Journal of Statistics Applications & Probability Letters
An International Journal

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

Improved Chain Ratio-Product Type Estimators Under Double Sampling Scheme

Sunil Kumar * *and Vishwantra Sharma*

Department of Statistics, University of Jammu, Jammu, India

Received: 17 May 2019, Revised: 22 Dec. 2019, Accepted: 18 Feb. 2020
Published online: 1 May 2020

Abstract: In the present paper, an improved chain ratio-product type estimator has been developed for estimating the population mean of the study variable using two auxiliary variables under double sampling scheme. Properties of the estimators have been addressed under two situations: i. when second phase sample is a sub-sample of the first-phase sample, and ii. when second phase sample is drawn independently of the first phase sample. The expressions for the biases and mean squared errors ($MSE's$) have been obtained up to the first order of approximation. Theoretical and empirical study has been conducted to demonstrate the performance of the proposed estimator compared to other estimators in terms of efficiency.

Keywords: Ratio estimator, product estimator, double sampling, chain ratio-product estimator and mean square error.

1 Introduction

In survey sampling, a great variety of techniques are available for using auxiliary information to obtain more efficient estimators. The classical ratio and product estimators are considered when the correlation between the study variable $'Y'$ and the auxiliary variable $'X'$ is positive (high) and negative (high) respectively. They have been discovered by Robson [1] and Murthy [2]. When the relation between study and auxiliary variable is a straight line and passing through the origin, the ratio and product estimators perform equal to the usual regression estimator in terms of efficiencies. However, in many situations, the line does not pass through the origin. Under such circumstances the ratio and product type estimators do not perform equivalent to the regression estimators. Under this situation, many researchers, such as Singh [3,4], Sahai [5], Sahai and Ray [6], Bahl and Tuteja [7], Mohanty and Sahoo [8], Upadhyaya and Singh [9], Singh and Tailor [10], Kadilar and Cingi [11], Singh et al. [12], have paid their attention towards the formulation of modified ratio and product estimators. Chand [13] as well as Sukhatme and Chand [14] proposed a technique of chaining the available information on auxiliary variable with the study variable. Various researchers, like Kiregyera [15,16], Singh et al. [17], Ahmed et al. [18], Isaki [19] have proposed chain type estimators by considering two auxiliary variables. Pal and Singh [20,21] investigated estimation of mean using auxiliary variable and non-response. Some researchers, such as Singh et al. [22] and Pal et al. [23,24] addressed chain ratio type exponential estimators. In this paper, we have proposed a modified chain-type estimator for estimating the population mean $\bar{Y}$ of the study variable using two auxiliary variables X and Z under double sampling technique.

2 The Proposed Chain-Type Estimator

Consider a situation, when population mean $\bar{X}$ of the auxiliary variable X is unknown but the population mean of another auxiliary variable Z is known and has high positive correlation with auxiliary variable X. Let $\bar{x}_1$ and $\bar{z}'$ be the sample means of X and Z respectively based on the preliminary sample of size n_1 drawn with SRSWOR technique to get estimate of $\bar{X}$. Under this set up, we consider a modified chain ratio-product type estimator for estimating the population mean $\bar{Y}$

* Corresponding author e-mail: sunilbhougal06@gmail.com

of the study variable Y as

$$T_s = \bar{y}\left[k\frac{\bar{x}_1^{\delta}}{\bar{x}}\frac{\bar{Z}}{\bar{z}'} + (1-k)\frac{\bar{x}}{\bar{x}_1^{\delta}}\frac{\bar{z}'}{\bar{Z}}\right], \quad (1)$$

where $\bar{x}_1^{\delta} = \frac{n_1\bar{x}_1 - n\bar{x}}{n_1-n}$, k is an suitable choose constant; $\bar{y}$ and $\bar{x}$ are the sample means of y and x, respectively based on sample of size n out of population of size N units; $\bar{X}$ and $\bar{Z}$ are the respective population mean of X and Z. The characteristics of the suggested estimator T_s will be discussed in two different situations:

Situation I: When the second phase sample of size n is a sub-sample of the first phase of size n_1 and

Situation II: When the second phase sample of size n is drawn independently of the first phase of size n_1 see Bose [25].

Situation I:

First, we proceed with situation I, to find the expression of the bias and mean squared error of the proposed estimator T_s, we assume

$e_0 = \frac{\bar{y}-\bar{Y}}{\bar{Y}}, e_1 = \frac{\bar{x}-\bar{X}}{\bar{X}}, e_1' = \frac{\bar{x}_1-\bar{X}}{\bar{X}}, e_2' = \frac{\bar{z}'-\bar{Z}}{\bar{Z}}$

Note that

$E(e_0) = E(e_1) = E(e_1') = E(e_2') = 0$,and

$E(e_0^2) = \lambda C_y^2; E(e_1^2) = \lambda C_x^2; E(e_1'^2) = \lambda_1 C_x^2; E(e_2'^2) = \lambda_1 C_z^2;$

$E(e_0e_1) = \lambda C_{yx}; E(e_0e_1') = \lambda_1 C_{yx}; E(e_0e_2') = \lambda_1 C_{yz}; E(e_1e_1') = \lambda_1 C_x^2; E(e_1e_2) = \lambda_1 C_{xz}; E(e_0'e_1') = \lambda_1 C_{yx};$

$C_y^2 = \frac{S_y^2}{\bar{Y}^2}; C_x^2 = \frac{S_x^2}{\bar{X}^2}; C_z^2 = \frac{S_z^2}{\bar{Z}^2};$

$\lambda = \left(\frac{1}{n} - \frac{1}{N}\right); \lambda_1 = \left(\frac{1}{n_1} - \frac{1}{n}\right);$

$\rho_{xy} = \frac{S_{xy}}{S_xS_y}; \rho_{yz} = \frac{S_{yz}}{S_yS_z}; \rho_{xz} = \frac{S_{xz}}{S_xS_z}; S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i-\bar{X})^2; S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i-\bar{Y})^2; S_z^2 = \frac{1}{N-1}\sum_{i=1}^{N}(z_i-\bar{Z})^2;$

$S_{yx} = \frac{1}{N-1}\sum_{i=1}^{N}(y_i-\bar{Y})(x_i-\bar{X}); S_{yz} = \frac{1}{N-1}\sum_{i=1}^{N}(y_i-\bar{Y})(z_i-\bar{Z}); S_{xz} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i-\bar{X})(z_i-\bar{Z}).$

Now, expressing (1) in terms of e_i's, we have

$T_s = \bar{y}\left[k\frac{n_1\bar{x}_1 - n\bar{x}}{(n_1-n)\bar{x}\bar{z}'}\frac{\bar{Z}}{\bar{z}'} + (1-k)\frac{(n_1-n)\bar{x}\bar{z}'}{n_1\bar{x}_1 - n\bar{x}}\frac{\bar{z}'}{\bar{Z}}\right],$

$$T_s = \bar{Y}(1+e_0)\left[k\left(1+\frac{n_1e_1'-ne_1}{n_1-n}\right)(1+e_1)^{-1}(1+e_2)^{-1} + (1-k)(1+e_1+e_2'+e_1e_2')\left(1+\frac{n_1e_1'-ne_1}{n_1-n}\right)^{-1}\right] \quad (2)$$

Assume that $|e_1| < 1$, and $|e_2'| < 1$, so $(1+e_1)^{-1}, (1+e_2)^{-1}$ and $\left(1+\frac{n_1e_1'-ne_1}{n_1-n}\right)^{-1}$ are expandable.

Now expanding the right hand side of (2), and neglecting the terms of $e's$ greater than second degree, one can obtain

$$T_s - \bar{Y} = \bar{Y}\left[e_0 + e_1 + e_2' + e_1e_2' - \left(\frac{n_1}{n_1-n}\right)(e_1' + e_1e_1' + e_1'e_2') + \left(\frac{n}{n_1-n}\right)(e_1 + e_1^2 + e_1e_2') + \left(\frac{1}{n_1-n}\right)^2 (n_1^2e_1'^2 + n^2e_1'^2 - 2nn_1e_1e_1') + e_0e_1 + e_0e_2' - \left(\frac{n_1}{n_1-n}\right)e_0e_1' + \left(\frac{n}{n_1-n}\right)e_0e_1 + k\left\{e_1^2 + e_2'^2 - 2e_1 - 2e_2' - 2e_0e_1 - 2e_0e_2' + 2\left(\frac{n_1}{n_1-n}\right)(e_1' + e_0e_1') - 2\left(\frac{n}{n_1-n}\right)(e_1 + e_0e_1) - \left(\frac{1}{n_1-n}\right)^2 (n_1^2e_1'^2 + n^2e_1^2 - 2nn_1e_1e_1')\right\}\right] \quad (3)$$

Taking expectation on both side of equation (3), we get the bias of the estimator T_s to the first degree of the approximation as

$$B(T_s) = \bar{Y}\left[\lambda_1C_{xz} - \left(\frac{n_1}{n_1-n}\right)\lambda_1(C_x^2 + C_{xz} + C_{yx}) + \left(\frac{n}{n_1-n}\right)\{\lambda(C_x^2 + C_{yx}) + \lambda_1C_{xz}\} + \lambda(C_{xy}) + \lambda_1C_{yz} + \frac{1}{(n_1-n)^2} (n_1^2\lambda_1 + n^2\lambda - 2nn_1\lambda_1)C_x^2 + k\left\{\lambda(C_x^2 - C_{yx}) + \lambda_1(C_z^2 + C_{yz}) + \left(\frac{n_1}{n_1-n}\right)2\lambda_1C_{yx} - 2\left(\frac{n}{n_1-n}\right)\lambda(C_{yx}) - \left(\frac{1}{n_1-n}\right)^2 (n_1^2\lambda_1 + n^2\lambda - 2nn_1\lambda)C_x^2\right\}\right] \quad (4)$$

Squaring both sides of equation (3) to the first degree of approximation, we get

$$(T_s - \bar{Y})^2 = \bar{Y}^2\left[e_0 + e_1 + e_2' + e_1e_2' - \left(\frac{n_1}{n_1-n}\right)e_1' + \left(\frac{n}{n_1-n}\right)e_1 - 2k\left\{e_1 + e_2' - \left(\frac{n_1}{n_1-n}\right)e_1' + \left(\frac{n}{n_1-n}\right)e_1\right\}\right]^2 \quad (5)$$

Taking expectations on both sides of equation (5) we get the $MSE's$ of T_s to the first degree of approximation as

$$MSE(T_s) = \bar{Y}^2[\lambda A_1 + \lambda_1 B_1 - kC + 4k^2 D] \tag{6}$$

where $A_1 = C_y^2 + (1+A)^2 C_x^2 + (1+A)C_{yx}$;
$B_1 = C_z^2 + (B^2 - 2B - 2AB)C_x^2 + 2(A-B)C_{xz} + 2(C_{xz} + C_{yz} - BC_{yx})$;
$C = 4\lambda\{(1+A)^2 C_x^2 + (1+A)C_{yx}\} - 4\lambda_1\{(2B - B^2 + 2AB)C_x^2 + 2(B - A - 1)C_{xz} - C_{yz} - BC_{yx}\}$;
$D = \lambda(1+A)^2 C_x^2 + \lambda_1\{C_z^2 + (B^2 - 2B - 2AB)C_x^2 + 2(1+A-B)C_{xz}\}; A = \frac{n}{n_1 - n}; B = \frac{n_1}{n_1 - n}$.
The MSE of T_s is minimized when

$$k = \frac{C}{8D} = k_{I(opt)} \tag{7}$$

Substitute the optimum value of k from eq (7) to eq (6) to obtain optimum MSE of T_s as

$$minMSE(T_s) = \bar{Y}^2\left[\lambda A_1 + \lambda_1 B_1 - \frac{C^2}{16D}\right] \tag{8}$$

3 Efficiency Comparison for Situation I

Comparison with chain ratio estimator
For $k = 1$, the estimator T_s in equation (1) tends towards the chain ratio-type estimator T_{s1} as

$$T_{s1} = \bar{y}\frac{\bar{x}_1^\delta}{\bar{x}}\frac{\bar{Z}}{\bar{z}'}$$

The MSE of T_{s1} can be obtained by substituting $k = 1$ in equation (6), as

$$MSE(T_{s1}) = \bar{Y}^2[\lambda A_1 + \lambda_1 B_1 - C + 4D] \tag{9}$$

From equation (6) and equation (9), we have

$$MSE(T_{s1}) - MSE(T_s) \geq 0$$

if either

$$1 < k < \frac{C-4D}{4D} \; or \; \frac{C-4D}{4D} < k < 1 \tag{10}$$

or equivalently,

$$min\left(1, \frac{C-4D}{4D}\right) < \; k \; < max\left(1, \frac{C-4D}{4D}\right)$$

Comparison with chain product estimator
For $k = 0$, the estimator T_s in equation (1) tends towards the chain product-type estimator T_{s2} as

$$T_{s2} = \bar{y}\frac{\bar{x}}{\bar{x}_1^\delta}\frac{\bar{z}'}{\bar{Z}}$$

The MSE of T_{s2} can be obtained by substituting $k = 1$ in equation (6), as

$$MSE(T_{s2}) = \bar{Y}^2[\lambda A_1 + \lambda_1 B_1] \tag{11}$$

From equation (6) and equation (11), we have

$$MSE(T_{s2}) - MSE(T_s) \geq 0$$

if either

$$0 < k < \frac{C}{4D} \; or \; \frac{C}{4D} < k < 0 \tag{12}$$

or equivalently,

$$min(0, \frac{C}{4D}) < k < max(0, \frac{C}{4D})$$

Comparison with sample mean per unit estimator
The variance of sample mean $\bar{y}$ under SRSWOR sampling scheme is given by

$$var(\bar{y}) = \lambda \bar{Y}^2 C_y^2 \tag{13}$$

From equation (6) and equation (13), we have

$$var(\bar{y}) - MSE(T_s) \geq 0,$$

if either

$$\frac{C-\sqrt{C^2-16DE}}{8D} < k < \frac{C+\sqrt{C^2-16DE}}{8D} or \frac{C+\sqrt{C^2-16DE}}{8D} < k < \frac{C-\sqrt{C^2-16DE}}{8D} \tag{14}$$

or equivalently,

$$min\left(\frac{C-\sqrt{C^2-16DE}}{8D}, \frac{C+\sqrt{C^2-16DE}}{8D}\right) < k < max\left(\frac{C-\sqrt{C^2-16DE}}{8D}, \frac{C-\sqrt{C^2-16DE}}{8D}\right)$$

where $E = \lambda_1 B_1 - \lambda(C_y^2 - A_1)$.

Situation II:
Now in situation II, to obtain the bias and mean squared error of T_s, when the second phase sample of size n is drawn independently of the first phase of size n_1, we have
$E(e_0) = E(e_1) = E(e_1') = E(e_2') = 0$,and
$E(e_0^2) = \lambda C_y^2; E(e_1^2) = \lambda C_x^2; E(e_1'^2) = \lambda_1 C_x^2; E(e_2'^2) = \lambda_1 C_z^2;$
$E(e_0 e_1) = \lambda C_{yx}; E(e_1' e_2') = \lambda_1 C_{xz};$
$E(e_0 e_1') = E(e_0 e_2') = E(e_1 e_1') = E(e_1' e_2') = 0$
Taking expectations on both side of equation (3), we get the bias of the estimator T_s to the first degree of approximation as

$$B(T_s) = \bar{Y}\Bigg[\lambda_1 C_{yx} + \left(\frac{n}{n_1-n}\right)\lambda\left\{1+\left(\frac{n}{n_1-n}\right)+\beta_{yx}\right\}C_x^2 + \left(\frac{n_1}{n_1-n}\right)\lambda_1\left\{1+\left(\frac{n_1}{n_1-n}\right)+\beta_{xz}\right\}C_x^2 + k\left\{\lambda(1-2\beta_{yx})C_x^2 + \lambda_1 C_z^2 - \left(\frac{n_1}{n_1-n}\right)^2\lambda_1 C_x^2 - 2\left(\frac{n}{n_1-n}\right)\lambda\left(\beta_{yx}-\left(\frac{n}{n_1-n}\right)\right)C_x^2\right\}\Bigg] \tag{15}$$

where $\beta_{yx} = \frac{C_{yx}}{C_x^2}, \beta_{xz} = \frac{C_{xz}}{C_x^2}$.
Taking squares on both side of equation (3), then taking expectation, we obtain the mean squared error of T_s to the first degree of approximation as

$$MSE(T_s) = \bar{Y}^2[\lambda A_1' + \lambda_1 B_1' - 4kC_1' + 4k^2 D_1'] \tag{16}$$

where $A_1' = C_y^2 + (1+A)^2 C_x^2 + (1+A)C_{yx}$;
$B_1' = C_z^2 + B^2 C_x^2 + 2BC_{xz}$;
$C_1' = \lambda\{(1+A)^2 C_x^2 + (1+A)C_{yx}\} - \lambda_1\{C_z^2 + B^2 C_x^2 - 2BC_{xz}\}$;
$D_1' = \lambda(1+A)^2 C_x^2 + \lambda_1\{C_z^2 + B^2 C_x^2 - 2BC_{xz}\}$;
which will be minimum, when

$$k = \frac{C_1'}{2D_1'} = k_{II(opt)}. \tag{17}$$

Using equation(17) in equation (16), we get the minimum MSE of T_s as

$$minMSE(T_s) = \bar{Y}^2\left[\lambda A_1' + \lambda_1 B_1' - \frac{C_1'^2}{D_1'}\right] \tag{18}$$

4 Efficiency Comparison for Situation II

Comparison with chain ratio estimator
For $k=1$, the estimator T_s in equation (1) tends towards the chain ratio-type estimator T_{s1} as

$$T_{s1}=\bar{y}\frac{\bar{x}_1^{\delta}}{\bar{x}}\frac{\bar{Z}}{\bar{z}'}$$

The MSE of T_{s1} can be obtained by substituting $k=1$in equation (6), as

$$MSE(T_{s1})=\bar{Y}^2[\lambda A_1'+\lambda_1 B_1'-C'+4D'] \tag{19}$$

From equation (16) and equation (19), we have

$$MSE(T_{s1})-MSE(T_s)\geq 0$$

if

$$k=\frac{C'}{2D'} \tag{20}$$

Comparison with chain product estimator
For $k=0$, the estimator T_s in equation (1) tends towards the chain product-type estimator T_{s2} as

$$T_{s2}=\bar{y}\frac{\bar{x}}{\bar{x}_1^{\delta}}\frac{\bar{z}'}{\bar{Z}}$$

The MSE of T_{s2} can be obtained by substituting $k=0$ in equation (16), as

$$MSE(T_{s2})=\bar{Y}^2[\lambda A_1'+\lambda_1 B_1'] \tag{21}$$

From equation (16) and equation (21), we have

$$MSE(T_{s2})-MSE(T_s)\geq 0$$

if either

$$\frac{C_1'}{D_1'}<k<0\; or\; 0<k<\frac{C_1'}{D_1'} \tag{22}$$

or equivalently,

$$min(0,\frac{C_1'}{D_1'})<k<max(0,\frac{C_1'}{D_1'})$$

Comparison with sample mean per unit estimator
The variance of sample mean $\bar{y}$ under SRSWOR sampling scheme is given by

$$var(\bar{y})=\lambda\bar{Y}^2C_y^2 \tag{23}$$

From equation (16) and equation (23), we have

$$var(\bar{y})-MSE(T_s)\geq 0,$$

if either

$$\frac{C_1'-\sqrt{C_1^2-D_1'C^*}}{2D_1'}<k<\frac{C_1'+\sqrt{C_1^2-D_1'C^*}}{2D_1'}\; or\; \frac{C_1'+\sqrt{C_1^2-D_1'C^*}}{2D_1'}<k<\frac{C_1'-\sqrt{C_1^2-D_1'C^*}}{2D_1'} \tag{24}$$

or equivalently,

$$min\left(\frac{C_1'-\sqrt{C_1^2-D_1'C^*}}{2D_1'},\frac{C_1'+\sqrt{C_1^2-D_1'C^*}}{2D_1'}\right)<k<max\left(\frac{C_1'-\sqrt{C_1^2-D_1'C^*}}{2D_1'},\frac{C_1'+\sqrt{C_1^2-D_1'C^*}}{2D_1'}\right)$$

where $C^*=\lambda_1B_1'-\lambda(C_y^2-A_1')$.

5 Empirical Study

To illustrate the performance of the proposed estimator compared to the other estimators, we have considered the four natural population data sets. The source of the populations, the nature of the variables y, x and z and the values of the various parameters are given as follows:

Population I Source: Cochran [26]
Y : Number of placebo children.
X : Number of paralytic polio cases in the placebo group.
Z : Number of paralytic polio cases in the not inoculated group,
$N = 34, n = 10, n_1 = 15, \bar{Y} = 4.92, \bar{X} = 2.59, \bar{Z} = 2.91, \rho_{yx} = 0.7326, \rho_{yz} = 0.6430;$
$\rho_{xz} = 0.6837, C_y^2 = 1.0248, C_x^2 = 1.5175, C_z^2 = 1.1492.$

Population II Source: Srivastava et al. [27]
Y : The measurement of weight of children.
X : Mid-arm circumference of children.
Z : Skull circumference of children.
$N = 82, n = 25, n_1 = 43, \bar{Y} = 5.60Kg, \bar{X} = 11.90cm, \bar{Z} = 39.80, \rho_{yx} = 0.09, \rho_{yz} = 0.12;$
$\rho_{xz} = 0.86, C_y^2 = 0.0107, C_x^2 = 0.0052, C_z^2 = 0.0008.$

Population III Source: Srivastava et al. [27]
Y : The measurement of weight of children.
X : Mid-arm circumference of children.
Z : Skull circumference of children.
$N = 55, n = 18, n_1 = 30, \bar{Y} = 17.08Kg, \bar{X} = 16.92cm, \bar{Z} = 50.44, \rho_{yx} = 0.54, \rho_{yz} = 0.51;$
$\rho_{xz} = -0.08, C_y^2 = 0.0161, C_x^2 = 0.0049, C_z^2 = 0.0007.$

Population IV Source: Khare and Sinha [28]
Y : The measurement of weight of children.
X : Skull circumference of the children.
Z : Chest circumference of the children.
$N = 95, n = 35, n_1 = 70, \bar{Y} = 19.49Kg, \bar{X} = 51.17cm, \bar{Z} = 55.86, \rho_{yx} = 0.328, \rho_{yz} = 0.846;$
$\rho_{xz} = 0.297, C_y = 0.15613, C_x = 0.03006, C_z = 0.0586, C = Rs\,0.75, C_1 = Rs\,2.00, C_2 = Rs\,4.00.$

To reflect the gain in the efficiency of the proposed estimator T_s over the conventional estimators $\bar{y}$, T_{s1} and T_{s2}under situation I and II, the effective ranges along with optimum value of k are presented in Table 1 with respect to the above-mentioned population sets.
To observe the relative performance of different estimators of $\bar{Y}$, we have calculated the percent relative efficiencies of the proposed estimator T_s, chain ratio estimator T_{s1} and product estimator T_{s2} in double sampling with respect to the usual unbiased estimator $\bar{y}$ in situation I and situation II and the findings are represented in Table 2 and graphical representation of our findings are shown in Figure 1.

Table 1: Effective ranges and optimum value of k for the proposed estimator T_s.

Poulation data sets	Ranges for which T_s is better than			Optimum value of k
	T_{s1}	T_{s2}	$\bar{y}$	
Situation I				k_{Iopt}
Population I	(0.18,1.00)	(0.00,1.18)	(0.5,0.68)	0.59
Population II	(0.05,1.00)	(0.00,1.05)	(0.5,0.55)	0.52
Population III	(0.37,1.00)	(0.00,1.37)	(0.5,0.87)	0.68
Population IV	(-0.16,1.00)	(0.00,0.83)	(0.33,0.5)	0.41
Situation II				k_{IIopt}
Population I	(0.00,0.64)	(0.00,1.29)	(0.5,0.79)	0.64
Population II	(0.00,0.54)	(0.00,1.09)	(0.5,0.59)	0.54
Population III	(0.00,1.01)	(0.00,2.03)	(0.5,1.53)	1.01
Population IV	(-4.48,0.00)	(-8.96,0.00)	(-9.46,0.5)	-4.48

Table 1 indicates that the proposed estimator T_s is always better than the considered estimators viz. $\bar{Y}, T_{s1}, T_{s2}$, as the optimum values of k in both situations satisfy the conditions.

Table 2: Percent relative efficiencies of different estimators with respect to $\bar{y}$

Estimator	$\bar{y}$	T_{s1}		T_{s2}		T_s	
Polulation		Situation I	Situation II	Situation I	Situation II	Situation I	Situation II
I	100.00	*	*	*	*	317.30	**487.09**
II	100.00	*	*	*	*	101.11	**101.49**
III	100.00	*	437.80	*	*	172.52	**439.73**
IV	100.00	*	**136.10**	*	*	100.41	*

Table 2 reveals that the proposed estimator T_s performs better than the usual unbiased estimator $\bar{Y}, T_{s1}, T_{s2}$ under situation II when the second phase sample of size n is drawn independently of the first phase sample, for the population I, II and III. However, for population IV, the estimator T_{s1} performs better than the proposed estimator T_s, where the association between the variable Y and X is moderately positive.
Furthermore, for population IV, the proposed estimator T_s performs better in terms of mean squared error with the usual unbiased estimator $\bar{Y}, T_{s1}, T_{s2}$ under situation I i.e. when second phase sample is a sub sample of the first phase sample. But, for situation II, the estimator T_{s1} performs better than the proposed estimator T_s.

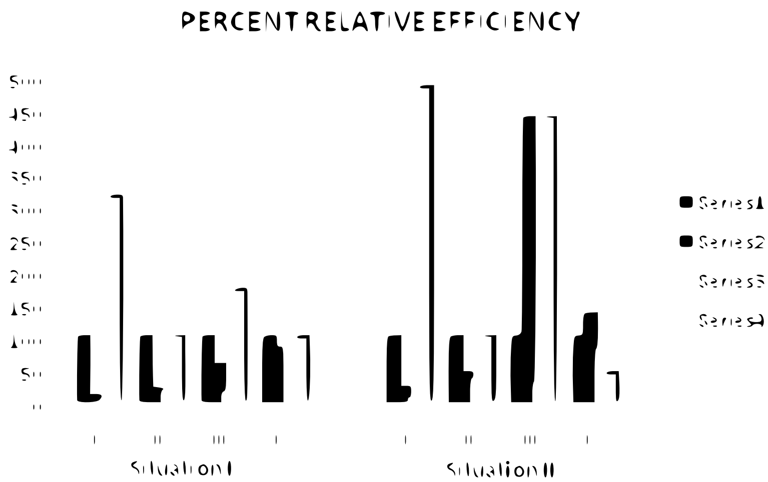

Fig. 1: PRE of the proposed estimator (s) with respect to $\bar{y}$.

6 Cost Function Analysis

We examine the cost function for situation I and situation II separately in case of population IV.
Situation I : Let the cost function be

$$C = C_1 n + C_2 n_1$$

where C_1 is the unit cost associated with second phase sample of size n and C_2 is the unit cost associated with first phase sample of size n_1. Ignoring fpc we write the variance expression of proposed estimator T_s for situation I

$$V = \frac{V_1}{n} + \frac{V_2}{n_1}$$

where $V_1 = A_1 - B_1 - 4k(1-k)(1+A)^2 C_x^2 - 4k(1-k)\{(2B - B^2 + 2AB)C_x^2 - C_z^2 + 2(B - A - 1)C_{xz}\}$
and $V_2 = B_1 + 4k(1-k)(2B - B^2 + 2AB)C_x^2 - C_z^2 + 2(B - A - 1)C_{xz} + 4k(C_{yz} + BC_{yx})$

To obtain the optimum values of n_1 and n with fixed cost, one can obtain the result using lagrangian multiplier as

$$\phi = Var(T_s) + \lambda(C_1 n + C_2 n_1) \tag{25}$$

To minimize the variance for fixed cost, partially differentiate equation (25) with respect to n and n_1, we obtain the optimum values of n and n_1 as

$$n_{(opt)} = \frac{C\sqrt{V_1/C_1}}{\sqrt{V_1 C_1} + \sqrt{V_2/C_2}}$$

$$n_{1(opt)} = \frac{C\sqrt{V_2/C_2}}{\sqrt{V_1 C_1} + \sqrt{V_2/C_2}}$$

The optimum variance T_s for fixed cost is

$$minVar(T_s) = \frac{\left(\sqrt{V_1/C_1} + \sqrt{V_2/C_2}\right)^2}{C}.$$

Situation II: Let the cost function be

$$C = C_1 n + C_2 n_1$$

where C equals total cost of the survey C_1 and C_2 which are the cost per unit. Ignoring fpc, we write the variance expression of proposed estimator T_s for situation II

$$V = \frac{V_1'}{n} + \frac{V_2'}{n_1}$$

where $V_1' = A_1' - B_1' - 4k(1-k)(1+A)^2 C_x^2 + 4k(1-k)\{BC_x^2 + C_z^2 - 2BC_{xz}\}$
and $V_2' = B_1 + 4k(1-k)(2B - B^2 + 2AB)C_x^2 - C_z^2 + 2(B - A - 1)C_{xz} + 4k(C_{yz} + BC_{yx})$
To obtain the optimum values of n_1 and n with fixed cost, one can obtain the result using lagrangian multiplier as

$$\phi = Var(T_s) + \lambda(C_1 n + C_2 n_1) \tag{26}$$

To minimize the variance for fixed cost, partially differentiate equation (26) with respect to n andn_1, we obtain the optimum values of n and n_1 as

$$n_{(opt)} = \frac{C\sqrt{V_1'/C_1}}{\sqrt{V_1' C_1} + \sqrt{V_2'/C_2}}$$

$$n_{1(opt)} = \frac{C\sqrt{V_2'/C_2}}{\sqrt{V_1' C_1} + \sqrt{V_2'/C_2}}$$

The optimum variance T_s for fixed cost is

$$minVar(T_s) = \frac{\left(\sqrt{V_1'/C_1} + \sqrt{V_2'/C_2}\right)^2}{C}.$$

Because of data unavailability, we have shown the performance of the proposed estimators in Situation I and II respectively only for the population IV. The results are shown in Table 3.

Table 3: Variance of estimators when cost is fixed

SITUATION I		SITUATION II		$V(\bar{Y})$
V_1	0.02131	V_1'	0.92751	
V_2	0.02743	V_2'	0.49248	
$Var(T_s)$	0.38553	$Var(T_s)$	10.1976	9.26615

Table 3 exhibits that the proposed estimator in situation I has less variance compared to the usual unbiased estimator. Whereas, situation II, the estimator has more variance with respect to the usual unbiased estimator.

7 Conclusion

We have investigated the improved chain ratio-product estimators under double sampling scheme and evaluated its bias and MSE equations for two different situations i.e. situation I and situation II. Table 1 manifests that our proposed estimator T_s is more efficient than the conventional estimators (unbiased estimator, ratio estimator and product estimator) for both situations under the effective ranges of k.
In Table 2, we have calculated the PRE of the proposed estimator T_s compared to conventional estimators $(\bar{Y}, T_{s1}, T_{s2})$ for both the situations. Table 2 indicates the existence of significant gain in efficiency using proposed estimator T_s compared to the conventional estimators under situation I. Moreover, in situation II proposed estimator T_s performs better than the usual unbiased estimator $(\bar{Y}, T_{s1}, T_{s2})$ for population I, II and III. However, for population IV, the ratio estimator T_{s1} performs better than the proposed estimator T_s and other estimators. Table 2, indicates that our proposed estimator T_s for situation I performs better only in case of population IV. In the other three populations i.e, in populations I, II and III, the proposed estimator T_s performs better.
Figure 1 shows the performance of PRE of different estimators with respect to the usual unbiased estimator $\bar{y}$ in situation I and situation II.

Acknowledgement

We are highly grateful to the reviewers for their thorough review, which improved the quality of the paper.

References

[1] D.S. Robson, Application of multivariate polykays to the theory of unbiased ratio-type estimation, *Journal of American Statistical Association*, **52**, 411-422 (1957).
[2] M.N. Murthy, Product method of estimation, *Sankhya*, **26(A)**, 294-307, (1964).
[3] M.P. Singh, On the estimation of ratio and product of population parameters, *Sankhya*, **B, 27**, 321-328 (1965).
[4] M.P. Singh, Ratio cum product method of estimation, *Sankhya*, **12**, 34-43 (1967).
[5] A. Sahai, An efficient variant of the product and ratio estimators, *Statistica Neelandica*, **33**, 27-35 (1979).
[6] A.Sahai and S.K. Ray, An efficient estimator using auxiliary information, *Metrika*, 27(1), 271-275 (1980).
[7] S. Bahl and R.K. Tuteja, Ratio and product type exponential estimator, *Information and Optimization Sciences*, **12**, 159-163 (1991).
[8] S. Mohanty and J. Sahoo, A note on improving the ratio method of estimation through linear transformation using certain known population parameters, *Sankhya Series*, B, **57(1)**, 93-102 (1995).
[9] L.N. Upadhyaya and H.P. Singh, Use of transformed auxiliary variable in estimating the finite population mean, *Biometrical journal*, **41(5)**, 627-636 (1999).
[10] H.P. Singh and R. Tailor, Use of known correlation coefficient in estimating the finite population mean, *Statistics in Transition*, **6(4)**, 555-560 (2003).
[11] C. Kadilar and H. Cingi, Ratio estimators for the population variance in simple and stratified random sampling, *Applied Mathematics and Computations*, **173**, 1047-1059 (2006).
[12] H.P. Singh, S. Singh and J.M. Kim, Efficient use of auxiliary variables in estimating finite population variance in two-phase sampling, *Communications of the Korean Statistical Society*, **17(3)**, 1-21 (2010).
[13] L. Chand, *Some ratio type estimator based on two or more auxiliary variables*, Ph.D. Dissertation. Lowa State University. Ames, Lowa (unpublished) (1975).
[14] B.V. Sukhatme and L. Chand, *Multivariate ratio-type estimators*, Proceedings of American Statistical Association, Social Statistics Section, 927-931 (1977).
[15] B. Kiregyera, A chain ratio-type estimator in finite population in double sampling using two auxiliary variables, *Metrika*, **27**, 217-223 (1980).
[16] B. Kiregyera, Regression-type estimators using two auxiliary variables and the model of double sampling from finite populations, *Metrika*, **31(3-4)**, 215-226 (1984).
[17] S. Singh, H.P. Singh and L.N. Upadhyaya, Chain ratio and regression type estimators for median estimation in sample survey, *Statistical Papers*, **48**, 23-46 (2006).
[18] M.S. Ahmed, M.S. Raman and M.I. Hossain, Some competitive estimators of finite population variance using multivariate auxiliary information, *Information and Management Sciences*, **11(1)**, 49-54 (2000).
[19] C.T. Isaki, Variance estimation using auxiliary information, *Journal of the American statistical association*, **78**, 117-123 (1983).
[20] S.K. Pal and H.P. Singh, A class of ratio-cum-ratio-type exponential estimators for population mean with sub-sampling the non-respondents, *Jordan Journal of Mathematics & Statistics*, **10(1)**, 73-94 (2017).
[21] S.K. Pal and H.P. Singh, Estimation of mean with two-parameter ratio-product-ratio estimator in double sampling using auxiliary information under non-response, *Journal of Modern Applied Statistical Methods*, **17(2)**, 2-32 (2018).

[22] H.P. Singh, S.K. Pal and A. Yadav, A study on the chain ratio-ratio type exponential estimators for finite population variance, *Communications in statistics Theory and Methods*, **47(6)**, 1442-1458, (2018).
[23] S.K. Pal, H.P. Singh and S. Srivastava, Some new development of chain type estimators for estimating population mean in cluster sampling, *Bulletin of Pure and Applied Sciences-Math* & Stat, **38(1)**, 1-16, (2019).
[24] S.K. Pal, H.P. Singh and F. Bagwala, A study based on chain ratio-type and chain ratio-ratio-type exponential estimator for estimating the population mean in systematic sampling, *International Journal of Agriculture Statistics Sciences*, (Accepted)(2019).
[25] C. BOSE, Note on the sampling error in the method of double sampling, *Sankhya*, **6**, 330, (1943).
[26] W. G. Cochran, *Sampling Techniques*, Third Edition, New York: John Wiley & Sons (1977).
[27] R.S. Srivastava, S.P. Srivastava and B.B. Khare, Chain ratio type estimator for ratio of two population means using auxiliary characters, *Communications in Statistics- Theory and Methods*, **18(9)**, 3917-3926, (1989).
[28] B.B. Khare and R.R. Sinha, *Estimation of the ratio of the two population means using multi- auxiliary characters in presence of non-response*, Proceedings of Statistical techniques in life testing, reliability, sampling theory and quality control, edited by B. N. Pandey Narosa publishing house, New Delhi, 163-171, (2007).

Sunil Kumar is Assistant Professor of Statistics at Department of Statistics, Unviersity of Jammu, Jammu and Kashmir (India). Visiting Scientist/Professor at Indian Statistical Institute, Kolkata (India). He received the Ph. D. degree in “Statistics” at Vikram University (India). He is referee and editorial member of several International journals in the frame of Statistics and Management studies. His main research interests are: sample survey, non-response, estimation theory, consumer behavior, latent class analysis and applications.

Vishwantra Sharma is Research Scholar of Statistics pursuing Ph.D at Department of Statistics, Unviersity of Jammu, Jammu and Kashmir (India). She has completed her M.Phill in ”Statistics” from University of Jammu (India). Her main research interests are: sample survey, non-response, estimation theory.