Estimation of Population Mean by Using a Generalized Family of Estimators Under Classical Ranked Set Sampling

Asad Ali, Muhammad Moeen Butt, Kanwal Iqbal, Muhammad Hanif and Muhammad Zubair

ABSTRACT
Estimation of population mean of study variable Y suffers loss of precision in the presence of high variation in the data set. The use of auxiliary information incorporated in construction of an estimator under Rank set sampling scheme results in efficient estimation of population mean. In this paper, we propose an efficient generalized family of estimators to estimate finite population mean of study variable under ranked set sampling utilizing information on an auxiliary variable. Bias and Mean Square Error (MSE) of the proposed generalized family of estimators are derived. The conditions of efficiency of proposed generalized family of estimators from competitor estimators are also derived. The applications of estimator are discussed using simulation study and real-life data sets for comparisons of efficiency. It is concluded that when correlation between study and auxiliary variables increases, the proposed generalized family of estimators proves to be the efficient estimator of population mean of the study variable.

1. Introduction
In many situations of practical interest, mainly in environmental and ecological studies, the variable of interest, say Y, is not easily observable in the sense that measurement may be expensive, time consuming, invasive or even destructive. Although data collection may be complex, ranking the potential sampled units with respect to an available auxiliary variable can often be relatively simple at no additional cost or for a very little cost. In those situations, where the variations in study variable is high and it is strongly correlated with auxiliary variable, the Ranked Set Sampling (RSS) proposed by McIntyre (McIntyre, 1952) is more efficient as compared to Simple Random Sampling (SRS) (Patil et al., 1993, 1994; Stokes, 1977).

Literature on RSS has rapidly grown and several estimators, originally conceived for the SRS, have been re-proposed to estimate the mean of the study variable by changing their sampling design into RSS framework (Ali et al., 2021; Iqbal et al., 2020; Kadilar et al., 2009; Khan & Shabbir, 2016a, 2016b, 2016c; Mandowara & Mehta, 2013, 2016; Mehta & Mandowara, 2016; Pelle & Perri, 2018; Samawi & Muttlak, 1996; Singh et al., 2014; Vishwakarma et al., 2017). Motivated by these studies, and in line with many other contributions present in the literature section of this article, we propose an efficient generalized family of estimators by changing the sampling design of the estimator proposed by Shahzad et al. (Shahzad et al., 2019).

Notations under Ranked Set Sampling Design
Let \( \Omega = \{1, 2, \ldots, N\} \) be a finite population of N units, Y the variable under study and X an auxiliary variable which is highly correlated with Y. Let \( \mu_Y \) and \( \mu_X \) denote the population means of Y and X, respectively, \( S^2_Y \) and \( S^2_X \) the variances, \( C_Y \) and \( C_X \) the coefficients of variation, \( \rho_{XY} \) the correlation coefficient between X and Y, \( \beta_{2(x)} \) and \( \beta_{1(X)} \) the kurtosis and skewness, and \( C_{xy} = \rho_{xy} C_Y C_X \). Let us denote \( (X_{j(y)}, Y_{j(y)}) \) as the pair of the \( j \)-th order statistics of X and the associated element Y in the \( j \)-th cycle. Then the ranked set sample is,

\[
\begin{align*}
(X_{1(1)}, Y_{1[1]}), \ldots, (X_{1[m]}, Y_{1[m]}), \\
(X_{2(1)}, Y_{2[1]}), \ldots, (X_{2[m]}, Y_{2[m]}), \\
\vdots \\
(X_{r(1)}, Y_{r[1]}), \ldots, (X_{r[m]}, Y_{r[m]}).
\end{align*}
\]

To obtain biases and mean square error, we consider following notations under RSS:

CONTACT Asad Ali, ranaasadali23@gmail.com Department of Quantitative Methods, School of Business and Economics, University of Management and Technology, Lahore, Pakistan

Reviewing editor: Guohua Zou, Chinese Academy of Sciences, Beijing, China

© 2021 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.

You are free to: Share — copy and redistribute the material in any medium or format. Adapt — remix, transform, and build upon the material for any purpose, even commercially. The licensor cannot revoke these freedoms as long as you follow the license terms. Under the following terms: Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use. No additional restrictions You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.
\[ e_0 = \frac{\bar{y}_{\text{true}} - \mu_x}{\mu_y}, \quad e_1 = \frac{x_{\text{true}} - \mu_x}{\mu_y}, \quad \bar{y}_{\text{true}} = \mu_y(1 + e_0), \]
\[ \sum_{i=1}^{m} x_{y(i)} \bar{x}_{i} = \frac{y_{i}}{r}, \quad \sum_{i=1}^{m} y_{i} = \frac{y_{i}}{r}, \]
\[ \tau_{x(i)} = \bar{x}_{i} - \mu_x; \text{ Deviation of } i \text{th cycle ranked mean from population mean } \mu_x. \]
\[ \tau_{y[i]} = \bar{y}_{[i]} - \mu_y; \text{ Deviation of } i \text{th cycle ranked mean from population mean } \mu_y. \]
\[ \tau_{xy} = (\bar{x}_{i} - \mu_x)(\bar{y}_{i} - \mu_y); \text{ Cross product of the deviations.} \]
\[ y = \frac{1}{n} - \frac{1}{N} = \frac{1}{m}, \quad C_{xy} = \frac{S_{xy}}{\mu_y^2}, \quad C_{y} = \frac{S_{yy}}{\mu_y^2}. \]
\[ W_{s(x)}^2 = \frac{\sum_{i=1}^{m} \tau_{x(i)}^2}{m \mu_x}, \quad W_{s(y)}^2 = \frac{\sum_{i=1}^{m} \tau_{y[i]}^2}{m \mu_y}, \quad W_{s(xy)} = \frac{\sum_{i=1}^{m} \tau_{xy}}{m \mu_x \mu_y}. \]

To obtain biases and mean square error, we consider following notations under SRS:
\[ \vartheta_0 = \frac{y_{\text{true}} - \mu_x}{\mu_y}, \quad \vartheta_1 = \frac{x_{\text{true}} - \mu_x}{\mu_x}. \]
\[ \bar{y}_{\text{true}} = \mu_y(1 + \vartheta_0), \quad \bar{x}_{\text{true}} = \mu_x(1 + \vartheta_1). \]
\[ E(e_0) = E(e_1) = 0, E(e_1^2) = V_{yy} - W_{s(y)}^2 = V_{xy}, E(e_1^2) = V_{xx} - W_{s(x)}^2 = V_{xx}, E(e_1^2) = V_{xy}. \]
\[ E(\vartheta_0) = 0, E(\vartheta_1) = \gamma C_y = V_{yy}, E(\vartheta_1^2) = \gamma C_x = V_{xx}, E(\vartheta_1^2) = V_{xy}. \]

(1.1)

2. Some Existing Estimators under SRS and RSS

Following are given some famous estimators along with their mean square errors.

Mandowara and Mehta (Mandowara & Mehta, 2013) proposed following estimators,
\[ \Delta_{\text{RSS}1} = \bar{y}_{\text{true}} \left\{ \frac{\mu_x C_x + \beta_2(x)}{\bar{x}_{\text{true}} C_x + \beta_2(x)} \right\} \delta. \]
\[ \Delta_{\text{RSS}2} = \bar{y}_{\text{true}} \left\{ \frac{\mu_x \beta_2(x) + C_x}{\bar{x}_{\text{true}} \beta_2(x) + C_x} \right\} \alpha. \]
\[ \Delta_{\text{RSS}3} = \bar{y}_{\text{true}} \left\{ \frac{\mu_x + \beta_2(x)}{\bar{x}_{\text{true}} + \beta_2(x)} \right\}. \]

The Bias and MSE of \( \Delta_{\text{RSS}1}, \Delta_{\text{RSS}2} \) and \( \Delta_{\text{RSS}3} \) are,

\[ B(\Delta_{\text{RSS}1}) = \frac{\mu_y}{2} \left[ \delta \frac{1}{(\delta + 1)} V_{xx} - 2 V_{xy} \right], \]
\[ \text{MSE}(\Delta_{\text{RSS}1}) = \mu_y^2 \left[ (V_{yy} + \delta^2 V_{xx} - 2 \delta \phi_2 V_{xy}) \right]. \]

(2.4)

\[ B(\Delta_{\text{RSS}2}) = \frac{\mu_y}{2} \left[ (\alpha + 1) V_{xx} - 2 V_{xy} \right], \]
\[ \text{MSE}(\Delta_{\text{RSS}2}) = \mu_y^2 \left[ V_{yy} + \alpha^2 \phi_2 V_{xx} - 2 \alpha \phi_1 V_{xy} \right]. \]

(2.5)

\[ B(\Delta_{\text{RSS}3}) = \mu_y \left( \lambda_2^2 V_{xx} - \lambda_2 V_{xy} \right), \]
\[ \text{MSE}(\Delta_{\text{RSS}3}) = \mu_y^2 \left( V_{yy} + \lambda_2^2 V_{xx} - 2 \lambda_2 V_{xy} \right). \]

Which are minimum for, \( \delta_{opt} = \rho_{xy} \phi_{C_C}, \phi_2 = \frac{\mu_Y C_C}{\rho_{xy} \phi_{C_C}}, \alpha_{opt} = \rho_{xy} \phi_{C_C}, \phi_1 = \frac{\mu_Y \beta_2(x)}{\rho_{xy} \beta_2(x) + C_x}, \lambda_2 = \frac{\mu_X}{\rho_{xy} \beta_2(x) + C_x}. \]

Vishwakarma, Zeeshan and Bouza (Vishwakarma et al., 2017) developed the following exponential estimator,
\[ \Delta_{\text{RSS}ext} = \bar{y}_{\text{true}} \exp \left\{ \frac{\mu_x - \bar{x}_{\text{true}}}{\mu_x^2 + \bar{x}_{\text{true}}^2} \right\}. \]

(2.7)

The MSE of \( \Delta_{\text{RSS}ext} \) is,
\[ \text{MSE}(\Delta_{\text{RSS}ext}) = \mu_y^2 \left[ \frac{V_{xx}}{4} + V_{yy} - V_{xy} \right]. \]

Shahzad et al. (Shahzad et al., 2019) introduced the generalized form of the estimators under simple random sampling given as,
\[ \Delta_{\text{RSS}SH} = \left[ \frac{w_{sh}}{\eta_{sh}} \right] \left[ \frac{\mu_x - \bar{x}}{\eta_{sh}(2 \mu_x - \phi \mu_x + \bar{x} + 2 \lambda)} \right] \left[ \eta_{sh} \left( \mu_x - \bar{x} \right) \right] ^{\alpha}. \]

(2.9)

Where \( w_{sh}, \eta_{sh}, a, b, \phi, \lambda \) are all constants. The bias and MSE of \( \Delta_{\text{RSS}SH} \) are,
\[ B(\Delta_{\text{RSS}SH}) = w_{sh} \mu_x(1 - k_1 V_{xy} + \left( \frac{g(2 + \alpha) + \alpha \gamma \phi_2 + \frac{3}{2} \lambda \gamma \phi_2}{2} \right) V_{xx} + w_{sh} \mu_y V_{xy} - \mu_y), \]
\[ \text{MSE}(\Delta_{\text{RSS}SH}) = \mu_y^2 + w_{sh}^2 \theta_{SH} + w_{sh}^2 \theta_{Sh} + w_{sh} \mu_y \theta_{Osh} + w_{sh} \mu_y \theta_{Dsh} + w_{sh} \mu_y \theta_{Esh}. \]
Where, 
\[ k_1 = a\Gamma + y_{ih}, \quad y_{ih} = \frac{(b_k+1)\eta \mu_y}{\eta \mu_y + 2\lambda}, \quad b_k = 2\xi - \phi, \]
\[ k_2 = \frac{g}{2} \alpha^2 \Gamma^2 + \alpha y_{ih} \Gamma g + \frac{1}{2} y_{ih}^2, \]
\[ \Theta_{Ath} = \mu_y^2 \left[ 1 + V_{yy} + (k_1^2 + 2k_2) V_{xx} - 4k_1 V_{xy} \right], \]
\[ \Theta_{Bth} = \mu_y^2 \left( k_1 V_{xy} - k_2 V_{xx} - 1 \right), \]
\[ \Theta_{Eth} = -2\mu_y V_{xx}. \]

Which is minimum for,
\[ w_{sh1}^{opt} = \frac{\Theta_{Ath} \Theta_{Bth} - \Theta_{Eth} \Theta_{Ath} \Theta_{Eth}}{4\Theta_{Ath} \Theta_{Bth} - \Theta_{Eth}^2}. \]

The minimum MSE of \( \Delta_{(RSS)ih} \) is given by,
\[ \text{MSE}(\Delta_{(RSS)ih})_{min} = \frac{\mu_y^2}{\Theta_{Bth} \Theta_{Eth}^2 - \Theta_{Bth}^2 \Theta_{Eth}}. \]

**3. Proposed Generalized Family of Estimators under RSS**

Motivated from Shahzad et al. (Shahzad et al., 2019), we propose the following generalized family of estimators under Ranked set sampling,

\[ \Delta_{(RSS)G} = \left[ \frac{w_1 y_{[r_1]} \left( \frac{ap_i + b}{a(x_{(n)}) + b} + (1-a)(ap_i + b) \right)^g}{w_2 \left( \mu_x - \bar{x}_{(r_1)} \right)} \right]. \]

**Derivation of Bias and Mean Square Error**

Rewriting the above estimator with “e” terms under first order of approximation we get,

\[ \Delta_{(RSS)G} = \left[ \frac{w_1 \mu_y \left( 1 + e_0 - ag \Gamma e_1 + \frac{g}{2} \alpha^2 \Gamma^2 e_1^2 \right)}{-ag \Gamma e_0 e_1} \right]. \]

Where,
\[ \Gamma = \frac{(ap_i + b)}{a(x_{(n)}) + b}, k_1 = (ag \Gamma + \nu), k_2 = \frac{g}{2} \alpha^2 \Gamma^2 + vag + \frac{3}{2} \nu^2, b_k = 2\xi - \phi, \nu = \frac{\eta \mu_y}{\eta \mu_y + 2\lambda}. \]

Subtracting \( \mu_y \) on both sides,
\[ \Delta_{(RSS)G} - \mu_y = w_1 \mu_y (1 + e_0 - k_1 e_1 + k_2 e_1^2 - k_1 e_0 e_1) \]
\[ - w_2 \mu_y (e_1 - e_1^2) - \mu_y. \]

**Bias**

For bias, we apply expectation on both sides of 3.2, the expression for bias of \( \Delta_{(RSS)G} \) is given as,
\[ \text{Bias}(\Delta_{(RSS)G}) = \left[ w_1 \mu_y (1 - k_1 V_{xy} + k_2 V_{xx}) + w_2 \mu_y V_{xx} - \mu_y \right]. \]

For MSE, we apply square and expectation on both sides of 3.2, the expression for MSE of \( \Delta_{(RSS)G} \) is given as,
\[ \text{MSE}(\Delta_{(RSS)G}) = \left[ \mu_y^2 + w_1^2 \Theta_{Aam} + w_2^2 \Theta_{Bam} + w_1 w_2 \Theta_{Cam} \right]. \]

Where,
\[ \Theta_{Aam} = \left[ \mu_y^2 (1 + V_{yy} + (k_1^2 + 2k_2) V_{xx} - 4k_1 V_{xy}) \right], \]
\[ \Theta_{Bam} = \mu_y^2 V_{xx}, \Theta_{Cam} = 2\mu_y \left( k_1 V_{xy} - k_2 V_{xx} - 1 \right), \]
\[ \Theta_{Eam} = -2\mu_y V_{xx}. \]

For minimizing MSE, we obtained the optimum values of \( w_1 \) and \( w_2 \) as follows:
\[ w_1^{opt} = \frac{-2\Theta_{Bam} \Theta_{Eam} + \Theta_{Cam} \Theta_{Cam} - \Theta_{Bam}}{4 \Theta_{Aam} \Theta_{Bam} - \Theta_{Cam}^2}, \]
\[ w_2^{opt} = \frac{\Theta_{Cam} \Theta_{Eam} - 2\Theta_{Aam} \Theta_{Eam}}{4 \Theta_{Aam} \Theta_{Bam} - \Theta_{Cam}^2}. \]

Hence, the minimum Bias and MSE are given by,
\[ \text{Bias}(\Delta_{(RSS)G})_{min} = \left[ w_1^{opt} \mu_y \left( 1 - k_1 V_{xy} + k_2 V_{xx} \right) + w_2^{opt} \mu_y V_{xx} - \mu_y \right]. \]
\[ \text{MSE}(\Delta_{(RSS)G})_{min} = \left[ \mu_y^2 - \Theta_{Bam} \Theta_{Eam} - \Theta_{Bam} + \Theta_{Cam} \Theta_{Cam} \right]. \]

The estimator of \( \text{Bias}(\Delta_{(RSS)G})_{min} \) and \( \text{MSE}(\Delta_{(RSS)G})_{min} \) based on sample measurements are given as follows:
\[ \text{Bias}(\Delta_{(RSS)G})_{min} = \left[ w_1^{opt} \mu_y \left( 1 - k_1 V_{xy} + k_2 V_{xx} \right) + w_2^{opt} \mu_y V_{xx} - \mu_y \right], \]
4. Efficiency Comparison

We derive the theoretical conditions to compare the efficiency of our proposed generalized family of estimators to their competitor estimators.

(i) By (3.3) and (2.4),

\[
MSE(\Delta_{(RSS)G})_{\text{min}} < MSE(\Delta_{(RSS)mm1})
\]

if

\[
\left[ \mu_y^2 - \frac{\Theta_{Bam}\Theta_{Dam}^2 + \Theta_{Aam}\Theta_{Eam}^2 - \Theta_{Cam}\Theta_{Dam}\Theta_{Eam}}{4\Theta_{Aam}\Theta_{Bam} - \Theta_{Cam}^2} \right. - \left. \mu_y^2 \left( V_{yy} + \delta^2 \phi_2 V_{xx} - 2 \phi_2 V_{xy} \right) \right] < 0
\]

(i) By (3.3) and (2.5),

\[
MSE(\Delta_{(RSS)G})_{\text{min}} < MSE(\Delta_{(RSS)mm2})
\]

if

\[
\left[ \mu_y^2 - \frac{\Theta_{Bam}\Theta_{Dam}^2 + \Theta_{Aam}\Theta_{Eam}^2 - \Theta_{Cam}\Theta_{Dam}\Theta_{Eam}}{4\Theta_{Aam}\Theta_{Bam} - \Theta_{Cam}^2} \right. - \left. \mu_y^2 \left( V_{yy} + \alpha^2 \phi_1^2 V_{xx} - 2 \phi_1 \phi_2 V_{xy} \right) \right] < 0
\]

(i) By (3.3) and (2.6),

\[
MSE(\Delta_{(RSS)G})_{\text{min}} < MSE(\Delta_{(RSS)mm4})
\]

if

\[
\left[ \mu_y^2 - \frac{\Theta_{Bam}\Theta_{Dam}^2 + \Theta_{Aam}\Theta_{Eam}^2 - \Theta_{Cam}\Theta_{Dam}\Theta_{Eam}}{4\Theta_{Aam}\Theta_{Bam} - \Theta_{Cam}^2} \right. - \left. \mu_y^2 \left( V_{yy} + \lambda_2^2 V_{xx} - 2 \lambda_2 V_{xy} \right) \right] < 0
\]

Note: When these conditions are satisfied, the proposed estimators will perform more efficiently as compared to their competitor estimators.

5. Simulation Study

A hypothetical data for the study variable (Y) and auxiliary variable (X) is generated by using Bivariate Normal Distribution with parameters,

\[
N = 1500, n = 20, 30, m = 5, 10, r = 4, 3, \mu_x = 850, \mu_y = 550, C_y = 1.25, C_x = 1.5, \rho_{xy} = \{0.4, 0.7, 0.8, 0.9, 0.99\}
\]

Samples of different values of n have been simulated 50,000 to calculate their average mean square errors and percent relative efficiencies. Percent Relative Efficiencies (PRe) of our proposed generalized family of estimators along with competitor estimators from literature have been presented in Tables 1, 2, 3, 4 and 5 for different values of n and \(\rho_{xy}\).

Table 1 shows that, when correlation coefficient of x and y equals to 0.4 and \(n = 20\), our proposed estimator will be 257.31% more efficient than \(\Delta_{(SRS)sh}\). In the same situation proposed estimator is 204.05%, 201.47%, 223.89% and 224.67% more efficient than \(\Delta_{(RSS)mm1}\), \(\Delta_{(RSS)mm2}\), \(\Delta_{(RSS)mm4}\) and \(\Delta_{(RSS)yz}\) respectively. At the same correlation coefficient, when we increase the sample size by \(n = 30\), our proposed estimator will be 262.10% more efficient then \(\Delta_{(SRS)sh}\). In the same situation proposed estimator is 201.59%, 198.80%, 237.06% and
### Table 1. PRE of Estimators by Simulation Study with $\rho_{xy} = 0.4$

| $n$ | $\Delta_{(RSS)sh}$ | $\Delta_{(RSS)mm1}$ | $\Delta_{(RSS)mm2}$ | $\Delta_{(RSS)mm4}$ | $\Delta_{(RSS)vz}$ | $\Delta_{(RSS)G}$ |
|-----|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| 20  | 100                  | 100                  | 100                  | 100                  | 100                  | 100                  |
|     | 126.10               | 101.28               | 89.98                | 100                  | 99.66                | 100                  |
|     | 114.92               | 91.14                | 89.68                | 100                  | 99.66                | 100                  |
|     | 114.53               | 90.82                | 89.68                | 100                  | 100                  | 100                  |
|     | 257.31               | 204.05               | 201.47               | 223.89               | 224.67               | 100                  |
| 30  | 100                  | 100                  | 100                  | 100                  | 100                  | 100                  |
|     | 130.02               | 101.40               | 100                  | 100                  | 100                  | 100                  |
|     | 131.84               | 85.04                | 83.86                | 100                  | 100                  | 100                  |
|     | 112.82               | 86.78                | 85.57                | 102.04               | 100                  | 100                  |
|     | 262.10               | 201.59               | 198.80               | 237.06               | 232.31               | 100                  |

### Table 2. PRE of Estimators by Simulation Study with $\rho_{xy} = 0.7$

| $n$ | $\Delta_{(RSS)sh}$ | $\Delta_{(RSS)mm1}$ | $\Delta_{(RSS)mm2}$ | $\Delta_{(RSS)mm4}$ | $\Delta_{(RSS)vz}$ | $\Delta_{(RSS)G}$ |
|-----|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| 20  | 100                  | 100                  | 100                  | 100                  | 100                  | 100                  |
|     | 134.21               | 101.47               | 100                  | 100                  | 100                  | 100                  |
|     | 117.70               | 87.70                | 86.44                | 100                  | 100                  | 100                  |
|     | 118.76               | 88.49                | 87.21                | 100.89               | 100                  | 100                  |
|     | 274.65               | 204.64               | 201.69               | 233.34               | 231.27               | 100                  |
| 30  | 100                  | 100                  | 100                  | 100                  | 100                  | 100                  |
|     | 152.31               | 103.50               | 100                  | 100                  | 100                  | 100                  |
|     | 157.63               | 83.10                | 80.29                | 100                  | 100                  | 100                  |
|     | 132.20               | 86.80                | 83.87                | 104.46               | 100                  | 100                  |
|     | 319.29               | 206.63               | 202.35               | 252.28               | 241.51               | 100                  |

### Table 3. PRE of Estimators by Simulation Study with $\rho_{xy} = 0.8$

| $n$ | $\Delta_{(RSS)sh}$ | $\Delta_{(RSS)mm1}$ | $\Delta_{(RSS)mm2}$ | $\Delta_{(RSS)mm4}$ | $\Delta_{(RSS)vz}$ | $\Delta_{(RSS)G}$ |
|-----|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| 20  | 100                  | 100                  | 100                  | 100                  | 100                  | 100                  |
|     | 167.83               | 104.24               | 100                  | 100                  | 100                  | 100                  |
|     | 147.11               | 87.65                | 84.09                | 100                  | 100                  | 100                  |
|     | 151.53               | 90.29                | 86.62                | 103.00               | 100                  | 100                  |
|     | 366.52               | 218.39               | 209.32               | 249.15               | 241.89               | 100                  |
| 30  | 100                  | 100                  | 100                  | 100                  | 100                  | 100                  |
|     | 206.64               | 102.62               | 100                  | 100                  | 100                  | 100                  |
|     | 158.83               | 76.86                | 74.90                | 100                  | 100                  | 100                  |
|     | 179.49               | 86.86                | 84.65                | 113.01               | 100                  | 100                  |
|     | 394.62               | 190.97               | 186.10               | 240.46               | 219.85               | 100                  |

### Table 4. PRE of Estimators by Simulation Study with $\rho_{xy} = 0.9$

| $n$ | $\Delta_{(RSS)sh}$ | $\Delta_{(RSS)mm1}$ | $\Delta_{(RSS)mm2}$ | $\Delta_{(RSS)mm4}$ | $\Delta_{(RSS)vz}$ | $\Delta_{(RSS)G}$ |
|-----|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| 20  | 100                  | 100                  | 100                  | 100                  | 100                  | 100                  |
|     | 207.81               | 108.82               | 100                  | 100                  | 100                  | 100                  |
|     | 149.67               | 72.02                | 66.18                | 100                  | 100                  | 100                  |
|     | 184.58               | 88.82                | 81.62                | 123.32               | 100                  | 100                  |
|     | 437.43               | 210.49               | 193.42               | 292.25               | 236.98               | 100                  |
| 30  | 100                  | 100                  | 100                  | 100                  | 100                  | 100                  |
|     | 278.87               | 103.85               | 100                  | 100                  | 100                  | 100                  |
|     | 162.36               | 58.22                | 56.06                | 100                  | 100                  | 100                  |
|     | 199.94               | 71.70                | 69.04                | 123.14               | 100                  | 100                  |
|     | 516.48               | 185.20               | 178.34               | 318.10               | 258.32               | 100                  |
232.31% more efficient than $\Delta_{(RSS)mm1}$, $\Delta_{(RSS)mm2}$, $\Delta_{(RSS)mm4}$ and $\Delta_{(RSS)vz}$ respectively. Table 2 shows that, when correlation coefficient of x and y equals to 0.7 and $n = 20$, our proposed estimator will be 274.65% more efficient then $\Delta_{(SRS)sh}$. In the same situation proposed estimator is 204.64%, 201.69%, 233.34% and 231.27% more efficient than $\Delta_{(RSS)mm1}$, $\Delta_{(RSS)mm2}$, $\Delta_{(RSS)mm4}$ and $\Delta_{(RSS)vz}$ respectively. At the same correlation coefficient, when we increase the sample size by $n = 30$, our proposed estimator will be 319.29% more efficient then $\Delta_{(SRS)sh}$. In the same situation proposed estimator is 209.63%, 202.55%, 252.28% and 241.51% more efficient than $\Delta_{(RSS)mm1}$, $\Delta_{(RSS)mm2}$, $\Delta_{(RSS)mm4}$ and $\Delta_{(RSS)vz}$ respectively. Table 3 shows that, when correlation coefficient of x and y equals to 0.8 and $n = 20$, our proposed estimator will be 366.52% more efficient than $\Delta_{(SRS)sh}$. In the same situation proposed estimator will be 218.39%, 209.52%, 249.15% and 241.89% more efficient then $\Delta_{(RSS)mm1}$, $\Delta_{(RSS)mm2}$, $\Delta_{(RSS)mm4}$ and $\Delta_{(RSS)vz}$ respectively. At the same correlation coefficient, when we increase the sample size by $n = 30$, our proposed estimator will be 394.62% more efficient then $\Delta_{(SRS)sh}$. In the same situation proposed estimator is 190.97%, 186.10%, 248.46% and 219.85% more efficient than $\Delta_{(RSS)mm1}$, $\Delta_{(RSS)mm2}$, $\Delta_{(RSS)mm4}$ and $\Delta_{(RSS)vz}$ respectively.

Table 4 shows that, when correlation coefficient of x and y equals to 0.9 and $n = 20$, our proposed estimator will be 437.43% more efficient then $\Delta_{(SRS)sh}$. In the same situation proposed estimator will be 210.49%, 193.42%, 292.25% and 236.98% more efficient then $\Delta_{(RSS)mm1}$, $\Delta_{(RSS)mm2}$, $\Delta_{(RSS)mm4}$ and $\Delta_{(RSS)vz}$ respectively. At the same correlation coefficient, when we increase the sample size by $n = 30$, our proposed estimator will be 516.48% more efficient then $\Delta_{(SRS)sh}$. In the same situation proposed estimator is 185.20%, 178.34%, 318.10% and 258.32% more efficient than $\Delta_{(RSS)mm1}$, $\Delta_{(RSS)mm2}$, $\Delta_{(RSS)mm4}$ and $\Delta_{(RSS)vz}$ respectively. Table 5 shows that, when correlation coefficient of x and y equals to 0.99 and $n = 20$, our proposed estimator will be 614.42% more efficient then $\Delta_{(SRS)sh}$. In the same situation proposed estimator is 278.19%, 269.29%, 367.86% and 331.27% more efficient then $\Delta_{(RSS)mm1}$, $\Delta_{(RSS)mm2}$, $\Delta_{(RSS)mm4}$ and $\Delta_{(RSS)vz}$ respectively. At the same correlation coefficient, when we increase the sample size by $n = 30$, our proposed estimator will be 680.86% more efficient then $\Delta_{(SRS)sh}$. In the same situation proposed estimator is 227.26%, 216.05%, 351.55% and 305.36% more efficient than $\Delta_{(RSS)mm1}$, $\Delta_{(RSS)mm2}$, $\Delta_{(RSS)mm4}$ and $\Delta_{(RSS)vz}$ respectively. Simulated results in Tables 1, 2, 3, 4 and 5 show the trend that when we increase the sample size, efficiency of proposed estimators under RSS design also increases as compare to estimator under SRS design.

Results also revealed that as we increase the $p_{xy}$, proposed estimator in RSS performs more efficiently as compared to its competitor estimator in SRS (i.e. $\Delta_{(SRS)sh}$). Therefore, we may say that as the correlation coefficient of x and y increases, the use of RSS is more appropriate as compared to SRS.

6. Real-Life Applications

To observe performances of the estimators, we use the following real-life data sets. The descriptions of these populations are given below.

**Population 1** [source: (James et al., 2013)]

The summary statistics are given below.

- **Y**: Acceleration of automobiles
- **X**: Engine horsepower of automobiles

Objective: To estimate population mean of Acceleration of automobiles.

| $n$ | $\Delta_{(SRS)sh}$ | $\Delta_{(RSS)mm1}$ | $\Delta_{(RSS)mm2}$ | $\Delta_{(RSS)mm4}$ | $\Delta_{(RSS)vz}$ | $\Delta_{(RSS)G}$ |
|-----|-------------------|---------------------|---------------------|---------------------|---------------------|-------------------|
| 20  | 100               | 220.87              | 228.16              | 167.03              | 185.47              | 614.42           |
| 30  | $\Delta_{(SRS)sh}$ | $\Delta_{(RSS)mm1}$ | $\Delta_{(RSS)mm2}$ | $\Delta_{(RSS)mm4}$ | $\Delta_{(RSS)vz}$ | $\Delta_{(RSS)G}$ |
|     | 100               | 229.60              | 105.19              | 315.14              | 193.67              | 222.97           |
|     | 200               | 351.55              | 305.36              | 305.36              | 305.36              | 305.36           |
Population II [source: (Multiple Indicator Cluster Survey (MICS, 2018–19))]
The summary statistics are given as below.
Y: Body Mass Index (BMI)
X: Weight
Objective: To estimate population mean of Body Mass Index (BMI).

N = 39118, n = 30, m = 10, r = 3,
μx = 12.1883, μy = 16.8151, Sy = 10.8438, Sx = 10.7911,
Cx = 0.8854, Cy = 0.6449, Cxy = 0.4877,
β2(x) = 54.4802, β1(x) = 7.0622, ρxy = 0.5542,

Population III [source: (Multiple Indicator Cluster Survey (MICS, 2018–19))]
The summary statistics are:
Y: Weight
X: Height
Objective: To estimate population mean of Weight.

N = 39118, n = 30, m = 10, r = 3,
μx = 94.6221, μy = 12.1883, Sy = 10.7911,
Sx = 101.0391, Cx = 1.0678, Cy = 0.8853,
Cxy = 0.7016, β2(x) = 74.6241,
β1(x) = 8.658, ρxy = 0.7421,

Population IV [source: (Daly et al., 2001)]
The summary statistics are:
Y: Body Mass Index (BMI) of Crohn’s disease patients
X: Weight of Crohn’s disease patients
Objective: To estimate population mean of Body Mass Index (BMI) of Crohn’s disease patients.

N = 117, n = 20, m = 5, r = 4,
μx = 69.0256, μy = 26.0624, Sy = 4.9888, Sx = 14.2438,
Cx = 0.2063, Cy = 0.1914, Cxy = 0.0325, β2(x) = 0.7746,
β1(x) = 0.6571, ρxy = 0.8222,

Population V [source: (Husby et al., 2005)]
The summary statistics are:
Y: Body Mass Index (BMI)
X: Thigh Circumference
Objective: To estimate population mean of Body Mass Index (BMI).

N = 36, n = 8, m = 4, r = 2,
μx = 49.3806, μy = 25.678, Sy = 3.8198, Sx = 3.7599,
Cx = 0.0761, Cy = 0.1488, Cxy = 0.0066, β2(x) = -0.6159,
β1(x) = -0.0607, ρxy = 0.9848,

Percent Relative Efficiencies (PREs) of our proposed generalized family of estimators along with competitor estimators from literature have been presented in Tables 6, 7, 8, 9 and 10 for different real-life populations. Table 6 shows that, when we consider the population I, our proposed estimator will be 363.74% more efficient than Δ1(RSS)dr. In the same situation the proposed estimator is 150.18%, 150.18%, 150.18% and 153.49% more efficient than Δ1(RSS)mm1, Δ1(RSS)mm2, Δ1(RSS)mm4 and Δ1(RSS)ve respectively. Table 7 shows that, when we consider the population II, our proposed estimator will be 418.12% more efficient than Δ1(RSS)dr. In the same situation proposed estimator will be 140.35%, 140.35%, 177.79% and 151.89% more efficient than Δ1(RSS)mm1, Δ1(RSS)mm2, Δ1(RSS)mm4 and Δ1(RSS)ve respectively. Table 8 shows that, when we consider the population III, our proposed estimator will be 149.93% more efficient than Δ1(RSS)dr. In the same situation proposed estimator will be 118.81%, 118.81%, 127.66% and 121.52% more efficient than Δ1(RSS)mm1, Δ1(RSS)mm2, Δ1(RSS)mm4 and Δ1(RSS)ve respectively. Table 9 shows that, when we consider the population IV, our proposed estimator will be 221.91% more efficient than Δ1(RSS)dr. In the same situation proposed estimator will be 180.39% 

| Δ1(RSS)dr | Δ1(RSS)mm1 | Δ1(RSS)mm2 | Δ1(RSS)mm4 | Δ1(RSS)ve | Δ1(RSS)d |
|-----------|-------------|-------------|-------------|-----------|-----------|
| 100       | 242.19      | 242.19      | 234.45      | 236.97    | 363.74    |
| 100       | 100         | 100         | 96.81       | 97.84     | 150.18    |
| 100       | 100         | 100         | 100         | 101.08    | 155.15    |
| 100       | 100         | 100         | 100         | 100       | 153.49    |
| 100       | 100         | 100         | 100         | 100       | 100       |
7. Conclusion

In this study, we proposed generalized family of estimators under RSS to estimate the finite population mean motivated from Shahzad et al. (Shahzad et al., 2019). The biases and MSEs of the proposed estimators were derived up to first order of approximation. The efficiency conditions for the proposed generalized estimator were also derived. On the basis of simulation study and real-life data sets, MSEs of all estimators have been computed and it is shown that the proposed generalized family of estimators are more efficient than the competitor estimators under SRS and RSS. It may concluded that with an increase in sample size and $\rho_{xy}$ the proposed estimator in RSS performs more efficiently as compared to its competitor estimators in SRS (i.e. $\Delta_{(SRS)sh}$).

### Table 7. PRE of Estimators for Population II

| $\Delta_{(RSS)sh}$ | $\Delta_{(RSS)mm1}$ | $\Delta_{(RSS)mm2}$ | $\Delta_{(RSS)mm4}$ | $\Delta_{(RSS)vz}$ | $\Delta_{(RSS)G}$ |
|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| 100               | 100               | 100               | 100               | 100               | 100               |
| 297.91            | 100               | 100               | 100               | 100               | 100               |
| 297.91            | 78.94             | 78.94             | 100               | 100               | 100               |
| 235.16            | 92.41             | 92.41             | 117.06            | 100               | 100               |
| 275.28            | 140.35            | 140.35            | 177.79            | 151.89            | 100               |
| 418.12            |                   |                   |                   |                   |                   |

### Table 8. PRE of Estimators for Population III

| $\Delta_{(RSS)sh}$ | $\Delta_{(RSS)mm1}$ | $\Delta_{(RSS)mm2}$ | $\Delta_{(RSS)mm4}$ | $\Delta_{(RSS)vz}$ | $\Delta_{(RSS)G}$ |
|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| 100               | 100               | 100               | 100               | 100               | 100               |
| 126.19            | 100               | 100               | 100               | 100               | 100               |
| 126.19            | 93.07             | 93.07             | 100               | 100               | 100               |
| 117.44            | 97.77             | 97.77             | 105.06            | 100               | 100               |
| 123.38            | 118.81            | 118.81            | 127.66            | 121.52            | 100               |
| 149.93            |                   |                   |                   |                   |                   |

### Table 9. PRE of Estimators for Population IV

| $\Delta_{(RSS)sh}$ | $\Delta_{(RSS)mm1}$ | $\Delta_{(RSS)mm2}$ | $\Delta_{(RSS)mm4}$ | $\Delta_{(RSS)vz}$ | $\Delta_{(RSS)G}$ |
|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| 100               | 100               | 100               | 100               | 100               | 100               |
| 207.18            | 100               | 100               | 100               | 100               | 100               |
| 207.18            | 106.77            | 106.77            | 100               | 100               | 100               |
| 221.19            | 88.72             | 88.72             | 83.09             | 100               | 100               |
| 183.81            | 107.11            | 107.11            | 100.32            | 120.72            | 100               |
| 221.91            |                   |                   |                   |                   |                   |

### Table 10. PRE of Estimators for Population V

| $\Delta_{(RSS)sh}$ | $\Delta_{(RSS)mm1}$ | $\Delta_{(RSS)mm2}$ | $\Delta_{(RSS)mm4}$ | $\Delta_{(RSS)vz}$ | $\Delta_{(RSS)G}$ |
|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| 100               | 100               | 100               | 100               | 100               | 100               |
| 393.48            | 100               | 100               | 100               | 100               | 100               |
| 393.48            | 84.04             | 84.04             | 100               | 100               | 100               |
| 330.69            | 55.85             | 55.85             | 66.45             | 100               | 100               |
| 219.75            | 104.32            | 104.32            | 124.13            | 186.79            | 100               |
| 410.47            |                   |                   |                   |                   |                   |

**estimator will be 107.11%, 107.11%, 100.32% and 120.72% more efficient then $\Delta_{(RSS)mm1}$, $\Delta_{(RSS)mm2}$, $\Delta_{(RSS)mm4}$ and $\Delta_{(RSS)vz}$ respectively.**
Acknowledgements

The authors wish to thanks DG-Bureau of Statistics Punjab for providing the data about Multiple Indicator Cluster Survey (MICS) for the year 2018–19.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The authors received no direct funding for this research.

ORCID

Asad Ali http://orcid.org/0000-0001-9032-9878
Muhammad Moeen Butt http://orcid.org/0000-0003-1736-8998
Kanwal Iqbal http://orcid.org/0000-0001-5186-1845
Muhammad Hanif http://orcid.org/0000-0002-0775-3457
Muhammad Zubair http://orcid.org/0000-0001-7142-9929

Public interest statement

Estimation of population parameters with minimum mean square error is very important issue of survey sampling. Different sampling designs and estimators have been proposed by researchers to deal with this issue. In this study, we proposed a generalized family of estimators for estimating population mean under classical ranked set sampling. Mathematical comparison, simulation study and real-life applications have been utilized for comparison of efficiency.

References

Ali, A., Butt, M. M., Azad, M. D., Ahmed, Z., & Hanif, M. (2021). Stratified Extreme-cum-Median Ranked Set sampling. Pakistan Journal of Statistics, 37(3), 215–235.
Daly, M. J., Rioux, J. D., Schaffner, S. F., Hudson, T. J., & Lander, E. S. (2001). High-resolution haplotype structure in the human genome. Nature Genetics, 29(2), 229–232. https://doi.org/10.1038/ng1001-229
Husby, C. E., Stasny, E. A., & Wolfe, D. A. (2005). An application of ranked set sampling for mean and median estimation using USDA crop production data. Journal of Agricultural, Biological, and Environmental Statistics, 10 (3), 354–373. https://doi.org/10.1198/108571105X58234
Iqbal, K., Moeen, M., Ali, A., & Iqbal, A. (2020). Mixture regression cum ratio estimators of population mean under stratified random sampling. Journal of Statistical Computation and Simulation, 90(5), 854–868. https://doi.org/10.1080/00949655.2019.1710149
James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An introduction to statistical learning (Vol. 112, pp. 18). Springer.
Kadilar, C., Unyazici, Y., & Cingi, H. (2009). Ratio estimator for the population mean using ranked set sampling. Statistical Papers, 50(2), 301–309. https://doi.org/10.1007/s00362-007-0079-y
Khan, L., & Shabbir, J. (2016a). A class of Hartley-Ross type unbiased estimators for population mean using ranked set sampling. Hacettepe Journal of Mathematics and Statistics, 45(3), 917–928.
Khan, L., & Shabbir, J. (2016b). An efficient class of estimators for the finite population mean in ranked set sampling. Open Journal of Statistics, 6(3), 426–435. https://doi.org/10.4236/ojs.2016.63038
Khan, L., & Shabbir, J. (2016c). Hartley-Ross type unbiased estimators using ranked set sampling and stratified ranked set sampling. The North Carolina Journal of Mathematics and Statistics, 2, 10–22.
Mandowara, V. L., & Mehta, N. (2013). Efficient generalized ratio-product type estimators for finite population mean with ranked set sampling. Austrian Journal of Statistics, 42 (3), 137–148. https://doi.org/10.17713/ajs.v42i3.147
Mandowara, V. L., & Mehta, N. (2016). On the improvement of product method of estimation in ranked set sampling. Chilean Journal of Statistics (Chis), 7(1), 43–53.
McIntyre, G. A. (1952). A method for unbiased selective sampling, using ranked sets. Australian Journal of Agricultural Research, 3(4), 385–390. https://doi.org/10.1017/S000492795200034X
Mehta, N., & Mandowara, V. L. (2016). A is the modified ratio-cum-product estimator of finite population mean using ranked set sampling. Communications in Statistics-Theory and Methods, 45(2), 267–276. https://doi.org/10.1080/03610926.2013.830748
Multiple Indicator Cluster Survey (MICS, 2018-19). Unpublished Survey, Bureau of Statistics Punjab. [http://bos.gop.pk/mics ]
Patil, G. P., Sinha, A. K., & Taille, C. (1993). Relative precision of ranked set sampling: A comparison with the regression estimator. Environmetrics, 4(4), 399–412. https://doi.org/10.1002/env.3170040404
Patil, G. P., Sinha, A. K., & Taille, C. (1994). Ranked set sampling. Handbook of Statistics, 12, 167–200.
Pelle, E., & Perri, P. F. (2018). Improving mean estimation in ranked set sampling using the Rao regression-type estimator. Brazilian Journal of Probability and Statistics, 32(3), 467–496. https://doi.org/10.1214/17-BJPS350
Samawi, H. M., & Muttlak, H. A. (1996). Estimation of ratio using rank set sampling. Biometrical Journal, 38(6), 753–764. https://doi.org/10.1002/bimj.4710380616
Shahzad, U., Hanif, M., Koyuncu, N., Luengo, A. V. G., & Khan, N. (2019). An efficient generalized family of estimators for mean estimation under simple random sampling. Investigación Operacional, 40(1), 28–45.
Singh, H. P., Tailor, R., & Singh, S. (2014). General procedure for estimating the population mean using ranked set sampling. Journal of Statistical Computation and Simulation, 84(5), 931–945. https://doi.org/10.1080/00949655.2012.733395
Stokes, L. S. (1977). Ranked set sampling with concomitant variables. Communications in Statistics-Theory and Methods, 6(12), 1207–1211. https://doi.org/10.1080/03610927708827563
Vishwakarma, G. K., Zeeshan, S. M., & Bouza, C. N. (2017). Ratio and product type exponential estimators for population mean using ranked set sampling. Investigación Operacional, 38(3), 266–271.