https://doi.org/10.56093/JISAS.V79I3.10
Available online at http://isas.org.in/isa/jisas
JOURNAL OF THE INDIAN SOCIETY OF
AGRICULTURAL STATISTICS 79(3) 2025 303–314

Some Novel Logarithmic Type Estimators for the Finite Population Mean using Auxiliary Information

Mukesh Kumar[1], Anchal Yadav[1] and Shobh Nath Tiwari[2]
[1]*University of Lucknow, Lucknow*
[2]*Banaras Hindu University, Varanasi*

Received 24 October 2025; Revised 11 January 2026; Accepted 12 February 2026

SUMMARY

The incorporation of auxiliary information has become a vital component in the process of statistical estimation, as it substantially enhances the accuracy of estimators for population parameters such as the mean and variance of the study variable. Over the years, several methodologies including the ratio, product, and regression estimators have gained recognition for their effectiveness in this regard. Building on these classical approaches, the present paper introduces some novel and efficient logarithmic ratio-type estimators for estimating the finite population mean under simple random sampling. To analyze the performance of the proposed estimator, mathematical expressions for its bias and mean squared error (MSE) are derived, up to the first-order approximation. Furthermore, the specific conditions under which the proposed estimator demonstrates greater efficiency than the conventional estimators are clearly established. To support the theoretical results, an empirical investigation using two datasets is carried out, together with a simulation study. The findings of this analysis strongly indicate that the proposed logarithmic ratio-type estimator consistently outperforms the competing estimators considered in the study, thereby validating its practical utility.

Keywords: Auxiliary information; Logarithmic-type estimators; Mean squared error; Bias; Simple random sampling.

1. INTRODUCTION

Survey sampling has long been recognized as a cornerstone of statistical inference, providing a systematic framework for estimating population characteristics such as the population mean and other related measures. Over the years, statisticians have consistently explored methods to improve the precision and efficiency of these estimates by incorporating auxiliary information that is correlated with the study variable. When a strong positive linear relationship exists between the study and auxiliary variables, ratio estimators are often employed, particularly under the assumption that the regression line passes through the origin. The classical ratio estimator remains one of the most widely applied techniques in finite population sampling. Conversely, when the relationship between the study and auxiliary variables is negative, product-type estimators are considered more suitable for achieving improved accuracy. In situations where the regression line does not pass through the origin, regression-type estimators have proven to be more effective. These estimators adjust for deviations from proportionality and yield more reliable estimates under non-linear conditions. Over time, several scholars have introduced modified and generalized versions of these estimators to enhance their performance across diverse sampling schemes such as stratified, systematic, and two-stage sampling. More recently, logarithmic-type estimators have attracted growing attention in the field of survey sampling. By applying logarithmic transformation to both study and auxiliary variables, these estimators are capable of stabilizing relationships, reducing skewness, and improving estimation efficiency when associations between variables are non-linear or multiplicative in nature. Within the broader framework

Corresponding author: Anchal Yadav
E-mail address: yanchal635@gmail.com

of sample surveys, the integration of auxiliary information through such estimators contributes substantially to the improvement of population mean estimation. Modern developments in this area highlight the role of logarithmic and exponential-type estimators in addressing the limitations of traditional ratio and regression estimators. Consequently, the evolving body of research underscores the dynamic interplay between theoretical advancement and practical application in the continuous pursuit of unbiased and efficient estimators for finite population parameters.

The traditional ratio estimator was initially presented by Cochran (1940) three on wheat, three on barley and one on oats, the yields of grain and straw per plot were estimated by weighing the total produce on each plot and taking samples, usually from the sheaves, to estimate the ratio of grain to total produce. This paper discusses the sampling errors of this method. The method proved considerably less accurate than was anticipated from previous calculations made by Yates & Zacopanay. Amongst the reasons which are suggested to account for this are the larger sizes of plot and sampling unit used in these experiments and the additional variability introduced by the presence of weeds, undergrowth and moisture. Nevertheless, the method appears to be substantially superior to the older method of cutting small areas from the standing crop, without weighing total produce, only about one-quarter of the number of samples being required to obtain results of equal precision. The samples were taken both by an approximately random process and by grabbing a few shoots haphazardly from each of several sheaves. The grab samples gave on the whole a slightly higher yield of grain, the greatest positive bias being 6%, but were otherwise just as accurate as the random samples. Since the grab samples can be selected and bagged in about one-third of the time required for random samples, their use is recommended for the majority of the samples required in any experiment. The validity of an approximate formula for calculating the variance of a ratio (in the present instance the ratio of grain to total produce and initiated the application of auxiliary information, while Murthy (1964) introduced the classical product-type estimator. A significant advancement was made by Srivastava (1967) who introduced a power transformation estimator, which generalized both ratio and product estimators. An improvement was made by Bahl and Tuteja [4], who recommended exponential ratio and product-type estimators to estimate the unknown population mean of the study variable, based on auxiliary information.

Numerous researchers have since made significant contributions to this area, including Kadilar and Cingi (2004), Khoshnevisan *et al.* (2007), Onyeka (2012), Chaun and Singh (2014), Yadav *et al.* (2014), Onyeka *et al.* (2015), Etebong (2016), Subramani (2016). The arithmetic mean is seen as the basic measure of a population and is more effective in yielding precise results when the population is homogeneous. Numerous efficient estimators for the finite population mean under simple random sampling have been introduced over years by researchers. As an illustration, Singh and Tailor (2005) introduced efficient estimators that make use of the known values of the correlation coefficient. Kadilar and Cingi (2006) proposed techniques that integrate the coefficient of kurtosis and median of an auxiliary variable for estimation of the population variance, while Singh and Vishwakarma (2009) formulated modified ratio and product exponential estimators using a double sampling framework. Haq and Shabbir (2013) introduced an improved family of ratio-type estimators, whereas Singh and Solanki created efficient class of estimators that utilizes auxiliary information. A modified ratio-type estimator, including the known median of an auxiliary variable, was proposed by Subramani and Kumarpandiyan (2012), while Muneer *et al.* (2017) developed estimators for the population mean under both simple and stratified random sampling frameworks, incorporating two auxiliary variables and Izunobi and Onyeka (2018), presented efficient logarithmic ratio and product estimators in the context of simple random sampling.

Brar *et al.* (2014) proposed new logarithmic-type estimators their biases and MSEs are obtained up to a defined order of approximation. They are more efficient than traditional estimators such as mean per unit, ratio, and product types under certain conditions Yusuna *et al.* (2021) Proposes a logarithmic ratio-type estimator for the population coefficient of variation using log-transformed variances. Cekim *et al.* (2020) propose a family of ln-type estimators for the simple random sampling for the estimation of variance and show that they are more efficient than the other types of estimators under certain conditions obtained theoretically and by simulation study. Yusuna *et al.* (2022) Suggested the logarithmic-product-cum-ratio type estimator for the estimation of finite population coefficient of variation. Theoretical and empirical results show that it is more efficient and gives estimates closer to the true value than existing methods. Zaman *et al.* (2023) The study proposes a logarithmic ratio-type estimator for estimating the finite population mean using auxiliary information under simple random sampling. Theoretical MSE expressions, along with numerical and simulation results, show that the proposed estimator is more efficient than existing ones. Adejumobi *et al.* (2023) Proposes a logarithmic ratio-type estimator for finite population mean under simple random sampling. Theoretical bias and MSE are derived, and empirical results from three datasets show it is more efficient than existing estimators.

Gupta *et al*. (2024) Introduce a logarithmic ratio type estimator, recognized for its efficiency, to estimate the mean of a finite population using simple Random Sampling. Qureshi *et al.* (2024) Offered two ln-type estimators for the population mean estimation of a sensitive study variable by using the auxiliary information under the design of basic probability sampling. Alafara *et al.* (2025) proposes an improved logarithmic product cum-ratio estimator for the population Coefficient of Variation using power transformation. Simulation results show that it significantly reduces bias and mean squared error, making it more robust than traditional estimators, especially with outlier data. The main objective of this study is to develop logarithmic ratio-type estimators that effectively utilize auxiliary information to achieve improved estimation of the finite population mean of the study variable. The use of auxiliary information has been a central aspect of survey sampling, enhancing the efficiency and reliability of population mean estimation. Classical ratio estimators, though widely used, often face limitations when the relationship between study and auxiliary variables deviates from linearity. To address such complexities, the present study explores the potential of logarithmic transformation in refining estimation procedures. By integrating the properties of logarithmic functions with ratio-type approaches, the proposed estimators aim to produce more consistent and precise estimates. This methodological advancement is expected to strengthen the application of auxiliary information in complex population settings and broaden the scope of estimation techniques in survey research. This study builds upon the previous foundation,proposing novel estimators that incorporate logarithmic transformations of ratio and product forms to enhance the finite population mean estimation with the effective use of auxiliary information.

2. MATERIALS AND METHODS

In this current study, we proposed an efficient logarithmic ratio type estimator for finite population mean under simple random sampling. Consider a finite population U=$\left(u_1, u_{2,\dots\dots,} u_N\right)$ contains N distinct units. A sample of n units is selected from this population using simple random sampling without replacement. Let y represent the study variable, x represents the auxiliary variable and y_i and x_i stand for the observations on the i^{th} unit. The symbols employed in this context are:

N= Size of the population.

n= Size of the sample.

$\bar{Y}$ = Population mean of study variable.

$\bar{X}$ = Population mean of auxiliary variable.

$$S_y^2 = \left(N-1\right)^{-1}\sum_{i=1}^{N}\left(y_i-\bar{Y}\right)^2$$

$$S_x^2 = \left(N-1\right)^{-1}\sum_{i=1}^{N}\left(x_i-\bar{X}\right)^2$$

$s_y^2 = \left(n-1\right)^{-1}\sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2$: sample variance of study variable.

$s_x^2 = \left(n-1\right)^{-1}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2$: sample variance of auxiliary variable.

ρ_{xy} = Correlation coefficient between X and Y.

$C_y = \dfrac{S_y}{\bar{Y}}$: Coefficient of variation (CV) of study variable.

$C_x = \dfrac{S_x}{\bar{X}}$: Coefficient of variation (CV) of auxiliary variable.

To obtain the bias and mean squared error (MSE) for the proposed estimators and existing ones considered here, we defined the following sample error terms.

Let, $e_0 = \bar{Y}^{-1}\left(\bar{y}-\bar{Y}\right)$ and $e_1 = \bar{X}^{-1}\left(\bar{x}-\bar{X}\right)$ such that E (e_i) =0 for $(i=0,1)$, E $(e_o^2) = \lambda C_y^2$,

E $(e_1^2) = \lambda C_x^2$, E $(e_0 e_1) = \lambda\rho C_y C_x$, where $\lambda = \left(\frac{1}{n}-\frac{1}{N}\right)$.

3. REVIEW OF PRIOR RESEARCHERS

In this section, a few established estimators are discussed along with their sampling properties.

The conventional sample mean is expressed as,

3.1 Usual Mean Estimator

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{3.1}$$

Variance of $\bar{y}$ can be written as,

$$\text{Var}\ (\bar{y}) = \lambda \bar{Y}^2 C_y^2. \tag{3.2}$$

3.2 Cochran (1940) Ratio Estimator

Initially a ratio estimator was suggested by Cochran [1] three on wheat, three on barley and one on oats, the yields of grain and straw per plot were estimated by weighing the total produce on each plot and taking samples, usually from the sheaves, to estimate the ratio of grain to total produce. This paper discusses the sampling errors of this method. The method proved considerably less accurate than was anticipated from previous calculations made by Yates & Zacopanay. Amongst the reasons which are suggested to account for this are the larger sizes of plot and sampling unit used in these experiments and the additional variability introduced by the presence of weeds, undergrowth and moisture. Nevertheless, the method appears to be substantially superior to the older method of cutting small areas from the standing crop, without weighing total produce, only about one-quarter of the number of samples being required to obtain results of equal precision. The samples were taken both by an approximately random process and by grabbing a few shoots haphazardly from each of several sheaves. The grab samples gave on the whole a slightly higher yield of grain, the greatest positive bias being 6%, but were otherwise just as accurate as the random samples. Since the grab samples can be selected and bagged in about one-third of the time required for random samples, their use is recommended for the majority of the samples required in any experiment. The validity of an approximate formula for calculating the variance of a ratio (in the present instance the ratio of grain to total produce and it is expressed as,

$$t_1 = \left(\frac{\bar{y}}{\bar{x}}\right)\bar{X} \tag{3.3}$$

The mean square error (MSE) for t_1, is given by,

$$\text{MSE}\ (t_1) = \lambda\bar{Y}^2\,[\,C_y^2 + C_x^2 \text{ -2 } \rho C_y C_x\,]. \tag{3.4}$$

3.3 Sisodia and Dwivedi (1981) Estimator

Sisodia and Dwivedi (1981) put forward the given estimator,

$$t_2 = \bar{y}\left[\frac{\bar{X}+C_x}{\bar{x}+C_x}\right] \tag{3.5}$$

The mean squared error for estimator t_2, approximating to the first order yields:

$$\text{MSE}\ (t_2) = \lambda\bar{Y}^2\,[\,C_y^2 + \theta_1^2 C_x^2 \text{ -2 } \theta_1 \rho C_y C_x\,]. \tag{3.6}$$

Where, $\theta_1 = \dfrac{\bar{X}}{\bar{X}+C_x}$.

3.4 Bahl and Tuteja (1991) Estimator

Bahl and Tuteja (1991) suggested an estimator,

$$t_3 = \bar{y}exp\left[\frac{\bar{X}-\bar{x}}{\bar{X}+\bar{x}}\right] \tag{3.7}$$

The MSE of t_3, is given by,

$$\text{MSE}\ (t_3) = \lambda\bar{Y}^2\,[\,C_y^2 + \frac{C_x^2}{4} - \rho C_x C_y\,]. \tag{3.8}$$

3.5 Brar *et al.* (2014) Estimator

Bral *et al.* (2014) suggested a functional form of ratio estimator,

$$t_4 = \bar{y}\frac{\bar{X}}{\bar{X}+\bar{x}}\,l\,\text{n}\left(\frac{\bar{x}}{\bar{X}}\right) \tag{3.9}$$

The MSE (t_4), is given by,

$$\text{MSE}\ (t_4) = \lambda\bar{Y}^2\,[\,C_y^2 + \frac{C_x^2}{4} - \rho C_x C_y\,]. \tag{3.10}$$

3.6 Adejumobi *et al.* (2023) Estimator

Adejumobi *et al.* (2023) proposed a logarithmic ratio type estimator as,

$$t_5 = \frac{\left[\bar{y} - ln\left(\frac{\bar{x}}{\bar{X}}\right)\right]}{\bar{x}}\bar{X} \tag{3.11}$$

The MSE (t_5), is given by,

$$\text{MSE }(t_5) = \lambda\left[\bar{Y}^2\left(C_y^2 + C_x^2 - 2\rho C_y C_x\right) + \left(1+2\bar{Y}\right)C_x^2 - 2\bar{Y}\rho C_y C_x\right]. \tag{3.12}$$

4. PROPOSED LOGARITHMIC ESTIMATORS

Motivated by Adejumobi *et al.* (2023), in this article we propose three logarithmic estimators, which are shown through theoretical derivations, numerical analysis, and simulation studies to perform better than the existing logarithmic-type estimators.

4.1 Proposed Estimator-I

$$\theta_1^* = k\frac{\left[\bar{y} - \left\{ln\left(\frac{\bar{x}}{\bar{X}}\right)^{\alpha}\right\}\right]}{\bar{x}}\bar{X} \tag{4.1}$$

Where, 'α' refers to the statistical constants and 'k' is characterizing scalar to be found such that the MSE of θ_1^* is minimum.

By expressing Eq. (4.1) in terms of errors up to the first-order approximation gives us the following result,

$$\theta_1^* = \text{k}\frac{\left[\bar{Y}\left(1+e_0\right) - ln\left(1+e_1\right)^{\alpha}\right]}{\left(1+e_1\right)} \tag{4.2}$$

$$\theta_1^* = \text{k}\left[\left\{\bar{Y}\left(1+e_0\right) - \alpha ln\left(1+e_1\right)\right\}\right]\left(1+e_1\right)^{-1} \tag{4.3}$$

After expanding and simplifying the right-hand side of the Eq. (4.3), while retaining terms for the first-order approximation, we obtain,

$$\theta_1^* = \text{k}\left[\bar{Y}\left(1+e_0\right) - \alpha\left(e_1 - \frac{e_1^2}{2}\right)\right]\left(1 - e_1 + e_1^2\right) \tag{4.4}$$

$$\theta_1^* = \text{k}\bar{Y}\left(1 - e_1 + e_1^2 + e_0 - e_0 e_1\right) - \text{k}\,\alpha\left(e_1 - \frac{3e_1^2}{2}\right) \tag{4.5}$$

By subtracting $\bar{Y}$ (population mean) from both sides of the Eq. (4.5), we obtain,

$$\theta_1^* - \bar{Y} = \left(k-1\right)\bar{Y} + k\bar{Y}\left(e_0 - e_1 + e_1^2 - e_0 e_1\right) - \text{k}\,\alpha\left(e_1 - \frac{3e_1^2}{2}\right). \tag{4.6}$$

By applying expectations to both sides of Eq. (4.6) and substituting the respective expectation values, the bias of (θ_1^*) is obtained as,

$$E(\theta_1^* - \bar{Y}) = \left(k-1\right)\bar{Y} + \text{k}\,\bar{Y}\left(E(e_0) - E(e_1)\right) + \left(E(e_1^2) - E(e_0 e_1)\right) - \text{k}\,\alpha\left(E\left(e_1\right) - \frac{3E(e_1^2)}{2}\right) \tag{4.7}$$

$$\text{Bias }(\theta_1^*) = \left(k-1\right)\bar{Y} + \text{k}\,\bar{Y}(\lambda C_x^2 - \lambda\rho C_y C_x) + \frac{3}{2}\alpha\lambda\text{ k }C_x^2. \tag{4.8}$$

Using Eq. (4.6), we obtain,

$$\theta_1^* - \bar{Y} = \left(k-1\right)\bar{Y} - \text{k}\,\bar{Y}e_1 + \text{k}\,\bar{Y}e_0 - \alpha\,\text{k}\,e_1. \tag{4.9}$$

By squaring both sides of Eq. (4.9) and computing the expected values, we derive the MSE of the estimator θ_1^* to the first-order approximation as,

$$\text{MSE}(\theta_1^*) = \left[\left(k-1\right)^2\bar{Y}^2 + k^2\bar{Y}^2\lambda C_y^2 + \left(\bar{Y}^2 + \alpha^2 + 2\alpha\bar{Y}\right)k^2\lambda C_x^2 - 2k^2\bar{Y}^2\lambda\rho C_y C_x - 2k^2\bar{Y}\lambda\rho C_y C_x\right]. \tag{4.10}$$

$$\text{MSE }(\theta_1^*) = \bar{Y}^2 + Ak^2 - Bk. \tag{4.11}$$

Where,

$$\text{A} = \left[\bar{Y}^2 + \bar{Y}^2\lambda C_y^2 + \left(\bar{Y}^2 + \alpha^2 + 2\alpha\bar{Y}\right)\lambda C_x^2 - 2\bar{Y}^2\lambda\rho C_y C_x - 2\bar{Y}\lambda\rho C_y C_x\right].$$

$$\text{B} = 2\bar{Y}^2.$$

The minimum MSE of the proposed estimator (θ_1^*) is achieved when the constants take the following optimal value, respectively,

$$k = \frac{B}{2A} = k_{opt}\ \text{(say)}.$$

Using the optimal value for k_{opt}, the estimator's minimum MSE is defined as,

$$MSE\left(\theta_1^*\right) = \left[\bar{Y}^2 - \frac{B^2}{4A}\right]. \tag{4.12}$$

4.2 Proposed Estimator-II

$$\theta_2^* = \left[\bar{y} - \alpha\left\{\frac{ln(\bar{x})}{ln(\bar{X})}\right\}\right]\frac{\bar{X}}{\bar{x}} \tag{4.13}$$

Adapting the similar procedure to calculate MSE and Bias as we did for estimator θ_1^*. Writing estimator θ_2^*, using the approximations given in the notation section,

By expressing Eq.(4.13) in terms of errors up to the first-order approximation gives us the following result,

$$\theta_2^* = \left[\bar{Y}(1+e_0) - \alpha\left\{\frac{\ln(\bar{X}) + ln(1+e_1)}{ln(\bar{X})}\right\}\right](1+e_1)^{-1} \tag{4.14}$$

After expanding and simplifying the right-hand side of the Eq. (4.14), while retaining terms for first-order approximation, we obtain,

$$\theta_2^* = \left[\bar{Y}(1+e_0) - \alpha\left\{1+\theta\left(e_1 - \frac{e_1^2}{2}\right)\right\}\right](1-e_1+e_1^2) \tag{4.15}$$

$$\theta_2^* = \left[\bar{Y}(1-e_1+e_1^2+e_0-e_0e_1) - \alpha\left(1-e_1+e_1^2+\theta e_1 - \frac{3\theta e_1^2}{2}\right)\right] \tag{4.16}$$

Where, $\theta = \frac{1}{ln(\bar{X})}$.

By subtracting $\bar{Y}$ (population mean) from both sides of the Eq. (4.16), we obtain,

$$\theta_2^* - \bar{Y} = \bar{Y}(e_0 - e_1 + e_1^2 - e_0e_1) - \alpha\left(1-e_1+e_1^2+\theta e_1 - \frac{3\theta e_1^2}{2}\right) \tag{4.17}$$

By applying expectations to both sides of Eq. (4.17) and substituting the respective expectation values, the bias of (θ_1^*) is obtained as,

$$E(\theta_2^* - \bar{Y}) = \left[\bar{Y}\left(E(e_0) - E(e_1) + E(e_1^2) - E(e_0e_1)\right) - \alpha\left(1 - E(e_1) + E(e_1^2) + \theta E(e_1) - \frac{3\theta E(e_1^2)}{2}\right)\right] \tag{4.18}$$

$$Bias\ (\theta_2^*) = \bar{Y}\lambda C_x^2 - \bar{Y}\lambda\rho C_y C_x - \alpha - \alpha\lambda C_x^2 - \frac{3\theta\alpha\lambda C_x^2}{2}. \tag{4.19}$$

Using Eq. (4.17), we obtain,

$$\theta_2^* - \bar{Y} = \bar{Y}e_0 - \bar{Y}e_1 - \alpha + \alpha e_1 - \alpha\theta e_1 \tag{4.20}$$

By squaring both sides of Eq. (4.20) and computing the expected values, we derive the MSE of the estimator θ_2^* to the first-order approximation as,

$$MSE(\theta_2^*) = \left[\bar{Y}^2\lambda C_y^2 + \bar{Y}^2\lambda C_x^2 + \alpha^2 + \alpha^2\lambda C_x^2 + \alpha^2\theta^2\lambda C_x^2 - 2\bar{Y}\alpha\lambda C_x^2 - 2\alpha^2\theta\lambda C_x^2 + 2\alpha\theta\bar{Y}\lambda C_x^2 - 2\bar{Y}^2\lambda\rho C_y C_x + 2\alpha\bar{Y}\lambda\rho C_y C_x - 2\bar{Y}\alpha\theta\lambda\rho C_y C_x\right] \tag{4.21}$$

$$MSE(\theta_2^*) = \alpha^2 P + \alpha Q + R. \tag{4.22}$$

Where,

$P = \left(1 + \lambda C_x^2 + \theta^2\lambda C_x^2 - 2\theta\lambda C_x^2\right)$.

$Q = \left(2\theta\bar{Y}\lambda C_x^2 - 2\bar{Y}\lambda C_x^2 + 2\bar{Y}\lambda\rho C_y C_x - 2\bar{Y}\theta\lambda\rho C_y C_x\right)$.

$R = \left(\bar{Y}^2\lambda C_y^2 + \bar{Y}^2\lambda C_x^2 - 2\bar{Y}^2\lambda\rho C_y C_x\right)$.

The minimum MSE of the proposed estimator (θ_2^*) is achieved when the constants take the following optimal value, respectively,

$$\alpha = -\frac{Q}{2P} = \alpha_{opt} \text{ (say)}$$

Using the optimal value for α_{opt}, the estimator's minimum MSE is defined as,

$$\text{MSE}\ (\theta_2^*) = \left[\text{R} - \frac{\text{Q}^2}{4\text{P}}\right]. \tag{4.23}$$

4.3 Proposed Estimator-III

$$\theta_3^* = \bar{y}\left\{\frac{ln(\bar{X})}{ln(\bar{x})}\right\}^{\alpha} \tag{4.24}$$

Adapting the similar procedure to calculate MSE and Bias as we did for the estimator θ_1^* and θ_2^* writing estimator θ_3^*, using the approximations given in the notation section, we get:

$$\theta_3^* = \bar{Y}\left(1+e_0\right)\left\{\frac{ln\left(\bar{X}\right)}{ln\left(\bar{X}\left(1+e_1\right)\right)}\right\}^{\alpha} \quad (4.25)$$

$$\theta_3^* = \bar{Y}\left(1+e_0\right)\left\{1+kln\left(1+e_1\right)\right\}^{-\alpha} \quad (4.26)$$

Where, $k = \dfrac{1}{ln\left(\bar{X}\right)}$.

After expanding and simplifying the right-hand side of the Eq. (4.26), while retaining terms for the first-order approximation, we obtain,

$$\theta_3^* = \bar{Y}\left(1+e_0\right)\left\{1+k\left(e_1-\frac{e_1^2}{2}\right)\right\}^{-\alpha} \quad (4.27)$$

$$\theta_3^* = \bar{Y}\left(1+e_0\right) - \alpha\ k\ e_1\bar{Y}\left(1+e_0\right) + \alpha k\bar{Y}\left(1+e_0\right) + \alpha^2k^2\bar{Y}\left(1+e_0\right)\frac{e_1^2}{2} + \alpha k^2\bar{Y}\left(1+e_0\right)\frac{e_1^2}{2}. \quad (4.28)$$

Subtracting $\bar{Y}$ (population mean) from both sides of the Eq. (4.28) we have,

$$\theta_3^* - \bar{Y} = \begin{bmatrix} \bar{Y}\left(1+e_0\right) - \alpha ke_1\bar{Y}\left(1+e_0\right) + \alpha k\bar{Y}\left(1+e_0\right) + \\ \alpha^2k^2\bar{Y}\left(1+e_0\right)\dfrac{e_1^2}{2} + \alpha k^2\bar{Y}\left(1+e_0\right)\dfrac{e_1^2}{2} - \bar{Y} \end{bmatrix} \quad (4.29)$$

By applying expectations of both sides of the Eq. (4.29) and computing expectation values, the bias of (θ_3^*) is obtained as,

$$E(\theta_3^* - \bar{Y}) = \bar{Y}\left(1+E(e_0)\right) - \alpha\ \mathrm{k}\ e_1\bar{Y}\left(1+E(e_0)\right) + \alpha k\bar{Y}\left(1+E(e_0)\right) + \alpha^2k^2\bar{Y}\left(1+E(e_0)\right)\frac{e_1^2}{2} + \alpha k^2\bar{Y}\left(1+E(e_0)\right)\frac{E\left(e_1^2\right)}{2} - \bar{Y} \quad (4.30)$$

$$\text{Bias}\ (\theta_3^*) = \bar{Y}\lambda\left[\alpha^2k^2C_x^2 + \frac{C_x^2}{2}k\alpha\left(1+\alpha\right) - k\alpha\rho C_yC_x\right], \quad (4.31)$$

Now, squaring both sides of (θ_1^* - $\bar{Y}$) and taking expectation, we get the expression for the MSE

$$\text{MSE}\ (\theta_3^*) = \bar{Y}^2\lambda\left(C_y^2 + \alpha^2k^2C_x^2 - 2k\alpha\rho C_yC_x\right). \quad (4.32)$$

Where, $\lambda = \left(\dfrac{1}{n} - \dfrac{1}{N}\right)$.

Differentiate above equation with respect to α_{opt} and equate it equal to zero, we get

$$\alpha = \frac{\rho C_y}{kC_x} = \alpha_{opt}\ \text{(say)}$$

$$\text{MSE}\ (\theta_3^*) = \bar{Y}^2\lambda C_y^2\left(1-\rho^2\right). \quad (4.33)$$

5. COMPARISONS OF EFFICIENCIES

The comparative assessment of efficiencies between the proposed and existing estimators in survey sampling plays a vital role in identifying the estimator that yields more precise and reliable results. In this section, a class of estimators for the population mean $\bar{Y}$ is introduced, and the conditions under which they surpass the performance of existing estimators are summarized in the Table 1 as given below, and the proposed logarithmic ratio-type estimator θ_3^* demonstrates superiority only when certain efficiency criteria are satisfied.

Table 1. Theoretical Efficiency Comparison of the existing estimator with proposed estimator θ_3^*.

Estimator	Efficiency Condition
Sample mean per unit estimator ($\bar{y}$)	$\lambda\bar{Y}^2C_y^2 - \left[\bar{Y}^2\lambda C_y^2\left(1-\rho^2\right)\right] > 0$
Cochran [1] estimator (t_1)	$\lambda\bar{Y}^2[C_y^2 + \theta_1^2C_x^2 + 2\rho C_yC_x] - \left[\bar{Y}^2\lambda C_y^2\left(1-\rho^2\right)\right] > 0$
Sisodia and Dwivedi [30] estimator (t_2)	$\lambda\bar{Y}^2[C_y^2 + \frac{C_x^2}{4} - \rho C_xC_y] - \left[\bar{Y}^2\lambda C_y^2\left(1-\rho^2\right)\right] > 0$
Bahl and Tuteja [4] estimator (t_3)	$\lambda\bar{Y}^2[C_y^2 + \frac{C_x^2}{4} - \rho C_xC_y] - \left[\bar{Y}^2\lambda C_y^2\left(1-\rho^2\right)\right] > 0$
Adejumobi [26] estimator (t_5)	$\left[C - \frac{B^2}{A}\right] - \left[\bar{Y}^2\lambda C_y^2\left(1-\rho^2\right)\right] > 0$

6. EMPIRICAL STUDY

The performance of proposed and existing estimators examined through two real data sets.

Data set: I **(*Source:* Sukhatme & Chand)**

The data set pertains to a study on apple production. In this data, the study variable $\left(Y\right)$ represents the number of apple trees of bearing age in the year 1964, while the auxiliary variable (X) denotes the number of bushels of apples harvested during the same year (1964).

The relevant summary statistics, computed from this data set, are reported below and are used in the empirical study for comparison and efficiency assessment of the proposed estimators:

$N=200; n=20; \bar{Y}=1031.82; \bar{X}=2934.58; \rho=0.93; C_y=1.59775; C_x=2.00625.$

Table 2. MSE Values of the Estimators Along with PRE Values for Dataset I

Estimator	MSE	PRE
$\bar{y}$	122303.3	100.00
t_1	29494.93	414.6585
t_2	29426.69	415.6203
t_3	27689.83	441.6901
t_4	27689.83	441.6901
t_5	29592.06	413.2976
θ_1^*	28791.79	424.7852
θ_2^*	28405.11	430.5678
θ_3^*	**16523.17**	**740.1925**

***Data set: II (Source:* Kadilar and Cingi (2004)**

We have obtained this dataset from Kadilar and Cingi (2004) mean square error (MSE. This data relates to 106 villages in the Aegean Region of Turkey in 1999 (Source: Institute of Statistics, Republic of Turkey), where simple random sampling was considered. The study involves apple production amount (taken as the study variate) and the number of apple trees (taken as the auxiliary variate).

X: Number of apple trees.

Y: Apple production amount.

Here are the parameters taken from data set I for the empirical study purpose:

$N=106; n=20; \bar{Y}=2212.59; \bar{X}=27421.70; \rho=0.86; C_y=5.22; C_x=2.10.$

Table 3. MSE Values of the Estimators Along with PRE Values for Dataset II

Estimator	MSE	PRE
$\bar{y}$	5411348	100.00
t_1	2542740	212.82
t_2	2542893	212.80
t_3	3758095	143.99
t_4	3758095	143.99
t_5	2541840	212.89
θ_1^*	1673128	323.43
θ_2^*	2432834	222.43
θ_3^*	**1409115**	**384.02**

Table 2 and Table 3 present the MSEs and PREs of both the proposed and existing estimators on Dataset I and Dataset II, respectively. The findings indicate that the proposed estimator yields the lowest MSE and the highest PRE across both populations, suggesting its superior efficiency over the other estimators examined in the study.

The computation of Percent Relative Efficiency (PRE) is carried out with the formula given below:

$$PRE = \frac{Var(\bar{y})}{MSE(Proposed\ estimators)} \times 100.$$

7. SIMULATION ANALYSIS

We have performed a simulation analysis to assess the relative efficiency and mean square error of the proposed estimators. The simulation process involved the steps mentioned below:

(1) We simulated 1000 observations from a Bivariate Normal distribution with the mean vector $\mu=[26,32]$ and covariance matrix

$$\Sigma = \begin{bmatrix} \sigma_y^2 & \sigma_y\sigma_x\rho_{xy} \\ \sigma_y\sigma_x\rho_{xy} & \sigma_x^2 \end{bmatrix}$$

using parameter values; $\sigma_x^2=4$, $\sigma_y^2=9$ and the correlation coefficient (ρ_{xy}) was varied as 0.3, 0.5, 0.7 and 0.9.

Table 4. MSE and PRE Comparison of Existing and Proposed Estimators for Various Correlations and Sample Sizes

$N = 1000$				
ρ_{xy}	***n***	**Estimators**	**MSE**	**PRE**
0.3	100	$\bar{y}$	0.08104	100.00
		t_1	0.09571	84.67
		t_2	0.09551	84.67
		t_3	0.07474	108.43
		t_4	0.07475	108.43
		t_5	0.09792	82.76
		θ_1^*	0.09791	82.77
		θ_2^*	0.09571	84.67
		θ_3^*	**0.07368**	**109.99**
	200	$\bar{y}$	0.03602	100.00
		t_1	0.04253	84.67
		t_2	0.04244	84.67
		t_3	0.03322	108.43
		t_4	0.04352	108.43
		t_5	0.04352	82.76
		θ_1^*	0.04351	82.77
		θ_2^*	0.04253	84.67
		θ_3^*	**0.03274**	**109.99**
	300	$\bar{y}$	0.02101	100.00
		t_1	0.02481	84.67
		t_2	0.02476	84.67
		t_3	0.01937	108.43
		t_4	0.01937	108.43
		t_5	0.02538	82.76
		θ_1^*	0.02538	82.76
		θ_2^*	0.02481	84.67
		θ_3^*	**0.01910**	**109.99**
0.5	100	$\bar{y}$	0.08104	100.00
		t_1	0.06912	117.27
		t_2	0.06898	117.49
		t_3	0.06144	131.90
		t_4	0.06144	131.90
		t_5	0.07049	114.97
		θ_1^*	0.07048	114.98
		θ_2^*	0.06910	117.27
		θ_3^*	**0.06072**	**133.46**
	200	$\bar{y}$	0.03602	100.00
		t_1	0.03071	117.27
		t_2	0.03065	117.49
		t_3	0.02730	131.90
		t_4	0.02730	131.90
		t_5	0.03132	114.97
		θ_1^*	0.03132	114.97
		θ_2^*	0.03071	117.27
		θ_3^*	**0.02699**	**133.46**
	300	$\bar{y}$	0.02101	100.00
		t_1	0.01791	117.27
		t_2	0.01788	117.49
		t_3	0.01593	131.90
		t_4	0.01593	131.90
		t_5	0.01827	114.97
		θ_1^*	0.01827	114.97
		θ_2^*	0.01791	117.27
		θ_3^*	**0.01574**	**133.46**

0.7	100	$\bar{y}$	0.08104	100.00
		t_1	0.04250	190.66
		t_2	0.04246	190.87
		t_3	0.04814	168.35
		t_4	0.04814	168.35
		t_5	0.04306	188.21
		θ_1^*	0.04305	188.22
		θ_2^*	0.04250	190.66
		θ_3^*	**0.04129**	**196.27**
	200	$\bar{y}$	0.03602	100.00
		t_1	0.01889	190.66
		t_2	0.01887	190.87
		t_3	0.02139	168.35
		t_4	0.02139	168.35
		t_5	0.01914	188.21
		θ_1^*	0.01913	188.22
		θ_2^*	0.01889	190.66
		θ_3^*	**0.01835**	**196.27**
	300	$\bar{y}$	0.02101	100.00
		t_1	0.01102	190.66
		t_2	0.01100	190.87
		t_3	0.01248	168.35
		t_4	0.01248	168.35
		t_5	0.01116	188.21
		θ_1^*	0.01116	188.22
		θ_2^*	0.01102	190.66
		θ_3^*	**0.01070**	**196.27**

0.9	100	$\bar{y}$	0.08104	100.00
		t_1	0.01591	509.37
		t_2	0.01594	508.35
		t_3	0.03483	232.61
		t_4	0.03483	232.61
		t_5	0.01563	518.41
		θ_1^*	0.01563	518.42
		θ_2^*	0.01590	509.38
		θ_3^*	**0.01538**	**526.89**
	200	$\bar{y}$	0.03601	100.00
		t_1	0.00707	509.37
		t_2	0.00708	508.35
		t_3	0.01548	232.61
		t_4	0.01548	232.61
		t_5	0.00694	518.41
		θ_1^*	0.00694	518.42
		θ_2^*	0.00706	509.38
		θ_3^*	**0.00683**	**526.89**
	300	$\bar{y}$	0.02101	100.00
		t_1	0.00412	509.37
		t_2	0.00413	508.35
		t_3	0.00903	232.61
		t_4	0.00903	232.61
		t_5	0.00405	518.41
		θ_1^*	0.00404	518.42
		θ_2^*	0.00412	509.38
		θ_3^*	**0.00398**	**526.89**

(2) The simulated population was used to obtain samples consisting of 100, 200, and 300 units.

(3) Sample statistics, including the mean, variance, and other relevant parameters, were derived for further evaluation.

(4) The analysis involved calculating the MSE and PRE of both proposed and existing estimators for multiple ρ_{xy} values using the simulated samples.

(5) The procedure of obtaining the MSEs of the estimators under simple random sampling was replicated 10,000 times.

8. RESULTS AND DISCUSSION

Table 2 indicates that the proposed estimators achieve lower mean squared error (MSE) than the usual sample mean ($\bar{y}$) when the correlation between the study and auxiliary variables is sufficiently high. Among them, the estimator θ_3^* shows the best performance, attaining the lowest MSE and the highest percentage relative efficiency (PRE). Similar efficiency gains for the estimator θ_3^* are also observed in Table 3.

A simulation study assuming bivariate normal populations was conducted by varying the correlation coefficient $\left(\rho_{xy} = 0.3, 0.5, 0.7, 0.9\right)$ and sample sizes $\left(n = 100, 200, 300\right)$. The results in Table 4 indicate that the performance of the proposed logarithmic ratio-type estimators depends strongly on the level of correlation. For moderate to high correlation $\left(\rho_{xy} \geq 0.5\right)$, the proposed estimators, particularly the estimator θ_3^*, generally outperform the sample mean in terms of MSE and PRE. However, under weak correlation ($\rho_{xy} = 0.3$), the proposed estimators do not perform well, and their efficiency advantage over the usual sample mean either diminishes or disappears. This outcome represents a clear limitation of the proposed logarithmic ratio-type estimators, as their effectiveness depends on the availability of auxiliary information that is sufficiently correlated with the study variable. Consequently, their practical applicability is restricted in low-correlation settings.

9. CONCLUSION

This study proposes novel logarithmic ratio-type estimators for estimating the finite population mean under simple random sampling by incorporating auxiliary information. The bias and mean squared error expressions of the proposed estimators have been derived up to the first-order approximation. The empirical and simulation studies, conducted using bivariate normal populations, provide insights into the performance of these estimators under varying correlation structures. The results indicate that the efficiency gains of the proposed estimators are closely linked to the strength of correlation between the study variable and the auxiliary variable. In particular, the estimators θ_1^* and θ_2^* do not outperform the sample mean $\bar{y}$ when the correlation is weak, which highlights a limitation of logarithmic ratio-type estimators in such scenarios. However, when moderate to high correlation exists, the proposed estimators exhibit improved performance compared to conventional estimators. Among the proposed estimators, θ_3^* consistently demonstrates superior efficiency and lower MSE relative to θ_1^* and θ_2^* across most practical situations considered in this study. Therefore, the use of θ_3^* is recommended in applications where reliable auxiliary information with sufficiently strong correlation is available. Overall, the study demonstrates that logarithmic ratio-type estimators provide meaningful efficiency improvements in suitable correlation settings, offering valuable guidance for their application in real-world estimation problems.

10. CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

Adejumobi, A., Yunusa, M.A., Erinola, Y.A., and Abubakar, K. (2023). An efficient logarithmic ratio type estimator of finite population mean under simple random sampling. *International Journal of Engineering and Applied Physics,* **3**(2), 700-705.

Alafara, A.A., Imam, I.A., and Yunusa, M.A. (2025). Improved logarithmic product cum ratio-type estimators for population coefficient of variation using known power transformation. *International Journal of Scientific Research in Mathematical and Statistical Sciences,* **12**(2), 15-25.

Bahl, S., and Tuteja, R.K. (1991). Ratio and product type exponential estimators. *Journal of Information and Optimization Sciences,* **12**(1), 159-164. https://doi.org/10.1080/02522667.1991.10699058

Brar, S.S., Bajwa, J.S., and Kaur, J. (2014). Some new functional forms of the ratio and the product estimators of the population mean. *Pakistan Journal of Statistics,* **30**(4), 495-506.

Cekim, H.O., and Kadilar, C. (2020). Ln-type variance estimators in simple random sampling. *Pakistan Journal of Statistics and Operations Research,* **16**(4), 689-696. https://doi.org/10.18187/pjsor.v16i4.3072

Chanu, W.W., and Singh, B.K. (2014). Improved class of ratio-cum-product estimators of finite population mean in two-phase sampling. *Global Journal of Science Frontier Research: Mathematics and Decision Sciences,* **14**(2), 69-81.

Clement, E.P. (2016). An improved ratio estimator for population mean in stratified random sampling. *European Journal of Statistics and Probability,* **4**(4), 12-17.

Cochran, W.G. (1940). The estimation of the yields of cereal experiments by sampling for the ratio of grain to total produce. *Journal of Agricultural Science,* **30**(2), 262-275. https://doi.org/10.1017/S0021859600048012

Gupta, R., Ali, M., Yadav, S.K., Kumar, G., and Kumar, S. (2024). Improved estimation of the population mean by integrating additional auxiliary parameters. *Korean Journal of Physiology and Pharmacology,* **28**(1), 205-210.

Izunobi, C.H., and Onyeka, A.C. (2018). Logarithmic ratio and product-type estimators of population mean in simple random sampling. *International Journal of Mathematics and Statistics Studies,* **6**(2), 45-53.

Kadilar, C. (2006). New ratio estimators using correlation coefficient. *Hacettepe Journal of Mathematics and Statistics,* **35**(1), 103-109.

Kadilar, C., and Cingi, H. (2004). Ratio estimators in simple random sampling. *Applied Mathematics and Computation,* **151**(3), 893-902. https://doi.org/10.1016/S0096-3003(03)00803-8

Khoshnevisan, M., Singh, R., Chauhan, P., and Sawan, N. (2007). A general family of estimators for estimating population mean using known value of some population parameter(s). *Infinite Study*.

Mishra, M., Singh, B.P., and Singh, R. (2017). Estimation of population mean using two auxiliary variables in stratified random sampling. *Journal of Reliability and Statistical Studies,* **10**(2), 59-68.

Muneer, S., Shabbir, J., and Khalil, A. (2017). Estimation of finite population mean in simple random sampling and stratified random sampling using two auxiliary variables. *Communications in Statistics—Theory and Methods,* **46**(5), 2181-2192. https://doi.org/10.1080/03610926.2015.1035394

Murthy, M.N. (1964). Product method of estimation. *Sankhyā: The Indian Journal of Statistics,* **26**(1), 69-74.

Onyeka, A.C. (2012). Estimation of population mean in post-stratified sampling using known value of some population parameters. *Statistical Transition – New Series,* **13**(1), 65-78.

Onyeka, A.C., Izunobi, C.H., and Iwueze, I.S. (2015). Estimation of population ratio in post-stratified sampling using variable transformation. *Open Journal of Statistics,* **5**(1), 1-9. https://doi.org/10.4236/ojs.2015.51001

Onyeka, A.C., Nlebedim, V. U., and Izunobi, C.H. (2013). Exponential estimators of population mean in post-stratified sampling using known value of some population parameters. *Global Journal of Science Frontier Research: Mathematics and Decision Sciences,* **13**(9).

Qureshi, M.N., Faizan, Y., Shetty, A., Ahelali, M.H., Hanif, M., and Alamri, O.A. (2024). Ln-type estimators for the estimation of the population mean of a sensitive study variable using auxiliary information. *Heliyon,* **10**(1), 1-7.

Singh, H. P., and Tailor, R. (2005). Estimation of finite population mean with known coefficient of variation of an auxiliary character. *Statistica,* **65**(3), 301-313. https://doi.org/10.6092/issn.1973-2201/1299

Singh, H.P., and Vishwakarma, G.K. (2009). A general procedure for estimating population mean in successive sampling. *Communications in Statistics—Theory and Methods,* **38**(2), 293-308. https://doi.org/10.1080/03610920802169594

Sisodia, B.V.S., and Dwivedi, V.K. (1981). A modified ratio estimator using coefficient of variation of auxiliary variable. *Journal of the Indian Society of Agricultural Statistics,* **33**(1), 13-18.

Srivastava, S.K. (1967). An estimator using auxiliary information in sample surveys. *Journal of the Indian Society of Agricultural Statistics,* **19**(2), 1-10. https://doi.org/10.1177/0008068319670205

Subramani, J. (2016). Modified ratio-cum-product estimator for estimation of finite population mean with known correlation coefficient. *Biometrics & Biostatistics International Journal,* **4**(6), 1-5. https://doi.org/10.15406/bbij.2016.04.00113

Subramani, J., and Kumarapandiyan, G. (2012). Variance estimation using median of the auxiliary variable. *International Journal of Probability and Statistics,* **1**(3), 36-40.

Yadav, S.K., Mishra, S. S., and Shukla, A.K. (2014). Improved ratio estimators for population mean based on median using linear combination of population mean and median of an auxiliary variable. *American Journal of Operations Research,* **4**(2), 21-27.

Yunusa, M.A., Adejumobi, A., Erinola, Y.A., and Abubakar, K. (2021). Logarithmic ratio-type estimator of population coefficient of variation. *Asian Journal of Probability and Statistics,* **14**(2), 13-22. https://doi.org/10.9734/ajpas/2021/v14i230323

Yusuna, Y.M.A., Audu, A., and Adejumobi, A. (2022). Logarithmic product-cum-ratio type estimator for estimating finite population coefficient of variation. *Oriental Journal of Physical Sciences,* **7**(2), 82-87.

Zaman, T., and Iftikhar, S. (2023). A new logarithmic ratio type estimator of population mean for simple random sampling: A simulation study. *Journal of Science and Arts,* **23**(4), 839-848.