Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v8n1p67

Generalized chain ratio in regression estimators for population mean in the presence of non-response

By Khare, Kumar, Srivastava

Published: 26 April 2015

Electronic Journal of Applied Statistical Analysis
Vol. 08, Issue 01, April 2015, 67-83
DOI: 10.1285/i20705948v8n1p67

Generalized chain ratio in regression estimators for population mean in the presence of non-response

B.B.Khare[a], Kamlesh Kumar *[a], and U.Srivastava[b]

[a]*Department of Statistics, Banaras Hindu University, , Varanasi, India-221005*
[b]*Statistics Section, MMV, Banaras Hindu University, , Varanasi, India-221005*

Published: 26 April 2015

Two generalized chain ratio in regression estimators for population mean using two auxiliary characters in the presence of non-response have been proposed and their properties have been studied. Relative efficiency of the proposed estimators are obtained in the case of fixed first phase sample, second phase sample and sub-sample fraction and also in the case of the fixed cost. Comparison of the proposed estimators has been carried out with the relevant estimators. Expected cost is also obtained in the case of the specified variance. The performance of the proposed estimators in comparison to the relevant estimators has been made with the help of empirical study.

Keywords: Non-response, two phase sampling, bias, mean square error, auxiliary characters.

1 Introduction

Sample surveys are generally used in the field of agricultural, socio-economic and medical sciences. During sample surveys, sometimes information on some units in the selected sample is not obtained due to the problem of non-response. To deal with the problem of non-response, Hansen and Hurwitz (1946) have first suggested a technique of sub-sampling from non-respondents.

The information on the auxiliary character provides a very important contribution in the field of sample surveys. In the case when the population mean of the auxiliary

*Corresponding author: kamalbhu03@gmail.com.

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

character is known, Khare and Srivastava (1997) proposed transformed ratio estimators in the presence of non-response. Sometimes, the population mean of the auxiliary character is not known, in this situation, two phase sampling ratio, product and regression type estimators have been suggested by Khare and Srivastava (1995, 2010), Khare and Kumar (2011) and Khare et al. (2012).
In the case when the population mean of the main auxiliary character is not known but the population mean of another additional auxiliary character is known, which is cheaper than the main auxiliary character but less correlated to the study character than the main auxiliary character. In this situation, Kiregyera (1980, 1984) proposed chain ratio to regression, ratio in regression and regression in regression estimators for the population mean of study character. Further, some other chain regression type estimators for population mean of study character have been proposed by Mishra and Rout (1997) and Dash and Mishra (2011). In such situation, Khare and Kumar (2010) and Khare et al. (2011) have proposed chain regression type estimators and generalized chain estimators for the population mean in the presence of non-response.

In the present paper, we have proposed generalized chain ratio in regression estimators for the population mean using two auxiliary characters in the presence of non-response. We have obtained the expressions for bias and mean square error of the proposed estimators for the fixed first phase sample $n^{'}$, second phase sample n and the optimum values of constants. A comparative study of the mean square error of the proposed estimators is made with the relevant estimators. The optimum values of $n^{'}$ and n have been obtained for the fixed cost $(C \leq C_0)$and also for the specified variance V_0. The minimum value of the mean square error and the minimum value of the total cost incurred in the survey for the proposed estimators have been obtained for the fixed cost $(C \leq C_0)$ and for the specified variance V_0 respectively. An empirical study has been given to show the performance of the proposed estimators for fixed sample sizes $(n^{'}, n)$, for the fixed cost $(C \leq C_0)$ and also for the specified variance V_0.

2 The estimators

Let $\bar{Y}$, $\bar{X}$ and $\bar{Z}$ denote the population means of study character y , auxiliary character x and additional auxiliary character z having j^{th} values Y_j, X_j and Z_j : $j = 1, 2, ..., N$. The population of size N is supposed to be divided in N_1 responding units and N_2 non-responding units. According to Hansen and Hurwitz (1946), a sample of size n is drawn from the population of size N by using simple random sampling without replacement (SRSWOR) method of sampling and it has been observed that n_1 units respond and n_2 units do not respond. Again, by making extra effort, a sub-sample of size $r(= n_2 k^{-1})$ is drawn from n_2 non-responding units using SRSWOR method of sampling and the information on r units is collected by personal interview for study character y. Hence, the estimator for $\bar{Y}$ based on $n_1 + r$ units on study character y is given by Hansen and Hurwitz (1946) as follows:

$$\bar{y}^{*} = \frac{n_1}{n}\bar{y}_1 + \frac{n_2}{n}\bar{y}_2^{'}, \tag{1}$$

where $\bar{y}_1$ and $\bar{y}_2^{'}$ are the means of study character y based on n_1 and r units respectively. The variance of the estimator is given by.

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

where $W_2 = \frac{N_2}{N}$, $f = 1 - \frac{n}{N}$ and $(S_y^2, S_{y(2)}^2)$ are the population mean squares of study character y for the entire population and for the non-responding part of the population. In the case when the population mean of the auxiliary character is not known, we draw a first phase sample of size $n^{'}(< N)$ from the population of size N by using simple random sampling without replacement (SRSWOR) method of sampling and estimate the population mean $\bar{X}$ by first phase sample mean $\bar{x}^{'}$ based on $n^{'}$ units. Further, we draw a second phase sample of size $n(< n^{'})$ from first phase sample of size $n^{'}$ by using SRSWOR method of sampling and observe that n_1 units respond and n_2 units do not respond for the study character y. Again, we draw a sub-sample of size $r(= n_2k^{-1})$ from n_2 non-responding units by using SRSWOR method of sampling and collect the information on r units by making extra effort. So, the mean of y based on $n_1 + r$ units is define by (1) and the mean of values of x corresponding to the incomplete information on y is defined by

$$\bar{x}^* = \frac{n_1}{n}\bar{x}_1 + \frac{n_2}{n}\bar{x}_2^{'}, \tag{3}$$

where $\bar{x}_1$ and $\bar{x}_2^{'}$ are the means of auxiliary character x based on n_1 and r units corresponding to n_1 and r units on study character y respectively.
Using $\bar{x}^{'}$, $\bar{x}$, $\bar{x}^*$ and $\bar{y}^*$, Khare and Srivastava (1995) proposed the conventional (T_{l1}) and the alternative (T_{l2}) regression type estimators for the population mean in the presence of non-response, which are given as follows:

$$T_{l1} = \bar{y}^* + b_{yx}^*(\bar{x}^{'} - \bar{x}^*) \tag{4}$$

and

$$T_{l2} = \bar{y}^* + b_{yx}^{**}(\bar{x}^{'} - \bar{x}), \tag{5}$$

where $\bar{x} = \sum_{j=1}^{n} \frac{x_j}{n}$, $b_{yx}^* = \frac{\hat{S}_{yx}}{\hat{S}_x^2}$, $b_{yx}^{**} = \frac{\hat{S}_{yx}}{s_x^2}$, $s_X^2 = \dfrac{1}{n-1}\sum_{j=1}^{n}(x_j - \bar{x})^2$, $\hat{S}_{yx}$ and $\hat{S}_x^2$ denote the estimates of S_{yx} and S_x^2 based on $n_1 + r$ units. In the case when the population mean $\bar{X}$ of an auxiliary character x is not known but the population mean $\bar{Z}$ of another additional auxiliary character z is known which may be cheaper and less correlated to the study character y in comparison to main auxiliary character x i.e. $\rho_{yx} > \rho_{yz}$, a first phase sample of size $n^{'}(< N)$ is drawn from the population of size N by using SRSWOR method of sampling and the population mean $\bar{X}$ is estimated by $\hat{\bar{X}} = \bar{x}^{'}\frac{\bar{Z}}{\bar{z}^{'}}$ which is more efficient in comparison to $\bar{x}^{'}$ if $\rho_{xz} = \frac{C_z}{2C_x}$, where ρ_{xz} is the correlation coefficient between x and z, (C_x, C_z) are the coefficient of variations of X and Z and $(\bar{x}^{'}, \bar{z}^{'})$ are the means of the auxiliary character x and additional auxiliary character z based on $n^{'}$ units.
Now, using $\bar{Z}$, $\bar{z}^{'}$, $\bar{x}^{'}$, $\bar{x}$, $\bar{x}^*$ and $\bar{y}^*$, we propose the conventional (T_{l5}) and the alternative

(T_{l6}) generalized chain ratio in regression estimators for the population mean using two auxiliary characters in the presence of non-response, which are given as follows:

$$T_{l5} = \bar{y}^* + b^*_{yx}\left[\bar{x}'\left(\frac{\bar{Z}}{\bar{z}'}\right)^{\alpha_1} - \bar{x}^*\right] \tag{6}$$

and

$$T_{l6} = \bar{y}^* + b^{**}_{yx}\left[\bar{x}'\left(\frac{\bar{Z}}{\bar{z}'}\right)^{\alpha_2} - \bar{x}\right] \tag{7}$$

where α_1 and α_2 are constants.

3 Bias and mean square error of the estimators (T_{l5}) and (T_{l6})

The expressions for bias and mean square error of the estimators (T_{l5}) and (T_{l6}) up to the terms of order n^{-1} are given by

$$Bias(T_{l5}) = \theta_1\left[-\mu_{15} + \mu_{16} + \mu_{35} - \mu_{36} - \alpha_1\mu_{45} + \alpha_1\mu_{46} + \frac{f'}{n'}\left\{\frac{\alpha_1(\alpha_1+1)}{2}C_z^2 - \alpha_1 C_{xz}\right\}\right], \tag{8}$$

$$Bias(T_{l6}) = \theta_1\left[-\mu_{25} + \mu_{27} + \mu_{35} - \mu_{37} - \alpha_2\mu_{45} + \alpha_2\mu_{47} + \frac{f'}{n'}\left\{\frac{\alpha_2(\alpha_2+1)}{2}C_z^2 - \alpha_2 C_{xz}\right\}\right] \tag{9}$$

and

$$\begin{aligned} MSE(T_{l5}) =& V(\bar{y}^*) - \bar{Y}^2\left[\left(\frac{1}{n} - \frac{1}{n'}\right)\rho^2_{yx}C^2_y - \frac{W_2(k-1)}{n}\left\{B^2C^2_{x(2)} - 2BC_{yx(2)}\right\}\right. \\ &\left. - \frac{f'}{n'}\left(\alpha_1^2\frac{\rho^2_{yx}C^2_yC^2_z}{C^2_x} - 2\alpha_1\frac{\rho_{yx}\rho_{yz}C^2_yC_z}{C^2_x}\right)\right], \end{aligned} \tag{10}$$

$$MSE(T_{l6}) = V(\bar{y}^*) - \bar{Y}^2\left[\left(\frac{1}{n} - \frac{1}{n'}\right)\rho^2_{yx}C^2_y - \frac{f'}{n'}\left(\alpha_2^2\frac{\rho^2_{yx}C^2_yC^2_z}{C^2_x} - 2\alpha_2\frac{\rho_{yx}\rho_{yz}C^2_yC_z}{C^2_x}\right)\right] \tag{11}$$

where $C_{yx} = \rho_{yx}C_yC_x$, $C_{yx(2)} = \rho_{yx(2)}C_{y(2)}C_{x(2)}$, $C_{yz} = \rho_{yz}C_yC_z$, $C_{xz} = \rho_{xz}C_xC_z$, $\theta_1 = \bar{Y}\frac{C_{yx}}{C_x^2}$, $f' = 1 - \frac{n'}{N}$, $B = \frac{\beta}{R}$, $R = \frac{\bar{Y}}{\bar{X}}$, $\beta = \frac{S_{yx}}{S_x^2}$, $C_y = \frac{S_y}{\bar{Y}}$, $C_x = \frac{S_x}{\bar{X}}$, $C_{y(2)} = \frac{S_{y(2)}}{\bar{Y}}$, $C_{x(2)} = \frac{S_{x(2)}}{\bar{X}}$, $(S_x^2, S_z^2, \rho_{yx}, \rho_{yz}, \rho_{xz})$ are the population mean squares of auxiliary characters (x,z), correlation coefficient between (y,x),(y,z) and (x,z) for the entire population, $(S^2_{x(2)}, \rho_{yx(2)})$ are the population mean square of x and correlation coefficients between (y, x) for the non-responding part of the population.
The optimum values of $\alpha_1$and $\alpha_2$which minimize $MSE(T_{l5})$and $MSE(T_{l6})$are given as follows: $\alpha_{1opt} = \frac{\rho_{yz}}{\rho_{yx}}\frac{C_x}{C_z}$ and $\alpha_{2opt} = \frac{\rho_{yz}}{\rho_{yx}}\frac{C_x}{C_z}$.

The minimum mean square errors of the estimators T_{l5} and $T_{l6}$for the optimum values of $\alpha_1$and $\alpha_2$are given as

$$MSE(T_{l5})_{\min} = V(\bar{y}^*) - \bar{Y}^2 \left[\left(\frac{1}{n} - \frac{1}{n'} \right) \rho_{yx}^2 C_y^2 - \frac{W_2(k-1)}{n} \left\{ B^2 C_{x(2)}^2 - 2BC_{yx(2)} \right\} + \frac{f'}{n'} \rho_{yz}^2 C_y^2 \right] \tag{12}$$

and

$$MSE(T_{l6})_{\min} = V(\bar{y}^*) - \bar{Y}^2 \left[\left(\frac{1}{n} - \frac{1}{n'} \right) \rho_{yx}^2 C_y^2 + \frac{f'}{n'} \rho_{yz}^2 C_y^2 \right] \tag{13}$$

The optimum values of $\alpha_1$and $\alpha_2$may be obtained from past data. If past data is not available then one may estimate it on the basis of sample values without having any loss in the efficiency of the estimators. If we estimate the optimum values of the constants by using the sample values, the minimum values of the mean square error of the estimators up to the terms of order (n^{-1}) are unchanged.

Remark 3.1: For $\alpha_1 = 0$ and $\alpha_2 = 0$, T_{l5} reduces to $T_{l1} = \bar{y}^* + b_{yx}^*(\bar{x}' - \bar{x}^*)$ and T_{l6} reduces to $T_{l2} = \bar{y}^* + b_{yx}^{**}(\bar{x}' - \bar{x})$. The MSE of T_{l1} and T_{l2} can be obtained by putting $\alpha_1 = 0$ and $\alpha_2 = 0$ in equations (10) and (11), which are given as

$$MSE(T_{l1}) = V(\bar{y}^*) - \bar{Y}^2 \left[\left(\frac{1}{n} - \frac{1}{n'} \right) \rho_{yx}^2 C_y^2 - \frac{W_2(k-1)}{n} \left\{ B^2 C_{x(2)}^2 - 2B\, C_{yx(2)} \right\} \right] \tag{14}$$

and

$$MSE(T_{l2}) = V(\bar{y}^*) - \bar{Y}^2 \left(\frac{1}{n} - \frac{1}{n'} \right) \rho_{yx}^2 C_y^2 \tag{15}$$

Remark 3.2: For $\alpha_1 = 1$ and $\alpha_2 = 1$, T_{l5} reduces to $T_{l3} = \bar{y}^* + b_{yx}^* \left[\bar{x}' \frac{\bar{Z}}{\bar{z}'} - \bar{x}^* \right]$ and T_{l6} reduces to $T_{l4} = \bar{y}^* + b_{yx}^{**} \left[\bar{x}' \frac{\bar{Z}}{\bar{z}'} - \bar{x} \right]$, The MSE of T_{l3} and T_{l4}can be obtained by putting $\alpha_1 = 1$ and $\alpha_2 = 1$ in equations (10) and (11), which are given as

$$MSE(T_{l3}) = MSE(T_{l1}) + \bar{Y}^2 \frac{f'}{n'} \left[\frac{\rho_{yx}^2 C_y^2 C_z^2}{C_x^2} - 2 \frac{\rho_{yx}\rho_{yz} C_y^2 C_z}{C_x} \right] \tag{16}$$

and

$$MSE(T_{l4}) = MSE(T_{l2}) + \bar{Y}^2 \frac{f'}{n'} \left[\frac{\rho_{yx}^2 C_y^2 C_z^2}{C_x^2} - 2 \frac{\rho_{yx}\rho_{yz} C_y^2 C_z}{C_x} \right] \tag{17}$$

Remark 3.3: For $\alpha_1 = -1$ and $\alpha_2 = -1$, T_{l5} reduces to $T_{l3}' = \bar{y}^* + b_{yx}^* \left[\bar{x}' \frac{\bar{z}'}{\bar{Z}} - \bar{x}^* \right]$ and T_{l6} reduces to $T_{l4}' = \bar{y}^* + b_{yx}^{**} \left[\bar{x}' \frac{\bar{z}'}{\bar{Z}} - \bar{x} \right]$. The MSE of T_{l3}' and T_{l4}' can be obtained by putting $\alpha_1 = -1$ and $\alpha_2 = -1$ in equations (10) and (11) which are given as

$$MSE(T'_{l3}) = MSE(T_{l1}) + \bar{Y}^2 \frac{f'}{n'} \left[\frac{\rho_{yx}^2 C_y^2 C_z^2}{C_x^2} + 2 \frac{\rho_{yx}\rho_{yz} C_y^2 C_z}{C_x} \right] \tag{18}$$

and

$$MSE(T'_{l4}) = MSE(T_{l2}) + \bar{Y}^2 \frac{f'}{n'} \left[\frac{\rho_{yx}^2 C_y^2 C_z^2}{C_x^2} + 2 \frac{\rho_{yx}\rho_{yz} C_y^2 C_z}{C_x} \right] \tag{19}$$

4 Comparison of the proposed estimators $T_{l\,5}$ and $T_{l\,6}$ with $T_{l\,1}$, $T_{l\,2}$, $T_{l\,3}$, $T_{l\,4}$,$T'_{l\,3}$,$T'_{l\,4}$

$$MSE(T_{l5}) < MSE(T_{l1}), if 0 < \alpha_1 < 2Q_1 \tag{20}$$

$$MSE(T_{l5}) < MSE(T_{l3}), if 2Q_1 - 1 < \alpha_1 < 1 \tag{21}$$

$$MSE(T_{l5}) < MSE(T'_{l3}) if -1 < \alpha_1 < 2Q_1 + 1 \tag{22}$$

and

$$MSE(T_{l6}) < MSE(T_{l2}), if 0 < \alpha_2 < 2Q_2 \tag{23}$$

$$MSE(T_{l6}) < MSE(T_{l4}), if 2Q_2 - 1 < \alpha_2 < 1 \tag{24}$$

$$MSE(T_{l6}) < MSE(T'_{l4}), if -1 < \alpha_2 < 2Q_2 + 1 \tag{25}$$

where $Q_1 = \frac{\rho_{yz}}{\rho_{yx}} \frac{C_x}{C_z}$ and $Q_2 = \frac{\rho_{yz}}{\rho_{yx}} \frac{C_x}{C_z}$.

5 Determination of n', n and k for the fixed cost $C \leq C_0$

Let C_0 denotes the total cost (fixed) of the survey apart from overhead cost. The cost function C' can be expressed by

$$C' = (p'_1 + p'_2)n' + p_1 n + p_2 n_1 + p_3 \frac{n_2}{k}. \tag{26}$$

The expected cost of survey apart from overhead cost is given as follows:

$$C = E(C') = (p'_1 + p'_2)n' + n \left(p_1 + p_2 W_1 + p_3 \frac{W_2}{k} \right), \tag{27}$$

where

p'_1 - the cost per unit of obtaining information on auxiliary character x at the first phase,

p'_2 - the cost per unit of obtaining information on additional auxiliary character z at the first phase,

p_1 - the cost per unit of mailing questionnaire/visiting the units at the second phase,

p_2 - the cost per unit of collecting, processing data for the study character y obtained from n_1 responding units,

p_3 - the cost per unit of obtaining and processing data (after extra efforts) for the study character y from sub-sampling units

and $W_1 = N_1/N$,$W_2 = N_2/N$.

It is to be noted that $p'_2 < p'_1 < p_1 < p_2 < p_3$.

The expression of $MSE(T(i))$: i=1,2,3,4 can be expressed in terms of B_{0i}, B_{1i}, B_{2i} and B_{3i}, which is given as

$$MSE(T(i)) = \frac{B_{0i}}{n} + \frac{B_{1i}}{n'} + \frac{k\ B_{2\mathrm{i}}}{n} - \frac{B_{3i}}{N}, \tag{28}$$

where B_{oi}, B_{1i},B_{2i} and B_{3i} are respectively the coefficients of the terms of n^{-1},n'^{-1}, $k\ n^{-1}$ and N^{-1} in the expression of $MSE(T(i))$ and $T(1) = T_{l3}$, $T(2) = T_{l4}$, $T(3) = T_{l5}$, $T(4) = T_{l6}$.

Now, we define a function ϕ to minimize $MSE(T(i))$ for the fixed cost $C \leq C_0$, which is given as follows:

$$\phi = MSE(T(i)) + \lambda_i \left(C - C_0\right), \tag{29}$$

where λ_i is the Lagrange's multiplier.

Differentiating equation (29) with respect to n', n and k, we get the optimum values of n', n and k, which are given as

$$n' = \sqrt{\frac{B_{1i}}{\lambda_i(p'_1 + p'_2)}}, \tag{30}$$

$$n = \sqrt{\frac{(B_{0i} + k\ B_{2i})}{\lambda_i\left(p_1 + p_2 W_1 + p_3 \frac{W_2}{k}\right)}} \tag{31}$$

and

$$k_{opt} = \sqrt{\frac{B_{0i} p_3 W_2}{B_{2i}(p_1 + p_2 W_1)}}, \tag{32}$$

where

$$\sqrt{\lambda_i} = \frac{1}{C_0}\left[\sqrt{B_{1i}(p'_1 + p'_2)} + \sqrt{(B_{0i} + k_{opt} B_{2i})\left(p_1 + p_2 W_1 + p_3 \frac{W_2}{k_{opt}}\right)}\right]. \tag{33}$$

By putting the optimum values of n',n and k from equations (30), (31) and (32) in equation (28) and neglecting the term of order (N^{-1}), the minimum value of $MSE\ (T(i))$ for the fixed cost $C \leq C_0$ is obtained as.

$$MSE\ (T(i))_{\min} = \frac{1}{C_0}\left[\sqrt{B_{1i}(p'_1 + p'_2)} + \sqrt{(B_{0i} + k_{opt} B_{2i})\left(p_1 + p_2 W_1 + p_3 \frac{W_2}{k_{opt}}\right)}\right]^2 \tag{34}$$

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

Let V_0 denotes the variance of the estimator $T(i)$ which is fixed in advance. So, we have

$$V_0 = \frac{B_{0i}}{n} + \frac{B_{1i}}{n'} + \frac{kB_{2i}}{n} - \frac{B_{3i}}{N}. \tag{35}$$

For minimizing the average total cost C for the specified variance ($i.e.$ $MSE(T(i)) = V_0$) of the estimator $T(i)$ and also for obtaining the optimum values of n', n and k, we define a function ψ which is given as follows:

$$\psi = (p_1' + p_2')n' + n\left(p_1 + p_2 W_1 + p_3 \frac{W_2}{k}\right) + \mu_i(MSE(T(i)) - V_0), \tag{36}$$

where μ_i is the Lagrange's multiplier.

Now, differentiating equations (36) with respect to n', n and k, we get the optimum values of n', n and k which are given as

$$n' = \sqrt{\frac{\mu_i B_{1i}}{(p_1' + p_2')}}, \tag{37}$$

$$n = \sqrt{\frac{\mu_i(B_{0i} + kB_{2i})}{(p_1 + p_2 W_1 + p_3 \frac{W_2}{k})}} \tag{38}$$

and

$$k_{opt} = \sqrt{\frac{B_{0i} W_2 p_3}{B_{2i}(p_1 + p_2 W_1)}}, \tag{39}$$

where

$$\sqrt{\mu_i} = \frac{\left[\sqrt{B_{1i}(p_1' + p_2')} + \sqrt{(B_{0i} + k_{opt} B_{2i})\left(p_1 + p_2 W_1 + p_3 \frac{W_2}{k_{opt}}\right)}\right]}{V_0 + \frac{B_{3i}}{N}}. \tag{40}$$

By putting the optimum values of n', n and k from equations (37), (38) and (39) in equation (27) and neglecting the terms of order (N^{-1}), the minimum expected total cost for the specified variance V_0 is obtained as.

$$C(T(i))_{\min} = \frac{\left[\sqrt{B_{1i}(p_1' + p_2')} + \sqrt{(B_{0i} + k_{opt} B_{2i})\left(p_1 + p_2 W_1 + p_3 \frac{W_2}{k_{opt}}\right)}\right]^2}{V_0} \tag{41}$$

7 An empirical study

The present data has been taken from paper of Khare and Kumar (2010). The data from the population of 100 records of resale of homes from Feb 15 to Apr 30, 1993 from the files maintained by the Albuquerque Board of Realtors. On selling price ($hundreds) as a study character (y), square feet of living space as an auxiliary character (x) and annual taxes ($) as an additional auxiliary character (z) have been taken.
The values of the population parameters are given as follows:

$$\bar{Y} = 1093.41, \bar{X} = 1697.44, \bar{Z} = 801.58, S_y = 391.90, S_x = 535.01, S_z = 316.62,$$

$$\rho_{yx} = 0.84, \rho_{y\ z} = 0.64, \rho_{x\ z} = 0.86.$$

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:

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

The optimum values of α_1 and α_2 are obtained as follows:
$\alpha_{1opt} = 0.608$ and $\alpha_{2opt} = 0.608$.

Table 1: Relative efficiency (in %) of the estimators with respect to $\bar{y}^*$for the fixed values of n', n and different values of k and N =*100*, n' =*70* and n =*40*

	$1/k$		
Estimators	1/4	1/3	1/2
$\bar{y}^*$	100.00 (4757.47)	100.00 (3939.57)	100.00 (3121.68)
T_{l1}	243.77 (1951.58)	234.23 (1681.94)	221.03 (1412.31)
T_{l2}	132.28 (3596.36)	141.79 (2778.46)	159.22 (1960.57)
T_{l3}	265.17 (1794.07)	258.43 (1524.44)	248.77 (1254.81)
T_{l4}	138.34 (3438.86)	150.31 (2620.96)	173.13 (1803.07)
T_{l5}	282.85 (1681.97)	278.94 (1412.34)	273.18 (1142.70)
T_{l6}	143.00 (3326.75)	157.03 (2508.86)	184.61 (1690.96)

Figures in parenthesis give the MSE (.)

From table 1, It is observed that for the fixed sample sizes, the proposed estimators (T_{l5}, T_{l6}) are more efficient than corresponding estimators (T_{l1}, T_{l2}) and (T_{l3}, T_{l4}). It is also observed that the estimator T_{l5} is more efficient than the estimator T_{l6}. The values of mean square errors of all the estimators $\bar{y}^*, T_{l1}, T_{l2}, T_{l3}$, T_{l4},T_{l5} and T_{l6} decrease as the value of $1/k$ increases.

Table 2: Relative efficiency (in %) of the estimators with respect to$\bar{y}^*$for the fixed cost $C \leq C_0$ = *Rs. 280*, p_1' =*Rs. 0.90*, p_2' =*Rs. 0.20*, p_1 =*Rs.1.5*, p_2 =*Rs. 3*, p_3 =*Rs. 65*

Estimators	k_{opt}	n'_{opt}	n_{opt}	RE(in %)
$\bar{y}^*$	4.00	—-	36	100.00 (7023.79)
T_{l1}	3.72	89	25	165.49 (4245.31)
T_{l2}	1.29	77	13	124.28 (5651.46)
T_{l3}	3.72	67	25	175.48 (4002.62)
T_{l4}	1.29	58	13	130.76 (5370.89)
T_{l5}	3.72	57	27	196.04 (3582.79)
T_{l6}	1.29	49	14	143.85 (4882.74)

Figures in parenthesis give the MSE (.), Rs. : Rupees (Indian Currency) and RE:Relative Efficiency

From table 2, It is observed that for the fixed cost, the proposed estimators (T_{l5}, T_{l6}) have less mean square error in comparison to the corresponding estimators (T_{l1}, T_{l2}) and (T_{l3}, T_{l4}). It is also observed that the estimator T_{l5} has less mean square error in comparison to the estimator T_{l6}.

Table 3: Expected cost (in Rs.) of the estimators for the specified variance $V_0 = 5789$, p'_1 *=Rs. 0.90*, p'_2 *=Rs. 0.20*, p_1 *=Rs.1.5*, p_2 *=Rs. 3*, p_3 *=Rs. 65*

Estimators	k_{opt}	n'_{opt}	n_{opt}	EC (in Rs.)
$\bar{y}^*$	4.00	—-	43	339.72
T_{l1}	3.72	65	18	205.33
T_{l2}	1.29	75	13	273.35
T_{l3}	3.72	47	18	193.59
T_{l4}	1.29	54	12	259.78
T_{l5}	3.72	35	17	173.29
T_{l6}	1.29	41	12	236.16

Here EC: Expected Cost

From table 3, It is observed that for the specified variance, the proposed estimators (T_{l5}, T_{l6}) have less cost in comparison to corresponding estimators (T_{l1}, T_{l2}) and (T_{l3}, T_{l4}). It is also observed that the estimator T_{l5} has less cost in comparison to the estimator T_{l6}.

8 Conclusion

Hence, we conclude that the proposed conventional generalized chain ratio in regression estimator T_{l5} is more efficient than the conventional estimators (T_{l1},T_{l3}) for fixed sample sizes (n', n) and for the fixed cost $C \leq C_0$. Further, for specified variance, the cost incurred in the survey for the estimator T_{l5} is less than the corresponding estimators (T_{l1},T_{l3}). Similarly, the proposed alternative generalized chain ratio in regression estimator T_{l6} is also more efficient than the alternative estimators (T_{l2},T_{l4}) for fixed sample sizes (n', n) and for the fixed cost $C \leq C_0$. Further, for specified variance, the cost incurred in the survey for the estimator T_{l6} is less than the corresponding estimators (T_{l2},T_{l4}). However, in the present study, it has been observed that the conventional estimator T_{l5} is more efficient than the alternative estimator T_{l6} for fixed sample sizes (n', n) and for the fixed cost $C \leq C_0$ and the cost incurred during the survey for T_{l5} is also found to be less than the alternative estimator T_{l6}.

Aknowledgment

The authors are highly grateful to the referees for his valuable suggestions to improve the quality of the paper.

Appendix

Let
$\bar{y}^* = \bar{Y}(1+\varepsilon_0)$, $\bar{x}^* = \bar{X}(1+\varepsilon_1)$, $\bar{x} = \bar{X}(1+\varepsilon_2)$, $\bar{x}' = \bar{X}(1+\varepsilon_3)$, $\bar{z}' = \bar{Z}(1+\varepsilon_4)$, $\hat{S}_{yx} = S_{yx}(1+\varepsilon_5)$, $\hat{S}_x^2 = S_x^2(1+\varepsilon_6)$, $s_x^2 = S_x^2(1+\varepsilon_7)$, such that $\mathrm{E}(\varepsilon_\ell) = 0$ and $|\varepsilon_\ell| < 1$ $\forall$ ℓ= 0, 1,. . .,7.

Now, using SRSWOR method of sampling, we have

$$E(\varepsilon_0^2) = \frac{V(\bar{y}^*)}{\bar{Y}^2} = \frac{f}{n}C_y^2 + \frac{W_2(k-1)}{n}C_{y(2)}^2, E(\varepsilon_1^2) = \frac{V(\bar{\text{x}}^*)}{\bar{X}^2} = \frac{f}{n}C_x^2 + \frac{W_2(\text{k}-1)}{n}C_{x(2)}^2, \tag{42}$$

$$E(\varepsilon_2^2) = \frac{V(\bar{x})}{\bar{X}^2} = \frac{f}{n}C_x^2, E(\varepsilon_3^2) = \frac{V(\bar{x}')}{\bar{X}^2} = \frac{f'}{n'}C_x^2, \tag{43}$$

$$E(\varepsilon_4^2) = \frac{V(\bar{\text{z}}')}{\bar{Z}^2} = \frac{f'}{n'}C_z^2, E(\varepsilon_5^2) = \frac{V(\hat{S}_{yx})}{S_{yx}^2}, E(\varepsilon_6^2) = \frac{\text{V}(\hat{S}_x^2)}{S_x^4}, \tag{44}$$

$$E(\varepsilon_7^2) = \frac{\text{V}(s_x^2)}{S_x^4}, E(\varepsilon_0\varepsilon_1) = \frac{Cov(\bar{y}^*, \bar{x}^*)}{\bar{Y}\ \bar{X}} = \frac{f}{n}C_{yx} + \frac{W_2(k-1)}{n}C_{yx(2)}, \tag{45}$$

$$E(\varepsilon_0\varepsilon_2) = \frac{Cov(\bar{y}^*, \bar{x})}{\bar{Y}\ \bar{X}} = \frac{f}{n}C_{yx}, E(\varepsilon_0\varepsilon_3) = \frac{Cov(\bar{y}^*, \bar{x}')}{\bar{Y}\ \bar{X}} = \frac{f'}{n'}C_{yx}, \tag{46}$$

$$E(\varepsilon_0\varepsilon_4) = \frac{Cov(\bar{y}^*, \bar{z}')}{\bar{Y}\ \bar{Z}} = \frac{f'}{n'}C_{yz}, E(\varepsilon_0\varepsilon_5) = \frac{Cov(\bar{y}^*, \hat{S}_{yx})}{\bar{Y}S_{yx}} = \mu_{05}, \tag{47}$$

$$E(\varepsilon_0\varepsilon_6) = \frac{Cov(\bar{y}^*, \hat{S}_x^2)}{\bar{Y}S_x^2} = \mu_{06}, E(\varepsilon_0\varepsilon_7) = \frac{Cov(\bar{y}^*, s_x^2)}{\bar{Y}S_x^2} = \mu_{07}, \tag{48}$$

$$E(\varepsilon_1\varepsilon_3) = \frac{Cov(\bar{x}^*, \bar{x}')}{\bar{X}^2} = \frac{f'}{n'}C_x^2, E(\varepsilon_1\varepsilon_4) = \frac{Cov(\bar{x}^*, \bar{z}')}{\bar{X}\ \bar{Z}} = \frac{f'}{n'}C_{xz}, \tag{49}$$

$$E(\varepsilon_1\varepsilon_5) = \frac{Cov(\bar{x}^*, \hat{S}_{yx})}{\bar{X}S_{yx}} = \mu_{15}, E(\varepsilon_1\varepsilon_6) = \frac{Cov(\bar{x}^*, \hat{S}_x^2)}{\bar{X}S_x^2} = \mu_{16}, \tag{50}$$

$$E(\varepsilon_2\varepsilon_3) = \frac{Cov(\bar{x}, \bar{x}')}{\bar{X}^2} = \frac{f'}{n'}C_x^2, E(\varepsilon_2\varepsilon_4) = \frac{Cov(\bar{x}, \bar{z}')}{\bar{X}\ \bar{Z}} = \frac{f'}{n'}C_{xz}, \tag{51}$$

$$E(\varepsilon_2\varepsilon_5) = \frac{Cov(\bar{x}, \hat{S}_{yx})}{\bar{X}S_{yx}} = \mu_{25}, E(\varepsilon_2\varepsilon_7) = \frac{Cov(\bar{x}, s_x^2)}{\bar{X}S_x^2} = \mu_{27}, \tag{52}$$

$$E(\varepsilon_3\varepsilon_4) = \frac{Cov(\bar{x}', \bar{z}')}{\bar{X}\ \bar{Z}} = \frac{f'}{n'}C_{xz}, E(\varepsilon_3\varepsilon_5) = \frac{Cov(\bar{x}', \hat{S}_{yx})}{\bar{X}S_{yx}} = \mu_{35}, \tag{53}$$

$$E(\varepsilon_3\varepsilon_6) = \frac{Cov(\bar{x}', \hat{S}_x^2)}{\bar{X}S_x^2} = \mu_{36}, E(\varepsilon_3\varepsilon_7) = \frac{Cov(\bar{x}', s_x^2)}{\bar{X}S_x^2} = \mu_{37}, \tag{54}$$

$$E(\varepsilon_4\varepsilon_5) = \frac{Cov(\bar{z}', \hat{S}_{yx})}{\bar{Z}S_{yx}} = \mu_{45}, E(\varepsilon_4\varepsilon_6) = \frac{Cov(\bar{z}', \hat{S}_x^2)}{\bar{Z}S_x^2} = \mu_{46}, \tag{55}$$

$$E(\varepsilon_4\varepsilon_7) = \frac{Cov(\bar{z}', s_x^2)}{\bar{Z}S_x^2} = \mu_{47}, E(\varepsilon_5\varepsilon_6) = \frac{Cov(\hat{S}_{yx}, \hat{S}_x^2)}{S_{yx}S_x^2} = \mu_{56}, \tag{56}$$

$$E(\varepsilon_5\varepsilon_7) = \frac{Cov(\hat{S}_{yx}, s_x^2)}{S_{yx}S_x^2} = \mu_{57} \tag{57}$$

Theorem 1: The bias of the estimators $T_{l\,5}$ and $T_{l\,6}$ up to the terms of order (n^{-1}) are given by

$$Bias(T_{l5}) = \theta_1\left[-\mu_{15} + \mu_{16} + \mu_{35} - \mu_{36} - \alpha_1\mu_{45} + \alpha_1\mu_{46} + \frac{f'}{n'}\left\{\frac{\alpha_1(\alpha_1+1)}{2}C_z^2 - \alpha_1 C_{xz}\right\}\right] \tag{58}$$

and

$$Bias(T_{l6}) = \theta_1\left[-\mu_{25} + \mu_{27} + \mu_{35} - \mu_{37} - \alpha_2\mu_{45} + \alpha_2\mu_{47} + \frac{f'}{n'}\left\{\frac{\alpha_2(\alpha_2+1)}{2}C_z^2 - \alpha_2 C_{xz}\right\}\right] \tag{59}$$

where $\theta_1 = \bar{Y}\frac{C_{yx}}{C_x^2}$

Proof: The estimator T_{l5} is given as

$$T_{l5} = \bar{y}^* + b_{yx}^*\left[\bar{x}'\left(\frac{\bar{Z}}{\bar{z}'}\right)^{\alpha_1} - \bar{x}^*\right]$$

$$= \bar{y}^* + \frac{\hat{S}_{yx}}{\hat{S}_x^2}\left[\bar{x}'\left(\frac{\bar{Z}}{\bar{z}'}\right)^{\alpha_1} - \bar{x}^*\right]$$

The estimator T_{l5} can be expressed in terms of ε_l's as

$$T_{l5} = \bar{Y}(1+\varepsilon_0) + \frac{S_{yx}(1+\varepsilon_5)}{S_x^2(1+\varepsilon_6)}\left[\bar{X}(1+\varepsilon_3)\left(\frac{\bar{Z}}{\bar{Z}(1+\varepsilon_4)}\right)^{\alpha_1} - \bar{X}(1+\varepsilon_1)\right]$$

$$= \bar{Y}(1+\varepsilon_0) + \bar{X}\frac{S_{yx}}{S_x^2}(1+\varepsilon_5)(1+\varepsilon_6)^{-1}\left[(1+\varepsilon_3)(1+\varepsilon_4)^{-\alpha_1} - (1+\varepsilon_1)\right]$$

$$\begin{aligned} =&\bar{Y}(1+\varepsilon_0) + \theta_1(1+\varepsilon_5)(1-\varepsilon_6+\varepsilon_6^2,...)(1+\varepsilon_3)\left(1-\alpha_1\varepsilon_4+\frac{\alpha_1(\alpha_1+1)}{2}\varepsilon_4^2,...\right) \\ &-\theta_1(1+\varepsilon_5)(1-\varepsilon_6+\varepsilon_6^2,...)(1+\varepsilon_1) \end{aligned} \tag{60}$$

After solving and neglecting the terms of ε_l's having power more than two, we get

$$\begin{aligned} T_{l5} - \bar{Y} =&\bar{Y}\varepsilon_0 + \theta_1\{\varepsilon_3 - \varepsilon_1 - \alpha_1\varepsilon_4 - \varepsilon_1\varepsilon_5 + \varepsilon_1\varepsilon_6 + \varepsilon_3\varepsilon_5 - \varepsilon_3\varepsilon_6 - \alpha_1\varepsilon_4\varepsilon_5\} \\ &+ \theta_1\{\alpha_1\varepsilon_4\varepsilon_6 - \alpha_1\varepsilon_3\varepsilon_4 + \frac{\alpha_1(\alpha_1+1)}{2}\varepsilon_4^2\} \end{aligned} \tag{61}$$

On taking the expectation on both side of equation (61), we get the bias of the estimator T_{l5} which is given as

$$\begin{aligned} Bias(T_{l5}) =& \bar{Y}E(\varepsilon_0) + \theta_1 \{E(\varepsilon_3) - E(\varepsilon_1) - \alpha_1 E(\varepsilon_4) - E(\varepsilon_1\varepsilon_5) + E(\varepsilon_1\varepsilon_6) + E(\varepsilon_3\varepsilon_5) \\ & -E(\varepsilon_3\varepsilon_6) - \alpha_1 E(\varepsilon_4\varepsilon_5) + \alpha_1(\varepsilon_4\varepsilon_6) - \alpha_1 E(\varepsilon_3\varepsilon_4) + \frac{\alpha_1(\alpha_1+1)}{2}E(\varepsilon_4^2)\Big\} \end{aligned} \tag{62}$$

Now, using the results on the expectations from equation (42-57) and after solving, we get the expression for the bias of the estimator T_{l5} which is given in equation (58).

The estimator T_{l6} is given as

$$\begin{aligned} T_{l6} &= \bar{y}^* + b_{yx}^{**}\left[\bar{x}'\left(\frac{\bar{Z}}{\bar{z}'}\right)^{\alpha_2} - \bar{x}\right] \\ &= \bar{y}^* + \frac{\hat{S}_{yx}}{s_x^2}\left[\bar{x}'\left(\frac{\bar{Z}}{\bar{z}'}\right)^{\alpha_2} - \bar{x}\right] \end{aligned}$$

The estimator T_{l6} can be expressed in terms of ε_l's as
$T_{l6} = \bar{Y}(1+\varepsilon_0) + \frac{S_{yx}(1+\varepsilon_5)}{S_x^2(1+\varepsilon_7)}\left[\bar{X}(1+\varepsilon_3)\left(\frac{\bar{Z}}{\bar{Z}(1+\varepsilon_4)}\right)^{\alpha_2} - \bar{X}(1+\varepsilon_2)\right]$

$$= \bar{Y}(1+\varepsilon_0) + \bar{X}\frac{S_{yx}}{S_x^2}(1+\varepsilon_5)(1+\varepsilon_7)^{-1}\left[(1+\varepsilon_3)(1+\varepsilon_4)^{-\alpha_2} - (1+\varepsilon_2)\right]$$

$$= \bar{Y}(1+\varepsilon_0)+\theta_1(1+\varepsilon_5)(1-\varepsilon_7+\varepsilon_7^2,...)\left[(1+\varepsilon_3)\left(1-\alpha_2\varepsilon_4+\frac{\alpha_2(\alpha_2+1)}{2}\varepsilon_4^2,...\right) - (1+\varepsilon_2)\right]$$

After solving and neglecting the terms of ε_l's having power more than two, we get

$$\begin{aligned} T_{l6} - \bar{Y} =& \bar{Y}\varepsilon_0 + \theta_1\{\varepsilon_3 - \varepsilon_2 - \alpha_2\varepsilon_4 - \varepsilon_2\varepsilon_5 + \varepsilon_2\varepsilon_7 + \varepsilon_3\varepsilon_5 - \varepsilon_3\varepsilon_7 - \alpha_2\varepsilon_4\varepsilon_5 \\ & +\alpha_2\varepsilon_4\varepsilon_7 - \alpha_2\varepsilon_3\varepsilon_4 + \frac{\alpha_2(\alpha_2+1)}{2}\varepsilon_4^2\Big\} \end{aligned} \tag{63}$$

On taking the expectation on both side of equation (63), we get the bias of estimator T_{l6} which is given as

$$\begin{aligned} Bias(T_{l6}) =& \bar{Y}E(\varepsilon_0) + \theta_1 \{E(\varepsilon_3) - E(\varepsilon_2) - \alpha_2 E(\varepsilon_4) - E(\varepsilon_2\varepsilon_5) + E(\varepsilon_2\varepsilon_7) + E(\varepsilon_3\varepsilon_5) \\ & -E(\varepsilon_3\varepsilon_7) - \alpha_2 E(\varepsilon_4\varepsilon_5) + \alpha_2(\varepsilon_4\varepsilon_7) - \alpha_2 E(\varepsilon_3\varepsilon_4) + \frac{\alpha_2(\alpha_2+1)}{2}E(\varepsilon_4^2)\Big\} \end{aligned} \tag{64}$$

Now, using the results on the expectations from equation (42-57) and after solving, we get the expression for the bias of the estimator $T_{l\,6}$ which is given in equation (59). Hence the proof.

Theorem 2: The expressions for mean square errors of the estimators $T_{l\,5}$ and $T_{l\,6}$ up to the terms of order (n^{-1}) are given by

$$\begin{aligned} MSE(T_{l5}) =& V(\bar{y}^*) - \bar{Y}^2 \left[\left(\frac{1}{n} - \frac{1}{n'} \right) \rho_{yx}^2\, C_y^2 - \frac{W_2(k-1)}{n} \left\{ B^2 C_{x(2)}^2 - 2BC_{yx(2)} \right\} \right. \\ & \left. - \frac{f'}{n'} \left(\alpha_1^2 \frac{\rho_{yx}^2 C_y^2 C_z^2}{C_x^2} - 2\alpha_1 \frac{\rho_{yx}\rho_{yz} C_y^2 C_z}{C_x} \right) \right] \end{aligned} \tag{65}$$

and

$$MSE(T_{l6}) = V(\bar{y}^*) - \bar{Y}^2 \left[\left(\frac{1}{n} - \frac{1}{n'} \right) \rho_{yx}^2 C_y^2 - \frac{f'}{n'} \left(\alpha_2^2 \frac{\rho_{yx}^2 C_y^2 C_z^2}{C_x^2} - 2\alpha_2 \frac{\rho_{yx}\rho_{yz} C_y^2 C_z}{C_x} \right) \right] \tag{66}$$

where $B = \frac{\beta}{R}$, $\beta = \frac{S_{yx}}{S_x^2}$, $R = \frac{\bar{Y}}{\bar{X}}$.

Proof: By squaring and taking the expectation on both sides of equation (61), we get

$$\begin{aligned} MSE(T_{l5}) = E[T_{l5} - \bar{Y}]^2 =& E \left[\bar{Y}\varepsilon_0 + \theta_1 \left\{ \varepsilon_3 - \varepsilon_1 - \alpha_1\varepsilon_4 - \varepsilon_1\varepsilon_5 + \varepsilon_1\varepsilon_6 + \varepsilon_3\varepsilon_5 - \varepsilon_3\varepsilon_6 \right. \right. \\ & \left. \left. - \alpha_1\varepsilon_4\varepsilon_5 + \alpha_1\varepsilon_4\varepsilon_6 - \alpha_1\varepsilon_3\varepsilon_4 + \frac{\alpha_1(\alpha_1+1)}{2}\varepsilon_4^2 \right\} \right]^2 \end{aligned} \tag{67}$$

$$\begin{aligned} =& E[\bar{Y}^2\varepsilon_0^2 + \theta_1^2\varepsilon_3^2 + \theta_1^2\varepsilon_1^2 + \theta_1^2\alpha_1^2\varepsilon_4^2 + 2\bar{Y}\theta_1\varepsilon_0\varepsilon_3 - 2\bar{Y}\theta_1\varepsilon_0\varepsilon_1 - 2\bar{Y}\theta_1\alpha_1\varepsilon_0\varepsilon_4 - 2\theta_1^2\varepsilon_1\varepsilon_3 \\ & - 2\theta_1^2\alpha_1\varepsilon_3\varepsilon_4 + 2\theta_1^2\alpha_1\varepsilon_1\varepsilon_4 + ...] \end{aligned}$$

Neglecting the terms of ε_l's having power more than two, we get

$$\begin{aligned} MSE(T_{l5}) =& \bar{Y}^2 E(\varepsilon_0^2) + \theta_1^2 E(\varepsilon_3^2) + \theta_1^2 E(\varepsilon_1^2) + \theta_1^2\alpha_1^2 E(\varepsilon_4^2) + 2\bar{Y}\theta_1 E(\varepsilon_0\varepsilon_3) - 2\bar{Y}\theta_1 E(\varepsilon_0\varepsilon_1) \\ & - 2\bar{Y}\theta_1\alpha_1 E(\varepsilon_0\varepsilon_4) - 2\theta_1^2 E(\varepsilon_1\varepsilon_3) - 2\theta_1^2\alpha_1\varepsilon_3\varepsilon_4 + 2\theta_1^2\alpha_1 E(\varepsilon_1\varepsilon_4)] \end{aligned}$$

Now, using the results on the expectations from equation (42-57) and after solving, we get the expression for mean square error of the estimators $T_{l\,5}$ which is given in equation (65).

Similarly by squaring and taking the expectation on both sides of equations (63), we get

$$\begin{aligned} MSE(T_{l6}) = E[T_{l6} - \bar{Y}]^2 =& E[\bar{Y}\varepsilon_0 + \theta_1\{\varepsilon_3 - \varepsilon_2 - \alpha_2\varepsilon_4 - \varepsilon_2\varepsilon_5 + \varepsilon_2\varepsilon_7 + \varepsilon_3\varepsilon_5 - \varepsilon_3\varepsilon_7 \\ & - \alpha_2\varepsilon_4\varepsilon_5 + \alpha_2\varepsilon_4\varepsilon_7 - \alpha_2\varepsilon_3\varepsilon_4 + \frac{\alpha_2(\alpha_2+1)}{2}\varepsilon_4^2\}]^2 \end{aligned} \tag{68}$$

$$\begin{aligned} =& E[\bar{Y}^2\varepsilon_0^2 + \theta_1^2\varepsilon_3^2 + \theta_1^2\varepsilon_2^2 + \theta_1^2\alpha_2^2\varepsilon_4^2 + 2\bar{Y}\theta_1\varepsilon_0\varepsilon_3 - 2\bar{Y}\theta_1\varepsilon_0\varepsilon_2 - 2\bar{Y}\theta_1\alpha_2\varepsilon_0\varepsilon_4 - 2\theta_1^2\varepsilon_2\varepsilon_3 \\ & - 2\theta_1^2\alpha_2\varepsilon_3\varepsilon_4 + 2\theta_1^2\alpha_2\varepsilon_2\varepsilon_4 + ...] \end{aligned}$$

Neglecting the terms of ε_l's having power more than two, we get

$$MSE(T_{l6}) = \bar{Y}^2 E(\varepsilon_0^2) + \theta_1^2 E(\varepsilon_3^2) + \theta_1^2 E(\varepsilon_2^2) + \theta_1^2 \alpha_2^2 E(\varepsilon_4^2) + 2\bar{Y}\theta_1 E(\varepsilon_0\varepsilon_3) - 2\bar{Y}\theta_1 E(\varepsilon_0\varepsilon_2)$$
$$-2\bar{Y}\theta_1\alpha_2 E(\varepsilon_0\varepsilon_4) - 2\theta_1^2 E(\varepsilon_2\varepsilon_3) - 2\theta_1^2\alpha_2 E(\varepsilon_3\varepsilon_4) + 2\theta_1^2\alpha_2 E(\varepsilon_2\varepsilon_4)]$$

Now, using the results on the expectations from equation (42-57) and after solving, we get the expression for mean square error of the estimators $T_{l\,6}$, which is given in equation (66). Hence the proof.

Theorem 3: The minimum mean square errors of the estimators T_{l5} and T_{l6} are given by

$$MSE(T_{l5})_{\min} = V(\bar{y}^*) - \bar{Y}^2 \left[\left(\frac{1}{n} - \frac{1}{n'}\right)\rho_{yx}^2\, C_y^2 - \frac{W_2(k-1)}{n}\left\{B^2 C_{x(2)}^2 - 2BC_{yx(2)}\right\} + \frac{f'}{n'}\rho_{yz}^2 C_y^2\right] \tag{69}$$

and

$$MSE(T_{l6})_{\min} = V(\bar{y}^*) - \bar{Y}^2 \left[\left(\frac{1}{n} - \frac{1}{n'}\right)\rho_{yx}^2 C_y^2 + \frac{f'}{n'}\rho_{yz}^2 C_y^2\right] \tag{70}$$

Proof: On differentiating the expressions of $MSE(T_{l5})$ and $MSE(T_{l6})$ given in equations (65) and (66) with respect to α_1 and α_2 we get
$\alpha_{1opt} = \frac{\rho_{yz}}{\rho_{yx}}\frac{C_x}{C_z}$ and $\alpha_{2opt} = \frac{\rho_{yz}}{\rho_{yx}}\frac{C_x}{C_z}$
Now, after putting the optimum values of α_1 and α_2 in the expressions of $MSE(T_{l5})$ and $MSE(T_{l6})$ which are given in equations (65) and (66), we get the minimum mean square errors of the estimators T_{l5} and T_{l6}, which are given in equations (69) and (70). Hence the proof.

References

Dash, P. R. and Mishra, G. (2011). An improved class of estimators in two phase sampling using two auxiliary variables. *Communication in Statistics - Theory Methods*, 24(40):4347–4352.

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

Khare, B. and Kumar, S. (2010). Chain regression type estimators using additional auxiliary variable in two phase sampling in the presence of non-response. *National Academy Science Letters, India*, 33(11):369–375.

Khare, B. and Kumar, S. (2011). Estimation of population mean using known coefficient of variation of the study character in the presence of non-response. *Communication in Statistics - Theory Methods*, 40(11):2044–2058.

Khare, B. and Srivastava, S. (1995). Study of conventional and alternative two phase sampling ratio, product and regression estimators in presence of nonresponse. *Proceedings of the National Academy Sciences, India*, 65(11):195–203.

Khare, B. and Srivastava, S. (1997). Transformed ratio type estimators for population mean in the presence of nonresponse. *Communication in Statistics - Theory Methods*, 26(7):1779–1791.

Khare, B. and Srivastava, S. (2010). Generalized two phase sampling estimators for the population mean in the presence of nonresponse. *Aligarh Journal of Statistics*, 30:39–54.

Khare, B., Srivastava, U., and Kumar, K. (2011). Generalized chain estimators for the population mean in the presence of non-response. *Proceedings of the National Academy Sciences, India*, 81:231–238.

Khare, B., Srivastava, U., and Kumar, K. (2012). An improved class of estimators for population mean using auxiliary character in the presence of non-response. *National Academy Science Letters, India*, 35:361–366.

Kiregyera, B. (1980). A chain ratio type estimator in finite population double sampling using two auxiliary variables. *Metrika*, 27(7):217–223.

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

Mishra, G. and Rout, K. (1997). A regression estimator in two phase sampling in presence of two auxiliary variables. *Metron*, 12:177–186.