



---

*Research article*

## A new class of ratio type estimators in single- and two-phase sampling

Amber Yousaf Dar<sup>1</sup>, Nadia Saeed<sup>1</sup>, Moustafa Omar Ahmed Abu-Shawiesh<sup>2,\*</sup>, Saman Hanif Shahbaz<sup>3</sup> and Muhammad Qaiser Shahbaz<sup>3</sup>

<sup>1</sup> College of Statistical and Actuarial Sciences, University of the Punjab, Lahore 54000, Pakistan

<sup>2</sup> Department of Mathematics, Faculty of Science, The Hashemite University, P.O. Box 330127, Zarqa 13133, Jordan

<sup>3</sup> Department of Statistics, King Abdulaziz University, Jeddah, Saudi Arabia

\* **Correspondence:** Email: [mabushawiesh@hu.edu.jo](mailto:mabushawiesh@hu.edu.jo); Tel: +962795976630.

**Abstract:** The estimation of a certain population characteristics is required for several situations. The estimates are built so that the error of estimation is minimized. In several situations estimation of the population mean is required. Different estimators for the mean are available but, there is still room for improvement. In this paper, a new class of ratio-type estimators is proposed for the estimation of the population mean. The estimators are proposed for single- and two-phase sampling schemes. The expressions for bias and mean square error are obtained for single-phase and two-phase sampling estimators. Mathematical comparison of the proposed estimators has been achieved by using some existing single-phase and two-phase sampling estimators. Extensive simulations have been conducted to compare the proposed estimators with some available single- and two-phase sampling estimators. It has been observed that the proposed estimators are better than the existing estimators. Consequently, the proposed ratio estimators are recommended for use by the practitioners in various fields of industry, engineering and medical and physical sciences.

**Keywords:** ratio-type estimators; two-phase sampling; bias; mean square error; families of estimators

**Mathematics Subject Classification:** 62D05, 62G30, 62P99

---

### 1. Introduction

The efficient estimation of population parameters has been a challenging job within the domain of statistics. Sampling methods have played an important role in developing various estimators for

various situations. Simple random sampling has been a useful method to draw a sample from a population that is homogeneous with respect to the characteristics under study. The estimation of the population mean or total is of interest in several situations. For example, we may be interested to estimate the average or total yearly income of a household; hence the estimation of the average or total income of the locality might be required. The estimate can be improved by using some auxiliary variable(s) that is (are) highly correlated with the variable for which the estimation is required. A classical method of estimation in such a situation is the ratio method of estimation, proposed by [1]. The method has been modified from time to time in order to improve the efficiency by reducing the mean square error (*MSE*). A modification of the ratio estimator was proposed in [2], whereas an extensive numerical study of the ratio estimator has been performed by the authors of [3]. More modifications of the ratio estimator were proposed in [4–6]. The product method of estimation has also been used by some authors when the variable of interest and auxiliary variable are negatively correlated. The ratio and product estimators have been combined by some authors to propose some more efficient estimators. The estimator proposed in [7] combines ratio and product estimators and has a smaller *MSE* than the classical ratio estimator. Another ratio-type estimator was proposed by the authors of [8] who used the ideas presented in [4–6].

In recent years, some work has been done to propose families of estimators. Some families of estimators were proposed in [9–11]. These families of estimators provide other different estimators as a special case. Ratio and ratio-type estimators have been developed by some authors by using some measures of auxiliary variables other than the mean. Some notable references in this regard are [12–14]. The idea of using a ratio estimator in ranked set sampling was proposed in [15].

In some situations, the population information of auxiliary variables is unknown; hence, ratio- or product-type estimators cannot be used. For such situations, a useful sampling technique, known as two-phase sampling, has been proposed. The method is described in [16] and [17], and some ratio- and regression-type estimators have been introduced by the authors of [18–20]. More details on various estimators in single- and two-phase sampling can be found in [21] and [22]. There is always room to suggest some more efficient estimators, and we have proposed some new estimators for the population mean in single- and two-phase sampling in this paper. The plan of the paper follows.

A brief description of the ratio estimator in single- and two-phase sampling is given in Section 2 alongside some notations. Some existing single- and two-phase sampling estimators are given in Section 3. A new estimator for single phase sampling is proposed in Section 4 alongside the expressions for the bias and *MSE* of the proposed estimator. In Section 5, a two-phase sampling version of the proposed estimator is presented. The expression for the bias and *MSE* of the two-phase sampling estimator are also given in Section 5. In Section 6, a numerical study is detailed for the proposed estimators. The numerical study comprised a simulation and applications using some real populations. The conclusions and recommendations are given in Section 7.

## 2. Methodology and notations

In this section the methodology and notations are given. Suppose the units of a population of size  $N$  are  $U_1, U_2, \dots, U_N$  and the values of some variable of interest are  $Y_1, Y_2, \dots, Y_N$ . Suppose further that the estimation of the population mean  $\bar{Y} = N^{-1} \sum_{i=1}^N Y_i$  is required. The estimation is done by using a random sample of size  $n$ . The conventional single phase sampling estimator of  $\bar{Y}$  is the sample mean

$\bar{y} = n^{-1} \sum_{i=1}^n y_i$ . When information on an auxiliary variable is available then the ratio estimator of the population mean is given as

$$\bar{y}_R = \frac{\bar{y}}{\bar{x}} \bar{X}, \quad (1)$$

where  $\bar{X} = N^{-1} \sum_{i=1}^N X_i$  and  $\bar{x} = n^{-1} \sum_{i=1}^n x_i$  are population and sample means of the auxiliary variable, respectively. The *MSE* of the ratio estimator is

$$MSE(\bar{y}_R) = \theta \bar{Y}^2 (C_y^2 + C_x^2 - 2\rho C_x C_y), \quad (2)$$

where  $\theta = n^{-1} - N^{-1}$ ,  $C_y$  is the population coefficient of variation for  $Y$ ,  $C_x$  is the population coefficient of variation for  $X$  and  $\rho$  is the population correlation coefficient between  $X$  and  $Y$ . The classical ratio estimator has been very popular for a long period of time.

In some situations, the population information for the auxiliary variable is not available and in such situations the ratio estimator (1) cannot be used. The problem can be solved by using a two-phase sampling technique. In two-phase sampling a first-phase sample of size  $n_1$  is drawn from a population of size  $N$ , and information about the auxiliary variable is recorded. A sub-sample of size  $n_2 < n_1$  is drawn from the first-phase sample and information about the auxiliary variable and study variable are recorded. The conventional ratio estimator in two-phase sampling is given as

$$\bar{y}_{R(2)} = \frac{\bar{y}_2}{\bar{x}_2} \bar{x}_1, \quad (3)$$

where  $\bar{y}_2 = n_2^{-1} \sum_{i=1}^{n_2} y_i$  is the second-phase sample mean of the study variable  $Y$ ,  $\bar{x}_2 = n_2^{-1} \sum_{i=1}^{n_2} x_i$  is second-phase sample mean of the auxiliary variable  $X$  and  $\bar{x}_1 = n_1^{-1} \sum_{i=1}^{n_1} x_i$  is the first-phase sample mean of the auxiliary variable  $X$ . The *MSE* of two-phase sampling ratio estimator is

$$MSE(\bar{y}_{R(2)}) = \bar{Y}^2 [\theta_2 C_y^2 + (\theta_2 - \theta_1)(C_x^2 - 2\rho C_x C_y)], \quad (4)$$

where  $\theta_2 = n_2^{-1} - N^{-1}$  and  $\theta_1 = n_1^{-1} - N^{-1}$ . Several modifications of the two-phase sampling ratio estimator have been proposed from time to time; see for example [21].

The *MSE* of an estimator in single-phase sampling is usually obtained by writing  $\bar{y} = \bar{Y}(1 + e_y)$  and  $\bar{x} = \bar{X}(1 + e_x)$  where  $e_y$  and  $e_x$  are errors in estimation such that

$$E(e_y) = E(e_x) = 0, E(e_y^2) = \theta C_y^2, E(e_x^2) = \theta C_x^2 \text{ and } E(e_y e_x) = \theta \rho C_x C_y. \quad (5)$$

The *MSE* of an estimator in two-phase sampling is usually obtained by writing  $\bar{y}_2 = \bar{Y}(1 + e_{y_2})$ ,  $\bar{x}_2 = \bar{X}(1 + e_{x_2})$  and  $\bar{x}_1 = \bar{X}(1 + e_{x_1})$  where  $e_{y_2}$ ,  $e_{x_2}$  and  $e_{x_1}$  are errors in estimation such that

$$\left. \begin{aligned} E(e_{x_1}) = E(e_{x_2}) = E(e_{y_2}) = 0; E(e_{y_2}^2) = \theta_2 C_y^2; E(e_{x_2}^2) = \theta_2 C_x^2; E(e_{x_1}^2) = \theta_1 C_x^2 \\ E(e_{y_2} e_{x_2}) = \theta_2 \rho C_x C_y; E(e_{y_2} e_{x_1}) = \theta_1 \rho C_x C_y; E(e_{x_1} e_{x_2}) = \theta_1 C_x^2; E[e_{y_1} (e_{x_1} - e_{x_2})] = 0 \\ E(e_{x_1} - e_{x_2})^2 = (\theta_2 - \theta_1) C_x^2; E[e_{y_2} (e_{x_1} - e_{x_2})] = (\theta_1 - \theta_2) \rho C_x C_y \end{aligned} \right\} \quad (6)$$

We will now discuss some important single- and two-phase sampling estimators in the following section.

### 3. Some existing estimators

Several estimators have been proposed from time to time for the estimation of the population mean in single- and two-phase sampling. Some of these are given in the following subsections.

#### 3.1 Some single phase sampling estimators

Some popular single-phase sampling estimators will now be discussed.

- Sisodia and Dwivedi estimator [2]

$$\text{Estimator: } \bar{y}_{SD} = \bar{y} \left( \frac{\bar{X} + C_x}{\bar{x} + C_x} \right). \quad (7)$$

MSE:

$$MSE(\bar{y}_{SD}) = \theta \bar{Y}^2 (C_y^2 + m_1^2 C_x^2 - 2m_1 \rho C_x C_y); m_1 = \bar{X} / (\bar{X} + C_x). \quad (8)$$

- Singh estimator [23]

$$\text{Estimator: } \bar{y}_{S04} = \bar{y} \left( \frac{\bar{X} + \beta_2}{\bar{x} + \beta_2} \right). \quad (9)$$

MSE:

$$MSE(\bar{y}_{S04}) = \theta \bar{Y}^2 (C_y^2 + m_2^2 C_x^2 - 2m_2 \rho C_x C_y); m_2 = \bar{X} / (\bar{X} + \beta_2), \quad (10)$$

where  $\beta_2$  is the coefficient of kurtosis for the auxiliary variable  $X$ .

- Kadilar and Cingi estimator [9]

$$\text{Estimator: } \bar{y}_{KC} = \left[ \bar{y} + b(\bar{X} - \bar{x}) \right] \frac{\bar{X}}{\bar{x}}. \quad (11)$$

MSE:

$$MSE(\bar{y}_{KC}) = \theta \bar{Y}^2 \left[ C_x^2 + C_y^2 (1 - \rho^2) \right], \quad (12)$$

where  $b$  is the regression coefficient between  $x$  and  $y$ .

- Yan and Tian estimator [24]

Estimator:

$$\bar{y}_{YT} = \bar{y} \left( \frac{\bar{X} + \beta_1}{\bar{x} + \beta_1} \right). \quad (13)$$

MSE:

$$MSE(\bar{y}_{YT}) = \theta \bar{Y}^2 (C_y^2 + m_3^2 C_x^2 - 2m_3 \rho C_x C_y); m_3 = \bar{X} / (\bar{X} + \beta_1), \quad (14)$$

where  $\beta_1$  is the coefficient of skewness for the auxiliary variable  $X$ .

- Singh estimator [25]

Estimator:

$$\bar{y}_{S03} = \bar{y} \left( \frac{\bar{X} + S_x}{\bar{x} + S_x} \right). \quad (15)$$

MSE:

$$MSE(\bar{y}_{S03}) = \theta \bar{Y}^2 (C_y^2 + m_4^2 C_x^2 - 2m_4 \rho C_x C_y); m_4 = \bar{X} / (\bar{X} + S_x), \quad (16)$$

where  $S_x$  is the standard deviation of the auxiliary variable  $X$ .

- Subramani and Kumarpandiyan estimator [26]

Estimator:

$$\bar{y}_{SK} = \bar{y} \left( \frac{\bar{X} + M_d}{\bar{x} + M_d} \right). \quad (17)$$

MSE:

$$MSE(\bar{y}_{SK}) = \theta \bar{Y}^2 (C_y^2 + m_5^2 C_x^2 - 2m_5 \rho C_x C_y); m_5 = \bar{X} / (\bar{X} + M_d), \quad (18)$$

where  $M_d$  is the mean deviation of the auxiliary variable  $X$ .

For the above estimators, it is assumed that the quantities  $C_x, C_y, M_d, \beta_1$  and  $\beta_2$  are known.

### 3.2 Some two-phase sampling estimators

Some popular two-phase sampling estimators will now be discussed.

- Mohanty estimator [27]

Estimator:

$$\bar{y}_{M(2)} = \left[ \bar{y}_2 + \beta_{yx} (\bar{x}_1 - \bar{x}_2) \right] \frac{\bar{z}_1}{\bar{z}_2}. \quad (19)$$

*MSE:*

$$MSE(\bar{y}_{M(2)}) = \bar{Y}^2 \left[ \theta_2 C_y^2 + (\theta_2 - \theta_1) \left\{ \rho_{xz}^2 C_z^2 - (\rho_{xy} C_y - \rho_{xz} C_z)^2 + (C_z - C_y \rho_{yz})^2 - C_y^2 \rho_{yz}^2 \right\} \right], \quad (20)$$

where  $Z$  is another auxiliary variable.

- Hanif, Hammad and Shahbaz estimator [28]

Estimator:

$$\bar{y}_{HHS(2)} = [\bar{y}_2 + \beta_{yx}(\bar{x}_1 - \bar{x}_2)] [k(\bar{z}_1/\bar{z}_2) + (1-k)(\bar{z}_2/\bar{z}_1)], \quad (21)$$

where  $k$  is the weighting constant such that  $0 < k < 1$ .

*MSE:*

$$MSE(\bar{y}_{HHS(2)}) = \bar{Y}^2 C_y^2 \left[ \theta_2 - (\theta_2 - \theta_1) \left\{ \rho_{xy}^2 + (\rho_{yz} - \rho_{xy} \rho_{xz})^2 \right\} \right], \quad (22)$$

where  $\rho_{xy}$  is the correlation coefficient between  $(X, Y)$ ,  $\rho_{yz}$  is the correlation coefficient between  $(Y, Z)$  and  $\rho_{xz}$  is the correlation coefficient between  $(X, Z)$ .

More details about single- and two-phase sampling estimators can be found in [21]. We will now propose a new ratio-type estimator in single-phase sampling.

#### 4. A new class of estimators for the mean in single-phase sampling

The proposed class of estimators for single phase sampling is

$$\bar{y}_{DS_i^3} = \alpha \bar{y} \left( \frac{\bar{X} + A_i}{\bar{x} + A_i} \right) + (1 - \alpha) [\bar{y} + b(\bar{X} - \bar{x})] \frac{\bar{X}}{\bar{x}}; \quad i = 1, 2, 3, 4, 5 \quad (23)$$

where  $A_1 = C_x$ ,  $A_2 = \beta_2$ ,  $A_3 = \beta_1$ ,  $A_4 = S_x$  and  $A_5 = M_d$ . We can see that the proposed class of estimators can be written as

$$\bar{y}_{DS_i^3} = \alpha t_i + (1 - \alpha) \bar{y}_{KC}; \quad i = 1, 2, 3, 4, 5$$

where  $t_1 = \bar{y}_{SD}$ ,  $t_2 = \bar{y}_{S04}$ ,  $t_3 = \bar{y}_{YT}$ ,  $t_4 = \bar{y}_{S03}$  and  $t_5 = \bar{y}_{SK}$ . From above we can see that the proposed class of estimators is a weighted sum of various ratio type estimators. The bias and *MSE* of the proposed class of estimators are derived in the following subsections.

##### 4.1 Bias and MSE

Using the notations from (5), the class of estimators given in (23) can be written as

$$\begin{aligned}\bar{y}_{DS_i^3} &= \alpha \bar{Y} (1 + e_y) \left( \frac{\bar{X} + A_i}{\bar{X} + \bar{X}e_x + A_i} \right) + (1 - \alpha) [\bar{Y} (1 + e_y) - b\bar{X}e_x] \frac{\bar{X}}{\bar{X} (1 + e_x)} \\ &= \alpha \bar{Y} (1 + e_y) (1 + m_i e_x)^{-1} + (1 - \alpha) (\bar{Y} + \bar{Y}e_y - b\bar{X}e_x) (1 + e_x)^{-1}; m_i = \frac{\bar{X}}{\bar{X} + A_i}.\end{aligned}$$

Expanding the negative powers and retaining the linear terms only, we have

$$\bar{y}_{DS_i^3} = \alpha \bar{Y} (1 + e_y) (1 - m_i e_x) + (1 - \alpha) (\bar{Y} + \bar{Y}e_y - b\bar{X}e_x) (1 - e_x). \quad (24)$$

Expanding and applying expectation, the bias of the proposed class of estimators is

$$\text{Bias}(\bar{y}_{DS_i^3}) = E(\bar{y}_{DS_i^3} - \bar{Y}) = -\theta \alpha \rho m_i \bar{Y} C_x C_y.$$

Again, expanding (24) and retaining only the linear terms we have

$$\bar{y}_{DS_i^3} - \bar{Y} = (\bar{Y}e_y - \bar{Y}e_x + \alpha \bar{Y}e_x - b\bar{X}e_x + ab\bar{X}e_x - \alpha m_i \bar{Y}e_x).$$

Squaring and applying expectation, the *MSE* of the proposed class of estimators is

$$MSE(\bar{y}_{DS_i^3}) = E(\bar{y}_{DS_i^3} - \bar{Y})^2 = E(\bar{Y}e_y - \bar{Y}e_x + \alpha \bar{Y}e_x - b\bar{X}e_x + ab\bar{X}e_x - \alpha m_i \bar{Y}e_x)^2.$$

Expanding the square, applying expectation to individual terms and ignoring  $E(b) - \beta$ , where  $\beta$  is the population regression coefficient between  $X$  and  $Y$ , the *MSE* of the proposed class of estimators is

$$MSE(\bar{y}_{DS_i^3}) = \theta \bar{Y}^2 \left[ C_y^2 \{1 - \rho^2 (1 - \alpha^2)\} + C_x^2 (1 - \alpha + \alpha m_i)^2 - 2\alpha \rho C_y C_x (1 - \alpha + \alpha m_i) \right]. \quad (25)$$

The optimum value of  $\alpha$  that minimizes the *MSE* is obtained by differentiating (25) with respect to  $\alpha$  and equating the resulting derivative to zero. The optimum value thus obtained is

$$\alpha = \frac{C_x}{C_x (1 - m_i) + \rho C_y}.$$

Substituting the optimum value of  $\alpha$  in (25), the minimum *MSE* of the proposed class of estimators is

$$MSE_{\min}(\bar{y}_{DS_i^3}) = \theta \bar{Y}^2 C_y^2 (1 - \rho^2). \quad (26)$$

#### 4.2 Comparison of the proposed class of estimators

A comparison of the proposed class of estimators is given below. For the comparison we have compared the *MSE* of the proposed class of estimators with some existing estimators. The proposed class of estimators will be better than the estimators compared if

$$MSE(t) - MSE(\bar{y}_{DS_i^3}) > 0,$$

where  $t$  is any estimator compared. We first compare the proposed class of estimators with the classical ratio estimator. For this we have

$$\begin{aligned} MSE(\bar{y}_R) - MSE(\bar{y}_{DS_i^3}) &= \theta \bar{Y}^2 (C_y^2 + C_x^2 - 2\rho C_x C_y) - \theta \bar{Y}^2 C_y^2 (1 - \rho^2) \\ &= \theta \bar{Y}^2 (\rho C_y - C_x)^2. \end{aligned} \quad (27)$$

From above, it is easy to see that  $MSE(\bar{y}_R) \geq MSE(\bar{y}_{DS_i^3})$ ; hence the proposed class of estimators is always more precise than the ratio estimator. The equality holds when  $\rho C_y = C_x$ .

Again, we will compare the proposed class of estimators with the estimators given in Subsection 3.1. The proposed class of estimators will be more efficient than the estimators given in Subsection 3.1 if

$$MSE(t_i) - MSE(\bar{y}_{DS_i^3}) \geq 0; t_i = \bar{y}_{SD}, \bar{y}_{S04}, \bar{y}_{YT}, \bar{y}_{S03} \text{ and } \bar{y}_{SK}.$$

Now,

$$MSE(t_i) = \theta \bar{Y}^2 (C_y^2 + m_i^2 C_x^2 - 2m_i \rho C_x C_y); i = 1, 2, 3, 4, 5,$$

and  $m_i$  is suitably defined. Comparing the  $MSE$  of  $t_i$  with  $MSE$  of proposed class of estimators, we have

$$MSE(t_i) - MSE(\bar{y}_{DS_i^3}) = \theta \bar{Y}^2 (C_y^2 + m_i^2 C_x^2 - 2m_i \rho C_x C_y) - \theta \bar{Y}^2 C_y^2 (1 - \rho^2)$$

or

$$MSE(t_i) - MSE(\bar{y}_{DS_i^3}) = \theta \bar{Y}^2 (C_y^2 \rho^2 - m_i^2 C_x^2 + 2\rho C_y C_x). \quad (28)$$

From (28) we can see that the proposed class of estimators will be more efficient than any of the estimators given in Subsection 3.1 if  $C_y^2 \rho^2 + 2\rho C_y C_x > m_i^2 C_x^2$ .

Finally, we will compare the proposed class of estimators with the estimator proposed in [9]. For this we have

$$MSE(\bar{y}_{KC}) - MSE(\bar{y}_{DS_i^3}) = \theta \bar{Y}^2 [C_x^2 + C_y^2 (1 - \rho^2)] - \theta \bar{Y}^2 C_y^2 (1 - \rho^2)$$

or

$$MSE(\bar{y}_{KC}) - MSE(\bar{y}_{DS_i^3}) = \theta \bar{Y}^2 C_x^2 > 0; \quad (29)$$

hence the proposed class of estimators will always be more efficient than the estimator proposed in [9].

We will now propose a class of estimators for two-phase sampling in the following section.



## 5. A new class of estimators for the mean in two-phase sampling

The proposed class of estimators for two-phase sampling is

$$\bar{y}_{DS_i^3(2)} = \alpha \bar{y}_2 \left( \frac{\bar{x}_1 + A_i}{\bar{x}_2 + A_i} \right) + (1 - \alpha) \left[ \bar{y}_2 + b(\bar{x}_1 - \bar{x}_2) \right] \frac{\bar{x}_1}{\bar{x}_2}; i = 1, 2, 3, 4, 5 \quad (30)$$

where  $A_i$  has been defined in Section 4. Also  $\bar{x}_1$  and  $\bar{x}_2$  are the first- and second-phase sample means of the auxiliary variable  $X$  and  $\bar{y}_2$  is the second-phase sample mean of the study variable  $Y$ . We can see that the proposed class of estimators for two-phase sampling can be written as

$$\bar{y}_{DS_i^3(2)} = \alpha t_{i(2)} + (1 - \alpha) \bar{y}_{KC(2)}; i = 1, 2, 3, 4, 5$$

where

$$t_{i(2)} = \bar{y}_2 \left( \frac{\bar{x}_1 + A_i}{\bar{x}_2 + A_i} \right); i = 1, 2, 3, 4, 5$$

represents the two-phase sampling counterparts of the estimators  $\bar{y}_{SD}, \bar{y}_{S04}, \bar{y}_{YT}, \bar{y}_{S03}$  and  $\bar{y}_{SK}$ . Also  $\bar{y}_{KC(2)}$  is the two-phase sampling counterpart of the estimator proposed in [9]; it is given as

$$\bar{y}_{KC(2)} = \left[ \bar{y}_2 + b(\bar{x}_1 - \bar{x}_2) \right] \frac{\bar{x}_1}{\bar{x}_2}.$$

From above we can see that the proposed class of estimators is a weighted sum of various ratio-type estimators in two-phase sampling. The bias and  $MSE$  of the proposed class of estimators are derived in the following subsections.

### 5.1 Bias and MSE

Using the notations from (6), the class of estimators given by (30) can be written as

$$\begin{aligned} \bar{y}_{DS_i^3(2)} &= \alpha \frac{(\bar{Y} + \bar{Y}e_{y_2})(\bar{X} + \bar{X}e_{x_1} + A_i)}{(\bar{X} + \bar{X}e_{x_2} + A_i)} + (1 - \alpha) \left[ (\bar{Y} + \bar{Y}e_{y_2}) + b\bar{X}(e_{x_1} - \bar{X}e_{x_2}) \right] \frac{1 + e_{x_1}}{1 + e_{x_2}} \\ &= \alpha \bar{Y} (1 + e_{y_2}) (1 + m_i e_{x_1}) (1 + m_i e_{x_2})^{-1} + (1 - \alpha) \left[ (\bar{Y} + \bar{Y}e_{y_2}) + b\bar{X}(e_{x_1} - e_{x_2}) \right] \\ &\quad \times (1 + e_{x_1}) (1 + e_{x_2})^{-1}. \end{aligned}$$

Expanding the negative powers and retaining the linear terms only, we have

$$\begin{aligned} \bar{y}_{DS_i^3(2)} &= \alpha \bar{Y} (1 + e_{y_2}) (1 + m_i e_{x_1}) (1 - m_i e_{x_2}) + (1 - \alpha) \left[ \bar{Y} (1 + e_{y_2}) + b\bar{X}(e_{x_1} - e_{x_2}) \right] \\ &\quad \times (1 + e_{x_1}) (1 - e_{x_2}) \end{aligned} \quad (31)$$

Expanding (31) and applying expectation, the bias of the proposed class of estimators for two-phase sampling is

$$\text{Bias}\left(\bar{y}_{DS_i^3(2)}\right) = E\left(\bar{y}_{DS_i^3(2)} - \bar{Y}\right) = -(\theta_2 - \theta_1)\alpha m_i \bar{Y} \rho C_y C_x.$$

Again, expanding (31) and retaining only the linear terms, we have

$$\bar{y}_{DS_i^3(2)} - \bar{Y} = \bar{Y}e_{y_2} + \bar{Y}(e_{x_1} - e_{x_2}) - \alpha\bar{Y}(e_{x_1} - e_{x_2}) + \alpha m_i \bar{Y}(e_{x_1} - e_{x_2}) + b\bar{X}(e_{x_1} - e_{x_2}) - ab\bar{X}(e_{x_1} - e_{x_2}).$$

Squaring and applying expectation, the *MSE* of the proposed class of estimators is

$$\begin{aligned} \text{MSE}\left(\bar{y}_{DS_i^3(2)}\right) &= E\left(\bar{y}_{DS_i^3(2)} - \bar{Y}\right)^2 = E\left[\bar{Y}e_{y_2} + \bar{Y}(e_{x_1} - e_{x_2}) - \alpha\bar{Y}(e_{x_1} - e_{x_2}) + \alpha m_i \bar{Y}(e_{x_1} - e_{x_2})\right. \\ &\quad \left.+ b\bar{X}(e_{x_1} - e_{x_2}) - ab\bar{X}(e_{x_1} - e_{x_2})\right]^2. \end{aligned}$$

Expanding the square and applying expectation to individual terms, the *MSE* of the proposed class of estimators is

$$\begin{aligned} \text{MSE}\left(\bar{y}_{DS_i^3(2)}\right) &= \bar{Y}^2 C_y^2 \left[ \theta_2 + (1 - \alpha^2)(\theta_2 - \theta_1)\rho^2 \right] - 2\alpha(\theta_2 - \theta_1)\rho\bar{Y}^2 C_y C_x (1 - \alpha - m_i\alpha) \\ &\quad + (\theta_2 - \theta_1)\bar{Y}^2 C_x^2 \left[ 1 - \alpha(1 - m_i) \right]^2. \end{aligned} \quad (32)$$

The optimum value of  $\alpha$  that minimizes the mean square error is obtained by differentiating (32) with respect to  $\alpha$  and equating the resulting derivative to zero. The optimum value thus obtained is

$$\alpha = \frac{C_x^2(1 - m_i) + \rho C_x C_y}{C_x^2(1 - m_i)^2 + 2\rho C_y C_x(1 + m_i) - \rho^2 C_y^2}. \quad (33)$$

The minimum value of the *MSE* of the class of two-phase sampling estimators can be obtained by applying the optimum value of  $\alpha$  from (33) in (32), as follows:

$$\begin{aligned} \text{MSE}\left(\bar{y}_{DS_i^3(2)}\right) &= \frac{1}{\Delta} \left[ \bar{Y}^2 C_y \left\{ 4m_i \rho C_x^3 (\theta_2 - \theta_1) + \rho C_y^2 \left[ 2C_x(1 + m_i) - \rho C_y \right] \left[ \theta_2 + \rho^2 (\theta_2 - \theta_1) \right] \right. \right. \\ &\quad \left. \left. + C_y C_x^2 \left[ (1 - m_i)^2 + \theta_2 \left\{ 1 - \rho^2 + m(m - 2)(1 + \rho^2) \right\} \right] \right\} \right], \end{aligned} \quad (34)$$

where  $\Delta = C_x^2(1 - m_i)^2 + 2\rho C_y C_x(1 + m_i) - \rho^2 C_y^2$ .

## 6. Numerical study

In this section, we present the numerical study for the proposed estimators. The numerical study comprised a simulation for the single- and two-phase sampling estimators and an application of the proposed single-phase sampling estimator by using some real populations. We will first give the simulation study for the proposed estimators in the following subsection.

6.1 Simulation study

Here, we present the simulation study for the proposed single- and two-phase sampling estimators. The algorithm for the single phase sampling is given below.

- 1) Generate an artificial population of size 5000 from a bivariate normal distribution  $N_2(65, 50, 6^2, 5^2, \rho)$  by using different values of the correlation coefficient.
- 2) Generate random samples of sizes 50, 100, 200 and 500 from the generated population.
- 3) Compute different estimators by using the generated samples.
- 4) Repeat Steps 2 and 3; 20000 times for each sample size.
- 5) Compute the bias and *MSE* of each estimator for different sample sizes by using

$$Bias(\bar{y}_i) = \bar{\bar{y}}_i - 65 \text{ and } MSE(\bar{y}_i) = \frac{1}{20000} \sum_{j=1}^{20000} (\bar{y}_{ij} - \bar{\bar{y}}_i)^2 ; \bar{\bar{y}}_i = \frac{1}{20000} \sum_{j=1}^{20000} \bar{y}_{ij}$$

where  $i = R, SD, S04, KC, YT, S03, SK, DS_1^3, DS_2^3, DS_3^3, DS_4^3$  and  $DS_5^3$ .

Table 1 contains the bias of various estimators and Table 2 contains the *MSE* of various estimators used in the simulation study.

**Table 1.** Bias of various estimators.

	$\rho = -0.9$				$\rho = -0.5$			
	$n = 50$	$n = 100$	$n = 200$	$n = 500$	$n = 50$	$n = 100$	$n = 200$	$n = 500$
$\bar{y}_R$	0.1215	0.1304	-0.0082	0.0302	0.0672	0.0288	0.1273	-0.1214
$\bar{y}_{SD}$	0.1214	0.1304	-0.0082	0.0302	0.0671	0.0287	0.1273	-0.1214
$\bar{y}_{S04}$	0.1200	0.1301	-0.0086	0.0302	0.0659	0.0279	0.1269	-0.1216
$\bar{y}_{KC}$	0.1097	0.1311	-0.0117	0.0304	0.0626	0.0249	0.1249	-0.1224
$\bar{y}_{YT}$	0.1215	0.1304	-0.0082	0.0302	0.0672	0.0288	0.1273	-0.1214
$\bar{y}_{S03}$	0.1191	0.1300	-0.0088	0.0301	0.0652	0.0275	0.1266	-0.1217
$\bar{y}_{SK}$	0.1196	0.1300	-0.0087	0.0302	0.0655	0.0277	0.1267	-0.1216
$\bar{y}_{DS_1^3}$	0.0958	0.1319	-0.0159	0.0307	0.0530	0.0166	0.1198	-0.1246
$\bar{y}_{DS_2^3}$	0.0966	0.1324	-0.0157	0.0307	0.0546	0.0174	0.1202	-0.1244
$\bar{y}_{DS_3^3}$	0.0957	0.1319	-0.0159	0.0307	0.0529	0.0165	0.1198	-0.1246
$\bar{y}_{DS_4^3}$	0.0972	0.1326	-0.0156	0.0308	0.0558	0.0180	0.1204	-0.1243
$\bar{y}_{DS_5^3}$	0.0969	0.1325	-0.0156	0.0308	0.0552	0.0177	0.1203	-0.1244
	$\rho = 0.5$				$\rho = 0.9$			
	$n = 50$	$n = 100$	$n = 200$	$n = 500$	$n = 50$	$n = 100$	$n = 200$	$n = 500$
$\bar{y}_R$	-0.0360	0.0185	-0.1519	-0.0638	0.0337	0.0175	-0.0997	-0.1203
$\bar{y}_{SD}$	-0.0360	0.0185	-0.1519	-0.0638	0.0337	0.0175	-0.0997	-0.1203
$\bar{y}_{S04}$	-0.0381	0.0179	-0.1521	-0.0640	0.0323	0.0170	-0.1004	-0.1204
$\bar{y}_{KC}$	-0.0250	0.0211	-0.1512	-0.0625	0.0444	0.0206	-0.0920	-0.1196
$\bar{y}_{YT}$	-0.0360	0.0186	-0.1519	-0.0638	0.0337	0.0175	-0.0997	-0.1203
$\bar{y}_{S03}$	-0.0392	0.0175	-0.1523	-0.0642	0.0315	0.0167	-0.1008	-0.1204
$\bar{y}_{SK}$	-0.0386	0.0177	-0.1522	-0.0641	0.0319	0.0168	-0.1006	-0.1204

*Continued on next page*

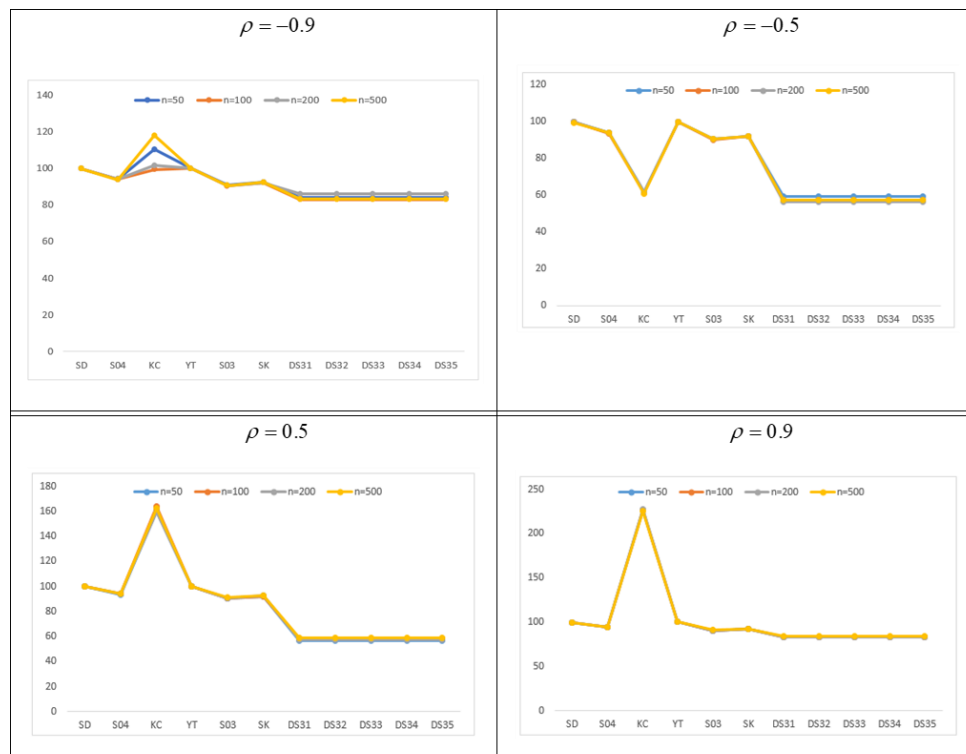
	$\rho = 0.5$				$\rho = 0.9$			
	$n = 50$	$n = 100$	$n = 200$	$n = 500$	$n = 50$	$n = 100$	$n = 200$	$n = 500$
$\bar{y}_{DS_1^3}$	-0.0493	0.0157	-0.1527	-0.0653	0.0315	0.0169	0.1013	-0.1204
$\bar{y}_{DS_2^3}$	-0.0506	0.0151	-0.1530	-0.0654	0.0307	0.0165	-0.1014	-0.1205
$\bar{y}_{DS_3^3}$	-0.0492	0.0157	-0.1527	-0.0653	0.0315	0.0169	-0.1013	-0.1204
$\bar{y}_{DS_4^3}$	-0.0512	0.0147	-0.1532	-0.0654	0.0303	0.0163	-0.1015	-0.1205
$\bar{y}_{DS_5^3}$	-0.0509	0.0149	-0.1531	-0.0654	0.0305	0.0164	-0.1015	-0.1205

**Table 2.** *MSE* of various estimators.

	$\rho = -0.9$				$\rho = -0.5$			
	$n = 50$	$n = 100$	$n = 200$	$n = 500$	$n = 50$	$n = 100$	$n = 200$	$n = 500$
$\bar{y}_R$	1.5992	0.7699	0.3732	0.1413	1.5658	0.7463	0.3722	0.1434
$\bar{y}_{SD}$	1.5957	0.7682	0.3723	0.1409	1.5624	0.7447	0.3714	0.1430
$\bar{y}_{S04}$	1.5037	0.7241	0.3515	0.1325	1.4727	0.6990	0.3507	0.1348
$\bar{y}_{KC}$	1.7655	0.7655	0.3793	0.1667	0.9687	0.4573	0.2299	0.0873
$\bar{y}_{YT}$	1.5986	0.7690	0.3729	0.1412	1.5672	0.7463	0.3719	0.1434
$\bar{y}_{S03}$	1.4504	0.6965	0.3379	0.1280	1.4192	0.6752	0.3378	0.1298
$\bar{y}_{SK}$	1.4771	0.7097	0.3442	0.1305	1.4453	0.6882	0.3440	0.1322
$\bar{y}_{DS_1^3}$	1.3436	0.6385	0.3210	0.1175	0.9292	0.4294	0.2108	0.0825
$\bar{y}_{DS_2^3}$	1.3435	0.6385	0.3210	0.1175	0.9292	0.4293	0.2108	0.0825
$\bar{y}_{DS_3^3}$	1.3436	0.6385	0.3210	0.1175	0.9292	0.4294	0.2108	0.0825
$\bar{y}_{DS_4^3}$	1.3435	0.6384	0.3210	0.1175	0.9291	0.4293	0.2108	0.0825
$\bar{y}_{DS_5^3}$	1.3435	0.6384	0.3210	0.1175	0.9291	0.4293	0.2108	0.0825
	$\rho = 0.5$				$\rho = 0.9$			
	$n = 50$	$n = 100$	$n = 200$	$n = 500$	$n = 50$	$n = 100$	$n = 200$	$n = 500$
$\bar{y}_R$	1.5602	0.7839	0.3858	0.1370	1.5661	0.7574	0.3738	0.1388
$\bar{y}_{SD}$	1.5568	0.7822	0.3850	0.1367	1.5627	0.7558	0.3730	0.1385
$\bar{y}_{S04}$	1.4633	0.7400	0.3629	0.1291	1.4712	0.7127	0.3518	0.1306
$\bar{y}_{KC}$	2.5034	1.2828	0.6160	0.2223	3.5518	1.7170	0.8512	0.3124
$\bar{y}_{YT}$	1.5611	0.7847	0.3856	0.1368	1.5662	0.7576	0.3740	0.1388
$\bar{y}_{S03}$	1.4125	0.7100	0.3501	0.1245	1.4181	0.6858	0.3386	0.1261
$\bar{y}_{SK}$	1.4394	0.7228	0.3566	0.1268	1.4452	0.6985	0.3450	0.1284
$\bar{y}_{DS_1^3}$	0.8868	0.4585	0.2218	0.0806	1.2971	0.6257	0.3113	0.1163
$\bar{y}_{DS_2^3}$	0.8865	0.4585	0.2218	0.0806	1.2969	0.6256	0.3113	0.1163
$\bar{y}_{DS_3^3}$	0.8868	0.4585	0.2218	0.0806	1.2971	0.6257	0.3113	0.1163
$\bar{y}_{DS_4^3}$	0.8863	0.4584	0.2218	0.0806	1.2968	0.6256	0.3113	0.1163
$\bar{y}_{DS_5^3}$	0.8864	0.4584	0.2218	0.0806	1.2968	0.6256	0.3113	0.1163

From Table 1, we can see that the bias remains negative for all of the sample sizes and all values of the correlation coefficients. From Table 2, we can see that all members of the proposed class of estimators are more efficient than the other estimators used in the study. We can also see that the *MSE*

of all of the estimators decreases with an increase in the sample size. The relative efficiency of various estimators relative to the classical ratio estimator was also computed; the results are plotted in Figure 1 given below.



**Figure 1.** Relative efficiency of various estimators.

From the above figure we can see that all the members of the proposed class of estimators are more efficient than the ratio estimator.

We also conducted the simulation study for the proposed class of estimators for two-phase sampling. The algorithm for the two-phase sampling simulation is given below.

- 1) Generate an artificial population of size 5000 from a bivariate normal distribution  $N_2(65, 50, 6^2, 5^2, \rho)$  by using different values of the correlation coefficient.
- 2) Generate first-phase random samples of sizes 500 and 1000 from the generated population.
- 3) Generate second-phase random samples with sizes 5%, 10% and 20% of the first-phase sample.
- 4) Compute different estimators by using the second-phase sample mean of  $Y$ , first- and second-phase sample means of the auxiliary variable  $X$  and some population measures for the auxiliary variable  $X$ .
- 5) Repeat Steps 2–4; 20000 times for each combination of first- and second-phase sample sizes.
- 6) Compute the bias and  $MSE$  of each estimator for different sample sizes by using

$$Bias(\bar{y}_i) = \bar{\bar{y}}_i - 65 \text{ and } MSE(\bar{y}_i) = \frac{1}{20000} \sum_{j=1}^{20000} (\bar{y}_{ij} - \bar{\bar{y}}_i)^2; \bar{\bar{y}}_i = \frac{1}{20000} \sum_{j=1}^{20000} \bar{y}_{ij}$$

where  $i = CR, DS_1^3(2), DS_2^3(2), DS_3^3(2), DS_4^3(2)$  and  $DS_5^3(2)$ .

Table 3 contains the bias of various estimators whereas Table 4 contains the *MSE* of various estimators.

**Table 3.** Bias of various two-phase sampling estimators.

$\rho$	$n_1$	$n_2$	$\bar{y}_{R(2)}$	$\bar{y}_{DS_1^3(2)}$	$\bar{y}_{DS_2^3(2)}$	$\bar{y}_{DS_3^3(2)}$	$\bar{y}_{DS_4^3(2)}$	$\bar{y}_{DS_5^3(2)}$
-0.9	500	25	0.0886	0.0989	0.0974	0.0990	0.0962	0.0968
		50	0.0159	0.0242	0.0231	0.0243	0.0223	0.0227
		100	0.1785	0.1811	0.1807	0.1811	0.1805	0.1806
	1000	50	-0.0637	-0.0598	-0.0605	-0.0598	-0.0609	-0.0607
		100	0.0124	0.0144	0.0141	0.0144	0.0138	0.0139
		200	-0.0481	-0.0474	-0.0475	-0.0474	-0.0476	-0.0475
-0.5	500	25	-0.0584	-0.0516	-0.0528	-0.0516	-0.0536	-0.0532
		50	0.0005	0.0047	0.0040	0.0048	0.0036	0.0038
		100	0.0611	0.0632	0.0629	0.0632	0.0627	0.0628
	1000	50	0.0267	0.0312	0.0304	0.0312	0.0299	0.0302
		100	0.0614	0.0643	0.0639	0.0643	0.0636	0.0638
		200	-0.0453	-0.0446	-0.0447	-0.0446	-0.0448	-0.0448
0.5	500	25	0.1086	0.1157	0.1144	0.1157	0.1136	0.1141
		50	0.1128	0.1144	0.1140	0.1144	0.1137	0.1139
		100	-0.0076	-0.0051	-0.0054	-0.0050	-0.0057	-0.0056
	1000	50	-0.0448	-0.0398	-0.0406	-0.0398	-0.0412	-0.0409
		100	-0.1156	-0.1126	-0.1130	-0.1126	-0.1134	-0.1132
		200	-0.0675	-0.0668	-0.0669	-0.0668	-0.0670	-0.0670
0.9	500	25	-0.0229	-0.0136	-0.0150	-0.0136	-0.0160	-0.0155
		50	-0.0940	-0.0857	-0.0869	-0.0857	-0.0876	-0.0872
		100	0.1172	0.1211	0.1206	0.1211	0.1202	0.1204
	1000	50	0.0821	0.0862	0.0855	0.0862	0.0850	0.0853
		100	0.0125	0.0139	0.0136	0.0139	0.0135	0.0136
		200	-0.0984	-0.0957	-0.0960	-0.0956	-0.0963	-0.0962

The above table shows that the bias of various estimators fluctuates and is on the relatively higher side.

**Table 4.** *MSE* of various two-phase sampling estimators.

$\rho$	$n_1$	$n_2$	$\bar{y}_{R(2)}$	$\bar{y}_{DS_1^3(2)}$	$\bar{y}_{DS_2^3(2)}$	$\bar{y}_{DS_3^3(2)}$	$\bar{y}_{DS_4^3(2)}$	$\bar{y}_{DS_5^3(2)}$
-0.9	500	25	2.9315	2.4527	2.3638	2.4561	2.3013	2.3317
		50	1.5120	1.2882	1.2404	1.2905	1.2059	1.2226
		100	0.6847	0.5644	0.5450	0.5649	0.5307	0.5375
	1000	50	1.5420	1.3049	1.2560	1.3068	1.2224	1.2392
		100	0.7237	0.6036	0.5813	0.6044	0.5666	0.5740
		200	0.3472	0.2867	0.2768	0.2872	0.2699	0.2733
-0.5	500	25	3.0511	2.5930	2.4944	2.5961	2.4319	2.4648
		50	1.4463	1.2070	1.1614	1.2094	1.1337	1.1488

*Continued on next page*

$\rho$	$n_1$	$n_2$	$\bar{y}_{R(2)}$	$\bar{y}_{DS_1^3(2)}$	$\bar{y}_{DS_2^3(2)}$	$\bar{y}_{DS_3^3(2)}$	$\bar{y}_{DS_4^3(2)}$	$\bar{y}_{DS_5^3(2)}$
-0.5	1000	100	0.7090	0.5796	0.5588	0.5803	0.5454	0.5525
		50	1.5738	1.3242	1.2749	1.3265	1.2400	1.2572
		100	0.7220	0.6083	0.5856	0.6087	0.5709	0.5785
		200	0.3383	0.2798	0.2696	0.2802	0.2631	0.2665
0.5	500	25	3.0231	2.5636	2.4679	2.5666	2.4058	2.4376
		50	1.4745	1.2240	1.1792	1.2261	1.1482	1.1638
	1000	100	0.6860	0.5672	0.5480	0.5683	0.5339	0.5407
		50	1.5257	1.2847	1.2366	1.2854	1.2040	1.2206
	1000	100	0.7258	0.6067	0.5853	0.6073	0.5689	0.5765
		200	0.3377	0.2774	0.2679	0.2777	0.2608	0.2641
0.9	500	25	2.9532	2.5239	2.4298	2.5265	2.3682	2.4001
		50	1.4656	1.2274	1.1804	1.2297	1.1538	1.1693
	1000	100	0.6928	0.5783	0.5581	0.5788	0.5436	0.5505
		50	1.5425	1.3028	1.2558	1.3044	1.2217	1.2381
	1000	100	0.7250	0.6113	0.5886	0.6114	0.5741	0.5817
		200	0.3324	0.2724	0.2631	0.2728	0.2564	0.2596

From the above tables, we can see that all members of the proposed class of estimators are more efficient than the classical two-phase sampling ratio estimator.

We also computed the relative efficiency of various two-phase sampling estimators relative to the two-phase sampling ratio estimator. The relative efficiencies are plotted for various combinations of  $\rho$ ,  $n_1$  and  $n_2$ . The plots of relative efficiencies are given in Figure 2 below.

From the above figure, we can see that the estimator  $\bar{y}_{DS_4^3(2)}$  is the most efficient among the members of the new class. The member  $\bar{y}_{DS_3^3(2)}$  is the worst in the class, but all members are more efficient than the ratio estimator. We will now present the numerical study of different estimators using some real populations.

### 6.2 Study using data from some real populations

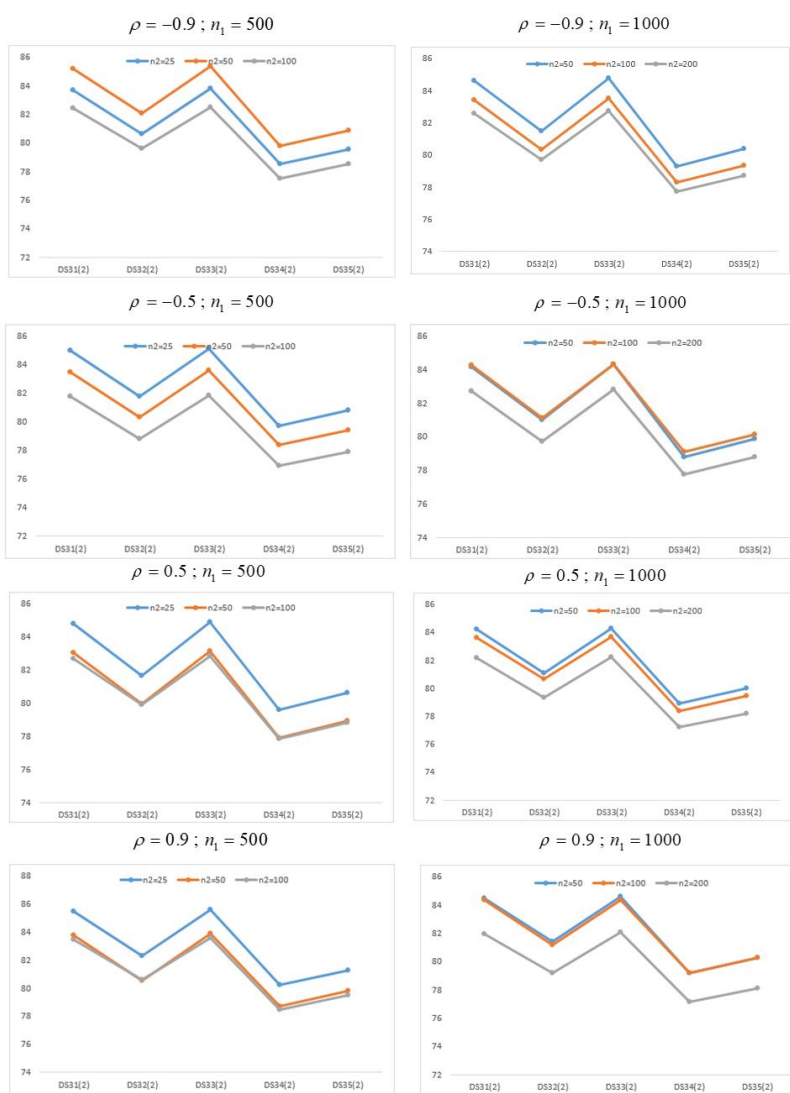
In this subsection, we present a numerical comparison of various estimators using some real populations. The description of the populations used in the study alongside some measures computed from these populations are given in Table 5 below.

**Table 5.** Description of various populations.

Pop.	1	2	3	4	5	6	7	8
Source	[29]	[30]	[31]	[32]	[32]	[31]	[33]	[34]
$Y$	Output	Wheat Area 74	Output	Output	Output	No. of Cultivator	Food Amount	Wheat Area 37
$X$	No. of Workers	Wheat Area 71	No. of Workers	No. of Workers	No. of Workers	No. of Cultivated Area	Total Income	Wheat Area 37
$N$	40	34	20	10	20	15	10	15

*Continued on next page*

Pop.	1	2	3	4	5	6	7	8
$n$	20	20	6	4	12	4	4	6
$\bar{X}$	230.31	208.88	440.81	362.43	202.54	2361.13	88.50	19.47
$\bar{Y}$	5078.61	856.41	6668.63	5729.22	4919.95	704.20	106.31	168.67
$M_d$	125	150	447.52	302.54	109.52	1881	49	131
$\beta_1$	0.97	0.98	0.04	0.91	1.26	0.57	2.18	1.85
$\beta_2$	-0.53	0.09	-1.05	-0.06	0.29	-1.11	4.78	2.83
$S_x$	193.60	150.51	114.69	318.15	187.28	1718.24	82.99	182.56
$C_x$	0.84	0.72	0.26	0.87	0.92	0.73	0.13	0.92
$S_y$	1673.52	733.14	514.52	2090.23	1915.34	735.12	92.21	126.10
$C_y$	0.33	0.86	0.08	0.37	0.39	1.04	0.18	0.75
$\rho$	0.80	0.45	0.95	0.92	0.69	0.18	0.44	0.91



**Figure 2.** Relative efficiency of various two-phase sampling estimators

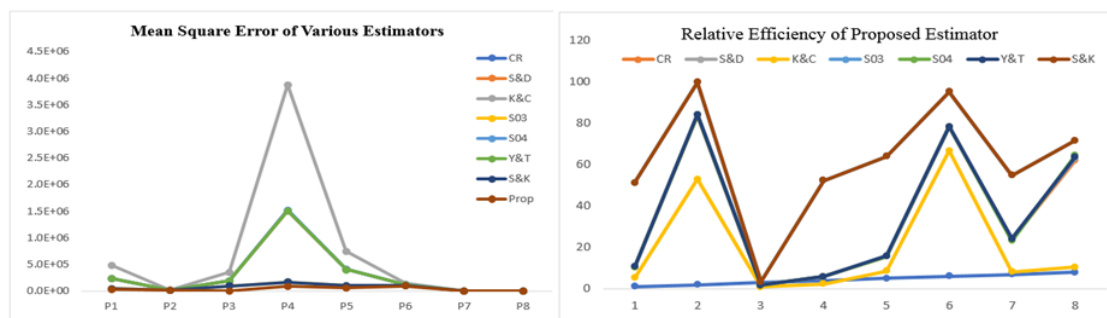
We computed the  $MSE$  of various estimators by using the data on the above-mentioned populations. The results are shown in Table 6 below.



**Table 6.** *MSE* of various estimators.

Pop	CR	S&D	K&C	S 03	S 04	Y&T	S&K	Proposed
1	239659.8	237392.1	480754.6	49130.7	241116.4	237034.8	49130.7	25134.9
2	10539.9	10514.9	16673.8	8853.7	10536.6	10506.1	8853.7	8835.0
3	185284.7	184986.6	354520.2	95865.3	186493.0	185241.7	95865.3	3311.3
4	1515718.0	1504491.0	3882710.0	168872.8	1516529.0	1504032.0	168872.8	88122.5
5	411331.2	406893.8	754125.1	100180.7	409900.2	405269.9	100180.7	64005.7
6	122361.6	122339.6	144004.2	100813.8	122395.1	122344.2	100813.8	95861.5
7	42.5	41.6	122.7	18.4	43.7	41.9	18.4	10.1
8	454.1	448.1	2712.9	392.8	436.2	442.2	392.8	281.5

From the above table we can see that the proposed estimator outperformed all of the other estimators used in the study as it has smallest *MSE* among all of the estimators. We also computed the relative efficiency of the proposed estimator relative to various estimators. The *MSE* and relative efficiency are shown in Figure 3 below.

**Figure 3.** *MSE* and relative efficiency of various estimators

The graph also shows the same results as are provided in Table 6.

## 7. Conclusions

In this paper we have proposed two new families of estimators for estimation of the population mean. The families of estimators have been proposed for single- and two-phase sampling. The expressions for the bias and *MSE* for the proposed families of estimators have been obtained. We have seen that the proposed family of estimators in single-phase sampling is more efficient than the other estimators to which they were compared. This has been shown through analytical and empirical comparisons. The results of the simulation study also support this comparison. We can thus conclude that the proposed single-phase sampling estimator will estimate the population mean with more precision. We have also found that the proposed family of estimators in two-phase sampling is also more efficient than the classical ratio estimator for two-phase sampling. This has been concluded by the results of an extensive simulation study on the two-phase sampling family of estimators. We can, therefore, conclude that the proposed families of estimators will provide more efficient results for estimation of the population mean in single- and two-phase sampling when information on a single auxiliary variable is available.

## Conflict of interest

The authors declare no conflict of interest.

## References

1. W. G. Cochran, The estimation of the yields of cereal experiments by sampling for the ratio gain to total produce, *J. Agric. Sic.*, **30** (1940), 262–275. <https://doi.org/10.1017/S0021859600048012>
2. B. V. S. Sisodia, V. K. Dwivedi, A modified ratio estimator using co-efficient of variation of auxiliary variable, *J. Indian Soc. Agric. Stat.*, **33** (1981), 13–18.
3. B. N. Pandey, V. Dubey, Modified product estimator using coefficient of variation of auxiliary variate, *Assam Stat. Rev.*, **2** (1988), 64–66.
4. S. K. Srivastava, An estimator using auxiliary information in sample surveys, *Calc. Stat. Assoc. Bull.*, **16** (1967), 121–132. <https://doi.org/10.1177/0008068319670205>
5. V. N. Reddy, On ratio and product methods of estimation, *Sankhya B*, **35** (1973), 307–316.
6. J. E. Walsh, Generalization of ratio estimator for population total, *Sankhyā: Indian J. Stat. Ser. B*, **32** (1970), 99–106.
7. H. P. Singh, M. R. Espejo, On linear regression and ratio–product estimation of a finite population mean, *J. Royal Stat. Soc: Ser. D*, **52** (2003), 59–67. <https://doi.org/10.1111/1467-9884.00341>
8. M. Khoshnevisan, R. Singh, P. Chauhan, N. Sawan, F. Smarandache, A general family of estimators for estimating population mean using known value of some population parameter(s), *Far East J. Theor. Stat.*, **22** (2007), 181–191.
9. C. Kadilar, H. Cingi, Ratio estimators in simple random sampling, *Appl. Math. Comput.*, **151** (2004), 893–902.
10. M. Abid, N. Abbas, R. A. K. Sherwani, H. Z. Nazir, Improved ratio estimators for the population mean using non-conventional measures of dispersion, *Pak. J. Stat. Oper. Res.*, **12** (2016), 353–367. <https://doi.org/10.18187/pjsor.v12i2.1182>
11. M. Subzar, S. Maqbool, T. A. Raja, S. A. Mir, M. I. Jeelani, M. A. Bhat, Improved family of ratio type estimators for estimating population mean using conventional and non-conventional location parameters, *Invest. Operacional*, **38** (2018), 510–524.
12. C. Unal, C. Kadilar, Improved family of estimators using exponential function for the population mean in the presence of non-response, *Commun. Stat.-Theory M.*, **50** (2021), 237–248. <https://doi.org/10.1080/03610926.2019.1634818>
13. T. Zaman, C. Kadilar, New class of exponential estimators for finite population mean in two-phase sampling, *Commun. Stat. Theory M.*, **50** (2021), 874–889. <https://doi.org/10.1080/03610926.2019.1643480>
14. T. Zaman, An efficient exponential estimator of the mean under stratified random sampling, *Math. Popul. Stud.*, **28** (2021), 104–121. <https://doi.org/10.1080/08898480.2020.1767420>
15. M. Tayyab, M. Noor-ul-Amin, M. Hanif, Quartile pair ranked set sampling: development and estimation, *Proc. Natl. Acad. Sci., India, Sect. A Phys. Sci.*, **91** (2021), 111–116. <https://doi.org/10.1007/s40010-019-00651-2>
16. J. N. K. Rao, On double sampling for stratification and analytic surveys, *Biometrika*, **60** (1973), 125–133. <https://doi.org/10.1093/biomet/60.1.125>

17. C. E. Sarndal, B. Swensson, A general view of estimation for two-phases of selection with applications to two-phase sampling and non-response, *Int. Stat. Rev.*, **55** (1987), 279–294. <https://doi.org/10.2307/1403406>
18. J. Sahoo, L. N. Sahoo, On the efficiency of four chain-type estimators in two-phase sampling under a model, *Statistics*, **25** (1994), 361–366. <https://doi.org/10.1080/02331889408802459>
19. G. N. Singh, L. N. Upadhyaya, A class of modified chain type estimators using two auxiliary variables in two-phase sampling, *Metron*, **1** (1995), 117–125.
20. M. Samiuddin, M. Hanif, Estimation of population mean in single- and two-phase sampling with or without additional information, *Pak. J. Stat.*, **23** (2007), 99–118.
21. Z. Ahmad, M. Q. Shahbaz, M. Hanif, *Two-phase sampling*, UK: Cambridge Scholars Publishing, 2013.
22. M. Hanif, M. Q. Shahbaz, M. Ahmed, *Sampling techniques: Methods and applications*, USA: Nova Science Publisher, 2018.
23. H. P. Singh, R. Tailor, M. S. Kakran, An improved estimator of population mean using power transformation, *J. Indian Soc. Agri. Stat.*, **58** (2004), 223–230.
24. Z. Z. Yan, B. Tian, Ratio method to the mean estimation using coefficient of skewness of auxiliary variable, In: *ICICA 2010: Information computing and applications*, Berlin, Heidelberg: Springer, 2010, 103–110. [https://doi.org/10.1007/978-3-642-16339-5\\_14](https://doi.org/10.1007/978-3-642-16339-5_14)
25. G. N. Singh, On the improvement of product method of estimation in sample surveys, *J. Indian Soc. Agri. Stat.*, **56** (2003), 267–275.
26. J. Subramani, G. Kumarapandiyan, A new modified ratio estimator for estimation of population mean when median of the auxiliary variable is known, *Pak. J. Stat. Oper. Res.*, **9** (2013), 137–145. <https://doi.org/10.18187/pjsor.v9i2.486>
27. S. Mohanty, Combination of regression and ratio estimate, *J. Ind. Statist. Assoc.*, **5** (1967), 16–19.
28. M. Hanif, N. Hamad, M. Q. Shahbaz, Some new regression types estimators in two-phase sampling, *World Appl. Sci. J.*, **8** (2010), 799–803.
29. P. Mukhopandhyay, *Theory and methods of survey sampling*, 2 Eds., 2009.
30. D. Singh, F. S. Chaudhary, *Theory and analysis of sample survey designs*, New Delhi: New Age International Publisher, 1986.
31. M. N. Murthy, *Sampling theory and methods*, statistical publishing society, 1967.
32. P. Mukhopandhyay, *Theory and methods of survey sampling*, New Delhi: Prentice Hall of India, 1998.
33. L. R. Schaeffer, Pseudo expectation approach to variance component estimation, *J. Dairy Sci.*, **69** (1986), 2884–2889.
34. R. K. Som, *A manual of sampling techniques*, London: Heineman Educational Books Limited, 1973.



AIMS Press

© 2022 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)