Jordan Journal of Mathematics and Statistics.
Yarmouk University

DOI:https://doi.org/10.47013/18.4.1

A Novel Logarithmic-Exponential Cum Ratio-Type Estimator Under Simple Random Sampling

SM Afsar Basha [1], *Mahamood Usman* [2,*]

Department of Mathematics (School of Advanced Sciences), Vellore Institute of Technology, Vellore 632014, India

Received: April 1, 2024 Accepted: Oct. 9, 2025.

Abstract: In sample surveys, the use of auxiliary variables to estimate the population mean has become crucial for improving the efficiency of the estimators, including traditional ratio, product and regression estimators. This paper introduces a new logarithmic-exponential cum ratio-type estimator for the elevated estimation of population mean under simple random sampling. We have obtained the bias and mean squared error (MSE) of the proposed estimator up to the first order of approximation and identified the situations in which it performs more efficiently than existing estimators. To verify the theoretical results, we have conducted numerical study based on eight real data sets belonging from the clinical, agricultural and business fields. Their performances have also been evaluated through simulation study that utilized two artificially generated datasets. A sensitivity analysis based on the sample estimates has been investigated to reassert the behaviours of proposed estimators.

Keywords: Study variable; Auxiliary variable; Population mean; Mean squared error; Percent Relative Efficiency.

2010 Mathematics Subject Classification. 62D05.

1 Introduction

In sample surveys, the primary objective is to make reliable population inferences while minimizing costs and time. The efficiency of estimators can be substantially improved by incorporating an auxiliary variable that is highly correlated with the study variable. The strategic use of auxiliary variables is thus a critical practice that allows researchers to draw stronger conclusions and make more informed decisions based on survey data. For example, in agricultural surveys, this approach helps produce accurate crop yield estimates, while in economic studies, it can be used to analyze fixed capital and income indicators. Cochran (1991) propounded ratio estimator using auxiliary variable, which performs better in case of highly positive correlation between study and auxiliary variable. Numerous research papers have been published on estimators based on various transformations, including product-type, exponential ratio-type, regression-type, logarithmic type and logarithmic ratio and product-type estimators. The primary objective of this paper is to develop a population mean estimator that surpasses the accuracy of the ratio estimator of Cochran (1977), the product estimator of Murthy (1967), the regression estimator of Hansen et al. (1953), the modified regression estimator of Grover and Kaur (2011), the exponential estimation methods proposed by Bahl and Tuteja (1991), the modified exponential type estimator of Kadilar (2016), the logarithmic type of Bhushan et al. (2015), the logarithmic ratio and product estimators of Brar et al. (2020) and the logarithmic ratio-type estimator of Adejumobi et al. (2023). To estimate the mean finite population under simple random sampling, this study has developed an effective logarithmic-exponential cum ratio-type estimator. It is anticipated that the proposed estimators will outperform the comparable estimators currently discussed in the literature, providing more accurate estimates of the population mean.
Let $U = \{U_1, U_2, \ldots, U_N\}$ denote a finite population of size N. Let Y represent the study variable and X the auxiliary variable. A sample of size n is drawn from the population using simple random sampling without replacement (SRSWOR). The corresponding paired observations of the study and auxiliary variables in the sample are denoted by (y_i, x_i), for

* Corresponding author e-mail: mahmoodu33@gmail.com

$i = 1, 2, \ldots, n$.
Let the sample mean of the study variable y and the auxiliary variable x be defined as

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \quad \text{and} \quad \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i,$$

which represent the mean of the respective variables based on a sample of size n. Similarly, the population mean of the study variable Y and the auxiliary variable X are given by

$$\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i \quad \text{and} \quad \bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i,$$

where N is the population size.
The sample variance for y and x are expressed as

$$s_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2 \quad \text{and} \quad s_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2,$$

while the corresponding population variance are given by

$$S_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})^2 \quad \text{and} \quad S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \bar{X})^2.$$

Furthermore, let $\rho_{X,Y}$ denote the Pearson correlation coefficient between the population variables X and Y. The coefficient of variation for Y and X are defined as

$$C_y = \frac{S_y}{\bar{Y}} \quad \text{and} \quad C_x = \frac{S_x}{\bar{X}},$$

respectively.
To compute the bias and MSE for the proposed estimator as well as the existing estimators under discussion. The following sample error terms are introduced:

$$\bar{y} = \bar{Y}(1 + e_0) \quad \text{and} \quad \bar{x} = \bar{X}(1 + e_1), \quad \text{such that } E(e_i) = 0 \text{ for } i = 0, 1.$$

Additionally,

$$E\left(e_0^2\right) = \phi C_y^2, \quad E\left(e_1^2\right) = \phi C_x^2, \quad \text{and} \quad E(e_1 e_0) = \phi\,\rho\, C_y C_x,$$

where

$$\phi = \left(\frac{1-f}{n}\right); \quad f = \frac{n}{N}.$$

In this paper, we present a highly effective logarithmic-exponential cum ratio-type estimator for estimating the finite population mean. The proposed estimator is shown to be more efficient than all other existing estimators.

2 The Existing estimators

This section discusses several existing estimators for the estimation of population mean.
The conventional sample mean is described as:

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{1}$$

The estimator $\bar{y}$ whose variance is expressed as :

$$\mathrm{Var}(\bar{y}) = \phi \bar{Y}^2 C_y^2 \tag{2}$$

Cochran (1977) introduced an ratio type estimator defined as:

$$T_R = \bar{y}\left(\frac{\bar{X}}{\bar{x}}\right) \tag{3}$$

The approximate expression for the MSE of T_R at the first order is:

$$\text{MSE}(T_R) = \phi \bar{Y}^2 \left(C_y^2 + C_x^2 - 2\rho C_y C_x\right) \tag{4}$$

Murthy (1967) introduced the product estimator for $\bar{Y}$, specifically in situations characterized by a strong negative correlation between Y and X.

$$T_P = \bar{y}\left(\frac{\bar{x}}{\bar{X}}\right) \tag{5}$$

The approximate expression for the MSE of T_P at the first order is:

$$\text{MSE}(T_P) = \phi \bar{Y}^2 \left(C_y^2 + C_x^2 + 2\rho C_y C_x\right) \tag{6}$$

Hansen et al. (1953) introduced the regression estimator for the population mean $\bar{Y}$ as

$$T_{RE} = \bar{y} + \beta\left(\bar{X} - \bar{x}\right) \tag{7}$$

The approximate expression for the MSE of T_{RE} at the first order is:

$$\text{MSE}(T_{RE}) = \phi \bar{Y}^2 \left[C_y^2 - \frac{2}{\bar{Y}}\beta\rho\, C_y C_x + \frac{1}{\bar{Y}^2}\beta^2 C_x^2\right]$$

where $\beta^{opt} = \bar{Y}\,\rho\,\frac{C_y}{C_x}$, therefore

$$min.\,\text{MSE}(T_{RE}) = \phi \bar{Y}^2 C_y^2 \left(1 - \rho^2\right) \tag{8}$$

Grover and kaur (2011) suggested an modified exponential-type estimator of $\bar{Y}$:

$$T_G = \left[k_1\bar{y} + k_2\left(\bar{X} - \bar{x}\right)\right]\exp\left(\frac{\bar{X} - \bar{x}}{\bar{X} + \bar{x}}\right) \tag{9}$$

The approximate expression for the MSE of T_G at the first order is:

$$\begin{aligned}\text{MSE}(T_G) = \bar{Y}^2\left[(k_1 - 1)^2 + k_1^2\phi L_1 - \frac{\phi k_1}{2}\left\{L_2 + \left(\frac{C_x^2}{2} - \rho C_y C_x\right)\right\}\right] + k_2^2\bar{X}^2\phi C_x^2 \\ + 2k_2\bar{X}\bar{Y}\phi\left(k_1 L_2 - \frac{C_x^2}{2}\right)\end{aligned} \tag{10}$$

The optimal values for k_1 and k_2 are as follows:

$$k_{1(opt)} = \frac{-C_x^2\left[2 - \frac{\phi}{2}L_2 + \frac{\phi}{2}\left(\frac{C_x^2}{2} - \rho C_y C_x\right)\right]}{2\left[\phi L_2^2 - (1 + \phi L_1)\,C_x^2\right]}$$

and

$$k_{2(opt)} = \frac{\bar{Y}\left[L_2\left\{2 + \frac{\phi}{2}L_2 + \frac{\phi}{2}\left(\frac{C_x^2}{2} - \rho C_y C_x\right)\right\} - C_x^2\left(1 + \phi L_1\right)\right]}{2\bar{X}\left[\phi L_2^2 - (1 + \phi L_1)\,C_x^2\right]}$$

where $L_1 = C_y^2 + C_x^2 - 2\rho C_y C_x$ and $L_2 = C_x^2 - \rho C_y C_x$.

$$min.\,\text{MSE}(T_G) = \frac{\phi\bar{Y}^2 C_y^2\left(1 - \rho^2\right)}{1 + \phi C_y^2\left(1 - \rho^2\right)} - \frac{\phi^2\bar{Y}^2 C_x^2\left[4C_y^2\left(1 - \rho^2\right) + \frac{C_x^2}{4}\right]}{16\left[1 + \phi C_y^2\left(1 - \rho^2\right)\right]}$$

Or

$$min.\,\text{MSE}(T_G) = min.\,\text{MSE}(T_{RE}) - M_1 - M_2 \tag{11}$$

where

$$M_1 = \frac{\frac{[min.\,\mathrm{MSE}(T_{RE})]^2}{\bar{Y}^2}}{1+\frac{min.\,\mathrm{MSE}(T_{RE})}{\bar{Y}^2}}$$

$$M_2 = \frac{\phi C_x^2\left[min.\,\mathrm{MSE}(T_{RE})+\frac{\phi\bar{Y}^2C_x^2}{16}\right]}{4\left[1+\frac{min.\,\mathrm{MSE}(T_{RE})}{\bar{Y}^2}\right]}$$

Bahl and Tuteja (1991) proposed estimators for $\bar{Y}$ based on the ratio and product type, are as follows:

$$T_{ER} = \bar{y}\exp\left[\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right] \tag{12}$$

and

$$T_{EP} = \bar{y}\exp\left[\frac{\bar{x}-\bar{X}}{\bar{x}+\bar{X}}\right] \tag{13}$$

The approximate expression for the MSE of T_{ER} and T_{EP} at the first order is:

$$\mathrm{MSE}(T_{ER}) = \phi\bar{Y}^2\left[C_y^2+\frac{C_x^2}{4}-\rho C_yC_x\right] \tag{14}$$

and

$$\mathrm{MSE}(T_{EP}) = \phi\bar{Y}^2\left[C_y^2+\frac{C_x^2}{4}+\rho C_yC_x\right] \tag{15}$$

Kadilar (2016) Suggested a newly exponential estimator is expressed as:

$$T_K = \bar{y}\left(\frac{\bar{x}}{\bar{X}}\right)^{\alpha}\exp\left[\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right] \tag{16}$$

The approximate expression for the MSE of T_K at the first order is:

$$\mathrm{MSE}\,(T_K) = \phi\bar{Y}^2\left(C_y^2+C_x^2/4+2\alpha\rho C_xC_y+\rho C_xC_y+\alpha^2C_x^2+\alpha C_x^2\right)$$

the optimum value of α is given by $\alpha^{\mathrm{opt}} = \frac{(C_x-2\rho C_y)}{2C_x}$.

$$min.\,\mathrm{MSE}(T_K) = \phi\bar{Y}^2C_y^2\left(1-\rho^2\right) \tag{17}$$

Bhushan et al. (2016) proposed a class of logarithmic-type estimator in the form of :

$$T_B = \bar{y}\left\{1+\log\left(\frac{\bar{x}}{\bar{X}}\right)\right\}^{\alpha} \tag{18}$$

where α is a chosen scalar.
The approximate expression for the MSE of T_B at the first order is:

$$\mathrm{MSE}(T_B) = \phi\bar{Y}^2(C_y^2+\alpha^2C_x^2+2\alpha\rho_{xy}C_xC_y)$$

The minimum MSE at the optimal value of $\alpha_{\mathrm{opt}} = -\rho_{xy}\left(\frac{C_y}{C_x}\right)$ is given by

$$min.\,\mathrm{MSE}(T_B) = \phi\bar{Y}^2C_y^2(1-\rho_{xy}^2) \tag{19}$$

Brar et al. (2020) proposed ratio and product estimators of the population mean was given new functional representations as follows:

$$T_{BR} = \bar{y}\left(\frac{\bar{X}}{\bar{x}-\bar{X}}\right)\ln\left(\frac{\bar{x}}{\bar{X}}\right) \tag{20}$$

$$T_{BP} = \bar{y}\left(\frac{\bar{X}}{\bar{x}-\bar{X}}\right)\ln\left(\frac{2\bar{X}-\bar{x}}{\bar{X}}\right) \tag{21}$$

The approximate expression for the MSEs of T_{BR} and T_{BP} respectively, expressed as:

$$\text{MSE}(T_{BR}) = \phi\bar{Y}^2\left[C_y^2 + \frac{C_x^2}{4} - \rho C_y C_x\right] \tag{22}$$

$$\text{MSE}(T_{BP}) = \phi\bar{Y}^2\left[C_y^2 + \frac{C_x^2}{4} + \rho C_y C_x\right] \tag{23}$$

Adejumobi et al. (2023) suggested a new estimator based on logarithmic ratios for the estimation of the finite population mean as:

$$T_A = \frac{\left[\bar{y} - Ln\left(\frac{\bar{x}}{\bar{X}}\right)\right]}{\bar{x}}\bar{X} \tag{24}$$

The approximate expression for the MSE of T_A at the first order is:

$$\text{MSE}(T_A) = \phi\left[\bar{Y}^2\left(C_y^2 + C_x^2 - 2\rho C_y C_x\right) + (1+2\bar{Y})C_x^2 - 2\bar{Y}\rho C_y C_x\right] \tag{25}$$

3 Proposed Estimator

Inspired by the above researchers, we have proposed a new class of logarithmic-exponential cum ratio-type Estimator for the population mean $\bar{Y}$, using the information on a single auxiliary variable x with a known population mean $\bar{X}$ which is described as :

$$T^*_{PR} = \bar{y}\left\{\delta\left(\frac{\bar{X}}{\bar{x}}\right)^{\alpha}\exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right) + \beta\left(\frac{\bar{X}}{\bar{x}}\right)\log\left(\frac{\bar{x}}{\bar{X}}\right)\right\}\left\{\exp\left(\gamma\frac{(a\bar{X}+b)-(a\bar{x}+b)}{(a\bar{X}+b)+(a\bar{x}+b)}\right)\right\} \tag{26}$$

In this situation, δ and β stand for constants whose optimal value will be determined later, while α and γ is a fixed constant that specifically takes on the values 1, -1, or 0, a and b is a chosen constant it may be correlation coefficient ρ_{yx}, coefficient of variation (C_y or C_x).
To derive the MSE of T^*_{PR},
let's express $\bar{y} = \bar{Y}(1+e_0)$ and $\bar{x} = \bar{X}(1+e_1)$, where $E(e_0) = E(e_1) = 0$ and $E(e_0^2) = \phi C_y^2$, $E(e_1^2) = \phi C_x^2$, $E(e_0 e_1) = \phi\rho C_y C_x$.
Reframing equation (26) in terms of e's, we have:

$$\begin{aligned} T^*_{PR} &= \bar{Y}\left\{\delta(1+e_0)(1+e_1)^{-\alpha}\exp\left(\frac{\bar{X}-\bar{X}(1+e_1)}{\bar{X}+\bar{X}(1+e_1)}\right) + \beta(1+e_0)\left(\frac{\bar{X}}{\bar{X}(1+e_1)}\right)\right. \\ &\left.\times\log\left(\frac{\bar{X}(1+e_1)}{\bar{X}}\right)\right\} \times \left\{\exp\left(\gamma\frac{a\bar{X}-a\bar{X}(1+e_1)}{a\bar{X}+2b+a\bar{X}(1+e_1)}\right)\right\} \\ &= \bar{Y}\left\{\delta\left(1-\alpha e_1 + \frac{(\alpha(\alpha+1))}{2}e_1^2\right)\exp\left(-\frac{e_1}{2}+\frac{e_1^2}{4}\right) + \beta(1+e_0)(1+e_1)^{-1}\right. \\ &\left.\times\log(1+e_1)\right\} \times \left\{\exp\left(-\gamma\theta e_1 + \gamma\theta^2 e_1^2\right)\right\} \end{aligned} \tag{27}$$

Expanding the expression on the right side of (27) and keeping terms up to the second power of e's, we obtain:

$$\begin{aligned} &= \bar{Y}\left\{\delta\left(1-\alpha e_1+\frac{\alpha(\alpha+1)}{2}e_1^2\right)\left(1-\frac{e_1}{2}+\frac{3e_1^2}{2}\right)+\beta\left(e_1+e_0e_1-\frac{3}{2}e_1^2\right)\right\} \\ &\times\left\{1-\gamma\theta e_1+\left(\gamma+\frac{\gamma^2}{2}\right)\theta^2e_1^2\right\} \end{aligned} \tag{28}$$

where in (28),

$$\theta = \frac{a\bar{X}}{2(a\bar{X}+b)}$$

$$\begin{aligned} T^*_{PR} = \bar{Y}&\left\{\delta-\left(\delta\alpha+\frac{\delta}{2}-\beta\right)e_1+\left(\frac{\alpha^2\delta}{2}+\delta\alpha+\frac{3\delta}{2}-\frac{3\beta}{2}\right)e_1^2\right. \\ &\left.+\delta e_0-\left(\delta\alpha+\frac{\delta}{2}-\beta\right)e_0e_1\right\}\times\left\{1-\gamma\theta e_1+\left(\gamma+\frac{\gamma^2}{2}\right)\theta^2e_1^2\right\} \end{aligned} \tag{29}$$

To obtain the bias of T_{PR}^{*}, $B(T_{PR}^{*}) = E(T_{PR}^{*}) - \bar{Y}$, we get

$$B(T_{PR}^{*}) = \bar{Y}\left[\delta - \phi E_1 \rho C_x C_y + \phi E_2 C_x^2 - 1\right] \tag{30}$$

where

$$E_1 = \left(\delta\alpha + \frac{\delta}{2} - \beta + \delta\gamma\theta\right)$$
$$E_2 = \left(\frac{\alpha^2\delta}{2} + \delta\alpha + \frac{3\delta}{2} - \frac{3\beta}{2} + \gamma\delta\alpha\theta + \frac{\delta\gamma\theta}{2} - \gamma\theta\beta + \theta^2\delta\gamma + \frac{\gamma^2\delta\theta^2}{2}\right)$$

MSE(T_{PR}^{*}) is obtained by $E\left[T_{PR}^{*} - \bar{Y}\right]^2 is$

$$\text{MSE}(T_{PR}^{*}) = \bar{Y}^2\left\{\delta^2 - 2\delta + 1 + \phi\left(2\mu_2\delta - 2\mu_2 + \mu_1^2\right)C_x^2 + \phi\delta^2 C_y^2 - \phi\left(4\delta\mu_1 - 2\mu_1\right)\rho C_x C_y\right\} \tag{31}$$

Or

$$\text{MSE}(T_{PR}^{*}) = \bar{Y}^2\left\{D + \phi A C_x^2 + \phi\delta^2 C_y^2 - \phi B\rho C_x C_y\right\} \tag{32}$$

where

$$A = \left(2\mu_2\delta - 2\mu_2 + \mu_1^2\right)$$
$$B = \left(4\delta\mu_1 - 2\mu_1\right)$$
$$D = \delta^2 - 2\delta + 1$$
$$\mu_1 = \left(\delta\alpha + \frac{\delta}{2} - \beta + \delta\gamma\theta\right)$$
$$\mu_2 = \left(\frac{\alpha^2\delta}{2} + \delta\alpha + \frac{3\delta}{2} - \frac{3\beta}{2} + \gamma\delta\alpha\theta + \frac{\delta\gamma\theta}{2} - \gamma\theta\beta + \theta^2\delta\gamma + \frac{\gamma^2\delta\theta^2}{2}\right)$$

To find the optimal values for δ and β to minimize MSE(T_{PR}^{*}). By differentiating MSE(T_{PR}^{*}) with respect to δ and β, setting the derivatives to zero, the optimal values of δ and β can be determined.

$$\beta_{\text{opt}} = \frac{(A_1 + A_2 - A_3 + A_4)}{2(A_5 - A_6)} = \beta^{*}(\text{say}) \tag{33}$$

$$\delta_{\text{opt}} = \frac{B_1 + B_2}{B_3 + B_4} = \delta^{*}(\text{say}) \tag{34}$$

Where

$$A_1 = \left[\left(4\alpha^2 - 8\alpha - 8\right)\alpha - \left(12\alpha\gamma^2 - 8\alpha + 8\gamma^2\theta + 24\gamma + 8\right)\gamma\theta^2 - (36\alpha + 18)\gamma\theta - 15\right]C_x^2$$
$$A_2 = \left[(24\alpha + 48)\gamma\theta + 36\alpha + 24\theta^2\gamma^2 + 18\right]\rho C_x C_y$$
$$A_3 = \left[16\alpha\rho^2 + 16\gamma\theta e^2 + 8e^2 + 8\gamma\theta - \frac{1}{\phi}8\rho\frac{C_y}{C_x} + 12\right]C_y^2$$
$$A_4 = \frac{1}{\phi}\left[8\alpha^2 + 8\gamma\theta + 4 - \rho\frac{C_y}{C_x}\right]$$
$$A_5 = \left[\left(4\alpha^2 - 4\alpha - 8\gamma^2\theta^2 + 8\gamma\theta^2 - 24\gamma\theta - 3\right)C_x^2\right] + \left[(16\gamma\theta + 24)\right]\rho C_x C_y$$
$$A_6 = \left[\left(16\rho^2 + 4\right)C_y^2 - \frac{1}{\phi}4\right]$$

$$B_1 = \left(2\alpha^2 - 2\alpha - 6\gamma^2\theta^2 + 4\gamma\theta^2 - 18\gamma\theta - 6\right)C_x^2$$
$$B_2 = (12\gamma\theta + 18)\rho C_y C_x - 8\rho^2 C_y^2 + \frac{1}{\phi}4$$
$$B_3 = \left(4\alpha^2 - 4\alpha - 8\gamma^2\theta^2 + 8\gamma\theta^2 - 24\gamma\theta - 3\right)C_x^2$$
$$B_4 = (16\gamma\theta + 24)\rho C_y C_x + \left(4 - 16\rho^2\right)C_y^2 + \frac{1}{\phi}4$$

Upon substituting the optimal values of β_{opt} and δ_{opt} into (32), the resulting minimum value of MSE(T^*_{PR}) is:

$$\min.\text{MSE}(T^*_{PR}) = \bar{Y}^2\left\{D^* + \phi A^* C_x^2 + \phi(\delta^*)^2 C_y^2 \; -\phi B^* \rho C_x C_y\right\} \tag{35}$$

where

$$A^* = \left(2\mu_2^*\delta^* - 2\mu_2^* + (\mu_1^*)^2\right)$$
$$B^* = (4\delta^*\mu_1^* - 2\mu_1^*)$$
$$D^* = (\delta^*)^2 - 2\delta^* + 1$$
$$\mu_1^* = \left(\alpha + \frac{1}{2} - \beta^* + \gamma\theta\right)\delta^*$$
$$\mu_2^* = \left(\frac{\alpha^2}{2} + \alpha + \frac{3}{2} + \theta^2\gamma + \frac{\gamma^2\theta^2}{2} + \gamma\alpha\theta + \frac{\gamma\theta}{2}\right)\delta^* - \left(\frac{3}{2} + \gamma\theta\right)\beta^*$$

Some special propertices of proposed estimator:

(a)In cases of positive correlation between populations, setting the values $[\alpha,\gamma] = [1,1]$ is suitable, provided that appropriate values are assigned to a and b in the proposed class of estimators T^*_{PR} to attain the maximum efficiency gain over the estimator T_G. Hence, the suggested class of estimators for positively correlated populations is expressed as:

$$T^*_{PR(1)} = \bar{y}\left\{\delta\left(\frac{\bar{X}}{\bar{x}}\right)\exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right) + \beta\left(\frac{\bar{X}}{\bar{x}}\right)\log\left(\frac{\bar{x}}{\bar{X}}\right)\right\}\left\{\exp\left(\frac{(a\bar{X}+b)-(a\bar{x}+b)}{(a\bar{X}+b)+(a\bar{x}+b)}\right)\right\} \tag{36}$$

Note : If we put $[\alpha,\gamma] = [1,0]$ in (26), we obtain the result.

$$T^*_{PR(2)} = \bar{y}\left\{\delta\left(\frac{\bar{X}}{\bar{x}}\right)\exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right) + \beta\left(\frac{\bar{X}}{\bar{x}}\right)\log\left(\frac{\bar{x}}{\bar{X}}\right)\right\} \tag{37}$$

Therefore,the suitable selections are $[\alpha,\gamma] = [1,1],[1,0]$ for positively correlated populations, the estimators $T^*_{PR(1)}$ and $T^*_{PR(2)}$ perform more effectively than T_G.

(b)In cases of negative correlation between populations, choosing the value $(\alpha,\gamma) = (-1,1)$ is recommended, with the adjustment of suitable values for a and b in the proposed class of estimators T^*_{PR} to improve accuracy over the estimator T_G. Consequently, the applicable class of estimators is defined as:

$$T^*_{PR(3)} = \bar{y}\left\{\delta\left(\frac{\bar{x}}{\bar{X}}\right)\exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right) + \beta\left(\frac{\bar{X}}{\bar{x}}\right)\log\left(\frac{\bar{x}}{\bar{X}}\right)\right\}\left\{\exp\left(\frac{(a\bar{X}+b)-(a\bar{x}+b)}{(a\bar{X}+b)+(a\bar{x}+b)}\right)\right\} \tag{38}$$

Note : If we put $[\alpha,\gamma] = [-1,0]$ in (26), we obtain

$$T^*_{PR(4)} = \bar{y}\left\{\delta\left(\frac{\bar{x}}{\bar{X}}\right)\exp\left(\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right) + \beta\left(\frac{\bar{X}}{\bar{x}}\right)\log\left(\frac{\bar{x}}{\bar{X}}\right)\right\} \tag{39}$$

Therefore, the suitable selections are $[\alpha,\gamma] = [-1,1],[-1,0]$ for negatively correlated data, the estimators $T^*_{PR(3)}$ and $T^*_{PR(4)}$ perform more effectively than T_G with the appropriate values of a and b to acquire suitable estimators.

4 Efficiency Comparisons

In this section, criteria for the effectiveness of the newly proposed estimators compared to certain existing estimators are outlined. The MSEs of commonly used estimators such as $\bar{y}, T_R, T_P, T_{RE}, T_G, T_{ER}, T_{EP}, T_k, T_B, T_{BR}, T_{BP}$ and T_A are contrasted with the MSE of proposed estimator T^*_{PR}.
T^*_{PR} is more effective than $\bar{y}$ when:

$$Var(\bar{y}) - min.\,\text{MSE}\,(T^*_{PR}) > 0$$

$$\phi\bar{Y}^2\left(A^*C_x^2 + \left((\delta^*)^2 - 1\right)C_y^2 - B^*\rho C_xC_y\right) + \bar{Y}^2D^* > 0 \tag{40}$$

T^*_{PR} is more effective than T_R when:

$$\text{MSE}\,(T_R) - min.\,\text{MSE}\,(T^*_{PR}) > 0$$

$$\phi\bar{Y}^2\left[(A^* - 1)C_x^2 + (2 - B^*)\rho C_xC_y + \left((\delta^*)^2 - 1\right)C_y^2\right] + \bar{Y}^2D^* > 0 \tag{41}$$

T^*_{PR} is more effective than T_P when:

$$\text{MSE}\,(T_P) - min.\,\text{MSE}\,(T^*_{PR}) > 0$$

$$\phi\bar{Y}^2\left[(A^* - 1)C_x^2 - (2 + B^*)\rho C_xC_y + \left((\delta^*)^2 - 1\right)C_y^2\right] + \bar{Y}^2D^* > 0 \tag{42}$$

T^*_{PR} is more effective than T_{RE} when:

$$min.\,\text{MSE}\,(T_{RE}) - min.\,\text{MSE}\,(T^*_{PR}) > 0$$

$$\phi\bar{Y}^2\left[A^*C_x^2 + \left((\delta^*)^2 + \rho^2 - 1\right)C_y^2 - B^*\rho C_xC_y\right] + \bar{Y}^2D^* > 0 \tag{43}$$

T^*_{PR} is more effective than T_G when:

$$\text{MSE}(T_G) - min.\,\text{MSE}(T^*_{PR}) > 0$$

$$\phi\bar{Y}^2\left[A^*C_x^2 + \left((\delta^*)^2 + \rho^2 - 1\right)C_y^2 - B^*\rho C_xC_y\right] + \bar{Y}^2D^* + M_1 + M_2 > 0 \tag{44}$$

T^*_{PR} is more effective than T_{ER} when:

$$\text{MSE}\,(T_{ER}) - min.\,\text{MSE}\,(T^*_{PR}) > 0$$

$$\phi\bar{Y}^2\left[\left(A^* - \frac{1}{4}\right)C_x^2 + \left((\delta^*)^2 - 1\right)C_y^2 + (1 - B^*)\rho C_xC_y\right] + \bar{y}^2D^* > 0 \tag{45}$$

T^*_{PR} is more effective than T_{EP} when:

$$\text{MSE}\,(T_{EP}) - min.\,\text{MSE}\,(T^*_{PR}) > 0$$

$$\phi\bar{Y}^2\left[\left(A^* - \frac{1}{4}\right)C_x^2 + \left((\delta^*)^2 - 1\right)C_y^2 - (1 + B^*)\rho C_xC_y\right] + \bar{y}^2D^* > 0 \tag{46}$$

T^*_{PR} is more effective than T_K when:

$$min.\,\text{MSE}(T_K) - min.\,\text{MSE}(T^*_{PR}) > 0$$

$$\phi\bar{Y}^2\left[A^*C_x^2+\left((\delta^*)^2+\rho^2-1\right)C_y^2-B^*\rho C_xC_y\right]+\bar{Y}^2D^*>0 \quad (47)$$

T_{PR}^* is more effective than T_B when:

$$min.\,\text{MSE}(T_B)-min.\,\text{MSE}(T_{PR}^*)>0$$

$$\phi\bar{Y}^2\left[A^*C_x^2+\left((\delta^*)^2+\rho^2-1\right)C_y^2-B^*\rho C_xC_y\right]+\bar{Y}^2D^*>0 \quad (48)$$

T_{PR}^* is more effective than T_{BR} when:

$$\text{MSE}(T_{BR})-min.\,\text{MSE}(T_{PR}^*)>0$$

$$\phi\bar{Y}^2\left[\left(A^*-\frac{1}{4}\right)C_x^2+\left((\delta^*)^2-1\right)C_y^2+(1-B^*)\rho C_xC_y\right]+\bar{Y}^2D^*>0 \quad (49)$$

T_{PR}^* is more effective than T_{BP} when:

$$\text{MSE}(T_{BP})-min.\,\text{MSE}(T_{PR}^*)>0$$

$$\phi\bar{Y}^2\left[\left(A^*-\frac{1}{4}\right)C_x^2+\left((\delta^*)^2-1\right)C_y^2+(1+B^*)\rho C_xC_y\right]+\bar{Y}^2D^*>0 \quad (50)$$

T_{PR}^* is more effective than T_A when:

$$min.\,\text{MSE}(T_A)-min.\,\text{MSE}(T_{PR}^*)>0$$

$$\phi\bar{Y}^2\left[(A^*-1)C_x^2+\left((\delta^*)^2-1\right)C_y^2+(2-B^*)\rho C_xC_y\right]$$
$$+\phi\left(2\bar{Y}\rho C_xC_y-(1-2\bar{Y})C_x^2\right)+\bar{Y}^2D^*>0 \quad (51)$$

It is noted that T_{PR}^* consistently outperforms the existing estimators $\bar{y},T_R,T_P,T_{RE},T_G$, $T_{ER},T_{EP},T_k,T_B,T_{BR},T_{BP}$ and T_A as the conditions are satisfied in equations (40) to (51) are consistently met.

5 Numerical study

The following data sets have been used to confirm the suitability of the suggested estimator T_{PR}. We have collected nine sets of data, identified as Data Set-1, Data Set-2, Data Set-3, Data Set-4, Data Set-5, Data Set-6, Data Set-7 and Data Set-8. Which are as follows:

Data set-1: [Murthy (1967)]: Consider the study variable Y as the fixed capital, and the auxiliary variable X represents the output of the 80 factories. The parameters characterizing this dataset are: $N=80,n=20,\bar{Y}=11.264,\bar{X}=51.826,\rho=0.941,C_y=0.750,C_x=0.354$

Data set-2: [Shabbir et al. (2014)]: Consider the study variable Y as the level of apple production (in 1000 tons), and let the auxiliary variable X represent the number of apple trees in 104 villages in 1999. The parameters characterizing this data are as follows: $N=104,n=20,\bar{Y}=6.254,\bar{X}=13931.680,\rho=0.860,C_y=1.860,C_x=1.650$

Data set-3: [US Environmental Protection Agency 1991]: Let the study variable Y be the average miles per gallon and X be the engine horsepower (as an auxiliary variate). The parameters that describe this data are as follows: $N=80,n=5,\bar{Y}=33.5175,\bar{X}=118.1625,\rho=-0.801,C_y=0.2877,C_x=0.4804$

Data set-4: [Murthy (1967)]: The number of workers is denoted by the auxiliary variable X, and the study variable Y corresponds to the output of 80 factories in a region. The parameters that describe this data are as follows: $N=34,n=15,\bar{Y}=199.44,\bar{X}=208.88,\rho=0.98,C_y=0.75,C_x=0.72$

Data set-5: [Singh and Chaudhary (1986)]: the parameters that describe these data are as follows: $N=204,n=50,\bar{Y}=26441,\bar{X}=966,\rho=0.71,C_y=2.4739,C_x=1.7171$

Data set-6: [Das (1988)]: Let Y be the number of agricultural labourers in 1971 and X be the number of agricultural labourers in 1971: $N=278,n=25,\bar{Y}=39.068,\bar{X}=25.111,\rho=0.7213,C_y=1.4451,C_x=1.6198$

Data set-7: [Singh and Chaudhary(1986)]: The parameters that describe this data are as follows: $N=34,n=20,\bar{Y}=84.6412,\bar{X}=19.9441,\rho=0.4453,C_y=0.7532,C_x=2$

Data set-8: [Gupta and Kothwala (1990)]: Let Y denote the proportion of the irrigated area, and let X represent the

area under the crop gram and mixture. The parameters that describe this data are as follows: $N = 400, n = 100, \bar{Y} = 36.71438, \bar{X} = 6.56383, \rho = -0.4020, C_y = 0.9928, C_x = 0.9617$
The percentage relative efficiency (PRE) of the estimators can be calculated using the formula

$$\text{PRE}(\Theta, T) = \frac{Var(T)}{\text{MSE}(\Theta)} \times 100 \tag{52}$$

Where, $\Theta = T$ or $\bar{y}$, $T_R, T_P, T_{RE}, T_G, T_{ER}, T_{EP}, T_K, T_B, T_{BR}, T_{BP}, T_A$ and T_{PR}^*. The results are presented in Table 1 and 2 at $\alpha = 1, \gamma = \rho, a = -1$ and $b = 1$ for the proposed estimator. For a clearer and more immediate interpretation of Table 1 and 2, the results have been presented graphically in Figure 1 in terms of PREs.

Interpetation of the results:

It is observed from Table 1 and 2:

(i)The MSE and PRE values of the proposed estimator T_{PR}^* and existing estimators have been evaluated across Data Sets 1 to 10 to examine comparative performance.

(ii)For Data Sets 1 and 2, T_{PR}^* attains the lowest MSE values (0.30532 and 0.54111, respectively) and the highest PRE values (876.552 and 1009.9), clearly outperforming all the existing estimators.

(iii)In Data Set 3, where the correlation between the variables is negative, the proposed estimator still demonstrates remarkable efficiency, with an MSE of 1.91849 and a PRE of 908.791—significantly better than the classical product estimator T_P.

(iv)For Data Sets 4, 5, and 7, the estimator T_{PR}^* consistently maintains superior performance. In Data Set 4, it achieves an MSE of 13.1386 and PRE of 6344.25 compared to the baseline estimator $\bar{y}$ with an MSE of 833.548.

(v)In Data Set 5, despite the presence of high variability in data, the proposed estimator shows excellent performance with an MSE of 25981934 and a PRE of 248.639, outperforming other estimators with notably larger MSEs.

(vi)In Data Set 7, the estimator again demonstrates dominance with an MSE of 11.2414 and PRE of 762.054, outperforming all estimators, including T_{RE}, T_G, T_K, and T_A.

(vii)In Data Set 7, the correlation is again negative, yet the proposed estimator yields the lowest MSE (8.13122) and highest PRE (122.547), surpassing both product-type and generalized estimators in performance.

(viii)Overall, across Data Sets 1 to 10, the proposed estimator T_{PR}^* consistently achieves the lowest MSE and highest PRE, establishing its efficiency and robustness in comparison to the existing methods.

Table 1: MSEs and PREs of proposed estimator and existing estimators for data sets 1 to 5

Estimator	Data 1		Data 2		Data 3		Data 4		Data 5	
	MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE
$\bar{y}$	2.67633	100	5.4646	100	17.4351	100	833.548	100	64601200	100
T_R	0.895178	298.972	1.427	382.94	112.687	15.4721	33.3419	2500	32052095	201.551
T_P	5.64996	47.3689	18.103	30.186	19.4088	89.8308	3170.15	26.2936	159394301	40.5292
T_{RE}	0.30649	873.218	1.423	384.02	6.24872	279.019	33.0085	2525.25	32035735	201.654
T_G	0.305349	876.482	1.3282	411.44	6.11424	285.155	32.5903	2557.66	30270495	213.413
T_{ER}	1.63669	163.521	2.3707	230.5	52.9078	32.9537	241.395	345.304	40546148	159.328
T_{EP}	4.01408	66.6734	10.709	51.03	6.26875	278.127	1809.8	46.0575	104217251	61.9871
T_K	0.30649	873.218	1.423	384.02	6.24872	279.019	33.0085	2525.25	32035735	201.654
T_B	0.30649	873.218	1.423	384.02	6.24872	279.019	33.0085	2525.25	32035735	201.654
T_{BR}	1.63669	163.521	2.3707	230.5	52.9078	32.9537	241.395	345.304	40546148	159.328
T_{BP}	4.014083	66.6734	10.70862	51.030	6.26875	278.1270	1809.79887	46.0575	104217251	61.9871
T_A	0.794683	336.779	1.5789	346.09	117.022	14.89	33.2007	2510.63	32052041	201.551
T_{PR}^*	**0.30532**	**876.552**	**0.54111**	**1009.9**	**1.91849**	**908.791**	**13.1386**	**6344.25**	**25981934**	**248.639**

Table 2: MSEs and PREs of proposed estimator and existing estimators for data sets 6 to 10

Estimator	Data 6		Data 7		Data 8		Data 9		Data 10	
	MSE	**PRE**	**MSE**	**PRE**	**MSE**	**PRE**	**MSE**	**PRE**	**MSE**	**PRE**
$\bar{y}$	116.0310	100.0000	85.6654	100	9.96454	100	0.588426	100	0.00590638	100
T_R	74.1901	156.3967	487.091	17.5871	27.0751	36.8033	0.210335	279.756	0.0280158	21.0823
T_P	449.4337	25.8171	892.261	9.60093	11.554	86.2429	2.95985	19.8803	0.00181148	326.052
T_{RE}	55.6631	208.4522	68.6786	124.734	8.35423	119.275	0.114356	514.554	0.00114171	517.327
T_G	52.2123	222.2292	65.8706	130.051	8.28737	120.238	0.114293	514.84	0.00114114	517.585
T_{ER}	58.6653	197.7846	135.375	63.2798	16.1823	61.5767	0.150214	391.725	0.0147093	40.1542
T_{EP}	246.2871	47.1121	337.961	25.3477	8.42178	118.319	1.52497	38.586	0.00160712	367.513
T_K	55.6631	208.4522	68.6786	124.734	8.35423	119.275	0.114356	514.554	0.00114171	517.327
T_B	55.6631	208.4522	68.6786	124.734	8.35423	119.275	0.114356	514.554	0.00114171	517.327
T_{BR}	58.6653	197.7846	135.375	63.2798	16.1823	61.5767	0.150214	391.725	0.0147093	40.1542
T_{BP}	246.2871	47.1121	337.9606	25.34774	8.42178	118.3186	1.524970	38.5860	0.00160712	367.5133
T_A	498.913	17.1704	498.913	17.1704	27.8027	35.8401	0.232054	253.573	0.0393767	14.9997
T^*_{PR}	**27.7616**	**417.9549**	**11.2414**	**762.054**	**8.13122**	**122.547**	**0.112126**	**524.789**	**0.00114112**	**517.597**

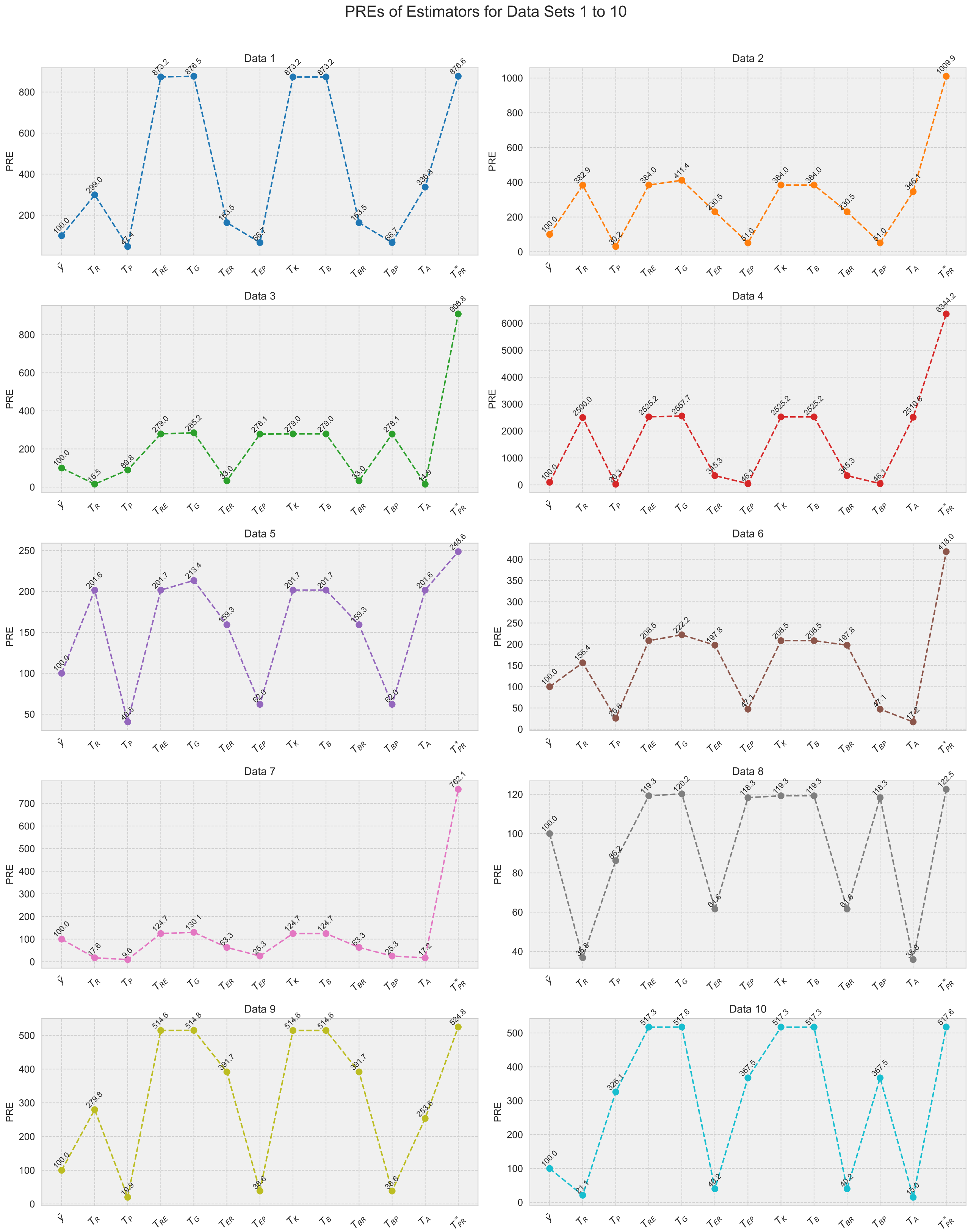

Figure 1. The PRE values of proposed estimator T_{PR}^{*} and existing esitmators

6 Simulation study

To evaluate the effectiveness of the estimators, we conducted a simulation involving two synthetic populations generated using the R programming language to imitate real-world data characteristics provided below. The majority of social and economic surveys often show that the variable under consideration, denoted as y, follows a normal distribution. As a result, we have evaluated the performance of the suggested class of estimators T^*_{PR}.
Data set-9: [Artificial data]: The pair (y,x) is generated with a Population size of $N = 5000$ using a bivariate normal distribution $(Y,X) \sim \mathcal{N}(30,10,0.9,35^2,15^2)$. The sample size of $(n = 1500)$ is specifically considered. $N = 5000, n = 1500, \bar{Y} = 30.0105, \bar{X} = 10.0315, \rho = 0.897584, C_y = 1.18323, C_x = 1.53992$
Data set-10: [Artificial data]: A data set of size $(N = 5000)$ with two variables (y,x) is generated using a bivariate normal distribution $(Y,X) \sim \mathcal{N}(3,1,-0.9,3.5^2,1.5^2)$. A sample size of $n = 1500$ has been selected. $N = 5000, n = 1500, \bar{Y} = 3.00293, \bar{X} = 1.00157, \rho = -0.898164, C_y = 1.18471, C_x = 1.46301$
The absolute bias (ABS) of proposed estimator can be obtained using equation (30), we have calculated the ABS and PRE of the suggested class of estimators with respect to T for various values of $[(\alpha,\gamma),(a,b)]$. The results are displayed in Table 3 and 4.

Table 3: The ABS for the proposed class of estimator T^*_{PR} with respect to $\bar{y}$

	$ABS(\cdot)$ for $\gamma = \rho$									
	Data Sets									
$(\alpha,[a,b])$	1	2	3	4	5	6	7	8	9	10
$(1,[-1,1])$	0.0271062	0.086522	0.173106	0.0658776	0.627382	0.710597	0.329378	0.225248	0.00376007	0.00129651
$(-1,[-1,1])$	0.0265519	0.1049	0.0572386	0.164165	0.372445	0.27854	1.05167	0.219526	0.00363761	0.00129652
$(-1,[0,1])$	0.0268057	0.11117	0.126612	0.192263	0.423219	0.399943	1.3274	0.221472	0.00359564	0.000380001
$(1,[-1,0])$	0.0271092	0.086523	0.173214	0.0658987	0.629001	0.710308	0.324613	0.225334	0.0037591	0.000380185
$(0,[-1,1])$	0.0271062	0.086522	0.173106	0.0658776	0.627382	0.710597	0.319868	0.225248	0.00376007	0.00129651
$(1,[0,1])$	0.0271902	0.077632	0.186044	0.0484058	0.537021	0.593894	0.131261	0.226248	0.00373623	0.000378533
	$ABS(\cdot)$ for $\gamma = C_x$									
	Data Sets									
$(\alpha,[a,b])$	1	2	3	4	5	6	7	8	9	10
$(1,[-1,1])$	0.0271606	0.11675	0.168678	0.057407	0.816994	0.986772	0.678977	0.227539	0.00378459	0.00610211
$(-1,[-1,1])$	0.026696	0.072704	0.164148	0.178236	0.038749	0.0434	0.373301	0.225941	0.00368625	0.00610171
$(-1,[0,1])$	0.0268057	0.11117	0.126612	0.192263	0.423219	0.399943	1.3274	0.221472	0.00359564	0.000380001
$(1,[-1,0])$	0.0271616	0.11675	0.168885	0.0574614	0.823178	0.975526	0.732428	0.227539	0.00378152	0.000364492
$(0,[-1,1])$	0.0271606	0.11675	0.168678	0.057407	0.816994	0.986772	0.678977	0.227539	0.00378459	0.00610211
$(1,[0,1])$	0.0271902	0.11675	0.186044	0.0484058	0.537021	0.593894	0.131261	0.226248	0.00373623	0.000378533
	$ABS(\cdot)$ for $\gamma = C_y$									
	Data Sets									
$(\alpha,[a,b])$	1	2	3	4	5	6	7	8	9	10
$(1,[-1,1])$	0.0271232	0.12811	0.179337	0.0582533	0.742007	0.923349	0.480564	0.227528	0.00377077	0.00131865
$(-1,[-1,0])$	0.0265946	0.059853	0.179406	0.17675	0.036751	0.0412442	0.832121	0.225635	0.00365418	0.000368766
$(-1,[-1,1])$	0.0265882	0.059851	0.150233	0.176848	0.029653	0.0339548	0.81822	0.226056	0.00365788	0.00131865
$(-1,[0,1])$	0.0268057	0.11117	0.126612	0.192263	0.423219	0.399943	1.3274	0.221472	0.00359564	0.000380001
$(1,[-1,0])$	0.0265946	0.12811	0.179406	0.0583051	0.733452	0.91498	0.470643	0.227546	0.00376886	0.000368766
$(0,[-1,1])$	0.0271257	0.12811	0.179337	0.0582533	0.742007	0.923349	0.480564	0.227528	0.00377077	0.00131865
$(0,[0,1])$	0.0271902	0.077632	0.186044	0.0484058	0.537021	0.593894	0.131261	0.226248	0.00373623	0.000378533
	$ABS(\cdot)$ for $\gamma = 0$									
	Data Sets									
α	1	2	3	4	5	6	7	8	9	10
-1	0.0268057	0.11117	0.126612	0.192263	0.423219	0.399943	1.3274	0.221472	0.00359564	0.000380001
0 or 1	0.0271902	0.077632	0.186044	0.0484058	0.537021	0.593894	0.131261	0.226248	0.00373623	0.000378533

Main Findings:
Table 4 presents the PREs for the proposed class of estimator T^*_{PR} with respect to the unbiased estimator $\bar{y}$ under various combinations of α, $\gamma \in \{\rho, C_x, C_y, 0\}$, and (a,b). It is observed that T^*_{PR} demonstrates notably higher efficiency than the generalized estimator T_G in many cases, particularly when $\alpha = 0$ or 1, and γ is closer to 1.
(i)$[(\alpha,\gamma),(a,b)] = [(0,\rho),(-1,1)]$ in Data Set 1 shows maximum efficiency (PRE = 876.552).
(ii)$[(\alpha,\gamma),(a,b)] = [(0,\rho),(-1,0)],[(1,\rho),(-1,0)],[(0,\rho),(-1,1)],[(0,\rho),(0,1)],$
$[(1,\rho),(0,1)],[(0,C_x),(0,1)],[(1,C_x),(0,1)],[(0,C_y),(0,1)],[(1,C_y),(0,1)],$
$[(0,0),(0,0)],[(1,0),(0,0)]$ in Data Set 2 consistently yield high PREs ($\geq$ 1009.9).

Table 4: The PREs for the proposed class of estimator T_{PR}^{*} with respect to $\bar{y}$

$PRE\,(T_{PR}^{*},\bar{y})$ for $\gamma=\rho$										
	Data Sets									
$[\alpha,(a,b)]$	1	2	3	4	5	6	7	8	9	10
$[-1,(-1,0)]$	*****	*****	**904.532**	*****	*****	*****	*****	**123.546**	*****	**520.895**
$[-1,(-1,1)]$	*****	*****	**908.791**	*****	*****	*****	*****	**123.622**	*****	455.557
$[-1,(0,1)]$	*****	*****	**410.845**	*****	*****	*****	*****	**122.547**	*****	**517.597**
$[z,(-1,0)]$	876.454	**1009.9**	*****	**6342.22**	**248.635**	**418.1247**	**308.146**	*****	**521.597**	*****
$[0,(-1,1)]$	**876.552**	**1009.9**	*****	**6344.25**	**248.639**	**417.9549**	**303.689**	*****	**521.462**	*****
$[z,(0,1)]$	873.844	**1125.5**	*****	**8634.17**	**250.196**	**500.0847**	**762.054**	*****	**524.789**	*****
$PRE\,(T_{PR}^{*},\bar{y})$ for $\gamma=C_x$										
	Data Sets									
$[\alpha,(a,b)]$	1	2	3	4	5	6	7	8	9	10
$[-1,(-1,0)]$	*****	*****	**317.261**	*****	*****	*****	*****	**120.344**	*****	**524.111**
$[-1,(-1,1)]$	*****	*****	**316.896**	*****	*****	*****	*****	120.123	*****	96.7988
$[-1,(0,1)]$	*****	*****	**410.845**	*****	*****	*****	*****	**122.547**	*****	**517.597**
$[z,(-1,0)]$	874.765	748.41	*****	**7273.48**	**237.865**	**304.4486**	**136.571**	*****	**518.503**	*****
$[0,(-1,1)]$	874.797	748.4	*****	**7280.36**	**237.857**	**300.9788**	**147.322**	*****	**518.084**	*****
$[z,(0,1)]$	873.844	**1125.5**	*****	**8634.17**	**250.196**	**500.0847**	**762.054**	*****	**524.789**	*****
$PRE\,(T_{PR}^{*},\bar{y})$ for $\gamma=C_y$										
	Data Sets									
$[\alpha,(a,b)]$	1	2	3	4	5	6	7	8	9	10
$[-1,(-1,0)]$	*****	*****	**346.541**	*****	*****	*****	*****	120.286	*****	**521.005**
$[-1,(-1,1)]$	*****	*****	**346.248**	*****	*****	*****	*****	120.062	*****	447.947
$[-1,(0,1)]$	*****	*****	**410.854**	*****	*****	*****	*****	**122.547**	*****	**517.597**
$[z,(-1,0)]$	875.924	682.05	*****	**7168.23**	**225.459**	**324.5945**	**212.535**	*****	**520.245**	*****
$[0,(-1,1)]$	876.001	682.03	*****	**7174.61**	**225.435**	**321.6523**	**208.147**	*****	**519.981**	*****
$[z,(0,1)]$	873.844	**1125.5**	*****	**8634.17**	**250.196**	**500.0847**	**762.054**	*****	**524.789**	*****
$PRE\,(T_{PR}^{*},\bar{y})$ for $\gamma=0$										
	Data Sets									
α	1	2	3	4	5	6	7	8	9	10
-1	*****	*****	**410.845**	*****	*****	*****	*****	**122.547**	*****	**517.597**
z	873.844	**1125.5**	*****	**8634.17**	**250.196**	**500.0847**	**762.054**	*****	**524.789**	*****

(iii) $[(\alpha,\gamma),(a,b)]$, for all $\alpha=-1$, $\gamma\in\{\rho,C_x,C_y\}$, $a\in\{-1,0\}$, and $b=1$, give strong performance in Data Set 3 with PREs such as 904.532, 908.791, 410.845, 317.261, 346.541, and 410.854.

(iv) $[(\alpha,\gamma),(a,b)]$, for all $\alpha\in\{0,1\}$, $\gamma\in\{\rho,C_x,C_y\}$, $a\in\{-1,0\}$, and $b=1$, demonstrate high efficiency in Data Sets 4, 5, 6, and 8, with PREs like 6342.22, 248.639, 417.9549, 500.0847, and 762.054.

(v) $[(\alpha,\gamma),(a,b)]=[(-1,\rho),(-1,0)],[(-1,\rho),(-1,1)],[(-1,\rho),(0,1)],$ $[(-1,C_x),(-1,0)],[(-1,C_x),(0,1)],[(-1,C_y),(0,1)]$ are efficient combinations in Data Set 7, with values such as 308.146, 303.689, 136.571, 147.322, and 212.535.

(vi) $[(\alpha,\gamma),(a,b)]=[(-1,\rho),(-1,0)],[(-1,\rho),(0,1)],[(-1,C_x),(-1,0)],$ $[(-1,C_y),(-1,0)],[(-1,C_y),(0,1)],[(-1,0),(0,0)]$ perform well in Data Set 9 with PREs like 521.597, 518.503, 524.789, and 521.005.

(vii) The pair $[(\alpha,\gamma),(a,b)]=[(1,\rho),(-1,0)]$ indicates that T_{PR}^{*} is more efficient than T_G in six positively correlated data sets (1, 2, 4, 5, 6, 8). However, its superiority is less pronounced at $[(\alpha,\gamma),(a,b)]=[(1,C_x \text{ or } C_y),(-1,0)]$. Therefore, for higher gains in efficiency, it is advisable to choose γ values closer to 1 when y and x are positively correlated.

(viii) The case $[(\alpha,\gamma),(a,b)]=[(-1,\rho),(0,1)]$ is significant in negatively correlated data sets (3, 7, 9), where T_{PR}^{*} consistently outperforms T_G. It is recommended to select γ values near -1 for such cases.

The suggested class of estimators T_{PR}^{*} perform better than the usual unbiased estimator $\bar{y}$ in Table 4. The suggested estimators T_{PR}^{*} differ in their biases up to the first-order approximation. The biases and MSEs measured up to the first-order approximation are sufficient for formulating a population mean estimator, with very few biases. These findings can be broadly applied to various real populations. The extensive simulation study in Section 6 substantiates our theoretical assertions favoring the proposed estimators.

7 Sensitivity analysis

In numerical and simulation studies, the MSEs and PREs are computed using the known values of population parameters C_y, C_x and $\rho_{X,Y}$, whereas in this sensitivity analysis, these metrics are derived from sample estimates. In real-life survey sampling, the parameters C_y, C_x and $\rho_{X,Y}$ are unknown for population. Therefore, we have used the estimates c_y, c_x and $\rho_{x,y}$ of these parameters based on the sample to calculate the MSEs and PREs of the estimators under this sensitivity analysis.

7.1 For real data sets

We have taken two populations (data sets 1 and 2), which are already given in Section 5. We computed MSEs and PREs of the estimators based on c_y, c_x and $\rho_{x,y}$ estimated from the sample data, where C_y, C_x and $\rho_{X,Y}$ are unknown for population parameters. The results are presented in Table 5.

Table 5: MSEs and PREs of estimators for population 1 and 2

Estimators	Population 1		Population 2	
	MSE	**PRE**	**MSE**	**PRE**
$\bar{y}$	33946.74	100.00	65485.30	100.00
T_R	9398.54	361.19	302865.80	21.62
T_P	78582.30	43.20	173187.00	37.81
T_{RE}	4161.87	815.66	59393.80	110.26
T_G	4136.99	820.57	59153.60	110.70
T_{ER}	19161.72	177.16	141040.20	46.43
T_{EP}	53753.60	63.15	76200.90	85.94
T_K	4161.87	815.66	59393.80	110.26
T_B	4161.87	815.66	59393.80	110.26
T_{BR}	19161.72	177.16	141040.20	46.43
T_{BP}	53753.60	63.15	76200.90	85.94
T_A	9385.16	361.71	302944.30	21.62
T^*_{PR}	**4075.37**	**832.97**	**58407.30**	**112.12**

7.2 For Artifically generated data sets

The procedure for sensitivity analysis as follows:
Generate a population $N = 10000$ using bivariate normal distribution $(Y, X) \sim \mathcal{N}(30, 10, \rho_{X,Y}, 8, 4)$. Draw a sample of size $n = 1500$ using SRSWOR. Estimate c_x, c_y and $\rho_{x,y}$ from sample data. We have computed MSEs and PREs for proposed estimator and existing estimators with different amount of correlation $\rho_{x,y} = -0.9, -0.5, 0.5, 0.9$ to evaluate there effectiveness in highly positive and negative correlated sample data. The results are displayed in Table 5.

Table 6: MSEs and PREs of estimators for various correlation values ρ

Estimators	$\rho = -0.9$		$\rho = -0.5$		$\rho = 0.5$		$\rho = 0.9$	
	MSE	**PRE**	**MSE**	**PRE**	**MSE**	**PRE**	**MSE**	**PRE**
$\bar{y}$	0.0351	100.0000	0.0352	100.0000	0.0349	100.0000	0.0351	100.0000
T_R	0.2114	16.6119	0.1774	19.8394	0.0672	51.9760	0.2114	16.6119
T_P	0.0199	176.1793	0.0668	52.6932	0.1778	19.6427	0.0199	176.1793
T_{RE}	0.0067	526.3158	0.0264	133.3333	0.0262	133.3333	0.0067	526.3158
T_G	0.0067	526.3403	0.0264	133.3411	0.0262	133.3411	0.0067	526.3403
T_{ER}	0.1031	34.0545	0.0846	41.6155	0.0292	119.7437	0.1031	34.0545
T_{EP}	0.0074	475.3341	0.0293	120.2489	0.0845	41.3460	0.0074	475.3341
T_K	0.0067	526.3158	0.0264	133.3333	0.0262	133.3333	0.0067	526.3158
T_B	0.0067	526.3158	0.0264	133.3333	0.0262	133.3333	0.0067	526.3158
T_{BR}	0.1031	34.0545	0.0846	41.6155	0.0292	119.7437	0.1031	34.0545
T_{BP}	0.0074	475.3341	0.0293	120.2489	0.0845	41.3460	0.0074	475.3341
T_A	0.2201	15.9594	0.1851	19.0114	0.0713	48.9934	0.2201	15.9594
T^*_{PR}	**0.0067**	**526.3465**	**0.0264**	**133.3596**	**0.0262**	**133.3968**	**0.0067**	**526.3465**

Interpetataion of the results:

The sensitivity analysis results, as detailed in Tables 5 and 6, reveals that the proposed estimator T^*_{PR} significantly outperforms existing estimators in terms of MSE and PRE across both real and simulated datasets.

For the real data sets (Populations 1 and 2) in Table 5, the proposed estimator T^*_{PR} consistently outperforms the other considered estimators in terms of minimum MSE and highest PRE. Notably, for Population 1, it achieves the lowest MSE (4075.37) and the highest PRE (832.97), highlighting its efficiency. A similar pattern is observed for Population 2, with T^*_{PR} maintaining superior performance (MSE = 58407.30, PRE = 112.12).

In the artificially generated data sets (Table 6) across varying levels of correlation ρ, T^*_{PR} consistently yield the lowest MSEs and highest PREs, especially when correlation is strong ($|\rho| = 0.9$). The performance remains efficitive for both highly negative and highly positive correlations, demonstrating the stability and adaptability of the proposed estimator under different dependency structures.

Thus from the above findings, we can reassert that the effectiveness of the proposed estimator remains consistent and exhibits no sensitivity for the considered parameters C_y, C_x, and $\rho_{X,Y}$ at their known values or estimated values.

8 Conclusions

This paper proposes a novel logarithmic-exponential cum ratio-type estimator that targets the estimation of population mean. To verify the theoretical results, we have taken eight real and two artificially generated datasets in numerical and simulation studies, respectively, based on the assumption that the population parameters C_y, C_x, and $\rho_{X,Y}$ are known. In light of this, we also investigated situations where the population parameters are unknown through sensitivity analysis. This analysis, conducted using two real and two artificially generated datasets, focused on cases with highly positive and negative correlation between the study and auxiliary variables. For this, the parameters c_y, c_x, and $\rho_{x,y}$ were estimated from sample data. The proposed estimator demonstrates its practical applicability and is widely used across various fields, including population studies (where the objective is to estimate demographic characteristics); agriculture (to assess crop yields and labor statistics); economics and industry (for analyzing fixed capital, factory output, and income indicators); and environmental and engineering studies (focusing on vehicle efficiency and horsepower). It is also useful for social and demographic research for understanding regional and village-level characteristics. The findings suggest promising directions for advancing estimator performance under different sampling schemes, notably ranked set sampling and stratified random sampling. Based on these truthful findings, this may encourage survey researchers to implement the proposed estimator in real-world scenarios for estimation of the population mean.

Declarations

Competing interests: No

Authors' contributions: SM Afsar Basha: Writing – original draft, Methodology, Conceptualization, Analysis. Mahamood Usman: Supervision, Investigation, Validation, Conceptualization, Analysis, review and editing

Funding: No

Availability of data and materials: Data sets 1 to 8 - Source: Murthy (1967), Shabbir et al. (2014), US Environmental Protection Agency 1991, Murthy (1967), Singh and Chaudhary (1986), Das (1988), Singh and Chaudhary (1986) and Gupta and Kothwala (1990), respectively.

Acknowledgments: Authors are highly thankful to Vellore Institute of Technology for providing financial assistance and the necessary infrastructure to accomplish the present work, authors are thankful for VIT SEED GRANT (RGEMS) Sanction Order No.SG20230035 for technical assistance.

References

[1] A. Adejumobi, M. A. Yunusa, Y. A. Erinola and K. Abubakar, An efficient logarithmic ratio type estimator of finite population mean under simple random sampling. *International Journal of Engineering and Applied Physics*, 3(2)(2023), 700-705.
[2] S. Bahl and R. Tuteja, Ratio and product type exponential estimators. *Journal of Information and Optimization Sciences*, 12(1)(1991), 159-164.
[3] S. Bhushan, R. Gupta and S. K. Pandey, Some log-type classes of estimators using auxiliary information. *International Journal of Scientific Engineering and Research*, 11(2015), 487-491.
[4] S. S. Brar, J. S. Bajwa and J. Kaur, Some new functional forms of the ratio and the product estimators of the population mean. *Investigación Operacional*, 41(3) (2020), 416-425.
[5] W. G. Cochran, Sampling techniques. *john wiley and sons*, (1977).
[6] A. K. Das, Contributions to the theory of sampling strategies based on auxiliary information. Ph. D. Thesis, Bidhan Chandra Krishi Vishwavidyalaya, (1988).
[7] L. K. Grover and P. Kaur, An improved estimator of the finite population mean in simple random sampling. *Model Assisted Statistics and Applications*, 6(1)(2011), 47-55.
[8] M. N. Murthy, Sampling theory and methods, (1967).
[9] M. H. Hansen, W. N. Hurwitz and W. G. Madow, Sample survey methods and theory. Vol. I. Methods and applications, (1953).
[10] G. Ö. Kadilar, A new exponential type estimator for the population mean in simple random sampling. *Journal of Modern Applied Statistical Methods*, 15 (2016), 207-214.
[11] J. Shabbir, A. Haq and S. Gupta, A new difference-cum-exponential type estimator of finite population mean in simple random sampling. *Revista Colombiana de Estadística*, 37(1)(2014), 199-211.
[12] D. Singh and F. S. Chaudhary, Theory and analysis of sample survey designs, (1986).
[13] P. C. Gupta and N. H. Kothwala, A study of second order approximation for some product type estimators. *Indian Society of Agricultural Statistics*, 42(2)(1990), 171-185.