Maejo Int. J. Sci. Technol. **2026**, *20*(01), 91-107

Maejo International
Journal of Science and Technology

e-ISSN 2697-4746
Available online at www.mijst.mju.ac.th

Full Paper

Efficient class of hybrid estimators for population variance in two-phase sampling

Iqra Niaz [1], Aamir Sanaullah [2], Sat Gupta [3], Iram Saleem [4], Prayas Sharma [5,*] and Mohammad M. A. Almazah [6,7]

[1] Government Associate College for Women, Heir Bedian, Lahore, Pakistan
[2] COMSATS University Islamabad, Lahore Campus, Pakistan
[3] University of North Carolina, Greensboro, USA
[4] Forman Christian College Lahore, Pakistan
[5] Department of Statistics, Babasaheb Bhimrao Ambedkar University Lucknow, India
[6] Department of Mathematics and Computer, College of Sciences, Ibb University, 70270, Yemen
[7] Department of Mathematics, College of Sciences and Arts (Muhyil), King Khalid University, Muhyil 61421, Saudi Arabia
* Corresponding author, e-mail: prayassharma02@gmail.com

Received: 17 May 2025 / Accepted: 28 March 2026 / Published: 7 April 2026

Abstract: Auxiliary information is used to estimate population parameters and enhance the efficiency of estimators. However, when such information is partially or entirely unavailable, or when surveying a large sample is prohibitively expensive, two-phase sampling becomes practical. Unlike previous approaches, this study introduces a generalised hybrid estimator that integrates ratio, product and exponential-type strategies to achieve consistently lower mean squared errors. This unified framework extends existing variance estimation methods in two-phase sampling and demonstrates clear efficiency gains over conventional estimators. Therefore, this study aims at addressing the problem of variance estimation under such a context. This study suggests a class of efficient hybrid estimators for estimating finite population variance in two-phase sampling. The bias and mean squared error expressions are derived up to the first order of approximation and compared with those of existing estimators. A special case of the proposed class of estimators is also discussed. The performance of the proposed estimator is evaluated empirically using two real-world data sets. A simulation study is conducted using three bivariate normal populations with varying correlations between the auxiliary and study variables. To evaluate estimator performance, theoretical percentage relative efficiencies are computed. The results of both simulation and empirical studies demonstrate significant efficiency gains of the proposed estimator over existing approaches.

Keywords: auxiliary variable, two-phase sampling, ratio estimator, generalised hybrid estimator, variance estimation, mean square error

INTRODUCTION

Collecting basic information from all elements in a sample, along with additional data from a subset, is often feasible and economically advantageous. This approach of data collection involving two samples is recognised as double sampling or two-phase sampling, which is usually advised when countless elements is needed to achieve the desired accuracy across highly varying entries. It also helps improve the accuracy of data gathered in the second phase. Although population values are approximated using first-phase sample values, which may be prone to errors, the estimation techniques in two-phase sampling are like those used in single-phase sampling. Two-phase sampling designs were introduced by Neyman [1]. Rao [2] adopted this design to deal with the problems related to stratification and non-response. Cochran [3] and Tabassum and Khan [4, 5] formulated several estimates employing two-phase sampling. Auxiliary information is efficaciously utilised in two-phase sampling to estimate population characteristics. Bose [6], Patterson [7] and Yates [8] examined the application of the regression estimation method in two-phase sampling. Hanif et al. [9] and Hammad et al. [10] proposed regression-type estimators for situations where auxiliary information is available in the second-phase sample. Bahl and Tuteja [11] proposed exponential-type estimators under simple random sampling without replacement, which were later modified by Singh et al. [12] under two-phase sampling. Sanaullah et al. [13] proposed several generalised exponential estimators for stratified two-phase sampling. Recent developments under double sampling include efficient classes of estimators for the population mean proposed. Bhushan et al. [14] proposed a generalised class of estimators for the population mean under double sampling using auxiliary information, demonstrating how suitably chosen characterising parameters can unify several classical estimators. This construction strategy motivates the generalised framework adopted in the present variance-estimation study. In a recent contribution, Sher et al. [15] developed hybrid-type estimators for the finite population mean under two-phase sampling by combining ratio, product and exponential forms, highlighting the effectiveness of hybrid structures under auxiliary information, an idea conceptually aligned with the generalised hybrid variance estimator proposed in the present study.

Variability is observed across many real-world problems of societal importance. For instance, an agriculturist may require detailed information on climatic variations to decide where and when to plant crops. Similarly, an industrialist must continuously monitor the variation in consumer responses to a product to determine whether its features need improvement. Therefore, it is often essential to practice techniques that result in reduced variation and improved quality. Isaki [16] suggested the ratio estimator for population variance. Prasad and Singh [17] suggested an enhanced ratio-type estimator to estimate population variance based on Isaki's estimator. Subsequent research has produced a variety of ratio, exponential and hybrid estimators for finite population variance under different sampling designs. For example, Yadav and Kadilar [18] proposed an exponential ratio-type estimator; Asghar et al. [19] developed generalised exponential-type estimators; Sanaullah et al. [20] presented a class of efficient hybrid-type estimators; and Niaz et al. [21] introduced improved estimator classes targeting reduced mean squared error (MSE). Recent developments emphasise optimal use of auxiliary variables and robustification against model departures. Grover [22] clarified corrections that reduce bias and MSE when auxiliary information is incorporated while Shabbir and Gupta [23] extended theoretical foundations for auxiliary-based variance estimators. Subramani and Kumarapandiyan [24] demonstrated the utility of quartiles and related functions of auxiliary variables to construct robust estimators across diverse sampling situations.

Sanaullah et al. [25] introduced a generalised exponential function class of estimators utilising the mean of auxiliary variables for estimating population variance. Yaqub and Shabbir [26] contributed further work in this area with methodologies addressing population variance estimation. Asghar et al. [27] developed a multivariate regression-cum-exponential estimator designed specifically for vector population variance estimation. Extensive recent research has explored the use of auxiliary information to enhance estimation of finite population variance, particularly within single- and two-phase sampling frameworks. Muneer et al. [28] proposed an improved ratio–product exponential estimator that integrates auxiliary information to reduce bias and MSE. Al-Marshadi et al. [29] developed estimators for single- and two-phase designs that exploit multiple auxiliary variables while Abid et al. [30] constructed innovative variance estimators based on midrange and inter-decile range summaries of auxiliary data. Singh and Khalid [31] examined variance estimation under nonresponse using two-phase, two-occasion sampling and provided a detailed analysis of the properties of their estimator. Yasmeen et al. [32] introduced a generalised exponential estimator leveraging two transformed auxiliary variables, and Sharma and Singh [33] proposed a difference-type imputation procedure to mitigate random nonresponse in two-phase variance estimation. Cekim and Kadilar [34] contributed ln-type estimators under simple random sampling and Gulzar et al. [35] advanced estimators grounded in alternative location measures such as the tri-mean, Hodges-Lehmann estimator and decile mean. Shahzad et al. [36-38] further improved variance estimation through calibration estimators that incorporate L-moments and auxiliary information.

Additional contributions include formulation of an efficient estimator that systematically exploits auxiliary variables by Bhushan et al. [39], and a family of estimators by Shahbaz et al. [40] for single- and two-phase sampling that utilise both single and multiple auxiliary variables to estimate general population parameters. Asghar et al. [41] addressed estimation of the population variance vector when multiple auxiliary variables are available. Daraz et al. [42] proposed a novel class of estimators that employ extreme values and ranks of an auxiliary variable while Hussain et al. [43] developed an optimal family of exponential variance estimators validated on both real and simulated data sets. Audu et al. [44] investigated variance estimation in successive sampling under various response scenarios and presented logarithmic-type estimators. More rexcently, Daraz et al. [45] introduced double exponential-type estimators that incorporate outliers and ranks from auxiliary variables. Recently, Basit and Bhatti [46] developed efficient classes of estimators for population variance under two-phase successive sampling in the presence of random non-response, extending variance estimation theory to rotation designs. However, their framework differs from the present study, which considers single-occasion two-phase sampling under complete response. More recently, Daraz et al. [47] introduced variance estimators under two-phase sampling by incorporating ranks and extreme values of auxiliary variables, providing an alternative auxiliary-information mechanism. In contrast, the present study develops a generalised hybrid variance estimator without relying on rank-based or extreme-value information.

Recent methodological advances in survey estimation have focused on exploiting auxiliary information and on developing robust functional transformations to improve estimator efficiency. Ratio and exponential-type estimators for means under two-phase sampling have demonstrated notable efficiency gains when auxiliary variables are appropriately transformed [48]. Complementary strategies, including non-parametric regression estimators augmented by probability-weighted moments, have been proposed to increase robustness against outliers and model misspecification [49]. These developments motivate extending hybrid and adaptive estimation frameworks to the problem of population variance estimation under two-phase sampling.

Although several authors have applied auxiliary information to improving estimation of population parameters, the bulk of the literature emphasises mean estimation and considers a limited family of estimator forms. In particular, recent contributions under two-phase sampling primarily focus on population mean estimation, e.g. those from Bhushan et al. [14] and Sher et al. [15], whereas variance estimation under comparable designs has received relatively limited attention. Existing contributions introduce ratio, product and exponential strategies under simple and two-phase designs, typically targeting bias reduction and MSE minimisation through functional transformations of auxiliary variables. Such methodological structures provide a natural foundation for generalising hybrid approaches to variance estimation by combining complementary transformation components within a unified class.

Many studies address single-phase designs, assume complete response or restrict attention to a narrow selection of auxiliary variables. Despite these advances, extant work on variance estimation remains limited in scope with many studies confined to single-phase designs, successive or rotation sampling frameworks, or settings involving non-response rather than single-occasion two-phase sampling under complete response. Recent improvements reported by Al-Marshadi et al. [29], Abid et al. [30] and Daraz et al. [42, 45, 47] demonstrate efficiency gains under specific assumptions but stop short of proposing a generalised hybrid architecture that encapsulates multiple estimation strategies. For example, Basit and Bhatti [46] considered variance estimation under two-phase successive sampling and random non-response. Similarly, Daraz et al. [42, 45, 47] exploited rank and extreme-value information of auxiliary variables. These studies, although effective within their respective frameworks, do not develop a generalised hybrid class that simultaneously integrates ratio, product and exponential components under a single-occasion two-phase design. Prior investigations have also seldom undertaken analytical optimisation of estimator parameters to attain minimum MSE, instead relying predominantly on ad hoc or purely numerical tuning.

The present paper proposes an efficient class of hybrid estimators for population variance tailored to two-phase sampling. The proposed class integrates ratio, product and exponential components through a small number of tuning constants, yielding a flexible family that nests many existing estimators as special cases. For this class we derive first-order expressions for bias and MSE under the two-phase sampling design and identify conditions under which asymptotic efficiency is achieved. Where closed-form minimisation of MSE is attainable, we provide explicit analytic formulae for optimal tuning constants and where analytic solutions are intractable, we develop stable numerical procedures for parameter optimisation.

The objective of this investigation is to develop an efficient and flexible class of generalised hybrid-type, auxiliary information-based estimators for population variance in two-phase sampling and to evaluate their efficiency. Another objective is to derive closed-form expressions for the bias and MSE of the proposed class of hybrid estimators. Subsequently, the proposed class of hybrid estimators $u_{prp(d)}$ is compared systematically with other existing estimators. Both empirical and simulation studies demonstrate that the proposed hybrid estimators outperform several existing estimators including u_0'', $u_{r(d)}$, $u_{re(d)}$, $u_{reg(d)}$, $u_{d(d)}$ and $u_{de(d)}$.

NOTATIONS AND SOME EXISTING ESTIMATORS UNDER TWO-PHASE SAMPLING

Consider a finite population of N distinct units. Let y be the study variable and x the auxiliary variables. Let $\bar{Y}=\frac{1}{N}\sum_{i=1}^{N} y_i$ and $\bar{X}=\frac{1}{N}\sum_{i=1}^{N} x_i$ be the population means, and $S_y^2=\frac{1}{N-1}\sum_{i=1}^{N}(y_i-\bar{Y})^2$ and $S_x^2=\frac{1}{N-1}\sum_{i=1}^{N}(x_i-\bar{X})^2$ be the population variances of y and x respectively. Some basic notations used in the subsequent sections are defined below.

n': first-phase sample size,

n'': sub-sample \ second-phase sample size,

$\bar{x}'=\frac{1}{n'}\sum_{i=1}^{n'} x_i$: sample mean estimator for x based on n' units,

$s_x^{2\prime}=\frac{1}{n'-1}\sum_{i=1}^{n'}(x_i-\bar{x}')^2$: sample variance estimator for x based on n' units,

$\bar{y}''=\frac{1}{n''}\sum_{i=1}^{n''} y_i$ and $\bar{x}''=\frac{1}{n''}\sum_{i=1}^{n''} x_i$: sample mean estimator for y based on n' units,

$s_y^{2\prime\prime}=\frac{1}{n''-1}\sum_{i=1}^{n''}(y_i-\bar{y}'')^2$ and $s_x^{2\prime\prime}=\frac{1}{n''-1}\sum_{i=1}^{n''}(x_i-\bar{x}'')^2$: sample variance estimators for y and x respectively, based on n'' units.

Two-phase sampling is preferred when the required parameters of an auxiliary variable (e.g. $\bar{X}, S_x^2$) are not available in advance. In the first-phase a relatively large sample of size $n' \subset N$ is taken using simple random sampling without replacement. In the second-phase a sub-sample of size $n'' \subset n'$ is taken from the first-phase sample.

Consider notions of the error terms:

$e_0''=\frac{s_y^{2\prime\prime}-S_y^2}{S_y^2}$, $e_1'=\frac{s_x^{2\prime}-S_x^2}{S_x^2}$ and $e_1''=\frac{s_x^{2\prime\prime}-S_x^2}{S_x^2}$ such that $E(e_0'')=E(e_1')=E(e_1'')=0$;

$E(e_0''^2)=\gamma''(\varphi_{40}-1)$, $E(e_1'^2)=\gamma'(\varphi_{04}-1)$, $E(e_1''^2)=\gamma''(\varphi_{04}-1)$;

$E(e_0''e_1')=\gamma'(\varphi_{22}-1)$, $E(e_0''e_1'')=\gamma''(\varphi_{22}-1)$ and $E(e_1'e_1'')=\gamma'(\varphi_{04}-1)$,

where

$\gamma'=\left(\frac{1}{n'}-\frac{1}{N}\right)$, $\gamma''=\left(\frac{1}{n''}-\frac{1}{N}\right)$, $\varphi_{rs}=\frac{\mu_{rs}}{\mu_{20}^{r/2}\mu_{02}^{s/2}}$, and $\mu_{rs}=\frac{1}{N}\sum_{i=1}^{N}(Y_i-\bar{Y})^r(X_i-\bar{X})^s$.

The usual unbiased estimator of the population variance is given by

$$u_0''=s_y^{2\prime\prime}. \tag{1}$$

The variance of u_0'' is given by

$$Var(u_0'')=\gamma'' S_y^4(\varphi_{40}-1). \tag{2}$$

Grover [22] discussed the two-phase version of the ratio estimator of the population variance by Isaki [16]. The ratio estimator takes the form

$$u_{r(d)}=s_y^{2\prime\prime}\frac{s_x^{2\prime}}{s_x^{2\prime\prime}} \tag{3}$$

and its MSE up to the first order of approximation is

$$MSE(u_{r(d)})=S_y^4[\gamma''(\varphi_{40}-1)+(\gamma''-\gamma')\{(\varphi_{04}-1)-2(\varphi_{22}-1)\}]. \tag{4}$$

We then define a product estimator in two-phase sampling as

$$u_{p(d)}=s_y^{2\prime\prime}\frac{s_x^{2\prime\prime}}{s_x^{2\prime}} \tag{5}$$

Maejo Int. J. Sci. Technol. **2026**, *20*(01), 91-107

and its MSE up to the first order of approximation is

$$MSE(u_{p(d)}) = S_y^4[\gamma''(\varphi_{40} - 1) + (\gamma'' - \gamma')\{(\varphi_{04} - 1) + 2(\varphi_{22} - 1)\}]. \tag{6}$$

Singh et al. [12] proposed ratio and product exponential-type estimators. Modified forms of these estimators in two-phase sampling are given by

$$u_{re(d)} = s_y^{2''} \exp\left[\frac{s_x^{2'} - s_x^{2''}}{s_x^{2'} + s_x^{2''}}\right] \tag{7}$$

and

$$u_{pe(d)} = s_y^{2''} \exp\left[\frac{s_x^{2''} - s_x^{2'}}{s_x^{2''} + s_x^{2'}}\right]. \tag{8}$$

The MSE of $u_{re(d)}$ and $u_{pe(d)}$ respectively are given up to the first order of approximation by

$$MSE(u_{re(d)}) = S_y^4\left[\gamma''(\varphi_{40} - 1) + (\gamma'' - \gamma')\left\{\frac{(\varphi_{04} - 1)}{4} - (\varphi_{22} - 1)\right\}\right] \tag{9}$$

and

$$MSE(u_{pe(d)}) = S_y^4\left[\gamma''(\varphi_{40} - 1) + (\gamma'' - \gamma')\left\{\frac{(\varphi_{04} - 1)}{4} + (\varphi_{22} - 1)\right\}\right]. \tag{10}$$

Grover [22] proposed a difference-type estimator in two-phase sampling for the population variance. The form of the estimator is given by

$$u_{d(d)} = k_1 s_y^{2''} + k_2(s_x^{2'} - s_x^{2''}). \tag{11}$$

The estimator $u_{d(d)}$ performs better than the usual regression estimator $u_{reg(d)}$. The minimum MSE of $u_{d(d)}$ up to the first order of approximation is given by

$$MSE(u_{d(d)})_{min} = S_y^4\left[1 - \frac{1}{1 + \left\{\gamma''(\varphi_{40} - 1) - (\gamma'' - \gamma')\frac{(\varphi_{22} - 1)^2}{(\varphi_{04} - 1)}\right\}}\right]. \tag{12}$$

Grover [22] adapted the variance estimator in two-phase sampling of Shabbir and Gupta [23]. The estimator is given by

$$u_{de(d)} = \left[k_1 s_y^{2''} + k_2(s_x^{2'} - s_x^{2''})\right] \exp\left(\frac{s_x^{2'} - s_x^{2''}}{s_x^{2'} + s_x^{2''}}\right). \tag{13}$$

The minimum MSE of $u_{de(d)}$ up to the first order of approximation is given by

$$MSE(u_{de(d)})_{min} = MSE(u_{reg}) - \frac{S_y^4\left[(\gamma'' - \gamma')(\varphi_{04} - 1) + 8\left(\gamma''(\varphi_{40} - 1) - (\gamma'' - \gamma')\frac{(\varphi_{22} - 1)^2}{(\varphi_{04} - 1)}\right)\right]^2}{64\left\{1 + \left\{\gamma''(\varphi_{40} - 1) - (\gamma'' - \gamma')\frac{(\varphi_{22} - 1)^2}{(\varphi_{04} - 1)}\right\}\right\}}. \tag{14}$$

It is obvious from the above equations that the estimator $u_{de(d)}$ performs better than $u_{reg(d)}$, so $u_{de(d)}$ will also perform better than the estimators u_0'', $u_{r(d)}$ and $u_{re(d)}$ [22].

PROPOSED METHODOLOGY OF VARIANCE ESTIMATION

Singh et al. [12] suggested a ratio-product exponential-type estimator for population variance in two-phase sampling. It is given by

$$u_{rpe(d)} = s_y^{2''}\left[\alpha \exp\left[\frac{s_x^{2'}-s_x^{2''}}{s_x^{2'}+s_x^{2''}}\right] + (1-\alpha)\exp\left[\frac{s_x^{2''}-s_x^{2'}}{s_x^{2'}+s_x^{2''}}\right]\right]. \tag{15}$$

When α =1, the proposed estimator of Eq. (15) reduces to the exponential-type ratio estimator, and for α =0, it becomes the exponential-type product estimator [12].

The minimum MSE of $u_{rpe(d)}$ at the optimum value $\alpha_{opt} = \frac{1}{2} + \frac{(\varphi_{22}-1)}{(\varphi_{04}-1)}$ is given by

$$MSE\left(u_{rpe(d)}\right)_{min} = S_y^4\left[\gamma''(\varphi_{40}-1) - (\gamma''-\gamma')\frac{(\varphi_{22}-1)^2}{(\varphi_{04}-1)}\right]. \tag{16}$$

It corresponds to the following regression estimator's MSE [22]:

$$u_{reg(d)} = s_y^{2''} + b\left(s_x^{2'} - s_x^{2''}\right), \qquad \text{where } b = \frac{s_y^{2''}(\varphi_{22}-1)}{s_x^{2''}(\varphi_{04}-1)}.$$

It is therefore clear that the estimator of Singh et al. [12] is not superior to the regression estimator by Grover [22]. We can note that substituting α =0.5 into Eq. (15) gives an exponential-type ratio-product estimator:

$$u_{rpe(d1)} = \frac{s_y^2}{2}\left[\exp\left[\frac{s_x^{2'}-s_x^{2''}}{s_x^{2'}+s_x^{2''}}\right] + \exp\left[\frac{s_x^{2''}-s_x^{2'}}{s_x^{2''}+s_x^{2'}}\right]\right]. \tag{17}$$

Based on Eq. (17) and Grover [22], we propose a hybrid estimator for population variance in two-phase sampling as given below:

$$u_{dce(d)} = \left[\frac{s_y^{2''}}{2}\left\{\exp\left(\frac{s_x^{2'}-s_x^{2''}}{s_x^{2'}+s_x^{2''}}\right) + \exp\left(\frac{s_x^{2''}-s_x^{2'}}{s_x^{2''}+s_x^{2'}}\right)\right\} + k_1\left(s_x^{2'} - s_x^{2''}\right) + k_2 s_y^{2''}\right]\exp\left(\frac{s_x^{2'}-s_x^{2''}}{s_x^{2'}+s_x^{2'}}\right). \tag{18}$$

The minimum MSE of $u_{dce(d)}$ up to the first order of approximation $O(n^{-1})$ is given by

$$MSE\left(u_{dce(d)}\right)_{min} = MSE\left(u_{reg(d)}\right) - \frac{S_y^4\left[\gamma''\left\{(\varphi_{40}-1)-(\gamma''-\gamma')\frac{(\varphi_{22}-1)^2}{(\varphi_{04}-1)}\right\}+\frac{1}{4}(\gamma''-\gamma')(\varphi_{04}-1)^2\right]}{\left\{1+\left\{\gamma''(\varphi_{40}-1)-(\gamma''-\gamma')\frac{(\varphi_{22}-1)^2}{(\varphi_{40}-1)}\right\}\right\}}. \tag{19}$$

From Eq. (14) and (19) it is evident that the proposed hybrid estimator $u_{dce(d)}$ has the least MSE among other estimators discussed earlier and can also perform better than the regression estimator as well as the estimator of Grover [22].

Motivated by the performance of the estimator in Eq. (18), we propose a more generalised and flexible class of hybrid estimators for estimating population variance in two-phase sampling. It is given by

$$u_{prp(d)} = \left(s_y^{2''}\left(\alpha \exp\left(\frac{\mathrm{s_x^{2'}} - \mathrm{s_x^{2''}}}{\mathrm{s_x^{2'}} + \mathrm{s_x^{2''}}}\right) + (1-\alpha)\exp\left(\frac{\mathrm{s_x^{2''}} - \mathrm{s_x^{2'}}}{\mathrm{s_x^{2'}} + \mathrm{s_x^{2''}}}\right)\right) + k_{1(d)}\left(\mathrm{s_x^{2'}} - \mathrm{s_x^{2''}}\right) + \mathrm{k_{2(d)}}\mathrm{s_y^{2''}}\right)$$
$$\left(\delta \exp\left(\frac{\mathrm{s_x^{2'}}-\mathrm{s_x^{2''}}}{\mathrm{s_x^{2'}}+\mathrm{s_x^{2''}}}\right) + (1-\delta)\exp\left(\frac{\mathrm{s_x^{2''}}-\mathrm{s_x^{2'}}}{\mathrm{s_x^{2'}}+\mathrm{s_x^{2''}}}\right)\right), \tag{20}$$

where α and δ are simply generalising constants which can assume any value within the range of 0 to 1, and $k_{1(d)}$ and $k_{2(d)}$ are constants that require to be optimised and provide the minimum MSE of the proposed hybrid estimator. Note that the proposed hybrid estimator $u_{prp(d)}$ reduces to $u_{dce(d)}$ if we set (α, $k_{1(d)}$, $k_{2(d)}$, δ) = (0.5, $k_{1(d)}$, $k_{2(d)}$, 1) in Eq. (20). Similarly, $u_{prp(d)}$ can provide different estimators for different values of (α, $k_{1(d)}$, $k_{2(d)}$, δ).

Expressing the estimator in Eq. (20) in terms of e_i, we have

Maejo Int. J. Sci. Technol. **2026**, *20*(01), 91-107

$$u_{prp(d)} = [S_y^2(1+e_0'')\{\alpha \exp\left(\frac{e_1'-e_1''}{2+e_1'+e_1''}\right) + (1-\alpha)\exp\left(\frac{e_1''-e_1'}{2+e_1'+e_1''}\right)\} - k_{1(d)}S_x^2(e_1'-e_1'')$$
$$+k_{2(d)}S_y^2(1+e_0'')][\delta \exp\left(\frac{e_1'-e_1''}{2+e_1'+e_1''}\right) + (1-\delta)\exp\left(\frac{e_1''-e_1'}{2+e_1'+e_1''}\right)\}. \tag{21}$$

Expanding Eq. (21) to the order $O(n^{-1})$, we obtain

$$u_{prp(d)} - S_y^2 = S_y^2[e_0'' + \Omega(e_1''-e_1') + \Omega(e_0''e_1''-e_0''e_1')$$
$$+\alpha\delta e_1^{2'} + \Omega e_1^{2'} + \alpha\delta e_1^{2''}$$
$$-(\Omega+2\alpha\delta)e_1'e_1'' - k_{1(d)}R\left\{e_1''-e_1' + \frac{e_1^{2'}}{2} + \frac{e_1^{2''}}{2} - e_1'e_1'' + 2\delta e_1'e_1''\right.$$
$$\left. -\delta(e_1^{2'}+e_1^{2''})\right\}$$
$$+k_2\left\{1+e_0'' + \frac{(e_1''-e_1')}{2} - \frac{e_1^{2''}}{8} + \frac{3e_1^{2'}}{8} - \frac{e_1'e_1''}{4} + \delta(e_1'-e_1'') + \frac{\delta}{2}(e_1^{2''}-e_1^{2'}) + \frac{e_1''e_0''}{2} - \right.$$
$$\left.\frac{e_1'e_0''}{2} + \delta(e_1'e_0''-e_1''e_0'')\right\}\Bigg], \tag{22}$$

where $\Omega = (1-\alpha-\delta)$ and $R = \frac{S_x^2}{S_y^2}$.

For the proposed estimator $u_{prp(d)}$, the bias and MSE are given by

$$Bias(u_{prp(d)}) = S_y^2\left[\begin{matrix}(\gamma''-\gamma')\Omega(\varphi_{22}-1) + \alpha\delta(\gamma''-\gamma')(\varphi_{04}-1)\\ -k_{1(d)}R\left\{(\gamma''-\gamma')\left(\frac{1-2\delta}{2}\right)(\varphi_{04}-1)\right\}\end{matrix} + \right.$$
$$\left. k_{2(d)}\left\{\begin{matrix}1+(\gamma''-\gamma')\left(\frac{1-2\delta}{2}\right)(\varphi_{22}-1)\\ +(\gamma''-\gamma')\left(\frac{4\delta-1}{8}\right)(\varphi_{04}-1)\end{matrix}\right\}\right], \tag{23}$$

$$MSE(u_{prp(d)}) = S_y^4\left[\begin{matrix}\gamma''\{(\varphi_{40}-1) + (\gamma''-\gamma')\Omega^2(\varphi_{04}-1)\\ +2\Omega(\gamma''-\gamma')(\varphi_{22}-1)\}\\ +k_{1(d)}^2D_{1(d)} + k_{2(d)}^2D_{2(d)} \begin{matrix}-2k_{1(d)}k_{2(d)}D_{3(d)}\\ -2k_{1(d)}D_{4(d)} + 2k_{2(d)}D_{5(d)}\end{matrix}\end{matrix}\right], \tag{24}$$

where

$D_{1(d)} = (\gamma''-\gamma')R^2(\varphi_{04}-1)$,

$D_{2(d)} = \left[1+\left\{\begin{matrix}\gamma''(\varphi_{40}-1) + 2(1-2\delta)(\gamma''-\gamma')(\varphi_{22}-1)\\ +\delta^2(\gamma''-\gamma')(\varphi_{04}-1)\end{matrix}\right\}\right]$,

$D_{3(d)} = (\gamma''-\gamma')R[(\varphi_{22}-1) + (1-2\delta)(\varphi_{04}-1)]$,

$D_{4(d)} = (\gamma''-\gamma')R[(\varphi_{22}-1) + \Omega(\varphi_{04}-1)]$,

$D_{5(d)} = \left[\begin{matrix}\gamma''(\varphi_{40}-1) + (\gamma''-\gamma')\left\{2\Omega + \left(\frac{1-2\delta}{2}\right)\right\}(\varphi_{22}-1)\\ +(\gamma''-\gamma')\left\{\alpha\delta + \left(\frac{1-2\delta}{2}\right)\right\}(\varphi_{04}-1)\end{matrix}\right]$.

The MSE of $u_{prp(d)}$ is minimised for (for instance)

$$k_{1(d)} = \frac{D_{2(d)}D_{4(d)} - D_{3(d)}D_{5(d)}}{D_{1(d)}D_{2(d)} - D_{3(d)}^2} = k_{1(d)opt}, \tag{25}$$

$$k_{2(d)} = \frac{D_{3(d)}D_{4(d)}-D_{1(d)}D_{5(d)}}{D_{1(d)}D_{2(d)}-D_{3(d)}^2} = k_{2(d)opt}. \quad (26)$$

The minimum MSE of $u_{prp(d)}$ is obtained by inputting the optimum values of $k_{1(d)}$ and $k_{2(d)}$ into Eq. (24):

$$MSE\left(u_{prp(d)}\right)_{min} = S_y^4[A_1 - A_2], \quad (27)$$

where

$$A_1 = \gamma''((\varphi_{40}-1) + \Omega^2(\gamma''-\gamma')(\varphi_{04}-1) + 2\Omega(\gamma''-\gamma')(\varphi_{22}-1)),$$

$$A_2 = \frac{D_{1(d)}D_{5(d)}^2 + D_{2(d)}D_{4(d)}^2 - 2D_{3(d)}D_{4(d)}D_{5(d)}}{D_{1(d)}D_{2(d)}-D_{3(d)}^2}.$$

Particular Case

By taking $k_2(d) = 0$ in Eq. (20), the class of estimators $u_{prp}(d)$ reduces to $u_{prp(d)1}$, and the resulting estimator takes the form

$$u_{prp(d)1} = \left[s_y^{2''}\left(\alpha \exp\left(\frac{s_x^{2'}-s_x^{2''}}{s_x^{2'}+s_x^{2''}}\right) + (1-\alpha)\exp\left(\frac{s_x^{2''}-s_x^{2'}}{s_x^{2'}+s_x^{2''}}\right)\right) + k_{1(d)}\left(s_x^{2'}-s_x^{2''}\right)\right]$$
$$\left(\delta \exp\left(\frac{s_x^{2'}-s_x^{2''}}{s_x^{2'}+s_x^{2''}}\right) + (1-\delta)\exp\left(\frac{s_x^{2''}-s_x^{2'}}{s_x^{2'}+s_x^{2''}}\right)\right). \quad (28)$$

Correspondingly the estimator $u_{prp(d)1}$ under the particular case gives the bias and MSE respectively as

$$Bias\left(u_{prp(d)1}\right) = S_y^2(\gamma''-\gamma')\left[\Omega(\varphi_{22}-1) + \alpha\delta(\varphi_{40}-1) - k_{1(d)}R\left\{\left(\frac{1-2\delta}{2}\right)(\varphi_{04}-1)\right\}\right], \quad (29)$$

$$MSE\left(u_{prp(d)1}\right) = S_y^4[\gamma''(\varphi_{40}-1) + (\gamma''-\gamma')\{\Omega^2(\varphi_{04}-1) + 2\Omega(\varphi_{22}-1)\} + k_{1(d)}^2 D_{1(d)} - 2k_{1(d)}D_{4(d)}]. \quad (30)$$

The $MSE\left(u_{prp(d)1}\right)$ is minimised for

$$k_{1(d)opt}^{+} = \frac{D_{4(d)}}{D_{1(d)}}. \quad (31)$$

Substituting $k_{1(d)opt}^{+}$ for $k_{1(d)}$, we obtain minimum $MSE\left(u_{prp(d)1}\right)$ as

$$MSE\left(u_{prp1}\right)_{min} = S_y^4\left[\gamma''(\varphi_{40}-1) + (\gamma''-\gamma')\{\Omega^2(\varphi_{04}-1) + 2\Omega(\varphi_{22}-1)\} - \frac{D_{4(d)}{}^2}{D_{1(d)}}\right], \quad (32)$$

or

$$MSE\left(u_{prp1}\right)_{min} = S_y^4\left[\gamma''(\varphi_{40}-1) - (\gamma''-\gamma')\frac{(\varphi_{22}-1)^2}{(\varphi_{04}-1)}\right]. \quad (33)$$

REAL-LIFE APPLICATIONS

This section interprets the empirical and simulation findings, explaining the statistical reasons behind the performance of the proposed and existing estimators.

To demonstrate the practical utility of the proposed hybrid variance estimator for the estimation of population variance, two real-life data sets are considered where two-phase sampling is appropriate due to cost or availability constraints of auxiliary information.

Application I: Recyclable Waste Collection in Italy [50]

- **Population I:** N = 80 cities in Italy in 2003

- **Study variable (Y):** Total amount of recyclable waste collection (in tons) in each city in Italy in 2003
- **Auxiliary variable (X):** Number of inhabitants in each city in 2003
- **Descriptive statistics:** $\bar{Y}$=51.83, $\bar{X}$=11.26, S_y=18.36,S_x=8.46,φ_{04}=2.8, φ_{40}=2.27, φ_{22}=2.22 and ρ=0.94 based on *N*
- **Sampling design:** A first-phase sample of n'= 20 cities is drawn to collect the auxiliary variable followed by a second-phase sample of n''= 8 cities for the study variable.

Application II: Recyclable Waste Collection in Italy [24]

- **Population II:** *N* = 103 cities in Italy in 2003
- **Study variable (Y):** Total amount of recyclable waste collection (in tons) in each city in Italy in 2003
- **Auxiliary variable (X):** Total amount of recyclable waste collection (in tons) in each city in Italy in 2002
- **Descriptive statistics:** $\bar{Y}$=62.62, $\bar{X}$=556.55, S_y=91.35, S_x=610.16, φ_{04}=17.87, φ_{40}=37.13, φ_{22}=17.22 and ρ=0.73 based on *N*
- **Sampling design:** A first-phase sample of n' = 60 cities is drawn to collect X, followed by a second-phase sample of n''= 40 cities for Y.

Estimation and Evaluation

For both applications, the relative efficiency (RE) of each estimator is calculated with respect to the usual unbiased variance estimator u_0'':

$$RE = \frac{Var(u_0'')}{MSE(u_{i(d)}) or MSE(u_{prp(d)(\alpha,\delta)})}, \tag{34}$$

where *i = r(d), re(d), reg(d), d(d), de(d), dce(d),* $prp(d)$ and α & δ take distinct values. Since both populations show positive correlation between *x* and *y*, the estimators $u_{p(d)}$ and $u_{pe(d)}$ are excluded from this empirical study. Table 1 presents the RE values for all estimators across the two populations. From the results presented in Table 1, the following points are made:

- Compared to the usual regression estimator $u_{reg(d)}$, the estimators $u_{d(d)}$, $u_{de(d)}$ and $u_{dce(d)}$ exhibit better performance.
- The proposed estimator $u_{prp(d)}$ consistently outperforms $u_{dce(d)}$.
  i. Under population I, higher efficiency is observed for $(\alpha,\delta) = \left(\frac{1}{4},\frac{1}{2}\right), \left(\frac{1}{4},\frac{3}{4}\right), \left(\frac{1}{4}, 1\right), \left(\frac{1}{2},\frac{3}{4}\right), \left(\frac{3}{4},\frac{3}{4}\right)$.
  ii. Under population II, better efficiency is noted for $(\alpha,\delta) = \left(\frac{1}{4},\frac{1}{2}\right), \left(\frac{1}{4},\frac{3}{4}\right), \left(\frac{1}{4},1\right), \left(\frac{1}{2},\frac{3}{4}\right)$.
- For both populations, $u_{prp(d)}$ is noted to be more efficient than $u_{dce(d)}$ for $(\alpha,\delta) = \left(\frac{1}{4},\frac{3}{4}\right), \left(\frac{1}{4},1\right)$ and $\left(\frac{1}{2},\frac{3}{4}\right)$.
- For $(\alpha,\delta) = \left(\frac{1}{2},1\right)$, efficiency of the proposed estimator reflects the efficiency of $u_{dce(d)}$ since the estimator $u_{dce(d)}$ is the unique case of $u_{prp(d)}$.
- The maximum performance is noted for $(\alpha,\delta) = \left(\frac{1}{4},\frac{3}{4}\right)$ across both populations, making it a suitable option for (α,δ) to achieve the optimal efficiency gain.

Maejo Int. J. Sci. Technol. **2026**, *20*(01), 91-107

- For $(\alpha, \delta) = \left(1, \frac{1}{4}\right), \left(\frac{3}{4}, \frac{1}{4}\right), \left(\frac{1}{2}, \frac{1}{4}\right), \left(\frac{1}{4}, \frac{1}{4}\right)$, the proposed estimator appears slightly less efficient, yet its *RE* values are closer to those of $u_{dce(d)}$ and u_{de}.

The proposed generalised hybrid estimator $u_{\text{prp}}(d)$ consistently outperforms conventional estimators, demonstrating substantial efficiency gains. For instance, for Population I, $u_{\text{prp}}(d; \alpha = 1/4, \delta = 3/4)$ achieves an RE of 1.953, which is higher than the regression estimator ($u_{\text{reg}}(d) = 1.641$) and the exponential-difference estimator ($u_{de}(d) = 1.823$). Similar patterns are observed in Population II, where $u_{\text{prp}}(d; \alpha = 1/4, \delta = 3/4)$ reaches an RE of 2.096, indicating almost double efficiency compared to the usual estimator u_0''.

Table 1. RE of $u_{i(d)}$ with respect to u_0''

Estimator	Population I	Population II
u_0''	1.000	1.000
$u_{r(d)}$	1.392	1.307
$u_{re(d)}$	1.585	1.221
$u_{reg(d)}$	1.641	1.308
$u_{d(d)}$	1.766	1.860
$u_{de(d)}$	1.823	1.929
$u_{dce(d)}$	**1.894**	**2.007**
$u_{prp(d)\left(\frac{1}{4},\frac{1}{4}\right)}$	1.815	**1.980**
$u_{prp(d)\left(\frac{1}{4},\frac{1}{2}\right)}$	**1.912**	**2.058**
$u_{prp(d)\left(\frac{1}{4},\frac{3}{4}\right)}$	**1.953**	**2.096**
$u_{prp(d)\left(\frac{1}{4},1\right)}$	**1.920**	**2.088**
$u_{prp(d)\left(\frac{1}{2},\frac{1}{4}\right)}$	1.796	1.910
$u_{prp(d)\left(\frac{1}{2},\frac{1}{2}\right)}$	**1.887**	**1.981**
$u_{prp(d)\left(\frac{1}{2},\frac{3}{4}\right)}$	**1.925**	**2.015**
$u_{prp(d)\left(\frac{1}{2},1\right)}$	**1.895**	**2.007**
$u_{prp(d)\left(\frac{3}{4},\frac{1}{4}\right)}$	1.779	1.847
$u_{prp(d)\left(\frac{3}{4},\frac{1}{2}\right)}$	**1.864**	**1.913**
$u_{prp(d)\left(\frac{3}{4},\frac{3}{4}\right)}$	**1.900**	**1.943**
$u_{prp(d)\left(\frac{3}{4},1\right)}$	**1.871**	**1.935**
$u_{prp(d)\left(1,\frac{1}{4}\right)}$	1.763	1.790
$u_{prp(d)\left(1,\frac{1}{2}\right)}$	**1.843**	**1.851**
$u_{prp(d)\left(1,\frac{3}{4}\right)}$	**1.876**	**1.879**
$u_{prp(d)(1,1)}$	**1.849**	**1.870**

Maejo Int. J. Sci. Technol. **2026**, *20*(01), 91-107

These empirical applications highlight the practical relevance of the proposed hybrid estimator. By effectively incorporating auxiliary information in a two-phase design, the estimator reduces the MSE and improves the precision in variance estimation. The results validate the theoretical derivations and confirm that the proposed framework is adaptable to real-world survey scenarios where auxiliary information is available.

The proposed estimator $u_{prp}(d)$ provides marginal but consistent improvements over the regression-based estimator $u_{reg}(d)$ and previously proposed hybrid estimator $u_{dce}(d)$, validating the effectiveness of combining a classical ratio estimator, an exponential estimator, and a difference-type estimator within a generalised hybrid framework.

As regards the effects of α and δ parameters, adjusting the α and δ parameters of $u_{prp}(d)$ affects efficiency. Smaller α values with larger δ tend to yield higher RE in Population II, whereas moderate values of both parameters are optimal in Population I. This demonstrates the flexibility of the proposed estimator, allowing users to tune parameters based on the characteristics of the population and auxiliary variable correlation.

As regards the influence of correlation, Population I exhibits a strong correlation (ρ=0.94) between Y and X, resulting in higher efficiencies across all auxiliary-assisted estimators. Population II, with moderate correlation (ρ=0.73), shows slightly lower RE values, but the proposed hybrid estimators still outperform existing estimators, confirming their robustness across varying correlation levels.

The real-world applications demonstrate that the proposed class of hybrid estimators provides significant efficiency gains for estimating population variance under two-phase sampling. By adjusting tuning parameters α and δ, practitioners can achieve optimal performance tailored to the correlation structure and variability of their data. Overall, the proposed estimator offers a flexible and superior alternative compared to classical and existing hybrid estimators in practice.

SIMULATION STUDY

A simulation study was conducted to evaluate the finite-sample performance of the proposed estimator $u_{prp}(d)$ in comparison with existing variance estimators under two-phase sampling. The primary objective of this study is to assess how efficiently the proposed estimator estimates the population variance of the study variable Y when auxiliary information X is available at different levels of correlation.

Three artificial finite populations are generated from a bivariate normal distribution of size N=5000. These populations are constructed to represent low, moderate and high correlation between the study variable Y and the auxiliary variable X. The mean vectors and variance–covariance matrices used for data generation are summarised in Table 2 along with the corresponding correlation coefficients (ρ=0.23, 0.53 and 0.81).

Although the populations are generated from known distributions, all population parameters (means and variances) are computed directly from the generated finite populations. This ensures that the simulation reflects a finite population framework consistent with the theoretical development of the estimators.

Maejo Int. J. Sci. Technol. **2026**, *20*(01), 91-107

Table 2. Simulated populations description

Simulated Population	Mean Vector of $(Y, X) = \mu$	variance-covariance matrix σ^2	Correlation ρ_{xy}
I		$\begin{bmatrix} 16 & 3 \\ 3 & 11 \end{bmatrix}$	0.23
II	$[2,2]$	$\begin{bmatrix} 16 & 7.03 \\ 7.03 & 11 \end{bmatrix}$	0.53
III		$\begin{bmatrix} 16 & 10.75 \\ 10.75 & 11 \end{bmatrix}$	0.81

For each simulated population, the following steps are carried out:

(i) A first-phase random sample of size n'=500 is drawn from the population.

(ii) From this first-phase sample, a second-phase subsample of size n''=200 is selected.

(iii) Sample statistics are computed and all competing estimators $u_i(d)$ are evaluated using the corresponding finite population parameters.

(iv) Steps (i) - (iii) are repeated independently, k=30,000 times.

(v) The MSE of each estimator is computed as

$\text{MSE}(u_i) = \frac{1}{k}\sum_{j=1}^{k}\left(u_{ij} - \bar{U}_i\right)^2$, where $\bar{U}_i$ denotes the average of the estimator over all replications. The RE of each estimator is then calculated with respect to the usual unbiased estimator u_0'' using Eq. (34).

For the proposed estimator $u_{prp}(d)$, the tuning parameters α and δ are selected by evaluating the MSE over a grid of candidate values and choosing those yielding the minimum MSE for each population.

Table 3 reports both simulated and empirical relative efficiency values for all estimators across the three populations. The empirical RE values are based on large-sample approximations while the simulated RE values are obtained directly from Monte Carlo replications. The close agreement between these two sets of values confirms the accuracy and stability of the theoretical expressions derived earlier. Several important observations emerge from Table 3:

Consistent superiority of the proposed estimator: Across all three populations, the proposed hybrid estimator $u_{prp}(d)$ achieves higher relative efficiency than traditional estimators such as $u_r(d)$, $u_{re}(d)$ and the regression estimator $u_{reg}(d)$.

Effect of correlation: When the correlation between Y and X is weak (ρ =0.23), the efficiency gains are modest but still evident, demonstrating the robustness of the proposed estimator. As the correlation increases to moderate and high levels (ρ = 0.53 and 0.81), the efficiency gains become substantial, indicating that the estimator effectively exploits auxiliary information.

Comparison with existing hybrid estimators: The estimator $u_{dce}(d)$ performs better than classical estimators. However, the proposed estimator $u_{prp}(d)$ either matches or surpasses its efficiency for several combinations of α and δ, particularly in moderate- and high-correlation settings.

Maejo Int. J. Sci. Technol. **2026**, *20*(01), 91-107

Stability across simulated and empirical results: The negligible differences between simulated and empirical RE values demonstrate that the proposed estimator performs consistently in both theoretical and finite-sample settings.

Table 3. Simulated and empirical RE values for the estimator u_i with respect to u_0

Estimators	Population I		Population II		Population III	
	Simulated RE	Empirical RE	Simulated RE	Empirical RE	Simulated RE	Empirical RE
u_0''	1.000	1.000	1.000	1.000	1.000	1.000
$u_{r(d)}$	0.973	0.970	1.208	1.196	1.455	1.454
$u_{re(d)}$	1.034	1.034	1.226	1.233	1.399	1.400
$u_{reg(d)}$	1.114	1.116	1.721	1.731	1.864	1.858
$u_{d(d)}$	1.231	1.231	1.854	1.850	1.995	1.989
$u_{de(d)}$	1.338	1.338	2.323	2.318	2.482	2.490
$u_{dce(d)}$	**1.443**	**1.437**	**2.855**	**2.848**	**3.994**	**3.990**
$u_{prp(d)\left(\frac{1}{4},\frac{1}{4}\right)}$	1.429	1.430	2.841	2.844	3.845	3.848
$u_{prp(d)\left(\frac{1}{4},\frac{1}{2}\right)}$	1.394	1.390	2.661	2.660	3.778	3.779
$u_{prp(d)\left(\frac{1}{4},\frac{3}{4}\right)}$	1.146	1.144	2.683	2.680	3.844	3.843
$u_{prp(d)\left(\frac{1}{4},1\right)}$	1.415	1.424	2.852	2.857	3.927	3.922
$u_{prp(d)\left(\frac{1}{2},\frac{1}{4}\right)}$	1.445	1.437	2.637	2.635	3.071	3.069
$u_{prp(d)\left(\frac{1}{2},\frac{1}{2}\right)}$	1.414	1.410	2.660	2.665	3.203	3.207
$u_{prp(d)\left(\frac{1}{2},\frac{3}{4}\right)}$	1.424	1.422	2.774	2.768	3.970	3.975
$u_{prp(d)\left(\frac{1}{2},1\right)}$	**1.443**	**1.437**	**2.855**	**2.848**	**3.994**	**3.990**
$u_{prp(d)\left(\frac{3}{4},\frac{1}{4}\right)}$	1.378	1.376	2.658	2.657	3.053	3.059
$u_{prp(d)\left(\frac{3}{4},\frac{1}{2}\right)}$	1.328	1.323	2.691	2.690	3.353	3.351
$u_{prp(d)\left(\frac{3}{4},\frac{3}{4}\right)}$	1.428	1.427	2.846	2.856	3.701	3.700
$u_{prp(d)\left(1,\frac{1}{4}\right)}$	1.131	1.137	2.657	2.660	3.156	3.161
$u_{prp(d)\left(1,\frac{1}{2}\right)}$	1.413	1.419	2.791	2.795	3.339	3.344
$u_{prp(d)\left(1,\frac{3}{4}\right)}$	1.282	1.287	3.236	3.234	3.446	3.452
$u_{prp(d)(1,1)}$	**1.497**	**1.495**	**3.443**	**3.441**	**4597**	**4.601**

CONCLUSIONS

Both simulation and real-world application results confirm that the proposed estimator consistently outperforms classical and existing hybrid estimators including $u_{dce}(d)$. Notably, the estimator maintains strong performance even under weak correlation between the study and auxiliary variables while exhibiting substantial efficiency gains as the correlation increases. The close agreement between simulated and empirical relative efficiency values further validates the theoretical developments.

Despite these encouraging results, some limitations remain. The proposed estimator has been evaluated under specific sampling schemes and assumptions which may not fully capture the complexity of real-world survey designs. Its performance under non-normal data distributions, stratified designs or in the presence of outliers warrants further investigation. Additionally, the computational search for α and δ could be improved through better systematic optimisation methods or adaptive algorithms. Future research may extend this work by incorporating imputation-based robust estimators and developing imputation-based quantile estimators for improved population variance estimation. It may also investigate Bayesian or machine-learning-based approaches to enhance efficiency in two-phase sampling frameworks. Exploring the applicability of the proposed estimator to large-scale, real-world data sets and complex survey frameworks would also be valuable to further demonstrate its practical utility.

REFERENCES

1. J. Neyman, "Contribution to the theory of sampling human populations", *J. Am. Stat. Assoc.*, **1938**, *33*, 101-116.
2. J. N. K. Rao, "On double sampling for stratification and analytic surveys", *Biometrika*, **1973**, *60*, 125-133.
3. W. G. Cochran, "Sampling Techniques", 3rd Edn., John Wiley and Sons, New York, **1977**.
4. R. Tabassum and I. A. Khan, "Double sampling ratio estimation with non-response", *J. Indian. Soc. Agric. Stat.*, **2004**, *58*, 300-306.
5. R. Tabassum, and I.A. Khan, "Double sampling ratio estimator for the population mean in presence of non-response", *Assam. Stat. Rev.*, **2006**, *20*, 73-83.
6. C. Bose, "Note on the sampling error in the method of double sampling", *Sankhyā*, **1943**, *6*, 330.
7. H. D. Patterson, "Sampling on successive occasions with partial replacement of units", *J. R. Stat. Soc. Ser. B*, **1950**, *12*, 241-255.
8. F. Yates, "Sampling Methods for Censuses and Surveys", Charles Griffin and Co., London, **1949**.
9. M. Hanif, N. Hamad and M. Q. Shahbaz, "Some new regression type estimators in two-phase sampling", *World Appl. Sci. J.*, **2010**, *8*, 799-803.
10. N. Hamad, M. Hanif and N. Haider, "A regression type estimator with two auxiliary variables for two-phase sampling", *Open J. Stat.*, **2013**, *3*, 74-78.
11. S. Bahl and R. K. Tuteja, "Ratio and product type exponential estimators", *J. Inform. Optim. Sci.*, **1991**, *12*, 159-164.
12. R. Singh, P. Chauhan, N. Sawan and F. Smarandache, "Improved exponential estimator for population variance using two auxiliary variables", in "Studies in Statistical Inference, Sampling Techniques and Demography" (Ed. R. Singh, J. Singh and F. Smarandache), InfoLearnQuest, Ann Arbor (USA), **2009**, pp.36-44.
13. A. Sanaullah, H. A. Ali, M. Noor-ul-Amin and M. Hanif, "Generalized exponential chain ratio estimators under stratified two-phase random sampling", *Appl. Math. Comput.*, **2014**, *226*, 541-547.
14. S. Bhushan, A. Kumar, and S. Kumar, "Efficient class of estimators of population mean under double sampling", *Thailand Stat.*, 2023, *21*, 498-509.
15. K. Sher, M. Ameeq, S. Naz, B. A. Alkhaleel, M. M. Hassan, and O. Albalawi, "Developing and evaluating efficient estimators for finite population mean in two-phase sampling", *AIMS Math.*, 2025, *10*, 8907-8925.

16. C. T. Isaki, "Variance estimation using auxiliary information", *J. Am. Stat. Assoc.*, **1983**, *78*, 117-123.
17. B. Prasad and H. P. Singh, "Some improved ratio-type estimators of finite population variance in sample surveys", *Commun. Stat .– Theory Meth.*, **1990**, *19*, 1127-1139.
18. S. K. Yadav and C. Kadilar, "Improved exponential type ratio estimator of population variance", *Rev. Colomb. Estad.*, **2013**, *36*, 145-152.
19. A. Asghar, A. Sanaullah and M. Hanif, "Generalized exponential type estimator for population variance in survey sampling", *Rev. Colomb. Estad.*, **2014**, *37*, 213-224.
20. A. Sanaullah, I. Niaz, J. Shabbir and I. Ehsan, "A class of hybrid type estimators for variance of a finite population in simple random sampling", *Commun. Stat. Simul. Comput.*, **2022**, *51*, 5609-5619.
21. I. Niaz, A. Sanaullah, I. Saleem and J. Shabbir, "An improved efficient class of estimators for the population variance", *Concurr. Comput. Pract. Exp.*, **2021**, *34*, e6620.
22. L. K. Grover, "A correction note on improvement in variance estimation using auxiliary information", *Commun. Stat. – Theory Meth.*, **2010**, *39*, 753-764.
23. J. Shabbir and S. Gupta, "On improvement in variance estimation using auxiliary information", *Commun. Stat. – Theory Meth.*, **2007**, *36*, 2177-2185.
24. J. Subramani and G. Kumarapandiyan, "Variance estimation using quartiles and their functions of an auxiliary variable", *Int. J. Stat. Appl.*, **2012**, *2*, 67-72.
25. A. Sanaullah, M. Hanif and A. Asghar, "Generalized exponential estimators for population variance under two-phase sampling", *Int. J. Appl. Comput. Math.*, **2016**, *2*, 75-84.
26. M. Yaqub and J. Shabbir, "An improved class of estimators for finite population variance", *Hacet. J. Math. Stat.*, **2016**, *45*, 1641-1660.
27. A. Asghar, A. Sanaullah and M. Hanif, "A multivariate regression-cum-exponential estimator for population variance vector in two phase sampling", *J. King Saud Univ. Sci.*, **2018**, *30*, 223-228.
28. S. Muneer, A. Khalil, J. Shabbir and G. Narjis, "New improved ratio-product type exponential estimator of finite population variance using auxiliary information", *J. Stat. Comput. Simul.*, **2018**, *88*, 3179-3192.
29. A. H. Al-Marshadi, A. H. Alharby and M. Q. Shahbaz, "On some new estimators of population variance in single and two-phase sampling", *Maejo Int. J. Sci. Technol.*, **2018**, *12*, 272-281.
30. M. Abid, S. Ahmed, M. Tahir, H. Z. Nazir and M. Riaz, "Improved ratio estimators of variance based on robust measures", *Sci. Iran.*, **2019**, *26*, 2484-2494.
31. 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.*, **2019**, *13*, Art.no.4.
32. U. Yasmeen, M. Noor-ul-Amin and M. Hanif, "Exponential estimators of finite population variance using transformed auxiliary variables", *Proc. Natl. Acad. Sci., India Sect. A Phys. Sci.*, **2019**, *89*, 185-191.
33. A. K. Sharma and A. K. Singh, "Estimation of population variance under an imputation method in two-phase sampling", *Proc. Natl. Acad. Sci., India Sect. A Phys. Sci.*, **2020**, *90*, 185-191.
34. H. O. Cekim and C. Kadilar, "*ln*-Type variance estimators in simple random sampling", *Pak. J. Stat. Oper. Res.*, **2020**, *16*, 689-696.
35. M. A. Gulzar, M. Abid, H. Z. Nazir, F. M. Zahid and M. Riaz, "On enhanced estimation of population variance using unconventional measures of an auxiliary variable", *J. Stat. Comput. Simul.*, **2020**, *90*, 2180-2197.

36. U. Shahzad, I. Ahmad, I. M. Almanjahie, N. H. Al-Noor and M. Hanif, "A new class of L-moments based calibration variance estimators", *Comput. Mater. Contin.*, **2020**, *66*,3013-3028.
37. U. Shahzad, I. Ahmad, I. M. Almanjahie and N. H. Al-Noor, "L-moments based calibrated variance estimators using double stratified sampling", *Comput. Mater. Contin.*, **2021**, *68*, 3411-3430.
38. U. Shahzad, I. Ahmad, I. M. Almanjahie, N. H. Al-Noor and M. Hanif, "A novel family of variance estimators based on L-moments and calibration approach under stratified random sampling", *Commun. Stat. Simul. Comput.*, **2023**, *52*, 3782-3795.
39. S. Bhushan, A. Kumar, A. Alsubie and S. A. Lone, "Variance estimation under an efficient class of estimators in simple random sampling", *Ain Shams Eng. J.*, **2023**, *14*, Art.no.102012.
40. M. H. Shahbaz, A. Fayomi and M. Q. Shahbaz, "Estimation of the general population parameter in single- and two-phase sampling", *AIMS Math.*, **2023**, *8*, 14951-14977.
41. M. B. Bantu, A. Asghar, A. Sanaullah, M. Hanif and L. A. Al-Essa, "Enhancing precision in population variance vector estimation: A two-phase sampling approach with multi-auxiliary information", *Sains Malaysiana*, **2024**, *53*, 1693-1702.
42. U. Daraz, M. A. Alomair, O. Albalawi and A. S. Al Naim, "New techniques for estimating finite population variance using ranks of auxiliary variable in two-stage sampling", *Math.*, **2024**, *12*, Art.no.2741.
43. M. Hussain, Q. Zaman, H. Ahmad, O. Albalawi and S. Iftikhar, "Improved exponential type variance estimators for population utilizing supplementary information", *Heliyon*, **2024**, *10*, Art.no.e31529.
44. A. Audu, M. Aphane, O. O. Ishaq and R. V. K. Singh, "New estimators of population variance based on logarithmic transformation in the presence of random non-response and measurement errors under successive sampling", *Afr. Mat.*, **2025**, *36*, Art.no.104.
45. U. Daraz, D. Agustiana, J. Wu and W. Emam, "Twofold auxiliary information under two-phase sampling: An improved family of double-transformed variance estimators", *Axioms*, **2025**, *14*, Art.no.64.
46. Z. Basit and M. I. Bhatti, "Efficient classes of estimators of population variance in two-phase successive sampling under random non-response", *Statistica*, **2022**, *82*, 177-198.
47. U. Daraz, M. A. Alomair, O. Albalawi and A. S. Al Naim, "New techniques for estimating finite population variance using ranks of auxiliary variable in two-stage sampling", *Math.*, **2024**, *12*, Art.no.2741.
48. A. M. Alomair, U. Shahzad, N. H. Al-Noor and M. A. Alomair, "Minimum-covariance-determinant-based mean estimators under systematic sampling", *Maejo Int. J. Sci. Technol.*, **2025**, *19*, 67-79.
49. A. M. Alomair, U. Shahzad, N. H. Al-Noor and H. Zhu, "Probability weighted moments and family of non-parametric regression estimators", *Maejo Int. J. Sci. Technol.*, **2025**, *19*, 160-170.
50. M. N. Murthy, "Sampling Theory and Methods", 2nd Edn., Statistical Publishing Society, Calcultta, **1967**.