ProbStat Forum, Volume 08, July 2015, Pages 95–102
ISSN 0974-3235

ProbStat Forum is an e-journal.
For details please visit www.probstat.org.in

Chain ratio-to-regression estimators in two-phase sampling in the presence of non-response

B.B. Khare[a], **Sanjay Kumar**[b]

[a] *Department of statistics, Banaras Hindu University Varanasi (U.P.)*-221005, *India*
[b] *Department of Statistics, Central University of Rajasthan Kishangarh*-305801, *Ajmer, Rajasthan*

Abstract. Using double sampling, this paper presents conventional and alternative chain ratio-to-regression estimators for the finite population mean when the population of the main auxiliary variable x is unknown but that of an additional auxiliary variable z is known. The proposed estimators are found to be more efficient than the relevant estimators in case of fixed cost. An empirical study has been made for the support of the problem.

1. Introduction

Information on variables correlated with the main variable under study is popularly known as auxiliary information which may be utilized either at planning stage or at design stage or at the information stage to arrive at improved estimator compared to those, not utilizing auxiliary information. Use of auxiliary information for forming ratio and regression method of estimation were introduced during 1930s with a comprehensive theory provided by Cochran (1942). Chaed (1975) and Kiogyera (1980, 1984) have proposed chain ratio type estimators using additional variable with known population mean. If the population mean of the auxiliary variable (x) is not know 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. $\rho_{yx} > \rho_{yz}$) sometimes it may not be possible to collect the compete information for all the units selected in the sample due to non-response. Estimation of the population mean in sample surveys when some observations are missing due to non-response not at random has been considered by Hansen and Hurwitz (1946) and Rao (1986, 1987).

In this paper, we have proposed conventional and alternative ratio-to-regression estimators for the population mean of the study variable in the presence of non-response. The expressions for mean square error of the proposed estimators are obtained and a comparison of the proposed estimators has been made with the relevant estimators.

2. The estimators

Let X_i, Y_i and Z_i be the non negative values for the i^{th} unit of the population $U = U_1, U_2, \ldots, U_N$ on the study variable y, the auxiliary variable x and the additional auxiliary variable z with their population means $\overline{Y}$, $\overline{X}$ and $\overline{Z}$. Here $\overline{X}$ is unknown, but $\overline{Z}$, the population mean of additional auxiliary variable z (closely

2010 *Mathematics Subject Classification.* Primary 62D05
Keywords. Chain ratio-to-regression estimator, Population mean, Study variable, Auxiliary variable, Additional variable.
Received: 13 August 2014; Accepted: 29 June 2015
Email addresses: bbkhare56@yahoo.com (B.B. Khare), sanjay.kumar@curaj.ac.in (Sanjay Kumar)

related to x) is known, which may be cheaper and less correlated to the study variable (z) in comparison to the main auxiliary variable (x). Now a first phase sample of size n' is selected from the population of size N using SRSWOR and we estimate the population mean $\overline{X}$ using additional auxiliary character with known population mean $\overline{Z}$ and n' observations on x and z. Again subsample of size n is selected from n' first phase sample by using SRSWOR scheme of sampling 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 $r = \frac{n_2}{K}$ $(K > 1)$ from n_2 non-responding units has been drawn by using SRSWOR method of sampling by making extra effort.

In some situations, $\overline{X}$ is unknown, then we select a larger sample of size $n'(> n)$ from the population of size N by using SRSWOR method of sampling and we estimate $\overline{X}$ by $\overline{x}' = \frac{1}{n'}\sum_{i=1}^{n'} x_i$ and again draw a sub sample of size n and observe y variable. Now, it has been observed that n_1 units respond and n_2 units do not respond and then we select a sub-sample of size $r = \frac{n_2}{K}$ $(K > 1)$ from n_2 non responding units by using SRSWOR method of sampling and consequently the conventional and alternative two phase sampling ratio (t_1, t_2) for population mean in the presence of non-response proposed by Khare and Srivastava (1993) are given as follows:

$$t_1 = \frac{\overline{y}^*}{\overline{x}^*}\,\overline{x}' \text{ and } t_2 = \frac{\overline{y}^*}{\overline{x}}\overline{x}',$$

where,

$$\overline{y}^* = \frac{n_1}{n}\overline{y}_1 + \frac{n_2}{n}\overline{y}_2',\ \overline{x}^* = \frac{n_1}{n}\overline{x}_1 + \frac{n_2}{n}\overline{x}_2',\ \overline{x} = \sum_{i=1}^{n}\frac{x_i}{n},$$

The estimator $\overline{y}^*$ is unbiased and the $V(\overline{y}^*)$ is given by

$$V(\overline{y}^*) = \frac{f}{n}S_y^2 + \frac{W_2(K-1)}{n}S_{y(2)}^2$$

where, $f = \frac{N-n}{N}$, $W_i = \frac{N_i}{N}\,(i = 1, 2)$, S_y^2 and $S_{y(2)}^2$ are the population mean squares of the character y for the whole population and for the non responding part of the population.

If $\overline{X}$ is not known, but $\overline{Z}$, the population mean of the additional auxiliary character z (closely related to x) is known, which may be cheaper and less correlated to the study character(y) in comparison to main auxiliary character (x). In this situation, we take a first phase sample $n'$from the population of size N with SRSWOR and we estimate $\overline{X}$ by $\overline{x}' = \frac{1}{n'}\sum_{i=1}^{n'} x_i$ by using the sample means $\overline{x}'$ and $\overline{z}' = \frac{1}{n'}\sum_{i=1}^{n'} z_i$ based on n' units and the known additional population mean $\overline{Z}$.

It is well known result that ratio type estimator of $\overline{X}$ using auxiliary information z is optimum if the regression of x on z is linear and passes through the origin and the variance of x about this regression line is proportional to z. However, the regression of x on z is linear but does not go through the origin. In such situation, Kiogyera (1980) suggested a chain ratio-to-regression (e_1) estimator, which is given as follows:

$$e_1 = \frac{\overline{y}}{\overline{x}}\left\{\overline{x}' + b_2(\overline{Z} - \overline{z}')\right\}$$

where $b_2 = \frac{\hat{S}_{xz}}{s_z^2}$, $S_{xz} = \frac{1}{N-1}\sum_{i=1}^{N}\left(Z_i - \overline{Z}\right)(X_i - \overline{X})$ and$s_z^2 = \frac{1}{n-1}\sum_{i=1}^{n}(z_i - \overline{z})^2$. The estimate $\hat{S}_{xz}$ is based on the available data under the given sampling design.

Singh and Kumar (2010) proposed improved estimators of population mean under two-phase sampling with sub-sampling the non-respondents, which are given as follows:

$$t_1' = \{\overline{y}^* + b_{yx}(\overline{x}' - \overline{x}^*)\}\,\frac{\overline{Z}}{\left\{\overline{Z} + \delta_1(\overline{z}^* - \overline{Z})\right\}} \text{ and } t'_2 = \{\overline{y}^* + b_{yx}(\overline{x}' - \overline{x})\}\,\frac{\overline{Z}}{\left\{\overline{Z} + \delta_1'(\overline{z} - \overline{Z})\right\}},$$

where

$$\delta_{1(opt.)} = \left\{ \frac{M^*}{\left(\frac{\overline{Y}}{\overline{Z}} D^*\right)} \right\},\ \delta'_{1(opt.)} = \left[\frac{Q^*}{\left\{ \left(\frac{1}{n} - \frac{1}{N}\right) \frac{\overline{Y}}{\overline{Z}} \right\}} \right],$$

$$D^* = \left\{ \left(\frac{1}{n} - \frac{1}{N}\right) S_z^2 + \frac{W_2(K-1)}{n} S_{z(2)}^2 \right\},$$

$$M^* = \left\{ \left(\frac{1}{n'} - \frac{1}{N}\right) \beta_1 S_z^2 + \left(\frac{1}{n} - \frac{1}{n'}\right) A^* S_z^2 + \frac{W_2(K-1)}{n} B^* S_{z(2)}^2 \right\},$$

$$A^* = (\beta_1 - \beta\beta_2),\ B^* = (\beta_3 - \beta\beta_4),\ Q^* = \left\{ \left(\frac{1}{n} - \frac{1}{n'}\right) A^* + \left(\frac{1}{n'} - \frac{1}{N}\right) \beta_1 \right\},$$

$$\beta = \frac{S_{yx}}{S_x^2}, \beta_1 = \frac{S_{yz}}{S_z^2}\ \ \beta_2 = \frac{S_{xz}}{S_z^2},\ \beta_3 = \frac{S_{yz(2)}}{S_{z(2)}^2},\ \beta_4 = \frac{S_{xz(2)}}{S_{z(2)}^2},\ \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i,$$

$$b_{yx} = \frac{\hat{S}_{yx}}{\hat{S}_x^2},\ \ b'_{yx} = \frac{\hat{S}_{yx}}{s_x^2}, s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2, S_z^2 = \frac{1}{N-1}\sum_{i=1}^{n}(Z_i - \overline{Z})^2,$$

$$S_{xz} = \frac{1}{N-1}\sum_{i=1}^{N}\left\{(X_i - \overline{X})(Z_i - \overline{Z})\right\}, S_x^2 = \frac{1}{N-1}\sum_{i=1}^{n}(X_i - \overline{X})^2,$$

$$S_{yx} = \frac{1}{N-1}\sum_{i=1}^{N}\left\{(Y_i - \overline{Y})(X_i - \overline{X})\right\}, S_{yz} = \frac{1}{N-1}\sum_{i=1}^{N}\left\{(Y_i - \overline{Y})(Z_i - \overline{Z})\right\}$$

and $(S_{z(2)}^2, S_{yz(2)},\ S_{xz(2)})$ denotes population mean squares of the character z, the covariance (y, z), covariance (x, z) for the non-response group of the population.

The expressions for the $MSE(t_i)$ and $MSE\,(t'_1)_{min.}$ (i=1,2) up to the terms of order n^{-1} and $n'^{-1}$are given as follows:

$$\begin{aligned} MSE\,(t_1) = &\ \left(\tfrac{1}{n} - \tfrac{1}{n'}\right)\left[S_y^2 + R^2 S_x^2 - 2R\rho_{yx} S_y S_x\right] + \left(\tfrac{1}{n'} - \tfrac{1}{N}\right) S_y^2 \\ &+ \tfrac{W_2(K-1)}{n}\left[S_{y(2)}^2 + R^2 S_{x(2)}^2 - 2R\rho_{yx(2)} S_{y(2)} S_{x(2)}\right], \\ MSE\,(t_2) = &\ \left(\tfrac{1}{n} - \tfrac{1}{n'}\right)\left[S_y^2 + R^2 S_x^2 - 2R\rho_{yx} S_y S_x\right] + \left(\tfrac{1}{n'} - \tfrac{1}{N}\right) S_y^2 \\ &+ \tfrac{W_2(K-1)}{n} S_{y(2)}^2, \end{aligned}$$

$$\begin{aligned} MSE\,(t'_1)_{min.} &= MSE\,(t_1) - \frac{M^{*2}}{D^*}, \\ MSE\,(t'_2)_{min.} &= MSE\,(t_2) - \left\{ \frac{Q^{*2}}{\left(\frac{1}{n} - \frac{1}{N}\right)} \right\} S_z^2, \end{aligned}$$

where, $\rho_{yx} = \frac{S_{yx}}{S_y S_x}$, $\rho_{yx(2)} = \frac{S_{yx(2)}}{S_{y(2)} S_{x(2)}}$, $\rho_{yz} = \frac{S_{yz}}{S_y S_z}$, $R_1 = \frac{\overline{Y}}{\overline{Z}}$, $R_2 = \frac{\overline{X}}{\overline{Z}}$, $S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \overline{Y})^2$, $\beta = \frac{S_{yx}}{S_x{}^2}$, $R = \frac{\overline{Y}}{\overline{X}}$ and $(S_{yx(2)},\ \rho_{yx(2)})$ denotes the covariance and correlation between y and x variables for the non-response group of the population.

3. The proposed estimators and their mean square error (MSE)

Following the estimator in Kiogyera (1980), we propose the conventional and the alternative chain ratio-to-regression estimators $(T_1,\ T_2)$ for population mean $\overline{Y}$ using two phase sampling in the presence of non-response, which are given as follows:

$$T_1 = \frac{\overline{y}^*}{\overline{x}^*}\left\{\overline{x}' + b_3(\overline{Z} - \overline{z}')\right\} \quad \text{and} \quad T_2 = \frac{\overline{y}^*}{\overline{x}}\left\{\overline{x}' + b_4(\overline{Z} - \overline{z}')\right\}$$

where $b_3 = \frac{\hat{S}_{xz}}{\hat{S}_z^2}, b_4 = \frac{\hat{S}_{xz}}{s_z^2}, S_{xz} = \frac{1}{N-1}\sum_{i=1}^{N}\left(Z_i - \overline{Z}\right)\left(X_i - \overline{X}\right), S_z^2 = \frac{1}{N-1}\sum_{i=1}^{n}\left(Z_i - \overline{Z}\right)^2$

The estimate $\hat{S}_z^2$ are based on the available data under the given sampling design.

The expressions for the MSE of the proposed estimators $T_i,\ (i = 1, 2)$ up to the terms of order n^{-1} and $n'^{-1}$are given as follows:

$$\begin{aligned} MSE\,(T_1) &= MSE\,(t_1) + \left(\frac{1}{n'} - \frac{1}{N}\right) R\rho_{xz}S_x(R\rho_{xz}S_x - 2\rho_{yz}S_y), \\ MSE\,(T_2) &= MSE\,(2) + \left(\frac{1}{n'} - \frac{1}{N}\right) R\rho_{xz}S_x(R\rho_{xz}S_x - 2\rho_{yz}S_y). \end{aligned}$$

Many estimators turn out as special cases of T_i, i=1, 2 which are given as follows:

1. T_1 reduces to the estimators t_1 for $b_3 = 0$.
2. T_2 reduces to the estimators t_2 for $b_4 = 0$.

4. Comparison of the proposed estimators with the relevant estimators

On comparing the $MSE\,(T_i)$, with $MSE(\overline{y}^*)$, $MSE(t_i)$ and $MSE\,(t_i')_{min.}$ $(i = 1, 2).$ we see that

$$MSE\,(T_1) < MSE\,(t_1) < MSE(\overline{y}^*) \quad \text{if } \rho_{yx} > \frac{1}{2}\frac{C_x}{C_y},\ \rho_{yx(2)} > \frac{1}{2}\frac{R}{R_2}\frac{C_{x(2)}}{C_{y(2)}} \text{ and } \rho_{yz} > \frac{1}{2}\rho_{xz}\frac{C_x}{C_y},$$

$$MSE\,(T_2) < MSE\,(t_2) < MSE(\overline{y}^*) \quad \text{if} \quad \rho_{yx} > \frac{1}{2}\frac{C_x}{C_y} \quad \text{and} \quad \rho_{yz} > \frac{1}{2}\rho_{xz}\frac{C_x}{C_y}$$

$$MSE\,(T_1) < MSE\,(T_2) \quad \text{if} \quad \rho_{yx(2)} > \frac{1}{2}\frac{R}{R_2}\frac{C_{x(2)}}{C_{y(2)}}$$

$$MSE\,(T_1) < MSE\,(t_1') \quad \text{if} \quad \frac{M^{*2}}{D^*} + \left(\frac{1}{n} - \frac{1}{n'}\right)E^* + \left(\frac{1}{n'} - \frac{1}{N}\right)F^* + \frac{W_2(K-1)}{n}G^* < 0$$

$$MSE\,(T_2) < MSE\,(t_2') \quad \text{if} \quad \left\{\frac{M^{*2}}{\left(\frac{1}{n} - \frac{1}{N}\right)}\right\}S_z^2 + \left(\frac{1}{n} - \frac{1}{n'}\right)E^* < 0$$

where $E^* = R^2S_x^2 + (\rho_{yx}S_y - 2RS_x)\,\rho_{yx}S_y$, $F^* = R\rho_{xz}S_x\,(R\rho_{xz}S_x - 2\rho_{yz}S_y)$ and $G^* = \left(R^2 - \beta^2\right)S_{x(2)}^2 - 2\,(R - \beta)\,\rho_{yx(2)}S_{y(2)}S_{x(2)}$.

5. Determination of n', n and K for fixed cost $C \leq C_0$

Let C_0 be the total cost (fixed) of the survey apart from overhead cost. The expected total of the survey apart from overhead cost is given by

$$C = (C_1' + C_2')n' + n'\left(C_1 + C_2W_1 + C_3\frac{W_2}{K}\right), \tag{1}$$

where C_1' is the cost per unit of identifying and observing main auxiliary character x,
C_2' is he cost per unit of identifying and observing additional auxiliary character z,
C_1 is the cost per unit of mailing questionnaire/visiting the unit at the second phase,
C_2 is the cost per unit of collecting and processing data from n_1 responding units,
C_3 is the cost per unit of obtaining and processing data after extra effort from the sub-sampled units,
$W_i = \frac{N_i}{N}(i = 1, 2)$ denotes the response and non-response rate in the population respectively.

The expression for $MSE(T_i)$ can be written as follows:

$$MSE\,(T_i) = \frac{M_{0i}}{n} + \frac{M_{1i}}{n'} + \frac{K}{n}M_{2i} + \text{terms independent from } n \text{ and } K,$$

where M_{0i} M_{1i} and M_{2i} are the coefficients of the terms of $\frac{1}{n}$, $\frac{1}{n'}$ and $\frac{K}{n}$ respectively in the expression of $MSE\left(T_i\right)$.

Let us define a function φ which is given by

$$\varphi = MSE\left(T_i\right) + \alpha_i \left\{ (C_1' + C_2')n' + n\left(C_1 + C_2W_1 + C_3\frac{W_2}{K}\right)\right\},$$

where α_i is the Lagrange's multiplier.

Now, differentiating φ with respect to n', n and K and equating them to zero, we get,

$$n' = \sqrt{\frac{M_{1i}}{\alpha_i(C_1' + C_2')}} \tag{2}$$

$$n = \sqrt{\frac{M_{0i} + KM_{2i}}{\alpha_i\left(C_1 + C_2W_1 + C_3\frac{W_2}{K}\right)}} \tag{3}$$

and

$$\frac{n}{K} = \sqrt{\frac{M_{2i}}{\alpha_i C_3 W_2}} \tag{4}$$

Now, putting the value of n in (4), we get,

$$K_{opt.} = \sqrt{\frac{C_3W_2M_{0i}}{\left(C_1 + C_2W_1\right)M_{2i}}} \tag{5}$$

Putting the values of n', n and $K_{opt.}$ from (2), (3) and (5) in (1), we get

$$\sqrt{\alpha_i} = \frac{1}{C_0}\left[\sqrt{(C_1' + C_2')M_{1i}} + \sqrt{\left[C_1 + C_2W_1 + \frac{C_3W_2}{K_{opt.}}\right]\left[M_{0i} + K_{opt.}M_{2i}\right]}\right] \tag{6}$$

It has also been seen that the determinant of the matrix of second order derivative of φ with respect to n', n and $K_{opt.}$ is negative for the optimum values of n', n and $K_{opt.}$, which shows that the solution for n', n given by (2) and (3) using (5), (6) and the optimum value of K for C for $C \leq C_0$ minimizes the $V(T_i)$. The minimum value of $MSE\left(T_i\right)$ for the optimum value of n', n and $K_{opt.}$ are given by

$$\begin{aligned} MSE\left(T_i\right)_{min.} =& \frac{1}{C_0}\left[\sqrt{(C_1' + C_2')M_{1i}} + \sqrt{\left[C_1 + C_2W_1 + \frac{C_3W_2}{K_{opt.}}\right]\left[M_{0i} + K_{opt.}M_{2i}\right]}\right]^2 \\ &- \frac{1}{N}\left[S_y^2 + R\rho_{xz}S_x(R\rho_{xz}S_x - 2\rho_{yz}S_y)\right] \end{aligned}$$

Now, neglecting the term of order N^{-1}, we have

$$MSE\left(T_i\right)_{min.} = \frac{1}{C_0}\left[\sqrt{(C_1' + C_2')M_{1i}} + \sqrt{\left[C_1 + C_2W_1 + \frac{C_3W_2}{K_{opt.}}\right]\left[M_{0i} + K_{opt.}M_{2i}\right]}\right]^2. \tag{7}$$

Now, putting the value of $K_{opt.}$ from (5) in (7), we have

$$MSE\left(T_i\right)_{min.} = \frac{1}{C_0}\left[\sqrt{(C_1' + C_2')M_{1i}} + \sqrt{M_{0i}\left(C_1 + C_2W_1\right)} + \sqrt{C_3W_2M_{2i}}\right]^2.$$

In the case of $\overline{y}^*$, we have

$$MSE\left(\overline{y}^*\right)_{min.} = \frac{1}{C_0}\left[\sqrt{M_0\left(C_1 + C_2W_1\right)} + \sqrt{C_3W_2M_2}\right]^2 - \frac{S_y^2}{N}.$$

Now, neglecting the term of order N^{-1}, we have

$$MSE\,(\overline{y}^{*})_{min.} = \frac{1}{C_0}\left[\sqrt{M_0\,(C_1 + C_2W_1)} + \sqrt{C_3W_2M_2}\,\right]^2,$$

where, M_0 and M_2 are the coefficients of the terms of $\frac{1}{n}$ and $\frac{K}{n}$ respectively in the expression of V$(\overline{y}^{*}) = \frac{f}{n}S_y^2 + \frac{W_2(K-1)}{n}S_y^2$.

Similarly, for the fixed cost C_0, the expression for $MSE\,(t_i)_{min.}$ is given by

$$MSE\,(t_i)_{min.} = \frac{1}{C_0}\left[\sqrt{(C_1' + C_2')V_{1i}} + \sqrt{V_{0i}\,(C_1 + C_2W_1)} + \sqrt{C_3W_2V_{2i}}\right]^2 - \frac{S_y^2}{N}.$$

Now, neglecting the term of order N^{-1}, we have

$$MSE\,(t_i)_{min.} = \frac{1}{C_0}\left[\sqrt{(C_1' + C_2')V_{1i}} + \sqrt{V_{0i}\,(C_1 + C_2W_1)} + \sqrt{C_3W_2V_{2i}}\right]^2,$$

where V_{0i} V_{1i} and V_{2i} are the coefficients of the terms of $\frac{1}{n}$, $\frac{1}{n'}$ and $\frac{K}{n}$ respectively in the expression of $MSE\,(t_i)$. The value of C_1' for which $MSE\,(t_i) < MSE(\overline{y}^{*})$ is given by

$$C_1' < \frac{1}{V_{1i}}\left[\sqrt{(C_1 + C_2W_1)}\left(\sqrt{M_0} - \sqrt{V_{0i}}\right) + \sqrt{C_3W_2}(\sqrt{M_2} - \sqrt{V_{2i}})\right]^2.$$

For the fixed value of C_0, the value of C_2' for which $MSE\,(T_i) < MSE\,(t_i)$ is given by

$$\frac{C_2'}{C_1'} < \frac{V_{1i} - M_{0i}}{M_{1i}}.$$

6. Empirical study

The data from the population of 100 records of resales of homes from Feb 15 to Apr 30, 1993 from the files maintained by the Albuquerque Board of Realtors on Selling price ($hundreds) as a study variable (y), Square feet of living space as an auxiliary variable (x) and annual taxes ($) as an additional variable (z) have been taken. The values of the parameters of the population are given as follows:

$$\overline{X} = 1697.44,\ \overline{Y} = 1093.41,\ \ \overline{Z} = 801.58,\ \ S_x = 535.01,\ n = 25, S_y = 391.90,$$

$$R = \frac{\overline{Y}}{\overline{X}}, S_z = 316.62, \rho_{yx} = 0.84, \rho_{yz} = 0.64,\ \rho_{xz} = 0.86,\ n' = 60.$$

The non-response rate in the population is considered to be 25%. So, the values of the population parameters based on the non-responding parts, which are taken as the last 25% units of the population are given as follows:

$$\overline{X}_2 = 1563.80,\ \overline{Y}_2 = 1017.04,\ S_{x(2)} = 383.44,\ S_{y(2)} = 361.75,\ \rho_{yx(2)} = 0.84,\ \ R_2 = \overline{Y}_2/\overline{X}_2.$$

From the table 1, we see that the proposed estimator T_1 is more efficient than $\overline{y}^{*}$, t_1, t_2, t_1', t_2' and T_2. The estimator T_2 is also more efficient than $\overline{y}^{*}$, t_2, t_1' and t_2'. It has been also observed that the mean square error decreases as K decreases.

From figure 1, we see that the proposed estimator T_1 is more efficient than $\overline{y}^{*}$, t_1, t_2, t_1', t_2' and T_2. The estimator T_2 is also more efficient than $\overline{y}^{*}$, t_2, t_1' and t_2'. It has been also observed that the mean square error decreases as K decreases.

Cost study has not be done for (t_1', t_2') due to the lack of information on coefficients of $\frac{1}{n}, \frac{1}{n'}$ and $\frac{K}{n}$.

From the table 2, we see that the proposed estimator T_1 is more efficient than $\overline{y}^{*}$, t_1, T_2 and t_2 for fixed cost $C \leq C_0$. The estimator T_2 is also more efficient than $\overline{y}^{*}$ and t_2 for fixed cost $C \leq C_0$.

Table 1: Relative efficiency of the proposed estimators w.r.t. $\overline{y}^*$.

	MSE (.)(R.E.(.) in % w.r.t. $\overline{y}^*$)		
Estimator	$K = 2.5$	$K = 2.0$	$K = 1.5$
$\overline{y}^*$	6603.23(100.00)	5938.00(100.00)	5272.78(100.00)
T_1	2319.36(284.70)	2106.21(281.93)	1893.06(278.53)
t_1	2724.49(242.37)	2511.34(236.45)	2298.19(229.43)
T_2	3675.58(179.65)	3010.35(197.25)	2345.13(224.84)
t_2	4080.71(161.82)	3415.49(173.86)	2750.26(191.72)
t_1'	2643.15(249.82)	2448.23(242.54)	2251.04(234.24)
t_2'	4046.44(163.19)	3381.23(175.62)	2716.02(194.14)

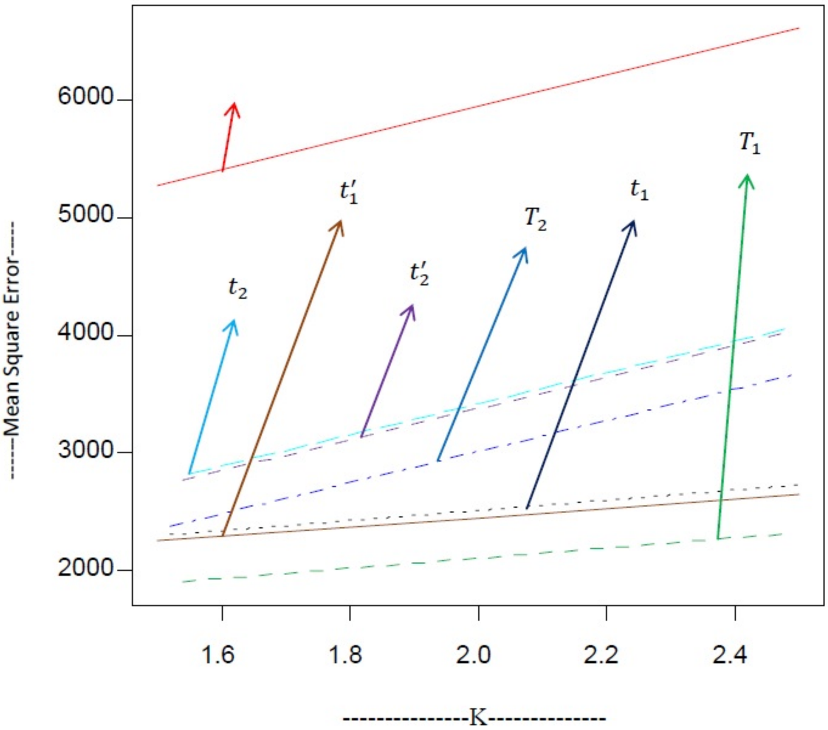

Figure 1: A graph for mean square errors of different estimators for different values of $K = 1.5,\ 2.0,\ 2.5$.

Table 2: Relative efficiency of the proposed estimators w.r.t. $\overline{y}^*$ for fixed cost.

C_0= Rs. 150.00 (fixed)				
Estimator	$K_{opt.}$	$n'_{opt.}$	$n_{opt.}$	R.E.(.)
$\overline{y}^*$	4.8	–	22	100.00
T_1	4.5	44	16	215.74
t_1	4.5	54	15	184.19
T_2	1.5	36	8	138.99
t_2	1.5	45	8	122.66

7. Conclusions

In this work, we have proposed conventional and alternative chain ratio-to-regression type estimators for the population mean using an additional auxiliary variable in the presence of non-response. Here, we conclude that the using information on an additional auxiliary variable are fruitful in increasing the precision of the estimators compared to those, not utilizing such information. For the support of the problem, an empirical study has been made. On the basis of the empirical study, we observe that use of an additional information in the proposed estimators for population mean in the presence of non-response is found to be more useful in increasing the precision of the proposed estimators with respect to the relevant estimators for the fixed cost $C \leq C_0$.The proposed estimators may be used in the field of Medical, Agricultural and Biological Sciences etc.

References

Cochran, W.G. (1942) Sampling thepry when the samoling units are of unequal sizes, Journal of American Statistical Association 37, 199–212.

Chaed, L. (1975) Some ratio-type Sstimotors based on two ar more auxiaiery variables, PhD Thesis submitted to Iowa etata University, Ames, IOWA.

Khare, B.B., Srivastava, S. (1993) Estimation of population mean using auxiliary character in presence of non-response, National Academy Science Letters, India, 16(3), 111–114.

Kiogyera, B. (1980) A chain ratio type estimator in finite population double sampling using two auxiliary variabres, Metrika 27, 217–223.

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

Hansen, M.H.A., Hurwitz, W.N. (1946) The problem of non-response in sample surveys, Journal of American Statistical Association 41, 517–629.

Rao, P.S.R.S. (1986) Ratio estimation with sub-sampling the non-respondents, Survey Methodology 12(2), 217–230.

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

Singh, H.P., Kumar, S. (2010) Improved estimation of population mean under two-phase sampling with subsampling the non-respondents, Journal of Statistical Planning and Inference 140(9), 2536–2550.