https://doi.org/10.56093/JISAS.V80I1.20
Available online at https://isas.live/newsite
JOURNAL OF THE INDIAN SOCIETY OF
AGRICULTURAL STATISTICS 80(1) 2026 235–248

On the Unification of Auxiliary Information Estimators in Survey Sampling

Lane Christiansen, Sarjinder Singh and Stephen A. Sedory

Texas A&M University, Kingsville, Texas, USA

Received 05 March 2026; Revised 30 April 2026; Accepted 30 April 2026

SUMMARY

The use of auxiliary information has led to the development of a large number of estimators for improving the efficiency of population mean estimation in survey sampling. This study establishes a unified theoretical framework by demonstrating that many ratio-type, product-type, exponential, logarithmic, and regression-type estimators proposed in recent decades are special cases of the pioneering general class of estimators introduced by Srivastava (1971) and the wider class proposed by Singh and Upadhyaya (1986). Analytical expressions of mean squared error are presented, and it is shown that the minimum mean squared error achievable under Srivastava's class is equal to that of the classical linear regression estimator, while the Singh and Upadhyaya class provides a broader framework with additional flexibility. A recently proposed log-ratio estimator is examined and shown to require continuity-based patching. The results confirm the fundamental role of the Srivastava, and Singh and Upadhyaya classes as the underlying unifying structure for auxiliary information estimators and provide practical guidance for selecting efficient estimators in finite and large population settings.

Keywords: Auxiliary information; General class of estimators; Regression estimator, Mean squared error; Survey sampling.

1. INTRODUCTION

Assume a sample, s, of n units is selected by simple random without replacement sampling (SRSWOR) from a finite population Ω of N units. Let (y_i, x_i), $i = 1, 2, ..., N$ be the values of the study variable, y, and the auxiliary variable, x, respectively in the population. Let (y_i, x_i), $i = 1, 2, ..., n$ be the values of the study variable, y, and the auxiliary variable, x, respectively, in the sample.

Let

$\bar{y}_n = \frac{1}{n}\sum_{i=1}^{n} y_i$ be the sample mean of the study variable, y,

$\bar{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i$ be the sample mean of the auxiliary variable, x,

$\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} y_i$ be the population mean of the study variable, y,

$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} x_i$ be the population mean of the auxiliary variable, x,

Cochran (1940) was the first who suggested the ratio estimator for estimating the population mean $\bar{Y}$ as:

$$\bar{y}_R = \bar{y}\left(\frac{\bar{X}}{\bar{x}}\right) \tag{1.1}$$

It is easy to verify that the ratio estimator is more efficient than the sample mean estimator $\bar{y}$ for a value of correlation coefficient between 0.5 to 1.0.

Hansen, Hurwitz, and Madow (1953) proposed a regression estimator for estimating $\bar{Y}$ as:

Corresponding author: Sarjinder Singh
E-mail address: sarjinder@yahoo.com

$$\bar{y}_{lr} = \bar{y} + \hat{\beta}\left(\bar{X} - \bar{x}\right) \tag{1.2}$$

where $\hat{\beta} = s_{xy}/s_x^2$ is the sample regression coefficient for the sample co-variance and sample variance, respectively, given by

$s_{xy} = (n-1)^{-1}\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})$ and $s_x^2 = (n-1)^{-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$.

The regression estimator is always more efficient than the sample mean estimator $\bar{y}$ for any non-zero value of correlation coefficient between the study variable and the auxiliary variable.

Further let

$s_y^2 = (n-1)^{-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$ be the sample variance of the study variable, y,

$S_y^2 = (N-1)^{-1}\sum_{i=1}^{N}(y_i - \bar{Y})^2$; be the population mean squared error of the study variable, y,

and

$S_x^2 = (N-1)^{-1}\sum_{i=1}^{N}(x_i - \bar{X})^2$ be the population mean squared error of the auxiliary variable, x,

Let us define

$\varepsilon = \frac{\bar{y}}{\bar{Y}} - 1$, and $\delta = \frac{\bar{x}}{\bar{X}} - 1$

such that

$$E(\varepsilon) = E(\delta) = 0$$

$E(\varepsilon^2) = \left(\frac{1-f}{n}\right)C_y^2$, $E(\delta^2) = \left(\frac{1-f}{n}\right)C_x^2$, and $E(\varepsilon\delta) = \left(\frac{1-f}{n}\right)\rho_{xy}C_xC_y$

where

$C_y^2 = \frac{S_y^2}{\bar{Y}^2}$, $C_x^2 = \frac{S_x^2}{\bar{X}^2}$, $\rho_{xy} = \frac{S_{xy}}{S_xS_y}$ and $f = \frac{n}{N}$ is called the finite population correction factor (f. p. c.).

The minimum variance of the linear regression estimator is given by:

$$Min.V(\bar{y}_{lr}) = \left(\frac{1-f}{n}\right)S_y^2\left(1-\rho_{xy}^2\right) \tag{1.3}$$

In the next section we discuss the general class of ratio type estimators introduced by Srivastava (1971), and then we will show several estimators as a special case of his class of estimators.

2. GENERAL CLASS OF RATIO TYPE ESTIMATORS

Srivastava (1971) defined a general class of ratio type estimators as:

$$\bar{y}_{class} = \bar{y}h(u) \tag{2.1}$$

where $h(u)$ with $u = \frac{\bar{x}}{\bar{X}}$ is a parametric function such that it satisfies certain regularity conditions:

(a) $h(1) = 1$ (2.2)

(b) The first and second order partial derivatives of h with respect to u exist.

Let $A = \frac{\partial h}{\partial u}|_{u=1}$ and $B = \frac{\partial^2 h}{\partial u^2}|_{u=1}$ be known constants.

By using the Taylor's series approximation, it is easy to write an approximate result as:

$$\bar{y}_{class} \approx \bar{y}\left[h(1) + (u-1)\frac{\partial h}{\partial u}|_{u=1} + \frac{(u-1)^2}{2}\frac{\partial^2 h}{\partial u^2}|_{u=1} + ...\right]$$

$$= \bar{y}\left[1 + (u-1)A + \frac{(u-1)^2}{2}B + ...\right]$$

$$= \bar{Y}(1+\varepsilon)\left[1 + \delta A + \frac{\delta^2}{2}B + ...\right]$$

$$= \bar{Y}\left[1 + \varepsilon + \delta A + \frac{\delta^2}{2}B + \varepsilon\delta A + O_p\left(n^{-1}\right)\right] \tag{2.3}$$

The mean squared error of the class $\bar{y}_{class}$ is given by

$$MSE(\bar{y}_{class}) = E\left[\bar{y}_{class} - \bar{Y}\right]^2$$

$$\simeq E\left[\bar{Y}\left\{1+\varepsilon+\delta A+\frac{\delta^2}{2}B+\varepsilon\delta A+O_p\left(n^{-1}\right)\right\}-\bar{Y}\right]^2$$

$$\approx \bar{Y}^2 E\left[\varepsilon+\delta A\right]^2$$

$$=\bar{Y}^2 E\left[\varepsilon^2+\delta^2 A^2+2A\varepsilon\delta\right]$$

$$=\bar{Y}^2\left(\frac{1-f}{n}\right)\left[C_y^2+A^2C_x^2+2A\rho_{xy}C_yC_x\right]$$

$$=\left(\frac{1-f}{n}\right)\left[S_y^2+A^2S_x^2+2AS_{xy}\right] \quad (2.4)$$

On differentiating (2.4) with respect to A and equating it equal to zero, we

$$A=-\frac{S_{xy}}{S_x^2} \quad (2.5)$$

On substituting (2.5) in to (2.4), the minimum mean squared error of the general class of estimators is given by

$$Min.MSE\left(\bar{y}_{class}\right)=\left(\frac{1-f}{n}\right)S_y^2\left(1-\rho_{xy}^2\right) \quad (2.6)$$

Thus, the general class of ratio type estimators have variance equal to the linear regression estimator. Srivastava (1971) showed that if we calibrate the sample mean estimator $\bar{y}$ with any function $h(u)$, then mean squared error of the resultant estimator will be greater than or equal to the mean squared error of the linear regression estimator.

In the next section, we show a few estimators which are special case of the general class of estimators.

2.1 Ratio Estimator

In the ratio estimator due to Cochran (1940) given by

$$\bar{y}_R=\bar{y}\left(\frac{\bar{X}}{\bar{x}}\right) \quad (2.1.1)$$

we have

$$h(u)=u^{-1}$$

with $A=h'(1)=-1$ and $B=h''(1)=2$.

2.2 Product Estimator

In the product estimator due to Murthy (1964) given by

$$\bar{y}_P=\bar{y}\left(\frac{\bar{x}}{\bar{X}}\right) \quad (2.2.1)$$

we have

$$h(u)=u$$

with $A=h'(1)=1$ and $B=h''(1)=0$.

2.3 Sisodia and Dwivedi Estimator

Sisodia and Dwivedi (1981) estimator is given by:

$$\bar{y}_{SD}=\bar{y}\left(\frac{\bar{X}+C_x}{\bar{x}+C_x}\right) \quad (2.3.1)$$

In the estimator in Equation (2.3.1), we have

$$h(u)=\left(1+\frac{C_x}{\bar{X}}\right)\left(u+\frac{C_x}{\bar{X}}\right)^{-1} \text{ with } A=-\left(1+\frac{C_x}{\bar{X}}\right)^{-1}$$

and $B=2\left(1+\frac{C_x}{\bar{X}}\right)^{-2}$.

Note that for this choice of A and B the class of estimators will not have minimum variance and hence the estimator due to Sisodia and Dwivedi (1981) can never be more efficient than the linear regression estimator.

2.4 Singh and Tailor Estimator

Singh and Tailor (2003) estimator is given by:

$$\bar{y}_{ST}=\bar{y}\left(\frac{\bar{X}+\rho_{xy}}{\bar{x}+\rho_{xy}}\right) \quad (2.4.1)$$

In the estimator in Equation (2.4.1), we have

$$h(u)=\left(1+\frac{\rho}{\bar{X}}\right)\left(u+\frac{\rho}{\bar{X}}\right)^{-1} \text{ with } A=-\left(1+\frac{\rho}{\bar{X}}\right)^{-1}$$

and $B=2\left(1+\frac{\rho}{\bar{X}}\right)^{-2}$. Again, note that such choice of A and B will not lead to the minimum mean squared error of this general class of estimators.

2.5 Yan and Tian Estimator

Yan and Tian (2010) proposed two estimators, the first one is given by:

$$\bar{y}_{YT(1)} = \bar{y}\left(\frac{\bar{X}+\beta_{1(x)}}{\bar{x}+\beta_{1(x)}}\right) \tag{2.5.1}$$

where $\beta_{1(x)}$ is the population coefficient of skewness of the auxiliary variable, x.

In the estimator in Equation (2.5.1), we have

$h(u) = \left(1+\frac{\beta_{1(x)}}{\bar{X}}\right)\left(u+\frac{\beta_{1(x)}}{\bar{X}}\right)^{-1}$ with $A = -\left(1+\frac{\beta_{1(x)}}{\bar{X}}\right)^{-1}$ and $B = 2\left(1+\frac{\beta_{1(x)}}{\bar{X}}\right)^{-2}$. Again, note that such a choice of A and B will not lead to the minimum mean squared error of this general class of estimators.

Yan and Tian (2010) second estimator is given by:

$$\bar{y}_{YT(2)} = \bar{y}\left(\frac{\bar{X}C_x+\beta_{1(x)}}{\bar{x}C_x+\beta_{1(x)}}\right) \tag{2.5.2}$$

In the estimator in Equation (2.5.2), we have

$h(u) = \left(1+\frac{\beta_{1(x)}}{\bar{X}C_x}\right)\left(u+\frac{\beta_{1(x)}}{\bar{X}C_x}\right)^{-1}$ with $A = -\left(1+\frac{\beta_{1(x)}}{\bar{X}C_x}\right)^{-1}$ and $B = 2\left(1+\frac{\beta_{1(x)}}{\bar{X}C_x}\right)^{-2}$.

Again, note that such a choice of A and B will not lead to the minimum mean squared error of this general class of estimators.

2.6 Zaman, Bulut and Yadav Estimator

Zaman, Bulut and Yadav (2022) defined an estimator of the population mean as:

$$\bar{y}_{ZBY(1)} = \bar{y}\left(\frac{\bar{X}+M_d}{\bar{x}+M_d}\right) \tag{2.6.1}$$

where M_d is the known population median of the auxiliary variable. In the estimator in Equation (2.6.1), we have

$h(u) = \left(1+\frac{M_d}{\bar{X}}\right)\left(u+\frac{M_d}{\bar{X}}\right)^{-1}$ with $A = -\left(1+\frac{M_d}{\bar{X}}\right)^{-1}$ and $B = 2\left(1+\frac{M_d}{\bar{X}}\right)^{-2}$. Again, note that such a choice of A and B will not lead to the minimum mean squared error of this general class of estimators.

2.7 Subramani and Kumarapandiyan Estimator

Subramani and Kumarapandiyan (2012) defined an estimator as:

$$\bar{y}_{ZBY(1)} = \bar{y}\left(\frac{\bar{X}C_x+M_d}{\bar{x}C_x+M_d}\right) \tag{2.7.1}$$

In the estimator in Equation (2.7.1), we have

$h(u) = \left(1+\frac{M_d}{\bar{X}C_x}\right)\left(u+\frac{M_d}{\bar{X}C_x}\right)^{-1}$ with $A = -\left(1+\frac{M_d}{\bar{X}C_x}\right)^{-1}$ and $B = 2\left(1+\frac{M_d}{\bar{X}C_x}\right)^{-2}$. Again, note that such a choice of A and B will not lead to the minimum mean squared error of the general class of estimators.

2.8 Bahal and Tuteja Estimators

Bahal and Tuteja (1991) proposed ratio and product type estimators. We consider here the first one as:

$$\bar{y}_{BT(1)} = \bar{y}\exp\left[\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right] \tag{2.8.1}$$

In the estimator in Equation (2.8.1), we have

$h(u) = \exp\left(\frac{u^{-1}-1}{u^{-1}+1}\right)$ then $h(1) = 1$, $A = -\frac{1}{2}$ and $B = \frac{3}{4}$. Again, note that such a choice of A and B will not lead to the minimum mean squared error of this general class of estimators.

The second estimator of Bahal and Tuteja (1991) is:

$$\bar{y}_{BT(2)} = \bar{y}\exp\left[\frac{\bar{x}-\bar{X}}{\bar{x}+\bar{X}}\right] \tag{2.8.2}$$

In the estimator in Equation (2.8.2), we have

$h(u) = \exp\left(\frac{1-u^{-1}}{1+u^{-1}}\right)$ then $h(1) = 1$, $A = \frac{1}{2}$ and $B = -\frac{1}{4}$. Again, note that such a choice of A and B

will not lead to the minimum mean squared error of this general class of estimators.

The third estimator of Bahal and Tuteja (1991) is:

$$\bar{y}_{BT(3)} = \bar{y}\exp\left[\frac{\bar{x}-\bar{X}}{\bar{x}+\bar{X}+2K}\right] \tag{2.8.3}$$

where K is a constant to be determined such that the mean squared error of the estimator is minimum.

In the estimator in Equation (2.8.3), we have

$$h(u) = \exp\left(\frac{1-u^{-1}}{1+u^{-1}\left(1+\frac{2K}{\bar{X}}\right)}\right) \quad \text{then} \quad h(1)=1,$$

$A = \frac{1}{2}\left(1+\frac{K}{\bar{X}}\right)^{-1}$ and $B = -\frac{1}{4}\left(1+\frac{K}{\bar{X}}\right)^{-2}$. Again, note that such a choice of A and B will not lead to the minimum mean squared error of this general class of estimators. But the estimator can be approximated as:

$$\bar{y}_{BT(3)} \simeq \bar{Y}\left[1+\varepsilon+\frac{1}{2}\delta\left(1+\frac{K}{\bar{X}}\right)^{-1}-\frac{\delta^2}{8}\left(1+\frac{K}{\bar{X}}\right)^{-2}+\frac{1}{2}\varepsilon\delta\left(1+\frac{K}{\bar{X}}\right)^{-1}+O_p\left(n^{-1}\right)\right] \tag{2.8.4}$$

It is easy to verify that for the optimum value of K given by

$$K = -\bar{X}\left(1+\frac{C_x}{2\rho_{xy}C_y}\right) \tag{2.8.5}$$

The minimum mean squared error of the estimator $\bar{y}_{BT(3)}$ can be approximated, to the first order of approximation, as:

$$Min.MSE\left(\bar{y}_{BT(3)}\right) = \left(\frac{1-f}{n}\right)S_y^2\left(1-\rho_{xy}^2\right) \tag{2.8.6}$$

which is same minimum mean squared error as that of the linear regression estimator. This has also been reported in Bahal and Tuteja (1991), but it seems that they overlooked the class of estimators by Srivastava (1971). Likewise, Bahal and Tuteja ‘s fourth estimator is also a special case of the general class of estimators.

In the next section, we discuss a wider class of estimators proposed by Srivastava (1980) which includes all regression type estimators as special cases.

3. WIDER CLASS OF ESTIMATORS

Singh and Upadhyaya (1986) defined a wider class of estimators of the population mean $\bar{Y}$ as:

$$\bar{y}_{wclass} = g\left(\bar{y},u\right) \tag{3.1}$$

where $g\left(\bar{y},u\right)$ is any parametric function that satisfies the following regularity conditions:

(a) $g\left(\bar{Y},1\right) = \alpha\bar{Y}$, where α is a constant.

(b) The first and second order partial derivatives of the function g exist, such that the function can be approximated as:

$$g\left(\bar{y},u\right) = g\left[\bar{Y}+\left(\bar{y}-\bar{Y}\right),1+\left(u-1\right)\right]$$

$$= g\left(\bar{Y},1\right)+\left(\bar{y}-\bar{Y}\right)\frac{\partial g}{\partial\bar{y}}|_{\bar{y}=\bar{Y}}+\left(u-1\right)\frac{\partial g}{\partial u}|_{u=1}+\frac{\left(\bar{y}-\bar{Y}\right)^2}{2}\frac{\partial^2 g}{\partial\bar{y}^2}|_{\bar{y}=\bar{Y}}+\frac{\left(u-1\right)^2}{2}\frac{\partial^2 g}{\partial\mu^2}|_{u=1}+\left(\bar{y}-\bar{Y}\right)\left(u-1\right)\frac{\partial^2 g}{\partial u\partial\bar{y}}|_{u=1,\bar{y}=\bar{Y}}+..$$

$$= g\left(\bar{Y},1\right)+\left(\bar{y}-\bar{Y}\right)\alpha+\left(u-1\right)\beta+\frac{\left(\bar{y}-\bar{Y}\right)^2}{2}\frac{\partial^2 g}{\partial\bar{y}^2}|_{\bar{y}=\bar{Y}}+\frac{\left(u-1\right)^2}{2}\frac{\partial^2 g}{\partial\mu^2}|_{u=1}+\left(\bar{y}-\bar{Y}\right)\left(u-1\right)\frac{\partial^2 g}{\partial u\partial\bar{y}}|_{u=1,\bar{y}=\bar{Y}}+..$$

$$= \alpha\bar{y}+\left(u-1\right)\beta+O_p\left(n^{-1}\right) \tag{3.2}$$

The values of α and β are optimized by finding those values where

$$MSE\left(\bar{y}_{wclass}\right) = E\left[\alpha\bar{y}+\left(u-1\right)\beta-\bar{Y}\right]^2 \tag{3.3}$$

is minimum.

The optimum values of α and β are given by

$$\alpha = \frac{1}{1+\left(\frac{1-f}{n}\right)C_y^2\left(1-\rho_{xy}^2\right)} \tag{3.4}$$

and

$$\beta = -\frac{\bar{Y}\,\rho_{xy}\,C_y}{C_x\left\{1+\left(\frac{1-f}{n}\right)C_y^2\left(1-\rho_{xy}^2\right)\right\}} \tag{3.5}$$

The minimum mean squared error of the wider class of estimators is given by

$$\text{Min.MSE}\left(\bar{y}_{wclass}\right)=\frac{\left(\frac{1-f}{n}\right)\bar{Y}^2C_y^2\left(1-\rho_{xy}^2\right)}{1+\left(\frac{1-f}{n}\right)C_y^2\left(1-\rho_{xy}^2\right)} \quad (3.6)$$

The expression (3.6) shows that the wider class of estimators proposed by Singh and Upadhyaya (1986) is more efficient than the linear regression estimators. Note that the wider class of estimators by Srivastava (1980) is a special case of Singh and Upadhyaya (1986) for $\alpha=1$.

Now we discuss a few estimators as special cases of the wider class of estimators.

3.1 Subramani and Kumarapandiyan Estimators

We consider the first estimator proposed by Subramani and Kumarapandiyan (2012) as:

$$\bar{y}_{SK(1)}=\frac{\bar{y}+b\left(\bar{X}-\bar{x}\right)}{\bar{x}+M_d}\left(\bar{X}+M_d\right) \quad (3.1.1)$$

The estimator (3.1.1) can be rewritten as:

$$\bar{y}_{SK(1)}=g\left(\bar{y},u\right)=\frac{\bar{y}+b\bar{X}\left(1-u\right)}{\bar{X}u+M_d}\left(\bar{X}+M_d\right) \quad (3.1.2)$$

where

$$g\left(\bar{Y},1\right)=\bar{Y} \quad (3.1.3)$$

$$\alpha=\frac{\partial g\left(\bar{y},u\right)}{\partial\bar{y}}|_{\left(\bar{Y},1\right)}=1 \quad (3.1.4)$$

and

$$\beta=\frac{\partial g\left(\bar{y},u\right)}{\partial u}|_{\left(\bar{Y},1\right)}=-b\bar{X}-\frac{\bar{Y}\ \bar{X}}{\bar{X}+M_d} \quad (3.1.5)$$

that is the estimator in (3.1.1) is a special case of both Srivastava (1980) as well as Singh and Upadhyaya (1986) classes of estimators.

The modified second estimator proposed by Subramani and Kumarapandiyan (2012) is given by

$$\bar{y}_{SK(2)}=\frac{\bar{y}+b\left(\bar{X}-\bar{x}\right)}{\bar{x}C_x+M_d}\left(\bar{X}C_x+M_d\right) \quad (3.1.6)$$

The estimator (3.1.6) can be rewritten as:

$$\bar{y}_{SK(2)}=g\left(\bar{y},u\right)=\frac{\bar{y}+b\bar{X}\left(1-u\right)}{\bar{X}C_x\,u+M_d}\left(\bar{X}C_x+M_d\right) \quad (3.1.7)$$

where

$$g\left(\bar{Y},1\right)=\bar{Y} \quad (3.1.8)$$

$$\alpha=\frac{\partial g\left(\bar{y},u\right)}{\partial\bar{y}}|_{\left(\bar{Y},1\right)}=1 \quad (3.1.9)$$

and

$$\beta=\frac{\partial g\left(\bar{y},u\right)}{\partial u}|_{\left(\bar{Y},1\right)}=-b\bar{X}-\frac{\bar{Y}\ \bar{X}C_x}{\bar{X}C_x+M_d} \quad (3.1.10)$$

that is the estimator in (3.1.1) is a special case of both Srivastava (1980) as well as Singh and Upadhyaya (1986) classes of estimators.

3.2 Mishra, Adichwal and Singh Estimators

We consider the first estimator proposed by Mishra, Adichwal and Singh (2017) as:

$$\bar{y}_{MAS(1)}=\bar{y}+K\log\left(\frac{\bar{x}}{\bar{X}}\right) \quad (3.2.1)$$

where K is a constant to be determined such that the mean squared error is minimum.

The estimator (3.2.1) can be rewritten as:

$$\bar{y}_{MAS(1)}=g\left(\bar{y},u\right)=\bar{y}+K\log\left(u\right) \quad (3.2.2)$$

where

$$g\left(\bar{Y},1\right)=\bar{Y} \quad (3.2.3)$$

$$\alpha=\frac{\partial g\left(\bar{y},u\right)}{\partial\bar{y}}|_{\left(\bar{Y},1\right)}=1 \quad (3.2.4)$$

and

$$\beta=\frac{\partial g\left(\bar{y},u\right)}{\partial u}|_{\left(\bar{Y},1\right)}=K \quad (3.2.5)$$

that is the estimator in (3.2.1) is a special case of both Srivastava (1980) as well as Singh and Upadhyaya (1986) classes of estimators.

We consider the second estimator proposed by Mishra, Adichwal and Singh (2017) as:

$$\bar{y}_{MAS(2)}=\bar{y}\left(w_1+1\right)+w_2\log\left(\frac{\bar{x}}{\bar{X}}\right) \quad (3.2.6)$$

where w_1 and w_2 are constants to be determined such that the mean squared error is minimum.

The estimator (3.2.6) can be rewritten as:

$$\bar{y}_{MAS(1)} = g(\bar{y},u) = \bar{y}(w_1+1) + w_2 \log(u) \quad (3.2.7)$$

where

$$g(\bar{Y},1) = (w_1+1)\bar{Y} \quad (3.2.8)$$

$$\alpha = \frac{\partial g(\bar{y},u)}{\partial \bar{y}}\Big|_{(\bar{Y},1)} = (w_1+1) \quad (3.2.9)$$

and

$$\beta = \frac{\partial g(\bar{y},u)}{\partial u}\Big|_{(\bar{Y},1)} = w_2 \quad (3.2.10)$$

that is the estimator in (3.2.6) is a special case of Singh and Upadhyaya (1986) class but not of Srivastava (1980) class of estimators.

3.3 Singh and Tiwari Estimators

In case of one auxiliary variable, we consider the estimator proposed by Singh and Tiwari (2025) as:

$$\bar{y}_{ST} = \bar{y} + K_1 \log\left(\frac{\bar{X}+M_d}{\bar{x}+M_d}\right) \quad (3.3.1)$$

where K_1 is a constant to be determined such that the mean squared error is minimum.

The estimator (3.3.1) can be rewritten as:

$$\bar{y}_{ST} = g(\bar{y},u) = \bar{y} + K_1 \log\left(\frac{1+\frac{M_d}{\bar{X}}}{u+\frac{M_d}{\bar{X}}}\right) \quad (3.3.2)$$

where

$$g(\bar{Y},1) = \bar{Y} \quad (3.3.3)$$

$$\alpha = \frac{\partial g(\bar{y},u)}{\partial \bar{y}}\Big|_{(\bar{Y},1)} = 1 \quad (3.3.4)$$

and

$$\beta = \frac{\partial g(\bar{y},u)}{\partial u}\Big|_{(\bar{Y},1)} = -\frac{K_1\bar{X}}{\bar{X}+M_d} \quad (3.3.5)$$

that is the estimator in (3.3.1) is a special case of both Srivastava (1980) as well as Singh and Upadhyaya (1986) classes of estimators. The optimum value of K_1 will lead to same minimum mean squared error as that of the linear regression estimator.

3.4 Gupta and Shabbir Estimators

Gupta and Shabbir (2008) have proposed a general class of estimators as:

$$\bar{y}_{gs} = \left[w_1\bar{y} + w_2(\bar{X}-\bar{x})\right]\left(\frac{\eta\bar{X}+\lambda}{\eta\bar{x}+\lambda}\right) \quad (3.4.1)$$

where w_1, w_2, η and λ are constant. They have shown that the minimum mean squared error of the estimator $\bar{y}_{gs}$ is same as that of the wider class of estimators

$$\text{Min.MSE}(\bar{y}_{gs}) = \frac{\left(\frac{1-f}{n}\right)\bar{Y}^2 C_y^2(1-\rho_{xy}^2)}{1+\left(\frac{1-f}{n}\right)C_y^2(1-\rho_{xy}^2)} = \text{Min.MSE}(\bar{y}_{wclass}) \quad (3.4.2)$$

Kadilar and Cingi (2004) estimators given by:

$$\bar{y}_{kc(t)} = \alpha_t\left[\bar{y} + b(\bar{X}-\bar{x})\right], \text{ for } t = 1,2,3,4,5, \quad (3.4.3)$$

where $\alpha_1 = \frac{\bar{X}}{\bar{x}}$, $\alpha_2 = \frac{\bar{X}+C_x}{\bar{x}+C_x}$, $\alpha_3 = \frac{\bar{X}+\beta_{2(x)}}{\bar{x}+\beta_{2(x)}}$, $\alpha_4 = \frac{\bar{X}\beta_{2(x)}+C_x}{\bar{x}\beta_{2(x)}+C_x}$, and $\alpha_5 = \frac{\bar{X}C_x+\beta_{2(x)}}{\bar{x}C_x+\beta_{2(x)}}$ Kadilar and Cingi (2006a, 2006b) estimators given by:

$$\bar{y}_{kc(t)} = \alpha_t^*\left[\bar{y} + b(\bar{X}-\bar{x})\right], \text{ for } t = 1,2,3,4,5, \quad (3.4.4)$$

where

$\alpha_1^* = \frac{\bar{X}+\rho_{yx}}{\bar{x}+\rho_{yx}}$, $\alpha_2^* = \frac{\bar{X}C_x+\rho_{yx}}{\bar{x}C_x+\rho_{yx}}$, $\alpha_3^* = \frac{\bar{X}\rho_{yx}+C_x}{\bar{x}\rho_{yx}+C_x}$, $\alpha_4^* = \frac{\bar{X}\beta_{2(x)}+\rho_{yx}}{\bar{x}\beta_{2(x)}+\rho_{yx}}$, and $\alpha_5^* = \frac{\bar{X}\rho_{yx}+\beta_{2(x)}}{\bar{x}\rho_{yx}+\beta_{2(x)}}$ are also special cases of the wider class of estimators.

Kadilar and Cingi (2006c) also proposed a new estimator as:

$$\bar{y}_{kc(t)} = w_1\left[\bar{y} + b(\bar{X}-\bar{x})\right]\alpha_1 + w_2\left[\bar{y} + b(\bar{X}-\bar{x})\right]\alpha_t,$$

for $t = 2,3,4,5,$ (3.4.5)

where $w_1 + w_2 = 1$. The estimators by Kadilar and Cingi (2006a,2006b,2006c) are no more efficient than the estimator $\bar{y}_{gs}$ which is also not efficient than the wider class of estimators by Singh and Upadhyaya (1986).

In the next section, we consider a very interesting estimator which is not a special case of the classes of estimators proposed by Srivastava (1971, 1980) and Singh and Upadhyaya (1986) but has some special interesting observations.

3.5 Shakoor, Atrif, Ali, Alshammari, Himmat and Kefi Estimators

Shakoor *et al.* (2026) proposed first log-ratio type estimators as:

$$\bar{y}_{p1} = \left(k_1\bar{y} + k_2\right)\left(\frac{\bar{X}}{\bar{x} - \bar{X}}\right)\ln\left(\frac{\bar{x}}{\bar{X}}\right) \tag{3.5.1}$$

where k_1 and k_2 ae constants to be determined such that the mean squared error is minimum. The estimator in (3.5.1) has some technical issues. Mathematically it is a good estimator because the Taylor's series expression is working well as shown in (3.5.9). It may have computational problems especially in case of finite populations. For example, while using SRSWOR, there may be many samples where the sample mean $\bar{x}$ may take the same value as the population mean $\bar{X}$ making it non-functional. Thus, it can be used by adding a discontinuity patch as follows:

$$\bar{y}_{p1} = \begin{cases} \left(k_1\bar{y} + k_2\right)\left(\dfrac{\bar{X}}{\bar{x} - \bar{X}}\right)\ln\left(\dfrac{\bar{x}}{\bar{X}}\right) & \text{if } \bar{x} \neq \bar{X} \\ \left(k_1\bar{y} + k_2\right) & \text{if } \bar{x} = \bar{X} \end{cases} \tag{3.5.2}$$

It is worth pointing out that if $\bar{x} = \bar{X}$ then the ratio, product and regression estimator also reduce to sample mean estimator but there is no computational discontinuity. Thus,to remove the computational discontinuity a special patch is required as defined above.

The estimator (3.5.1) can be rewritten as:

$$\bar{y}_{p1} = g\left(\bar{y}, u\right) = \left(k_1\bar{y} + k_2\right)\left(\frac{1}{u-1}\right)\ln\left(u\right) \tag{3.5.3}$$

where

$$g\left(\bar{Y}, 1\right) = k_1\bar{Y} + k_2 \tag{3.5.4}$$

$$\alpha = \frac{\partial g\left(\bar{y}, u\right)}{\partial \bar{y}}\Big|_{\left(\bar{Y},1\right)} = k_1 \tag{3.5.5}$$

and

$$\beta = \frac{\partial g\left(\bar{y}, u\right)}{\partial u}\Big|_{\left(\bar{Y},1\right)} = -\frac{1}{2}\left(k_1\bar{Y} + k_2\right) \tag{3.5.6}$$

that is the estimator in (3.5.1) is a special case neither of Srivastava (1980) nor of Singh and Upadhyaya (1986) classes of estimators unless $k_2 = 0$. The estimator in (3.5.1) is mathematically correct because by L-Hospital Rule, we have

$$\lim_{u \to 1} \frac{\ln\left(u\right)}{u-1} = 1 \tag{3.5.7}$$

It is also easy to verify that:

$$\ln\left(u\right) = \left(u-1\right) - \frac{1}{2}\left(u-1\right)^2 + \frac{1}{3}\left(u-1\right)^3 + \ldots \tag{3.5.8}$$

So, we have

$$\frac{\ln\left(u\right)}{\left(u-1\right)} = 1 - \frac{1}{2}\left(u-1\right) + \frac{1}{3}\left(u-1\right)^2 + \ldots \tag{3.5.9}$$

Thus

$$\bar{y}_{p1} = \left(k_1\bar{y} + k_2\right)\left(\frac{\bar{X}}{\bar{x} - \bar{X}}\right)\ln\left(\frac{\bar{x}}{\bar{X}}\right)$$

$$= \left(k_1\bar{y} + k_2\right)\left[1 - \frac{1}{2}\left(u-1\right) + \frac{1}{3}\left(u-1\right)^2 + \ldots.\right]$$

$$= \left[k_1\bar{Y}\left(1+\varepsilon\right) + k_2\right]\left[1 - \frac{1}{2}\delta + \frac{1}{3}\delta^2 + \ldots.\right]$$

$$= k_1\bar{Y}\left(1+\varepsilon\right) + k_2 - \frac{1}{2}\delta\left[k_1\bar{Y}\left(1+\varepsilon\right) + k_2\right] + \frac{1}{3}\delta^2\left[k_1\bar{Y}\left(1+\varepsilon\right) + k_2\right] + O_p\left(n^{-1}\right)$$

$$= k_1\bar{Y}\left(1+\varepsilon\right) + k_2 - \frac{k_1\bar{Y}}{2}\delta - \frac{k_1\bar{Y}}{2}\varepsilon\delta - \frac{k_2}{2}\delta + \frac{k_1\bar{Y}}{3}\delta^2 + \frac{k_2}{3}\delta^2 + O_p\left(n^{-1}\right)$$

$$= k_1\bar{Y} + k_1\bar{Y}\varepsilon - \frac{k_1\bar{Y}}{2}\delta - \frac{k_2}{2}\delta + k_2 - \frac{k_1\bar{Y}}{2}\varepsilon\delta + \frac{k_1\bar{Y}}{3}\delta^2 + \frac{k_2}{3}\delta^2 + O_p\left(n^{-1}\right)$$

$$= k_1\bar{Y} + k_1\bar{Y}\left(\varepsilon - \frac{\delta}{2}\right) - \frac{k_2}{2}\delta + k_1\bar{Y}\left(\frac{\delta^2}{3} - \frac{\varepsilon\delta}{2}\right) + k_2\left(1 + \frac{\delta^2}{3}\right) + O_p\left(n^{-1}\right) \quad (3.5.10)$$

From (3.5.10), the bias in the estimator $\bar{y}_{p_1}$ is given by

$$B\left(\bar{y}_{p_1}\right) = E\left(\bar{y}_{p_1}\right) - \bar{Y}$$

$$= E\left[k_1\bar{Y} + k_1\bar{Y}\left(\varepsilon - \frac{\delta}{2}\right) - \frac{k_2}{2}\delta + k_1\bar{Y}\left(\frac{\delta^2}{3} - \frac{\varepsilon\delta}{2}\right) + k_2\left(1 + \frac{\delta^2}{3}\right) + O_p\left(n^{-1}\right)\right] - \bar{Y}$$

$$= \left(k_1 - 1\right)\bar{Y} + k_1\bar{Y}\left(\frac{E\left(\delta^2\right)}{3} - \frac{E\left(\varepsilon\delta\right)}{2}\right) + k_2\left(1 + \frac{E\left(\delta^2\right)}{3}\right) + O\left(n^{-1}\right)$$

$$= \left(k_1 - 1\right)\bar{Y} + k_2 + \left(\frac{1-f}{n}\right)\left[k_1\bar{Y}\left(\frac{C_x^2}{3} - \frac{\rho_{xy}C_xC_y}{2}\right) + k_2\left(\frac{C_x^2}{3}\right) + O\left(n^{-1}\right)\right]$$

If $n \to N$, then the bias in the estimator $\bar{y}_{p_1}$ is given by

$$B\left(\bar{y}_{p_1}\right) = \left(k_1 - 1\right)\bar{Y} + k_2 \quad (3.5.11)$$

If $k_1 \to 1$ and $k_2 \to 0$, then the bias expression $B\left(\bar{y}_2\right) \to 0$. Thus the estimator $\bar{y}_{p_1}$ is a biased estimator like Searls (1964) estimator but has additional factor of k_2. Otherwise, it is an inconsistent estimator.

The mean squared error of the estimator $\bar{y}_{p_1}$ is given by

$$MSE\left(\bar{y}_{p_1}\right) = E\left[\bar{y}_{p_1} - \bar{Y}\right]^2$$

$$= E\left[k_1\bar{Y}\left(1 + \varepsilon - \frac{\delta}{2} + \frac{\delta^2}{3} - \frac{\varepsilon\delta}{2} + ..\right) + k_2\left(1 - \frac{\delta}{2} + \frac{\delta^2}{3} + ..\right) - \bar{Y}\right]^2$$

$$= k_1^2\bar{Y}^2A_1 + k_2^2B_1 - 2k_1\bar{Y}^2C_1 - 2k_2\bar{Y}D_1 + 2k_1k_2\bar{Y}E_1 + \bar{Y}^2 \quad (3.5.12)$$

where

$$A_1 = E\left[1 + \varepsilon - \frac{\delta}{2} + \frac{\delta^2}{3} - \frac{\varepsilon\delta}{2}\right]^2 \approx 1 + \left(\frac{1-f}{n}\right)\left\{C_y^2 + \frac{11}{12}C_x^2 - 2\rho_{xy}C_yC_x\right\} \quad (3.5.13)$$

$$B_1 = E\left[1 - \frac{\delta}{2} + \frac{\delta^2}{3}\right]^2 \approx 1 + \frac{11}{12}\left(\frac{1-f}{n}\right)C_x^2 \quad (3.5.14)$$

$$C_1 = E\left[1 + \varepsilon - \frac{\delta}{2} + \frac{\delta^2}{3} - \frac{\varepsilon\delta}{2}\right] \approx 1 + \left(\frac{1-f}{n}\right)\left(\frac{C_x^2}{3} - \frac{\rho_{xy}C_yC_x}{2}\right) \quad (3.5.15)$$

$$D_1 = E\left[1 - \frac{\delta}{2} + \frac{\delta^2}{3}\right] = 1 + \frac{1}{3}\left(\frac{1-f}{n}\right)C_x^2 \quad (3.5.16)$$

and

$$E_1 = E\left[\left(1 + \varepsilon - \frac{\delta}{2} + \frac{\delta^2}{3} - \frac{\varepsilon\delta}{2}\right)\left(1 - \frac{\delta}{2} + \frac{\delta^2}{3}\right)\right] \approx 1 + \left(\frac{1-f}{n}\right)\left(\frac{11C_x^2}{12} - \rho_{xy}C_yC_x\right) \quad (3.5.17)$$

The mean squared error in (3.5.12) is minimum for

$$k_1 = \frac{C_1B_1 - D_1E_1}{A_1B_1 - E_1^2} \quad (3.5.18)$$

and

$$k_2 = \frac{\bar{Y}\left(A_1D_1 - C_1E_1\right)}{A_1B_1 - E_1^2} \quad (3.5.19)$$

The minimum mean squared error, to the first order of approximation, is given by

$$Min.MSE\left(\bar{y}_{p_1}\right)=\bar{Y}^2\left(1-\frac{A_1D_1^2+B_1C_1^2-2C_1D_1E_1}{A_1B_1-E_1^2}\right) \tag{3.5.20}$$

The approximate percent relative efficiency of the estimator $\bar{y}_{p_1}$ with respect to the wider class of estimators can be written as:

$$RE\left(\bar{y}_{p_1}\right)=\left[\frac{\left(\frac{1-f}{n}\right)C_y^2\left(1-\rho_{xy}^2\right)}{1+\left(\frac{1-f}{n}\right)C_y^2\left(1-\rho_{xy}^2\right)}\right]\left(1-\frac{A_1D_1^2+B_1C_1^2-2C_1D_1E_1}{A_1B_1-E_1^2}\right)^{-1} \tag{3.5.21}$$

Note that the values of A_1, B_1, C_1, D_1 and E_1 depend on the five parameters in any bivariate population viz, C_y, C_x, ρ_{xy}, f and n. For demonstration purposes, we considered nine values of C_y between 0.1 to 0.9 with a skip of 0.1, nine values of C_x between 0.1 to 0.9 with a skip of 0.1, nine values of ρ_{xy} between 0.1 to 0.9 with a skip of 0.1, nine values of f between 0.01 to 0.20 with a skip of 0.01, and $n=100$. Based on 14580 combinations, the average relative efficiency value is 1078.1 with standard error of 21.8. Thus, it is not a surprise that this estimator proposed by Shakoor *et al.* (2026) is not a member of the general class of the estimators proposed by Srivastava (1971, 1980).

To make the estimator $\bar{y}_{p_1}$ functional, it is suggested to use the following jackknifing mechanism:

$$\bar{y}_{p1}(j)=\begin{cases}\left(\hat{k}_1\bar{y}(j)+\hat{k}_2\right)\left(\dfrac{\bar{X}}{\bar{x}(j)-\bar{X}}\right)\ln\left(\dfrac{\bar{x}(j)}{\bar{X}}\right) & \text{if } \bar{x}(j)\neq\bar{X}\\ \left(\hat{k}_1\bar{y}(j)+\hat{k}_2\right) & \text{if } \bar{x}(j)=\bar{X}\end{cases} \tag{3.5.22}$$

where $\hat{k}_1$ and $\hat{k}_2$ are consistent estimators of k_1 and k_2 obtained by the method of moments.

The jackknifed sample means:

$$\bar{y}(j)=\frac{n\bar{y}-y_j}{n-1} \tag{3.5.23}$$

and

$$\bar{x}(j)=\frac{n\bar{x}-x_j}{n-1} \tag{3.5.24}$$

are the sample means obtained by dropping the jth unit from the sample.

Then the mean squared error of the estimator $\bar{y}_{p_1}$ can be estimated as;

$$v\left(\bar{y}_{p_1}\right)_{Jack}=\frac{(n-1)}{n}\sum_{j=1}^{n}\left[\bar{y}_{p_1}(j)-\frac{1}{n}\sum_{j=1}^{n}\bar{y}_{p_1}(j)\right]^2 \tag{3.5.25}$$

A $(1-\alpha)100\%$ confidence interval estimate is computed as:

$$\bar{y}_{p_1}^{Jack}\mp t_{\frac{\alpha}{2}}(df=n-3)\sqrt{v\left(\bar{y}_{p_1}\right)_{Jack}} \tag{3.5.26}$$

where

$$\bar{y}_{p_1}^{Jack}=\frac{1}{n}\sum_{j=1}^{n}\bar{y}_{p_1}(j) \tag{3.5.27}$$

Note the use of degree of freedom $df=n-3$ because we are estimating three parameters $\bar{Y}$, k_1 and k_2.

In the next section we demonstrate with an example that could occur with finite populations.

3.5.1. Exact Computational Results

In *Amy's laboratory*, a controlled biological experiment was conducted on a finite population of $N=20$ plants to investigate efficient estimation of a destructively measured response variable. The primary study variable Y represents plant yield, whose direct measurement requires harvesting and therefore cannot be obtained for all experimental units. An auxiliary variable X corresponding to leaf chlorophyll content, was concurrently available; this variable is inexpensive, non-destructive to measure, and its population mean was known in advance from complete routine assessment of all plants in the lab.

Table 3.5.1. Ordered paired values of (x_i, y_i) in the population

(40,65)	(42,75)	(44,60)	(46,50)	(48,55)	(50,68)	(52,74)	(54,71)	(56,63)	(58,78)
(60,84)	(62,74)	(64,76)	(66,73)	(70,75)	(77,91)	(72,96)	(79,93)	(78,89)	(82,95)

Fig. 3.5.1. Lab's controlled experiment

A simple random sample of $n = 6$ plants was selected without replacement, and both X and Y were observed for the sampled units. The true values of all the $N = 20$ ordered pair values (x_i, y_i) are given in Table 3.5.1.

At the end of the season, when all the crop was harvested from the $N = 20$ plots by Amy's lab administration and all the values of the study variable and auxiliary variable were made available, Amy wonders what was the accuracy of the estimates that she reported from the sample.

A statistician *Michael* in the lab looked at the data and realized that the Shakoor *et al.* (2026) estimator canbe implemented with patches as defined in the present paper. This laboratory setting demonstrates how principled use of auxiliary information can reduce biological loss and experimental cost while achieving statistically efficient inference in small-population biological systems. Michael selected all possible samples of sizes n from the population of N units which results in total $\binom{N}{n}$ samples. From each sample, he computed estimates the point estimates $\bar{y}_{p_1}^{Jack}$, $v(\bar{y}_{p_1})_{Jack}$ and the 95% confidence interval estimates (LL, UL). Then was counted the number of times the true population mean $\bar{Y}$ fell in the nominal 95% coverage with $t_{0.025}(df = n-3)$. It was also noted how many times the same mean $\bar{x}$ was equal to the population mean $\bar{X}$ which leads to computational issues in case of finite populations. We compared the estimator $\bar{y}_{p_1}$ with respect to the three estimators:

$$\hat{\theta}_1 = \bar{y} \tag{3.5.28}$$

The generalized regression estimator (GREG) due to Deville and Sarndal (1992):

$$\hat{\theta}_2 = \bar{y}_{greg} = \bar{y} + \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}\left(\bar{X} - \bar{x}\right) \tag{3.5.29}$$

and the linear regression estimator:

$$\hat{\theta}_3 = \bar{y}_{ylr} = \bar{y} + \frac{n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}\left(\bar{X} - \bar{x}\right) \tag{3.5.30}$$

Note that the linear regression estimator adjusts the correction factors in the estimation of regression coefficient as reported in Singh (2006).

The exact relative efficiency of the estimator $\bar{y}_{p_1}$ with respect to the other three estimators was computed as:

$$RE(k) = \frac{\sum_{t=1}^{nitr}\left[\hat{\theta}_{k|t} - \bar{Y}\right]^2}{\sum_{t=1}^{nitr}\left[\bar{y}_{p_1|t}^{Jack} - \bar{Y}\right]^2} \times 100\% \tag{3.5.31}$$

with $nitr = \binom{N}{n}$ that is all possible without replacement samples and $k = 1,2,3$. Let $nmatch$ denote the number of samples in which the sample mean $\bar{x}$ is same as the population mean $\bar{X}$, that is the computation of the estimator $\bar{y}_{p_1}$ is not feasible. The findings are reported in Table 3.5.2 which are guidelines to Amy for making choice of an estimator in future studies.

It is noted that while fitting the linear regression model the y-intercept is 25.75 with p-value 0.0051 and slope 0.8253 with p-value 6.37E-06 which means both the y-intercept and slop are significantly different from zero. The value of population correlation between y and x is 0.8287.

Table 3.5.2. Performance of the patched estimator $\bar{y}_{p_1}$

n	*nitr*	*nmatch*	$t_{0.025}$ ($df = n-3$)	95% coverage	$RE(1)$	$RE(2)$	$RE(3)$
5	15504	265	4.3027	0.9327	280.96	144.23	127.98
6	38760	570	3.1824	0.9027	281.82	142.28	116.38
7	77520	1007	2.7764	0.8957	282.42	141.26	111.68

When considering all 15504 possible samples each of $n=5$ units, there are 265 cases where the sample mean $\bar{x}$ matched with the population mean $\bar{X}$ thus leading to computational problem and requiring patches. The nominal 95% coverage is underestimated as 93.27%. The estimator $\bar{y}_{p_1}$ is more efficient than the sample mean, the GREG, and the traditional linear regression estimator due to Hansen, Hurwitz and Madow (1953) with relative efficiency values of 280.96%, 144.23%, and 127.98% respectively.

Next, considering all 38760 possible samples each of $n=6$ units, there were 570 cases where the sample mean $\bar{x}$ matched with the population mean $\bar{X}$. The nominal 95% coverage is underestimated as 90.27%. The values of the percent relative efficiencies are 281.82%, 142.48%, and 116.38%, respectively.

Finally, considering all 77520 samples each of $n=7$ units, there are 1070 cases where the sample mean $\bar{x}$ matched with the population mean $\bar{X}$. The nominal 95% coverage is underestimated as 89.57% which is under coverage. The values of the percent relative efficiencies are 282.42%, 141.26%, and 111.68%, respectively.

Following Stearns and Singh (2008), efficiency gains from auxiliary information depend critically on the structural adequacy of the working model rather than on calibration alone. In particular, when the relationship between the study variable Y and the auxiliary variable X is well represented by a linear model with a nonzero intercept, ratio-calibrated generalized regression (GREG) estimators, which implicitly impose a zero-intercept constraint, become structurally mis-specified. In such cases, the protection afforded by calibration introduces additional variability without compensatory bias reduction. Consequently, in finite populations and for small sample sizes, the linear regression estimator that explicitly accommodates a nonzero intercept may achieve lower mean squared error than the corresponding GREG estimator when evaluated over the complete set of possible samples. This behavior is consistent with the general efficiency robustness tradeoff articulated by Stearns and Singh (2008). Amy is happy that she selected a sample of $n=6$ plants, so her estimation process is more efficient than all the competitors. From this study, Amy is advised to use the patched estimator in her future estimation strategies.

3.5.2. Suggested Wider Class of Estimators

Shakoor *et al.* (2026) proposed a second estimator of the population mean as:

$$\bar{y}_{p_2} = \left(k_1^*\bar{y} + k_2^*\right)\ln\left(\frac{\bar{x}}{\bar{X}}\right)\exp\left(\frac{\bar{x}-\bar{X}}{\bar{x}+\bar{X}}\right) \tag{3.5.32}$$

The estimator $\bar{y}_{p_2}$ reduces to zero if $\bar{x}=\bar{X}$, thus there is also need of a patch attached to remove the computational discontinuity of the estimator as:

$$\bar{y}_{p_2} = \begin{cases} \left(k_1^*\bar{y} + k_2^*\right)\ln\left(\frac{\bar{x}}{\bar{X}}\right)\exp\left(\frac{\bar{x}-\bar{X}}{\bar{x}+\bar{X}}\right) & \text{if } \bar{x}\neq\bar{X} \\ k_1^*\bar{y} + k_2^* & \text{if } \bar{x}=\bar{X} \end{cases} \tag{3.5.33}$$

Similar jackknifing process can be implemented if needed to construct confidence interval estimates. Although we suggest jackknifing, this estimator seems impractical.

In case one wishes to use these types of inconsistent estimators, then we suggest a new wider class of estimators as:

$$\bar{y}_{nwclass} = \left(k_1\bar{y} + k_2\right)G(u) \tag{3.5.34}$$

where $G(u)$ is a function of u which may or may not need patching depending on its choice as discussed.

3.6 Conclusion and further studies

This study establishes a comprehensive unification of estimators utilizing auxiliary information in survey sampling under the general class of estimators proposed by Srivastava (1971) and the wider class developed by Singh and Upadhyaya (1986). The analytical results demonstrate that many ratio-type, product-type, exponential-type, logarithmic-type, and regression-type estimators proposed in the literature are special cases of these foundational classes. In particular, it has been shown that the minimum mean squared error attainable under the Srivastava class coincides with that of the classical linear regression estimator. The wider class of Singh and Upadhyaya (1986) provides additional flexibility and includes regression-type estimators and other modified estimators as special cases.

The exact finite population enumeration and large-scale simulation studies presented in this paper provide strong empirical support for the theoretical findings. These studies confirm that while several alternative estimators may perform better than the simple sample mean estimator, the classical linear regression estimator and estimators belonging to the Srivastava class remain highly efficient, stable, and reliable under realistic conditions. Furthermore, the analysis of recently proposed estimators highlights the importance of computational stability and the need for continuity-based patching in certain cases to ensure proper implementation in finite populations.

After spending considerable time and effort on this article, we arrived at the important conclusion that researchers should use estimators that belong to the class of estimators proposed by Srivastava (1971), because those estimators are consistent estimators. The purpose of publishing this paper is to emphasize that researchers should not overlook the Srivastava (1971) class of estimators, which provides a general and theoretically sound framework for efficient estimation using auxiliary information. Likewise, several other estimators proposed in the literature, including Singh (1986), Biradar and Singh (1997-1998), Singh and Upadhyaya (1986), Singh, Upadhyaya and Iachan (1990), Singh, Gangele and Biradar (1994-1995), and Upadhyaya, Singh and Vos (1985), are also special cases of the Srivastava (1971) class of estimators. In short, it can be concluded that most of the estimators published in various journals during the last several decades are special cases of the Srivastava (1971) class of estimators.

Overall, the unified framework presented in this study provides both theoretical clarity and practical guidance for researchers and practitioners in survey sampling. The results reinforce the central importance of the Srivastava and Singh–Upadhyaya classes as the underlying foundation for auxiliary information-based estimation and highlight their continuing relevance for modern statistical applications. Future research may build upon this framework to develop efficient estimators that maintain optimal theoretical properties while ensuring computational stability and practical applicability. It is also worth pointing out that the recent work of Swain, Dihidar and Mitra (2026) may also be adjusted by the use of patching since their method fails for any pair for which $x_i = x_j$ for $i \neq j$ in a sample, which has quite high chance of happening when sample size is large from a finite population. The idea of patching may allow researchers to consider it since it is same concept as that of removing discontinuity in a function.

ACKNOWLEDGEMENTS

The authors sincerely thank the Guest Editor, Dr. Amrit Kumar Paul, the Administrator, Mr. Satish Kumar, Varma and the reviewer for their valuable comments, which greatly improved the manuscript and helped shape it into its present form.

REFERENCES

Bahl, S., and Tuteja, R.K. (1991). Ratio and product type exponential estimators. *Journal of Information and Optimization Sciences*, **12**(1), 159-163.

Biradar, R.S. and Singh, H.P. (1997-1998). A class of estimators for population parameter using supplementary information. *Aligarh Journal of Statistics*, **17**&**18**, 54-71.

Cochran, W.G. (1940). Some properties of estimators based on sampling scheme with varying probabilities. *J. Agric. Science*, **30**, 262-275.

Deville, J.C. and Särndal, C.E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association*, **87**, 376-382.

Hansen, M.H., Hurwitz, W.N., and Madow, W.G. (1953). Sample Survey Methods and Theory, Volume I&II: Methods and Applications. New York: John Wiley & Sons.

Gupta, S. and Shabbir, J. (2008). On improvement in estimating the population mean in simple random sampling. *Journal of Applied Statistics*, **35**(5), 559-566.

Kadilar, C. and Cingi, H. (2004). Ratio estimators in simple random sampling, Appl. Math. Comput., **151**(3), 893-902.

Kadilar, C. and Cingi, H. (2006a). Improvement in estimating the population mean in simple random sampling. *Appl. Math.Lett.,* **19**(1), 75-79.

Kadilar, C. and Cingi, H. (2006b). Improvement in estimating population mean using correlation coefficient. *Hacet. J. Math. Stat*., **35**(1)(2006), pp. 103-109.

Kadilar, C. and Cingi, H. (2006c). New ratio estimators using correlation coefficient, *Inter Stat.* **3**, 1-11.

Mishra, P., Adichwal, N. K. & Singh, R. (2017). A new log-product-type estimator using auxiliary information. Journal of Scientific Research, **61**, 179.

Murthy, M.N. (1964). Product method of estimation. Sankhya, **26**, 69-74.

Searls, D.T. (1964). The utilization of a known coefficient of variation in the estimation procedure. *Journal of the American Statistical Association*, **59**(308), 1225-1226.

Shakoor, F., Atif, M., Ali, H. *et al.* (2026). Log-ratio type estimation for the finite population mean under simple random sampling without replacement with theory, simulation and application. *Sci Rep*., **16**, (29), https://doi.org/10.1038/s41598-025-29127-7.

Singh, H.P. (1986). A generalized class of estimators of ratio, product and mean utilizing supplementary information on an auxiliary character in PPSWR sampling scheme.\\ Gujarat Statistical Review, **13**(2), 1-30.

Singh, H.P., Gangele, R.K. and Biradar, R.S. (1994-1995). An efficient generalized class of estimators for finite population means using two auxiliary variables. *Gujarat Statistical Review*, (**21**-**22**), 3-16.

Singh, H.P. and Tailor, R. (2003). Use of known correlation coefficient in estimating the finite population mean. *Statistics in Transition*, **6**, 555-560.

Singh, H.P. and Upadhyaya, L.N. (1986). On a class of estimators of the population mean in survey sampling using auxiliary information. *Journal of Indian Society of Agricultural Statistics*, **38**(1), 100-104.

Singh, H.P., Upadhyaya, L.N. and Iachan, R. (1990). An efficient class of estimators using supplementary information in sample surveys. *Aligarh Journal of Statistics*, **10**, 37-50.

Singh, R. and Tiwari, S. N. (2025). Improved Estimator for Population Mean Utilizing Known Medians of Two Auxiliary Variables under Neutrosophic Framework. Neutrosophic syst. appl. **25**, 38-52.

Singh, S. (2006). Survey statisticians celebrate golden jubilee year 2003 of the linear regression estimator. *Metrika*, **63**(1), 1-18

Sisodia, B.V.S. and Dwivedi, V.K. (1981). A modified ratio estimator using co-efficient of variation of auxiliary variable. *Journal of the Indian Society of Agricultural Statistics*, **33**, 13-18.

Srivastava, S.K. (1971). A generalized estimator for the mean of a finite population using multi-auxiliary information. *J. Amer. Statist. Assoc.,* **66**, 404-407.

Srivastava, S.K. (1980). A class of estimators using auxiliary information in sample surveys. *Canad. J. Statist.,* **8**(2), 253-254.

Stearns, M. and Singh, S. (2008). On the estimation of the general parameter. *Computational Statistics and Data Analysis.,* **52**(9), 4253-4271.

Subramani, J. and Kumarapandiyan, G. (2012). Estimation of population mean using coefficient of variation and median of an auxiliary variable. *International Journal of Probability and Statistics*, **1**(4), 111-118.

Swain, A.K.P.C., Dihidar, K. and Mitra, A. (2026). On estimating the finite population variance by utilizing auxiliary information. *Communications in Statistics-Theory and Methods* (Online available)

Upadhyaya, L.N., Singh, H.P. and Vos, J.W.E. (1985). On the estimation of means and ratios using supplementary information. *Statistica Neerlandica*, **39**(3), 309-318.

Yan, Z. and Tian, B. (2010). Ratio method to the mean estimation using co-efficient of skewness of auxiliary variable. *Information Computing and Applications*, II, 103-110.

Zaman, T., Bulut, H. and Yadav, S.K. (2022). Robust ratio-type estimators for finite population mean in simple random sampling: A simulation study. *Concurrency and Computation: Practice and Experience*, **34**(25), e7273.