Heliyon 10 (2024) e27522

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Enhanced exponential ratio-cum-ratio estimator in ranked set sampling using transformed auxiliary information

Lakhkar Khan [a], Sohaib Ahmad [b], Abdullah Mohammed Alomair [c,*], Mohammed Ahmed Alomair [c]

[a] Department of Statistics, Government Post Graduate College, Mardan, Pakistan
[b] Department of Statistics, Abdul Wali Khan University, Mardan, Pakistan
[c] Department of Quantitative Mehtods, School of Business, King Faisal University, Al-Ahsa, 31982, Saudi Arabia

ARTICLE INFO

Keywords:
Ratio estimator
Exponential estimator
Transformed auxiliary variables
MSE
PRE

ABSTRACT

Estimation of population mean is a determined subject issue in sampling surveys and many efforts have been paid by various researchers to enhance the precision of the estimates by utilizing the correlated auxiliary information. In connection with this, we suggest an improved exponential ratio-cum-ratio estimator using transformed auxiliary variables under ranked set sampling scheme. Theoretical comparison between estimators is made in terms of mean square errors (*MSEs*), percentage relative efficiencies (*PREs*), and percentage relative root mean squared error (*PRRMSE*). The numerical expression for the bias and *MSE* of the suggested estimator is derived up to first order of approximation. Based on the results of actual data sets and a simulation study, it is found that the suggested estimator perform well as compared to its existing counterparts.

1. Introduction

Ranked set sampling (*RSS*) is the modification of simple random sampling for bringing better precision in estimation. In sample survey, major concern of the Researchers is cost-effective sampling, particularly, when the features of variable of interest is time consuming and costly. In view of [1–3], sampling methods based on ranking are designed to utilize extra information from easily obtained and inexpensive sources to collection best representative samples compared to other methods. The authors in Ref. [4] adapted the efficient ratio estimators in rank set sampling. The authors in Ref. [5] recommended an unbiased estimators for population mean using ranked set sampling. The authors in Refs. [6,7] recommended an efficient class of estimators using idea of ranked set sampling. In this paper the properties of the proposed estimator are investigated against usual *RSS* mean estimator in Refs. [8–10]. The authors in Ref. [11] recommended a family of estimators for population proportion using idea of rank. The [12] was the first who placed his consistent efforts to introduce this method for efficient estimating the average pasture yield in forest. The [13] suggested a modified estimator for population mean utilizing ranked set sampling. The [14,15] adapted the efficient ratio estimators in *RSS* scheme for mean and compared it to its counterpart estimators under simple random sampling and proved that *RSS* ratio estimators are more efficient than the later. The [16] was the first who used auxiliary information for the purpose of perfect ordering in *RSS*. The [17] catered the mathematical theory for the ranked set sampling for the first time. The [18] give idea of ratio and product type estimators,

* Corresponding author.
E-mail addresses: l.khan@stat.qau.edu.pk (L. Khan), sohaib_ahmad@awkum.edu.pk (S. Ahmad), ama.alomair@kfu.edu.sa (A.M. Alomair), ma.alomair@kfu.edu.sa (M.A. Alomair).

https://doi.org/10.1016/j.heliyon.2024.e27522
Received 5 November 2023; Received in revised form 23 February 2024; Accepted 1 March 2024
Available online 7 March 2024
2405-8440/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

when there is positive or negative correlation occur between the study and the auxiliary variable.

The [19] suggested a novel estimator for population mean using idea of rank under stratified random sampling. The [20] discussed a new family of estimators using rank idea for population distribution function. Some latest works done using rank set sampling see Refs. [21–26]. In this article we have suggested an enhanced exponential ratio-cum-ratio estimator in ranked set sampling using transformed auxiliary information. In terms of mean squared errors, our suggested estimator has minimum mean squared error and higher percentage relative efficiency as compared to existing estimators.

2. Ranked set sampling in case of two auxiliary variables and its algorithm for sample selection

Ranked set sampling is a procedure either it is the case of one auxiliary variable or two auxiliary variables. It the first stage, units are selected randomly through *SRS* and then it 2nd stage, the units are ranked by any suited concomitant variable. Finally, units are drawn from the selected units.

Under this case, it first stage, from a tri-variate finite population, draw independent m^2 random samples, each of volume *m*, through *SRS* without replacement along with equal probability. Then in each set, order the units either by correlated variable *X or Z,* where *X and Z* are concomitant variables, but the common practice is done according to the variable *X*. *RSS* procedure is then encounter to take the desired sample of size $n = rm$ for the actual measurement. From the first set *m,* draw the unit of *Y and Z* attached with the smallest unit of *X*. From the next sample, take the value of *Y and Z* attached with the second smallest unit of *X*. Following the same approach, the procedure is maintained until the units of *Y and Z* attached with the highest unit of *X* from the set *m.* This make up one cycle of the procedure. The procedure is replicate *r times* to achieve the desired sample size of *n-rm* units for analysis.

In bullet form, the procedure in case of two auxiliary variables may be elaborated as under:

a. Select randomly m^2 tri-variate sample units from the tri-variate finite population where *m* is set size that denotes the ranked number of samples in each set.
b. Allocate randomly m^2 chosen units into *m* sets having size *m.*
c. Within each set rank the units by either the auxiliary variable *X or Z* correlated with the variable of interest *Y,* but here is done according to variable *X.*
d. Now start choosing units of *Y and Z* as choose the smallest ordered unit *X* in the set 1, then 2nd smallest *X* in set 2 along-with the associated values of *Y and Z,* and so on until the largest ordered unit *X* is selected for the last set. This make up one *cycle.*
e. Replicate the above - mentioned steps for *r cycles* until the needed sample size i.e. $n = mr$ is achieved.

3. General procedure for obtaining bias and mean square error

To derive the bias and *MSE* of the suggested estimators, we proceed as:

$$\overline{y}_{[rss]} = \overline{Y}(1+e_0), \quad \overline{x}_{(rss)} = \overline{X}(1+e_1), \overline{z}_{(rss)} = \overline{Z}(1+e_2),$$

Such that $(e_p) = 0, \ (p = 0,1,2)$,
and

$$E\left(e_0^2\right) = \gamma C_y^2 - W_{(y)}^2 = V_{200} \quad E\left(e_1^2\right) = \gamma C_x^2 - W_{(x)}^2 = V_{020},$$

$$E\left(e_2^2\right) = \gamma C_z^2 - W_{(z)}^2 = V_{002} \quad E(e_0e_1) = \gamma C_{yx} - W_{(yx)} = V_{110},$$

$$E(e_0e_2) = \gamma C_{yz} - W_{(yz)} = V_{101} \quad E(e_1e_2) = \gamma C_{xz} - W_{(xz)} = V_{011},$$

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

$$W_{(z)}^2 = \frac{1}{m^2r\overline{Z}^2}\sum_{i=1}^{m}\tau_{z(i:m)}^2 \quad W_{(yx)} = \frac{1}{m^2r\overline{YX}}\sum_{i=1}^{m}\tau_{yx(i:m)} \quad W_{(yz)} = \frac{1}{m^2r\overline{ZY}}\sum_{i=1}^{m}\tau_{yz(i:m)} \quad W_{(xz)} = \frac{1}{m^2r\overline{XZ}}\sum_{i=1}^{m}\tau_{xz(i:m)} \quad \tau_{y(i:m)} = \left(\mu_{y(i:m)} - \overline{Y}\right)\tau_{x(i:m)} = \left(\mu_{x(i:m)} - \overline{X}\right),$$

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

$\tau_{yz(i:m)} = (\mu_{y(i:m)} - \overline{Y})(\mu_{z(i:m)} - \overline{Z})$, $\tau_{xz(i:m)} = (\mu_{z(i:m)} - \overline{Z})(\mu_{x(i:m)} - \overline{X})$, $\gamma = \left(\frac{1}{mr}\right)$, $C_{yx} = \rho C_y C_x$, $C_{yz} = \rho C_y C_z$ & $C_{xz} = \rho C_x C_z$, where C_x, C_y and C_z are the coefficients of variation of *X*, *Y* and *Z* respectively. The values of $\mu_{y(i:m)}$, $\mu_{z(i:m)}$ and $\mu_{x(i:m)}$ depend on order statistics from some specific distributions (See Ref. [3]).

4. Existing estimators

The usual per unit mean estimator with its *MSE* is given as under:

$$\bar{y}_{1(RSS)} = \frac{1}{mr}\sum_{j=1}^{r}\sum_{i=1}^{m} y_{[i:m]j},$$

The MSE of $\bar{y}_{1(RSS)}$ is given in equation (1):

$$MSE\left(\bar{y}_{1(RSS)}\right) = \bar{Y}^2\left(\gamma C_y^2 - W_{[y]}^2\right) = \bar{Y}^2 V_{200}. \tag{1}$$

Some of the members of [9] under *RSS* using two auxiliary variables as under:

$$\bar{y}_{2(RSS)} = \bar{y}_{[rss]}\left(\frac{\bar{X}}{\bar{x}_{(rss)}}\right)\exp\left(\frac{\bar{Z}-\bar{z}_{(rss)}}{\bar{Z}+\bar{z}_{(rss)}}\right),$$

The MSE of $\bar{y}_{2(RSS)}$ is given in equation (2):

$$MSE\left(\bar{y}_{2(RSS)}\right) = \bar{Y}^2\left[V_{200} + V_{020} + \frac{1}{4}V_{002} - 2V_{110} - V_{101} + V_{011}\right]. \tag{2}$$

$$\bar{y}_{3(RSS)} = \bar{y}_{[rss]}\left(\frac{\bar{Z}}{\bar{z}_{(rss)}}\right)\exp\left(\frac{\bar{X}-\bar{x}_{(rss)}}{\bar{X}+\bar{x}_{(rss)}}\right),$$

The MSE of $\bar{y}_{3(RSS)}$ is given in equation (3):

$$MSE\left(\bar{y}_{3(RSS)}\right) = \bar{Y}^2\left[V_{200} + \frac{1}{4}V_{020} + V_{002} - V_{110} - 2V_{101} + V_{011}\right]. \tag{3}$$

The authors in Ref. [10] also developed the following estimators under *RSS* in case of two auxiliary variables:

$$\bar{y}_{4(RSS)} = \bar{y}_{[rss]}\ \ exp\left(\frac{\bar{X}-\bar{x}_{(rss)}}{\bar{X}+\bar{x}_{(rss)}}\right)\exp\left(\frac{\bar{Z}-\bar{z}_{(rss)}}{\bar{Z}+\bar{z}_{(rss)}}\right),$$

The MSE of $\bar{y}_{4(RSS)}$ is given in equation (4):

$$MSE\left(\bar{y}_{4(RSS)}\right) = \bar{Y}^2\left[V_{200} + \frac{1}{4}V_{020} + \frac{1}{4}V_{002} - V_{110} - V_{101} + \frac{1}{2}V_{011}\right]. \tag{4}$$

$$\bar{y}_{5(RSS)} = \bar{y}_{[rss]}\ \ exp\left(\frac{\bar{X}-\bar{x}_{(rss)}}{\bar{X}}\right)\exp\left(\frac{\bar{Z}-\bar{z}_{(rss)}}{\bar{Z}+\bar{z}_{(rss)}}\right),$$

The MSE of $\bar{y}_{5(RSS)}$ is given in equation (5):

$$MSE\left(\bar{y}_{5(RSS)}\right) = \bar{Y}^2\left[V_{200} + V_{020} + \frac{1}{4}V_{002} - 2V_{110} - V_{101} + V_{011}\right]. \tag{5}$$

$$\bar{y}_{6(RSS)} = \bar{y}_{[rss]}\ \ exp\left(\frac{\bar{X}-\bar{x}_{(rss)}}{\bar{X}+\bar{x}_{(rss)}}\right) exp\left(\frac{\bar{Z}-\bar{z}_{(rss)}}{\bar{Z}}\right),$$

The MSE of $\bar{y}_{6(RSS)}$ is given in equation (6):

$$MSE\left(\bar{y}_{6(RSS)}\right) = \bar{Y}^2\left[V_{200} + \frac{1}{4}V_{020} + V_{002} - V_{110} - 2V_{101} + V_{011}\right]. \tag{6}$$

5. Suggested estimator

In this article we have suggested exponential ratio-cum-ratio estimator in ranked set sampling using transformed auxiliary information.

$$\text{Let } \bar{T}_{x(rss)} = \frac{N\bar{X} - n\bar{x}_{(rss)}}{N-n},$$

$$\bar{T}_{z(rss)} = \frac{N\bar{Z} - n\bar{z}_{(rss)}}{N-n},$$

$$\bar{T}_{x(rss)} = (1+g)\bar{X} - g\bar{x}_{(rss)},$$

$$\overline{T}_{z(rss)} = (1+g)\overline{Z} - g\overline{z}_{(rss)},$$

where $g = \frac{n}{N-n}$.

Using the transformation $\overline{T}_{x(rss)}$ and $\overline{T}_{z(rss)}$, the suggested exponential ratio-cum-ratio estimator in *RSS* is given in equation (7):

$$\overline{y}_{L(RSS)} = \overline{y}_{[rss]} \exp\left[a\left(\frac{\overline{T}_{x(rss)} - \overline{X}}{\overline{T}_{x(rss)} + \overline{X}}\right) + b\left(\frac{\overline{T}_{z(rss)} - \overline{Z}}{\overline{T}_{z(rss)} + \overline{Z}}\right)\right], \tag{7}$$

where

$$\overline{y}_{(rss]} = \frac{\sum_{j=1}^{r}\sum_{i=1}^{m} y_{[i:m]j}}{mr} \quad \overline{x}_{(rss)} = \frac{\sum_{j=1}^{r}\sum_{i=1}^{m} x_{[i:m]j}}{mr} \;\&\; \overline{z}_{(rss)} = \frac{\sum_{j=1}^{r}\sum_{i=1}^{m} z_{[i:m]j}}{mr}.$$

Here a and b are some suitable chosen scalars whose values are to be determined, so that *MSE* of $\overline{y}_{L(RSS)}$ is minimized. Also, $\overline{X}, \overline{Y}$ and $\overline{Z}$ are the population means of X, Y and Z respectively.

In error terms, we have the expression for equation (7), up to first order of approximation: $\overline{y}_{L(RSS)} = \overline{Y}(1+e_0)\{1 - \frac{1}{2}age_1 - \frac{1}{4}ag^2e_1^2 - \frac{1}{2}bge_2 - \frac{1}{4}bg^2e_2^2 + \frac{1}{8}a^2g^2e_1^2 + \frac{1}{8}b^2g^2e_2^2 + \frac{1}{4}abg^2e_1e_2\}$.

The simplified form is given in equation (8):

$$\left(\overline{y}_{L(RSS)} - \overline{Y}\right) = \overline{Y}\left[e_0 - \frac{1}{2}age_1 - \frac{1}{2}bge_2 - \frac{1}{4}ag^2e_1^2 - \frac{1}{4}bg^2e_2^2 + \frac{1}{8}a^2g^2e_1^2 + \frac{1}{8}b^2g^2e_2^2 + \frac{1}{4}abg^2e_1e_2 - \frac{1}{2}age_0e_1 - \frac{1}{2}bge_0e_2\right] \tag{8}$$

The bias of $\overline{y}_{L(RSS)}$, is given in equation (9):

$$Bias\left(\overline{y}_{L(RSS)}\right) = \overline{Y}\left[\frac{1}{8}a^2g^2V_{020} + \frac{1}{8}b^2g^2V_{002} - \frac{1}{4}ag^2V_{020} - \frac{1}{4}bg^2V_{002} + \frac{1}{4}abg^2V_{011} - \frac{1}{2}agV_{110} - \frac{1}{2}bgV_{101}\right] \tag{9}$$

Squaring and taking expectation of equation (8):

$$MSE\left(\overline{y}_{L(RSS)}\right) = \overline{Y}^2\left[V_{200} + \frac{1}{4}a^2g^2V_{020} + \frac{1}{4}b^2g^2V_{002} - agV_{110} - bgV_{101} + \frac{1}{2}abg^2V_{011}\right]. \tag{10}$$

To get the a and b optimal values, differentiate equation (10) with respect to a and b respectively and equate to zero i.e

$$a_{(opt)} = \frac{2(V_{002}V_{110} - V_{011}V_{101})}{g\left(V_{020}V_{002} - V_{011}^2\right)},$$

and

$$b_{(opt)} = \frac{2(V_{020}V_{101} - V_{011}V_{110})}{g\left(V_{020}V_{002} - V_{011}^2\right)}$$

Putting the optimum values of a and b in equation (10), we get the optimum $MSE(\overline{y}_{L(RSS)})$, which is given by:

$$MSE\left(\overline{y}_{L(RSS)}\right)_{(opt)} = \overline{Y}^2\left[V_{200} + \frac{\Omega_1 - \Omega_2 + \Omega_3 - \Omega_4 - \Omega_5 + \Omega_6}{\left(V_{020}V_{002} - V_{011}^2\right)^2}\right].$$

where

$$\Omega_1 = V_{002}V_{110}^2V_{011}^2 \quad \Omega_2 = V_{002}V_{020}^2V_{101}^2 \quad \Omega_3 = V_{020}V_{101}^2V_{011}^2 \quad \Omega_4 = V_{020}V_{002}^2V_{110}^2 \quad \Omega_5 = 2V_{110}V_{101}V_{011}^3 \text{ and } \quad \Omega_6 = 2V_{110}V_{101}V_{011}V_{020}V_{002}$$

6. Special cases of the proposed estimator

If $a = 0$ and $b = 0$ in equation (7), the proposed estimator reduced to per unit mean estimator of population mean under *RSS i.e:*

$$\overline{y}_{(rss)} = \frac{1}{mr}\sum_{j=1}^{r}\sum_{i=1}^{m} y_{[i:m]j},$$

The MSE of $\overline{y}_{(rss)}$ is given in equation (11):

$$MSE\left(\overline{y}_{(rss)}\right) = \overline{Y}^2V_{200}. \tag{11}$$

If $a = 1$ and $b = 1$ in equation (7), the following exponential ratio-cum ratio estimator may be obtained

$$\bar{y}_{L_1(RSS)} = \bar{y}_{[rss]} \ \exp\left[\left(\frac{\overline{T}_{x(rss)} - \overline{X}}{\overline{T}_{x(rss)} + \overline{X}}\right) + \left(\frac{\overline{T}_{z(rss)} - \overline{Z}}{\overline{T}_{z(rss)} + \overline{Z}}\right)\right],$$

The bias of $\bar{y}_{L_1(RSS)}$ is given by:

$$Bias\left(\bar{y}_{L_1(RSS)}\right) = \overline{Y}\left[-\frac{1}{4}g^2V_{020} - \frac{1}{4}g^2V_{002} - \frac{1}{2}gV_{110} - \frac{1}{2}gV_{101} + \frac{1}{4}g^2V_{011}\right],$$

The *MSE* of $\bar{y}_{L_1(RSS)}$ is given in equation (12):

$$MSE\left(\bar{y}_{L_1(RSS)}\right) = \overline{Y}^2\left[\begin{array}{c} V_{200} + \frac{1}{4}g^2V_{020} + \frac{1}{4}g^2V_{002} - gV_{110} - gV_{101} \\ +\frac{1}{2}g^2V_{011} \end{array}\right]. \tag{12}$$

If $a = 1$ and $b = 0$ in equation (7), the following exponential ratio estimator may be obtained

$$\bar{y}_{L_2(RSS)} = \bar{y}_{[rss]} \ \exp\left[\left(\frac{\overline{T}_{x(rss)} - \overline{X}}{\overline{T}_{x(rss)} + \overline{X}}\right)\right],$$

The bias of $\bar{y}_{L_2(RSS)}$ is given by:

$$Bias\left(\bar{y}_{L_2(RSS)}\right) = \overline{Y}\left[-\frac{1}{8}g^2V_{020} - \frac{1}{2}V_{110}\right]$$

The MSE of $\bar{y}_{L_2(RSS)}$ is given in equation (13) by:

$$MSE\left(\bar{y}_{L_2(RSS)}\right) = \overline{Y}^2\left[V_{200} + \frac{1}{4}g^2V_{020} - gV_{110}\right]. \tag{13}$$

If $a = 0$ and $b = 1$ in equation (7), the following exponential ratio estimator may be obtained

$$\bar{y}_{L_3(RSS)} = \bar{y}_{[rss]} \ \exp\left[\left(\frac{\overline{T}_{z(rss)} - \overline{Z}}{\overline{T}_{z(rss)} + \overline{Z}}\right)\right],$$

The bias of $\bar{y}_{L_3(RSS)}$ is given by:

$$Bias\left(\bar{y}_{L_3(RSS)}\right) = \overline{Y}\left[-\frac{1}{8}g^2V_{002} - \frac{1}{2}gV_{101}\right],$$

The *MSE* of $\bar{y}_{L_3(RSS)}$ is given in equation (14):

$$MSE\left(\bar{y}_{L_3(RSS)}\right) = \overline{Y}^2\left[V_{200} + \frac{1}{4}g^2V_{002} - gV_{101}\right]. \tag{14}$$

7. Theoretical conditions

In this section, we compare the proposed estimator with some existing estimators which are considered in this article, is given by.

1. By comparing equation (1) and equation (12),

$$MSE\left(\bar{y}_{L_1(RSS)}\right) < MSE\left(\bar{y}_{1(rss)}\right), \text{if}$$

$$\frac{g^2V_{020} + g^2V_{002}}{4gV_{110} + 4gV_{101} + 2g^2V_{011}} < 1$$

2. By comparing equation (2) and equation (12),

$$MSE\left(\bar{y}_{L_1(RSS)}\right) < MSE\left(\bar{y}_{2(RSS)}\right), \text{if}$$

$$\frac{(g^2-4)V_{020} + (g^2-1)V_{002}}{4(g-2)V_{110} + 4(g-1)V_{101} + 2(2-g^2)V_{011}} < 1$$

3. By comparing equation (3) and equation (12),

$$MSE\left(\overline{y}_{L_1(RSS)}\right) < MSE\left(\overline{y}_{3(RSS)}\right), \text{if}$$

$$\frac{(g^2-1)V_{020}+(g^2-4)V_{002}}{4(g-1)V_{110}+4(g-2)V_{101}+2(2-g^2)V_{011}} < 1$$

4. By comparing equation (4) and equation (12),

$$MSE\left(\overline{y}_{L_1(RSS)}\right) < MSE\left(\overline{y}_{4(RSS)}\right), \text{if}$$

$$\frac{(g+1)V_{020}+(g+1)V_{002}}{4V_{110}+4V_{101}-2(g+1)V_{011})} < 1$$

5. By comparing equation (5) and equation (12),

$$MSE\left(\overline{y}_{L_1(RSS)}\right) < MSE\left(\overline{y}_{5(RSS)}\right), \text{if}$$

$$\frac{(g^2-4)V_{020}+(g^2-1)V_{002}}{4(g-2)V_{110}+4(g-1)V_{101}+2(2-g^2)V_{011}} < 1$$

6. By comparing equation (6) and equation (12),

$$MSE\left(\overline{y}_{L_1(RSS)}\right) < MSE\left(\overline{y}_{6(RSS)}\right), \text{if}$$

$$\frac{(g^2-1)V_{020}+(g^2-4)V_{002}}{4(g-1)V_{110}+4(g-2)V_{101}+2(2-g^2)V_{011}} < 1$$

8. Simulation study

To check the efficiency of the suggested estimator, a simulation study is designed to estimate *MSEs*, *PREs*, *PRBs* and *PRRMSEs*. With respect to the correlated auxiliary variable X, rank is executed. Tri-variate random observations (X, Y, Z), were produced from a multivariate normal distribution with known population correlation coefficients $\rho_{yx} = 0.90$, $\rho_{yz} = 0.80$ and $\rho_{xz} = 0.70$. Using $10,000$ simulations, estimates of *MSEs*, *PREs*, *PRRMSEs* and *PRBs* are computed using *R-Language* and presented in Tables 1–4.

The following formulae are used for point estimations. The expression for $PRB(\overline{y}_{P(RSS)})$, $MSE(\overline{y}_{P(RSS)})$, $PRRMSE_{(P)}$ and $PR\,E_{(P)}$ are given in equation (15), equation (16) and equation (17):

$$PRB\left(\overline{y}_{P(RSS)}\right) = \frac{1}{\overline{Y}}\left[\frac{1}{10000}\sum_{i=1}^{10000}\left(\overline{y}_{P(RSS)i}-\overline{Y}\right)\right] \times 100, \tag{15}$$

Table 1
The simulated *MSEs* of different estimators.

m	r	n	$\overline{y}_{1(RSS)}$	$\overline{y}_{2(RSS)}$	$\overline{y}_{3(RSS)}$	$\overline{y}_{4(RSS)}$	$\overline{y}_{5(RSS)}$	$\overline{y}_{6(RSS)}$	$\overline{y}_{L(RSS)}$
3	3	9	0.1807	0.0832	0.1124	0.0449	0.0914	0.1350	**0.0393**
	4	12	0.1379	0.0605	0.0819	0.0330	0.0649	0.0924	**0.0295**
	5	15	0.1104	0.0487	0.0668	0.0259	0.0521	0.0748	**0.0228**
	10	30	0.0555	0.0240	0.0333	0.0130	0.0247	0.0354	**0.0114**
	15	45	0.0373	0.0162	0.0227	0.0088	0.0165	0.0235	**0.0077**
	20	60	0.0274	0.0122	0.0171	0.0066	0.0125	0.0177	**0.0058**
4	3	12	0.1203	0.0557	0.0797	0.0332	0.0590	0.0909	**0.0292**
	4	16	0.0903	0.0421	0.0593	0.0248	0.0439	0.0647	**0.0219**
	5	20	0.0701	0.0336	0.0489	0.0196	0.0349	0.0525	**0.0172**
	10	40	0.0353	0.0171	0.0249	0.0101	0.0175	0.0259	**0.0087**
	15	60	0.0239	0.0112	0.0163	0.0065	0.0113	0.0166	**0.0057**
	20	80	0.0180	0.0084	0.0122	0.0048	0.0085	0.0124	**0.0042**
5	3	15	0.0853	0.0429	0.0640	0.0268	0.0443	0.0696	**0.0233**
	4	20	0.0647	0.0313	0.0466	0.0195	0.0319	0.0498	**0.0172**
	5	25	0.0514	0.0254	0.0386	0.0156	0.0260	0.0408	**0.0134**
	10	50	0.0258	0.0123	0.0188	0.0076	0.0125	0.0194	**0.0067**
	15	75	0.0170	0.0081	0.0123	0.0051	0.0082	0.0126	**0.0044**
	20	100	0.0127	0.0062	0.0095	0.0038	0.0062	0.0096	**0.0033**

Table 2
The Simulated *PREs* of Different Estimators with respect to usual *RSS* mean estimator.

m	*r*	*n*	$\overline{y}_{1(RSS)}$	$\overline{y}_{2(RSS)}$	$\overline{y}_{3(RSS)}$	$\overline{y}_{4(RSS)}$	$\overline{y}_{5(RSS)}$	$\overline{y}_{6(RSS)}$	$\overline{y}_{L(RSS)}$
3	3	9	100	217	160	402	197	133	**459**
	4	12	100	227	168	416	212	149	**466**
	5	15	100	226	165	425	211	147	**483**
	10	30	100	230	166	425	224	156	**484**
	15	45	100	229	164	420	226	158	**481**
	20	60	100	223	159	412	219	154	**471**
4	3	12	100	216	150	361	203	132	**412**
	4	16	100	214	152	364	205	139	**410**
	5	20	100	208	143	356	200	133	**406**
	10	40	100	206	141	350	201	136	**406**
	15	60	100	212	146	367	210	143	**420**
	20	80	100	214	147	370	211	144	**425**
5	3	15	100	198	133	317	192	122	**365**
	4	20	100	206	138	330	202	129	**374**
	5	25	100	202	133	328	197	125	**381**
	10	50	100	209	137	336	206	132	**385**
	15	75	100	208	138	333	206	135	**380**
	20	100	100	204	134	330	203	131	**378**

Table 3
The simulated ***PRRMSEs*** of different estimators.

m	*r*	*n*	$\overline{y}_{1(RSS)}$	$\overline{y}_{2(RSS)}$	$\overline{y}_{3(RSS)}$	$\overline{y}_{4(RSS)}$	$\overline{y}_{5(RSS)}$	$\overline{y}_{6(RSS)}$	$\overline{y}_{L(RSS)}$
3	3	9	21.503	14.593	16.962	10.933	15.453	18.832	**9.037**
	4	12	18.787	12.444	14.482	9.208	12.972	15.496	**7.014**
	5	15	16.814	11.170	13.075	8.160	11.621	13.951	**6.028**
	10	30	11.919	7.849	9.246	6.782	7.991	9.576	**4.440**
	15	45	9.781	6.453	7.625	5.972	6.519	7.792	**4.070**
	20	60	8.377	5.603	6.629	4.626	5.668	6.747	**3.139**
4	3	12	17.551	11.94	14.285	9.236	12.391	15.403	**7.319**
	4	16	15.203	10.37	12.321	7.969	10.653	12.954	**6.862**
	5	20	13.395	9.282	11.190	7.190	9.487	11.641	**6.181**
	10	40	9.514	6.624	7.997	5.084	6.704	8.156	**4.737**
	15	60	7.834	5.375	6.467	4.089	5.408	6.542	**3.018**
	20	80	6.802	4.648	5.597	3.534	4.679	5.655	**2.303**
5	3	15	14.776	10.48	12.802	8.299	10.697	13.432	**7.762**
	4	20	12.868	8.953	10.927	7.088	9.089	11.355	**5.667**
	5	25	11.472	8.071	9.942	6.338	8.197	10.278	**4.882**
	10	50	8.138	5.625	6.952	5.435	5.676	7.076	**3.149**
	15	75	6.613	4.579	5.625	3.620	4.605	5.688	**2.392**
	20	100	5.709	3.996	4.931	3.140	4.011	4.977	**2.934**

Table 4
The simulated *PRBs* of different estimators.

m	*r*	*n*	$\overline{y}_{1(RSS)}$	$\overline{y}_{2(RSS)}$	$\overline{y}_{3(RSS)}$	$\overline{y}_{4(RSS)}$	$\overline{y}_{5(RSS)}$	$\overline{y}_{6(RSS)}$	$\overline{y}_{L(RSS)}$
3	3	9	0	2.1983	3.0396	0.5515	0.0261	0.2753	−0.8576
	4	12	0	1.4248	1.8640	0.3369	−0.2062	−0.1097	−0.4044
	5	15	0	1.2970	1.7598	0.3450	−0.0342	0.1337	−0.1356
	10	30	0	0.6725	0.9914	0.2649	0.0276	0.2011	−0.5135
	15	45	0	0.4879	0.7140	0.1700	0.0579	0.1912	−0.3642
	20	60	0	0.3789	0.4786	0.0888	0.0581	0.0872	−0.2824
4	3	12	0	1.5350	2.1025	0.4256	0.1779	0.2599	−0.1355
	4	16	0	1.0077	1.4225	0.2233	−0.0138	0.0949	−0.9594
	5	20	0	0.8045	1.0704	0.2317	0.0073	−0.0025	−0.6733
	10	40	0	0.3755	0.4896	0.0515	−0.0279	−0.0331	−0.4061
	15	60	0	0.3001	0.3356	0.0716	0.0309	−0.0131	−0.2204
	20	80	0	0.2153	0.2777	0.0365	0.0108	0.0152	−0.2032
5	3	15	0	1.0073	1.5071	0.3741	0.0822	0.1725	−0.7652
	4	20	0	0.8611	1.2349	0.3336	0.1687	0.2443	−0.5282
	5	25	0	0.7337	1.0110	0.3036	0.1747	0.2032	−0.3855
	10	50	0	0.3676	0.4971	0.1431	0.0905	0.0978	−0.2008
	15	75	0	0.2230	0.2954	0.0753	0.0404	0.0388	−0.1473
	20	100	0	0.2266	0.2764	0.0991	0.0902	0.0803	−0.0685

$$MSE\left(\overline{y}_{P(RSS)}\right)=\frac{1}{10000}\sum_{i=1}^{10000}\left(\overline{y}_{P(RSS)i}-\overline{Y}\right)^2, \tag{16}$$

$$PRRMSE_{(P)}=\frac{1}{\overline{Y}}\left[\frac{1}{10000}\sum_{i=1}^{10000}\left(\overline{y}_{P(SRSS)i}-\overline{Y}\right)^2\right]^{\frac{1}{2}}\times 100, \tag{17}$$

and

$$PR\,E_{(P)}=\frac{MSE\left(\overline{y}_{(RSS)}\right)}{MSE\left(\overline{y}_{P(RSS)}\right)}\times 100, \tag{18}$$

$$P=(1,2...6,L)$$

9. Real data set

In addition to simulation study, to observe the performance of the proposed estimator in real life application, a real data set is used. The data set contains three variables, whereas Y is variable of interest while X and Z are correlated auxiliary variables. From this real population set, number of samples were taken for comparing the efficiency of newly developed estimator with the existing counterpart estimators available in the literature stock in case of two auxiliary variables under *RSS*. Via *R-Language, MSEs* and *PREs* are obtained, using equation (16) and equation (18), for the proposed and exiting estimators, which is given in Table 5.

9.1. Population [source: [18]]

The summary statistics for the aforementioned population as follows:
y: Output for 80 factories in a region
x: Number of workers
z: Fixed capital

$$N=80 \quad m=4 \quad r=5$$

$$n=20 \quad \overline{Y}=285.125 \quad \overline{X}=1126.463$$

$$\overline{Z}=5182.637 \quad C_y=0.9484 \quad C_x=0.7507$$

$$C_z=0.3542 \quad \rho_{xy}=0.98 \quad \rho_{zy}=0.91$$

$$\rho_{xz}=0.94$$

10. Conclusion

From Tables 1 and 5 and it is shown that the suggested class of estimator $\overline{y}_{L(RSS)}$, have minimum *MSE* values as to $\overline{y}_{(rss)}$, [9,10]. Also, with increasing the sample size the values of *MSE* decreases. The proposed estimator has reasonable biases, since the values of *PRB* listed in Table 4 are all below 1 % in absolute form. Amongst the competitor estimators, as shown in Tables 2 and 5, $\overline{y}_{L(RSS)}$ has highest *PREs*. Similarly from Table 3 it is shown that the proposed estimator has minimum *PRRMSE* amongst counterpart estimators. It is noticed that the efficiency of the proposed estimator increases in real life survey. So, we reason out that, the developed estimator is strongly preferable over [6,7] estimators in real life application for better estimation of population mean. The present work can be extended to estimate population mean under stratified random sampling, probability proportional to size using predicative approach and non-response using ranked set sampling.

Data availability

Data will be made available on request.

CRediT authorship contribution statement

Lakhkar Khan: Software, Formal analysis. **Sohaib Ahmad:** Writing – original draft, Conceptualization. **Abdullah Mohammed Alomair:** Validation, Data curation. **Mohammed Ahmed Alomair:** Investigation, Formal analysis.

Table 5
MSEs and (*PREs*) of different estimators.

Sample	$\bar{y}_{1(RSS)}$	$\bar{y}_{2(RSS)}$	$\bar{y}_{3(RSS)}$	$\bar{y}_{4(RSS)}$	$\bar{y}_{5(RSS)}$	$\bar{y}_{6(RSS)}$	$\bar{y}_{L(RSS)}$
1	1732.91	105.74	301.57	114.76	449.90	302.65	**93.23**
	(100)	(1641)	(574)	(1508)	(385)	(573)	**(1846)**
2	1808.90	98.57	307.18	106.65	461.67	306.44	**88.41**
	(100)	(1835)	(588)	(1696)	(391)	(590)	**(2046)**
3	1575.82	108.11	281.65	121.05	417.99	283.45	**103.22**
	(100)	(1457)	(559)	(1301)	(377)	(557)	**(1526)**
4	1736.13	113.57	291.35	123.70	436.34	291.16	**99.23**
	(100)	(1528)	(596)	(1403)	(398)	(596)	**(1749)**
5	1784.78	109.25	308.13	122.23	462.64	308.96	**97.64**
	(100)	(1633)	(579)	(1460)	(387)	(578)	**(1825)**

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Deanship of Scientific Research, Vice Presidency for Graduate Studies and Scientific Research, King Faisal University, Saudi Arabia [Grant No. 5123]

References

[1] A.I. Al-Omari, Ratio estimation of the population mean using auxiliary information in simple random sampling and median ranked set sampling, Stat. Probab. Lett. 82 (11) (2012) 1883–1890.
[2] S. Amiri, M.J. Jozani, R. Modarres, Randomly selected order statistics in ranked set sampling: a less expensive comparable alternative to simple random sampling, Environ. Ecol. Stat. 25 (2) (2018) 237–256.
[3] B.C. Arnold, N. Balakrishnan, H.N. Nagaraja, A First Course in Order Statistics, vol. 54, Siam, 1992.
[4] C. Kadilar, Y. Unyazici, H. Cingi, Ratio estimator for the population mean using ranked set sampling, Stat. Pap. 50 (2) (2009) 301–309.
[5] L. Khan, J. Shabbir, A class of Hartley-Ross type unbiased estimators for population mean using ranked set sampling, Hacettepe Journal of Mathematics and Statistics 44 (3) (2015) 1–12.
[6] L. Khan, J. Shabbir, An efficient class of estimators for the finite population mean in ranked set sampling, Open J. Stat. 6 (3) (2016) 426–435.
[7] L. Khan, J. Shabbir, Hartley-Ross type unbiased estimators using ranked set sampling and stratified ranked set sampling, The North Carolina Journal of Mathematics and Statistics 2 (2016) 10–22.
[8] L. Khan, J. Shabbir, Improved ratio-type estimators of population mean in ranked set sampling using two concomitant variables, Pak. J. Statistics Oper. Res. 12 (3) (2016) 507–518.
[9] L. Khan, J. Shabbir, S. Gupta, Unbiased ratio estimators of mean in stratified ranked set sampling, Hacettepe Journal of Mathematics and Statistics 46 (6) (2016) 1151–1158.
[10] L. Khan, J. Shabbir, Generalized exponential-type ratio-cum ratio estimators of population mean in ranked set and stratified ranked set sampling, J. Stat. Manag. Syst. 20 (1) (2017) 133–151.
[11] S. Ahmad, E. Zahid, J. Shabbir, M. Aamir, W. Emam, Y. Tashkandy, S. Ahmad, A simulation study: an enhanced generalized class of estimators for estimation of population proportion using twofold auxiliary attribute, Heliyon 9 (2023), 2405-844.
[12] G. McIntyre, A method for unbiased selective sampling, using ranked sets, Aust. J. Agric. Res. 3 (4) (1952) 385–390.
[13] N. Mehta, V. Mandowara, A modified ratio-cum-product estimator of finite population mean using ranked set sampling, Commun. Stat. Theor. Methods 45 (2) (2016) 267–276.
[14] R. Murray, M. Ridout, J. Cross, The use of ranked set sampling in spray deposit assessment, Aspect Appl. Biol. 57 (2000) 141–146.
[15] H.M. Samawi, H.A. Muttlak, Estimation of ratio using rank set sampling, Biom. J. 38 (6) (1996) 753–764.
[16] S.L. Stokes, Ranked set sampling with concomitant variables, Commun. Stat. Theor. Methods 6 (12) (1977) 1207–1211.
[17] K. Takahasi, K. Wakimoto, On unbiased estimates of the population mean based on the sample stratified by means of ordering, Ann. Inst. Stat. Math. 20 (1) (1968) 1–31.
[18] M.N. Murthy, Sampling Theory and Methods, 1967. Sampling theory and methods.
[19] S. Ahmad, S. Hussain, U. Yasmeen, M. Aamir, J. Shabbir, M. El-Morshedy, Z. Ahmad, A simulation study: using dual ancillary variable to estimate population mean under stratified random sampling, PLoS One 17 (11) (2022) e0275875.
[20] S. Ahmad, K. Ullah, E. Zahid, J. Shabbir, M. Aamir, H.M. Alshanbari, A.A.A.H. El-Bagoury, A new improved generalized class of estimators for population distribution function using auxiliary variable under simple random sampling, Sci. Rep. 13 (1) (2023) 5415.
[21] S. Bhushan, A. Kumar, T. Zaman, A. Al Mutairi, Efficient difference and ratio-type imputation methods under ranked set sampling, Axioms 12 (6) (2023) 558.
[22] S. Bhushan, A. Kumar, Novel predictive estimators using ranked set sampling, Concurrency Comput. Pract. Ex. 35 (3) (2023) e7435.
[23] S. Bhushan, A. Kumar, Predictive estimation approach using difference and ratio type estimators in ranked set sampling, J. Comput. Appl. Math. 410 (2022) 114214.
[24] S. Bhushan, A. Kumar, S.A. Lone, On some novel classes of estimators using ranked set sampling, Alex. Eng. J. 61 (7) (2022) 5465–5474.
[25] A.M. Alomair, U. Shahzad, Optimizing mean estimators with calibrated minimum covariance determinant in median ranked set sampling, Symmetry 15 (8) (2023) 1581.
[26] A.M. Abbasi, M. Yousaf Shad, Estimation of population proportion using concomitant based ranked set sampling, Commun. Stat. Theor. Methods 51 (9) (2022) 2689–2709.