Pak. J. Statist.
2023 Vol. 39(3), 271-288

RATIO AND REGRESSION ESTIMATORS OF THE GENERAL POPULATION PARAMETER USING TWO AUXILIARY VARIABLES

Muhammad Qaiser Shahbaz[2] and **Muhammad Hanif**[2]
[1] Department of Statistics, King Abdulaziz University
Jeddah, Saudi Arabia. Email: mkmohamad@kau.edu.sa
[2] National College of Business Administration & Economics
Lahore, Pakistan. Email: drhanif@ncbae.edu.pk

ABSTRACT

Some new ratio and regression type estimators of the general population parameter has been proposed by using the information of two auxiliary variables. The estimators have been proposed for single- and two-phase sampling. The proposed estimators can be used to estimate population mean, population variance and any combination of the two parameters. The mean square error of the proposed estimators has been obtained. The numerical study has been conducted via simulation and real populations to see the performance of the proposed estimators. It is found that the proposed estimators perform better than the existing estimators in estimation of population mean and variance.

KEYWORDS

General population parameter, Auxiliary variable, Ratio and Regression estimators, Mean Square Error, Simulation.

Mathematics Subject Classification: 62D05, 62F10, 62J05.

1. INTRODUCTION

Parameter estimation has been an interesting area of research within the domain of survey sampling. Various researchers have proposed different estimators for estimation of the population mean and the population variance by using information of the auxiliary variables. Cochran (1940) have used the classical ratio estimator to estimate the yield of cereal. The classical regression estimator given as

$$\bar{y}_{lr} = \bar{y} + \beta\left(\bar{X} - \bar{x}\right), \tag{1}$$

where $\bar{y}$ and $\bar{x}$ are sample means of the study and auxiliary variables respectively and $\bar{X}$ is the population mean of the auxiliary variable, is a popular estimator of mean using information of a single auxiliary variable. Isaki (1983) has proposed the classical ratio and regression estimators of the population variance. The estimators are given as

$$s_R^2 = \frac{s_y^2}{s_x^2} \times S_x^2 \text{ and } s_{lr}^2 = s_y^2 + \gamma\left(S_x^2 - s_x^2\right), \tag{2}$$

where s_y^2 and s_x^2 are sample variances of the study and auxiliary variable and S_x^2 is the population variance of the auxiliary variable. The ratio and regression estimators of mean and variance has been extended by several authors by using information of more than one auxiliary variables. Kiregyera (1984) has proposed a regression type estimator on the basis of two auxiliary variables. Abu-Dayyeh et al. (2003) and Abu-Dayyeh and Ahmed (2005) have proposed ratio and regression estimators of mean and variance using two auxiliary variables. Another estimator of population mean by using information of two auxiliary variables has been proposed by Kadilar and Cingi (2005). Some linear ratio and regression type estimators of the population mean are given by Singh and Espejo (2003). Samiuddin and Hanif (2007) have suggested some ratio and regression type estimators of population mean by using information of two auxiliary variables. Grover and Kaur (2021) have used robust and non-robust regression methods to propose some estimators of mean using two auxiliary variables.

Various authors have paid the attention to propose estimators of variance with the view to improve the efficiency of the estimates. A general class of estimators of population variance by using information of two auxiliary variables has been proposed by Adichwal et al. (2015). Al-Marshadi et al. (2018) have proposed a logarithmic type estimator of variance using information of single and multiple auxiliary variables. Kumar and Choudhry (2023) have suggested an improved estimator of variance on the basis of two auxiliary variables.

Some authors have proposed the estimators for some combinations of population mean and variance in single-phase sampling. Singh et al. (2018) has suggested an improved estimator of population coefficient of variation in single-phase sampling. Adichwal et al. (2022) have proposed an estimator for general population parameter in single-phase sampling. Shahbaz et al. (2023) have proposed some estimators of the general population parameter in single-phase and two-phase sampling. More details about estimators of mean and variance can be found in Ahmed et al. (2013) and Hanif et al. (2018).

In this paper we have proposed some ratio and regression type estimators for estimation of the general population parameter by using information of two auxiliary variables. These estimators have been proposed for single- and two-phase sampling. The development of these estimators requires certain methodology and notations which are discussed in the following section.

2. METHODOLOGY AND NOTATIONS

In this section, we have discussed the methodology and notations to be used in this paper. Suppose that the units of a population are labeled as $U_1, U_2, \ldots, U_N$ while the values of some variable of interest are $Y_1, Y_2, \ldots, Y_N$. Suppose, further, that the estimation of some general population parameter $t_{(a,b)} = \bar{Y}^a (S_y^2)^{b/2}$ is required, where $\bar{Y} = N^{-1}\sum_{i=1}^{N} Y_i$ and $S_y^2 = \left(N-1\right)^{-1}\sum_{i=1}^{N}\left(Y_i - \bar{Y}\right)^2$ are, respectively, the population mean and variance of Y. It is interesting to note that the general parameter $t_{(a,b)}$ becomes to the population mean for $a = 1$ and $b = 0$, it transforms to the population variance for $a = 0$ and $b = 2$, and to the coefficient of variation for $a = -1$ and $b = 1$. A simple estimator of $t_{(a,b)}$ is immediately written as $\hat{t}_{(a,b)} = \bar{y}^a (s_y^2)^{b/2}$, where $\bar{y}$ and s_y^2 are, respectively, the sample mean and variance of the study variable Y.

The efficiency of the estimates can always be increased by using information of some auxiliary variables. When information of a single auxiliary variable is known, then the conventional regression estimator, using a sample of size *n*, is given in (1). The mean square error (MSE) of (1) is

$$MSE\left(\bar{y}_{lr}\right) = \theta S_y^2\left(1-\rho_{yx}^2\right), \quad (3)$$

where $\theta = n^{-1} - N^{1}$ and $\rho = S_{yx}\big/\sqrt{S_x^2 S_y^2}$ is population correlation coefficient between X and Y.

In some situations, the population information of auxiliary variable is not available and in such situations the regression 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 of an auxiliary variable is recorded. A sub-sample of size $n_2 < n_1$ is drawn from the first-phase sample, and information of the auxiliary variable and the study variable is recorded. The conventional regression estimator, in two phase sampling, is given as

$$\bar{y}_{lr(2)} = \bar{y}_{(2)} + \beta\left[\bar{x}_{(1)} - \bar{x}_{(2)}\right], \quad (4)$$

where $\bar{y}_{(2)} = n_2^{-1}\sum_{i=1}^{n_2} y_i$ is the second phase sample mean of study variable Y, $\bar{x}_{(2)} = n_2^{-1}\sum_{i=1}^{n_2} x_i$ is the second-phase sample mean of auxiliary variable X, and $\bar{x}_1 = n_{(1)}^{-1}\sum_{i=1}^{n_1} x_i$ is the first-phase sample mean of auxiliary variable X. The *MSE* of two-phase sampling regression estimator is

$$MSE\left(\bar{y}_{lr(2)}\right) = \bar{Y}^2 C_y^2\left[\theta_2\left(1-\rho_{yx}^2\right) + \theta_1\rho_{yx}^2\right], \quad (5)$$

where $\theta_2 = n_2^{-1} - N^{-1}$ and $\theta_1 = n_1^{-1} - N^{-1}$.

The regression estimator of population variance is given by Isaki (1981) as

$$s_{y(lr)}^2 = s_y^2 + \gamma\left(S_x^2 - s_x^2\right), \quad (6)$$

where γ is a constant, S_x^2 and s_x^2 are, respectively, the population and the sample variances of the auxiliary variable, and s_y^2 is the sample variance of *Y*. The estimator for two-phase sampling can be easily written. Several modifications of the two-phase sampling regression estimator of mean are given in Ahmed at al. (2013).

The derivation of bias and *MSE* of the estimators of the mean and the variance require certain notations. In this paper, we will assume that the sample mean and the sample variance of study and auxiliary variable are connected with the population mean and the population variance as $\bar{y} = \bar{Y}\left(1+\varepsilon_y\right)$, $\bar{x} = \bar{X}\left(1+\varepsilon_x\right)$, $s_y^2 = S_y^2\left(1+e_y\right)$ and $s_z^2 = S_z^2\left(1+e_z\right)$.

The relation between sample estimates and the population parameters in case of two-phase sampling is $\bar{y}_{(2)} = \bar{Y}\left(1+\varepsilon_{y(2)}\right)$, $\bar{x}_{(1)} = \bar{X}\left(1+\varepsilon_{x(1)}\right)$, $\bar{x}_{(2)} = \bar{X}\left(1+\varepsilon_{x(2)}\right)$, $s_{y(2)}^2 = S_y^2\left(1+e_{y(2)}\right)$ and $s_{z(2)}^2 = S_z^2\left(1+e_{z(2)}\right)$.

The expected values of error terms $\varepsilon's$ and $e's$ are all zero. Some additional expectations for single- and two-phase sampling and for single auxiliary variable, are

$$\left.\begin{array}{l} E\left(\varepsilon_y^2\right) = \theta C_y^2 \;;\; E\left(\varepsilon_x^2\right) = \theta C_x^2 \;;\; E\left(e_y^2\right) = \theta\varphi_{400}^* \;;\; E\left(e_z^2\right) = \theta\varphi_{004}^* \\ E\left(\varepsilon_y\varepsilon_x\right) = \theta\rho_{yx}C_xC_y \;;\; E\left(\varepsilon_x e_z\right) = \theta\varphi_{012}C_x \;; E\left(\varepsilon_y e_y\right) = \theta\varphi_{300}C_y \\ E\left(e_y\varepsilon_x\right) = \theta\varphi_{210}C_x \;; E\left(\varepsilon_y e_z\right) = \theta\varphi_{102}C_y \;; E\left(e_y e_z\right) = \theta\varphi_{202}^* \end{array}\right\} \quad (7)$$

$$\left.\begin{array}{l} E\left(\varepsilon_{y(2)}^2\right) = \theta_2 C_y^2 \;; E\left(e_{y(2)}^2\right) = \theta_2\varphi_{400}^* \;; E\left(\varepsilon_{y(2)}e_{y(2)}\right) = \theta_2\varphi_{300}C_y \\ E\left[\left(\varepsilon_{x(2)} - \varepsilon_{x(1)}\right)^2\right] = \left(\theta_2 - \theta_1\right)C_x^2 \;;\; E\left[\left(e_{z(2)} - e_{z(1)}\right)^2\right] = \left(\theta_2 - \theta_1\right)\varphi_{004}^* \\ E\left[\varepsilon_{y(2)}\left(\varepsilon_{x(2)} - \varepsilon_{x(1)}\right)\right] = \left(\theta_2 - \theta_1\right)\rho_{yx}C_xC_y \;; \\ \qquad E\left[e_{y(2)}\left(e_{z(2)} - e_{z(1)}\right)\right] = \left(\theta_2 - \theta_1\right)\varphi_{202}^* \\ E\left[\varepsilon_{y(2)}\left(e_{z(2)} - e_{z(1)}\right)\right] = \left(\theta_2 - \theta_1\right)\varphi_{102}C_y \;; \\ \qquad E\left[e_{y(2)}\left(\varepsilon_{x(2)} - \varepsilon_{x(1)}\right)\right] = \left(\theta_2 - \theta_1\right)\varphi_{210}C_x \\ E\left[\varepsilon_{x(2)}\left(e_{z(2)} - e_{z(1)}\right)\right] = \left(\theta_2 - \theta_1\right)\varphi_{012}C_x \;; \\ \qquad \varphi_{rst}^* = \left(\varphi_{rst} - 1\right) \;;\; \varphi_{rst} = \mu_{rst} \Big/ \left(\mu_{200}^{r/2}\mu_{020}^{s/2}\mu_{002}^{t/2}\right) \end{array}\right\} \quad (8)$$

Also, $\mu_{rst} = \left(N-1\right)^{-1}\sum_{i=1}^{N}\left(y_i - \bar{Y}\right)^r\left(x_i - \bar{X}\right)^s\left(z_i - \bar{Z}\right)^t$. We will now, propose some new estimators for single-phase sampling.

3. ESTIMATORS FOR SINGLE-PHASE SAMPLING

In this section, we will propose ratio and regression type estimators of general population parameter for single-phase sampling. These estimators are proposed in the following sub-sections.

3.1 The Regression Type Estimator

The proposed regression estimator of the general population parameter in single-phase sampling is

$$t_{lr(1)} = \bar{y}^a\left(s_y^2\right)^{b/2}\left[1 + \alpha\left(\bar{X} - \bar{x}\right) + \beta\left(S_z^2 - s_z^2\right)\right]. \quad (9)$$

We will, now, obtain the bias and mean square error of the proposed estimator. For this, we write the estimator (9) as

$$t_{lr(1)} = \left[\bar{Y}^a\left(1+\varepsilon_y\right)^a\left(S_y^2\right)^{b/2}\left(1+e_y\right)^{b/2}\right]\left[1-\alpha\bar{X}\varepsilon_x-\beta S_z^2 e_z\right].$$

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

$$t_{lr(1)} = t_{(a,b)}\left(1+a\varepsilon_y+\frac{b}{2}e_y+\frac{ab}{2}\varepsilon_y e_y\right)\left(1-\alpha\bar{X}\varepsilon_x-\beta S_z^2 e_z\right)$$

or

$$\begin{aligned} t_{lr(1)} - t_{(a,b)} = t_{(a,b)}\Bigg[& a\varepsilon_y+\frac{b}{2}e_y+\frac{ab}{2}\varepsilon_y e_y-\alpha\bar{X}\varepsilon_x-a\alpha\bar{X}\varepsilon_y\varepsilon_x \\ & -\frac{\alpha b}{2}\bar{X}e_y\varepsilon_x-\beta S_z^2 e_z-a\beta S_z^2\varepsilon_y e_z-\frac{b\beta}{2}S_z^2 e_y e_z\Bigg]. \end{aligned} \tag{10}$$

Applying expectation, and using (7), the bias of the proposed regression estimator of the general population parameter is

$$\begin{aligned} Bias\left(t_{lr(1)}\right) = t_{(a,b)}\Bigg[& \frac{ab}{2}\theta\varphi_{300}C_y-a\alpha\theta\bar{X}\rho_{yx}C_xC_y \\ & -\frac{\alpha b}{2}\theta\bar{X}\varphi_{210}C_x-a\beta\theta S_z^2\varphi_{102}C_y-\frac{b\beta}{2}\theta S_z^2\varphi_{202}^*\Bigg]. \end{aligned}$$

Now, to derive the mean square error, we have

$$\begin{aligned} \left(t_{lr(1)}-t_{(a,b)}\right)^2 = t_{(a,b)}^2\Big[& a^2\varepsilon_y^2+\left(b^2/4\right)e_y^2+\alpha^2\bar{X}^2\varepsilon_x^2+\beta^2 S_z^4 e_z^2+ab\varepsilon_y e_y \\ & -2a\alpha\bar{X}\varepsilon_y\varepsilon_x-2a\beta S_z^2\varepsilon_y e_z-\alpha b\bar{X}e_y\varepsilon_x-b\beta S_z^2 e_y e_z+2\alpha\beta\bar{X}S_z^2\varepsilon_x e_z\Big]. \end{aligned}$$

Applying expectation, and using (7), the mean square error of the proposed regression estimator of the general population parameter is

$$\begin{aligned} MSE\left(t_{lr(1)}\right) = \theta t_{(a,b)}^2\Big[& a^2C_y^2+\left(b^2/4\right)\varphi_{400}^*+\alpha^2\bar{X}^2C_x^2+\beta^2 S_z^4\varphi_{004}^* \\ & +ab\varphi_{300}C_y-2a\alpha\bar{X}\rho_{yx}C_xC_y-2a\beta S_z^2\varphi_{102}C_y \\ & -\alpha b\bar{X}\varphi_{210}C_x-b\beta S_z^2\varphi_{202}^*+2\alpha\beta\bar{X}S_z^2\varphi_{012}C_x\Big]. \end{aligned} \tag{11}$$

We will now obtain the optimum values of α and β which minimizes (11). For this, we differentiate (11) with respect to α and β, equate the derivatives to zero and solve the resulting equations. The derivatives of (11) with respect to α and β are

$$\frac{\partial}{\partial\alpha}MSE\left(t_{lr(1)}\right) = \theta t_{(a,b)}^2\bar{X}C_x\left(2\bar{X}C_x\alpha+2S_z^2\varphi_{201}\beta-2a\rho_{yx}C_y-b\varphi_{210}\right)$$

$$\frac{\partial}{\partial\beta}MSE\left(t_{lr(1)}\right) = \theta t_{(a,b)}^2 S_z^2\left(2\bar{X}C_x\varphi_{012}\alpha+2S_z^2\varphi_{004}\beta-2aC_y\varphi_{102}-b\varphi_{202}\right).$$

Equating the above derivatives to zero, the normal equations for α and β are

$$2\bar{X}C_x\alpha + 2S_z^2\varphi_{012}\beta = 2a\rho_{yx}C_y + b\varphi_{201}$$

and

$$2\bar{X}C_x\varphi_{012}\alpha + 2S_z^2\varphi_{004}\beta = 2aC_y\varphi_{102} + b\varphi_{202} .$$

Solving the above equations, the optimum values of α and β which minimizes (11) are

$$\alpha_0 = \frac{2aC_y\left(\rho_{yx}\varphi_{004}^* - \varphi_{012}\varphi_{102}\right) + b\left(\varphi_{004}^*\varphi_{210} - \varphi_{012}\varphi_{202}^*\right)}{2\bar{X}C_x\left(\varphi_{004}^* - \varphi_{012}^2\right)} \tag{12}$$

and

$$\beta_0 = \frac{2aC_y\left(\varphi_{102} - \rho_{yx}\varphi_{012}\right) + b\left(\varphi_{202}^* - \varphi_{012}\varphi_{210}\right)}{2S_z^2\left(\varphi_{004}^* - \varphi_{012}^2\right)} . \tag{13}$$

Using these values in (11), the minimum mean square error of $t_{lr(1)}$ is

$$MSE\left(t_{lr(1)}\right) = \frac{\theta t_{(a,b)}^2}{\left(\varphi_{004}^* - \varphi_{012}^2\right)}\left(a^2C_y^2 f_1^* + abC_y f_2^* + \frac{b^2}{4} f_3^*\right), \tag{14}$$

where

$$f_1^* = \varphi_{004}^*\left(1-\rho_{yx}^2\right) - \varphi_{012}^2 + 2\rho_{yx}\varphi_{012}\varphi_{102} - \varphi_{102}^2$$

$$f_2^* = \varphi_{300}\left(\varphi_{004}^* - \varphi_{012}^2\right) - \rho_{yx}\left(\varphi_{004}^*\varphi_{210} - \varphi_{012}\varphi_{202}^*\right) - \varphi_{102}\left(\varphi_{202}^* - \varphi_{012}\varphi_{210}\right)$$

and $$f_3^* = \varphi_{400}^*\left(\varphi_{004}^* - \varphi_{012}^2\right) + \varphi_{210}\left(2\varphi_{012}\varphi_{202}^* - \varphi_{004}^*\varphi_{210}\right) - \varphi_{202}^{*2} .$$

Some special cases of the estimators are readily written. For example, if we have $a=1$ and $b=0$ in (9) then we have a regression type estimator of mean with

$$MSE\left(t_{lr(1)M}\right) = \theta\bar{Y}^2C_y^2 f_1^*\left(\varphi_{004}^* - \varphi_{012}^2\right)^{-1} . \tag{15}$$

Again, if we have $a=0$ and $b=2$ in (9) then we have a regression type estimator of mean with

$$MSE\left(t_{lr(1)M}\right) = \theta S_y^4 f_3^*\left(\varphi_{004}^* - \varphi_{012}^2\right)^{-1} . \tag{16}$$

The expression for the mean square error for other combinations of a and b can be similarly obtained. We will now propose a ratio type estimator of the general population parameter in single-phase sampling.

3.2 The Ratio Type Estimator

The proposed ratio type estimator of the general population parameter is

$$t_{R(1)} = \bar{y}^a\left(s_y^2\right)^{b/2}\left(\bar{X}/\bar{x}\right)^{\gamma}\left(S_z^2/s_z^2\right)^{\delta} . \tag{17}$$

Using the error representations, the above estimator can be written as

$$t_{R(1)} = \left[\bar{Y}^a \left(S_y^2\right)^{b/2} \left(1+\varepsilon_y\right)^a \left(1+e_y\right)^{b/2}\right]\left(1+\varepsilon_x\right)^{-\gamma}\left(1+e_z\right)^{-\delta}.$$

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

$$t_{R(1)} = \left[\bar{Y}^a \left(S_y^2\right)^{b/2} \left(1+a\varepsilon_y\right)\left(1+\frac{b}{2}e_y\right)\right]\left(1-\gamma\varepsilon_x\right)\left(1-\delta e_z\right)$$

or

$$\begin{aligned} t_{R(1)} - t_{(a,b)} = t_{(a,b)}\Big[& a\varepsilon_y + \frac{b}{2}e_y + \frac{ab}{2}\varepsilon_y e_y - \gamma\varepsilon_x - a\gamma\varepsilon_y\varepsilon_x - \gamma\frac{b}{2}e_y\varepsilon_x \\ & -\delta e_z - a\delta\varepsilon_y e_z - \frac{b}{2}\delta e_y e_z + \gamma\delta\varepsilon_x e_z \Big]. \end{aligned} \tag{18}$$

Applying expectations and using (7), the bias of $t_{R(1)}$ is

$$\begin{aligned} Bias\left(t_1^*\right) = \theta t_{(a,b)}\Big[& \frac{ab}{2}\varphi_{300}C_y - a\gamma\rho_{yx}C_yC_x - \gamma\frac{b}{2}\varphi_{210}C_x \\ & -a\delta\varphi_{102}C_y - \frac{b}{2}\delta\varphi_{202}^* + \theta\gamma\delta\varphi_{012}C_x \Big]. \end{aligned}$$

The mean square error of $t_{R(1)}$ is obtained by squaring (18), applying expectation and using (7). Now,

$$\begin{aligned} \left(t_{R(1)} - t_{(a,b)}\right)^2 = t_{(a,b)}^2\Big[& a^2\varepsilon_y^2 + \frac{b^2}{4}e_y^2 + \gamma^2\varepsilon_x^2 + \delta^2 e_z^2 + ab\varepsilon_y e_y \\ & -2a\gamma\varepsilon_y\varepsilon_x - 2a\delta\varepsilon_y e_z - b\gamma e_y\varepsilon_x - b\delta e_y e_z + 2\gamma\delta\varepsilon_x e_z \Big], \end{aligned}$$

so

$$\begin{aligned} MSE\left(t_{R(1)}\right) = E\left(t_{R(1)} - t_{(a,b)}\right)^2 = \theta t_{(a,b)}^2\Big[& a^2C_y^2 + \frac{b^2}{4}\varphi_{400}^* + \gamma^2C_x^2 \\ & +\delta^2\varphi_{004}^* + ab\varphi_{300}C_y - 2a\gamma\rho_{yx}C_yC_x - 2a\delta\varphi_{102}C_y \\ & -b\gamma\varphi_{210}C_x - b\delta\varphi_{202}^* + 2\gamma\delta\varphi_{012}C_x \Big]. \end{aligned} \tag{19}$$

The optimum values of γ and δ which minimizes (17) are $\gamma_0 = \bar{X}\alpha_0$ and $\delta_0 = S_z^2\beta_0$, where α_0 and β_0 are given in (12) and (13), respectively. Using the optimum values of γ and δ in (17), it is easy to show that the minimum mean square error of (15) is same as given in (14).

We will now propose the ratio and regression type estimators of the general population parameter in two-phase sampling.

4. ESTIMATORS FOR TWO-PHASE SAMPLING

In this section, we will propose ratio and regression type estimators of general population parameter for two-phase sampling. These estimators are proposed in the following sub-sections.

4.1 The Regression Type Estimator

The proposed regression estimator of the general population parameter in two-phase sampling is

$$t_{lr(2)} = \bar{y}_{(2)}^{2}\left(s_{y(2)}^{2}\right)^{b/2}\left[1+\alpha\left(\bar{x}_{(1)}-\bar{x}_{(2)}\right)+\beta\left(s_{z(1)}^{2}-s_{z(2)}^{2}\right)\right] \tag{20}$$

It is easy to see that the proposed estimator (20) reduces to the regression type estimator of mean in two-phase sampling for $\left(a,b,\alpha_{(2)},\beta_{(2)}\right)=\left(1,0,\alpha_{(2)},0\right)$. The estimator (20) reduces to the regression type estimator of variance in two-phase sampling for $\left(a,b,\alpha_{(2)},\beta_{(2)}\right)=\left(0,2,0,\beta_{(2)}\right)$.

Now, to derive the bias and *MSE* of (27), we write the estimator (27), using error notations, as

$$t_{lr(2)}=\left[\bar{Y}^{a}\left(S_{y}^{2}\right)^{b/2}\left(1+\varepsilon_{y(2)}\right)^{a}\left(1+e_{y(2)}\right)^{b/2}\right]\left[1-\alpha\bar{X}\left(\varepsilon_{x(2)}-\varepsilon_{x(1)}\right)-\beta S_{z}^{2}\left(e_{z(2)}-e_{z(1)}\right)\right].$$

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

$$\begin{aligned} t_{lr(2)} = t_{(a,b)}\Bigg[&1+a\varepsilon_{y(2)}+\frac{b}{2}e_{y(2)}+\frac{ab}{2}\varepsilon_{y(2)}e_{y(2)}-\alpha_{(2)}\bar{X}\left(\varepsilon_{x(2)}-\varepsilon_{x(1)}\right)\\ &-a\alpha_{(2)}\bar{X}\varepsilon_{y(2)}\left(\varepsilon_{x(2)}-\varepsilon_{x(1)}\right)-\frac{b}{2}\alpha_{(2)}\bar{X}e_{y(2)}\left(\varepsilon_{x(2)}-\varepsilon_{x(1)}\right)\\ &-\beta_{(2)}S_{z}^{2}\left(e_{z(2)}-e_{z(1)}\right)-a\beta_{2}S_{z}^{2}\varepsilon_{y(2)}\left(e_{z(2)}-e_{z(1)}\right)\\ &-\frac{b}{2}\beta_{(2)}S_{x}^{2}e_{y(2)}\left(e_{z(2)}-e_{z(1)}\right)\Bigg]\end{aligned}$$

or

$$\begin{aligned} t_{lr(2)}-t_{(a,b)} = t_{(a,b)}\Bigg[&a\varepsilon_{y(2)}+\frac{b}{2}e_{y(2)}+\frac{ab}{2}\varepsilon_{y(2)}e_{y(2)}-\alpha_{(2)}\bar{X}\left(\varepsilon_{x(2)}-\varepsilon_{x(1)}\right)\\ &-a\alpha_{(2)}\bar{X}\varepsilon_{y(2)}\left(\varepsilon_{x(2)}-\varepsilon_{x(1)}\right)-\frac{b}{2}\alpha_{(2)}\bar{X}e_{y(2)}\left(\varepsilon_{x(2)}-\varepsilon_{x(1)}\right)\\ &-\beta_{(2)}S_{z}^{2}\left(e_{z(2)}-e_{z(1)}\right)-a\beta_{2}S_{z}^{2}\varepsilon_{y(2)}\left(e_{z(2)}-e_{z(1)}\right)\\ &-\frac{b}{2}\beta_{(2)}S_{x}^{2}e_{y(2)}\left(e_{z(2)}-e_{z(1)}\right)\Bigg].\end{aligned}$$

Applying expectation on both side of above equation, the bias of the proposed estimator in two-phase sampling is

$$Bias\left(t_{lr(2)}\right)=t_{(a,b)}\left[\frac{ab}{2}\theta_2\varphi_{300}C_y-a\left(\theta_2-\theta_1\right)\alpha_{(2)}\bar{X}\rho_{yx}C_yC_x\right.$$
$$-\frac{\alpha_{(2)}b}{2}\bar{X}\left(\theta_2-\theta_1\right)\varphi_{210}C_x-a\beta_2\left(\theta_2-\theta_1\right)S_z^2\varphi_{102}C_y$$
$$\left.-\frac{b}{2}\left(\theta_2-\theta_1\right)\beta_{(2)}S_z^2\varphi_{202}^*\right].$$

Again,

$$\left(t_{lr(2)}-t_{(a,b)}\right)^2=t_{(a,b)}^2\left[a^2\varepsilon_{y(2)}^2+\frac{b^2}{4}e_{y(2)}^2+\alpha_{(2)}^2\bar{X}^2\left(\varepsilon_{x(2)}-\varepsilon_{x(1)}\right)^2\right.$$
$$+\beta_{(2)}^2S_z^4\left(e_{z(2)}-e_{z(1)}\right)^2+ab\varepsilon_{y(2)}e_{y(2)}-2a\alpha_2\bar{X}\varepsilon_{y(2)}\left(\varepsilon_{x(2)}-\varepsilon_{x(1)}\right)$$
$$-2a\beta_{(2)}S_z^2\varepsilon_{y(2)}\left(e_{z(2)}-e_{z(1)}\right)-\alpha_{(2)}b\bar{X}e_{y(2)}\left(\varepsilon_{x(2)}-\varepsilon_{x(1)}\right)$$
$$\left.-b\beta_{(2)}S_z^2e_{y(2)}\left(e_{z(2)}-e_{z(1)}\right)+2\alpha_{(2)}\beta_{(2)}\bar{X}S_z^2\left(\varepsilon_{x(2)}-\varepsilon_{x(1)}\right)\left(e_{z(2)}-e_{z(1)}\right)\right].$$

Applying expectation and using (8), the mean square error of proposed regression type estimator in two-phase sampling is

$$MSE\left(t_{lr(2)}\right)=t_{(a,b)}^2\left[a^2\theta_2C_y^2+\left(b^2/4\right)\theta_2\varphi_{400}^*+\alpha_{(2)}^2\bar{X}^2\left(\theta_2-\theta_1\right)C_x^2\right.$$
$$+\beta_{(2)}^2S_z^4\left(\theta_2-\theta_1\right)\varphi_{004}^*+ab\theta_2\varphi_{300}C_y-2a\alpha_2\bar{X}\left(\theta_2-\theta_1\right)\rho_{yx}C_yC_x$$
$$-2a\beta_{(2)}S_z^2\left(\theta_2-\theta_1\right)\varphi_{102}C_y-\alpha_{(2)}b\bar{X}\left(\theta_2-\theta_1\right)\varphi_{210}C_x \tag{21}$$
$$\left.-b\beta_{(2)}S_z^2\left(\theta_2-\theta_1\right)\varphi_{202}^*+2\alpha_{(2)}\beta_{(2)}\bar{X}S_z^2\left(\theta_2-\theta_1\right)\varphi_{012}C_x\right].$$

The optimum values of $\alpha_{(2)}$ and $\beta_{(2)}$ which minimizes (21) are same as given in (12) and (13). Using these optimum values in (21), the mean square error of the proposed regression type estimator for two phase sampling is

$$MSE\left(t_{lr(2)}\right)=\frac{t_{(a,b)}^2}{\left(\varphi_{004}^*-\varphi_{012}^2\right)}\left(a^2C_y^2f_{1(2)}^*+abC_yf_{2(2)}^*+\frac{b^2}{4}f_{3(2)}^*\right), \tag{22}$$

where

$$f_{1(2)}^*=\theta_2\varphi_{004}^*\left(1-\rho_{yx}^2\right)+\theta_1\rho_{yx}^2\varphi_{004}^*-\theta_2\varphi_{012}^2$$
$$-\left(\theta_2-\theta_1\right)\left(\varphi_{102}^2-2\rho\varphi_{012}\varphi_{102}\right)\ f_{2(2)}^*=\theta_2\varphi_{300}\left(\varphi_{004}^*-\varphi_{012}^2\right)$$
$$-\left(\theta_2-\theta_1\right)\left[\rho_{xy}\left(\varphi_{004}^*\varphi_{210}-\varphi_{012}\varphi_{202}^*\right)+\varphi_{102}\left(\varphi_{202}^*-\varphi_{012}\varphi_{210}\right)\right]$$

and

$$f^*_{3(2)} = \theta_2\varphi^*_{400}\left(\varphi^*_{004} - \varphi^2_{012}\right) - \left(\theta_2 - \theta_1\right)\left(\varphi^*_{004}\varphi^2_{210} - 2\varphi_{012}\varphi_{210}\varphi^*_{202} + \varphi^{*2}_{202}\right).$$

The expression for the mean square error of specific cases of $t_{lr(2)}$ can be readily written. For example, using $(a,b)=(1,0)$ in (22) the mean square error of a regression type estimator of mean is obtained as

$$MSE\left(t_{lr(2)M}\right) = \bar{Y}^2 C_y^2 f^*_{1(2)}\left(\varphi^*_{004} - \varphi^2_{012}\right)^{-1}.$$

Similarly, the mean square error for a regression type estimator of variance can be easily written from (22) by using $(a,b)=(0,2)$.

4.2 The Ratio Type Estimator

In this sub-section we will propose a ratio type estimator of the general population parameter in two-phase sampling. The proposed estimator is

$$t_{R(2)} = \bar{y}^a\left(s_y^2\right)^{b/2}\left(\bar{x}_{(1)}/\bar{x}_{(2)}\right)^{\gamma_{(2)}}\left(s^2_{z(1)}/s^2_{z(2)}\right)^{\delta_{(2)}}. \quad (23)$$

Using the error notations, the above estimator can be written as

$$t_{R(2)} = \left[\bar{Y}^a\left(S_y^2\right)^{b/2}\left(1+\varepsilon_{y(2)}\right)^a\left(1+e_{y(2)}\right)^{b/2}\right]\left[\frac{\bar{X}\left(1+\varepsilon_{x(1)}\right)}{\bar{X}\left(1+\varepsilon_{x(2)}\right)}\right]^{\gamma_{(2)}}\left[\frac{S_z^2\left(1+e_{z(1)}\right)}{S_z^2\left(1+e_{z(2)}\right)}\right]^{\delta_{(2)}}.$$

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

$$t_{R(2)} = \left[\bar{Y}^a\left(S_y^2\right)^{b/2}\left(1+a\varepsilon_{y(2)}\right)\left(1+\frac{b}{2}e_{y(2)}\right)\right]$$
$$\left[1-\gamma_{(2)}\left(\varepsilon_{x(2)}-\varepsilon_{x(1)}\right)\right]\left[1+\delta_{(2)}\left(e_{z(2)}-e_{z(1)}\right)\right]$$

or

$$\begin{aligned} t_{R(2)} - t_{(a,b)} = t_{(a,b)}\Big[& a\varepsilon_{y(2)} + \frac{b}{2}e_{y(2)} + \frac{ab}{2}\varepsilon_{y(2)}e_{y(2)} - \gamma_2\left(\varepsilon_{x(2)}-\varepsilon_{x(1)}\right) \\ & - a\gamma_2\varepsilon_{y(2)}\left(\varepsilon_{x(2)}-\varepsilon_{x(1)}\right) - \gamma_{(2)}\frac{b}{2}e_{y(2)}\left(\varepsilon_{x(2)}-\varepsilon_{x(1)}\right) \\ & - \delta_{(2)}\left(e_{z(2)}-e_{z(1)}\right) - a\delta_{(2)}\varepsilon_{y(2)}\left(e_{z(2)}-e_{z(1)}\right) \\ & - \frac{b}{2}\delta_{(2)}e_{y(2)}\left(e_{z(2)}-e_{z(1)}\right) + \gamma_{(2)}\delta_{(2)}\left(\varepsilon_{x(2)}-\varepsilon_{x(1)}\right)\left(e_{z(2)}-e_{z(1)}\right)\Big]. \end{aligned} \quad (24)$$

Applying expectation and using (8), the bias of the proposed ratio type estimator for two-phase sampling is

$$Bias\left(t_{R(2)}\right)=t_{(a,b)}\left[\theta_2\frac{ab}{2}\varphi_{300}C_y-a\left(\theta_2-\theta_1\right)\gamma_{(2)}\rho_{yx}C_yC_x-\frac{b}{2}\gamma_{(2)}\left(\theta_2-\theta_1\right)\varphi_{210}C_x\right.$$
$$\left.-a\delta_{(2)}\left(\theta_2-\theta_1\right)\varphi_{102}C_y-\frac{b}{2}\delta_{(2)}\left(\theta_2-\theta_1\right)\varphi^*_{202}+\gamma_{(2)}\delta_{(2)}\left(\theta_2-\theta_1\right)\varphi_{012}C_x\right].$$

Again, squaring (24), we have

$$\left(t_{R(2)}-t_{(a,b)}\right)^2=t^2_{(a,b)}\left[a^2\varepsilon^2_{y(2)}+\left(b^2/4\right)e^2_{y(2)}+\gamma^2_{(2)}\left(\varepsilon_{x(2)}-\varepsilon_{x(1)}\right)^2\right.$$
$$+\delta^2_{(2)}\left(e_{z(2)}-e_{z(1)}\right)^2+ab\varepsilon_{y(2)}e_{y(2)}-2a\gamma_{(2)}\varepsilon_{y(2)}\left(\varepsilon_{x(2)}-\varepsilon_{x(1)}\right)$$
$$-2a\delta_{(2)}\varepsilon_{y(2)}\left(e_{z(2)}-e_{z(1)}\right)-b\gamma_{(2)}e_{y(2)}\left(\varepsilon_{x(2)}-\varepsilon_{x(1)}\right)$$
$$\left.-b\delta e_{y(2)}\left(e_{z(2)}-e_{z(1)}\right)+2\gamma_{(2)}\delta_{(2)}\left(\varepsilon_{x(2)}-\varepsilon_{x(1)}\right)\left(e_{z(2)}-e_{z(1)}\right)\right].$$

Applying expectation, and using (8), the mean square error of $t_{R(2)}$ is

$$MSE\left(t_{R(2)}\right)=t^2_{(a,b)}\left[\theta_2a^2C^2_y+\theta_2\frac{b^2}{4}\varphi^*_{400}+\left(\theta_2-\theta_1\right)\gamma^2_{(2)}C^2_x\right.$$
$$+\left(\theta_2-\theta_1\right)\delta^2_{(2)}\varphi^*_{004}+\theta_2ab\varphi_{300}C_y-2\left(\theta_2-\theta_1\right)a\gamma_{(2)}\rho_{yx}C_yC_x \qquad (25)$$
$$-2\left(\theta_2-\theta_1\right)a\delta_{(2)}\varphi_{102}C_y-\left(\theta_2-\theta_1\right)b\gamma_{(2)}\varphi_{210}C_x$$
$$\left.-\left(\theta_2-\theta_1\right)b\delta\varphi^*_{202}+2\left(\theta_2-\theta_1\right)\gamma_{(2)}\delta_{(2)}\varphi_{012}C_x\right].$$

The optimum values of $\gamma_{(2)}$ and $\delta_{(2)}$ which minimizes (24) are same as the optimum values of γ and δ for single-phase sampling ratio estimator. Using the optimum values of $\gamma_{(2)}$ and $\delta_{(2)}$ in (24), it is easy to show that the minimum mean square error of (23) is same as given in (22).

5. NUMERICAL STUDY AND DISCUSSION

In this section we have conducted the numerical study to see the performance of the proposed estimators in single- and two-phase sampling. The study has been conducted by using specific cases of the proposed estimators for mean and variance. The study has been conducted by using following estimators of mean and variance in single-phase sampling and their counterparts in two-phase sampling.

Table 1
The Estimators Compared

S#	Estimators of Mean	Estimators of Variance
1	$t_{1(M)} = \bar{y}\left(\bar{X}/\bar{x}\right)\left(\bar{Z}/\bar{z}\right)$	$t_{1(V)} = s_y^2\left(S_x^2/s_x^2\right)\left(S_z^2/s_z^2\right)$
2	$t_{2(M)} = \bar{y}\left(\bar{x}/\bar{X}\right)\left(\bar{z}/\bar{Z}\right)$	$t_{2(V)} = s_y^2\left(s_x^2/S_x^2\right)\left(s_z^2/S_z^2\right)$
3	$t_{3(M)} = \bar{y} + \beta_1\left(\bar{X} - \bar{x}\right) + \beta_2\left(\bar{Z} - \bar{z}\right)$	$t_{3(V)} = s_y^2 + \alpha_1\left(S_x^2 - s_x^2\right) + \alpha_2\left(S_z^2 - s_z^2\right)$
4	$t_{4(M)} = \left[\bar{y} + k_1\left(\bar{X} - \bar{x}\right)\right]\left(\bar{Z}/\bar{z}\right)$	$t_{4(V)} = s_y^2\left[\alpha \exp\left(\frac{S_x^2 - s_x^2}{S_x^2 + s_x^2}\right) + (1-\alpha)\exp\left(\frac{s_z^2 - S_z^2}{s_z^2 + S_z^2}\right)\right]$
5	$t_{R(1)M} = \bar{y}\left(\bar{X}/\bar{x}\right)^{\gamma}\left(S_z^2/s_z^2\right)^{\delta}$	$t_{R(1)V} = s_y^2\left(\bar{X}/\bar{x}\right)^{\gamma(2)}\left(S_z^2/s_z^2\right)^{\delta(2)}$
6	$t_{lr(1)M} = \bar{y}\begin{bmatrix} 1+\alpha\left(\bar{X} - \bar{x}\right) \\ +\beta\left(S_z^2 - s_z^2\right) \end{bmatrix}$	$t_{lr(1)V} = s_y^2\left[1 + \alpha_{(2)}\left(\bar{X} - \bar{x}\right) + \beta_{(2)}\left(S_z^2 - s_z^2\right)\right]$

The numerical study has been conducted on the basis of simulation. The simulation algorithm for single-phase sampling is given below.

1. Generate an artificial population of size 5000 from a trivariate normal distribution with mean vector and covariance matrices given as

$$\mu = \begin{bmatrix} 60 \\ 45 \\ 35 \end{bmatrix} \text{ and } \Sigma = \begin{bmatrix} 36 & (6\times 5)\rho_{yx} & (6\times 3)\rho_{yz} \\ & 25 & (5\times 3)\rho_{xz} \\ & & 9 \end{bmatrix}$$

   and by using different values of the correlation coefficients ρ_{yx}, ρ_{yz} and ρ_{xz}.
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 for 20000 times for each sample size.
5. Compute mean square error of each estimator of mean and variance at different sample sizes by using

$$MSE\left(t_{i(M)}\right) = \frac{1}{20000}\sum_{j=1}^{20000}\left(t_{ij} - \bar{t}_i\right)^2 \text{ ; } \bar{t}_i = \frac{1}{20000}\sum_{j=1}^{20000} t_{ij} \text{ ; } i = 1,2,3,4,R(1),lr(1)$$

$$MSE\left(t_{k(V)}\right) = \frac{1}{20000}\sum_{j=1}^{20000}\left(t_{kj}^* - \bar{t}_k^*\right)^2 \text{ ; } \bar{t}_k^* = \frac{1}{20000}\sum_{j=1}^{20000} t_{kj}^* \text{ ; } k = 1,2,3,4,R(1),lr(1)$$

The simulation algorithm for two phase sampling is given as

1. Generate an artificial population of size 5000 from a trivariate normal population as in case of simulation in single-phase sampling.

2. Generate first phase random samples of sizes 500 and 1000 from the generated population.
3. Generate second phase random samples of 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 auxiliary variable *X*, first and second-phase sample variance of the auxiliary variable *Z*, and some population measures of auxiliary variable *X*.
5. Repeat steps 2–4 for 20000 times for each combination of the first- and the second-phase sample size.
6. Compute bias and mean square error of each estimator at different sample sizes as given in Step 5 for single-phase case above.

The results of simulation study for single- and two-phase sampling are given in Tables 2 to 5 below:

Table 2
Mean Square Error of Estimators of Mean in Single-Phase Sampling

	$\rho_{yx} = 0.9$, $\rho_{yz} = 0.8$, $\rho_{xz} = 0.5$				$\rho_{yx} = 0.9$, $\rho_{yz} = –0.8$, $\rho_{xz} = –0.5$			
	$n = 50$	$n = 100$	$n = 200$	$n = 500$	$n = 50$	$n = 100$	$n = 200$	$n = 500$
$t_{1(M)}$	2.0021	0.9992	0.4509	0.1750	2.0143	1.0129	0.5010	0.1875
$t_{2(M)}$	2.0372	1.0077	0.4946	0.1838	2.0725	1.0247	0.5047	0.1894
$t_{3(M)}$	1.1155	0.5556	0.2781	0.1025	1.1508	0.5766	0.2796	0.1055
$t_{4(M)}$	1.3227	0.6522	0.3281	0.1212	2.4582	1.1900	0.5833	0.2190
$t_{R(1)M}$	1.1143	0.5385	0.2448	0.0955	1.1475	0.5695	0.2658	0.1003
$t_{lr(1)(M)}$	0.9677	0.4868	0.2406	0.0890	1.0081	0.4954	0.2413	0.0910
	$\rho_{yx} = –0.9$, $\rho_{yz} = 0.8$, $\rho_{xz} = –0.5$				$\rho_{yx} = –0.9$, $\rho_{yz} = –0.8$, $\rho_{xz} = 0.5$			
	$n = 50$	$n = 100$	$n = 200$	$n = 500$	$n = 50$	$n = 100$	$n = 200$	$n = 500$
$t_{1(M)}$	2.0323	1.0401	0.5068	0.1909	2.0955	1.1454	0.6196	0.2255
$t_{2(M)}$	2.1069	1.0513	0.5101	0.1914	2.0815	1.0613	0.5244	0.1940
$t_{3(M)}$	1.1780	0.5891	0.2847	0.1069	1.1633	0.5895	0.2928	0.1081
$t_{4(M)}$	1.3812	0.6879	0.3313	0.1256	2.4602	1.2235	0.6150	0.2276
$t_{R(1)M}$	1.1184	0.5387	0.2674	0.1008	1.0979	0.5675	0.2747	0.1016
$t_{lr(1)M}$	1.0258	0.5124	0.2494	0.0929	1.0201	0.5108	0.2561	0.0948

Table 3
Mean Square Error of Estimators of Variance in Single-Phase Sampling

	ρ_{yx} = 0.9, ρ_{yz} = 0.8, ρ_{xz} = 0.5				ρ_{yx} = 0.9, ρ_{yz} = –0.8, ρ_{xz} = –0.5			
	n = 50	n = 100	n = 200	n = 500	n = 50	n = 100	n = 200	n = 500
$t_{1(V)}$	139.7877	60.2638	32.2726	11.1710	191.6149	82.0928	36.7101	13.3691
$t_{2(V)}$	146.9988	69.7540	34.1012	12.9222	153.1466	71.5357	34.8559	13.1395
$t_{3(V)}$	72.8541	35.9331	17.5337	6.4471	71.9468	34.9008	17.1414	6.4418
$t_{4(V)}$	74.0471	35.0792	16.8858	6.1892	74.9150	35.3253	16.9390	6.3496
$t_{R(1)\mathrm{V}}$	68.5871	32.5289	15.5628	6.0852	64.8547	33.5287	16.0874	6.0896
$t_{lr(1)\mathrm{V}}$	61.0709	29.9825	14.3037	5.4668	61.0065	29.9660	14.3085	5.3337
	ρ_{yx} = –0.9, ρ_{yz} = 0.8, ρ_{xz} = –0.5				ρ_{yx} = –0.9, ρ_{yz} = –0.8, ρ_{xz} = 0.5			
	n = 50	n = 100	n = 200	n = 500	n = 50	n = 100	n = 200	n = 500
$t_{1(V)}$	196.8658	87.2749	39.5066	14.1016	203.7121	87.9975	39.7622	14.2298
$t_{2(V)}$	160.9924	76.7208	37.8723	13.4191	164.6783	78.3326	37.7447	14.199
$t_{3(V)}$	84.7500	42.1795	20.6543	7.7841	82.2643	40.2293	19.3902	7.3315
$t_{4(V)}$	77.8734	38.1840	18.2415	6.7952	80.9137	38.7421	18.5176	6.8718
$t_{R(1)\mathrm{V}}$	69.8743	33.1985	16.8971	6.0358	73.6187	34.0928	17.0491	6.0487
$t_{lr(1)\mathrm{V}}$	64.9631	32.2956	15.5493	5.7883	68.1011	32.8356	16.1142	5.9994

Table 4
Mean Square Error of Estimators of Mean in Two-Phase Sampling

Corr.	n_1	n_2	$t_{1(M)}^{2P}$	$t_{2(M)}^{2P}$	$t_{3(M)}^{2P}$	$t_{4(M)}^{2P}$	$t_{R(2)M}^{2P}$	$t_{lr(2)M}^{2P}$
$\rho_{yx} = 0.9$ $\rho_{yz} = 0.8$ $\rho_{xz} = 0.5$	500	25	4.0163	4.0597	2.2602	2.6639	2.2363	2.1992
		50	1.9120	1.9516	1.0791	1.2659	1.0750	1.0518
		100	0.8796	0.8876	0.5274	0.6106	0.5264	0.5078
	1000	50	1.9514	1.9787	1.1074	1.2956	1.1073	1.0929
		100	0.9451	0.9618	0.5578	0.6560	0.5513	0.5361
		200	0.4377	0.4432	0.2611	0.3038	0.2600	0.2510
$\rho_{yx} = 0.9$ $\rho_{yz} = -0.8$ $\rho_{xz} = -0.5$	500	25	4.0878	3.9902	2.2957	4.7435	2.2819	2.2517
		50	1.9508	1.9320	1.0982	2.2608	1.0860	1.0732
		100	0.9050	0.8902	0.5243	1.0427	0.5179	0.5094
	1000	50	1.9969	1.9987	1.1463	2.3203	1.1245	1.1018
		100	0.9621	0.9687	0.5432	1.1159	0.5392	0.5277
		200	0.4467	0.4473	0.2547	0.5150	0.2520	0.2474
$\rho_{yx} = -0.9$ $\rho_{yz} = 0.8$ $\rho_{xz} = -0.5$	500	25	4.0334	4.1274	2.3255	2.7085	2.3091	2.2784
		50	1.9761	1.9841	1.1230	1.3006	1.1181	1.1026
		100	0.9032	0.9234	0.5391	0.6137	0.5360	0.5291
	1000	50	2.0284	2.0777	1.1760	1.3634	1.1713	1.1555
		100	0.9887	0.9968	0.5698	0.6591	0.5670	0.5585
		200	0.4374	0.4502	0.2662	0.306	0.2613	0.2585
$\rho_{yx} = -0.9$ $\rho_{yz} = -0.8$ $\rho_{xz} = 0.5$	500	25	4.2473	4.1839	2.3745	4.9268	2.3633	2.342
		50	1.9955	1.9853	1.1314	2.3212	1.1255	1.1064
		100	0.9583	0.9342	0.5498	1.0536	0.5480	0.5282
	1000	50	2.0794	2.0576	1.1634	2.4523	1.1564	1.1448
		100	1.1106	1.0100	0.5750	1.1559	0.5691	0.5596
		200	0.4921	0.4611	0.2701	0.5295	0.2671	0.2623

Table 5
Mean Square Error of Estimators of Mean in Two-Phase Sampling

Corr.	n_1	n_2	$t_{1(V)}^{2P}$	$t_{2(V)}^{2P}$	$t_{3(V)}^{2P}$	$t_{4(V)}^{2P}$	$t_{R(2)V}^{2P}$	$t_{lr(2)V}^{2P}$
$\rho_{yx} = 0.9$, $\rho_{yz} = 0.8$, $\rho_{xz} = 0.5$	500	25	292.101	314.876	147.613	155.499	142.517	137.400
		50	132.509	140.142	70.500	71.382	68.267	66.025
		100	59.345	62.031	33.785	33.242	32.360	30.907
	1000	50	134.220	143.823	72.632	73.712	69.870	67.079
		100	67.847	68.811	35.539	35.038	34.166	32.776
		200	28.559	30.863	16.372	15.904	15.688	14.973
$\rho_{yx} = 0.9$, $\rho_{yz} = -0.8$, $\rho_{xz} = -0.5$	500	25	506.564	321.012	145.894	159.517	143.224	140.521
		50	176.977	142.606	69.262	71.647	67.788	66.287
		100	70.670	63.644	32.789	33.086	31.957	31.124
	1000	50	187.163	147.005	70.751	73.408	68.628	66.494
		100	75.484	69.057	34.197	34.334	33.004	31.798
		200	32.944	32.167	15.887	15.801	15.522	15.129
$\rho_{yx} = -0.9$, $\rho_{yz} = 0.8$, $\rho_{xz} = -0.5$	500	25	519.519	322.120	175.105	169.943	161.284	147.429
		50	185.176	150.024	82.894	77.268	77.414	71.909
		100	72.842	66.365	38.534	35.080	35.998	33.439
	1000	50	194.705	155.516	85.120	79.085	78.112	71.076
		100	80.160	72.988	40.908	36.755	37.714	34.510
		200	33.593	33.500	19.087	17.012	17.562	16.028
$\rho_{yx} = -0.9$, $\rho_{yz} = -0.8$, $\rho_{xz} = 0.5$	500	25	520.891	334.744	165.880	173.380	160.942	156.001
		50	188.809	154.230	78.807	78.671	75.267	71.722
		100	76.799	69.543	37.818	36.670	36.472	35.116
	1000	50	199.524	158.158	80.927	80.549	77.869	74.775
		100	80.728	74.076	38.776	37.391	37.270	35.751
		200	35.425	33.740	18.410	17.557	17.578	16.746

It is clear, from the above tables, that our proposed estimators of mean and variance performs better than the other competing estimators as it has the smallest mean square error at all the combinations and at all of the sample sizes. We have seen that the analytical mean square errors of the proposed ratio and regression type estimators are same but in simulation the performance of the proposed regression type estimator is better than the proposed ratio type estimator. The efficiencies of various estimators are also compared in comparison with the ratio type estimators of mean and variance. The graphs of the efficiency comparison are shown in Figures 1 and 2 below.

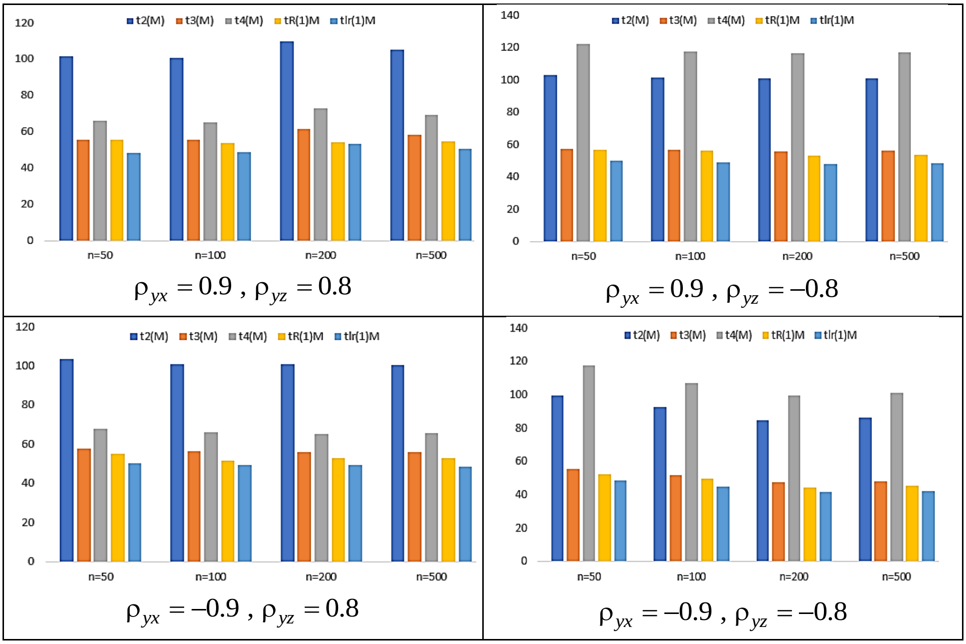

Figure 1: Relative Efficiencies of Various Estimators of the Mean

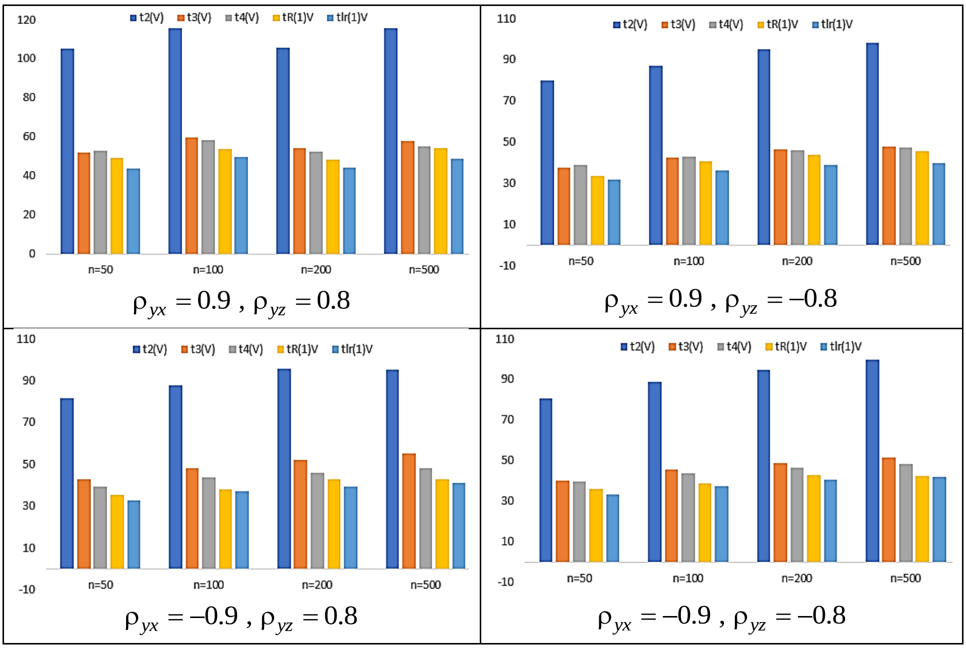

Figure 2: Relative Efficiencies of Various Estimators of the Variance

The figures above show the same results as found in tables above. We can also see from above figures that the ratio estimator of mean is a better estimator as compared with the product type estimator of mean when correlation of the response variable with one of the auxiliary variables is positive. We can also see that the estimator $t_{4(M)}$ is better than the ratio type estimator of mean when ρ_{yz} is positive. We can also see that all of the estimators of variance are better than the ratio type estimator of variance for all of the combinations of the correlation coefficients.

6. CONCLUSIONS AND RECOMMENDATIONS

In this paper we have proposed some ratio and regression type estimators of the general population parameter using two auxiliary variables. The estimators have been proposed for single- and two-phase sampling. We have obtained the expressions for the mean square errors of the proposed estimators. The proposed estimators have been compared with some of the existing estimators of mean and variance via a simulation study. It is found that our proposed estimators are better than the competing estimators in the study. As further research, the proposed estimators can be extended to the sets of auxiliary variables. The proposed estimators can also be study under some different sampling designs for example ranked set sampling. The estimators can also be study in presence of non – response.

REFERENCES

1. Abu-Dayyeh, W.A., Ahmed, M.S., Ahmed, R.A. and Mutlak, H.A. (2003). Some estimators of a finite population mean using auxiliary information. *App. Math. Comp.*, 139, 287-298.
2. Abu–Dayyeh, W.A. Ahmed, M.S. (2005). Ratio and regression estimator for the variance under two-phase sampling. *Int. J. of Stat. Sci.*, 4, 49-56.
3. Adichwal, N.K., Ahmadini, A.A.H., Raghav, Y.S., Singh, R. and Ali, I. (2022). Estimation of general parameters using auxiliary information in simple random sampling without replacement. *J. King Saud Uni. Sci.*, 34, Article 101754.
4. Adichwal, N.K., Sharma, P. and Singh, R. (2015). Generalized class of estimators for population variance using information on two auxiliary variables. *Int. J. App. Comp. Math.*, 3, 651-661.
5. Ahmad, Z. Shahbaz, M.Q. and Hanif, M. (2013). *Two Phase Sampling*, Cambridge Scholars.
6. Al-Marshadi, A.H., Al-Harby, A.H. and Shahbaz, M.Q. (2018). On some new estimators of population variance in single- and two-phase sampling. *Maejo. Int. J. Sci. & Tech.*, 12, 272-281.
7. Cochran, W.G. (1940). The estimation of the yields of cereal experiments by sampling for the ratio gain to total produce. *J. Agric. Soc*., 30, 262-275.
8. Grover, L.K. and Kaur, A. (2021). An improved regression type estimator of population mean with two auxiliary variables and its variant using robust regression method. *J. Comp. App. Math.*, 382, Article: 113072.
9. Hanif, M. Shahbaz, M.Q. and Ahmed, M. (2018). *Sampling Techniques: Methods and Applications*, Nova, USA.

10. Isaki, C. (1983). Variance estimation using auxiliary information. *J. Amer. Statist. Assoc.*, 78, 117-123.
11. Kadilar, C. and Cingi, H. (2005). A new estimator using two auxiliary variables. *App. Math. Comp.*, 162, 901-908.
12. Kiregyera, B. (1984). A Regression-type estimator using two auxiliary variables and model of double sampling from finite populations. *Metrika*, 31, 215-226.
13. Kumar, S. and Choudhary, M. (2023). An improved estimator of finite population variance using two auxiliary variables in SRS. *Rev. Colom. Estad.*, 46, 81-91.
14. Samiuddin, M. and Hanif, M. (2007). Estimation of population mean in single- and two-phase sampling with or without additional information. *Pak. J. Statist.*, 23, 1-9.
15. Shahbaz, S.H., Fayomi, A. and Shahbaz, M.Q. (2023). Estimation of the general population parameter in single- and two-phase sampling. *AIMS Math.*, 8, 14951-14977.
16. Singh, H.P. and Espejo, M.R. (2003). On linear regression and ratio-product estimation of a finite population mean. *J. of the Royal Stat. Soc., D*, 52, 59-67.
17. Singh, R., Mishra, M., Singh, B.P., Singh, P. and Adichwal, N.K. (2018). Improved estimators for population coefficient of variation using auxiliary variable. *J. Stat. Manage. Syst.*, 21, 1335-1355.