have

$$V(\bar{x}^*) = \frac{f}{n}S_x^2 + \frac{W_2(k-1)}{n}S_{x(2)}^2, \tag{2.4}$$

where $S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \bar{X})^2$ and $S_{x(2)}^2 = \frac{1}{N_2 - 1}\sum_{i=1}^{N_2}(X_{i(2)} - \bar{X}_{(2)})^2$ are population mean square of x for the entire population and non-responding part of the population.

In case when population mean $\bar{X}$ is not known, then, it is estimated by taking a preliminary sample of size $n'(n' < N)$ from the population of size N by using simple random sampling without replacement (SRSWOR) sampling scheme. In this situation when we have incomplete information both on the study variable y and incomplete/complete information on auxiliary variable x from the subsample of size n, Khare and Srivastava (1995) proposed conventional (T_1)/ alternative (T_2) two phase sampling ratio estimator for population mean in the presence of non-response, which are given as follows:

$$T_1 = \frac{\bar{y}^*}{\bar{x}^*}\bar{x}', \qquad T_2 = \frac{\bar{y}^*}{\bar{x}}\bar{x}', \tag{2.5}$$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, $\bar{x}' = \frac{1}{n'}\sum_{i=1}^{n'} x_i$

Now, we propose two chain ratio type estimators for the population mean in the presence of non-response, in the situation when $\bar{X}$ is unknown, but $\bar{Z}$ is known and we have incomplete information on study variable y and the incomplete /complete information on auxiliary variable x.

If $\bar{X}$ is not known, but $\bar{Z}$, the population mean of the additional auxiliary variable z (closely related to x) is known which may be cheaper and less correlated to the study variable (y) in comparison to main auxiliary variable (x), i.e. $\rho_{yx} > \rho_{yz}$ we take in this situation a preliminary sample of size $n'(n' < N)$ from the population of size N with SRSWOR scheme and estimate the population mean $\bar{X}$ by using the sample means $\bar{x}'$ and $\bar{z}'$ based on n' units and the known additional population mean $\bar{Z}$ .We see that $\hat{\bar{X}}_r = \frac{\bar{x}'}{\bar{z}'}\bar{Z}$ is more precise than preliminary sample mean $\bar{x}'$ if $\rho_{xz} > \frac{1}{2}\frac{C_z}{C_x}$, where $\bar{z}' = \frac{1}{n'}\sum_{i=1}^{n'} z_i$.

Now, we propose conventional (t_1) and alternative (t_2) chain ratio type estimators for $\bar{Y}$ using available information on two auxiliary variables x and z in the presence of non-response, which are given as follows:

$$t_1 = \frac{\bar{y}^*}{\bar{x}^*}\frac{\bar{x}'}{\bar{z}'}\bar{Z}, \qquad t_2 = \frac{\bar{y}^*}{\bar{x}}\frac{\bar{x}'}{\bar{z}'}\bar{Z} \tag{2.6}$$

3. Expressions for the Relative Bias (RB) and Approximated Mean Square Error (AMSE) of the proposed estimators

Using the large sample approximations, the expressions for the relative bias and approximated mean square error of the estimators' t_1 and t_2 up to the terms of order (n^{-1}) are given by (see appendix)

$$RB(t_1) = RB(T_1) + \frac{f'}{n'}(C_z^2 - \rho_{yz}C_yC_z), \tag{3.1}$$

$$RB(t_2) = RB(T_2) + \frac{f'}{n'}(C_z^2 - \rho_{yz}C_yC_z), \tag{3.2}$$

$$AMSE(t_1) = AMSE(T_1) + \bar{Y}^2\frac{f'}{n'}(C_z^2 - 2\rho_{yz}C_yC_z), \tag{3.3}$$

and

$$AMSE(t_2) = AMSE(T_2) + \bar{Y}^2 \frac{f'}{n'}(C_z^2 - 2\rho_{yz} C_y C_z), \tag{3.4}$$

where

$$RB(T_1) = \left(\frac{1}{n} - \frac{1}{n'}\right)(C_x^2 - \rho_{yx} C_y C_x) + \frac{W_2(k-1)}{n}(C_{x(2)}^2 - \rho_{yx(2)} C_{y(2)} C_{x(2)}), \tag{3.5}$$

$$RB(T_2) = \left(\frac{1}{n} - \frac{1}{n'}\right)(C_x^2 - \rho_{yx} C_y C_x), \tag{3.6}$$

$$AMSE(T_1) = \bar{Y}^2 \Bigg[\left(\frac{1}{n} - \frac{1}{n'}\right)(C_y^2 + C_x^2 - 2\rho_{yx} C_y C_x) + \frac{W_2(k-1)}{n}(C_{y(2)}^2 + C_{x(2)}^2 - 2\rho_{yx(2)} C_{y(2)} C_{x(2)}) + \frac{f'}{n'} C_y^2\Bigg], \tag{3.7}$$

and

$$AMSE(T_2) = \bar{Y}^2 \left[\left(\frac{1}{n} - \frac{1}{n'}\right)(C_y^2 + C_x^2 - 2\rho_{yx} C_y C_x) + \frac{W_2(k-1)}{n} C_{y(2)}^2 + \frac{f'}{n'} C_y^2\right]. \tag{3.8}$$

$$C_y = \frac{S_y}{\bar{Y}},\ C_x = \frac{S_x}{\bar{X}},\ C_z = \frac{S_z}{\bar{Z}},\ C_{y(2)} = \frac{S_{y(2)}}{\bar{Y}},\ C_{x(2)} = \frac{S_{x(2)}}{\bar{X}},\ f' = \left(1 - \frac{n'}{N}\right),$$

$$S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})^2,\quad S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \bar{X})^2,\quad S_z^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Z_i - \bar{Z})^2,$$

$$S_{y(2)}^2 = \frac{1}{N_2 - 1}\sum_{i=1}^{N_2}(Y_{i(2)} - \bar{Y}_{(2)})^2 \text{ and } S_{x(2)}^2 = \frac{1}{N_2 - 1}\sum_{i=1}^{N_2}(X_{i(2)} - \bar{X}_{(2)})^2.$$

4. The estimators of $RB(t_1)$, $RB(t_2)$, $AMSE(t_1)$ and $AMSE(t_2)$

$$R\hat{B}(t_1) = R\hat{B}(T_1) + \frac{f'}{n'}(c_z^2 - c_{yz}), \tag{4.1}$$

$$R\hat{B}(t_2) = R\hat{B}(T_2) + \frac{f'}{n'}(c_z^2 - c_{yz}), \tag{4.2}$$

$$AM\hat{S}E(t_1) = AM\hat{S}E(T_1) + \bar{y}\frac{f'}{n'}(c_z^2 - 2c_{yz}) \tag{4.3}$$

and $$AM\hat{S}E(t_2) = AM\hat{S}E(T_2) + \bar{y}\frac{f'}{n'}(c_z^2 - 2c_{yz}), \tag{4.4}$$

where

$$R\hat{B}(T_1) = \left(\frac{1}{n} - \frac{1}{n'}\right)(c_x^2 - c_{yx}) + \frac{W_2(k-1)}{n}(c_{x(2)}^2 - c_{yx(2)}), \tag{4.5}$$

$$R\hat{B}(T_2) = \left(\frac{1}{n} - \frac{1}{n'}\right)(c_x^2 - c_{yx}), \tag{4.6}$$

$$AM\hat{S}E(T_1) = \bar{y}^2\left[\left(\frac{1}{n} - \frac{1}{n'}\right)(c_y^2 + c_x^2 - 2c_{yx})\right.$$

$$\left. + \frac{W_2(k-1)}{n}(c_{y(2)}^2 + c_{x(2)}^2 - 2c_{yx(2)}) + \frac{f'}{n'}c_y^2\right] \tag{4.7}$$

$$AM\hat{S}E(T_2) = \bar{y}^2\left[\left(\frac{1}{n} - \frac{1}{n'}\right)(c_y^2 + c_x^2 - 2c_{yx}) + \frac{W_2(k-1)}{n}c_{y(2)}^2 + \frac{f'}{n'}c_y^2\right] \tag{4.8}$$

$$c_y = \frac{s_y}{\bar{y}},\ c_x = \frac{s_x}{\bar{x}},\ c_z = \frac{s_z}{\bar{z}}, c_{yx} = \frac{s_{yx}}{\bar{y}\,\bar{x}}, c_{yz} = \frac{s_{yz}}{\bar{y}\,\bar{z}}, c_{y(2)} = \frac{s_{y(2)}}{\bar{y}},\ c_{x(2)} = \frac{s_{x(2)}}{\bar{x}},$$

$$c_{yx(2)} = \frac{s_{yx(2)}}{\bar{y}\,\bar{x}}, s_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2, s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2, s_z^2 = \frac{1}{n-1}\sum_{i=1}^{n}(z_i - \bar{z})^2,$$

$$s_{yx} = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x}), s_{yz} = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})(z_i - \bar{z}), s_{y(2)}^2 = \frac{1}{r-1}\sum_{i=1}^{r}(y'_{i2} - \bar{y}'_2)^2,$$

$$s_{x(2)}^2 = \frac{1}{r-1}\sum_{i=1}^{r}(x'_{i2} - \bar{x}'_2)^2 \text{ and } s_{yx(2)} = \frac{1}{r-1}\sum_{i=1}^{r}(y'_{i2} - \bar{y}'_2)(x'_{i2} - \bar{x}'_2),$$

where (y_i, x_i, z_i) are the values of the i^{th} unit of the sample of size n for variables y, x and z respectively. (y'_{i2}, x'_{i2}) are the values of i^{th} unit in the subsample of size r drawn from n_2 non-responding units for the variables y and x respectively by using SRSWOR scheme of sampling.

The proposed estimators of $RB(t_1)$, $RB(t_2)$, $AMSE(t_1)$ and $AMSE(t_2)$ given by (4.1) to (4.4) are almost unbiased up to terms of order (n^{-1}) and the $AMSE$ of these proposed estimators will be of order (n^{-2}) and will be dependent upon the values of higher order moment(> order 2) involved in it.

5. Comparison of the proposed estimators t_1 and t_2 with respect to $\bar{y}^*, T_1$ and T_2

Comparing the estimator t_1 and t_2 with $\bar{y}^*$, T_1 and T_2 in terms of precision, we see

$$AMSE(t_1) < AMSE(\bar{y}^*) \text{ if } \rho_{yz} > \frac{1}{2}\frac{C_z}{C_y},\quad \rho_{yx} > \frac{1}{2}\frac{C_x}{C_y} \text{ and } \rho_{yx(2)} > \frac{1}{2}\frac{C_{x(2)}}{C_{y(2)}} \tag{5.1}$$

$$AMSE(t_2) < AMSE(\bar{y}^*) \text{ if } \rho_{yz} > \frac{1}{2}\frac{C_z}{C_y} \quad , \ \rho_{\text{yx}} > \frac{1}{2}\frac{C_x}{C_y} . \tag{5.2}$$

Under the condition (5.1), we have

$$AMSE(t_1) < AMSE(T_1) \text{ if } \rho_{yz} > \frac{1}{2}\frac{C_z}{C_y} . \tag{5.3}$$

Under the condition (5.2), we have

$$AMSE(t_2) < AMSE(T_2) \text{ if } \rho_{yz} > \frac{1}{2}\frac{C_z}{C_y} . \tag{5.4}$$

6. Determination of n', n and k (for the fixed cost $C \le C_0$)

Let C_0 be the total cost (fixed) of the survey apart from overhead cost.

The cost function C' can be written as:

$$C' = (c_1' + c_2')n' + c_1 n + c_2 n_1 + c_3 \frac{n_2}{k} , \tag{6.1}$$

since C' vary from sample to sample, so the expected cost can be written as:

$$C = E(C') = (c_1' + c_2')n' + n(c_1 + c_2 W_1 + c_3 \frac{W_2}{k}) , \tag{6.2}$$

where

c_1' - the cost per unit of obtaining information on auxiliary variable x,

c_2' - the cost per unit of obtaining information on additional auxiliary variable z,

c_1 - the cost per unit of mailing questionnaire/visiting the unit at the subsample,

c_2 - the cost per unit of collecting, processing data obtained from n_1 responding units,

c_3 - the cost per unit of obtaining and processing data (after extra efforts) for the subsampling units and $W_1 = N_1 / N$, response rate in the population.

Using the $AMSE(t_1)$ and $AMSE(t_2)$ from (3.3) and (3.4), the expression for $AMSE(t_i), i = 1, 2$ can be written in term of notations V_{0i}, V_{1i}, V_{2i} which is given as:

$$AMSE(t_i) = \frac{V_{0i}}{n} + \frac{V_{1i}}{n'} + \frac{k V_{2\text{i}}}{n} - \frac{1}{N}(S_y^2 + R^2 S_z^2 - 2R S_{yz}) , \tag{6.3}$$

where V_{oi}, V_{1i} and V_{2i} are respectively the coefficients of the terms of n^{-1}, n'^{-1}, and $k\, n^{-1}$ in the expression of $AMSE(t_i)$ and $R = \bar{Y} / \bar{Z}$. Now, for minimizing the $AMSE(t_i)$ for the fixed cost $C \le C_0$ and to obtain the optimum values of n', n and k , we define a function ϕ as:

$$\phi = AMSE(t_i) + \lambda_i \left\{ (c_1' + c_2')n' + n(c_1 + c_2 W_1 + c_3 \frac{W_2}{k}) - C_0 \right\}, \text{ i=1, 2,} \tag{6.4}$$

where λ_i is the Lagrange's multiplier. Now, differentiating ϕ with respect to n', n and k and equating to zero, we have

$$n' = \sqrt{\frac{V_{1i}}{\lambda_i (c_1' + c_2')}}, \qquad n = \sqrt{\frac{V_{0i} + k V_{2i}}{\lambda_i (c_1 + c_2 W_1 + c_3 \frac{W_2}{k})}} \tag{6.5}$$

and $k_{opt} = \sqrt{\dfrac{V_{0i} c_3 W_2}{V_{2i}(c_1 + c_2 W_1)}}$. (6.6)

Now, putting the values of n', n from (6.5) in (6.2) and using k_{opt} given by (6.6), we have

$$\sqrt{\lambda_i} = \frac{1}{C_0}\left[\sqrt{V_{1i}(c_1' + c_2')} + \sqrt{(V_{0i} + k_{opt} V_{2i})(c_1 + c_2 W_1 + c_3 \frac{W_2}{k_{opt}})}\right], \text{ i=1, 2.} \tag{6.7}$$

It has been observed that the determinant of the matrix of second order derivative of ϕ with respect to n', n and k is positive for the optimum values of n', n and k, which shows that the solutions for n', n given by (6.5) using (6.6) and (6.7), and the optimum of k under the condition $C \le C_0$ minimize the mean square error of t_i. It is also important to note here that the subsampling fraction (k_{opt}^{-1}) will decrease as $\sqrt{c_3 / (c_1 + c_2 W_1)}$ increases. The minimum value of $AMSE(t_i)$ can be obtained by putting the optimum values of n', n and k from (6.5), and (6.6) in the expression (6.3), so we have

$$AMSE(t_i)_{\min} = \frac{1}{C_0}\left[\sqrt{V_{1i}(c_1' + c_2')} + \sqrt{(V_{0i} + k_{opt} V_{2i})(c_1 + c_2 W_1 + c_3 \frac{W_2}{k_{opt}})}\right]^2 - \frac{1}{N}(S_y^2 + R^2 S_z^2 - 2RS_{yz}). \tag{6.8}$$

Now, neglecting the term of O (N^{-1}), we have

$$AMSE(t_i)_{\min} = \frac{1}{C_0}\left[\sqrt{\mathrm{V}_{1\mathrm{i}}(c_1' + c_2')} + \sqrt{(V_{0i} + k_{opt} V_{2i})(c_1 + c_2 W_1 + c_3 \frac{W_2}{k_{opt}})}\right]^2. \tag{6.9}$$

Now, putting the value of k_{opt} from (6.6) in (6.9), we have

$$AMSE(t_i)_{\min} = \frac{1}{C_0}\left[\sqrt{V_{1i}(c_1' + c_2')} + \sqrt{V_{0i}(c_1 + c_2 W_1)} + \sqrt{c_3 W_2 V_{2i}}\right]^2 \tag{6.10}$$

In case of $\bar{y}^*$, the expected total cost is given by

$$C = E(C') = n(c_1 + c_2 W_1 + c_3 \frac{W_2}{k}) \tag{6.11}$$

For the fixed cost C_0, the expression for $AMSE(\bar{y}^*)_{\min}$ will be given by

$$AMSE(\bar{y}^*)_{\min} = \frac{1}{C_0}\left[\sqrt{V_0(c_1 + c_2 W_1)} + \sqrt{c_3 W_2 V_2}\right]^2 - \frac{S_y^2}{N} \tag{6.12}$$

Now, neglecting the term of $O\,(N^{-1})$, we have

$$AMSE(\bar{y}^*)_{\min} = \frac{1}{C_0}\left[\sqrt{V_0(c_1 + c_2 W_1)} + \sqrt{c_3 W_2 V_2}\right]^2, \tag{6.13}$$

where $V_0 = S_y^2 - W_2\, S_{y(2)}^2$ and $V_2 = W_2 S_{y(2)}^2$ are the coefficients of the terms of n^{-1} and $k\, n^{-1}$ in the expression

$$V(\bar{y}^*) = \frac{f}{n} S_y^2 + \frac{W_2(k-1)}{n} S_{y(2)}^2 . \tag{6.14}$$

For the fixed cost C_0, the expression for $AMSE(T_i)_{\min}$ will be given by

$$AMSE(T_i)_{\min} = \frac{1}{C_0}\left[\sqrt{c_1' U_{1i}} + \sqrt{U_{0i}(c_1 + c_2 W_1)} + \sqrt{c_3 W_2 U_{2i}}\right]^2 - \frac{S_y^2}{N} \quad \text{i=1, 2.} \tag{6.15}$$

Now, neglecting the term of $O\,(N^{-1})$, we have

$$AMSE(T_i)_{\min} = \frac{1}{C_0}\left[\sqrt{c_1' U_{1i}} + \sqrt{U_{0i}(c_1 + c_2 W_1)} + \sqrt{c_3 W_2 U_{2i}}\right]^2, \tag{6.16}$$

where U_{0i} U_{1i} and U_{2i} are the coefficients of the terms of n^{-1}, n'^{-1} and $k\, n^{-1}$ in the expressions of $AMSE(T_i),\ i = 1, 2.$

It is important to obtain the values of c_1' for which the estimator T_i will be more efficient than $\bar{y}^*$, i. e. $AMSE(T_i)_{\min} < AMSE(\bar{y}^*)_{\min}$ is given by

$$c_1' < \frac{1}{U_{1i}}\left[\sqrt{(c_1 + c_2 W_1)}(\sqrt{V_0} - \sqrt{U_{0i}}) + \sqrt{c_3 W_2}(\sqrt{V_2} - \sqrt{U_{2i}})\right]^2. \tag{6.17}$$

For the fixed cost C_0, the value of c_2' for which the estimator t_i will be more efficient than T_i, is given by $\dfrac{c_2'}{c_1'} < \dfrac{(U_{1i} - V_{1i})}{V_{1i}}$.

Sometimes it may be required to obtain an estimator with the desired degree of precision with minimum cost C, so for the specified variance of the estimator, say, V_0', we obtain the optimum values of n', n and k, which have minimum cost.

7. Determination of n', n and k for the specified variance $V = V_0'$

Let V_0' be the variance of the estimator t_i fixed in advance, so we have

$$V_0' = \frac{V_{0i}}{n} + \frac{V_{1i}}{n'} + \frac{kV_{2i}}{n} - \frac{1}{N}(S_y^2 + R^2 S_z^2 - 2RS_{yz}) \quad \text{i=1, 2.} \tag{7.1}$$

For minimizing the average total cost C for the specified variance $(\cong AMSE(t_i),\ \ i = 1,2)$ of the estimator t_i, for obtaining the optimum values of $n',\ n$ and k, we define a function ψ given as:

$$\psi = (c_1' + c_2')n' + n(c_1 + c_2 W_1 + c_3 \frac{W_2}{k}) + \mu_i (AMSE(t_i) - V_0') \quad \text{i=1, 2,} \tag{7.2}$$

where μ_i is the Lagrange's multiplier. Now, differentiating ψ with respect to n', n and k and equal to zero, we have

$$n' = \sqrt{\frac{\mu_i V_{1i}}{(c_1' + c_2')}}, \qquad n = \sqrt{\frac{\mu_i (V_{0i} + kV_{2i})}{(c_1 + c_2 W_1 + c_3 \frac{W_2}{k})}}, \tag{7.3}$$

and $$k_{opt} = \sqrt{\frac{V_{0i} W_2 c_3}{V_{2i}(c_1 + c_2 W_1)}} . \tag{7.4}$$

Now, putting the value of n', n from (7.3) in (7.1) and using k_{opt} given by (7.4), we have

$$\sqrt{\mu_i} = \frac{\left[\sqrt{(V_{0i} + k_{opt} V_{2i})(c_1 + c_2 W_1 + c_3 \frac{W_2}{k_{opt}})} + \sqrt{V_{1i}(c_1' + c_2')}\right]}{V_0' + \frac{1}{N}(S_y^2 + R^2 S_z - 2R\, S_{yz})} . \tag{7.5}$$

The optimum values of n' and n are obtain by putting the value of k_{opt} and the optimum value of $\sqrt{\mu_i}$ from (7.4) and (7.5) in (7.3). Further, it has been observed that the determinant of the matrix of the second order derivatives of ψ with respect to $n',\ n$ and k is positive for the optimum values of $n',\ n$ and k. The minimum expected total cost to be incurred on the use of t_i for the specified variance V_0' will be given by

$$C_{i\min} = \frac{\left[\sqrt{(V_{0i} + k_{opt} V_{2i})(c_1 + c_2 W_1 + c_3 \frac{W_2}{k_{opt}})} + \sqrt{V_{1i}(c_1' + c_2')}\right]^2}{V_0' + \frac{1}{N}(S_y^2 + R^2 S_z^2 - 2R\, S_{yz})} \quad \text{i= 1, 2.} \tag{7.6}$$

Now, neglecting the terms of O (N^{-1}), we have

$$C_{i\min} = \frac{\left[\sqrt{(V_{0i} + k_{opt}V_{2i})(c_1 + c_2W_1 + c_3\dfrac{W_2}{k_{opt}})} + \sqrt{V_{1i}(c_1' + c_2')}\right]^2}{V_0'} \quad \text{i=1, 2.} \qquad (7.7)$$

8. An empirical study

The data on physical growth of upper socio-economic group of 95 school going children of Varanasi under an ICMR study, Department of Pediatrics, B.H.U., during 1983-84 has been taken under study, (Khare and Sinha (2007)).The last 25% (i.e. 24 children) of units have been considered as non-responding units. The study variable (y), auxiliary variable (x) and the additional auxiliary variable (z) are taken as follows:

y - weight (in kg) of the children,

x - chest circumference (in cm) of the children,

z - skull circumference (in cm) of the children.

The values of the parameters of the y, x and z variables for the given data are taken as follows:

$\overline{Y} = 21.9758$, $\overline{X} = 57.2116$, $\overline{Z} = 51.6084$, $C_y = 0.1905$, $C_x = 0.0705$, $C_z = 0.0322$, $C_{y(2)} = 0.1856$, $C_{x(2)} = 0.0752$, $\rho_{yx} = 0.8338$, $\rho_{yz} = 0.4274$, $\rho_{yx(2)} = 0.8426$.

Table 1. Absolute bias and Relative efficiency (with respect to $\bar{y}^*$) of the estimators $\bar{y}^*, T_1, T_2, t_1$ and t_2 for the fixed values of n', n and different values of k (N =*95*, n' =*70* and n =*50*)

Estimators	$1/k$		
	1/4	1/3	1/2
$\bar{y}^*$	00.00 100 (0.4181)*	00.00 100 (0.3340)	00.00 100 (0.2499)
T_1	0.00281 175 (0.2392)	0.00214 168 (0.1988)	0.00146 158 (0.1583)
T_2	0.00078 113 (0.3699)	0.00078 117 (0.2859)	0.00078 124 (0.2019)
t_1	0.00268 180 (0.2316)	0.00201 175 (0.1912)	0.00133 166 (0.1507)
t_2	0.00065 115 (0.3623)	0.00065 120 (0.2783)	0.00065 127 (0.1942)

**Figures in parenthesis give the AMSE(.)*.

From table 1, we observe that for the fixed value of n', n and $k = 2, 3, 4$, the absolute bias of the estimators t_1 and t_2 is less than the bias of the corresponding estimator T_1 and T_2. The amount of absolute bias for the estimators T_1 and t_1 decreases as values of k^{-1} increase but the amount of absolute bias for the estimators T_2 and t_2 remains constant for different values of k.We also observe that the estimator t_1 has maximum relative efficiency with respect to $\bar{y}^*$. Since the condition for preference of T_1 over T_2 is attained in the given data set, so we observe that the estimator T_1 is more efficient than T_2 and $\bar{y}^*$. However, the estimators t_1 and t_2 are more efficient than $\bar{y}^*$ and correspondingly with T_1 and T_2. In this case t_1 is found to be more efficient than t_2.

9. Monte Carlo simulation study

Case 1: In the present Monte Carlo simulation study, we consider the same data set as described in the previous section-8. From the population of 95 schools going children of Varanasi, two preliminary samples of different sizes 70 and 60 are taken by simple random sampling without replacement and the values of $\bar{x}'$ and $\bar{z}'$ based on 70 and 60 units are calculated. Again, we take two subsample of different size 50 and 40 from each preliminary sample of size 70 and 60 respectively using simple random sampling without replacement scheme. In each subsample, the last 25% (12 and 10 children respectively) of units have been considered as non-responding units. We again take a subsample of r units from non-responding units with simple random sampling without replacement and collect all information on $r = n_2 k^{-1}$ units. Here, n_2 =12, 10 is the non-responding unit in each subsample and k = 2, 3, 4 respectively. The above process is replicated 1000 times.

Simulated absolute bias and simulated mean square error of $t_i\,(i = 1, 2)$ are calculated as follows:

$$Absolute\ \ Bias(t_i) = \frac{1}{1000} \left| \sum_{j=1}^{1000} t_{i\,j} - \bar{Y} \right|,$$

$$MSE(t_i) = \frac{1}{1000} \sum_{j=1}^{1000} \left(t_{i\,j} - \bar{Y} \right)^2 .$$

Table 2. Simulated absolute bias and Simulated relative efficiency (with respect to $\bar{y}^*$) of the estimators $\bar{y}^*, T_1, T_2, t_1$ and t_2 for the fixed values of n', n and different values of k (N =95, n' =70 and n =50)

Estimators	$1/k$		
	1/4	1/3	1/2
$\bar{y}^*$	0.02313 100 (0.4248)*	0.02006 100 (0.3494)	0.01421 100 (0.2487)
T_1	0.02078 170 (0.2501)	0.01587 160 (0.2188)	0.00865 153 (0.1627)
T_2	0.02268 114 (0.3738)	0.01843 116 (0.3015)	0.01001 123 (0.2023)
t_1	0.02018 175 (0.2425)	0.01564 167 (0.2097)	0.00609 160 (0.1552)
t_2	0.02205 116 (0.3671)	0.01819 120 (0.2921)	0.00745 128 (0.1951)

**Figures in parenthesis give the* $MSE(.)$.

From table 2, we observe that for the fixed value of n', n and $k = 2, 3, 4$, the simulated absolute bias of the estimators (t_1, t_2) is less than the corresponding estimators (T_1, T_2) and $\bar{y}^*$. We also observe that the estimators (t_1, t_2) have less mean square error in comparison to the corresponding estimators (T_1, T_2) and $\bar{y}^*$.

The estimator t_1 has less mean square error in comparison to the estimator t_2. When k^{-1} increases, mean square error of all the estimators decreases.

From table 1-2, we observe that the amount of $Bias(.)$ and $AMSE(.)$ based on empirical study and the amount of $Bias(.)$ and $MSE(.)$ based on simulation study of different estimators considered under the study is almost same.

Table 3. Simulated absolute bias and Simulated relative efficiency (with respect to $\bar{y}^*$) of the estimators $\bar{y}^*, T_1, T_2, t_1$ and t_2 for the fixed values of n', n and different values of k (N =95, n' = 60 and n =40)

Estimators	$1/k$		
	1/4	1/3	1/2
$\bar{y}^*$	0.03455 100 (0.6589)*	0.02174 100 (0.5005)	0.01381 100 (0.3761)
T_1	0.01758 167 (0.3939)	0.01447 159 (0.3144)	0.00585 156 (0.2410)

Table 3. Simulated absolute bias and Simulated relative efficiency (with respect to $\bar{y}^*$) of the estimators $\bar{y}^*, T_1, T_2, t_1$ and t_2 for the fixed values of n', n and different values of k ($N = 95$, $n' = 60$ and $n = 40$) (cont.)

Estimators	$1/k$		
	1/4	1/3	1/2
T_2	0.02978 112 (0.5887)	0.01813 115 (0.4341)	0.01237 123 (0.3045)
t_1	0.01521 172 (0.3834)	0.01303 167 (0.2994)	0.00549 164 (0.2291)
t_2	0.02741 114 (0.5778)	0.01667 120 (0.4184)	0.01200 129 (0.2922)

* *Figures in parenthesis give the* $MSE(.)$.

From table 3, we observe that for the fixed value of n', n and $k = 2, 3, 4$, simulated absolute bias of the estimators (t_1, t_2) is less than the corresponding estimators (T_1, T_2) and $\bar{y}^*$. Simulated absolute bias of the estimators decreases as values of k^{-1} increase. We also observe that the estimators (t_1, t_2) have less mean square error in comparison to the corresponding estimators (T_1, T_2) and $\bar{y}^*$. The estimator t_2 has more mean square errors in comparison to the estimator t_1. Mean square error of all the estimators decreases as k^{-1} increases.

Case 2: For the simulation study of the estimators $R\hat{B}(.)$ and $AMS\hat{E}(.)$. In this case, we haven taken the sample of 50 units from the population of 95 school going children of Varanasi with simple random sampling without replacement and calculated $c_y^2, c_x^2, c_z^2, c_{xy}$ and c_{yz} based on 50 units**.** In sample of 50 units, the last 25% (12 children) of units have been considered as non-responding units. Again, we take a subsample of $r = 12k^{-1}$ units from non-responding units in sample of 50 units with simple random sampling without replacement and calculate $c_{y(2)}^2$, $c_{x(2)}^2$ and $c_{yx(2)}$ based on r units. Here, we take k = 2, 3, 4. After putting the values of c_y^2, c_x^2, c_z^2 , $c_{y(2)}^2, c_{x(2)}^2, c_{yx}, c_{yz}$ and $c_{yx(2)}$ in the expression of the estimators $R\hat{B}(.)$ and $AMS\hat{E}(.)$ and replicating above process 1000 times, we find the simulated values of the estimators $R\hat{B}(.)$ and $AMS\hat{E}(.)$.

Table 4. Simulated absolute value of the estimator $R\hat{B}(.)$ and Simulated relative efficiency (with respect to $AM\hat{S}E(\bar{y}^*)$) of the estimator $AM\hat{S}E(.)$ for the fixed values of n', n and different values of k (N =95, n' =70 and n =50)

$R\hat{B}(.)$ $AM\hat{S}E(.)$	$1/k$		
	1/4	1/3	1/2
$R\hat{B}(\bar{y}^*)$ $AM\hat{S}E(\bar{y}^*)$	00.00 100 (0.4263)*	00.00 100 (0.3418)	00.00 100 (0.2482)
$R\hat{B}(T_1)$ $AM\hat{S}E(T_1)$	1.1000×10^{-4} 177 (0.2414)	9.2073×10^{-5} 171 (0.1999)	6.3361×10^{-5} 157 (0.1579)
$R\hat{B}(T_2)$ $AM\hat{S}E(T_2)$	3.5282×10^{-5} 121 (0.3528)	3.4978×10^{-5} 122 (0.2794)	3.4784×10^{-5} 126 (0.1974)
$R\hat{B}(t_1)$ $AM\hat{S}E(t_1)$	1.2000×10^{-4} 183 (0.2336)	9.8001×10^{-5} 178 (0.1923)	6.9112×10^{-5} 165(0.1504)
$R\hat{B}(t_2)$ $AM\hat{S}E(t_2)$	4.1314×10^{-5} 124 (0.3451)	4.0906×10^{-5} 126 (0.2718)	4.0535×10^{-5} 131 (0.1899)

**Figures in parenthesis give the* $AM\hat{S}E(.)$.

From table 4, we observe that for n'=70, n=50 and $k = 2,3,4$, simulated values of the estimators $(R\hat{B}(t_1),\ R\hat{B}(t_2))$ are more than the values of the corresponding estimators $(R\hat{B}(T_1),\ R\hat{B}(T_2))$. We also observe that the estimators $(AM\hat{S}E(t_1),\ AM\hat{S}E(t_2))$ are more efficient in comparison to the corresponding estimators $(AM\hat{S}E(T_1), AM\hat{S}E(T_2))$ and $AM\hat{S}E(\bar{y}^*)$. The values of all the estimators decrease as k^{-1} increases.

Table 5. Absolute bias of the estimators $R\hat{B}(.)$ and $AM\hat{S}E(.)$ for the fixed values of n', n and different values of k (N =95, n' =70 and n =50)

$\lvert R\hat{B}(.) - RB(.) \rvert$ $\lvert AM\hat{S}E(.) - AMSE(.) \rvert$	$1/k$		
	1/4	1/3	1/2
$\lvert R\hat{B}(\bar{y}^*) - RB(\bar{y}^*) \rvert$ $\lvert AM\hat{S}E(\bar{y}^*) - AMSE(\bar{y}^*) \rvert$	00.00 0.00829	00.00 0.00785	00.00 0.00173

Table 5. Absolute bias of the estimators $R\hat{B}(.)$ and $AM\hat{S}E(.)$ for the fixed values of n', n and different values of k (N =95, n' =70 and n =50) (cont.)

$\mid R\hat{B}(.) - RB(.) \mid$ $\mid AM\hat{S}E(.) - AMSE(.) \mid$	$1/k$		
	1/4	1/3	1/2
$\mid R\hat{B}(T_1) - RB(T_1) \mid$ $\mid AM\hat{S}E(T_1) - AMSE(T_1) \mid$	1.358×10^{-5} 0.00212	5.1711×10^{-6} 0.00117	3.0452×10^{-6} 0.00043
$\mid R\hat{B}(T_2) - RB(T_2) \mid$ $\mid AM\hat{S}E(T_2) - AMSE(T_2) \mid$	2.888×10^{-7} 0.01708	5.9268×10^{-7} 0.00650	7.857×10^{-7} 0.00452
$\mid R\hat{B}(t_1) - RB(t_1) \mid$ $\mid AM\hat{S}E(t_1) - AMSE(t_1) \mid$	1.562×10^{-6} 0.00199	6.7111×10^{-6} 0.00115	8.6591×10^{-6} 0.00027
$\mid R\hat{B}(t_2) - RB(t_2) \mid$ $\mid AM\hat{S}E(t_2) - AMSE(t_2) \mid$	1.1697×10^{-5} 0.01721	1.1289×10^{-5} 0.00652	1.0919×10^{-7} 0.00436

From table 5, we observe that the differences between the estimated value and the actual value of $R\hat{B}(.)$ and $AM\hat{S}E(.)$ based on simulation technique are very small and likely to be neglected.

Case 3: For the simulation study of the estimators $\bar{y}^*, T_1, T_2, t_1$ and t_2 when some fraction of r value of y and x variables is not observed in the last phase.

In this case, we assumed $r/2$ units observed in the last phase and repeat all process according to case 1.

Simulated absolute bias and simulated mean square error of $t_i = 1, 2$ are calculated as follows:

$$Absolute \quad Bias(t_i) = \frac{1}{1000} \left| \sum_{j=1}^{1000} t_{i\,j} - \bar{Y} \right|,$$

$$MSE(t_i) = \frac{1}{1000} \sum_{j=1}^{1000} \left(t_{i\,j} - \bar{Y} \right)^2$$

Table 6. Simulated absolute bias and Simulated relative efficiency (with respect to $\bar{y}^*$) of the estimators $\bar{y}^*, T_1, T_2, t_1$ and t_2 (when $r/2$ units are observed) for the fixed values of n', n and different values of k (N =95, n' =70 and n =50)

Estimators	$1/k$		
	1/4	1/3	1/2
$\bar{y}^*$	0.04580 100 (1.1896)*	0.02465 100 (0.6566)	0.00918 100 (0.4187)
T_1	0.02859 197 (0.6042)	0.02396 179 (0.3664)	0.00391 169 (0.2480)
T_2	0.04191 105 (1.1325)	0.02272 108 (0.6064)	0.00379 113 (0.3718)
t_1	0.02821 200 (0.5950)	0.02359 183 (0.3595)	0.00275 175 (0.2387)
t_2	0.04149 106 (1.1218)	0.02237 109 (0.6007)	0.00264 115 (0.3627)

* *Figures in parenthesis give the* $MSE(.)$.

From table 6, we observe that for the fixed value of n', n and $k = 2, 3, 4$, simulated absolute bias of the estimators (t_1, t_2) is less than the corresponding estimators (T_1, T_2) and $\bar{y}^*$. Simulated absolute bias of the estimators decreases as values of k^{-1} increase. We also observe that the estimators (t_1, t_2) have less mean square error in comparison to the corresponding estimators (T_1, T_2) and $\bar{y}^*$.

Table 7. Simulated absolute bias and Simulated relative efficiency (with respect to $\bar{y}^*$) of the estimators $\bar{y}^*, T_1, T_2, t_1$ and t_2 (when $r/2$ units are observed) for the fixed values of n', n and different values of k (N =95, n' =60 and n =40)

Estimators	$1/k$		
	1/4	1/3	1/2
$\bar{y}^*$	0.03782 100 (1.2869)*	0.02297 100 (1.1901)	0.02117 100 (0.6753)
T_1	0.02079 188 (0.6830)	0.01992 175 (0.6789)	0.01023 170 (0.3983)
T_2	0.03383 106 (1.2118)	0.02076 106 (1.1213)	0.01655 112 (0.6024)
t_1	0.02069 191 (0.6729)	0.01963 179 (0.6661)	0.00827 174 (0.3877)
t_2	0.03369 107 (1.2003)	0.02049 107(1.1071)	0.01460 114 (0.5920)

* *Figures in parenthesis give the* $MSE(.)$.

From table 7, we observe that for the fixed value of n', n and $k = 2, 3, 4$, simulated absolute bias of the estimators (t_1, t_2) is less than the corresponding estimators (T_1, T_2) and $\bar{y}^*$. Simulated absolute bias of the estimators decreases as values of $1/k$ increase. We also observe that the estimators (t_1, t_2) have less mean square error in comparison to the corresponding estimators (T_1, T_2) and $\bar{y}^*$. Mean square error of all the estimators decreases as k^{-1} increases.

Case 4: For the simulation study of the estimators $R\hat{B}(.)$ and $AM\hat{S}E(.)$ when some fraction of r value of y and x variables is not observed in the last phase.

In this case, we assumed $r/2$ units observed in the last phase and repeat all process according to case 2.

Table 8. Simulated absolute value of the estimator $R\hat{B}(.)$ and Simulated relative efficiency (with respect to $AM\hat{S}E(\bar{y}^*)$) of the estimator $AM\hat{S}E(.)$ for the fixed values of n', n and different values of k and $r/2$ units is observed (N =*95*, n' =*70* and n =*50*)

$R\hat{B}(.)$ $AM\hat{S}E(.)$	$1/k$		
	1/4	1/3	1/2
$R\hat{B}(\bar{y}^*)$ $AM\hat{S}E(\bar{y}^*)$	00.00 100 (0.4482)*	00.00 100 (0.3379)	00.00 100 (0.2524)
$R\hat{B}(T_1)$ $AM\hat{S}E(T_1)$	1.2000×10^{-4} 185 (0.2428)	8.7005×10^{-5} 173 (0.1955)	6.2294×10^{-5} 159 (0.1585)
$R\hat{B}(T_2)$ $AM\hat{S}E(T_2)$	3.5567×10^{-5} 125 (0.3585)	3.5368×10^{-5} 126 (0.2689)	3.5054×10^{-5} 128 (0.1967)
$R\hat{B}(t_1)$ $AM\hat{S}E(t_1)$	1.2300×10^{-4} 191 (0.2352)	9.2965×10^{-5} 180 (0.1879)	6.8145×10^{-5} 167 (0.1509)
$R\hat{B}(t_2)$ $AM\hat{S}E(t_2)$	4.1547×10^{-5} 128 (0.3508)	4.1329×10^{-5} 129 (0.2612)	4.0904×10^{-5} 133 (0.1891)

**Figures in parenthesis give the* $AM\hat{S}E(.)$.

From table 8, we observe that for n'=70, n=50 and $k = 2,3,4$ simulated absolute value of the estimators $(R\hat{B}(t_1), R\hat{B}(t_2))$ is more than the corresponding estimators $(R\hat{B}(T_1),\ R\hat{B}(T_2))$. We also observe that the estimators $(AM\hat{S}E(t_1)$, $AM\hat{S}E(t_2))$ are more efficient in comparison to the corresponding estimators $(AM\hat{S}E(T_1), AM\hat{S}E(T_2))$ and $AM\hat{S}E(\bar{y}^*)$. The values of all the estimators decrease as k^{-1} increases.

Table 9. Absolute bias of the estimators $R\hat{B}(.)$ and $AM\hat{S}E(.)$ for the fixed values of n', n and different values of k and $r/2$ units is observed (N =*95*, n' =*70* and n =*50*)

$\mid R\hat{B}(.)-RB(.)\mid$ $\mid AM\hat{S}E(.)-AMSE(.)\mid$	$1/k$		
	1/4	1/3	1/2
$\mid R\hat{B}(\bar{y}^*)-RB(\bar{y}^*)\mid$ $\mid AM\hat{S}E(\bar{y}^*)-AMSE(\bar{y}^*)\mid$	00.00 0.0302	00.00 0.0039	00.00 0.0025
$\mid R\hat{B}(T_1)-RB(T_1)\mid$ $\mid AM\hat{S}E(T_1)-AMSE(T_1)\mid$	1.0891×10^{-5} 0.0036	1.0239×10^{-5} 0.0032	3.4624×10^{-6} 0.0002
$\mid R\hat{B}(T_2)-RB(T_2)\mid$ $\mid AM\hat{S}E(T_2)-AMSE(T_2)\mid$	2.9000×10^{-5} 0.0114	2.0185×10^{-7} 0.0171	5.1701×10^{-7} 0.0052
$\mid R\hat{B}(t_1)-RB(t_1)\mid$ $\mid AM\hat{S}E(t_1)-AMSE(t_1)\mid$	1.0695×10^{-5} 0.0035	1.6750×10^{-6} 0.0033	7.6916×10^{-6} 0.0002
$\mid R\hat{B}(t_2)-RB(t_2)\mid$ $\mid AM\hat{S}E(t_2)-AMSE(t_2)\mid$	1.1931×10^{-5} 0.0115	1.1712×10^{-5} 0.0171	1.1287×10^{-5} 0.0052

From table 9, we observe that the values of the estimator $R\hat{B}(.)$ based on simulation study are almost close to the $RB(.)$ of the estimators T_1, T_2, t_1 and t_2 . We also observe that the values of the estimator $AM\hat{S}E(.)$ based on simulation study are almost close to the $AMSE(.)$ of the estimators T_1, T_2, t_1 and t_2 based on empirical study.

Table 10. Relative efficiency of the estimators $\bar{y}^*, T_1, T_2, t_1$ and t_2 with respect to $\bar{y}^*$ (for the fixed cost $C \le C_0$ =*Rs.100*, c_1' =*Rs. 0.50*, c_2' =*Rs. 0.10*, c_1 =*Rs. 1*, c_2 =*Rs. 2*, c_3 =*Rs. 9*)

Estimators	k_{opt}	n'_{opt}	n_{opt}	R.E. (.) in %
$\bar{y}^*$	1.70	-	26	100 (0.7839) *
T_1	1.79	49	20	111 (0.7054)
T_2	1.03	47	16	105 (0.7464)
t_1	1.79	39	20	114 (0.6897)
t_2	1.03	38	16	107 (0.7302)

**Figures in parenthesis give the* $AMSE(.)$ *, Rs: Rupees (Indian currency)*

From table 10, we observe that for the fixed cost the estimators t_1 and t_2 have smaller approximated mean square error than that of $\bar{y}^*$ and (t_1, t_2) are also more efficient than the corresponding estimators (T_1, T_2). The estimator t_1 is more efficient than t_2.

Table 11. Expected cost of the estimators $\bar{y}^*, T_1, T_2, t_1$ and t_2 for the specified variance $V_0' = 0.4897$: (c_1' =*Rs. 0.50,* c_2' =*Rs. 0.10,* c_1 =*Rs. 1,* c_2 =*Rs. 2,* c_3 =*Rs. 9*)

Estimators	k_{opt}	n'_{opt}	n_{opt}	Expected cost (Rs.)
$\bar{y}^*$	1.70	-	42	160
T_1	1.79	70	29	144
T_2	1.03	72	25	152
t_1	1.79	55	29	141
t_2	1.03	57	24	149

From table 11, we observe that for the specified variance the expected cost is minimum for t_1 and t_2 with corresponding cost for the estimators T_1 and T_2 and also $\bar{y}^*$. The estimator t_1 has also less cost in comparison to t_2.

So, we conclude that the use of information on the additional auxiliary variable is useful in increasing the precision of the proposed estimators with respect to the relevant estimators for the fixed sample size (n', n) based on empirical study and also on Monte Carlo simulation study. The results derived from Monte Carlo simulation study based on empirical data were also found to be in congruence to the results based on empirical study. On the basis of empirical study, the use of additional variable in the proposed estimators for population mean in the presence of non-response is also found to be more useful in increasing the precision of the proposed estimators with respect to the relevant estimators for the fixed cost $C \le C_0$. For the specified variance, the total cost for the proposed estimators is also less than the corresponding relevant estimators.

Acknowledgement

The authors are highly grateful to the referee for his encouraging suggestions which were very much useful in improving the quality of the paper.

REFERENCES

BOWELY, A. L. (1926). Measurements of precision attained in sampling. Bull. Inter. Statist. Inst., 22: 1-62.

CHAND, L. (1975). Some ratio-type estimators based on two or more auxiliary variables. Ph.D. Thesis submitted to Iowa State University, Ames, IOWA.

COCHRAN, W. G. (1940). The estimation of the yields of the cereal experiments by sampling for the ratio of grain to total produce. Jour. Agri. Science, 30: 262-275.

COCHRAN, W. G. (1942). Sampling theory when the sampling units are of unequal sizes. Jour. Amer. Statist. Assoc., 37: 199-212.

HANSEN, M. H. and HURWITZ, W. N. (1946). The problem of non-response in sample surveys. Jour. Amer. Statist. Assoc., 41: 517-529.

HANSEN, M. H., HURWITZ, W. N. and MADOW, W. G. (1953). Sample Survey methods and theory Vol. 2: John Wiley and Sons, New York.

KHARE, B. B. and SRIVASTAVA, S. (1993). Estimation of population mean using auxiliary character in presence of non-response. Nat. Acad. Sc. Letters, India, 16(3): 111-114.

KHARE, B. B. and SRIVASTAVA, S. (1995). Study of conventional and alternative two phase sampling ratio, product and regression estimators in presence of nonresponse. Proc. Nat. Acad. Sci., India, 65(A) II: 195-203.

KHARE, B. B. and SINHA, R. R. (2007). Estimation of the ratio of the two population means using multi-auxiliary characters in the presence of non-response. Statistical Techniques in Life Testing, Reliability, Sampling Theory and Quality Control, 163-171.

KIREGYERA, B. (1980). A chain ratio type estimator in finite population double sampling using two auxiliary variables. Metrika, 27: 217-223.

KIREGYERA, B. (1984). Regression-type estimators using two auxiliary variables and the model of double sampling from finite populations. Metrika, 31: 215-226.

LITTLE, R. J. A. and RUBIN, D. B. (2002). Statistical Analysis with missing Data, New York: Wiley.

NEYMAN, J. (1934). On the two different aspects of representative method, the method of stratified sampling and the method of purposive selection. Jour. Roy. Statist. Soc., 97: 558-625.

NEYMAN, J. (1938). Contributions to the theory of sampling human population. J. Amer. Statist Assoc., 33: 101-116.

RAO, C. R. and TOUTENBURG , H. (1995). Linear Models: Least Squares and Alternatives. New York: Springer.

RAO, P. S. R. S. (1986). Ratio estimation with subsampling the nonrespondents. Survey Methodology, 12(2): 217-230.

RAO, P. S. R. S. (1987). Ratio and regression Estimates with Subsampling the non-respondents. Paper presented at a special contributed session of the International Statistical Association Meetings, September, 2-16, Tokyo, Japan.

ROBSON, D. S. (1957). Applications of multivariate polykays to the theory of unbiased ratio-type estimators. Jour. Amer. Statist. Assoc., 52: 511-522.

SRIVASTAVA, S. RANI, KHARE, B. B. and SRIVASTAVA, S. R. (1990). A generalised chain ratio estimator for mean of finite population. Jour. Ind. Soc. Ag. Statistics, 42(1): 108-117.

TOUTENBURG, H. and SRIVASTAVA, V. K. (1998). Estimation of ratio of population means in survey sampling when some observations are missing. Metrika, 48: 177-187.

TOUTENBURG, H. and SRIVASTAVA, V. K. (2003). Efficient estimation of population mean using incomplete survey data on study and auxiliary characteristics. Statistica, LXIII, 2: 2003.

WATSON, D. J. (1937). The estimation of leaf areas. Jour. Agril. Sci., 27: 474-481.

APPENDIX

Relative Bias and Approximated Mean Square Error (AMSE)

Let $\bar{y}^* = \bar{Y}(1+\varepsilon_0)$, $\bar{x}^* = \bar{X}(1+\varepsilon_1)$, $\bar{x} = \bar{X}(1+\varepsilon_2)$, $\bar{x}' = \bar{X}(1+\varepsilon_3)$, $\bar{z}' = \bar{Z}(1+\varepsilon_4)$

$E(\varepsilon_i) = 0$ and $|\varepsilon_i| < 1 \ \forall$ i=0,1,2,3,4.

$$E(\varepsilon_0^2) = \frac{f}{n}C_y^2 + \frac{W_2(k-1)}{n}C_{y(2)}^2,\ E(\varepsilon_1^2) = \frac{f}{n}C_x^2 + \frac{W_2(k-1)}{n}C_{x(2)}^2,\ E(\varepsilon_2^2) = \frac{f}{n}C_x^2$$

$$E(\varepsilon_3^2) = \frac{f'}{n'}C_x^2,\quad E(\varepsilon_4^2) = \frac{f'}{n'}C_z^2,\quad E(\varepsilon_0\varepsilon_1) = \frac{f}{n}C_{yx} + \frac{W_2(k-1)}{n}C_{yx(2)},$$

$$E(\varepsilon_0\varepsilon_2) = \frac{f}{n}C_{yx},\ E(\varepsilon_0\varepsilon_3) = \frac{f'}{n'}C_{yx},\ E(\varepsilon_0\varepsilon_4) = \frac{f'}{n'}C_{yz},\ E(\varepsilon_1\varepsilon_3) = \frac{f'}{n'}C_x^2,$$

$$E(\varepsilon_1\varepsilon_4) = \frac{f'}{n'}C_{xz},\ E(\varepsilon_2\varepsilon_3) = \frac{f'}{n'}C_x^2,\ E(\varepsilon_2\varepsilon_4) = \frac{f'}{n'}C_{xz},$$

$$t_1 = \frac{\bar{y}^*}{\bar{x}^*}\frac{\bar{x}'}{\bar{z}'}\bar{Z} = \frac{\bar{Y}(1+\varepsilon_0)}{\bar{X}(1+\varepsilon_1)}\frac{\bar{X}(1+\varepsilon_3)}{\bar{Z}(1+\varepsilon_4)}\bar{Z}$$

$$t_1 = \bar{Y}(1+\varepsilon_0)(1+\varepsilon_3)(1+\varepsilon_1)^{-1}(1+\varepsilon_4)^{-1}$$

$$t_1 = \bar{Y}(1+\varepsilon_0)(1+\varepsilon_3)(1-\varepsilon_1+\varepsilon_1^2 \ldots\ldots\ldots\ldots)(1-\varepsilon_4+\varepsilon_4^2 \ldots\ldots\ldots\ldots\ldots\ldots)$$

After neglecting the terms of ε_i up to the second degree, we have

$$t_1 = \bar{Y}[1+\varepsilon_0-\varepsilon_1+\varepsilon_3-\varepsilon_4-\varepsilon_0\varepsilon_1+\varepsilon_0\varepsilon_3-\varepsilon_0\varepsilon_4-\varepsilon_1\varepsilon_3+\varepsilon_1\varepsilon_4-\varepsilon_3\varepsilon_4+\varepsilon_1^2+\varepsilon_4^2] \tag{A.1}$$

$$t_2 = \frac{\bar{y}^*}{\bar{x}}\frac{\bar{x}'}{\bar{z}'}\bar{Z} = \frac{\bar{Y}(1+\varepsilon_0)}{\bar{X}(1+\varepsilon_2)}\frac{\bar{X}(1+\varepsilon_3)}{\bar{Z}(1+\varepsilon_4)}\bar{Z}$$

$$t_2 = \bar{Y}(1+\varepsilon_0)(1+\varepsilon_3)(1+\varepsilon_2)^{-1}(1+\varepsilon_4)^{-1}$$

$$t_2 = \bar{Y}(1+\varepsilon_0)(1+\varepsilon_3)(1-\varepsilon_2+\varepsilon_2^2 \ldots\ldots\ldots\ldots)(1-\varepsilon_4+\varepsilon_4^2 \ldots\ldots\ldots\ldots\ldots\ldots)$$

After neglecting the terms of ε_i up to the second degree, we have

$$t_2 = \bar{Y}[1+\varepsilon_0-\varepsilon_2+\varepsilon_3-\varepsilon_4-\varepsilon_0\varepsilon_2+\varepsilon_0\varepsilon_3-\varepsilon_0\varepsilon_4-\varepsilon_2\varepsilon_3+\varepsilon_2\varepsilon_4-\varepsilon_3\varepsilon_4+\varepsilon_2^2+\varepsilon_4^2] \tag{A.2}$$

1. Relative bias of the estimators t_1 and t_2

$$RB(t_1) = \frac{E(t_1)-\bar{Y}}{\bar{Y}}$$

$$= E[\bar{Y}+\bar{Y}(\varepsilon_0-\varepsilon_1+\varepsilon_3-\varepsilon_4-\varepsilon_0\varepsilon_1+\varepsilon_0\varepsilon_3-\varepsilon_0\varepsilon_4-\varepsilon_1\varepsilon_3+\varepsilon_1\varepsilon_4-\varepsilon_3\varepsilon_4+\varepsilon_1^2+\varepsilon_4^2) - \bar{Y}]/\bar{Y}$$

$$= -E(\varepsilon_0\varepsilon_1)+E(\varepsilon_0\varepsilon_3)-E(\varepsilon_0\varepsilon_4)-E(\varepsilon_1\varepsilon_3)+E(\varepsilon_1\varepsilon_4)-E(\varepsilon_3\varepsilon_4)+E(\varepsilon_1^2)+E(\varepsilon_4^2)$$

$$RB(t_1)=\left(\frac{1}{n}-\frac{1}{n'}\right)(C_x^2-\rho_{yx}C_yC_x)+\frac{W_2(k-1)}{n}(C_{x(2)}^2-\rho_{yx(2)}C_{y(2)}C_{x(2)})+$$

$$\frac{f'}{n'}(C_z^2-\rho_{yz}C_yC_z) \qquad \text{(A.3)}$$

$$RB(t_2)=\frac{E(t_2)-\bar{Y}}{\bar{Y}}$$

$$=E[\bar{Y}+\bar{Y}(\varepsilon_0-\varepsilon_2+\varepsilon_3-\varepsilon_4-\varepsilon_0\varepsilon_2+\varepsilon_0\varepsilon_3-\varepsilon_0\varepsilon_4-\varepsilon_2\varepsilon_3+\varepsilon_2\varepsilon_4-\varepsilon_3\varepsilon_4+\varepsilon_2^2+\varepsilon_4^2)-\bar{Y}]/\bar{Y}$$

$$=-E(\varepsilon_0\varepsilon_2)+E(\varepsilon_0\varepsilon_3)-E(\varepsilon_0\varepsilon_4)-E(\varepsilon_2\varepsilon_3)+E(\varepsilon_2\varepsilon_4)-E(\varepsilon_3\varepsilon_4)+E(\varepsilon_2^2)+E(\varepsilon_4^2)$$

$$RB(t_2)=\left(\frac{1}{n}-\frac{1}{n'}\right)(C_x^2-\rho_{yx}C_yC_x)+\frac{f'}{n'}(C_z^2-\rho_{yz}C_yC_z) \qquad \text{(A.4)}$$

2. Approximated Mean Square Error (AMSE) of the estimators t_1 and t_2

$$MSE(t_1)=E[t_1-\bar{Y}]^2$$

$$=E[\bar{Y}+\bar{Y}(\varepsilon_0-\varepsilon_1+\varepsilon_3-\varepsilon_4..........)-\bar{Y}]^2$$

$$=\bar{Y}^2E[\varepsilon_0^2+\varepsilon_1^2+\varepsilon_3^2+\varepsilon_4^2-2\varepsilon_0\varepsilon_1+2\varepsilon_0\varepsilon_3-2\varepsilon_0\varepsilon_4-2\varepsilon_1\varepsilon_3+2\varepsilon_1\varepsilon_4-2\varepsilon_3\varepsilon_4.....]$$

$$=\bar{Y}^2[E(\varepsilon_0^2)+E(\varepsilon_1^2)+E(\varepsilon_3^2)+E(\varepsilon_4^2)-2E(\varepsilon_0\varepsilon_1)+2E(\varepsilon_0\varepsilon_3)-2E(\varepsilon_0\varepsilon_4)$$

$$-2E(\varepsilon_1\varepsilon_3)+2E(\varepsilon_1\varepsilon_4)-2E(\varepsilon_3\varepsilon_4)..........]$$

$$AMSE(t_1)=\bar{Y}^2\left[\left(\frac{1}{n}-\frac{1}{n'}\right)\left\{C_y^2+C_x^2-2\rho_{yx}C_yC_x\right\}+\frac{W_2(k-1)}{n}\left\{C_{y(2)}^2+C_{x(2)}^2-\right.\right.$$

$$\left.\left.-2\rho_{yx(2)}C_{y(2)}C_{x(2)}\right\}+\frac{f'}{n'}(C_y^2+C_z^2-2\rho_{yz}C_yC_z)\right] \qquad \text{(A.5)}$$

$$MSE(t_2)=E[t_2-\bar{Y}]^2$$

$$=E[\bar{Y}+\bar{Y}(\varepsilon_0-\varepsilon_2+\varepsilon_3-\varepsilon_4..........)-\bar{Y}]^2$$

$$=\bar{Y}^2E[\varepsilon_0^2+\varepsilon_2^2+\varepsilon_3^2+\varepsilon_4^2-2\varepsilon_0\varepsilon_2+2\varepsilon_0\varepsilon_3-2\varepsilon_0\varepsilon_4-2\varepsilon_2\varepsilon_3+2\varepsilon_2\varepsilon_4-2\varepsilon_3\varepsilon_4.....]$$

$$=\bar{Y}^2[E(\varepsilon_0^2)+E(\varepsilon_2^2)+E(\varepsilon_3^2)+E(\varepsilon_4^2)-2E(\varepsilon_0\varepsilon_2)+2E(\varepsilon_0\varepsilon_3)-2E(\varepsilon_0\varepsilon_4)$$

$$-2E(\varepsilon_2\varepsilon_3)+2E(\varepsilon_2\varepsilon_4)-2E(\varepsilon_3\varepsilon_4)..........]$$

$$AMSE(t_2)=\bar{Y}^2\left[\left(\frac{1}{n}-\frac{1}{n'}\right)\left\{C_y^2+C_x^2-2\rho_{yx}C_yC_x\right\}+\frac{W_2(k-1)}{n}C_{y(2)}^2\right.$$

$$\left.+\frac{f'}{n'}(C_y^2+C_z^2-2\rho_{yz}C_yC_z)\right]. \qquad \text{(A.6)}$$