# ENHANCING ACCURACY IN ESTIMATING POPULATION MEAN THROUGH MODIFIED RATIO ESTIMATORS IN RANKED SET SAMPLING

S. A. Sabo[a], A. A. Osi*,[a], I. Z. Musa[a], H. U. Abubakar[a], A. Muhammad[a], U. Abubakar[b]

•

[a]Department of Statistics, Aliko Dangote University of Science and Technology, Wudil, Nigeria.
[b]Department of Statistics, Jigawa State Polytechnic Dutse, Nigeria.
Email: abuammarosi@gmail.com, aaosi@kust.edu.ng, usmanabubakar@jigpoly.edu.ng

**Abstract**

*This study aimed to improve the ratio estimation under rank set sampling by proposing some new modified estimators using auxiliary information on the size of the sample. The expressions for the bias and mean squared error of the proposed modified estimators are derived up to the first order of approximation using Tailor series expansion. A theoretical efficiency comparison between the proposed and competing estimators was done and the conditions upon which the proposed estimators expect to outperform the competing estimators were stated. Results from the numerical work indicate that the new proposed improved estimators are better than the already existing ones, showing a lower mean square error, coefficient of variation, and a considerable gain in efficiency. The new proposed estimators emerge as the optimal choice, outperforming the competing ones, and should therefore be used in applications.*

**Keywords:** Ranked set sampling, MSE, Sample size, Auxiliary variable, Ratio estimator

## 1. Introduction

Estimation of population parameters, particularly mean, is paramount and has received significant attention from researchers in the field of sampling theory over the past few decades [1]. Survey sampling studies often include additional auxiliary variables that, when correlated with the primary variable of interest, improve the efficiency of statistical inference by creating more effective designs and estimators [2]. Improved designs and efficient estimators are being developed to improve the precision of estimates. Techniques such as the ratio by [3], product by [4], and regression by [5] are widely used in the literature on sampling theory. The ratio method is recommended for positive correlations between the study and auxiliary variables. The random sample sampling according to [6] is a cost-effective method to classify experimental units, particularly in environmental monitoring and assessment. It uses a cheaply measurable covariate to rank units, resulting in a more representative sample of observations. This method offers greater confidence for a fixed number of observations or requires a smaller number for the desired confidence levels. This study proposes a ratio estimator for the population mean, using information on the sample size of the auxiliary variable through the ranked set under a simple random sampling scheme. In Simple Random Sampling, ratio estimators are widely used when the variable of interest is linearly related to an auxiliary variable. The classical ratio estimator of [3] is expressed as:

$$\overline{y}_R = \overline{y}\left[\frac{\overline{X}}{\overline{x}}\right] \tag{1}$$

Where $\overline{y}$ and $\overline{x}$ are the sample means of the study and the auxiliary variables, respectively, and the population mean of the auxiliary variable.

The mean squared error of (1) to the first degree of approximation is given as;

$$MSE(\overline{y_R}) = \theta\overline{Y}^2\left[C_y^2 + C_x^2 - 2\rho C_x C_y\right] \tag{2}$$

Where $\theta = \frac{1-f}{n}$, $f = \frac{n}{N}$, $C_y = \frac{S_y}{\overline{Y}}$, $C_x = \frac{S_x}{\overline{Y}}$, $\rho = \frac{S_{xy}}{S_x S_y}$, $S_y^2 = \frac{\sum_{i=1}^{N}(y_i-\overline{Y})^2}{N-1}$, $S_x^2 = \frac{\sum_{i=1}^{N}(x_i-\overline{X})^2}{N-1}$.

[7] highlighted that ratio estimators become highly efficient when there is a strong positive correlation between and variables. Modified ratio estimator for is given by [8] as;

$$\overline{y}_{SD} = \overline{y}\left[\frac{\overline{X}+C_x}{\overline{x}+C_x}\right] \tag{3}$$

Where is the coefficient of variation. The MSE of (3) can be expressed as;

$$MSE(\overline{y}_{SD}) = \theta\overline{Y}^2\left[C_y^2 + \psi_l C_x^2 - 2\psi_l\rho C_x C_y\right] \tag{4}$$

where $\psi_1 = \frac{\overline{X}}{\overline{X}+C_x}$

[9] proposed a modified ratio estimator for with known coefficient of kurtosis given by;

$$\overline{y}_{SK} = \overline{y}\left[\frac{\overline{X}+\beta_2(x)}{\overline{x}+\beta_2(x)}\right] \tag{5}$$

The mean squared error of (5) with first degree of approximation is given as;

$$MSE(\overline{y}_{SK}) = \theta\overline{Y}^2\left[C_y^2 + \psi_2 C_x^2 - 2\psi_2\rho C_x C_y\right] \tag{6}$$

Where $\psi_2 = \frac{\overline{X}}{\overline{X}+\beta_2(x)}$

Motivated by [7] and [8] a ratio estimator for suggested by [10] is expressed as;

$$\overline{y}_{UP} = \overline{y}\left[\frac{C_x\overline{X}+\beta_2(x)}{\overline{x}C_x+\beta_2(x)}\right] \tag{7}$$

The MSE of (7) to the first order of approximation can be expressed as;

$$MSE(\overline{y}_{UP}) = \theta\overline{Y}^2\left[C_y^2 + \psi_3 C_x^2 - 2\psi_3\rho C_x C_y\right] \tag{8}$$

Where, $\psi_3 = \frac{\overline{X}C_x}{\overline{X}C_x+\beta_2(x)}$

[11] utilized the information on the sample size of the auxiliary variable to improve the ratio estimation. The proposed estimator is given by;

$$\overline{y}_{JK} = \overline{y}\left[\frac{\overline{X}+n}{\overline{x}+n}\right] \tag{9}$$

To the first order of approximation, the MSE of (9) can be expressed as;

$$MSE(\overline{y}_{JK}) = \theta\overline{Y}^2\left[C_y^2 + \psi_4 C_x^2 - 2\psi_4\rho C_x C_y\right] \tag{10}$$

Where $\psi_4 = \frac{\overline{X}}{\overline{X}+n}$

The integration of ratio estimators with RSS has led to several theoretical advancements and efficiency improvements. Under RSS, ratio estimators can take advantage of both ranking information and the auxiliary variable, leading to more precise estimates. The classical ratio estimator under RSS suggested by [12] can be written as:

$$\bar{y}_{RSS} = \frac{\bar{y}_{[n]}}{\bar{x}_{[n]}} \tag{11}$$

Where $\bar{y}_{[n]} = \frac{1}{n}\sum_{i=1}^{n} y_{[i]} \cdot \bar{x}_{[n]} = \frac{1}{n}\sum_{i=1}^{n} x_{[i]}$ are the ranked set sample means for variables and respectively. The estimator in (11) is unbiased and its variance is given by;

$$V\left(\bar{y}_{RSS}\right) = \frac{\sigma^2}{rm} - \frac{1}{rm^2}\sum_{i=1}^{n}\tau_{y(i)}^2 \tag{12}$$

Where $\tau_{y(i)} = \left(\mu_{y(i)} - \bar{Y}\right)$.

Assuming that the population mean of the auxiliary variable is known, [13] suggested a ratio estimator using ranked set sampling as;

$$\bar{y}_{R,RSS} = \bar{y}_{[n]}\left(\frac{\bar{X}}{x_{[n]}}\right) \tag{13}$$

The MSE of (13) to the first order of approximation can be expressed as;

$$MSE(\bar{y}_{R,RSS}) = \frac{1}{mr}\left(S_y^2 + R^2S_x^2 - 2RS_{yx}\right) - \frac{1}{m^r}\left(\sum_{i=1}^{m}\tau_{y(i)}^2 + R^2\sum_{i=1}^{m}x_{(i)}^2 - 2R\sum_{i=1}^{m}\tau_{yx}\right) \tag{14}$$

where $\tau_{x(i)} = \mu_{x(i)} - \bar{X}$, $\tau_{y(i)} = \mu_{y(i)} - \bar{Y}$, $\tau_{yx(i)} = (\mu_{x(i)} - \bar{X})(\mu_{y(i)} - \bar{Y})$.

[**?**] proposed a ratio estimators based on linear combination of coefficient of skewness and quartile deviation by utilizing some known parameters of the concomitant variable given by;

$$\bar{y}_{JL,RSS} = \frac{\bar{y}_{[n]} + \beta_1\left(\mu_x - \bar{x}_{[n]}\right)}{QD\bar{x}_{[n]} + \beta_1}QD\mu_x + \beta_1 \tag{15}$$

The expression for mean squared error of (15) is given by;

$$MSE(\bar{y}_{JL,RSS}) = \bar{Y}^2\left[\theta\left(C_y^2 + \lambda_2^2C_x^2 - 2\lambda_2\rho_{yx}C_xC_y\right) - \left(W_{y(i)} - \lambda_2W_{x(i)}\right)^2\right] \tag{16}$$

where $\lambda_1 = \frac{\bar{Y}\beta_1}{\bar{X}\beta_1 + QD}$.

[15] suggests a modified ratio estimators of finite population mean utilizing information on coefficient of variation of an auxiliary variable in ranked set sampling given as;

$$\bar{y}_{ML,RSS} = \bar{y}_{[n]}\left[\frac{\bar{X} + C_x}{\bar{x}_{[n]} + C_x}\right] \tag{17}$$

$$\bar{y}_{M2,RSS} = \bar{y}_{[n]}\left[\frac{\bar{X}C_x + \beta_1(x)}{\bar{x}_{[n]}C_x + \beta_1(x)}\right] \tag{18}$$

Their mean squared error of (17) and (18) are given respectively as;

$$MSE(\bar{y}_{M1,RSS}) = \bar{Y}^2\left[\theta\left(C_y^2 + \lambda_1^2C_x^2 - 2\lambda_1\rho_{yx}C_xC_y\right) - \left(W_{y(i)} - \lambda_1W_{x(i)}\right)^2\right] \tag{19}$$

$$MSE(\bar{y}_{M2,RSS}) = \bar{Y}^2\left[\theta\left(C_y^2 + \lambda_2^2C_x^2 - 2\lambda_2\rho_{yx}C_xC_y\right) - \left(W_{y(i)} - \lambda_2W_{x(i)}\right)^2\right] \tag{20}$$

Where $\lambda_2 = \frac{\bar{X}}{\bar{X} + C_x}$ and $\lambda_3 = \frac{\bar{X}C_x}{\bar{X}C_x + \beta_2(x)}$.

Motivated by [14] and [8] a ratio estimator under ranked set sampling using information of population coefficient of variation and median of auxiliary was suggested by [16] which is expressed as;

$$\overline{y}_{RZ,RSS} = \overline{y}_{[n]} \left[ \frac{\overline{X} + C_x M_d}{x_{[n]} + C_x M_d} \right] \tag{21}$$

The mean squared error of (21) can be expressed as;

$$MSE(\overline{y}_{RZ,RSS}) = \overline{Y}^2 \left[ \theta \left( C_y^2 + \lambda_4^2 C_x^2 - 2\lambda_4 \rho_{yx} C_x C_y \right) - \left( W_{y_{(i)}} - \lambda_4 W_{x_{(i)}} \right)^2 \right] \tag{22}$$

where $\lambda_4 = \frac{\overline{X}}{\overline{X} + C_x M_d}$.

## 2. Methods

Motivated by [11], [16] and [17], after studying the works of [18], [19] and [20] we proposed some ratio estimators for population mean based on ranked set sampling and presented as;

$$t_{p1} = \overline{y}_{[n]} \left[ \frac{\overline{X} + n}{\overline{x}_{[n]} + n} \right] \tag{23}$$

$$t_{p2} = \overline{y}_{[n]} \left[ \alpha + (1 - \alpha) \left( \frac{\overline{X} + n}{\overline{x}_{[n]} + n} \right) \right] \tag{24}$$

Where is the scalar to be obtained such that, the mean squared error of the proposed estimator in (24) is minimum.

To study the large sample properties (Bias and Mean Squared Error) of the proposed modified ratio estimators, we have used the following approximations as:

$$\overline{y}_{[n]} = \overline{Y}(1 + e_0), \quad \overline{x}_{[n]} = \overline{X}(1 + e_1) \tag{25}$$

Such that; $E(e_0) = E(e_1) = 0$, $E(e_0^2) = \theta C_y^2 - W_{y_{(0)}}^2$, $E(e_1^2) = \theta C_x^2 - W_{x_{(0)}}^2$, $E(e_0 e_1) = \theta \rho C_x C_y - W_{y_{(0)}}$.

Where;

$$W_{yx(i)} = \frac{1}{m^2 r \overline{Y}\overline{X}} \sum \tau_{xy(i)}^2, \quad W_{x(i)}^2 = \frac{1}{m^2 r \overline{X}^2} \sum \tau_{x(i)}^2, \quad W_{y(i)}^2 = \frac{1}{m^2 r \overline{Y}^2} \sum \tau_{y(i)}^2$$

$$\tau_{x(i)} = \mu_{x(i)} - \overline{X}, \quad \tau_{y(i)} = \mu_{y(i)} - \overline{Y}, \quad \tau_{yx(i)} = \left( \mu_{y(i)} - \overline{Y} \right) \left( \mu_{x(i)} - \overline{X} \right)$$

Expressing $t_{p1}$ in terms of the definition in 25 gives;

$$t_{p1} = \overline{Y}(1 + e_0) \left[ \frac{\overline{X} + n}{\overline{X}(1 + e_1) + n} \right] \tag{26}$$

Simplifying 26 gives;

$$t_{p1} = \overline{Y}(1 + e_0)(1 + \gamma e_1)^{-1} \tag{27}$$

Where;

$$\gamma = \frac{\overline{X}}{\overline{X} + n}$$

Assume that $|e_1| < 1$ so that $(1 + \gamma_1 e_1)^{-1}$ may be expanded, retaining the terms with degree not more than two only gives;

$$t_{p1} = \overline{Y}(1 + e_0 - \gamma e_1 - \gamma e_0 e_1 + \gamma^2 e_1^2) \tag{28}$$

Subtracting $\overline{Y}$ both sides from equation 28 and taking expectations gives;

$$E\left(t_{p1} - \overline{Y}\right) = \overline{Y} E\left(e_0 - \gamma e_1 - \gamma e_0 e_1 + \gamma^2 e_1^2\right) \tag{29}$$

Applying the definitions of the expectations above gives;

$$B\left(t_{p1}\right) = \overline{Y}\left[\theta\left(\gamma^2 C_x^2 - \gamma\rho C_x C_y\right) - \left(\gamma^2 W_{x(i)}^2 - \gamma W_{yx(i)}\right)\right] \tag{30}$$

Squaring both sides of 30, expanding its RHS, neglecting powers greater than 2 and applying the definition of the expectations, we obtain the mean squared error of $t_{p1}$ as;

$$MSE\left(t_{p1} - \overline{Y}\right)^2 = \overline{Y}^2\left[\theta\left(C_y^2 + \gamma^2 C_x^2 - 2\gamma\rho C_x C_y\right) - \left(W_{y(i)} - \gamma W_{x(i)}\right)^2\right] \tag{31}$$

Similarly, the biased and mean square error of the second proposed estimator $t_{p2}$ can be respectively expressed as;

$$B\left(t_{p2}\right) = \overline{Y}\left[\theta\left(\gamma^2 C_x^2 - \gamma\rho C_x C_y - \alpha\gamma^2\rho C_x C_y\right) - \left(\gamma^2 W_{x(i)}^2 + \gamma W_{yx(i)} - \alpha\gamma^2 W_{yx(i)}\right)\right] \tag{32}$$

$$MSE\left(t_{p2}\right) = \overline{Y}^2\left[\begin{array}{l}\theta\left(C_y^2 - \gamma^2 C_x^2 - 2\gamma\rho C_x C_y - \alpha\gamma\rho C_x C_y - \alpha^2\gamma^2 C_x^2\right) \\ -W_{y(i)}^2 + 2\gamma W_{yx(i)} + \alpha\gamma W_{yx} + \alpha^2\gamma^2 W_{x(i)}^2\end{array}\right] \tag{33}$$

The optimum value of $\alpha$ that minimizes the MSE of the second proposed estimator ( $t_{p2}$ ) can easily be obtained by partially differentiating equation 33 with respect $\alpha$ and equating the result to zero. Thus;

$$\frac{\partial MSE\left(t_{p2}\right)}{\partial\alpha} = 0$$

Simplifying further gives;

$$\alpha_{opt} = \frac{\theta\gamma\rho C_x C_y - \gamma W_{y(x)}}{2\left(\gamma^2 W_{x(i)} - \theta\gamma^2 C_x^2\right)} \tag{34}$$

The minimum mean squared error of the proposed estimator $t_{p2}$ can be obtained by substituting $\alpha_{opt}$ in equation 33 as follows;

$$MSE\left(t_{p2}\right) = \overline{Y}^2\left[\begin{array}{l}\theta\left(C_y^2 - \gamma^2 C_x^2 - 2\gamma\rho C_x C_y - \alpha_{opt}\gamma\rho C_x C_y - \alpha_{opt}^2\gamma^2 C_x^2\right) \\ -W_{y(i)}^2 + 2\gamma W_{yx(i)} + \alpha_{opt}\gamma W_{yx} + \alpha_{opt}^2\gamma^2 W_{x(i)}^2\end{array}\right] \tag{35}$$

## 3. Results

### 3.1. Theoretical Efficiency Comparison

The efficiency of our proposed estimators was compared with that of other existing ratio estimators in the literature theoretically. The conditions to check the performance of the proposed estimators over the other competing estimators were strictly stated here.

The proposed estimator is more efficient than the usual sample mean if;

$$MSE\left(t_{pr2}\right) - V\left(y_{RSS}\right) < 0 \tag{36}$$

So that;

$$\overline{Y}^2\left[\begin{array}{l}\theta\left(C_y^2 - \gamma^2 C_x^2 - 2\gamma\rho C_x C_y - \alpha_{opt}\gamma\rho C_x C_y - \alpha_{opt}^2\gamma^2 C_x^2\right) \\ -W_{y(i)}^2 + 2\gamma W_{yx(i)} + \alpha_{opt}\gamma W_{yx} + \alpha_{opt}^2\gamma^2 W_{x(i)}^2\end{array}\right] - \left(\frac{\sigma^2}{rm} - \frac{1}{rm^2}\sum_{i=1}^{n}\tau_{y(i)}^2\right) < 0 \tag{37}$$

The proposed estimator is more efficient than the [?] ratio estimator if;

$$MSE\left(t_{pr2}\right) - MSE\left(\overline{y}_{RSS}\right) < 0 \tag{38}$$

so that;

$$\overline{Y}^2 \begin{bmatrix} \theta\left(C_y^2 - \gamma^2 C_x^2 - 2\gamma\rho C_x C_y - \alpha_{opt}\gamma\rho C_x C_y - \alpha_{opt}^2\gamma^2 C_x^2\right) \\ -W_{y(i)}^2 + 2\gamma W_{yx(i)} + \alpha_{opt}\gamma W_{yx} + \alpha_{opt}^2\gamma^2 W_{x(i)}^2 \end{bmatrix} - \theta\overline{Y}^2\left[C_y^2 + C_x^2 - 2\rho C_x C_y\right] < 0 \tag{39}$$

The proposed estimator is more efficient than the [8] ratio estimator if;

$$MSE\left(t_{pr2}\right) - MSE\left(\overline{y}_{SD}\right) < 0 \tag{40}$$

So that;

$$\overline{Y}^2 \begin{bmatrix} \theta\left(C_y^2 - \gamma^2 C_x^2 - 2\gamma\rho C_x C_y - \alpha_{opt}\gamma\rho C_x C_y - \alpha_{opt}^2\gamma^2 C_x^2\right) \\ -W_{y(i)}^2 + 2\gamma W_{yx(i)} + \alpha_{opt}\gamma W_{yx} + \alpha_{opt}^2\gamma^2 W_{x(i)}^2 \end{bmatrix} - \theta\overline{Y}^2\left[C_y^2 + \psi_1 C_x^2 - 2\psi_1\rho C_x C_y\right] < 0 \tag{41}$$

The proposed estimator is more efficient than the [9] ratio estimator if;

$$MSE\left(t_{pr2}\right) - MSE\left(\overline{y}_{SK}\right) < 0 \tag{42}$$

So that;

$$\overline{Y}^2 \begin{bmatrix} \theta\left(C_y^2 - \gamma^2 C_x^2 - 2\gamma\rho C_x C_y - \alpha_{opt}\gamma\rho C_x C_y - \alpha_{opt}^2\gamma^2 C_x^2\right) \\ -W_{y(i)}^2 + 2\gamma W_{yx(i)} + \alpha_{opt}\gamma W_{yx} + \alpha_{opt}^2\gamma^2 W_{x(i)}^2 \end{bmatrix} - \theta\overline{Y}^2\left[C_y^2 + \psi_2 C_x^2 - 2\psi_2\rho C_x C_y\right] < 0 \tag{43}$$

The proposed estimator is more efficient than the [10] ratio estimator if;

$$MSE\left(t_{pr2}\right) - MSE\left(\overline{y}_{UP}\right) < 0 \tag{44}$$

So that;

$$\overline{Y}^2 \begin{bmatrix} \theta\left(C_y^2 - \gamma^2 C_x^2 - 2\gamma\rho C_x C_y - \alpha_{opt}\gamma\rho C_x C_y - \alpha_{opt}^2\gamma^2 C_x^2\right) \\ -W_{y(i)}^2 + 2\gamma W_{yx(i)} + \alpha_{opt}\gamma W_{yx} + \alpha_{opt}^2\gamma^2 W_{x(i)}^2 \end{bmatrix} - \theta\overline{Y}^2\left[C_y^2 + \psi_3 C_x^2 - 2\psi_3\rho C_x C_y\right] < 0 \tag{45}$$

The proposed estimator is more efficient than the [11] estimator if;

$$MSE\left(t_{pr2}\right) - MSE\left(\overline{y}_{JK}\right) < 0 \tag{46}$$

So that;

$$\overline{Y}^2 \begin{bmatrix} \theta\left(C_y^2 - \gamma^2 C_x^2 - 2\gamma\rho C_x C_y - \alpha_{opt}\gamma\rho C_x C_y - \alpha_{opt}^2\gamma^2 C_x^2\right) \\ -W_{y(i)}^2 + 2\gamma W_{yx(i)} + \alpha_{opt}\gamma W_{yx} + \alpha_{opt}^2\gamma^2 W_{x(i)}^2 \end{bmatrix} - \theta\overline{Y}^2\left[C_y^2 + \psi_4 C_x^2 - 2\psi_4\rho C_x C_y\right] < 0 \tag{47}$$

The proposed estimator is more efficient than the [12] ratio estimator if;

$$MSE\left(t_{pr2}\right) - MSE\left(\overline{y}_{R,RSS}\right) < 0 \tag{48}$$

So that;

$$\begin{aligned} &\overline{Y}^2 \begin{bmatrix} \theta\left(C_y^2 - \gamma^2 C_x^2 - 2\gamma\rho C_x C_y - \alpha_{opt}\gamma\rho C_x C_y - \alpha_{opt}^2\gamma^2 C_x^2\right) \\ -W_{y(i)}^2 + 2\gamma W_{yx(i)} + \alpha_{opt}\gamma W_{yx} + \alpha_{opt}^2\gamma^2 W_{x(i)}^2 \end{bmatrix} - \frac{1}{mr}\left(S_y^2 + R^2 S_x^2 - 2RS_{yx}\right) \\ &- \frac{1}{m^2 r}\left(\sum_{i=1}^{m}\tau_{y(i)}^2 + R^2\sum_{i=1}^{m}\tau_{x(i)}^2 - 2R\sum_{i=1}^{m}\tau_{yx}\right) < 0 \end{aligned} \tag{49}$$

The proposed estimator is more efficient than the [?] ratio estimator if;

$$MSE\left(t_{pr2}\right) - MSE\left(\overline{y}_{R,RSS}\right) < 0 \tag{50}$$

So that;

$$\begin{array}{l} \overline{Y}^2 \left[ \begin{array}{l} \theta\left(C_y^2 - \gamma^2 C_x^2 - 2\gamma\rho C_x C_y - \alpha_{opt}\gamma\rho C_x C_y - \alpha_{opt}^2 \gamma^2 C_x^2\right) \\ -W_{y(i)}^2 + 2\gamma W_{yx(i)} + \alpha_{opt}\gamma W_{yx} + \alpha_{opt}^2 \gamma^2 W_{x(i)}^2 \end{array} \right] - \frac{1}{mr}\left(S_y^2 + R^2 S_x^2 - 2RS_{yx}\right) \\ -\frac{1}{m^2 r}\left(\sum_{i=1}^{m} \tau_{y(i)}^2 + R^2 \sum_{i=1}^{m} \tau_{x(i)}^2 - 2R\sum_{i=1}^{m} \tau_{yx}\right) \qquad < 0 \end{array} \tag{51}$$

The proposed estimator is more efficient than the [?] ratio estimator if;

$$MSE\left(t_{pr2}\right) - MSE\left(\overline{y}_{R,RSS}\right) < 0 \tag{52}$$

So that;

$$\begin{array}{l} \overline{Y}^2 \left[ \begin{array}{l} \theta\left(C_y^2 - \gamma^2 C_x^2 - 2\gamma\rho C_x C_y - \alpha_{opt}\gamma\rho C_x C_y - \alpha_{opt}^2 \gamma^2 C_x^2\right) \\ -W_{y(i)}^2 + 2\gamma W_{yx(i)} + \alpha_{opt}\gamma W_{yx} + \alpha_{opt}^2 \gamma^2 W_{x(i)}^2 \end{array} \right] - \frac{1}{mr}\left(S_y^2 + R^2 S_x^2 - 2RS_{yx}\right) \\ -\frac{1}{m^2 r}\left(\sum_{i=1}^{m} \tau_{y(i)}^2 + R^2 \sum_{i=1}^{m} \tau_{x(i)}^2 - 2R\sum_{i=1}^{m} \tau_{yx}\right) \qquad < 0 \end{array} \tag{53}$$

The proposed estimator is more efficient than the [14] ratio estimator if;

$$MSE\left(t_{pr2}\right) - MSE\left(\overline{y}_{JL,RSS}\right) < 0 \tag{54}$$

So that;

$$\overline{Y}^2 \left[ \begin{array}{l} \theta\left(C_y^2 - \gamma^2 C_x^2 - 2\gamma\rho C_x C_y - \alpha_{opt}\gamma\rho C_x C_y - \alpha_{opt}^2 \gamma^2 C_x^2\right) \\ -W_{y(i)}^2 + 2\gamma W_{yx(i)} + \alpha_{opt}\gamma W_{yx} + \alpha_{opt}^2 \gamma^2 W_{x(i)}^2 \end{array} \right] - \overline{Y}^2 \left[ \begin{array}{l} \theta\left(C_y^2 + \lambda_1^2 C_x^2 - 2\lambda_1 \rho_{yx} C_x C_y\right) \\ -\left(W_{y(i)} - \lambda_1 W_{x(i)}\right)^2 \end{array} \right] < 0 \tag{55}$$

The proposed estimator is more efficient than the [15] ratio estimator if;

$$MSE\left(t_{pr2}\right) - MSE\left(\overline{y}_{M1,RSS}\right) < 0 \tag{56}$$

So that;

$$\overline{Y}^2 \left[ \begin{array}{l} \theta\left(C_y^2 - \gamma^2 C_x^2 - 2\gamma\rho C_x C_y - \alpha_{opt}\gamma\rho C_x C_y - \alpha_{opt}^2 \gamma^2 C_x^2\right) \\ -W_{y(i)}^2 + 2\gamma W_{yx(i)} + \alpha_{opt}\gamma W_{yx} + \alpha_{opt}^2 \gamma^2 W_{x(i)}^2 \end{array} \right] - \overline{Y}^2 \left[ \begin{array}{l} \theta\left(C_y^2 + \lambda_2^2 C_x^2 - 2\lambda_2 \rho_{yx} C_x C_y\right) \\ -\left(W_{y(i)} - \lambda_2 W_{x(i)}\right)^2 \end{array} \right] < 0 \tag{57}$$

The proposed estimator is more efficient than the [15] ratio estimator if;

$$MSE\left(t_{pr2}\right) - MSE\left(\overline{y}_{M2,RSS}\right) < 0 \tag{58}$$

So that;

$$\overline{Y}^2 \left[ \begin{array}{l} \theta\left(C_y^2 - \gamma^2 C_x^2 - 2\gamma\rho C_x C_y - \alpha_{opt}\gamma\rho C_x C_y - \alpha_{opt}^2 \gamma^2 C_x^2\right) \\ -W_{y(i)}^2 + 2\gamma W_{yx(i)} + \alpha_{opt}\gamma W_{yx} + \alpha_{opt}^2 \gamma^2 W_{x(i)}^2 \end{array} \right] - \overline{Y}^2 \left[ \begin{array}{l} \theta\left(C_y^2 + \lambda_3^2 C_x^2 - 2\lambda_3 \rho_{yx} C_x C_y\right) \\ -\left(W_{y(i)} - \lambda_3 W_{x(i)}\right)^2 \end{array} \right] < 0 \tag{59}$$

The proposed estimator is more efficient than the [16] ratio estimator if;

$$MSE\left(t_{pr2}\right) - MSE\left(\overline{y}_{RZ,RSS}\right) < 0 \tag{60}$$

So that;

$$\overline{Y}^2 \left[ \begin{array}{l} \theta \left( C_y^2 - \gamma^2 C_x^2 - 2\gamma\rho C_x C_y - \alpha_{opt}\gamma\rho C_x C_y - \alpha_{opt}^2 \gamma^2 C_x^2 \right) \\ -W_{y(i)}^2 + 2\gamma W_{yx(i)} + \alpha_{opt}\gamma W_{yx} + \alpha_{opt}^2 \gamma^2 W_{x(i)}^2 \end{array} \right] - \overline{Y}^2 \left[ \begin{array}{l} \theta \left( C_y^2 + \lambda_4^2 C_x^2 - 2\lambda_4 \rho_{yx} C_x C_y \right) \\ -\left( W_{y(i)} - \lambda_4 W_{x(i)} \right)^2 \end{array} \right] < 0 \tag{61}$$

## 3.2. Numerical Example

To further investigate the efficiency (performance) of the suggested estimators over other existing estimators, an empirical study was conducted to support the theoretical comparison. In the empirical study, numerical values were used in the comparison. Hypothetical data from natural populations were employed for the comparison. Here, we present a summary of the descriptive statistics and the constant parameters for the population data used in the empirical study. The data from [7] between 1920 and 1930, consisting of information on the population size (measured in 1000s) of 49 cities was utilize by [16] and the summary statistics is presented in table 1.

**Table 1:** *Data Statistics used in the numerical example*

| | | | |
|---|---|---|---|
| $N = 4$ | $\overline{X} = 103.14$ | $\mu_3(x) = 76$ | $\beta_1(x) = 2.20\$$ |
| $n = 8$ | $\overline{Y} = 127.8$ | $\mu_4(x) = 144$ | $\beta_2(x) = 7.22$ |
| $m = 4$ | $S_y = 32.8720$ | $\mu_1(y) = 70$ | $C_y = 12.95$ |
| $R = 2$ | $S_x = 56.2342$ | $\mu_2(y) = 87$ | $C_x = 16.05$ |
| $\rho = 0.18$ | $\mu_1(x) = 70$ | $\mu_3(y) = 93.5$ | $QD = 37.68$ |
| $Md = 130.77$ | $\mu_2(x) = 42$ | $\mu_4(y) = 103.5$ | $\alpha_{opt} = -0.871$ |

Table 2 presents the summary result of the calculated Mean squared error (MSE), Percent relative efficiency (PRE) and Coefficient of variation (C.V) of both the proposed and competing estimators. The PRE and C.V can be calculated respectively as;

$$PRE\left(t_{j,\,\overline{y}_{srs}}\right) = \frac{V\left(\overline{y}_{srs}\right)}{MSE\left(t_j\right)} \times 100 \tag{62}$$

$$C.V = \frac{\sqrt{MSE(t_j)}}{\overline{X}} \times 100 \tag{63}$$

**Table 2:** *Numerical Comparison of the Proposed and Competing Estimators*

| Estimators | MSE | PRE | CV |
|---|---|---|---|
| $\overline{y}$ | 638.900 | 100.000 | 24.507 |
| $\overline{y}_{SD}$ | 838.138 | 76.228 | 28.069 |
| $\overline{y}_{SK}$ | 761.490 | 83.901 | 26.755 |
| $\overline{y}_{R}$ | 703.065 | 90.874 | 25.708 |
| $\overline{y}_{R,RSS}$ | 622.145 | 102.693 | 24.183 |
| $\overline{y}_{M1,RSS}$ | 617.114 | 103.530 | 24.086 |
| $\overline{y}_{M2,RSS}$ | 516.486 | 123.701 | 22.034 |
| $\overline{y}_{RZ,RSS}$ | 273.867 | 233.289 | 16.045 |
| $\overline{y}_{JL,RSS}$ | 356.987 | 178.970 | 18.319 |
| $\overline{y}_{JK}$ | 367.567 | 173.819 | 18.588 |
| $t_{p1}$ | 250.567 | 254.982 | 15.347 |
| $t_{p2}$ | 249.775 | 255.790 | 15.323 |

Results from table 2 reveals the performance of the proposed estimators compared to the existing ones. The MSEs and CVs of the proposed estimators are smaller than those of competing estimators. Additionally, the PREs of the proposed estimators are higher than those of competing estimators indicating a gain in efficiencies.

## 4. Conclusion

In this paper, we use the sample size information of the auxiliary variable to propose a modified ratio estimator under ranked set sampling. The statistical properties of the proposed estimators were derived and compared both theoretically and numerically with some existing estimators. The results presented in Tables 2 confirm the superior performance of the proposed estimators. The significantly lower mean squared errors (MSEs) and coefficients of variation (CVs) demonstrate their robustness and precision compared to competing estimators. Additionally, the higher percent relative efficiencies (PREs) of the proposed estimators outperform their enhanced accuracy and reliability. These findings highlight the practical utility and statistical soundness of the modified estimators for real-world applications, particularly when parameter information is available. Overall, the proposed estimators emerge as the optimal choice, outperforming alternatives both theoretically and numerically; as such, they should therefore be used in applications.

## References

[1] Liu, Q., Wang, Z., He, X., & Zhou, D. (2014). A survey of event-based strategies on control and estimation. *Systems Science & Control Engineering* 2(1): 90-97.

[2] Fermanagh, D., & Rao, C. R. (2009). Sample surveys: design, methods and applications. *Elsevier*,

[3] Cochran, W.G (1940). The estimation of the yields of the cereal experiments by sampling for the ratio of grain to total produce. *The Journal of Agricultural science*, 30(2): 262-275

[4] Murthy, M.N. (1964). Product method of estimation. *Sakhya: The Indian journal of Statistics, series A*, 69-74.

[5] Singh, R., Mangat, N.S. (1996). Regression Method of Estimation. In: Elements of Survey Sampling. *Kluwer Texts in the Mathematical Sciences*, vol 15. Springer, Dordrecht, 1996.

[6] McIntyre, G.A. (1952). A method of unbiased selective sampling using ranked sets. *Australian Journal of Agric. Res.*, 3: 385-390.

[7] Cochran, W.G.(1997). *Sampling techniques* 3rd edition. Wiley Eastern, New Delhi.

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

[9] Singh, H P., R. Tailor and M. S. Kakran (2004). An improved estimator of population mean using power transformation. *Journal of the Indian society of Agricultural statistics*, 58(2), 223-230.

[10] Upadhyaya, L.N., Singh, H.P. (1999). Use of transformed auxiliary variable in estimating the finite population mean. *Biometry Journal*, 41: 627-636.

[11] Jerajuddin, M., & Kishun, J. (2016). Modified ratio estimators for population mean using size of the sample, selected from population. *International Journal of Scientific Research in Science, Engineering and Technology*, 2(2), 10-16.

[12] Samawi, H.M and Muttlak, H.A (1996). Estimation of ration using ranked set sampling. *Biometrical journal* 38(6):753-764.

[13] Kadilar, C., Unyazici, Y. and Cingi, H. (2009). Ratio estimator for the population mean using ranked set sampling. *Statistical papers*, 50(20): 301-309.

[14] Jeelani, M. I., Bouza, C. N., & Sharma, M. (2017). Modified ratio estimator under rank set sampling. *Investigacin Operacional*, 38(1), 103-107.

[15] Mehta, N. and Mandowara, V, L. (2012). A better estimator of population mean with power transformation best on ranked set sampling. *Sampling method and estimation*, 13(3):551-558.

[16] Riyaz, S., Rather, K.U., Maqbool, S., & Jan, T.R. (2024). Ratio estimator of population mean using a linear combination under ranked set sampling. *RT&A*, 2(73): 479-485.

[17] Sabo, S. A., Musa, I. Z., & Kibiya, Y. I. (2020). Developed ratio estimator for estimating population mean from a finite population using sample size and correlation. *Int. J. Res. Appl. Sci. Eng. Tech*, 8(6): 904-911.

[18] Kumari, A., Singh, R. & Smarandache, F. (2024). New Modification of ranked set sampling for estimating population mean: Neutrosophic median ranked set sampling with an application to demographic data. *Int J Comput Intell Syst*, 17(210): 1-15.

[19] Rehman, S. A., Al-Essa, L. A., Shabbir, J., & Khan, Z. (2024). On the efficiency of paired ranked set sampling for estimating the population mean in the presence of non-response. *Scientific Reports*, 14(1), 25334.

[20] Shahzad, N., & Tshung, F. C. C. (2023). Estimating Population Mean Using Ratio Estimator for Mean is Less Than Variance in Simple Random Sampling. *Russian Law Journal*, 11(3), 2049-2061