International Journal of Advanced Statistics and Probability, 13 (1) (2026) 54-67

International Journal of Advanced Statistics and Probability

Website: www.sciencepubco.com/index.php/IJASP
https://doi.org/10.14419/y58gth40
Research paper

Enhanced Logarithmic-Type Estimators for Estimation of Population Variance Under Simple Random Sampling Framework: Computational Analysis and Simulation Evidence

Anchal Yadav *, Mukesh Kumar

Department of Statistics, University of Lucknow, Lucknow-226007, Uttar Pradesh, India
**Corresponding author E-mail: yanchal635@gmail.com*

Received: February 27, 2026, Accepted: April 19, 2026, Published: April 22, 2026

Abstract

Survey sampling continues to play a pivotal role in modern statistical research, offering a efficient means of estimating population parameters without requiring a complete enumeration. However, practical data collection frequently involves extreme or atypical observations that, if overlooked, can distort estimation accuracy and lead to biased conclusions. The former research primarily focuses on use of auxiliary information to estimate the population parameters of main variable. The class of estimators suggested in this category one ratio type estimators, product type estimators and regression type estimators. Later, Bahl and Tuteja (1991) suggested a new class of exponential and product estimators. Recent advancements have introduced logarithmic-type estimators, but their statistical properties remain only partially explored. Recognizing this challenge, the present study proposes a refined logarithmic-type estimator for finite population variance within the framework of simple random sampling without replacement (SRSWOR). The estimator incorporates auxiliary information to improve the precision and stability of variance estimation. Theoretical properties, including bias and mean squared error (MSE), are derived up to the first order of approximation, ensuring a rigorous analytical foundation. To assess its performance, comparative and simulation-based analyses are carried out against several existing estimators. The findings reveal that the proposed estimator consistently produces lower MSE values and exhibits greater robustness. These results confirm both the theoretical and empirical superiority of the proposed approach. Overall, the study contributes to the growing body of literature on efficient estimation techniques by presenting a more accurate and reliable alternative for population variance estimation in survey sampling and their applications.

Keywords: *Auxiliary Information; Logarithmic-Type Estimators; Mean Squared Error; Bias; Simple Random Sampling.*

1. Introduction

Survey sampling forms the cornerstone of statistical inquiry, aiming to obtain dependable, representative, and accurate information about population characteristics without conducting a full census. The primary objective is to develop efficient estimators capable of generating precise and unbiased results with minimal use of time, resources, and manpower (Yang et al., 2020). In practical scenarios, however, population data frequently include extreme or outlying observations that deviate markedly from the central tendency. Ignoring such outliers during estimation procedures can produce misleading results, potentially leading to the underestimation or overestimation of true population values. Therefore, incorporating auxiliary information related to these extreme values has been recognized as essential for improving the accuracy, stability, and efficiency of estimators. Over the years, researchers have developed advanced statistical methodologies that integrate extreme value information into estimator construction to mitigate these issues. Moreover, numerous studies have emphasized the pivotal role of population variance estimation in quantifying variability within a population. Such estimation is of immense practical relevance in domains like business, industry, services, pharmaceuticals, medicine, biology, and agriculture, where understanding dispersion assists in data-driven decision-making and quality enhancement. In addition, empirical findings demonstrate that variance estimators are extensively utilized in fields such as clinical research, economics, agriculture, security, and forensic science. In clinical contexts, they assess the variation in treatment effects, while in economics; they gauge market volatility and macroeconomic fluctuations such as GDP and inflation. In agriculture, they facilitate yield analysis, guiding farmers in optimizing resource use. In the realm of security, variance measures assist in assessing risks like cyber-attacks and natural disasters, and in forensic science, they aid in comparing biological evidence such as DNA and fingerprints with improved precision. Consequently, the pursuit of more efficient and reliable variance estimators continues to be a prominent focus within statistical research.

1.1. Literature review

Foundational contributions in this area were made by Isaki (1983) who proposed improved strategies for handling population variance, and by Bahl and Tuteja (1991), who introduced a exponential ratio and product type estimator that demonstrated superior performance over traditional estimators. Gracia (1996) also developed a ratio estimator for the estimation of population variance, Subsequent researchers, including Upadhyaya and Singh (1999),Upadhyaya and Singh (2004) , Chandra and Singh (2005), Arcos et al. (2005) Kadilar and Cingi (2006), H. P. Singh and Solanki (2013) and Yadav et al. (2015), further advanced this line of work by developing a wide range of modified and generalized estimators for more efficient estimation of finite population variance. Several researchers have contributed significantly to this area of study. Gupta and Shabbir (2010) introduced a regression-cum-exponential estimator for population variance under two-phase sampling and extended their work to stratified sampling designs, emphasizing the improved efficiency and robustness of their proposed methods over conventional estimators..

Singh and Khalid (2019) utilized ratio and product-type approaches to construct novel estimators for population variance under use of non-response. Kadilar and Cingi (2020) also contributed by designing new estimators that integrate multiple sampling strategies, examining their asymptotic behavior and practical applicability. Audu (2021) proposed Difference-cum-ratio- estimators for estimating the population coefficient of variation in SRSWOR. Nevertheless, the majority of the estimators developed in previous research was based on the premise of full response from all sampled units and were primarily unexamined under simple random sampling designs Methodology. Daraz et al. (2021) proposed estimators demonstrate remarkable performance, maintaining consistently lower MSE and higher efficiency than other established estimators, as verified by both theoretical and empirical evaluations. Ahmad et al. (2022) have improved the ratio-in-regression type variance estimator which is based on the dual use of the auxiliary variable. Daraz et al. (2024) Proposed double exponential ratio estimator of a finite population variance under SRSWOR. Adichwal (2017) proposed generalized class of estimators for the estimation of population variance by using auxiliary attributes. These studies have contributed to the theoretical enrichment of sampling methodologies, expanding their applicability under diverse population structures. Their contributions highlight the growing emphasis on designing estimators that are not only efficient under ideal conditions but also robust in the presence of data irregularities, thus laying a solid foundation for contemporary developments in survey sampling. Adejumobi et al. (2023) Develops a logarithmic ratio-type estimator for the population mean using auxiliary information under simple random sampling. The bias and MSE are derived, and results from three datasets show improved performance over existing estimators and Gupta et al. (2024) develops a logarithmic ratio-type estimator for the population mean. It derives bias and MSE and compares efficiency with existing estimators. Results from real data show improved performance of the proposed method

Most of the existing work in variance estimation under SRSWOR relies on traditional estimators like ratio, product, and regression forms that work effectively under simple linear relationships. However, their efficiency decreases when the underlying relationship is multiplicative or nonlinear. While logarithmic-type estimators exist, they are still restricted in adaptability, bias control, and efficiency across different data structures. The proposed estimators bridge this gap by offering improved efficiency, reduced mean squared error, and better performance for multiplicative relationships, as demonstrated through theoretical and empirical analysis.

A review of previous works indicates that many scholars have focused on developing efficient techniques for population variance estimation. Extending these efforts, the present study proposes an enhanced estimator that effectively utilizes auxiliary information also previous studies have shown that the use of auxiliary information can substantially enhance the precision of population parameter estimation. Motivated by this literature, the present research is designed to:

1) Develop an Enhanced logarithmic-type estimator for finite population variance under SRSWOR.
2) Examine its theoretical appropriateness, including bias and mean squared error, to the first order of approximation.
3) Demonstrate its applicability and performance using both empirical data and simulation experiments.
4) Conduct a comparative analysis with conventional estimators in terms of minimum MSE.
5) Provide an improved and efficient approach contributing to the ongoing development of population variance estimation techniques.

Adopting a structure consistent with established research practices, this paper is organized to ensure both conceptual clarity and methodological precision. The Introduction section 1 outlines the motivation behind the study and highlights the necessity of accurate population variance estimation in statistical research and section 1.1 elaborates the literature review. The Essential Notations and Descriptions section 2 presents foundational concepts and synthesize earlier work in the field. The Existing Estimator section 3 lays the groundwork for developing a novel estimator under the simple random sampling framework. Subsequently, the Proposed Estimator section 4 elaborates on the formulation and evaluates its efficiency in relation to existing estimators. The Efficiency Assessment section and Numerical Study sections demonstrate the estimator's performance through empirical and simulated analyses. The Simulation Study section 7 interprets the observed results, addresses methodological constraints, and proposes directions for further inquiry. Results and Their Implications section 8 shows the results ans implications and The paper concludes with the Discussion section 9, summarizing the key outcomes and offering final remarks on variance estimation within the SRSWOR context.

2. Essential Notations and Descriptions

Consider a finite population $\rho = \left(\rho_1, \rho_{2,}\rho_3, \dots, \rho_N\right)$ of size N units. Let y_i and x_i be the values of the study variable Y and the auxiliary variable X for the i^{th} units respectively. Let $\bar{Y} = \left(\frac{1}{N}\right)\sum_{i=1}^{N} Y_i$ and $\bar{X} = \left(\frac{1}{N}\right)\sum_{i=1}^{N} X_i$ be the population mean of the study and the auxiliary variable, respectively. It is further assumed that $S_y^2 = \frac{1}{(N-1)}\sum_{i=1}^{N}(Y_i - \bar{Y})^2$

and $S_x^2 = \frac{1}{(N-1)}\sum_{i=1}^{N}(X_i - \bar{X})^2$ be the population variances of the study as well as auxiliary variable, respectively. To estimate the unknown population parameter say $\bar{Y}$, we select a random sample of size n units from the population by using simple random sampling without replacement. Let $\bar{y} = \left(\frac{1}{n}\right)\sum_{i=1}^{n} y_i$ and $\bar{x} = \left(\frac{1}{n}\right)\sum_{i=1}^{n} x_i$ be the sample means of the study and the auxiliary variables, respectively, and their corresponding sample variances are $\hat{S}_y^2 = \frac{1}{(n-1)}\sum_{i=1}^{n}(y_i - \bar{y})^2$ and $\hat{S}_x^2 = \frac{1}{(n-1)}\sum_{i=1}^{n}(x_i - \bar{x})^2$ respectively.

To find bias and MSE for different estimators, Population variance we define the following terms.

$\vartheta_0 = \left(\frac{s_y^2 - S_y^2}{S_y^2}\right)$, $\vartheta_1 = \left(\frac{s_x^2 - S_x^2}{S_x^2}\right)$ and $\vartheta_2 = \left(\frac{\bar{x} - \bar{X}}{\bar{X}}\right)$, $E(\vartheta_i) = 0, i = 0,1,2$.

$E(e_0^2) = \theta\lambda_{40}^*$, $E(\vartheta_2^2) = \theta\lambda_{04}^*$, $E(\vartheta_1^2) = \theta C_x^2$

$E(\vartheta_0\vartheta_1) = \theta\lambda_{22}^*$, $E(\vartheta_0\vartheta_2) = \theta C_x\lambda_{21}$, $E(\vartheta_1\vartheta_2) = \theta C_x\lambda_{03}$,

Where, $\lambda_{40}^* = (\lambda_{40} - 1)$,$\lambda_{04}^* = (\lambda_{04} - 1)$, $\lambda_{22}^* = (\lambda_{22} - 1)$, $\theta = \left(\frac{1}{n} - \frac{1}{N}\right)$.Also

$$\lambda_{rs} = \frac{\mu_{rs}}{\mu_{20}^{\frac{r}{2}}\mu_{02}^{\frac{s}{2}}},$$

Where,

$$\mu_{rs} = \frac{\sum_{i=1}^{N}(Y_i - \bar{Y})^r\,(X_i - \bar{X})}{N-1}$$

Here $\lambda_{40}^* = \beta_{2(y)}$ and $\lambda_{04}^* = \beta_{2(x)}$ are the population coefficient of kurtosis.

3. Some Existing Estimators in Literature

3.1. Usual variance estimator

The usual estimator $\hat{S}_y^2$, is given by:

$$\hat{S}_y^2 = s_y^2$$

The usual variance estimator of $\hat{S}_y^2 = s_y^2$ for population variance is given by

$$\mathrm{Var}(\hat{S}_y^2) = \theta S_y^4\lambda_{40}^*. \quad (1)$$

3.2. Isaki estimator

Isaki (1983) developed a ratio-type estimator aimed at estimating the variance of the study variable Y. which is denoted by $\hat{S}_R^2$, and it is represented as follows

$$\hat{S}_R^2 = s_y^2\left(\frac{S_x^2}{s_x^2}\right),$$

Expressions for bias and MSE of $\hat{S}_R^2$, in simple random sampling (SRSWOR) are given by

$$\mathrm{Bias}\ (\hat{S}_R^2) \cong \theta S_y^4(\lambda_{04}^* - \lambda_{22}^*) \quad (2)$$

And

$$\mathrm{MSE}\ (\hat{S}_R^2) \cong \theta S_y^4(\lambda_{40}^* + \lambda_{04}^* - 2\lambda_{22}^*). \quad (3)$$

3.3. Bahl and tuteja estimator

Bahl and Tuteja (1991) developed exponential ratio-type estimator for the population variance of the study variable Y. Which is denoted by $\hat{S}_{BT}^2$ and it is represented as follows.

$$\hat{S}_{BTR}^2 = s_y^2\exp\left(\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2}\right)$$

The Bias and MSE corresponding to the estimator $\hat{S}_{BTR}^2$ is formulated as,

$$\mathrm{Bias}(\hat{S}_{BTR}^2) \cong \frac{1}{2}\theta S_y^2\left(\lambda_{40}^* + \frac{\lambda_{04}^*}{4} - \lambda_{22}^*\right) \quad (5)$$

And

$$\mathrm{MSE}(\hat{S}_{BTR}^2) \cong \theta S_y^2\left(\lambda_{40}^* + \frac{\lambda_{04}^*}{4} - \lambda_{22}^*\right) \quad (6)$$

3.4. Upadhyaya and singh estimator

Upadhyaya and Singh (1999) developed a ratio-type estimator $\hat{S}_{US}^2$ aimed at estimating population variance, that uses the kurtosis of an auxiliary variable in SRS is given by

$$\hat{S}_{US}^2 = s_y^2 x\left(\frac{S_x^2 + \lambda_{04}}{s_x^2 + \lambda_{04}}\right)$$

The Bias and MSE corresponding to the estimator $\hat{S}^2_{US}$ is formulated as,

$$\mathrm{Bias}(\hat{S}^2_{US}) \cong \theta S^2_y g_0(g_0\lambda^*_{04} - \lambda^*_{22}) \tag{7}$$

And

$$\mathrm{MSE}(\hat{S}^2_{US}) \cong \theta S^4_y(\lambda^*_{40} + g^2_0\lambda^*_{04} - 2g_0\lambda^*_{22}). \tag{8}$$

Where,

$$g_0 = \frac{S^2_x}{S^2_x + \lambda_{04}}$$

3.5. Kadilar and cingi estimator

Kadilar and Cingi (2006) developed a class of ratio estimators $\hat{S}^2_{KCi}$

$$\hat{S}^2_{KC1} = s^2_y\left(\frac{S^2_x + C_x}{s^2_x + C_x}\right)$$

$$\hat{S}^2_{KC2} = s^2_y\left(\frac{\lambda_{04}S^2_x + C_x}{\lambda_{04}s^2_x + C_x}\right)$$

$$\hat{S}^2_{KC3} = s^2_y\left(\frac{C_x S^2_x + \lambda_{04}}{C_x s^2_x + \lambda_{04}}\right)$$

Where $C_x = \frac{S_x}{\bar{X}}$ is the population coefficient of variation. The Bias and MSE corresponding to the estimator $\hat{S}^2_{KCi}$ is formulated as

$$\mathrm{Bias}(\hat{S}^2_{KCi}) \cong \theta S^2_y g_i(g_i\lambda^*_{04} - \lambda^*_{22}) \tag{9}$$

And

$$\mathrm{MSE}(\hat{S}^2_{KCi}) \cong \theta S^4_y(\lambda^*_{40} + g^2_i\lambda^*_{04} - 2g_i\lambda^*_{22}). \tag{10}$$

Where

$$g_1 = \frac{S^2_x}{S^2_x + C_x},\ g_2 = \frac{\lambda_{04}S^2_x}{\lambda_{04}S^2_x + C_x},\ g_3 = \frac{C_x S^2_x}{C_x S^2_x + \lambda_{04}}$$

4. Proposed Estimators of The Study

Motivated by the demonstrated effectiveness of auxiliary information in improving the estimation of the population mean Gupta et al.(2024) and Adejumobi et al.(2023) the present study extends this concept to the problem of population variance estimation under SRSWOR. In many practical survey situations, auxiliary information is not limited to a single parameter such as the mean, but may also include additional characteristics like variance, coefficient of variation, skewness, and other known population parameters. However, most existing variance estimators make only partial use of such information, which may lead to suboptimal efficiency. Recognizing this gap, the current research aims to enhance the estimation process by incorporating multiple auxiliary parameters into the estimator structure. By effectively utilizing this additional information, the proposed approach seeks to better capture the underlying population characteristics and improve estimation accuracy. Accordingly, three new population variance estimators are developed, each designed to progressively enhance performance through suitable modifications and parameter optimization. The proposed estimators are expected to yield reduced bias and lower mean squared error compared to traditional estimators, thereby providing a more reliable and efficient framework for variance estimation in practical survey applications.

4.1. Proposed estimator-I

$$t^*_{v1} = k\frac{\left[s^2_y - \ln\left\{\left(\frac{s^2_x}{S^2_x}\right)^{\alpha}\right\}\right]}{s^2_x}S^2_x \tag{11}$$

Here, α corresponds to the statistical constants, and k is a scalar that identified through optimization to ensure that the mean squared error (MSE) of t^*_{v1} is minimized.
By expanding Eq. (11) in terms of error components up to the first-order approximation, the following expression is obtained.

$$t^*_{v1} = k\frac{\left[S^2_y(1+\vartheta_0) - \ln\left\{\left(\frac{S^2_x(1+\vartheta_1)}{S^2_x}\right)^{\alpha}\right\}\right]}{S^2_x(1+\vartheta_1)}S^2_x \tag{12}$$

$$t^*_{v1} = k\left[S^2_y(1+\vartheta_0) - \alpha(1+\vartheta_1)\right](1+\vartheta_1)^{-1} \tag{13}$$

By expanding and simplifying the right-hand side of Eq. (13) and retaining terms up to the first-order approximation, the following expression is derived.

$$t_{v1}^{*} = k\left[S_y^2(1+\vartheta_0) - \alpha\left(\vartheta_1 - \frac{\vartheta_1^2}{2}\right)\right](1-\vartheta_1+\vartheta_1^2) \tag{14}$$

$$t_{v1}^{*} = k\left[S_y^2(1-\vartheta_1+\vartheta_1^2+\vartheta_0-\vartheta_0\vartheta_1) - \alpha\left(\vartheta_1 - \frac{3\vartheta_1^2}{2}\right)\right] \tag{15}$$

By subtracting S_y^2 (population variance) from each side of Eq. (15), the resulting expression is obtained as follows.

$$t_{v1}^{*} - S_y^2 = \left[(k-1)S_y^2 - kS_y^2\vartheta_1 + kS_y^2\vartheta_1^2 + kS_y^2\vartheta_0 - kS_y^2\vartheta_0\vartheta_1 - \alpha k\left(\vartheta_1 - \frac{3\vartheta_1^2}{2}\right)\right] \tag{16}$$

Taking the expectations of both sides of Eq. (16) and substituting the corresponding expected values, we derive the bias of (t_{v1}^{v}) as,

$$E\left(t_{v1}^{*} - S_y^2\right) = \begin{bmatrix} (k-1)S_y^2 - kS_y^2E(\vartheta_1) + kS_y^2E(\vartheta_1^2) + kS_y^2E(\vartheta_0) \\ -kS_y^2E(\vartheta_0\vartheta_1) - \alpha k\left(E(\vartheta_1) - \frac{3E(\vartheta_1^2)}{2}\right) \end{bmatrix} \tag{17}$$

$$\mathrm{Bias}(t_{v1}^{*}) = \left[(k-1)S_y^2 + kS_y^2\theta\lambda_{40}^{*} - kS_y^2\theta\lambda_{22}^{*} + \frac{3}{2}\alpha k\theta\lambda_{40}^{*}\right] \tag{18}$$

On the basis of Eq. (16), the following expression is obtained,

$$t_{v1}^{*} - S_y^2 = \left[(k-1)S_y^2 - kS_y^2\vartheta_1 + kS_y^2\vartheta_0 - \alpha k\vartheta_1\right] \tag{19}$$

Squaring both sides of Eq. (19) and taking the expectations, the mean squared error (MSE) of the estimator t_{v1}^{*} to the first-order approximation is obtained as,

$$\mathrm{MSE}(t_{v1}^{*}) = \left[(k-1)^2S_y^4 + k^2S_y^4\theta\lambda_{40}^{*} + \left(S_y^4 + \alpha^2 + 2\alpha S_y^2\right)k^2\theta\lambda_{40}^{*} - 2k^2S_y^2\left(S_y^2 + \alpha\right)\theta\lambda_{22}^{*}\right] \tag{20}$$

Or,

$$\mathrm{MSE}(t_{v1}^{*}) = \left[S_y^4 + Ak^2 - Bk\right] \tag{21}$$

Where,

$$A = \left[S_y^4 + S_y^4\theta\lambda_{40}^{*} + \left(S_y^4 + \alpha^2 + 2\alpha S_y^2\right)\theta\lambda_{04}^{*} - 2S_y^4\theta\lambda_{22}^{*} - 2S_y^2\alpha\theta\lambda_{22}^{*}\right].$$

$$B = 2S_y^4.$$

For the proposed estimator (t_{v1}^{*}), the MSE is minimized when the constants are set to the following optimal values,

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

By substituting the optimal value of k_{opt}, the minimum MSE of the estimator can be expressed as,

$$\min.\mathrm{MSE}(t_{v1}^{*}) = \left[S_y^4 - \frac{B^2}{4A}\right] \tag{22}$$

4.2. Proposed estimator-II

$$t_{v2}^{*} = s_y^2\left\{\frac{\ln(S_x^2)}{\ln(s_x^2)}\right\}^{\alpha} \tag{23}$$

Where,

$$\ln(S_x^2) \neq 0 \text{ and } \ln(s_x^2) \neq 0$$

Where, α corresponds to the optimizing constant for which MSE of the proposed estimator t_{v3}^{*} found to be minimum.
By applying the methodology used for the first estimator t_{v1}^{*} to determine their Bias and MSE, and incorporating the approximations described in the notation section, we derive the corresponding expressions for Estimator t_{v2}^{*} as:
Expressing Eq. (23) in terms of error components and retaining terms up to the first-order approximation yields the following expression:

$$t_{v2}^{*} = S_y^2(1+\vartheta_0)\left\{\frac{\ln(S_x^2)}{\ln\left(S_x^2(1+\vartheta_1)\right)}\right\}^{\alpha} \tag{24}$$

$$t_{v3}^{*} = S_y^2(1+\vartheta_0)\{1 + k\ln(1+\vartheta_1)\}^{-\alpha} \tag{25}$$

By expanding and simplifying the right-hand side of Eq. (25), and keeping terms up to the first-order approximation, the resulting expression is obtained as:

$$t_{v3}^{*}=S_y^2(1+\vartheta_0)\left\{1+k\left(\vartheta_1-\frac{\vartheta_1^2}{2}\right)\right\}^{-\alpha} \tag{26}$$

$$t_{v3}^{*}=S_y^2(1+\vartheta_0)\left\{1-\alpha k\vartheta_1+\alpha k\frac{\vartheta_1^2}{2}+\alpha^2k^2\frac{\vartheta_1^2}{2}+\alpha k^2\frac{\vartheta_1^2}{2}\right\} \tag{27}$$

By subtracting S_y^2 (population variance) from each side of Eq. (27), the resulting expression is obtained as follows.

$$t_{v3}^{*}-S_y^2=\begin{bmatrix}S_y^2\vartheta_0-\alpha kS_y^2\vartheta_1-\alpha kS_y^2\vartheta_0\vartheta_1+\alpha kS_y^2\frac{\vartheta_1^2}{2}+\alpha^2k^2S_y^2\frac{\vartheta_1^2}{2}+\\ \alpha k^2S_y^2\frac{\vartheta_1^2}{2}\end{bmatrix} \tag{28}$$

Taking the expectations of both sides of Eq. (28) and substituting the corresponding expected values, we derive the bias as,

$$E\left(t_{v2}^{*}-S_y^2\right)=\begin{bmatrix}S_y^2E(\vartheta_0)-\alpha kS_y^2E(\vartheta_1)-\alpha kS_y^2E(\vartheta_0\vartheta_1)+\alpha kS_y^2\frac{E(\vartheta_1^2)}{2}\\ +\alpha^2k^2S_y^2\frac{E(\vartheta_1^2)}{2}+\alpha k^2S_y^2\frac{E(\vartheta_1^2)}{2}\end{bmatrix} \tag{29}$$

$$\mathrm{Bias}(t_{v3}^{\nu})=\left[S_y^2\theta\left(\frac{\alpha}{2}k^2\lambda_{04}^{*}+\frac{\alpha^2}{2}k^2\lambda_{04}^{*}+\frac{\alpha k}{2}\lambda_{04}^{*}-k\alpha\lambda_{22}^{*}\right)\right] \tag{30}$$

On the basis of Eq. (27), the following expression is obtained,

$$t_{v2}^{*}-S_y^2=\left[S_y^2(\vartheta_0-\alpha k\vartheta_1)\right] \tag{31}$$

Squaring both sides of Eq. (31) and taking the expectations, the mean squared error (MSE) of the estimator t_{v2}^{*} to the first-order approximation is obtained as,

$$\left(t_{v2}^{*}-S_y^2\right)^2=\left[S_y^2(\vartheta_0-\alpha k\vartheta_1)\right]^2 \tag{32}$$

$$\left(t_{v2}^{*}-S_y^2\right)^2=S_y^4(\vartheta_0^2+\alpha^2k^2\vartheta_1^2-2\alpha k\vartheta_0\vartheta_1) \tag{33}$$

$$\mathrm{MSE}(t_{v2}^{*})=\left[S_y^4\theta(\lambda_{40}^{*}+\alpha^2k^2\lambda_{04}^{*}-2\alpha k\lambda_{22}^{*})\right]. \tag{34}$$

Where $\theta=\left(\frac{1}{n}-\frac{1}{N}\right)$
Differentiate above equation with respect to α_{opt} and equate it equal to zero, we get

$$\alpha=\frac{\rho C_y}{kC_x}=\alpha_{opt}\ \text{(Say)} \tag{35}$$

$$\min.\mathrm{MSE}(t_{v2}^{*})=\left[\theta S_y^4\lambda_{40}^{*}(1-\rho^2)\right]$$

4.3. Proposed estimator-III

$$t_{v3}^{*}=\ s_y^2\left[k_1\left\{\frac{S_x^2}{s_x^2}\right\}^{\alpha}+k_2\left\{\frac{\ln(S_x^2)}{\ln(s_x^2)}\right\}^{\beta}\right] \tag{36}$$

Where, $\ln(S_x^2)\neq 0$ and $\ln(s_x^2)\neq 0$.
Here, α and β corresponds to the statistical constants and k_1 and k_2 are scalars that identified through optimization to ensure that the mean squared error (MSE) of t_{v3}^{*} is minimized.
By applying the methodology used for the first two estimators (t_{v1}^{*} and t_{v2}^{*}) to determine their Bias and MSE, and incorporating the approximations described in the notation section, we derive the corresponding expressions for Estimator t_{v3}^{*} as:
By expanding Eq. (36) in terms of error components up to the first-order approximation, the following expression is obtained.

$$t_{v3}^{*}=S_y^2(1+\vartheta_0)\left[k_1\left(\frac{S_x^2}{S_x^2(1+\vartheta_1)}\right)^{\alpha}+k_2\left\{\frac{\ln(S_x^2)}{\ln\left(S_x^2(1+\vartheta_1)\right)}\right\}^{\beta}\right] \tag{37}$$

$$t_{v3}^{*}=S_y^2(1+\vartheta_0)\left[k_1(1+\vartheta_1)^{-\alpha}+k_2\{1+\theta\ln(1+\vartheta_1)\}^{-\beta}\right] \tag{38}$$

By expanding and simplifying the right-hand side of Eq. (38) and retaining terms up to the first-order approximation, the following expression is derived.

$$t_{v3}^{*}=S_y^2(1+\vartheta_0)\left[k_1\left(1-\alpha\vartheta_1+\frac{\alpha(\alpha+1)}{2}\vartheta_1^2\right)+k_2\left\{1-\beta\theta\left(\vartheta_1+\frac{\vartheta_1^2}{2}\right)+\frac{\beta(\beta+1)}{2}\theta^2\vartheta_1^2\right\}\right] \tag{39}$$

$$t_{v3}^{*}=S_y^2(1+\vartheta_0)\begin{bmatrix}k_1-\alpha k_1\vartheta_1+\frac{\alpha(\alpha+1)}{2}\vartheta_1^2k_1+k_2-k_2\beta\theta\left(\vartheta_1+\frac{\vartheta_1^2}{2}\right)+\\ k_2\frac{\beta(\beta+1)}{2}\theta^2\vartheta_1^2\end{bmatrix} \tag{40}$$

By subtracting S_y^2 (population variance) from each side of Eq. (40), the resulting expression is obtained as follows

$$t_{v3}^* - S_y^2 = \begin{bmatrix} S_y^2(k_1 + k_2 - 1) + S_y^2 k_1 \vartheta_0 - S_y^2 \alpha k_1 \vartheta_0 \vartheta_1 + \frac{\alpha^2}{2} k_1 \vartheta_1^2 + \frac{\alpha}{2} k_1 \vartheta_1^2 + \\ k_2 S_y^2 \vartheta_0 + k_2 \beta\theta S_y^2 \vartheta_1 - k_2 \beta\theta S_y^2 \frac{\vartheta_1^2}{2} + k_2 \beta\theta S_y^2 \vartheta_0 \vartheta_1 + S_y^2 \theta^2 \frac{\beta^2}{2} \vartheta_1^2 + \\ S_y^2 \theta^2 \frac{\beta}{2} \vartheta_1^2 \end{bmatrix} \tag{41}$$

Taking the expectations of both sides of Eq. (41) and substituting the corresponding expected values, we derive the bias of (t_{v3}^*) as,

$$E(t_{v3}^* - S_y^2) = \begin{bmatrix} S_y^2(k_1 + k_2 - 1) + S_y^2 k_1 E(\vartheta_0) - S_y^2 \alpha k_1 E(\vartheta_0 \vartheta_1) + \\ \frac{\alpha^2}{2} k_1 E(\vartheta_1^2) + \frac{\alpha}{2} k_1 E(\vartheta_1^2) + k_2 S_y^2 E(\vartheta_0) + k_2 \beta\theta S_y^2 E(\vartheta_1) - \\ k_2 \beta\theta S_y^2 \frac{E(\vartheta_1^2)}{2} + k_2 \beta\theta S_y^2 E(\vartheta_0 \vartheta_1) + S_y^2 \theta^2 \frac{\beta^2}{2} E(\vartheta_1^2) + \\ S_y^2 \theta^2 \frac{\beta}{2} E(\vartheta_1^2) \end{bmatrix} \tag{42}$$

$$\mathrm{Bias}(t_{v3}^*) = \left[S_y^2 \begin{pmatrix} (k_1 + k_2 - 1) - \alpha k_1 \theta \lambda_{22}^* + \frac{\alpha^2}{2} k_1 \theta \lambda_{04}^* + \frac{\alpha}{2} k_1 \theta \lambda_{04}^* - \beta\theta^2 k_2 \lambda_{22}^* \\ -\beta\theta^2 k_2 \frac{\lambda_{04}^*}{2} + \frac{\beta^2}{2} \theta^3 k_2 \lambda_{04}^* + \frac{\beta}{2} \theta^3 k_2 \lambda_{04}^* \end{pmatrix} \right] \tag{43}$$

On the basis of Eq. (41), the following expression is obtained,

$$t_{v3}^* - S_y^2 = \left[S_y^2 \{ (k_1 + k_2 - 1) + (\vartheta_0 - \alpha \vartheta_1) k_1 + (\vartheta_0 - \beta\theta \vartheta_1) k_2 \} \right] \tag{44}$$

Squaring both sides of Eq. (44) and taking the expectations, the mean squared error (MSE) of the estimator t_{v3}^* to the first-order approximation is obtained as,

$$\left(t_{v3}^* - S_y^2\right)^2 = \left[S_y^2 \{ (k_1 + k_2 - 1) + (\vartheta_0 - \alpha \vartheta_1) k_1 + (\vartheta_0 - \beta\theta \vartheta_1) k_2 \} \right]^2 \tag{45}$$

$$\left(t_{v3}^* - S_y^2\right)^2 = S_y^4 \begin{bmatrix} k_1^2 + k_2^2 - 2k_1 k_2 - 2k_1 - 2k_2 + k_1^2(\vartheta_0^2 + \alpha^2 \vartheta_1^2 - 2\beta \vartheta_0 \vartheta_1) \\ +k_2^2(\vartheta_0^2 + \beta^2 \vartheta_1^2 - 2\alpha \vartheta_0 \vartheta_1) + 2k_1(k_1 + k_2 - 1)(\vartheta_0 - \alpha \vartheta_1) \\ +2k_1 k_2 (\vartheta_0 - \alpha \vartheta_1)(\vartheta_0 - \beta\theta \vartheta_1) + 2(k_1 + k_2 - 1)(\vartheta_0 - \beta\theta \vartheta_1) \\ 2k_1 k_2 (\vartheta_0^2 - \beta \vartheta_0 \vartheta_1 - \alpha \vartheta_0 \vartheta_1 + \alpha\beta \vartheta_1^2) \end{bmatrix} \tag{46}$$

$$\left(t_{v3}^* - S_y^2\right)^2 = S_y^4 \begin{bmatrix} k_1^2 + k_2^2 - 2k_1 k_2 - 2k_1 - 2k_2 + k_1^2(\vartheta_0^2 + \alpha^2 \vartheta_1^2 - 2\beta \vartheta_0 \vartheta_1) \\ +k_2^2(\vartheta_0^2 + \beta^2 \vartheta_1^2 - 2\alpha \vartheta_0 \vartheta_1) + 2k_1(k_1 + k_2 - 1)(\vartheta_0 - \alpha \vartheta_1) \\ +2k_1 k_2 (\vartheta_0 - \alpha \vartheta_1)(\vartheta_0 - \beta\theta \vartheta_1) + 2(k_1 + k_2 - 1)(\vartheta_0 - \beta\theta \vartheta_1) \\ 2k_1 k_2 (\vartheta_0^2 - \beta \vartheta_0 \vartheta_1 - \alpha \vartheta_0 \vartheta_1 + \alpha\beta \vartheta_1^2) \end{bmatrix} \tag{47}$$

$$\mathrm{MSE}(t_{v3}^*) = \begin{bmatrix} S_y^4 - 2k_1 S_y^4 - 2k_2 S_y^4 + k_1^2 S_y^4 (\theta \lambda_{40}^* + \alpha^2 \theta \lambda_{04}^* - 2\alpha\theta \lambda_{22}^* + 1) + \\ k_2^2 S_y^4 (\theta \lambda_{40}^* + \beta^2 \theta^3 \lambda_{04}^* - 2\beta\theta^2 \lambda_{22}^* + 1) + 2k_1 k_2 S_y^4 \\ (\theta \lambda_{40}^* - \beta\theta^2 \lambda_{22}^* - \alpha\theta \lambda_{22}^* + \alpha\beta\theta^2 \lambda_{04}^* + 1) \end{bmatrix} \tag{48}$$

Or,

$$\mathrm{MSE}(t_{v3}^*) = [A k_1^2 + B k_2^2 + 2C k_1 k_2 - D k_1 - E k_2 + F] \tag{49}$$

Where,

$A = S_y^4 (\theta \lambda_{40}^* + \alpha^2 \theta \lambda_{04}^* - 2\alpha\theta \lambda_{22}^* + 1)$

$B = S_y^4 (\theta \lambda_{40}^* + \beta^2 \theta^3 \lambda_{04}^* - 2\beta\theta^2 \lambda_{22}^* + 1)$

$C = S_y^4 (\theta \lambda_{40}^* - \beta\theta^2 \lambda_{22}^* - \alpha\theta \lambda_{22}^* + \alpha\beta\theta^2 \lambda_{04}^* + 1)$

$D = 2S_y^4$

$E = 2S_y^4$

$F = S_y^4$

For the proposed estimator (t_{v3}^*), the MSE is minimized when the constants are set to the following optimal values,

$$k_{1(opt)} = \frac{BD - EC}{2(AB - C^2)}$$

And,

$$k_{2(\mathrm{opt})} = \frac{EA-DC}{2(AB-C^2)}$$

Putting the value of $k_{1(\mathrm{opt})}$ and $k_{2(\mathrm{opt})}$ in the Eq. (49) we get the minimum MSE.

$$\mathrm{MSE}(t_{v3}^{*}) = \begin{bmatrix} A\left(\frac{BD-EC}{2(AB-C^2)}\right)^2 + B\left(\frac{EA-DC}{2(AB-C^2)}\right)^2 + 2C\left(\frac{BD-EC}{2(AB-C^2)}\right) \\ \left(\frac{EA-DC}{2(AB-C^2)}\right) - D\left(\frac{BD-EC}{2(AB-C^2)}\right) - E\left(\frac{EA-DC}{2(AB-C^2)}\right) + F \end{bmatrix} \tag{50}$$

$$\mathrm{Min.\,MSE}(t_{v3}^{*}) = S_y^4\left[1 - \frac{(BD^2+AE^2-2CDE)}{4(AB-C^2)}\right] \tag{51}$$

Or,

$$\mathrm{Min.\,MSE}(t_{v3}^{*}) = S_y^4\left[1 - \frac{S}{T}\right] \tag{52}$$

Where,

$$S = BD^2 + AE^2 - 2CDE$$

$$T = 4(AB - C^2)$$

5. Theoretical Efficiency Comparison Analysis

In this part, we analytically derive and compare the efficiencies of the proposed estimators against the existing ones. The efficiency of an estimator signifies its relative precision in estimating the population parameter, where a higher efficiency corresponds to a lower mean squared error (MSE), indicating greater reliability. Moreover, robustness has been assessed through an extensive simulation study as well as a real data set analysis. The simulation experiments demonstrate that the estimator maintains good performance under varying conditions. Similarly, the results obtained from the real data set further support its effectiveness in practical applications. Together, these findings confirm the consistent and stable behavior of the proposed estimator.

5.1. Comparison with the first estimator

- Comparative analysis of the proposed logarithmic estimator t_{v1}^{*} with the sample variance $\hat{S}_y^2$:

$$\mathrm{Var}(\hat{S}_y^2) - \mathrm{min.\,MSE}(t_{v1}^{v}) > 0.$$

Or,

$$\theta S_y^4 \lambda_{40}^{*} - \left[S_y^4 - \frac{B^2}{4A}\right] > 0. \tag{53}$$

- Evaluating the performance of the proposed logarithmic estimator t_{v1}^{*} in relation to the $\hat{S}_R^2$:

$$\mathrm{MSE}\ (\hat{S}_R^2) - \mathrm{min.\,MSE}(t_{v1}^{*}) > 0.$$

Or,

$$[\theta S_y^4(\lambda_{40}^{*} + \lambda_{04}^{*} - 2\lambda_{22}^{*})] - \left[S_y^4 - \frac{B^2}{4A}\right] > 0. \tag{54}$$

- An evaluation is made to contrast the proposed logarithmic estimator t_{v1}^{*} with Bahl and Tuteja $\hat{S}_{BTR}^2$ estimator:

$$\mathrm{MSE}(\hat{S}_{BTR}^2) - \mathrm{min.\,MSE}(t_{v1}^{*}) > 0.$$

Or,

$$\left[\theta S_y^2\left(\lambda_{40}^{*} + \frac{\lambda_{04}^{*}}{4} - \lambda_{22}^{*}\right)\right] - \left[S_y^4 - \frac{B^2}{4A}\right] > 0. \tag{55}$$

- The effectiveness of the proposed logarithmic estimator t_{v1}^{*} is analyzed relative to Upadhyaya and Singh $\hat{S}_{US}^2$ estimator:

$$\mathrm{MSE}(\hat{S}_{US}^2) - \mathrm{min.\,MSE}(t_{v1}^{*}) > 0.$$

Or,

$$[\theta S_y^4(\lambda_{40}^{*} + g_0^2\lambda_{04}^{*} - 2g_0\lambda_{22}^{*})] - \left[S_y^4 - \frac{B^2}{4A}\right] > 0. \tag{56}$$

- Comparison of the proposed logarithmic estimator and the t^*_{v1} Kadilar and Cingi ($\hat{S}^2_{cki}$), estimator:

$$\mathrm{MSE}\left(\hat{S}^2_{cki}\right) - \min.\mathrm{MSE}(t^*_{v1}) > 0.$$

Or,

$$\left[\theta S_y^4(\lambda^*_{40} + (g_i)^2\lambda^*_{04} - 2(g_i)\lambda^*_{22})\right] - \left[S_y^4 - \frac{B^2}{4A}\right] > 0. i = 1,2,3 \quad (57)$$

5.2. Comparison with the second estimator

- Comparative analysis of the proposed logarithmic estimator t^*_{v2} with the sample variance $\hat{S}^2_y$:

$$\mathrm{Var}\left(\hat{S}^2_y\right) - \min.\mathrm{MSE}(t^*_{v2}) > 0.$$

Or,

$$\theta S_y^4\lambda^*_{40} - \left[\theta S_y^4\lambda^*_{40}(1 - \rho^2)\right] > 0. \quad (58)$$

- Evaluating the performance of the proposed logarithmic estimator t^*_{v2} in relation to the $\hat{S}^2_R$:

$$\mathrm{MSE}\ \left(\hat{S}^2_R\right) - \min.\mathrm{MSE}(t^*_{v2}) > 0.$$

Or,

$$\left[\theta S_y^4(\lambda^*_{40} + \lambda^*_{04} - 2\lambda^*_{22})\right] - \left[\theta S_y^4\lambda^*_{40}(1 - \rho^2)\right] > 0. \quad (59)$$

- An evaluation is made to contrast the proposed logarithmic estimator t^*_{v2} with Bahl and Tuteja $\hat{S}^2_{BTR}$ estimator:

$$\mathrm{MSE}\left(\hat{S}^2_{BTR}\right) - \min.\mathrm{MSE}(t^*_{v2}) > 0.$$

Or,

$$\Big[\theta S_y^2 \Big(\lambda^*_{40} + \frac{\lambda^*_{04}}{4} - \lambda^*_{22}\Big)\Big)\Big] - \left[\theta S_y^4\lambda^*_{40}(1 - \rho^2)\right] > 0. \quad (60)$$

- The effectiveness of the proposed logarithmic estimator t^*_{v2} is analyzed relative to Upadhyaya and Singh $\hat{S}^2_{US}$ estimator:

$$\mathrm{MSE}\left(\hat{S}^2_{US}\right) - \min.\mathrm{MSE}(t^*_{v2}) > 0.$$

Or,

$$\left[\theta S_y^4(\lambda^*_{40} + g_0^2\lambda^*_{04} - 2g_0\lambda^*_{22})\right] - \left[\theta S_y^4\lambda^*_{40}(1 - \rho^2)\right] > 0. \quad (61)$$

- Comparison of the proposed logarithmic estimator and the t^*_{v2} Kadilar and Cingi (kci) estimator:

$$\mathrm{MSE}\left(\hat{S}^2_{cki}\right) - \min.\mathrm{MSE}(t^*_{v2}) > 0.$$

Or,

$$\left[\theta S_y^4(\lambda^*_{40} + (g_i)^2\lambda^*_{04} - 2(g_i)\lambda^*_{22})\right] - \left[\theta S_y^4\lambda^*_{40}(1 - \rho^2)\right] > 0. i = 1,2,3 \quad (62)$$

5.3. Comparison with the third estimator

- Comparative analysis of the proposed logarithmic estimator t^*_{v3} with the sample variance $\hat{S}^2_y$:

$$\mathrm{Var}\left(\hat{S}^2_y\right) - \min.\mathrm{MSE}(t^*_{v3}) > 0.$$

Or,

$$\theta S_y^4\lambda^*_{40} - \left[S_y^4 - \frac{B^2}{4A}\right] > 0. \quad (63)$$

- Evaluating the performance of the proposed logarithmic estimator t^*_{v3} in relation to the $\hat{S}^2_R$:

$$\mathrm{MSE}\ \left(\hat{S}^2_R\right) - \min.\mathrm{MSE}(t^*_{v3}) > 0.$$

Or,

$$\left[\theta S_y^4(\lambda_{40}^* + \lambda_{04}^* - 2\lambda_{22}^*)\right] - \left[S_y^4 - \frac{B^2}{4A}\right] > 0. \quad (64)$$

- An evaluation is made to contrast the proposed logarithmic estimator t_{v3}^* with Bahl and Tuteja $\hat{S}_{BTR}^2$ estimator:

$$MSE(\hat{S}_{BTR}^2) - \min.MSE(t_{v3}^*) > 0.$$

Or,

$$\left[\theta S_y^2\left(\lambda_{40}^* + \frac{\lambda_{04}^*}{4} - \lambda_{22}^*\right)\right] - \left[S_y^4 - \frac{B^2}{4A}\right] > 0. \quad (65)$$

- The effectiveness of the proposed logarithmic estimator t_{v3}^* is analyzed relative to Upadhyaya and Singh $\hat{S}_{US}^2$ estimator:

$$MSE(\hat{S}_{US}^2) - \min.MSE(t_{v3}^*) > 0.$$

Or,

$$\left[\theta S_y^4(\lambda_{40}^* + g_0^2\lambda_{04}^* - 2g_0\lambda_{22}^*)\right] - \left[S_y^4 - \frac{B^2}{4A}\right] > 0 \quad (66)$$

- Comparison of the proposed logarithmic estimator and the t_{v3}^* Kadilar and Cingi (kc1) estimator:

$$MSE(\hat{S}_{ck1}^2) - \min.MSE(t_{v3}^*) > 0.$$

Or,

$$\left[\theta S_y^4(\lambda_{40}^* + (g_i)^2\lambda_{04}^* - 2(g_i)\lambda_{22}^*)\right] - \left[S_y^4 - \frac{B^2}{4A}\right] > 0. i = 1,2,3 \quad (67)$$

6. Empirical Investigation

To evaluate the performance of the proposed estimators, three real datasets were utilized.
A comparative analysis of Mean Squared Errors (MSEs) was conducted among different estimators. The datasets were carefully selected to ensure variability and reliability.
Their detailed descriptions and summary statistics are presented below:
Data set 1 :(Source Cochran (1963))
Y: Food expenses related to the family's employment and
X: Families weekly income.

Table 1: The Following Are the Summary Statistics

Parameter	Symbol	Value
Population size	N	33
Sample size	n	5
Mean of X	$\bar{X}$	72.55
Mean of Y	$\bar{Y}$	27.49
Maximum value of X	X_M	95
Minimum value of X	X_m	58
Standard deviation of X	S_x	10.58
Standard deviation of Y	S_y	10.13
Coefficient of variation of X	C_x	0.15
Coefficient of variation of Y	C_y	0.37
Correlation coefficient of Y and X	ρ_{yx}	0.25
Fourth-order moment of X	λ_{40}	3.25
Fourth-order moment of Y	λ_{04}	3.80
Third-order mixed moment	λ_{03}	1.00
Second-order mixed moment (2,1)	λ_{21}	2.91
Second-order mixed moment (2,2)	λ_{22}	2.22

Data Set 2: (Source Subramani and Kumarapandiyan (2012)).
Y: Total amount of recyclable-waste collection in Italy (2003).
X: Total amount of recyclable-waste collection in Italy (2002).

Table 2: The Following Are the Summary Statistics

Parameter	Symbol	Value
Population size	N	103
Sample size	N	40
Mean of X	$\bar{X}$	62.62
Mean of Y	$\bar{Y}$	556.55
Standard deviation of X	S_x	91.35
Standard deviation of Y	S_y	610.16
Correlation coefficient of Y and X	ρ_{yx}	0.72
Fourth-order moment of X	λ_{40}	37.12

Fourth-order moment of Y	λ_{04}	17.87
Second-order mixed moment (2,2)	λ_{22}	17.22

The following equations were employed to determine the PRE for the proposed as well as other estimators concerning the standard sample mean.

$$\text{PRE(Estimators)} = \frac{\text{Var}(\hat{S}_y^2)}{\text{MSE(Existing and Proposed estimators)}} \times 100$$

- $$\text{PRE}\left(\hat{S}_R^2, \hat{S}_y^2\right) = \frac{\text{Var}(\hat{S}_y^2)}{\text{MSE}\ (\hat{S}_R^2)} \times 100 = \frac{\theta S_y^4 \lambda_{40}^*}{[\theta S_y^4(\lambda_{40}^* + \lambda_{04}^* - 2\lambda_{22}^*)]} \times 100$$
- $$\text{PRE}\left(\hat{S}_{bt}^2, \hat{S}_y^2\right) = \frac{\text{Var}(\hat{S}_y^2)}{\text{MSE}(\hat{S}_{bt}^2)} \times 100 = \frac{\theta S_y^4 \lambda_{40}^*}{\left[\theta S_y^2\left(\lambda_{40}^* + \frac{\lambda_{04}^*}{4} - \lambda_{22}^*\right)\right]} \times 100$$
- $$\text{PRE}\left(\hat{S}_{US}^2, \hat{S}_y^2\right) = \frac{\text{Var}(\hat{S}_y^2)}{\text{MSE}(\hat{S}_{US}^2)} \times 100 = \frac{\theta S_y^4 \lambda_{40}^*}{[\theta S_y^4(\lambda_{40}^* + g_0^2\lambda_{04}^* - 2g_0\lambda_{22}^*)]} \times 100$$
- $$\text{PRE}\left(\hat{S}_{ck1}^2, \hat{S}_y^2\right) = \frac{\text{Var}(\hat{S}_y^2)}{\text{MSE}(\hat{S}_{ck1}^2)} \times 100 = \frac{\theta S_y^4 \lambda_{40}^*}{\left[\theta S_y^4\left(\lambda_{40}^* + \left(\frac{S_X^2}{S_X^2 + C_X}\right)^2 \lambda_{04}^* - 2\left(\frac{S_X^2}{S_X^2 + C_X}\right)\lambda_{22}^*\right)\right]} \times 100$$
- $$\text{PRE}\left(\hat{S}_{ck2}^2, \hat{S}_y^2\right) = \frac{\text{Var}(\hat{S}_y^2)}{\text{MSE}(\hat{S}_{ck2}^2)} \times 100 = \frac{\theta S_y^4 \lambda_{40}^*}{\left[\theta S_y^4\left(\lambda_{40}^* + \left(\frac{\lambda_{04} S_X^2}{\lambda_{04} S_X^2 + C_X}\right)^2 \lambda_{04}^* - 2\left(\frac{\lambda_{04} S_X^2}{\lambda_{04} S_X^2 + C_X}\right)\lambda_{22}^*\right)\right]} \times 100$$
- $$\text{PRE}\left(\hat{S}_{ck3}^2, \hat{S}_y^2\right) = \frac{\text{Var}(\hat{S}_y^2)}{\text{MSE}(\hat{S}_{ck3}^2)} \times 100 = \frac{\theta S_y^4 \lambda_{40}^*}{\left[\theta S_y^4\left(\lambda_{40}^* + \left(\frac{C_X S_X^2}{C_X S_X^2 + \lambda_{04}}\right)^2 \lambda_{04}^* - 2\left(\frac{C_X S_X^2}{C_X S_X^2 + \lambda_{04}}\right)\lambda_{22}^*\right)\right]} \times 100$$
- $$\text{PRE}\left(t_{v1}^*, \hat{S}_y^2\right) = \frac{\text{Var}(\hat{S}_y^2)}{\text{MSE}(t_{v1}^*)} \times 100 = \frac{\theta S_y^4 \lambda_{40}^*}{\left[S_y^4 - \frac{B^2}{4A}\right]} \times 100$$
- $$\text{PRE}\left(t_{v2}^*, \hat{S}_y^2\right) = \frac{\text{Var}(\hat{S}_y^2)}{\text{MSE}(t_{v2}^*)} \times 100 = \frac{\theta S_y^4 \lambda_{40}^*}{[\theta S_y^4 \lambda_{40}^*(1 - \rho^2)]} \times 100$$
- $$\text{PRE}\left(t_{v3}^*, \hat{S}_y^2\right) = \frac{\text{Var}(\hat{S}_y^2)}{\text{MSE}(t_{v3}^*)} \times 100 = \frac{\theta S_y^4 \lambda_{40}^*}{\left[S_y^4 - \frac{B^2}{4A}\right]} \times 100$$

Table 3: MSE and PRE Values Using Empirical Values:

Estimator	Data set-1		Data set-2	
	MSE	PRE	MSE	PRE
$\hat{S}_y^2$	4020.63	100	38461417	100
$\hat{S}_R^2$	4663.93	86.20	21882119	175.76
$\hat{S}_{bt}^2$	3091.42	130.05	25680876	149.76
$\hat{S}_{US}^2$	4529.11	88.77	21882052	175.76
$\hat{S}_{ck1}^2$	4656.39	86.34	21882114	175.76
$\hat{S}_{ck2}^2$	4661.94	86.24	21882118	175.76
$\hat{S}_{kc3}^2$	3959.72	101.53	21882058	175.76
t_{v1}^*	3258.87	123.37	16650171	230.99
t_{v2}^*	3070.74	130.93	21855451	175.98
t_{v3}^* $(\alpha = 1, \beta = 1)$	2377.44	169.11	16634631	231.21

MSE: Mean Squared Error ; PRE: Percentage Relative Efficiency

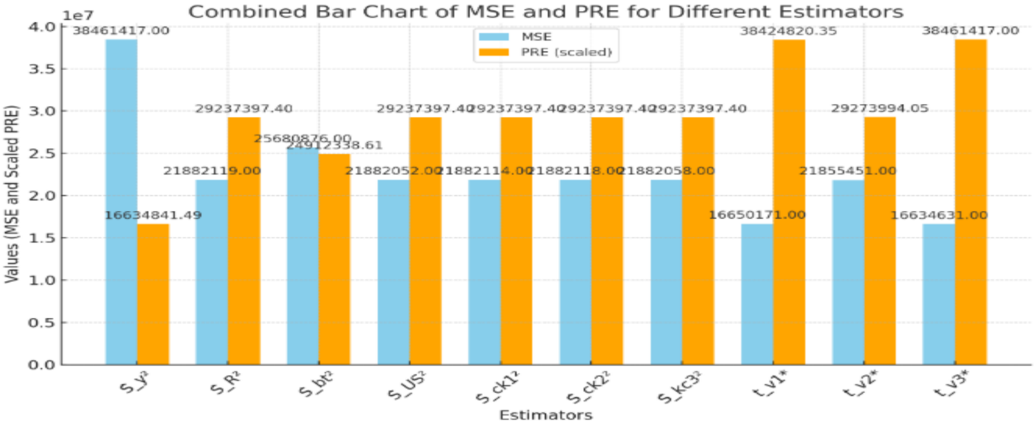

Fig. 1: Bar chart which depicted the MSE and PRE values along with the existing estimators for Data set-1

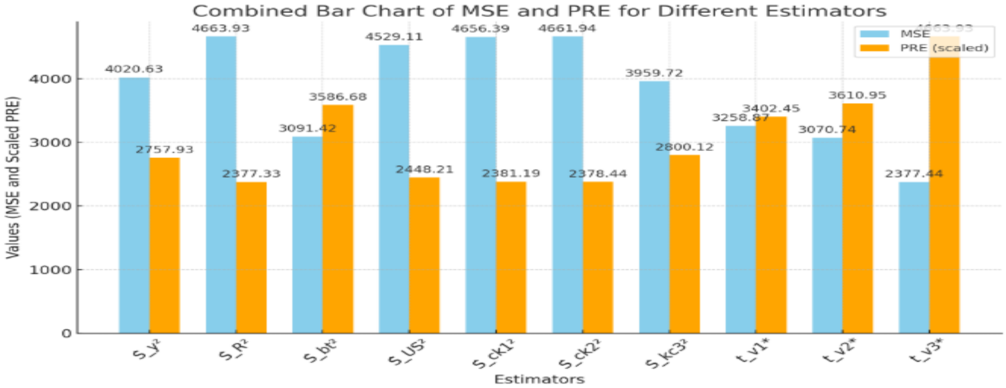

Fig. 2: Bar Chart Which Depicted the MSE and PRE Values Along with the Existing Estimators for Data Set-2.

7. Simulation-Based Comparative Study

A simulation study was conducted to evaluate the performance of the proposed estimators in terms of their relative efficiency and mean squared error. The procedure of the simulation followed the steps outlined below.

1) We simulated N= 2000 observations from a Bivariate Normal distribution with the mean vector $\mu = [8,7]$ and $\sigma_x = 10$,$\sigma_y = 11$ and with correlation coefficients $\rho_{xy} = 0.6, 0.7, 0.8, 0.9$.

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

2) Sample of sizes n=200 and 400 are drawn using simple random sampling without replacement (SRSWOR) and varying coefficients is 0.6, 0.7, 0.8, 0.9.
3) A total 20,000 repetitions are performed to compute the MSE of each estimator based on their corresponding mathematical formulations.
4) The mean values across 20,000 replications for different correlation levels and sample sizes.
5) Tables 4 and 5 provide the simulated estimates corresponding to the proposed estimators.

Table 4: MSE Values Across Different Correlation Values and Sample Sizes Using Artificially Generated Data for Sample Size (N=200)

Sample Size	n=200			
ρ	0.6	00.70000	0.8	0.9
Estimators				
$\hat{S}_y^2$	92.6121	92.7265	92.2088	92.1126
$\hat{S}_R^2$	132.9218	104.7643	74.6211	64.5299
$\hat{S}_{bt}^2$	86.9214	72.6357	63.1498	57.9921
$\hat{S}_{US}^2$	84.6215	102.4761	75.8765	38.9936
$\hat{S}_{ck1}^2$	128.9213	104.7643	72.5503	39.2234
$\hat{S}_{ck2}^2$	131.4286	74.5460	72.7702	38.1143
$\hat{S}_{kc3}^2$	127.6454	73.7890	64.1179	16.9925
t_{v1}^*	77.2303	73.6213	64.3836	16.7233
t_{v2}^*	77.1121	73.5062	64.2217	16.0592
t_{v3}^*	77.0026	73.3125	64.1125	16.0029

Table 5: MSE Values Across Different Correlation Values and Sample Sizes Using Artificially Generated Data for Sample Size (N=400)

Sample Size	n=400			
ρ	0.6	00.70000	0.8	0.9
Estimators				
$\hat{S}_y^2$	88.76	88.60	88.56	88.21
$\hat{S}_R^2$	102.65	100.98	98.36	92.24
$\hat{S}_{bt}^2$	99.98	92.91	94.73	89.61
$\hat{S}_{US}^2$	92.62	63.82	55.62	45.31
$\hat{S}_{ck1}^2$	89.62	55.92	45.79	36.93
$\hat{S}_{ck2}^2$	88.60	55.62	43.77	35.37
$\hat{S}_{kc3}^2$	87.59	55.45	43.70	34.21
t_{v1}^*	45.86	37.62	33.23	14.24
t_{v2}^*	44.62	37.52	33.21	14.02
t_{v3}^*	44.98	37.41	33.19	14.00

8. Results and Their Implications

The outcomes of the empirical investigation reported in Table 3 clearly indicate that the suggested logarithmic estimator's exhibits improved efficiency over the usual mean estimator, Isaki (1983) Estimator, Bahl and Tuteja (1991) Estimator, Upadhyaya and Singh (1999) Estimator, Kadilar and Cingi (2006) Estimator. The simulation results summarized in Table 4 and Table 5 indicate that the developed estimators surpass those considered in Section 3 for every level of correlation coefficient and for all examined sample sizes. Based on the results, the estimator t_{v3}^* $(\alpha = 1, \beta = 1)$ shows the best overall performance, achieving the lowest MSE in data set-1 and the highest PRE indicating strong efficiency. In data set-2, t_{v1}^* and t_{v3}^* perform almost similarly, with very close PRE values (230.99 and 231.21, respectively). The estimator t_{v2}^* performs relatively less efficiently compared to the other proposed estimators. t_{v3}^* Overall, can be considered the most efficient and stable estimator across both datasets. Furthermore, the simulation results presented in Table 4 reveal that the Mean Squared Error (MSE) tends to decline with an increase in sample size. It is also evident that MSE decreases as the correlation coefficient values become higher. In addition, a bar chart has been constructed to represent the empirical investigation of the MSE and PRE values. In order to examine the efficiency pattern and robustness of the introduced logarithmic type estimators for the estimation of population variance, different levels of correlation ($\rho_{xy} = 0.6,0.7,0.8,0.9$) and sample size (n= 200, 400) were considered. The corresponding Mean Squared Error (MSE) values derived from the simulation experiment are provided in Table 4 and Table 5. In addition, a bar chart has been constructed to provide a clear graphical representation of the findings, allowing easier comparison and interpretation of the simulated values of the estimator.

9. Discussion

This study introduces a new class of logarithmic ratio-type estimators for the estimation of finite population variance, demonstrating superior efficiency and precision over traditional estimation methods. Theoretical expressions for the bias and mean squared error (MSE) of the proposed estimators have been derived up to the first-order approximation. Their performance has been evaluated through both empirical analysis and simulation experiments based on artificially generated bivariate normal populations. The obtained results consistently reveal that the proposed estimators produce lower MSE values and improved accuracy compared with existing estimators in the literature. In addition, a bar chart representing the empirical MSE values and PRE values have been presented to visualize and validate the superiority of the proposed estimators under simple random sampling without replacement.

The logarithmic type estimator is based on the assumption that both the study variable Y and auxiliary variable X take positive values and that there exists a strong multiplicative or log-linear relationship between them, often represented as, $\log(Y) = a + b\log(X) + \epsilon$. It also assumes that simple random sampling (usually SRSWOR) is used, the auxiliary variable is known accurately for the population, and the error structure in the transformed (log) scale is well-behaved with roughly constant variance. However, its main limitations are that it cannot be applied when data contain zero or negative values, it may become inefficient or biased if the true relationship is not log-linear, and it can introduce back-transformation bias when converting results to the original scale. Additionally, it is sensitive to very small values, requires a highly correlated auxiliary variable to be effective, and is less intuitive to interpret compared to standard linear estimators like ratio or regression estimators. Logarithmic type estimators are widely applied in real-life survey work such as agricultural yield estimation, economic household surveys, health-related studies, and large national sample surveys where variables have a multiplicative relationship. Future research is directed toward making these estimators more robust to non-positive data, improving performance under complex sampling structures, addressing measurement errors, and combining classical survey theory with modern computational methods like Bayesian and machine learning approaches.

References

[1] C. T. Isaki, “Variance estimation using auxiliary information,” *J. Amer. Stat. Assoc.*, vol. 78, no. 381, pp. 117–123, 1983, https://doi.org/10.1080/01621459.1983.10477939.

[2] S. Bahl and R. K. Tuteja, “Ratio and product type exponential estimators,” *J. Inf. Optim. Sci.*, vol. 12, no. 1, pp. 159–164, Jan. 1991, https://doi.org/10.1080/02522667.1991.10699058.

[3] M. R. Garcia and A. A. Cebrian, “Repeated substitution method: The ratio estimator for the population variance,” *Metrika*, vol. 43, no. 1, pp. 101–105, Dec. 1996, https://doi.org/10.1007/BF02613900.

[4] L. N. Upadhyaya and H. P. Singh, “Use of transformed auxiliary variable in estimating the finite population mean,” *Biometrical J.*, vol. 41, no. 5, pp. 627–636, 1999. https://doi.org/10.1002/(SICI)1521-4036(199909)41:5<627::AID-BIMJ627>3.3.CO;2-N.

[5] L. N. Upadhyaya, H. P. Singh, and S. Singh, “A class of estimators for estimating the variance of the ratio estimator,” *J. Jpn. Stat. Soc.*, vol. 34, no. 1, pp. 47–63, 2004. https://doi.org/10.14490/jjss.34.47.

[6] P. Chandra and H. P. Singh, “A family of estimators for population variance using knowledge of kurtosis of an auxiliary variable in sample survey,” *Stat. Transit.*, vol. 7, no. 1, pp. 27–34, 2005.

[7] A. Arcos, M. Rueda, M. D. Martínez, S. González, and Y. Román, “Incorporating the auxiliary information available in variance estimation,” *Appl. Math. Comput.*, vol. 160, no. 2, pp. 387–399, 2005. https://doi.org/10.1016/j.amc.2003.11.010.

[8] C. Kadilar and H. Cingi, “New ratio estimators using correlation coefficient,” *InterStat,* vol. 4, pp. 1–11, Mar. 2006.

[9] H. P. Singh and R. S. Solanki, “A new procedure for variance estimation in simple random sampling using auxiliary information,” *Stat. Papers*, vol. 54, no. 2, pp. 479–497, May 2013, https://doi.org/10.1007/s00362-012-0445-2.

[10] S. K. Yadav, C. Kadilar, J. Shabbir, and S. Gupta, “Improved family of estimators of population variance in simple random sampling,” *J. Stat. Theory Pract.*, vol. 9, no. 2, pp. 219–226, Apr. 2015, https://doi.org/10.1080/15598608.2013.856359.

[11] J. Shabbir and S. Gupta, “Some estimators of finite population variance of stratified sample mean,” *Commun. Stat. Theory Methods*, vol. 39, no. 16, pp. 3001–3008, Aug. 2010, https://doi.org/10.1080/03610920903170384.

[12] G. N. Singh and M. Khalid, “Effective estimation strategy of population variance in two-phase successive sampling under random non-response,” *J. Stat. Theory Pract.*, vol. 13, no. 1, p. 4, Mar. 2019, https://doi.org/10.1007/s42519-018-0010-y.

[13] H. O. Cekim and C. Kadilar, “ln-type variance estimators in simple random sampling,” *Pak. J. Stat. Oper. Res.*, pp. 689–696, Dec. 2020, https://doi.org/10.18187/pjsor.v16i4.3072.

[14] A. Audu, M. D. Kareem, J. N. Eze, and U. Daraz, “Difference-cum-ratio estimators for estimating finite population coefficient of variation in simple random sampling,” *Asian J. Probab. Stat.*, vol. 13, no. 3, pp. 13–29, 2021. https://doi.org/10.9734/ajpas/2021/v13i330308.

[15] U. Daraz and M. Khan, “Estimation of variance of the difference-cum-ratio-type exponential estimator in simple random sampling,” *Res. Math. Stat.*, vol. 8, no. 1, p. 1899402, Jan. 2021, https://doi.org/10.1080/27658449.2021.1899402.

[16] S. Ahmad, M. Khan, U. Daraz, and M. Nawaz, “A simulation study: Improved ratio-in-regression type variance estimator based on dual use of auxiliary variable under simple random sampling,” *PLoS One*, vol. 17, no. 11, p. e0276540, 2022. https://doi.org/10.1371/journal.pone.0276540.

[17] U. Daraz, J. Wu, and O. Albalawi, “Double exponential ratio estimator of a finite population variance under extreme values in simple random sampling,” *Mathematics*, vol. 12, p. 1737, 2024. https://doi.org/10.3390/math12111737

[18] N. K. Adichwal, P. Sharma, and R. Singh, “Generalized class of estimators for population variance using information on two auxiliary variables,” *Int. J. Appl. Comput. Math.*, vol. 3, no. 2, pp. 651–661, Jun. 2017, https://doi.org/10.1007/s40819-015-0119-6.

[19] R. Gupta, S. K. Yadav, M. Ali, and G. Kumar, “Improved estimation of the population mean by integrating additional auxiliary parameters,” *Korean J. Physiol. Pharmacol.*, vol. 28, no. 1, pp. 205–210, Jan. 2024, https://doi.org/10.36106/ijar/7104032.

[20] A. Adejumobi, M. . Abiodun 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, vol. 3, no. 2, pp. 700–705, May 2023.

[21] R. A. Robb, “W. G. Cochran, *Sampling Techniques*, John Wiley & Sons, 1963, ix + 413 pp., 72 s.,” *Proc. Edinb. Math. Soc.*, vol. 13, no. 4, pp. 342–343, 1963. https://doi.org/10.1017/S0013091500025724.

[22] J. Subramani and G. Kumarapandiyan, “Variance estimation using quartiles and their functions of an auxiliary variable,” *Ratio J.*, vol. 1, no. 1, 2012. https://doi.org/10.5923/j.ijps.20120103.02.