Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/070205

A Modified Generalized Chain Regression Cum Ratio Estimator for Population Mean in the Presence of Nonresponse

B. B. Khare, Utkarsh Utkarsh and Supriya Khare*

Department of Statistics, Institute of Science, B.H.U, Varanasi-221005, India

Received: 24 Nov. 2017, Revised: 12 Feb. 2018, Accepted: 27 Feb. 2018
Published online: 1 Jul. 2018

Abstract: In this paper, a modified generalized chain regression cum ratio estimator for population mean in the presence of nonresponse has been proposed and its properties have been studied. A comparative study of the proposed estimator has been carried out with the relevant estimators and the optimum values of $n^{'}$, n and k have been obtained for the fixed cost $C \leq C_0$. For the several suitable choice of the constant a, the proposed estimator is found to be more efficient than the relevant estimators. The empirical and theoretical studies also justify the efficiency of the proposed estimator in comparison to the relevant estimators.

Keywords: Mean square error, Chain estimators, auxiliary character, additional auxiliary character, Non-response.

1 Introduction

The use of an auxiliary character to increase the efficiency of the estimator for population parameters is widely used in researches related to the field of socioeconomic, agriculture and biomedical sciences. The research work on the estimation of population parameters using auxiliary characters by several authors have been reviewed by [30] and [8]. For an example, in a forest surveys, the average amount of timber of a tree can be estimated by using the diameter of the tree as an auxiliary character. In some cases anyone may be interested in estimating the population mean of the study character (y) when the population mean of auxiliary character (x) is not known and the population mean of another additional auxiliary character (z) is known which may be cheaper and less correlated to y than x ($\rho_{yx} > \rho_{yz}$). For the first time, [1] has proposed an estimator for population mean using an additional auxiliary character. In this context, several estimators for population mean have been proposed by [17], [28], [21], [3] and [9] etc. Further the research work has been reviewed by [13].

While conducting a sample survey, it may not be possible to collect information on all the units selected in the sample due to nonresponse. In this case, [2] have suggested a method of sub-sampling from non-respondents and the estimator for population mean based on available information on responding units and sub sample units drawn from the non-responding units in the sample has been proposed. Further several problems on estimating the population mean through sample surveys in the presence of non-response have been considered by [18],[19], [4], [5], [6], [7], [22], [23], [24], [25] and [10]. The study of several chain type estimators and improved chain type estimators for population mean using auxiliary character and additional auxiliary character in the presence of nonresponse have been studied by [9], [11], [14]. The Generalized Chain ratio in regression estimators for the population mean using two phase sampling in the presence of non-response have been proposed by [14].

In the present context, we have proposed a modified generalized chain regression cum ratio estimator for population mean using two phase sampling in presence of nonresponse. The properties of the proposed estimator have been studied for fixed sample sizes ($n^{'}$, n) and also for the fixed cost $C \leq C_0$. Several estimators are found to be the member of the

* Corresponding author e-mail: utkarshparmar11@gmail.com

proposed estimator as a special case. A comparative study of the proposed estimator for the several suitable choice of the constant “a” has been made with the relevant estimators. An empirical study has also been given in the support of the problem for fixed sample sizes $(n^{'}, n)$ and also for fixed the cost $C \leq C_0$.

2 Proposed Estimators

Let y, x and z denote the study character, auxiliary character and additional auxiliary character respectively having $j-th$ values Y_j, X_j and Z_j $(j = 1,2,.....N)$ with their population means $\bar{Y}$, $\bar{X}$ and $\bar{Z}$ respectively. The population of size N is supposed to be divided into N_1 responding and N_2 non-responding units such that $N_1 + N_2 = N$. According to the [2], a sample of size n is drawn from the population of size N by using simple random sampling without replacement (SRSWOR) scheme, it has been observed that only n_1 units are responding and n_2 units are not responding in the sample of size n for the study character y. Further, by making some extra effort a sub-sample of size $r(r = n_2/k, k > 1)$ is drawn from $n_2$non-responding units by using SRSWOR sampling scheme and related information on available units are collected by personal interview for study character y.

Similarly, in case when the population mean of the auxiliary character is not known we draw the first phase sample of size $n^{'}(< N)$ from the population of size N by using simple random sampling without replacement (SRSWOR) scheme and estimate the population mean $\bar{X}$ by the first phase sample mean $\bar{x}^{'}$ based on $n^{'}$ units.

Further, we draw the second phase sample of size $n(< n^{'})$ from first phase sample of size $n^{'}$ by using SRSWOR sampling scheme and observed that only n_1 units are responding and n_2 units are not responding for the study character y. Again we draw the sub-sample of size $r(r = n_2/k, k > 1)$ is drawn from n_2 non-responding units by using SRSWOR sampling scheme and collect the related information on r units using personal interview method by making extra effort.

Using [2] technique of sub sampling from nonrespondents, the estimator for $\bar{Y}$ based on $n_1 + r$ observations on the study character y is given 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 character y based on n_1 and r units respectively. The variance of the estimator is given by

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

where $f = (1-\frac{n}{N})$, $W_2 = \frac{N_2}{N}$, S_y^2 and $S*_y^2$ are the population mean squares of study character y for the entire population and for the non-responding part of the population.
Now, the sample mean of values of x corresponding to $n_1 + r$ values on y is given as follows:

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

where $\bar{x}_1$,$\bar{x}_2^{'}$ denotes the means of auxiliary character (x) based on n_1 and r units . The variance of the estimator is given by

$$V(\bar{x}^{*}) = \frac{f}{n}S_x^2 + \frac{W_2(k-1)}{n}S_x^{*2}, \tag{4}$$

where S_x^2 and $S*_x^2$ are the population mean squares of study character x for the entire population and for the non-responding part of the population.
[25], have proposed some improved estimators using auxiliary information and additional auxiliary information under two phase sampling scheme which are given as follows:

$$t_{01} = \{\bar{y}^{*} + b_{yx}(\bar{x}^{'} - \bar{x}^{*})\}\left(\frac{\bar{Z}}{\bar{Z} + \delta_1(\bar{z}^{*} - \bar{Z})}\right), \tag{5}$$

where δ_1 is constant and b_{yx} denotes the estimate of regression coefficient of y on x.

[11] proposed general class of estimator for the population mean using main and additional auxiliary character which is given as follows:

$$t_{02} = \bar{y}^* \left(\frac{\bar{x}^*}{\bar{x}'}\right)^{\alpha_1} \left(\frac{\bar{z}'}{\bar{Z}}\right)^{\alpha_2} \tag{6}$$

and

$$t_{03} = \bar{y}^* \left(\frac{\bar{x}}{\bar{x}'}\right)^{\alpha_3} \left(\frac{\bar{z}'}{\bar{Z}}\right)^{\alpha_4} \tag{7}$$

where $\alpha_1, \alpha_2, \alpha_3$ *and* α_4 constants.
Further, [16] have proposed a generalized chain ratio and regression type estimator for population mean using auxiliary character in presence of nonresponse which is given as follows:

$$t_{04} = \bar{y}^* + b_{yx}\left\{\bar{x}' \left(\frac{\bar{Z}}{\bar{z}'}\right)^{\alpha_5} - \bar{x}^*\right\}, \tag{8}$$

where α_5 is constant. Further, [15] have extended t_{04} and proposed an improved ratio and regression type estimator for population mean using information on the coefficient of variation of the study character which is given as follows:

$$t_{05} = \bar{y}^{**} + b_{yx}\left\{\bar{x}' \left(\frac{\bar{Z}}{\bar{z}'}\right)^{\alpha_6} - \bar{x}^*\right\}, \tag{9}$$

where $\bar{y}^{**} = k\bar{y}^*$, k *and* α_6 are constants.

$$k_{opt.} = \left[1 + \frac{f}{n}C_y^2 + \frac{W_2(k-1)}{n}C{*}_y^2\right]^{-1}, \tag{10}$$

Further [14] have proposed some generalized chain ratio and regression estimators follows:

$$T_1 = \{\bar{y}^* + b_{yx}(\bar{x}' - \bar{x}^*)\}\left(\frac{\bar{Z}}{\bar{Z} + \delta_2(\bar{z}' - \bar{Z})}\right), \tag{11}$$

$$T_3 = \{\bar{y}^* + b_{yx}\left\{\bar{x}'\left(\frac{\bar{Z}}{\bar{Z} + \delta_3(\bar{z}' - \bar{Z})}\right) - \bar{x}^*\right\}, \tag{12}$$

and

$$T_5 = \{\bar{y}^* + b_{yx}(\bar{x}' - \bar{x}^*)\}\left(\frac{\bar{Z}}{\bar{z}'}\right)^{\delta_4} \tag{13}$$

where δ_1, δ_2, δ_3 *and* δ_4 are constants.
In the present context, we have proposed a modified generalized chain regression cum ratio estimator for population mean in the presence of nonresponse which is given as follows:

$$T = \left[a\bar{y}^* + b(\bar{x}' - \bar{x}^*)\right] \left(\frac{\bar{z}^*}{\bar{z}'}\right)^{\alpha_{01}} \left(\frac{\bar{z}'}{\bar{Z}}\right)^{\alpha_{02}} \tag{14}$$

where a, α_{01} *and* α_{02} are constants and b denotes the estimate of regression coefficient of y on x.

Case I: Using a=1, we find that T reduces to T_1 which is more efficient estimator in comparison to other estimators which is given as follows:

$$T_{a=1} = \left[\bar{y}^* + b(\bar{x}' - \bar{x}^*)\right] \left(\frac{\bar{z}^*}{\bar{z}'}\right)^{\alpha_{01}} \left(\frac{\bar{z}'}{\bar{Z}}\right)^{\alpha_{02}} \tag{15}$$

Case II: If the value of a, $\left(a = a_0 = \frac{n}{n+C_y^2}\right)$ is taken which minimizes the $MSE(a\bar{y})$ and T reduces to T_{a0} which is given as follows:

$$T_{a=a_0} = \left[a_0\bar{y}^* + b(\bar{x}' - \bar{x}^*)\right] \left(\frac{\bar{z}^*}{\bar{z}'}\right)^{\alpha_{01}} \left(\frac{\bar{z}'}{\bar{Z}}\right)^{\alpha_{02}} \tag{16}$$

Case III: If we take the optimum value of$a$, $\left(a = a_0^* = \frac{n}{n+C_y^2+W_2(k-1)C_y^{*2}}\right)$ is taken which minimizes the $V(a\bar{y}^*)$ then T reduces to $T_{a_0^*}$ which is given as follows:

$$T_{a=a_0^*} = \left[a\bar{y}^* + b(\bar{x}' - \bar{x}^*)\right] \left(\frac{\bar{z}^*}{\bar{z}'}\right)^{\alpha_{01}} \left(\frac{\bar{z}'}{\bar{Z}}\right)^{\alpha_{02}} \tag{17}$$

3 Bias and mean square error of the proposed estimator T

In order to derive the expression for Bias and Mean square error of the proposed estimator T using large sample approximation, we have:
Let$\bar{y}^* = \bar{Y}(1+\varepsilon_0)$,$\bar{x}^* = \bar{X}(1+\varepsilon_1)$,$\bar{x}' = \bar{X}(1+\varepsilon_2)$,$\bar{z}' = \bar{Z}(1+\varepsilon_3)$,$\bar{z}^* = \bar{Z}(1+\varepsilon_4)$,$b = \frac{s_{xy}}{s_{x^2}} = \frac{S_{xy}}{S_{x^2}}\frac{(1+\varepsilon_5)}{(1+\varepsilon_6)} = \beta\frac{(1+\varepsilon_5)}{(1+\varepsilon_6)}$, such that $E(\varepsilon_l) = 0$and $|\varepsilon_l|<1\ \forall \quad l = 0,1,2,3,4,5,6.$

$$T = [a\bar{Y}(1+\varepsilon_0) + b\bar{X}(\varepsilon_2-\varepsilon_1)]\ \left((1+\varepsilon_4)(1+\varepsilon_3)^{-1}\right)^{\alpha 01}(1+\varepsilon_3)^{\alpha 02} \tag{18}$$

$$\begin{aligned} T = &\left[a\bar{Y} + a\bar{Y}\varepsilon_0 + \beta(1+\varepsilon_5)(1+\varepsilon_6)^{-1}\bar{X}(\varepsilon_2-\varepsilon_1)\right] \\ [\quad &\left(1+\alpha_{01}\varepsilon_4 + \frac{\alpha_{01}(\alpha_{01}-1)}{2}\varepsilon_4^2 - \alpha_{01}\varepsilon_3 - \alpha_{01}^2\varepsilon_3\varepsilon_4 + \frac{\alpha_{01}(\alpha_{01}+1)}{2}\varepsilon_3^2\right)] \\ &\left((1+\alpha_{02}\varepsilon_3 + \frac{\alpha_{02}(\alpha_{02}-1)}{2}\varepsilon_3^2)\right) \end{aligned} \tag{19}$$

Now retaining the terms of $\varepsilon's$ up to second degree, we have:

$$\begin{aligned} T = &a\bar{Y} + a\bar{Y}\varepsilon_0 + \beta\bar{X}\varepsilon_2 - \beta\bar{X}\varepsilon_1 + a\bar{Y}(\alpha_{02}-\alpha_{01})\varepsilon_3 + a\bar{Y}(\alpha_{02}-\alpha_{01})\varepsilon_3\varepsilon_0 \\ &+\beta\bar{X}(\alpha_{02}-\alpha_{01})\varepsilon_3\varepsilon_2 - \beta\bar{X}(\alpha_{02}-\alpha_{01})\varepsilon_3\varepsilon_1 + a\bar{Y}\alpha_{01}\varepsilon_4 + a\bar{Y}\alpha_{01}\varepsilon_4\varepsilon_0 \\ &+\beta\bar{X}\alpha_{01}(\varepsilon_4\varepsilon_2 - \varepsilon_4\varepsilon_1)\ + a\bar{Y}(\alpha_{01}-\alpha_{02})\{\tfrac{1}{2}(\alpha_{01}-\alpha_{02}+1)\varepsilon_3^2 - \alpha_{01}\varepsilon_3\varepsilon_4\} \\ &+a\bar{Y}\frac{\alpha_{01}(\alpha_{01}-1)}{2}\varepsilon_4^2 \end{aligned} \tag{20}$$

The expression for $Bias(T)$ up to the terms of order n^{-1} is given by:

$$\begin{aligned} Bias(T) = &E[\{\bar{Y}(a-1)\} + a\bar{Y}(\alpha_{02}-\alpha_{01})(\varepsilon_0\varepsilon_3) + \alpha_{01}a\bar{Y}\varepsilon_0\varepsilon_4 - \alpha_{01}\beta\bar{X}(\varepsilon_1\varepsilon_4 - \varepsilon_2\varepsilon_4) \\ &-\beta\bar{X}(\alpha_{02}-\alpha_{01})(\varepsilon_3\varepsilon_1 - \varepsilon_3\varepsilon_2) + a\bar{Y}\frac{\alpha_{01}(\alpha_{01}-1)}{2}\varepsilon_4^2 \\ &+a\bar{Y}(\alpha_{01}-\alpha_{02})\{\tfrac{1}{2}(\alpha_{01}-\alpha_{02}+1)\varepsilon_3^2 - \alpha_{01}\varepsilon_3\varepsilon_4\}] \end{aligned} \tag{21}$$

$$\begin{aligned} &Bias(T) = [\{\bar{Y}(a-1)\} - a\bar{Y}(\alpha_{01}-\alpha_{02})\frac{f'}{n'}C_{yz} + \alpha_{01}a\bar{Y}(\frac{f}{n}C_{yz} + \frac{W_2(k-1)}{n}C_{yz}^*) \\ &-\alpha_{01}\beta\bar{X}((\frac{f}{n}-\frac{f'}{n'})C_{xz} + \frac{W_2(k-1)}{n}C_{xz}^*) + a\bar{Y}\frac{\alpha_{01}(\alpha_{01}-1)}{2}(\frac{f}{n}C_z^2 + \frac{W_2(k-1)}{n}C_z^{*2}) \\ &+a\bar{Y}(\alpha_{01}-\alpha_{02})\frac{f'}{n'}C_z^2(1-\alpha_{02})] \end{aligned} \tag{22}$$

The expression for $MSE(T)$ is given by:

$$\begin{aligned} MSE(T) = &E[\bar{Y}^2(a-1)^2 + \{\bar{Y}^2a^2\varepsilon_0^2 \\ &+\bar{Y}^2a^2\{(\alpha_{02}-\alpha_{01})^2\varepsilon_3^2 + 2\alpha_{01}(\alpha_{02}-\alpha_{01})\varepsilon_3\varepsilon_4 + \alpha_{01}^2\varepsilon_4^2 + 2(\alpha_{02}-\alpha_{01})\varepsilon_0\varepsilon_3 \\ &+2\alpha_{01}\varepsilon_0\varepsilon_4\} + \bar{X}^2\beta^2(\varepsilon_1^2 + \varepsilon_2^2 - 2\varepsilon_1\varepsilon_2) - 2\bar{X}\bar{Y}a\beta(\varepsilon_0\varepsilon_1 - \varepsilon_0\varepsilon_2 \\ &+\alpha_{01}(\varepsilon_1\varepsilon_4 - \varepsilon_2\varepsilon_4))\}\ + 2\bar{Y}(a-1)\{a\bar{Y}((\alpha_{02}-\alpha_{01})\varepsilon_0\varepsilon_3 + \alpha_{01}\varepsilon_0\varepsilon_4 \\ &+\tfrac{1}{2}(\alpha_{01}-\alpha_{02})(\alpha_{01}-\alpha_{02}+1)\varepsilon_3^2\ + \frac{\alpha_{01}(\alpha_{01}-1)}{2}\varepsilon_4^2 \\ &+\alpha_{01}(\alpha_{02}-\alpha_{01})\varepsilon_3\varepsilon_4) - \beta\bar{X}\alpha_{01}(\varepsilon_1\varepsilon_4 - \varepsilon_2\varepsilon_4)\}] \end{aligned} \tag{23}$$

$$\begin{aligned} &= \bar{Y}^2[(a-1)^2 + a^2\{(\frac{f}{n}C_y^2 + \frac{W_2(k-1)}{n}C_y^{*2}) + \alpha_{01}^2((\frac{f}{n}-\frac{f'}{n'})C_z^2 + \frac{W_2(k-1)}{n}C_z^{*2}) + \alpha_{02}^2(\frac{f'}{n'}C_z^2) + \\ &2(\alpha_{02}-\alpha_{01})(\frac{f'}{n'}C_{yz}) + 2\alpha_{01}(\frac{f}{n}C_{yz} + \frac{W_2(k-1)}{n}C_{yz}^*)\} + 2a(a-1)\{\alpha_{02}(\frac{f'}{n'}C_{yz}) + \alpha_{01}((\frac{f}{n}-\frac{f'}{n'})C_{yz} \\ &+\frac{W_2(k-1)}{n}C_{yz}^*) + \frac{\alpha_{02}(\alpha_{02}-1)}{2}(\frac{f'}{n'}C_z^2) + \frac{\alpha_{01}(\alpha_{01}-1)}{2}((\frac{f}{n}-\frac{f'}{n'})C_z^2 \\ &+\frac{W_2(k-1)}{n}C_z^{*2})\}] + \bar{X}^2\beta^2((\frac{f}{n}-\frac{f'}{n'})C_x^2 + \frac{W_2(k-1)}{n}C_x^{*2}) \\ &-2\bar{X}\bar{Y}\beta\{((\frac{f}{n}-\frac{f'}{n'})(\alpha_{01}(2a-1)C_{xz} + aC_{yx}) + \frac{W_2(k-1)}{n}(\alpha_{01}(2a-1)C_{xz}^* + aC_{yx}^*)\} \end{aligned} \tag{24}$$

Here we denote:

$$\begin{array}{ll} A=((\frac{f}{n}-\frac{f'}{n'})C_z^2+\frac{W_2(k-1)}{n}C_z^{*2}), & B=((\frac{f}{n}-\frac{f'}{n'})C_{yz}+\frac{W_2(k-1)}{n}C_{yz}^{*}), \\ C=((\frac{f}{n}-\frac{f'}{n'})C_x^2+\frac{W_2(k-1)}{n}C_x^{*2}), & D=((\frac{f}{n}-\frac{f'}{n'})C_{xz}+\frac{W_2(k-1)}{n}C_{xz}^{*}) \\ and \quad E=((\frac{f}{n}-\frac{f'}{n'})C_{xy}+\frac{W_2(k-1)}{n}C_{xy}^{*}) & \end{array} \tag{25}$$

The expression for $MSE(T)$ is given as follows:

$$\begin{aligned} MSE(T)=\bar{Y}^2[(a-1)^2+a^2\{(\tfrac{f}{n}C_y^2+\tfrac{W_2(k-1)}{n}C_y^{*2}) \\ +\alpha_{01}^2A+\alpha_{02}^2(\tfrac{f'}{n'}C_z^2)+2\alpha_{02}(\tfrac{f'}{n'}C_{yz})+2\alpha_{01}B\}+2a(a-1)\{\alpha_{02}(\tfrac{f'}{n'}C_{yz}) \\ +\alpha_{01}B+\tfrac{\alpha_{02}(\alpha_{02}-1)}{2}(\tfrac{f'}{n'}C_z^2)+\tfrac{\alpha_{01}(\alpha_{01}-1)}{2}A\}]\bar{X}^2\beta^2C-2\bar{X}\bar{Y}\beta\{\alpha_{01}(2a-1)D+aE\} \end{aligned} \tag{26}$$

The optimum values of α_{01} and α_{02} obtained by differentiating MSE (T), given by (26) , with respect to α_{01}& α_{02} and equating to zero, we have:

$$\alpha_{01opt.}=\frac{1}{A}\left(-B+\frac{\bar{X}\beta D}{\bar{Y}a}\right)+\frac{(a-1)}{(2a-1)}, \tag{27}$$

$$\alpha_{02opt.}=-\left(\frac{C_{yz}}{C_z^2}\right)+\frac{(a-1)}{2(2a-1)} \tag{28}$$

Noting that the second partial derivatives of MSE(T) with respect to α_{01} and α_{02} are greater than zero for α_{01opt} and α_{02opt} . The optimum values of α_{01opt} and α_{02opt} depends on the parameters of the population. So one may use on the bases of the prior information available for the parameters from the past data or experiences which does not change the minimum value of MSE (.) up to terms of order n^{-1} [20]. However, if no prior information is available then we may estimate the parameters on the bases of sample values and it has been observed that substitution of the constants based on these estimated values of the parameters does not affect the minimum value of Mean Square Error up to the term of order n^{-1} [27]. For α_{01opt} and α_{02opt} , the expression for $MSE(T)_{min}$ is given by:

$$\begin{aligned} MSE(T)_{min}=\bar{Y}^2[(a-1)^2+a^2\{(\tfrac{f}{n}C_y^2+\tfrac{W_2(k-1)}{n}C_y^{*2}) \\ +\alpha_{01(opt.)}^2A+\alpha_{02(opt.)}^2(\tfrac{f'}{n'}C_z^2)+2\alpha_{02(opt.)}(\tfrac{f'}{n'}C_{yz})+2\alpha_{01(opt.)}B\}+2a(a-1)\{\alpha_{02(opt.)}(\tfrac{f'}{n'}C_{yz}) \\ +\alpha_{01(opt.)}B+\tfrac{\alpha_{02(opt.)}(\alpha_{02(opt.)}-1)}{2}(\tfrac{f'}{n'}C_z^2)+\tfrac{\alpha_{01(opt.)}(\alpha_{01(opt.)}-1)}{2}A\}]\bar{X}^2\beta^2C-2\bar{X}\bar{Y}\beta\{\alpha_{01(opt.)}(2a-1)D+aE\} \end{aligned} \tag{29}$$

The value of $MSE(T_{a0})$ and $MSE(T_{a0}^{*})$ can be obtained easily by putting the value of constant “a” in the expression of MSE(T) given by (29).

4 Determination of optimum values of n', n and k for the proposed and relevant estimator for fixed cost $(C \leq C_0)$,

Let us assume that C_0 be the total cost (fixed) of the survey apart from overhead cost. The cost function C' is given by:

$$C'=(c_1'+c_2')n'+c_1n+c_2n_1+c_3\frac{n_2}{k} \tag{30}$$

The expected total cost of the survey apart from overhead cost is given by:

$$C=E(C')=(c_1'+c_2')n'+n(c_1+c_2W_1+c_3\frac{W_2}{k}) \tag{31}$$

It is to be noted here that $c_2'<c_1'<c_1<c_2<c_3$,
where

$c_1^{'}$: is the cost per unit of identifying and observing auxiliary character x at the first phase.
$c_2^{'}$: is the cost per unit of obtaining information on additional auxiliary character z at the first phase.
c_1 : is the cost per unit of mailing questionnaire/visiting the units at the second phase for the study character y.
c_2 : is the cost per unit of collecting and processing data obtained from n_1 responding units for the study character y.
c_3 : is the cost per unit of obtaining and processing data for the sub sampling units for the study character y **.**
The expression for $MSE(E_i), i = 1,2,3,4,5,6,7\&8$, can be expressed in terms of $V_0(E_i), V_1(E_i), V_2(E_i)$ and $V_3(E_i)$ which is given as:

$$MSE(E_i) = \frac{V_0(E_i)}{n} + \frac{V_1(E_i)}{n'} + \frac{kV_2(E_i)}{n} - \frac{V_3(E_i)}{N}, \tag{32}$$

where $V_0(E_i), V_1(E_i), V_2(E_i)$ and $V_3(E_i)$ are the coefficient of $\frac{1}{n}, \frac{1}{n'}, \frac{k}{n}$ *and* $\frac{1}{N}$ respectively in the expression of $MSE(E_i)$. Here we denote $E_1 = \bar{y}^*$, $E_2 = t_1$, $E_3 = T_1$, $E_4 = T_3$, $E_5 = T_5$, $E_6 = T_{a=1}$, $E_7 = T_{a=a0}$ and $E_8 = T_{a=a_0^*}$.
Now we define a function Φ for minimizing the $MSE(E_i)$ for the fixed cost to obtain the optimum values of n', n and k, The function Φ is given as:

$$\Phi = MSE(E_i) + \lambda_i\{(c_1^{'} + c_2^{'})n^{'} + n(c_1 + c_2W_1 + c_3\frac{W_2}{k})\}, \tag{33}$$

where $\lambda_i i = 1,2,3,4,5,6,7\&8$ is the Lagrange's multiplier.
Now differentiating Φ(33) with respect to n', n and k then equating to zero we get the optimum values of n ,$n^{'}$ and k which are given as follows:

$$n^{'}_{(opt.)} = \sqrt{\frac{V_1(E_i)}{\lambda_i(c_1^{'} + c_2^{'})}}, n_{(opt.)} = \sqrt{\frac{V_o(E_i) + kV_2(E_i)}{\lambda_i(c_1 + c_2W_1 + c_3\frac{W_2}{k})}} \& k_{(opt.)} = \sqrt{\frac{V_o(E_i)c_3W_2}{V_2(E_i)(c_1 + c_2W_1)}}, \tag{34}$$

where

$$\sqrt{\lambda_i} = \frac{1}{C_o}\left\{\sqrt{V_1(E_i)(c_1^{'} + c_2^{'})} + \sqrt{(V_0(E_i) + k_{opt}V_2(E_i))}\sqrt{(c_1 + c_2W_1 + c_3\frac{W_2}{k})}\right\}, \tag{35}$$

For the optimum values of $n^{'}, n$, k and neglecting the term of order (N^{-1}), the $MSE(E_i)_{\min.}$ is given as follows:

$$MSE(E_i)_{\min} = [\frac{1}{C_o}\left\{\sqrt{V_1(E_i)(c_1^{'} + c_2^{'})} + \sqrt{(V_0(E_i) + k_{opt}V_2(E_i))}\sqrt{(c_1 + c_2W_1 + c_3\frac{W_2}{k})}\right\}^2] \tag{36}$$

5 An Empirical Study

One Hundred and Nine Village/Town/ ward wise population of the urban area under police- station-Baria, Tahasil-Champua, Orissa has been taken under consideration from District Census Handbook, 1981, Orissa, published by Govt. of India, [26]. The last 25 percent villages (i.e. 27 villages) have been considered as non-response group of the population. The problem considered here is to estimate the average number of male workers in the village (y), using total population of the village (x) as auxiliary character and number of cultivators in the village (z) as additional auxiliary character. The values of the parameters of the population under study are given as follows:

$\bar{Y}$=165.2661, $\bar{X}$=485.9174, $\bar{Z}$=100.5505, C_y=0.6828, C_x=0.6590, C_z=0.7314, C_y^*=0.5769, C_x^*=0.4877, C_z^*=0.5678, ρ_{yx}=0.908, ρ_{yz}=0.841, ρ_{xz}=0.801, ρ_{yx}^*=0.907, ρ_{yz}^*=0.785, ρ_{xz}^*=0.654.

Figures in parentheses give the MSE(.).
From table 1, we observe that for the fixed sample sizes ($n^{'}, n$) and for different values of k, the proposed estimators $T_{a=1}$, $T_{a=a_0}$, and $T_{a=a_0^*}$ are more efficient than the relevant estimators $\bar{y}^*$, t_1, T_1, T_3, *and* T_5. It is observed that the MSE of the proposed estimator $T_{a=a_0^*}$ is minimum and comparatively more efficient than all relevant estimators. It is also observed that the value of MSE (.) of the proposed and relevant estimators decreases as the value of k decreases. The relative

efficiency of all estimators with respect to $\bar{y}^*$ decreases as k decreases. This is due to the fact that the MSE $(\bar{y}^*)$ decreases with the faster rate as other estimators as k decreases.

Table 1: Relative efficiency (in percentage) of the estimators with respect to $\bar{y}^*$ for the different values of k.

Estimators	1/4	1/3	1/2
$\bar{y}^*$	100 (532.0214)	100 (457.2255)	100 (382.4296)
t_1	254.7804 (208.8157)	234.1385 (195.2799)	210.4219 (181.7442)
T_1	477.9368 (111.3162)	467.6039 (97.7805)	453.9507 (84.2447)
T_3	420.8076 (126.4286)	405.0082 (112.8929)	384.9040 (99.3571)
T_5	477.9368 (111.3162)	467.6039 (97.7805)	453.9507 (84.2447)
$T_{a=1}$	494.0631(107.6829)	476.9807 (95.8583)	455.2835 (83.9981)
$T_{a=a_0}$	507.3863 (104.8553)	481.0052 (95.0562)	450.1805 (84.9503)
$T_{a=a_0^*}$	523.8823(101.5536)	500.0282 (91.4399)	472.0542 (81.0139)

Table 2: Relative efficiency (in percentage) of the estimators with respect to $(\bar{y}^*)$ for fixed cost $(C \leq C_0)$, C=Rs.300 respectively.where $c_1^{'} = Rs.0.40$, $c_2^{'} = Rs.0.30$, $c_1 = Rs.0.50$, $c_2 = Rs.1.75$, $c_3 = Rs.45$

Estimators	k_{opt}(approx.)	For fixed cost C_0	For fixed cost C_0	Relative efficiency (In %)
		$n_{opt}^{\prime}$ (approx.)	n_{opt} (approx.)	
$\bar{y}^*$	5.355994	0	76	100 (292.4312)
t_1	5.31691	173	45	196.8114 (148.5844)
T_1	5.31691	86	61	353.9355 (82.62272)
T_3	5.31691	112	56	302.8088 (96.57288)
T_5	5.31691	86	61	353.9355 (82.62272)
$T_{a=1}$	6.57019	88	67	391.6860 (74.65959)
$T_{a=a_0}$	6.64258	85	68	406.1209 (72.00593)
$T_{a=a_0^*}$	6.66068	85	69	409.8295 (71.35435)

Figures in parentheses give the MSE(.), Rs. =Rupees (Indian currency)

From table 2, it has been observed that for the fixed cost, the proposed estimators $T_{a=1}, T_{a=a_0}$ *and* $T_{a=a_0^*}$ have also less mean square error in comparison to the relevant estimators $\bar{y}^*$, t_1, T_1, T_3, *and* T_5. However $T_{a=a_0^*}$ is more efficient than $T_{a=1}$*and* $T_{a=a_0}$.

6 Perspective

On the basis of empirical and theoretical studies, we have observed that the proposed estimators $T_{a=1}, T_{a=a_0}$ *and* $T_{a=a_0^*}$ are more efficient than the relevant estimatorst_1, T_1, T_3, *and* T_5 for the fixed sample sizes $(n^{'}, n)$ and also for the fixed cost$(C \leq C_0)$,. Thus, we suggest to use the members $T_{a=1}, T_{a=a_0}$ *and* $T_{a=a_0^*}$of the proposed generalized chain regression cum ratio estimator for population mean in the presence of nonresponse for the different values of a as 1, $a_0$and a_0^*. The estimator $T_{a=1}$ is preferred in comparison to relevant estimators when no prior information on $C_y$is available. However if prior information on $C_y$is available then the use of $T_{a=a_0}$ is preferred and if we have prior information on C_y, W_2 *and* C_y^* then the use of $T_{a=a_0^*}$ is more preferred. Hence, we conclude to prefer the use of $T_{a=1}, T_{a=a_0}$ *and* $T_{a=a_0^*}$more efficiently than the relevant estimators according to the availability of the prior information onC_y in several areas of research in different fields of Science and Technology.

References

[1] L. Chand, Some ratio type estimators based on two or more auxiliary variables, Ph. D. thesis submitted to Iowa State University, Ames, IOWA (1975).

[2] M.H. Hansen and W.N. Hurwitz, The problem of non-response in sample surveys, J. Amer. Statis. Assoc. **41**, 517-529, (1946).

[3] C. Kadilar and H. Cingi, A New estimator using two auxiliary variables, Applied Mathematics and Computation, **162**, 901-908,(2005).

[4] B. B. Khare and S. Srivastava, Estimation of population mean using auxiliary character in presence of nonresponse, Nat. Acad. Sci. lett., India, **16(3)**, 111-114, (1993).

[5] B. B. Khare and S. Srivastava, Study of conventional and alternative two phase sampling ratio, product and regression estimators in presence of non-response. Proc. Nat. Acad. Sci., India, **65(A)**, II, 195-203, (1995).

[6] B. B. Khare and S. Srivastava, Transformed ratio type estimators for the population mean in the presence of nonresponse. Commun. Statist. - Theory Meth(USA),**26(7)**, 1779-1791, (1997).

[7] B. B. Khare and S. Srivastava, Generalized Estimator for population mean in presence of nonresponse, Int. J. Math. & Stat. Sc., **9(1)**, 75-87, (2000).

[8] B. B. Khare, Use of auxiliary information in sample survey upto 2000- A review, In "Mathematics and Statistics in Engineering, Biotechnology and Science" Edited by D.K. Sen and P.K. Mishra. M/S Centre for Bio-Mathematical Studies, India, 76-87, (2003).

[9] B. B. Khare and S. Kumar, Chain regression type estimators using additional auxiliary variables in two phase sampling in presence of nonresponse, Nat. Acad. Sci. let., India, **33(11-12)**, 369-375, (2010).

[10] B. B. Khare and S. Kumar, Estimation of the population mean using known coefficient of variation of the study character in presence of nonresponse, Commu. Statist. -Theory Meth.,**40(11)**, 2044-2058, (2011).

[11] B. B. Khare, U. Srivastava and K. Kumar, Generalized Chain estimators for the population mean in the presence of non-response, Proc. Nat. Acad. Sci. , India, **81A(III)**, 231-238, (2011).

[12] B. B. Khare, U. Srivastava and K. Kumar, A Generalized chain ratio in regression estimator for population mean using two auxiliary characters in sample surveys. J. Sci. Res.,BHU,**57**, 147-153, (2013).

[13] B. B. Khare, H. U. Rehman, P. S. Jha and Utkarsh, Chain Type Estimators for Population Parameters Using Auxiliary Variables and Addition Auxiliary Variables. Jour. of Nat. Acad. Math. India, **27**, 1-12, (2013).

[14] B. B. Khare and S. Kumar, Generalized Chain ratio in regression estimators for the population mean using two phase sampling in the presence of non-response, Jour. Info. & Opt. Sc., **36(4)**, 317-338, (2015).

[15] B. B. Khare and H. U. Rehman, Improved ratio in regression estimator for the population mean using known coefficient of variation of the study character in the presence of non-response, Int. J. of Tech. Innov. & Res., **14**, 1-7, (2015).

[16] B. B. Khare, K. Kumar and U. Srivastava, Generalized Chain ratio in regression estimators for the population mean using two auxiliary characters in the presence of non-response, Elect. J. of Appl. Stats. Anal., **8(1)**, 67-83, (2015).

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

[18] P. S. R. S. Rao, Ratio estimation with sub sampling the non respondents, Survey Methodology, **12**, 217-230, (1986).

[19] P. S. R. S. Rao, Regression estimators with sub sampling of non respondents, In-Data Quality Control, Theory and Pragmatics, (Eds.) Gunar E. Liepins and V.R.R. Uppuluri, Marcel Dekker, New York, 191-208, (1990).

[20] V.N. Reddy, A study of use of prior knowledge on certain population parameters in estimation, Sankhya, **B(40)**, 29-37, (1978).

[21] J. Sahoo, L. N. Sahoo and S. Mohanty, A regression approach to estimation in two-phase sampling using two auxiliary variables, Current Science, **65(1)**, 73-75,(1993).

[22] H.P. Singh and S. Kumar, A regression approach to the estimation of the finite population mean in the presence of non-response, Austr. & New Zeala. Jour. of Statist., **50(4)**, 395-408, (2008).

[23] H.P. Singh and S. Kumar, A general class of estimators of the population mean in survey sampling using auxiliary information with sub sampling the non-respondents, The Korean J. Appl. Statis., **22(2)**, 387-402, (2009a).

[24] H.P. Singh and S. Kumar, A general procedure of estimating the population mean in the presence of non-response under double sampling using auxiliary information, SORT-Statistics and Operation Research, **33(1)**, 71-83, (2009b).

[25] H.P. Singh and S. Kumar, Improved estimation of the population mean under two phase sampling with sub-sampling of non-respondents, Jour. of Stat. Planning and Inference, **140**, 2536-2550, (2010).

[26] R. R. Sinha, Some problems on the estimation of population parameters using multi auxiliary characters in presence of non-response, Ph. D. Thesis submitted to the Banaras Hindu University, Varanasi, India, (2001).

[27] S. K. Srivastava and S. S. Jhajj, A class of estimators of the population mean using multi auxiliary information, Calcutta Statistical Assoc. Bulletin, **32**, 47-56, (1983).

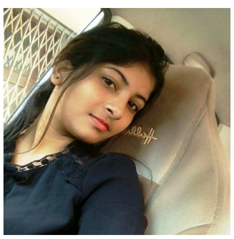
[28] S. Rani Srivastava, S. R. Srivastava, and B. B. Khare, Chain ratio-type estimators for ratio of two population means using auxiliary characters, Commun. Statist.-Theory Meth.,**18(10)**, 3917-3926, (1989).

[29] S. Rani Srivastava,B. B. Khare and S. R. Srivastava, A generalized chain ratio estimator for mean of finite population. Jour. Ind. Soc. Agri. Stat. **42(1)**, 108-117, (1990).

[30] T. P. Tripathi, A. K. Das and B. B. Khare, Use of auxiliary information in sample surveys-A Review, Aligarh Journal of Statist, **14** 79-134, (1994).

Brij Behari Khare (M.Sc., Ph.D., M.N.A.Sc. (India)), Department of Statistics, Institute of Science, BHU, Varanasi has 34 years experience of teaching in UG and PG, guidance of research and also established an active research group during last 34 years in BHU making significant contribution by publishing 117 research papers in several national and international journals of repute and produced 8 (7+1 Sub. June 2017) Ph.D. scholars working in different academic institutions as Assistant Professor, Associate Professor and Professor. Prof. Khare has made significant contribution in the field of sampling theory in statistics, Population studies (Auxology-Growth and Development, Breast feeding and Weaning Practices) and Statistical Inference.

Utkarsh Utkarsh (M.Sc.(Statistics)) joined the Department of Statistics, Institute of Science, BHU, Varanasi in June 2011 as a research scholar(BHU-UGC(RET), JRF) of Ph.D. in Statistics under the supervision of Prof. B. B. Khare in September 2011. During the period of research he has been awarded UGC BSR JRF and worked in the area of sampling theory in statistics till June 2017. During the period of research he has published 8 research papers in national and international journal of repute. He has also attended several workshops and presented research papers in national and international conferences during the period.



Supriya Khare (M.Sc.(Statistics), UGC NET(JRF)) persuing research work in the Department of Statistics, Institute of Science, BHU, Varanasi. Her research intrests in area of Sample survey and Demography. She has published several research papers in the journal of national and international repute and also presented four research papers in national and international conferences.