JOURNAL OF STATISTICS, OPTIMIZATION AND DATA SCIENCE

VOL NO.3 , ISSUE 1 (JUNE 2025), ,PAGE NO.25-40 ,ISSN:2583-9462

A novel approach to the almost filtration of bias precipitates

Chanda Rani[a], Sunil Kumar[b] and Med Ram Verma[c].

[a,b] *Department of Statistics, University of Jammu, Jammu, India*
[c] *Head Design of Experiments, ICAR-Indian Agricultural Statistics Research Institute, New Delhi, India*

ARTICLE HISTORY
Compiled 20 June 2025

.

ABSTRACT
Researchers often seek to improve estimator efficiency beyond the standard mean per unit approach, which considers only the study variable's observations. One common strategy involves incorporating data from a related auxiliary variable when drawing samples from a finite population to obtain a more accurate estimate of the study variable's population mean. This research study introduce an innovative approach using a funnel and filter paper to isolate bias precipitates that arise in the ratio and product methods of population mean estimation under two-phase sampling. Following filtration, we examine the resulting biases and mean squared error up to the first order of approximation. Furthermore, a simulation study is performed to illustrate and validate the effectiveness of our approach.

KEYWORDS
Filtration, Biases, Two-Phase Sampling, Linear Constraints.

Email: [a]chandaheer65@gmail.com, [b]sunilbhougal06@gmail.com and [c]mrverma19@yahoo.co.in

1. Introduction

In sample surveys, the main objective is often to estimate the population total or mean of a study variable (y). Cochran (1940) and Murthy (1964) introduced the standard ratio and product estimators for estimating the population total or mean of study variable. However, these estimators are biased and less efficient compared to the standard linear regression estimators, proposed by Hansen et al. (1953). Hartley and Ross (1954) managed to eliminate the bias from the ratio estimator, but their modified estimator remains less efficient than the linear regression approach. Singh (1989) introduced a class of almost unbiased estimators for the product of two population means. Similarly, Sahoo and Singh (1989) along with Sahoo (1983), suggested class of almost unbiased estimators for estimating population ratio and product. Despite these advancements, the linear regression estimators generally maintains greater efficiency than the almost unbiased estimators mentioned. Parsad [6], along with Prasad and Singh (1990) enhanced the ratio-type estimators from the mean square error perspective, but did not address the bias in the proposed estimators.

Srivatsava (1967, 1971), Chakrabarty (1968), Vos (1980), Walsh (1988) and Reddy (1973) demonstrated that the efficiency of Ratio and Product estimators can be enhanced to that of traditional linear regression estimators by making some optimum selections. Motivated by Singh and Singh (1993), we suggest an innovative methodology for separating bias precipitates from ratio and product type estimators using a funnel connected to filter paper. The apparatus has components made up of a linear variety and linear restrictions. It will be shown that the chemicals (statistical restrictions) used for bias filtration depend solely on the well-known optimal choices.

2. Notations and Expectations

Assume that $\omega = \omega_1, \omega_2, ..., \omega_N$ be a finite population of size N. In the first stage of two-phase sampling,, a random sample of size n' is drawn from the population ω then in the next stage, a sub sample of size n is further drawn from this initial sample n'. For the i^{th} unit $(i = 1, 2, .., N)$ of the population, let the values of the study variables Y_1, Y_2 and the auxiliary variable X be represented by Y_{1i}, Y_{2i}, and X_i; consequently, for the i^{th} unit in the sample $(i = 1, 2, .., n)$, let the values be represented by y_{1i}, y_{2i}, and x_i, respectively, based on the sample observations, we have

$\bar{y}_1 = \frac{1}{n}\sum_{i=1}^{n} y_{1i}$, $\bar{y}_2 = \frac{1}{n}\sum_{i=1}^{n} y_{2i}$, $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, $s_{y_1}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_{1i} - \bar{y}_1)^2$,

$s_{y_2}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_{2i} - \bar{y}_2)^2$, $s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})$, $s_{xy_1} = \frac{1}{(n-1)}\sum_{i=1}^{n}(y_{1i} - \bar{y}_1)(x_i - \bar{x})$

and $s_{xy_2} = \frac{1}{(n-1)}\sum_{i=1}^{n}(y_{2i} - \bar{y}_1)(x_i - \bar{x})$.

For the population, we have the analogue quantities

$\bar{Y}_1 = \frac{1}{n}\sum_{i=1}^{n} Y_{1i}$, $\bar{Y}_2 = \frac{1}{n}\sum_{i=1}^{n} Y_{2i}$, $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, $S_{y_1}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_{1i} - \bar{Y}_1)^2$,

$S_{y_2}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(Y_{2i} - \bar{Y}_2)^2$, $S_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})$, $S_{xy_1} = \frac{1}{(n-1)}\sum_{i=1}^{n}(Y_{1i} - \bar{Y}_1)(X_i - \bar{X})$,

$S_{xy_2} = \frac{1}{(n-1)} \sum_{i=1}^{n} (Y_{2i} - \bar{Y}_2)(X_i - \bar{X})$, $C_{y_1}^2 = \frac{S_{y_1}^2}{\bar{Y}_1^2}$, $C_{y_2}^2 = \frac{S_{y_2}^2}{\bar{Y}_2^2}$, and $C_x^2 = \frac{S_x^2}{\bar{X}^2}$.

Use the following transformation as an indicator for estimating the Bias and Mean Squared Error (MSE) of the proposed estimator up to the first-order approximation:

$\epsilon_0 = \frac{(\bar{y}_1 - \bar{Y}_1)}{\bar{Y}_1}$, $\epsilon_1 = \frac{(\bar{y}_2 - \bar{Y}_2)}{\bar{Y}_2}$, $\epsilon_2 = \frac{(\bar{x} - \bar{X})}{\bar{X}}$, $\epsilon_3 = \frac{(\bar{x}' - \bar{X})}{\bar{X}}$ such that

$E(\epsilon_0) = E(\epsilon_1) = E(\epsilon_2) = E(\epsilon_3) = 0,$

$E(\epsilon_0^2) = \lambda C_{y_1}^2$, $E(\epsilon_1^2) = \lambda C_{y_2}^2$, $E(\epsilon_2^2) = \lambda C_x^2$,

$E(\epsilon_3^2) = \lambda' C_x^2$, $E(\epsilon_0 \epsilon_2) = \lambda \rho_{xy_1} C_{y_1} C_x$, $E(\epsilon_0 \epsilon_3) = \lambda' \rho_{xy_1} C_{y_1} C_x$,

$E(\epsilon_2 \epsilon_3) = \lambda' C_x^2 = E(\epsilon_3^2)$, $E(\epsilon_0 \epsilon_1) = \lambda \rho_{y_1 y_2} C_{y_1} C_{y_2}$, $E(\epsilon_1 \epsilon_2) = \lambda \rho_{xy_2} C_{y_2} C_x$,

$E(\epsilon_1 \epsilon_3) = \lambda' \rho_{xy_2} C_{y_2} C_x$.

where $\lambda = \frac{(1-f)}{n}$, $\lambda' = \left(\frac{1}{n'} - \frac{1}{N}\right)$ and $f = \frac{n}{N}$ denotes the finite population correction factor.

3. Singh and Singh (1993) Estimators

In order to remove the bias precipitates from the ratio and product type estimators, Singh and Singh (1993) suggested using a funnel in conjunction with filter paper. The apparatus has two components: linear constraints and a linear variety of estimators.

Let us define

$$\hat{R}_j = \bar{y}_1 \left(\frac{\bar{X}}{\bar{x}} \right)^j, \quad such\ that\ \hat{R}_j \in G\ \ for\ \ j = 1, 2, 3, \tag{1}$$

where G is defined as the collection of all possible ratio estimators used to estimating the population mean $\bar{Y}_1$.

By definition, the set G will be the linear variety if

$$\hat{R}_s = \sum_{j=1}^{3} g_j \hat{R}_j \in G, \tag{2}$$

for

$$\sum_{j=1}^{3} g_j = 1\ \ and\ \ g_j \in R, \tag{3}$$

where g_j $(for\ j = 1, 2, 3)$ signifies the quantity of chemicals utilized for separating bias precipitates and R represents the set of real numbers. Based on equation (3), equation

(2) can be expressed in terms of ϵ_0 and ϵ_1 as follows

$$\hat{R}_s = \bar{Y}_1\left[1 + \epsilon_0 - K\epsilon_2 + O(n^{-2})\right] \tag{4}$$

where

$$K = (g_1 + 2g_2 + 3g_3). \tag{5}$$

Thus, the MSE, to the first-order approximation, can be represented as follows.

$$MSE(\hat{R}_s) = \lambda\bar{Y}_1^2\left[C_{y_1}^2 + KC_x^2 - 2K\rho_{xy_1}C_{y_1}C_x\right] \tag{6}$$

The MSE upto term of order $\left(\frac{1}{n}\right)$ is minimized for

$$K_{opt} = \rho\frac{C_{y_1}}{C_x} \tag{7}$$

Therefore, the minimum MSE upto first order of approximation is given by

$$MSE_{min}(\hat{R}_s) = \lambda S_{y_1}^2(1 - \rho_{xy_1}^2) \tag{8}$$

From (3), (5), and (7), Singh and Singh (1993) made a funnel consisting of two equations given as

$$\sum_{j=1}^{3} g_j = 1, \tag{9}$$

and

$$\sum_{j=1}^{3} jg_j = \rho_{xy_1}\frac{C_{y_1}}{C_x} \tag{10}$$

They need to figure out three unknowns in the two equations (9, 10) mentioned above. As a result, finding different values for the amount of chemicals g_1, g_2 and g_3 to be utilized for filtration is not feasible. They asserted that their funnel is handicapped and requires filter paper to function properly. To obtain unique chemical values, they recommended applying a filter paper with a funnel and implement linear constraints:

$$\sum_{j=1}^{3} g_j Bias(\hat{R}_j) = 0 \tag{11}$$

where $Bias(\hat{R}_j)$ represents the bias in $\hat{R}_j, j = 1, 2, 3,$ of the population mean. The aforementioned three equations can be expressed in the following form.

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ Bias(\hat{R}_1) & Bias(\hat{R}_2) & Bias(\hat{R}_3) \end{bmatrix} \begin{bmatrix} g_1 \\ g_2 \\ g_3 \end{bmatrix} = \begin{bmatrix} 1 \\ K \\ 0 \end{bmatrix} \tag{12}$$

Solving the system of equation (12) yields the values of $g_j, j = 1, 2, 3$ which allowing the separation of bias precipitates from the linear class of estimators, provided that $|A| \neq 0$, i.e. if

$$Bias(\hat{R}_2) \neq \frac{Bias(\hat{R}_1) + Bias(\hat{R}_3)}{2} \tag{13}$$

Thus, we have the following theorem

Theorem:- *The values of g_j, for $j = 1, 2, 3$ that filter the bias up to terms of first order are given as*

$g_1 = 3 - 3K + K^2$, $g_2 = -3 + 5K - 2K^2$ *and* $g_3 = 1 - 2K + K^2$

Proof: To the first order of approximation

$$Bias(\hat{R}_1) = \lambda\bar{Y}_1\big[C_x^2 - \rho_{xy_1}C_{y_1}C_x\big], \tag{14}$$

$$Bias(\hat{R}_2) = \lambda\bar{Y}_1\big[3C_x^2 - 2\rho_{xy_1}C_{y_1}C_x\big], \tag{15}$$

$$Bias(\hat{R}_3) = \lambda\bar{Y}_1\big[6C_x^2 - 3\rho_{xy_1}C_{y_1}C_x\big]. \tag{16}$$

On substituting these values of biases in (12), we develop the values of g_j given in the theorem.

4. Proposed Estimator

Motivated by Singh and Singh (1993), we propose a novel technique for removing bias precipitates from ratio and product type estimators that employs a funnel attached to filter paper. The apparatus's components consist of a linear variety and linear restrictions. Let

$$\hat{T}_j = \hat{T}_{(\alpha)}\left(\frac{\bar{x}}{\bar{x}'}\right)^j, \quad j = 1, 2, 3, \tag{17}$$

where $\hat{T}_{(\alpha)} = \left(\frac{\bar{y}_1}{\bar{y}_2^{\alpha}}\right)$; $\bar{y}_2 \neq 0$.

For different values of α, we have three situations.

- **Situation 1:** When $\alpha = 0$, equation (17) can be written as

$$\hat{T}_{1j} = \bar{y}_1\left(\frac{\bar{x}}{\bar{x}'}\right)^j, \quad \textit{such that } \hat{T}_{1j} \in G \textit{ for } j = 1, 2, 3, \tag{18}$$

By definition, the set G will be the linear variety if

$$\hat{T}_{s_1} = \sum_{j=1}^{3} g_j\hat{T}_{1j}, \quad \hat{T}_{s_1} \in G, \tag{19}$$

for

$$\sum_{j=1}^{3} g_j = 1 \;\; and \;\; g_j \in R, \tag{20}$$

where g_j $(j = 1, 2, 3)$ signifies the quantity of chemicals utilized for separating bias precipitates and R represents the set of real numbers. Based on equation (20), equation (19) can be expressed in terms of ϵ_0 and ϵ_1 as follows

$$\hat{T}_{s_1} = \bar{Y}_1\left[1 + \epsilon_0 + K_1(\epsilon_2 - \epsilon_3) + O(n^{-2})\right] \tag{21}$$

where

$$K_1 = (g_1 + 2g_2 + 3g_3) \tag{22}$$

By solving above equation (21), we get the MSE of $\hat{T}_s$ upto first order of approximation

$$MSE(\hat{T}_{s_1}) = \bar{Y}_1^2\left[\lambda C_{y_1}^2 + K_1^2 C_x^2(\lambda - \lambda') + 2K_1\rho_{xy_1}C_{y_1}C_x\right] \tag{23}$$

Differentiating w.r.t to K_1, we get the optimum value

$$K_{1(opt)} = -\rho_{xy_1}\frac{C_{y_1}}{C_x} \tag{24}$$

Therefore, the minimum MSE upto first order of approximation is given by

$$MSE_{min}(\hat{T}_{s_1}) = S_{y_1}^2\left[\lambda - \rho_{xy_1}^2(\lambda - \lambda')\right] \tag{25}$$

By considering the conditions on (20), (22) and (24), we generate a funnel consisting of two equations that are described as

$$\sum_{j=1}^{3} g_j = 1 \tag{26}$$

and

$$\sum_{j=1}^{3} j g_j = -\rho_{xy_1}\frac{C_{y_1}}{C_x} \tag{27}$$

Thus, there are three unknowns in the above two equations that need to be found. Therefore, it is impossible to determine the precise amounts of chemicals (g_1, g_2, g_3) that should be employed for filtration. Because it is stated that their funnel requires a filter paper to function properly and is disabled. They recommended utilizing a filter paper with a funnel and applying linear restrictions as follows in order to obtain unique results for these chemicals.

$$\sum_{j=1}^{3} g_j Bias(\hat{T}_{1j}) = 0, \tag{28}$$

where $Bias(\hat{T}_{1j})$ represents the bias in the $\hat{T}_{1j}, j = 1, 2, 3,$ of population mean. The aforementioned three equations can be expressed in the following form.

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ Bias(\hat{T}_{(11)}) & Bias(\hat{T}_{(12)}) & Bias(\hat{T}_{(13)}) \end{bmatrix} \begin{bmatrix} g_1 \\ g_2 \\ g_3 \end{bmatrix} = \begin{bmatrix} 1 \\ K_1 \\ 0 \end{bmatrix} \tag{29}$$

Solving the system of equation (29) yields the values of $g_j, j = 1, 2, 3$ which allowing the separation of bias precipitates from the linear class of estimators, provided that $|A| \neq 0$, i.e. if

$$Bias(\hat{T}_{(12)}) \neq \frac{Bias(\hat{T}_{(11)}) + Bias(\hat{T}_{(13)})}{2} \tag{30}$$

Thus, we have the following theorem.

Theorem 1:- *The values of $g_j, for\ j = 1, 2, 3$ that filter the bias up to terms of first order are given as*

$g_1 = \frac{1}{\lambda'}\left[3(\lambda - \lambda') - (3\lambda - 2\lambda')K_1 + (\lambda - \lambda')K_1^2\right],$

$g_2 = \frac{1}{\lambda'}\left[-3(2\lambda - \lambda') + 3(2\lambda - \lambda')K_1 - 2(\lambda - \lambda')K_1^2\right]$ *and*

$g_3 = \frac{1}{\lambda'}\left[(3\lambda - \lambda') - (3\lambda - \lambda')K_1 + (\lambda - \lambda')K_1^2\right].$

Proof: To the first order of approximation

$$Bias(\hat{T}_{(11)}) = \bar{Y}_1(\lambda - \lambda')\ell, \tag{31}$$

$$Bias(\hat{T}_{(12)}) = \bar{Y}_1\left[\lambda\{3C_x^2 + 2\ell\} - \lambda'\{C_x^2 + 2\ell\}\right], \tag{32}$$

$$Bias(\hat{T}_{(13)}) = \bar{Y}_1\left[3\lambda\{2C_x^2 + \ell\} - 3\lambda'\{C_x^2 + \ell\}\right], \tag{33}$$

where $\ell = \rho_{xy_1} C_{y_1} C_x$.

On substituting these values of biases in (29), we develop the values of g_j as stated in the theorem.

- **Situation 2:** When $\alpha = 1$, equation (17) can be written as

$$\hat{T}_{2j} = \frac{\bar{y}_1}{\bar{y}_2}\left(\frac{\bar{x}}{\bar{x}'}\right)^j, \quad such\ that\ \hat{T}_{2j} \in G\ \ for\ \ j = 1, 2, 3, \tag{34}$$

where G represents the collection of all possible ratio estimators for estimating the population mean of both the study variables $\bar{Y}_1$ and $\bar{Y}_2$.

By definition, the set G will be the linear variety if

$$\hat{T}_{s_2} = \sum_{j=1}^{3} g_j \hat{T}_{2j} \in G, \tag{35}$$

for

$$\sum_{j=1}^{3} g_j = 1 \;\; and \;\; g_j \in R, \tag{36}$$

where g_j $(j = 1, 2, 3)$ signifies the quantity of chemicals utilized for separating bias precipitates and R represents the set of real numbers. Based on equation (36), equation (35) can be expressed in terms of ϵ_0 and ϵ_1 as follows

$$\hat{T}_{s_2} = \frac{\bar{Y}_1}{\bar{Y}_2}\left[1 + \epsilon_0 - \epsilon_1 + K_2(\epsilon_2 - \epsilon_3) + O(n^{-2})\right], \tag{37}$$

where

$$K_2 = (g_1 + 2g_2 + 3g_3) \tag{38}$$

By solving above equation (37), we get the MSE of $\hat{T}_{s_2}$ upto first order of approximation

$$MSE(\hat{T}_{s_2}) = \left(\frac{\bar{Y}_1}{\bar{Y}_2}\right)^2 \Big[\lambda\{C_{y_1}^2 + C_{y_2}^2 - 2\rho_{y_1y_2}C_{y_1}C_{y_2}\} + K_2^2C_x^2(\lambda - \lambda') + 2K_2(\lambda - \lambda')\{\rho_{xy_1}C_{y_1}C_x - \rho_{xy_1}C_{y_2}C_x\}\Big] \tag{39}$$

Differentiating w.r.t to K_2, we get the optimum value

$$K_{2(opt)} = \left[\rho_{xy_2}\frac{C_{y_2}}{C_x} - \rho_{xy_1}\frac{C_{y_1}}{C_x}\right] \tag{40}$$

Therefore, the minimum MSE upto first order of approximation is given by

$$MSE_{min}(\hat{T}_{s_2}) = \left(\frac{\bar{Y}_1}{\bar{Y}_2}\right)^2 \left[\lambda\{C_{y_1}^2 + C_{y_2}^2 - 2\rho_{y_1y_2}C_{y_1}C_{y_2}\} - (\lambda - \lambda')\{\rho_{xy_2}C_{y_2} - \rho_{xy_1}C_{y_1}\}^2\right] \tag{41}$$

By considering the conditions on (36), (38) and (40), we develop a funnel consisting of two equations that describes as

$$\sum_{j=1}^{3} g_j = 1, \tag{42}$$

and

$$\sum_{j=1}^{3} j g_j = \left[\rho_{xy_2}\frac{C_{y_2}}{C_x} - \rho_{xy_1}\frac{C_{y_1}}{C_x}\right]. \tag{43}$$

Thus, there are three unknowns in the above two equations that need to be found. Therefore, it is impossible to determine the precise amounts of chemicals (g_1, g_2, g_3) that should be employed for filtration. Because it is stated that their funnel requires a filter paper to function properly and is disabled. They recommended utilizing a filter paper with a funnel and applying linear restrictions as follows in order to obtain unique results for these chemicals.

$$\sum_{j=1}^{3} g_j Bias(\hat{T}_{2j}) = 0, \tag{44}$$

where $Bias(\hat{T}_{2j})$ represents the bias in the $\hat{T}_{2j}, j = 1, 2, 3$, of population mean. The aforementioned three equations can be expressed in the following form.

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ Bias(\hat{T}_{(21)}) & Bias(\hat{T}_{(22)}) & Bias(\hat{T}_{(23)}) \end{bmatrix} \begin{bmatrix} g_1 \\ g_2 \\ g_3 \end{bmatrix} = \begin{bmatrix} 1 \\ K_2 \\ 0 \end{bmatrix} \tag{45}$$

Solving the system of equation (45) yields the values of $g_j, j = 1, 2, 3$ which allowing the separation of bias precipitates from the linear class of estimators, provided that $|A| \neq 0$, i.e. if

$$Bias(\hat{T}_{(22)}) \neq \frac{Bias(\hat{T}_{(21)}) + Bias(\hat{T}_{(23)})}{2} \tag{46}$$

Thus, we have the following theorem.

Theorem 2:- *The values of g_j, for $j = 1, 2, 3$ which filter the bias up to terms of first order are given as*

$g_1 = \frac{1}{\lambda'}\left[3(\lambda - \lambda') - (3\lambda - 2\lambda')K_2 + (\lambda - \lambda')\left(K_2^2 - \frac{\lambda\tau_1}{(\lambda-\lambda')}\right)\right],$

$g_2 = \frac{1}{\lambda'}\left[-3(2\lambda - \lambda') + 3(2\lambda - \lambda')K_2 - 2(\lambda - \lambda')\left(K_2^2 - \frac{\lambda\tau_1}{(\lambda-\lambda')}\right)\right]$ *and*

$g_3 = \frac{1}{\lambda'}\left[(3\lambda - \lambda') - (3\lambda - \lambda')K_2 + (\lambda - \lambda')\left(K_2^2 - \frac{\lambda\tau_1}{(\lambda-\lambda')}\right)\right],$

where $\tau_1 = \frac{\left(C_{y_2}^2 - \rho_{y_1y_2}C_{y_1}C_{y_2}\right)}{C_x^2}$.

Proof: To the first order of approximation

$$Bias(\hat{T}_{21}) = \left(\frac{\bar{Y}_1}{\bar{Y}_2}\right)\left[\lambda\left\{\kappa_1 + \kappa_2\right\} - \lambda'\kappa_2\right], \tag{47}$$

$$Bias(\hat{T}_{22}) = \left(\frac{\bar{Y}_1}{\bar{Y}_2}\right)\left[\lambda\{\kappa_1 + \big(3C_x^2 + 2\kappa_2\big)\} - \lambda'\{C_x^2 + 2\kappa_2\}\right], \tag{48}$$

$$Bias(\hat{T}_{23}) = \left(\frac{\bar{Y}_1}{\bar{Y}_2}\right)\left[\lambda\{\kappa_1 + 3\big(2C_x^2 + \kappa_2\big)\} - \lambda'\{3C_x^2 + 2\kappa_2\}\right], \tag{49}$$

where $\kappa_1 = \big(C_{y_2}^2 - \rho_{y_1y_2}C_{y_1}C_{y_2}\big)$ and $\kappa_2 = \big(\rho_{xy_1}C_{y_1}C_x - \rho_{xy_2}C_{y_2}C_x\big)$.
On substituting these values of biases in (45), We develop the values of g_j as stated in the theorem.

- **Situation 3:** When $\alpha = -1$, equation (17) can be written as

$$\hat{T}_{3j} = \bar{y}_1\bar{y}_2\left(\frac{\bar{x}}{\bar{x}'}\right)^j, \; such \; that \; \hat{T}_{3j} \in G \; for \; j = 1, 2, 3, \tag{50}$$

where G represents the collection of all possible product type estimators for estimating the population mean of both the study variables $\bar{Y}_1$ and $\bar{Y}_2$.

By definition, the set G will be the linear variety if

$$\hat{T}_{s_3} = \sum_{j=1}^{3} g_j\hat{T}_{3j} \in G, \tag{51}$$

for

$$\sum_{j=1}^{3} g_j = 1 \; and \; g_j \in R, \tag{52}$$

where g_j $(j = 1, 2, 3)$ signifies the quantity of chemicals utilized for separating bias precipitates and R represents the set of real numbers. Based on equation (52), equation (51) can be expressed in terms of ϵ_0 and ϵ_1 as follows

$$\hat{T}_{s_3} = \bar{y}_1\bar{y}_2\big[1 + \epsilon_0 - \epsilon_1 + K_3(\epsilon_2 - \epsilon_3) + O(n^{-2})\big] \tag{53}$$

where

$$K_3 = (g_1 + 2g_2 + 3g_3) \tag{54}$$

By solving above equation (53), we get the MSE of $\hat{T}_s$ upto first order of approximation

$$MSE(\hat{T}_{s_3}) = \big(\bar{y}_1\bar{y}_2\big)^2\Big[\lambda\{C_{y_1}^2 + C_{y_2}^2 + 2\rho_{y_1y_2}C_{y_1}C_{y_2}\} + {K_3}^2C_x^2(\lambda - \lambda') + 2K_3(\lambda - \lambda')\{\rho_{xy_2}C_{y_2}C_x + \rho_{xy_1}C_{y_1}C_x\}\Big] \tag{55}$$

Differentiating w.r.t to K_3, we get the optimum value

$$K_{3(opt)} = -\left[\rho_{xy_2}\frac{C_{y_2}}{C_x} + \rho_{xy_1}\frac{C_{y_1}}{C_x}\right] \tag{56}$$

Therefore, the minimum MSE upto first order of approximation is given by

$$MSE_{min}(\hat{T}_{s_3}) = \left(\bar{y}_1\bar{y}_2\right)^2\left[\lambda\left\{C_{y_1}^2+C_{y_2}^2+2\rho_{y_1y_2}C_{y_1}C_{y_2}\right\}-(\lambda-\lambda')\left\{\rho_{xy_2}C_{y_2}+\rho_{xy_1}C_{y_1}\right\}^2\right] \tag{57}$$

By considering the conditions on (52), (54) and (56), we develop a funnel consisting of two equations that describes as

$$\sum_{j=1}^{3} g_j = 1 \tag{58}$$

and

$$\sum_{j=1}^{3} j g_j = -\left[\rho_{xy_2}\frac{C_{y_2}}{C_x} + \rho_{xy_1}\frac{C_{y_1}}{C_x}\right] \tag{59}$$

Thus, there are three unknowns in the above two equations that need to be found. Therefore, it is impossible to determine the precise amounts of chemicals (g_1, g_2, g_3) that should be employed for filtration. Because it is stated that their funnel requires a filter paper to function properly and is disabled. They recommended utilizing a filter paper with a funnel and applying linear restrictions as follows in order to obtain unique results for these chemicals.

$$\sum_{j=1}^{3} g_j Bias(\hat{T}_{3j}) = 0, \tag{60}$$

where $Bias(\hat{T}_{3j})$ represents the bias in the $\hat{T}_{3j}, j = 1,2,3$, of population mean. The aforementioned three equations can be expressed in the following form.

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ Bias(\hat{T}_{(31)}) & Bias(\hat{T}_{(32)}) & Bias(\hat{T}_{(33)}) \end{bmatrix} \begin{bmatrix} g_1 \\ g_2 \\ g_3 \end{bmatrix} = \begin{bmatrix} 1 \\ K_3 \\ 0 \end{bmatrix} \tag{61}$$

Solving the system of equation (61) yields the values of $g_j, j = 1,2,3$ which allowing the separation of bias precipitates from the linear class of estimators, provided that $|A| \neq 0$, i.e. if

$$Bias(\hat{T}_{(32)}) \neq \frac{Bias(\hat{T}_{(31)}) + Bias(\hat{T}_{(33)})}{2} \tag{62}$$

Thus, we have the following theorem.

Theorem 3:- *The values of g_j, for $j = 1,2,3$ that filter the bias up to terms of first order are given as*

$g_1 = \frac{1}{\lambda'}\big[3(\lambda-\lambda') - (3\lambda - 2\lambda')K_3 + (\lambda-\lambda')\big(K_3^2 - \frac{\lambda\tau_2}{(\lambda-\lambda')}\big)\big]$,

$g_2 = \frac{1}{\lambda'}\big[-3(2\lambda-\lambda') + 3(2\lambda - \lambda')K_3 - 2(\lambda-\lambda')\big(K_3^2 - \frac{\lambda\tau_2}{(\lambda-\lambda')}\big)\big]$ *and*

$g_3 = \frac{1}{\lambda'}\big[(3\lambda-\lambda') - (3\lambda - \lambda')K_3 + (\lambda-\lambda')\big(K_3^2 - \frac{\lambda\tau_2}{(\lambda-\lambda')}\big)\big]$,

where $\tau_2 = \frac{\rho_{y_1y_2}C_{y_1}C_{y_2}}{C_x}$.

Proof: To the first order of approximation

$$Bias(\hat{T}_{31}) = \bar{y}_1\bar{y}_2\big[\lambda\{\phi_1 + \phi_2\} - \lambda'\{\phi_1\}\big], \tag{63}$$

$$Bias(\hat{T}_{32}) = \bar{y}_1\bar{y}_2\big[\lambda\{3C_x^2 + 2\phi_1 + \phi_2\} - \lambda'\{C_x^2 + 2\phi_1\}\big], \tag{64}$$

$$Bias(\hat{T}_{33}) = \bar{y}_1\bar{y}_2\big[\lambda\{6C_x^2 + 3\phi_1 + \phi_2\} - \lambda'\{3C_x^2 + 3\phi_1\}\big]. \tag{65}$$

Where $\phi_1 = \big(\rho_{xy_2}C_{y_2}C_x + \rho_{xy_1}C_{y_1}C_x\big)$ and $\phi_2 = \rho_{y_1y_2}C_{y_1}C_{y_2}$.

On substituting these values of biases in (61), we develop the values of g_j as stated in the theorem.

5. Efficiencies Comparison

After filtration of biases, the performance of the proposed estimator has been examined by comparing it with existing estimator.

Situation 1

- $Bias(\hat{T}_{11}) < Bias(\hat{R}_1)$

$$\bar{Y}_1(\lambda-\lambda')\ell < \lambda\bar{Y}_1\big[C_x^2 - \rho_{xy_1}C_{y_1}C_x\big]$$

- $Bias(\hat{T}_{12}) < Bias(\hat{R}_2)$ if

$$\bar{Y}_1\big[\lambda\{3C_x^2 + 2\ell\} - \lambda'\{C_x^2 + 2\ell\}\big] < \lambda\bar{Y}_1\big[3C_x^2 - 2\rho_{xy_1}C_{y_1}C_x\big]$$

- $Bias(\hat{T}_{13}) < Bias(\hat{R}_3)$ if

$$\bar{Y}_1\big[3\lambda\{2C_x^2 + \ell\} - 3\lambda'\{C_x^2 + \ell\}\big] < \lambda\bar{Y}_1\big[6C_x^2 - 3\rho_{xy_1}C_{y_1}C_x\big]$$

Situation 2

- $Bias(\hat{T}_{21}) < Bias(\hat{R}_1)$ if

$$\left(\frac{\bar{Y}_1}{\bar{Y}_2}\right)\Big[\lambda\{\kappa_1 + \kappa_2\} - \lambda'\kappa_2\Big] < \lambda\bar{Y}_1\big[C_x^2 - \rho_{xy_1}C_{y_1}C_x\big]$$

- $Bias(\hat{T}_{22}) < Bias(\hat{R}_2)$ if

$$\left(\frac{\bar{Y}_1}{\bar{Y}_2}\right)\left[\lambda\left\{\kappa_1 + \left(3C_x^2 + 2\kappa_2\right)\right\} - \lambda'\left\{C_x^2 + 2\kappa_2\right\}\right] < \lambda\bar{Y}_1\left[3C_x^2 - 2\rho_{xy_1}C_{y_1}C_x\right]$$

- $Bias(\hat{T}_{23}) < Bias(\hat{R}_3)$ if

$$\left(\frac{\bar{Y}_1}{\bar{Y}_2}\right)\left[\lambda\left\{\kappa_1 + 3\left(2C_x^2 + \kappa_2\right)\right\} - \lambda'\left\{3C_x^2 + 2\kappa_2\right\}\right] < \lambda\bar{Y}_1\left[6C_x^2 - 3\rho_{xy_1}C_{y_1}C_x\right]$$

Situation 3

- $Bias(\hat{T}_{31}) < Bias(\hat{R}_1)$ if

$$\bar{y}_1\bar{y}_2\left[\lambda\left\{\phi_1 + \phi_2\right\} - \lambda'\left\{\phi_1\right\}\right] < \lambda\bar{Y}_1\left[C_x^2 - \rho_{xy_1}C_{y_1}C_x\right]$$

- $Bias(\hat{T}_{32}) < Bias(\hat{R}_2)$ if

$$\bar{y}_1\bar{y}_2\left[\lambda\left\{3C_x^2 + 2\phi_1 + \phi_2\right\} - \lambda'\left\{C_x^2 + 2\phi_1\right\}\right] < \lambda\bar{Y}_1\left[3C_x^2 - 2\rho_{xy_1}C_{y_1}C_x\right]$$

- $Bias(\hat{T}_{33}) < Bias(\hat{R}_3)$ if

$$\bar{y}_1\bar{y}_2\left[\lambda\left\{6C_x^2 + 3\phi_1 + \phi_2\right\} - \lambda'\left\{3C_x^2 + 3\phi_1\right\}\right] < \lambda\bar{Y}_1\left[6C_x^2 - 3\rho_{xy_1}C_{y_1}C_x\right]$$

When certain conditions are satisfied, the biases of the proposed estimator under different situations shows improved performance relative to the conventional estimator. To evaluate the accuracy of these relationships, we performed a simulation study using R software, with details provided in the following section.

6. Simulation Study

To compare the biases of our proposed estimator with respect to existing estimator i.e., Singh and Singh (1993) estimator, we generate an artificial population to verify the theoretical findings through simulation in R under two-phase sampling.

Consider, an artificial population of size $N = 1000$ generated from a normal distribution. From this population, a sample of $n' = 650$ is selected, followed by subsamples of sizes $n = 200$ and 250, using a two-phase sampling approach. Further, the study variables $\bar{Y}_1$ and $\bar{Y}_2$ are obtained from a normal distribution using model $Y_1 \sim aX_1 + N(0.5, 0.2)$ and $Y_2 \sim aX_1 + N(1, 0.2)$, respectively, where $a = 0.025$ and $X_1 \sim N(0.8, 2)$ is an auxiliary variable that enhance the efficiency of the estimator of the parameters of the main study variables.

Table 1. Biases of the estimators when $n = 200$ and 250.

Situations	$n = 200$			$n = 250$		
–	$Bias(\hat{R}_1)$ 0.008998589	$Bias(\hat{R}_2)$ 0.02686598	$Bias(\hat{R}_3)$ 0.05360216	$Bias(\hat{R}_1)$ 0.00796032	$Bias(\hat{R}_2)$ 0.02405191	$Bias(\hat{R}_3)$ 0.04827477
$\alpha = 0$	$Bias(\hat{T}_{11})$ -0.0001123188	$Bias(\hat{T}_{12})$ 0.02518788	$Bias(\hat{T}_{13})$ 0.0492942	$Bias(\hat{T}_{11})$ 0.0001402681	$Bias(\hat{T}_{12})$ 0.02321489	$Bias(\hat{T}_{13})$ 0.04483006
$\alpha = 1$	$Bias(\hat{T}_{21})$ 2.146925e-05	$Bias(\hat{T}_{22})$ 0.0248805	$Bias(\hat{T}_{23})$ 0.04856989	$Bias(\hat{T}_{21})$ 0.0002056409	$Bias(\hat{T}_{22})$ 0.02308691	$Bias(\hat{T}_{23})$ 0.04452211
$\alpha = -1$	$Bias(\hat{T}_{31})$ -0.0001892478	$Bias(\hat{T}_{32})$ 0.02555973	$Bias(\hat{T}_{33})$ 0.05009009	$Bias(\hat{T}_{31})$ 0.000125269	$Bias(\hat{T}_{32})$ 0.02339487	$Bias(\hat{T}_{33})$ 0.04519149

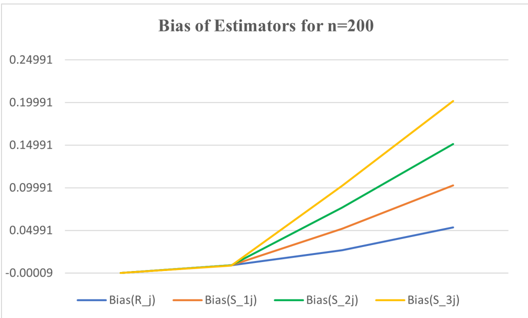

Figure 1. Biases of the estimators for n=200

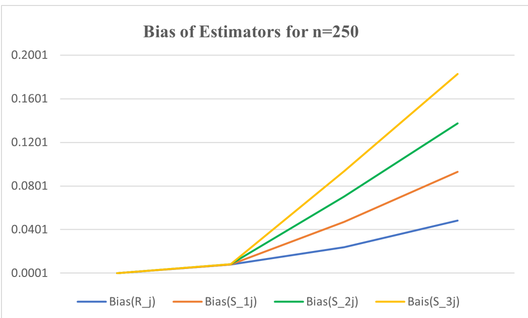

Figure 2. Biases of the estimators for n=250

From Table 1 and Figure 1 and Figure 2, it can be seen that the biases of the proposed estimator $\hat{T}_j$ under different situations i.e., $\alpha = 0, 1, -1$ is minimum in comparison to Singh and Singh's estimator $\hat{R}_j, j = 1, 2, 3$.

Based on theoretical finding and numerical study, we conclude that our suggested estimator is more reliable than the existing estimator by using filtration in order to approximate elimination of effect of bias.

7. Conclusion

A class of ratio and product type estimators after filtration of biases has been obtained using Singh and Singh (1993) estimator under three different situations. As we know that, whenever we drive a mean square error of an estimator, there is an effect of bias in an estimator. But in our proposed study, filtration is used to remove the approximate effect of bias. After deriving these situations, we justify that our proposed estimator after filtration performs better from the existing one and properties of the suggested estimator has been examined upto the first order of approximation. Thus, the simulation study confirms the theoretical results. Based on the findings in sections (5) and (6), clearly demonstrate that the proposed estimator exhibits superior performance compared to the existing one under two-phase sampling. Therefore, we strongly recommend that the researchers and practitioners adopt the suggested estimators in future studies.

References

[1] R. P. Chakrabarty. *Contribution to the theory of ratio type estimators.* Texas A&M University, 1968.

[2] W. G. Cochran. Some properties of estimators based on sampling scheme with varying probabilities. *Austrian Journal of Statistics*, 17:22–28, 1940.

[3] Hurwitz W. N. & Madow W. G. Hansen, M. H. *Sample Survey Methods and Theory.* John Wiley and Sons, New York, 1953.

[4] A Hartley, H. O. & Ross. Unbiased ratio estimators. *Nature*, 174:270–271, 1954.

[5] M. N. Murthy. Product method of estimation. *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)*, 26:69–74, 1964.

[6] B. Prasad. Some improved ratio type estimators of population mean and ratio in finite population sample surveys. *Communication in Statistics–Theory and Methods*, 18:379–392, 1989.

[7] H. P. Prasad, B. & Singh. Some improved ratio-type estimators of finite population variance in sample surveys. *Communication in Statistics–Theory and Methods*, 19:1127–1139, 1990.

[8] V. N. Reddy. On a transformed ratio method of estimation. *Sankhya*, 36:59–70, 1973.

[9] L. N. Sahoo. On a method of bias reduction in ratio estimation. *Journal of Statistical Research*, 17:1–6, 1983.

[10] H. P. Singh. A class of unbiased estimators of product of population means. *Journal of the Indian Society of Agricultural Statistics*, 41:113–118, 1989.

[11] L. N. Singh, H. P. & Sahoo. A class of almost unbiased estimators for population ratio and product. *Calcutta Statistical Association Bulletin*, 38:241–243, 1989.

[12] R. Singh, S. & Singh. A new method: Almost separation of bias precipitates in sample surveys. *Journal of Indian Statistical Association*, 31:99–105, 1993.

[13] S. K. Srivatsava. An estimator using auxiliary information in sample surveys. *Calcutta Statistical Association Bulletin*, 16:121–213, 1967.

[14] S. K. Srivatsava. A generalized estimator for the mean of finite population using multi-auxiliary information. *Journal of the American Statistical Association*, 66:404–407, 1971.

[15] J. W. E. Vos. Mixing of direct, ratio and product method estimators. *Statistica*, 34: 209–213, 1980.

[16] J. F. Walsh. Generalization of ratio estimator for population. *Sankhya*, 32:99–103, 1988.