

Research Article

Generalized Chain Regression-cum-Chain Ratio Estimator for Population Mean under Stratified Extreme-cum-Median Ranked Set Sampling

Asad Ali , Muhammad Moeen Butt , and Muhammad Zubair 

Department of Economics and Statistics, University of Management and Technology, Lahore, Pakistan

Correspondence should be addressed to Asad Ali; ranaasadali23@gmail.com

Received 1 October 2021; Revised 8 December 2021; Accepted 21 December 2021; Published 4 January 2022

Academic Editor: Filipe J. Marques

Copyright © 2022 Asad Ali et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

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 ranked set sampling scheme results in efficient estimation of population mean. In this paper, we propose an efficient generalized chain regression-cum-chain ratio type estimator to estimate finite population mean of study variable under stratified extreme-cum-median ranked set sampling utilizing information on two auxiliary variables. Mean square error (MSE) of the proposed generalized estimator is derived up to first order of approximation. The applications of the proposed estimator under symmetrical and asymmetrical probability distributions are discussed using simulation study and real-life data set for comparisons of efficiency. It is concluded that the proposed generalized estimator performs efficiently as compared to some existing estimators. It is also observed that the efficiency of the proposed estimator is directly proportional to the correlations between the study variable and its auxiliary variables.

1. Introduction

Survey sampling is a process to collect information on the subject under study from population by choosing and analyzing true subset from it [1]. The national and international agencies regularly present estimates for different indicators like family income, retail prices, poverty, inflation, and wages of employees. Survey sampling has many advantages over complete study of population (census), such as consuming fewer resources, less time, and less cost. Survey sampling provides precise and efficient estimates for parameters of interest. These advantages of survey sampling are achieved by incorporating suitable sample designs and estimation techniques.

Neyman [2] introduced stratified random sampling (StRS) for efficient estimation of the population parameters in heterogeneous environment. Neyman [3] proposed the procedure of estimating population parameters by utilizing auxiliary information in stratified random sampling. The sampling procedure of StRS helps to minimize biasness in

sample selection and ensures that every section of population gets appropriate representation in sample. This sampling design provides greater precision, so high level of accuracy can be achieved even for small sample size. However, the procedure of StRS requires more administrative work as compared to simple random sampling (SRS). The procedure of StRS is tedious and time consuming.

McIntyre [4] proposed a method for estimating mean of pasture yield and later named this method as ranked set sampling (RSS). This ranked set sampling procedure is also known as classical ranked set sampling. Takahasi and Wakimoto [5] proved that RSS will be more efficient and easily applicable in real life as compared to SRS when ranking is perfect (i.e., ranking is done on the basis of study variable itself). Dell and Clutter [6] introduced the RSS procedure when ranking is not perfect (i.e., ranking is not done on the basis of study variable itself). They conclude that RSS provide unbiased estimates for parameters of interest. Stokes [7] used the auxiliary variable in RSS and found that amount of increase in precision depends on the correlation

between study and auxiliary variable. Samawi [8] introduced stratified ranked set sampling (StRSS) for obtaining unbiased and efficient estimates of population mean. The unbiased StRSS estimator for population mean of variable of interest Y and its variance are given as

$$\widehat{\mu}_{y, \text{StRSS}} = \sum_{h=1}^k W_h \frac{\sum_{j=1}^r \sum_{i=1}^{m_h} Y_{i(i)jh}}{rm_h},$$

$$\text{var}(\widehat{\mu}_{y, \text{StRSS}}) = \sum_{h=1}^k W_h^2 \frac{1}{rm_h} \left[\sigma_{(y)h}^2 - \frac{1}{m_h} \sum_{i=1}^{m_h} \varphi_{(i)h}^2 \right], \quad (1)$$

or

$$\text{var}(\widehat{\mu}_{y, \text{StRSS}}) = \text{var}(\widehat{\mu}_{y, \text{StRS}}) - \sum_{h=1}^k W_h^2 \frac{1}{rm_h} \left[\frac{1}{m_h} \sum_{i=1}^{m_h} \varphi_{(i)h}^2 \right],$$

where $\varphi_{(i)h}^2 = (\mu_{y(i)h} - \mu_{(y)h})^2$.

RSS is the best competitor of SRS and StRS due to many advantages. Some of the most important advantages are saving resources, cost, and time, while increasing efficiency and precision.

Survey statisticians aim to increase efficiency and precision of the estimators. For this purpose, statisticians move toward the use of auxiliary variables. Use of single or more auxiliary variables depends only on study variable and easy availability of data. An auxiliary variable in survey sampling is assumed to be easily and cheaply available and highly correlated with study variable. Graunt [9] utilizes the auxiliary variable to estimate the population of London. The author applied this method for estimating the proportion of burials per year in families and considered the average family size as an auxiliary variable. To obtain more precise and efficient estimators, selection of proper estimation technique is very important. Some well-known estimation techniques in literature are ratio estimation, product estimation, regression estimation, exponential estimation, and mixture of at least two of the aforementioned estimation techniques.

Samawi and Saeid [10] introduced stratified extreme ranked set sampling (StERSS) by combining StRSS and ERSS. They showed that estimates of population mean computed by StERSS will be more efficient than those by StRS and SRS. Ibrahim et al. [11] suggested stratified median ranked set sampling (StMRSS) for estimating population mean. They showed that under symmetrical distribution, StMRSS will provide unbiased estimates of population mean. Khan et al. [12] introduced stratified double ranked set sampling (StDRSS) for estimating population mean. Ali et al. [13] introduced stratified extreme-cum-median ranked set sampling (StEMRSS) for estimation of mean of heterogeneous populations in the presence of outliers. They showed that StEMRSS performs efficiently as compared to other StRSS schemes. Iqbal et al. [14] proposed mixture regression-cum-ratio type estimator for population mean under stratified random sampling. Ali et al. [15] suggested generalized family of estimators for estimating population mean under classical RSS.

Olkin [16] concluded that efficiency of estimators increases by using two or more auxiliary variables in the construction of estimators. Therefore, there is a need to utilize two auxiliary variables in the construction of generalized estimators under RSS design to increase their efficiency. Khan and Shabbir [17] utilized the aforementioned theory of two auxiliary variables and suggested generalized exponential-type ratio-cum-ratio estimator to estimate population mean of study variable under StRSS scheme. They compared the proposed estimator with some existing estimators with the help of relative bias (RB), relative mean square error (RMSE), and percentage relative efficiency (PRE). They concluded that the proposed estimator performs efficiently when study variable and auxiliary variables follow trivariate normal distribution.

1.1. General Notations, Symbols, and Relations. Let $\Omega = \{1, 2, \dots, N\}$ be a finite population of N units and ' N_h ' be used as population size from h^{th} stratum, where ' k ' is the number of strata and $h = 1, 2, 3, \dots, k$. Y is the variable under study, and X and Z are auxiliary variables which are highly correlated with study variable Y . The sample size is ' n ,' and ' n_h ' will be used as sample points from h^{th} stratum. The set size in ranked set sampling schemes is ' m ,' and ' m_h ' will be used as set size from h^{th} stratum, where $j = 1, 2, 3, \dots, m$. The number of cycles in ranked set sampling schemes is ' r ,' where $i = 1, 2, 3, \dots, r$. Overall sample size will be denoted as $n = rm = \sum_{h=1}^k n_h = \sum_{h=1}^k rm_h$. The correction factor is $\gamma = 1/n - 1/N = 1/rm - 1/N \cong 1/rm$; for large population size, we ignore $1/N$ in the equation of γ . $W_h = N_h/N$ is the stratum weight. Population means of Y , X , and Z variables are denoted by μ_y , μ_x , and μ_z , respectively. Sample means of Y , X , and Z , variables are denoted by \bar{y} , \bar{x} , and \bar{z} , respectively. Population variances of Y , X , and Z variables are denoted by σ_y^2 , σ_x^2 , and σ_z^2 , respectively. Sample variances of Y , X , and Z variables are denoted by s_y^2 , s_x^2 , and s_z^2 , respectively. Population coefficients of variation of Y , X , and Z variables are denoted by $C_y^2 = \sigma_y^2/\mu_y^2$, $C_x^2 = \sigma_x^2/\mu_x^2$, and $C_z^2 = \sigma_z^2/\mu_z^2$, respectively. Population covariances between X , Y , and Z variables are denoted by $S_{xy} = \sum_{i=1}^N (y_i - \mu_y)(x_i - \mu_x)/N$, $S_{yz} = \sum_{i=1}^N (y_i - \mu_y)(z_i - \mu_z)/N$, and $S_{xz} = \sum_{i=1}^N (x_i - \mu_x)(z_i - \mu_z)/N$, respectively. Population coefficients of correlation between X , Y , and Z variables are denoted by $\rho_{xy} = S_{xy}/\sigma_x\sigma_y$, $\rho_{xz} = S_{xz}/\sigma_x\sigma_z$, and $\rho_{yz} = S_{yz}/\sigma_y\sigma_z$, respectively.

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

- (i) $\bar{y}_{\text{StEMRSS}(j)} = \mu_y (1 + e_{0(j)(\text{stemrss})})$, $\bar{x}_{\text{StEMRSS}(j)} = \mu_x (1 + e_{1(j)(\text{stemrss})})$, and $\bar{z}_{\text{StEMRSS}(j)} = \mu_z (1 + e_{2(j)(\text{stemrss})})$, for $j = \text{Even (E) or Odd (O)}$.
- (ii) $e_{0(j)(\text{stemrss})} = \bar{y}_{\text{StEMRSS}(j)} - \mu_y/\mu_y$, $e_{1(j)(\text{stemrss})} = \bar{x}_{\text{StEMRSS}(j)} - \mu_x/\mu_x$, and $e_{2(j)(\text{stemrss})} = \bar{z}_{\text{StEMRSS}(j)} - \mu_z/\mu_z$.
- (iii) $E(e_{0(j)(\text{stemrss})}) = E(e_{1(j)(\text{stemrss})}) = E(e_{2(j)(\text{stemrss})}) = 0$.

$$\begin{aligned}
 \text{(iv)} \quad E(e_{0(O)}^2(\text{stemrss})) &= \sum_{h=1}^k W_h^2/2rm_h [C_{yh}^2 - 1/4m_h \\
 &\{ (m_h - 1)(W_{y(1)h}^2 + W_{y(2m_h)h}^2) + (m_h + 1)(W_{y(m_h)h}^2 \\
 &+ W_{y(m_h+1)h}^2) \}] = V_{yy(O)}^{(\text{stemrss})}, E(e_{1(O)}^2(\text{stemrss})) = \sum_{h=1}^k \\
 &W_h^2/2rm_h [C_{xh}^2 - 1/4m_h \{ (m_h - 1)(W_{x(1)h}^2 + W_{x(2m_h)h}^2) \\
 &+ (m_h + 1)(W_{x(m_h)h}^2 + W_{x(m_h+1)h}^2) \}] = \\
 &V_{xx(O)}^{(\text{stemrss})}, E(e_{2(O)}^2(\text{stemrss})) = \sum_{h=1}^k W_h^2/2rm_h [\\
 &C_{zh}^2 - 1/4m_h \{ (m_h - 1)(W_{z(1)h}^2 + W_{z(2m_h)h}^2) + (m_h \\
 &+ 1)(W_{z(m_h)h}^2 + W_{z(m_h+1)h}^2) \}] = V_{zz(O)}^{(\text{stemrss})}, E \\
 &(e_{0(O)}(\text{stemrss})e_{1(O)}(\text{stemrss})) = \sum_{h=1}^k W_h^2/2rm_h [\rho_{yxh} \\
 &C_{yh}C_{xh} - 1/4m_h \{ (m_h - 1)(W_{yx(1)h} + W_{yx(2m_h)h}) \\
 &+ (m_h + 1)(W_{yx(m_h)h} + W_{yx(m_h+1)h}) \}] = V_{yx(O)}^{(\text{stemrss})}, E \\
 &(e_{0(O)}(\text{stemrss})e_{2(O)}(\text{stemrss})) = \sum_{h=1}^k W_h^2/2rm_h [\rho_{yzh} \\
 &C_{yh}C_{zh} - 1/4m_h \{ (m_h - 1)(W_{yz(1)h} + W_{yz(2m_h)h}) \\
 &+ (m_h + 1)(W_{yz(m_h)h} + W_{yz(m_h+1)h}) \}] = V_{yz} \\
 &(O)^{(\text{stemrss})}, E(e_{1(O)}(\text{stemrss})e_{2(O)}(\text{stemrss})) = \sum_{h=1}^k W_h^2/ \\
 &2rm_h [\rho_{xzh}C_{xh}C_{zh} - 1/4m_h \{ (m_h - 1)(W_{xz(1)h} + W_{xz} \\
 &(2m_h)h) + (m_h + 1)(W_{xz(m_h)h} + W_{xz(m_h+1)h}) \}] = \\
 &V_{xz(O)}^{(\text{stemrss})}. \\
 \text{(v)} \quad E(e_{0(E)}^2(\text{stemrss})) &= \sum_{h=1}^k W_h^2/2rm_h [C_{yh}^2 - 1/4\{ (W_{y(1)h}^2 + W_{y(2m_h)h}^2) \\
 &+ (W_{y(m_h)h}^2 + W_{y(m_h+1)h}^2) \}] \\
 &= V_{yy(E)}^{(\text{stemrss})}, E(e_{1(E)}^2(\text{stemrss})) = \sum_{h=1}^k W_h^2/2rm_h [C_{xh}^2 -
 \end{aligned}$$

$$\begin{aligned}
 &1/4\{ (W_{x(1)h}^2 + W_{x(2m_h)h}^2) + (W_{x(m_h)h}^2 + W_{x(m_h+1)h}^2) \}] = V_{xx} \\
 &(E)^{(\text{stemrss})}, E(e_{2(E)}^2(\text{stemrss})) = \sum_{h=1}^k W_h^2/2rm_h [C_{zh}^2 - \\
 &1/4\{ (W_{z(1)h}^2 + W_{z(2m_h)h}^2) \\
 &+ (W_{z(m_h)h}^2 + W_{z(m_h+1)h}^2) \}] = V_{zz(E)}^{(\text{stemrss})}, E(e_{0(E)}(\\
 &(\text{stemrss})e_{1(E)}(\text{stemrss})) = \sum_{h=1}^k W_h^2/2rm_h [\rho_{yxh} \\
 &C_{yh}C_{xh} - 1/4\{ (W_{yx(1)h} + W_{yx(2m_h)h}) + (W_{yx(m_h)h} + \\
 &W_{yx(m_h+1)h}) \}] = V_{xy(E)}^{(\text{stemrss})}, E(e_{0(E)}(\text{stemrss})e_{2(E)} \\
 &(\text{stemrss})) = \sum_{h=1}^k W_h^2/2rm_h [\rho_{yzh}C_{yh}C_{zh} - 1/4 \\
 &\{ (W_{yz(1)h} + W_{yz(2m_h)h}) + (W_{yz(m_h)h} + W_{yz(m_h+1)h}) \}] \\
 &= V_{yz(E)}^{(\text{stemrss})}, E(e_{1(E)}(\text{stemrss})e_{2(E)}(\text{stemrss})) = \\
 &\sum_{h=1}^k W_h^2/2rm_h [\rho_{xzh}C_{xh}C_{zh} - 1/4\{ (W_{xz(1)h} + W_{xz} \\
 &(2m_h)h) + (W_{xz(m_h)h} + W_{xz(m_h+1)h}) \}] = V_{xz(E)}^{(\text{stemrss})}.
 \end{aligned}$$

2. Existing Estimators under StRSS with Two Auxiliary Variables

Khan and Shabbir [17] suggested generalized exponential-type ratio-cum-ratio estimator to estimate population mean of study variable with two auxiliary variables under StRSS scheme. The mathematical expression of estimator under StRSS is given as

$$\hat{\mu}_{yKS,StRSS} = \sum_{h=1}^k W_h \bar{y}_{(StRSS)h} \exp \left[\frac{\sum_{h=1}^k W_h (\mu_{xh} - \bar{x}_{(StRSS)h})}{\sum_{h=1}^k W_h (\mu_{xh} + (\alpha_{12} - 1) \bar{x}_{(StRSS)h})} \right] \exp \left[\frac{\sum_{h=1}^k W_h (\mu_{zh} - \bar{z}_{(StRSS)h})}{\sum_{h=1}^k W_h (\mu_{zh} + (\alpha_{13} - 1) \bar{z}_{(StRSS)h})} \right], \tag{2}$$

where α_{12} and α_{13} were suitably chosen constants and their optimum values were given as

$$\begin{aligned}
 \alpha_{12opt} &= \frac{V_{xx}^{(stRSS)}V_{zz}^{(stRSS)} - V_{xz}^2(stRSS)}{V_{zz}^{(stRSS)}V_{xy}^{(stRSS)} - V_{xz}^{(stRSS)}V_{yz}^{(stRSS)}}, \\
 \alpha_{13opt} &= \frac{V_{xx}^{(stRSS)}V_{zz}^{(stRSS)} - V_{xz}^2(stRSS)}{V_{xx}^{(stRSS)}V_{yz}^{(stRSS)} - V_{xz}^{(stRSS)}V_{xy}^{(stRSS)}}.
 \end{aligned} \tag{3}$$

The mathematical expression of MSE was given as

$$\text{MSE}(\hat{\mu}_{yKS,StRSS})_{\min} = \mu_y^2 \left(\begin{aligned} &V_{yy}^{(stRSS)} + \frac{1}{\alpha_{12opt}^2} V_{xx}^{(stRSS)} + \frac{1}{\alpha_{13opt}^2} V_{zz}^{(stRSS)} \\ & - \frac{2}{\alpha_{12opt}} V_{xy}^{(stRSS)} - \frac{2}{\alpha_{13opt}} V_{yz}^{(stRSS)} + \frac{2}{\alpha_{12opt}\alpha_{13opt}} V_{xz}^{(stRSS)} \end{aligned} \right). \tag{4}$$

3. The Proposed Estimator

Motivated by Zubair and Ali [18], we have proposed a class of generalized chain regression-cum-chain ratio estimator

for population mean using two auxiliary variables under new modified ranked set sampling scheme called stratified extreme-cum-median ranked set sampling (StEMRSS). The proposed estimator is given as

$$\widehat{\mu}_{y,(StEMRSS),G3} = \left[\overline{Y}_{(StEMRSS)} + \widehat{\beta}_{xy}(\mu_x - \overline{x}_{(StEMRSS)}) + \widehat{\beta}_{yz}(\mu_z - \overline{z}_{(StEMRSS)}) \right] \left(\frac{\mu_x}{\overline{x}_{(StEMRSS)}} \right)^{\alpha_{19}} \left(\frac{\mu_z}{\overline{z}_{(StEMRSS)}} \right)^{\alpha_{20}}, \quad (5)$$

where $\widehat{\beta}_{xy}$ and $\widehat{\beta}_{yz}$ are estimates of coefficients of regression, and α_{19} and α_{20} are any suitable chosen constants to minimize MSE of estimator. For derivation of MSE, the

proposed estimator can be written in the form of $e_{0(i)}$, $e_{1(i)}$, and $e_{2(i)}$ as

$$\widehat{\mu}_{y,(StEMRSS),G3} = \left[\mu_y(1 + e_0) + \widehat{\beta}_{xy}(\mu_x - \mu_x(1 + e_1)) + \widehat{\beta}_{yz}(\mu_z - \mu_z(1 + e_2)) \right] (\mu_x/\mu_x(1 + e_1))^{\alpha_{19}} (\mu_z/\mu_z(1 + e_2))^{\alpha_{20}},$$

or

$$\widehat{\mu}_{y,(StEMRSS),G3} = \left[\mu_y + \mu_y e_0 - \widehat{\beta}_{xy} \mu_x e_1 - \widehat{\beta}_{yz} \mu_z e_2 \right] (\mu_x + \mu_x e_1/\mu_x)^{-\alpha_{19}} (\mu_z + \mu_z e_2/\mu_z)^{-\alpha_{20}}, \quad (6)$$

or

$$\widehat{\mu}_{y,(StEMRSS),G3} = \left[\mu_y + \mu_y e_0 - \widehat{\beta}_{xy} \mu_x e_1 - \widehat{\beta}_{yz} \mu_z e_2 \right] (1 + e_1)^{-\alpha_{19}} (1 + e_2)^{-\alpha_{20}}.$$

Using Taylor series expansion and exponential series expansion in (6) and ignoring higher order terms, we get

$$\widehat{\mu}_{y,(StEMRSS),G3} = \left[\mu_y + \mu_y e_0 - \widehat{\beta}_{xy} \mu_x e_1 - \widehat{\beta}_{yz} \mu_z e_2 \right] (1 - \alpha_{19} e_1 + \alpha_{19}^2 e_1^2) (1 - \alpha_{20} e_2 + \alpha_{20}^2 e_2^2),$$

or

$$\widehat{\mu}_{y,(StEMRSS),G3} = \left(\begin{array}{c} \mu_y + \mu_y e_0 - \alpha_{19} \mu_y e_1 - \alpha_{20} \mu_y e_2 - \widehat{\beta}_{xy} \mu_x e_1 - \widehat{\beta}_{yz} \mu_z e_2 \\ -\alpha_{19} \mu_y e_0 e_1 - \alpha_{20} \mu_y e_0 e_2 + \alpha_{19} \alpha_{20} \mu_y e_1 e_2 + \alpha_{19} \widehat{\beta}_{yz} \mu_z e_1 e_2 + \alpha_{20} \widehat{\beta}_{xy} \mu_x e_1 e_2 \\ + \alpha_{19} \widehat{\beta}_{xy} \mu_x e_1^2 + \frac{\alpha_{19}(\alpha_{19} + 1)}{2} \mu_y e_1^2 + \alpha_{20} \widehat{\beta}_{yz} \mu_z e_2^2 + \frac{\alpha_{20}(\alpha_{20} + 1)}{2} \mu_y e_2^2 \end{array} \right). \quad (7)$$

Subtracting μ_y from both sides of (7), we get

$$\widehat{\mu}_{y,(StEMRSS),G3} - \mu_y = \left(\begin{array}{c} \mu_y e_0 - \alpha_{19} \mu_y e_1 - \alpha_{20} \mu_y e_2 - \widehat{\beta}_{xy} \mu_x e_1 - \widehat{\beta}_{yz} \mu_z e_2 \\ -\alpha_{19} \mu_y e_0 e_1 - \alpha_{20} \mu_y e_0 e_2 + \alpha_{19} \alpha_{20} \mu_y e_1 e_2 + \alpha_{19} \widehat{\beta}_{yz} \mu_z e_1 e_2 + \alpha_{20} \widehat{\beta}_{xy} \mu_x e_1 e_2 \\ + \alpha_{19} \widehat{\beta}_{xy} \mu_x e_1^2 + \frac{\alpha_{19}(\alpha_{19} + 1)}{2} \mu_y e_1^2 + \alpha_{20} \widehat{\beta}_{yz} \mu_z e_2^2 + \frac{\alpha_{20}(\alpha_{20} + 1)}{2} \mu_y e_2^2 \end{array} \right). \quad (8)$$

Squaring both sides of (8), we get

$$(\widehat{\mu}_{y,(StEMRSS),G3} - \mu_y)^2 = \begin{bmatrix} \mu_y^2 e_0^2 + (\alpha_{19}^2 \mu_y^2 + \widehat{\beta}_{xy}^2 \mu_x^2 + 2\alpha_{19} \widehat{\beta}_{xy} \mu_y \mu_x) e_1^2 \\ + (\alpha_{20}^2 \mu_y^2 + \widehat{\beta}_{yz}^2 \mu_z^2 + 2\alpha_{20} \widehat{\beta}_{yz} \mu_y \mu_z) e_2^2 \\ - (2\alpha_{19} \mu_y^2 + 2\widehat{\beta}_{xy} \mu_y \mu_x) e_0 e_1 - (2\alpha_{20} \mu_y^2 + 2\widehat{\beta}_{yz} \mu_y \mu_z) e_0 e_2 \\ + (2\alpha_{19} \alpha_{20} \mu_y^2 + 2\alpha_{19} \widehat{\beta}_{yz} \mu_y \mu_z + 2\alpha_{20} \widehat{\beta}_{xy} \mu_y \mu_x + 2\widehat{\beta}_{xy} \widehat{\beta}_{yz} \mu_x \mu_z) e_1 e_2 \end{bmatrix}. \tag{9}$$

Applying expectation to both sides of (9), we get

$$MSE(\widehat{\mu}_{y,(StEMRSS),G3}) = E(\widehat{\mu}_{y,(StEMRSS),G3} - \mu_y)^2,$$

or

$$MSE(\widehat{\mu}_{y,(StEMRSS),G3}) = \begin{bmatrix} V_{yy}^{(stemrss)} + (\alpha_{19}^2 + \widehat{\beta}_{xy}^2 R_1^2 + 2\alpha_{19} \widehat{\beta}_{xy} R_1) V_{xx}^{(stemrss)} \\ + (\alpha_{20}^2 + \widehat{\beta}_{yz}^2 R_2^2 + 2\alpha_{20} \widehat{\beta}_{yz} R_2) V_{zz}^{(stemrss)} \\ - (2\alpha_{19} + 2\widehat{\beta}_{xy} R_1) V_{xy}^{(stemrss)} - (2\alpha_{20} + 2\widehat{\beta}_{yz} R_2) V_{yz}^{(stemrss)} \\ + (2\alpha_{19} \alpha_{20} + 2\alpha_{19} \widehat{\beta}_{yz} R_2 + 2\alpha_{20} \widehat{\beta}_{xy} R_1 + 2\widehat{\beta}_{xy} \widehat{\beta}_{yz} R_1 R_2) V_{xz}^{(stemrss)} \end{bmatrix}, \tag{10}$$

or

$$MSE(\widehat{\mu}_{y,(StEMRSS),G3}) = \begin{bmatrix} V_{yy}^{(stemrss)} + (\alpha_{19} + \widehat{\beta}_{xy} R_1)^2 V_{xx}^{(stemrss)} + (\alpha_{20} + \widehat{\beta}_{yz} R_2)^2 V_{zz}^{(stemrss)} \\ - 2(\alpha_{19} + \widehat{\beta}_{xy} R_1) V_{xy}^{(stemrss)} - 2(\alpha_{20} + \widehat{\beta}_{yz} R_2) V_{yz}^{(stemrss)} \\ + 2(\alpha_{19} + \widehat{\beta}_{xy} R_1)(\alpha_{20} + \widehat{\beta}_{yz} R_2) V_{xz}^{(stemrss)} \end{bmatrix},$$

where $R_1 = \mu_x/\mu_y$ and $R_2 = \mu_z/\mu_y$.

To minimize MSE of proposed estimator-I, optimum values of α_{19} and α_{20} have been derived by taking partial

derivative of (10) and equating to zero. First, we take partial derivative of (10) with respect to α_{19} as follows:

$$\frac{\partial MSE(\widehat{\mu}_{y,(StEMRSS),G3})}{\partial \alpha_{19}} = \frac{\partial}{\partial \alpha_{19}} \begin{bmatrix} V_{yy}^{(stemrss)} + (\alpha_{19}^2 + \widehat{\beta}_{xy}^2 R_1^2 + 2\alpha_{19} \widehat{\beta}_{xy} R_1) V_{xx}^{(stemrss)} \\ + (\alpha_{20}^2 + \widehat{\beta}_{yz}^2 R_2^2 + 2\alpha_{20} \widehat{\beta}_{yz} R_2) V_{zz}^{(stemrss)} \\ - (2\alpha_{19} + 2\widehat{\beta}_{xy} R_1) V_{xy}^{(stemrss)} - (2\alpha_{20} + 2\widehat{\beta}_{yz} R_2) V_{yz}^{(stemrss)} \\ + (2\alpha_{19} \alpha_{20} + 2\alpha_{19} \widehat{\beta}_{yz} R_2 + 2\alpha_{20} \widehat{\beta}_{xy} R_1 + 2\widehat{\beta}_{xy} \widehat{\beta}_{yz} R_1 R_2) V_{xz}^{(stemrss)} \end{bmatrix} = 0, \tag{11}$$

or

$$\alpha_{19} = \frac{-\alpha_{20} V_{xz}^{(stemrss)} + V_{xy}^{(stemrss)} - \widehat{\beta}_{xy} R_1 V_{xx}^{(stemrss)} - \widehat{\beta}_{yz} R_2 V_{zz}^{(stemrss)}}{V_{xx}^{(stemrss)}}.$$

On the same line, we take partial derivative of (10) with respect to α_{20} as follows:

$$\frac{\partial \text{MSE}(\widehat{\mu}_{y,(\text{StEMRSS}),G3})}{\partial \alpha_{20}} = \frac{\partial}{\partial \alpha_{20}} \left[\begin{array}{c} V_{yy}^{(\text{stemrSS})} + \left(\alpha_{19}^2 + \widehat{\beta}_{xy}^2 R_1^2 + 2\alpha_{19}\widehat{\beta}_{xy}R_1 \right) V_{xx}^{(\text{stemrSS})} \\ + \left(\alpha_{20}^2 + \widehat{\beta}_{yz}^2 R_2^2 + 2\alpha_{20}\widehat{\beta}_{yz}R_2 \right) V_{zz}^{(\text{stemrSS})} \\ - \left(2\alpha_{19} + 2\widehat{\beta}_{xy}R_1 \right) V_{xy}^{(\text{stemrSS})} - \left(2\alpha_{20} + 2\widehat{\beta}_{yz}R_2 \right) V_{yz}^{(\text{stemrSS})} \\ + \left(2\alpha_{19}\alpha_{20} + 2\alpha_{19}\widehat{\beta}_{yz}R_2 + 2\alpha_{20}\widehat{\beta}_{xy}R_1 + 2\widehat{\beta}_{xy}\widehat{\beta}_{yz}R_1R_2 \right) V_{xz}^{(\text{stemrSS})} \end{array} \right] = 0, \quad (12)$$

or

$$\alpha_{20} = \frac{-\alpha_{19}V_{xz}^{(\text{stemrSS})} + V_{yz}^{(\text{stemrSS})} - \widehat{\beta}_{yz}R_2V_{zz}^{(\text{stemrSS})} - \widehat{\beta}_{xy}R_1V_{xx}^{(\text{stemrSS})}}{V_{zz}^{(\text{stemrSS})}}.$$

Now, we put value of α_{20} in the value of α_{19} as

$$\alpha_{19} = -\alpha_{20} \frac{V_{xz}^{(\text{stemrSS})}}{V_{xx}^{(\text{stemrSS})}} + \frac{V_{xy}^{(\text{stemrSS})} - \widehat{\beta}_{xy}R_1V_{xx}^{(\text{stemrSS})} - \widehat{\beta}_{yz}R_2V_{xz}^{(\text{stemrSS})}}{V_{xx}^{(\text{stemrSS})}},$$

or

$$\alpha_{19opt} = \frac{\left(V_{xy}^{(\text{stemrSS})}V_{zz}^{(\text{stemrSS})} - V_{yz}^{(\text{stemrSS})}V_{xz}^{(\text{stemrSS})} + \widehat{\beta}_{xy}R_1V_{xz}^{(\text{stemrSS})}V_{xx}^{(\text{stemrSS})} - \widehat{\beta}_{xy}R_1V_{xx}^{(\text{stemrSS})}V_{zz}^{(\text{stemrSS})} \right)}{\left(V_{xx}^{(\text{stemrSS})}V_{zz}^{(\text{stemrSS})} - V_{xz}^{(\text{stemrSS})}V_{xz}^{(\text{stemrSS})} \right)}. \quad (13)$$

On the same line, we get

$$\alpha_{20opt} = \frac{\left(V_{yz}^{(\text{stemrSS})}V_{xx}^{(\text{stemrSS})} - V_{xy}^{(\text{stemrSS})}V_{xz}^{(\text{stemrSS})} - \widehat{\beta}_{yz}R_2V_{xx}^{(\text{stemrSS})}V_{zz}^{(\text{stemrSS})} + \widehat{\beta}_{yz}R_2V_{xz}^{(\text{stemrSS})}V_{xz}^{(\text{stemrSS})} \right)}{\left(V_{xx}^{(\text{stemrSS})}V_{zz}^{(\text{stemrSS})} - V_{xz}^{(\text{stemrSS})}V_{xz}^{(\text{stemrSS})} \right)}. \quad (14)$$

Substituting values of α_{19} and α_{20} in (9), we get

$$\text{MSE}(\widehat{\mu}_{y,(\text{StEMRSS}),G3})_{\min} = \left[\begin{array}{c} V_{yy}^{(\text{stemrSS})} + \left(\alpha_{19opt} + \widehat{\beta}_{xy}R_1 \right)^2 V_{xx}^{(\text{stemrSS})} + \left(\alpha_{20opt} + \widehat{\beta}_{yz}R_2 \right)^2 V_{zz}^{(\text{stemrSS})} \\ - 2\left(\alpha_{19opt} + \widehat{\beta}_{xy}R_1 \right) V_{xy}^{(\text{stemrSS})} - 2\left(\alpha_{20opt} + \widehat{\beta}_{yz}R_2 \right) V_{yz}^{(\text{stemrSS})} \\ + 2\left(\alpha_{19opt} + \widehat{\beta}_{xy}R_1 \right) \left(\alpha_{20opt} + \widehat{\beta}_{yz}R_2 \right) V_{xz}^{(\text{stemrSS})} \end{array} \right]. \quad (15)$$

The mathematical expression of minimum MSE in (15) will be used when population parameters are exactly known. However, this is very difficult in real-life situations.

Therefore, we present the mathematical form of estimated minimum MSE for real-life situations as follows:

$$MSE(\widehat{\mu}_{y,(StEMRSS,G3)})_{\min} = \begin{bmatrix} \widehat{V}_{yy}^{(stemrss)} + (\widehat{\alpha}_{19opt} + \widehat{\beta}_{xy}\widehat{R}_1)^2 \widehat{V}_{xx}^{(stemrss)} + (\widehat{\alpha}_{20opt} + \widehat{\beta}_{yz}\widehat{R}_2)^2 \widehat{V}_{zz}^{(stemrss)} \\ -2(\widehat{\alpha}_{19opt} + \widehat{\beta}_{xy}\widehat{R}_1) \widehat{V}_{xy}^{(stemrss)} - 2(\widehat{\alpha}_{20opt} + \widehat{\beta}_{yz}\widehat{R}_2) \widehat{V}_{yz}^{(stemrss)} \\ +2(\widehat{\alpha}_{19opt} + \widehat{\beta}_{xy}\widehat{R}_1)(\widehat{\alpha}_{20opt} + \widehat{\beta}_{yz}\widehat{R}_2) \widehat{V}_{xz}^{(stemrss)} \end{bmatrix}. \quad (16)$$

The mathematical expression of minimum MSE in (16) only depends on sample observations of study and auxiliary variables. Therefore, it is recommended that (16) is used for estimating minimum MSE of proposed estimator in real-life situations.

3.1. Special Cases of Proposed Estimator. In this section, we discuss some special cases of proposed estimator by putting different values of constants. In Table 1, some special cases of proposed estimator have been presented, but one may present more special cases by using other combinations of constants as well.

4. Simulation Study of Proposed Estimator

In this section, we conduct a Monte Carlo simulation for the efficiency comparison of proposed estimator. We compare the performance of proposed estimator with some existing estimators under RSS sampling schemes with stratification using two auxiliary variables. Percent relative efficiency (PRE) is used as performance criterion for estimators. In Monte Carlo simulation study, hypothetical data have been generated from any probability distribution by specifying values of its parameters. Mean square errors (MSEs) of desired estimator have been computed by using hypothetical data, and the results have been iterated by any desired number. In this study, Monte Carlo simulation has been carried out through the following steps:

- (i) We generate a hypothetical population of variables X and Z (auxiliary variables) of size 1000 from symmetric distributions (normal and uniform) and asymmetric distributions (Gamma and Weibull) with some specific values of parameters as described in Table 2.
- (ii) Study variable Y is computed using the following regression model:

$$Y = \rho_{xy}X + \rho_{yz}Z + e, \quad (17)$$

where ρ_{xy} and ρ_{yz} are coefficients of correlations and 'e' is the normally distributed error term having mean zero and variance one.

- (iii) The number of iterations is one million.

- (iv) The performance of estimators has been computed by taking different number of cycles r , set size m , and number of strata h .
- (v) Percent relative efficiencies (PREs) of estimators for symmetric distributions have been calculated by using the following equation:

$$PRE = \frac{\text{var}(\widehat{\mu}_{y,SRS})}{\text{var}(\widehat{\mu}_{y,i})} \times 100, \quad (18)$$

where $\text{var}(\widehat{\mu}_{y,SRS}) = \sigma_y^2/n$ and 'i' stand for any estimator or sampling scheme whose performance has to be compared. For asymmetric distributions, the equation for percent relative efficiency is

$$PRE = \frac{\text{var}(\widehat{\mu}_{y,SRS})}{MSE(\widehat{\mu}_{y,i})} \times 100, \quad (19)$$

where $MSE(\widehat{\mu}_{y,i})$ is the mean square error of estimators with 'i' standing for any estimator or sampling scheme whose performance has to be compared.

Tables 3–8 show percent relative efficiencies (PREs) of proposed estimator and existing estimators under stratified ranked set sampling scheme with respect to SRS. As we use auxiliary variable (X) for ranking purpose, there is chance of ranking error. To minimize the effect of ranking error, we should utilize strong positive correlations (ρ_{xy}) between study and auxiliary variable. Therefore, we calculate PREs of proposed and existing estimators for moderate and strong positive calculations to monitor their effect. In Table 3, hypothetical data of study and auxiliary variables with $r = 3$, correlation coefficient of 0.5, and $m = 10$ and 15 are generated from normal, uniform, Gamma, and Weibull distributions. Results show that under $m = 10$ and 15, proposed estimator under StEMRSS has the highest PREs (bold values) for all choices of probability distributions. Under normal distribution, PREs of proposed estimator for $m = 10$ and 15 are 324.718 and 352.862, respectively. Under uniform distribution, PREs of proposed estimator for $m = 10$ and 15 are 311.495 and 341.206, respectively. Under Gamma distribution, PREs of proposed estimator for $m = 10$ and 15 are 315.748 and 334.984, respectively. Under Weibull distribution, PREs of proposed estimator for $m = 10$ and 15 are 310.399 and 354.558, respectively. Results also show that

TABLE 1: Some special cases of proposed estimator.

α_{19}	α_{20}	Estimator
0	0	$\widehat{\mu}_{y,(StEMRSS),s1} = [\overline{Y}_{(StEMRSS)} + \widehat{\beta}_{xy}(\mu_x - \overline{x}_{(StEMRSS)}) + \widehat{\beta}_{yz}(\mu_z - \overline{z}_{(StEMRSS)})]$
1	0	$\widehat{\mu}_{y,(StEMRSS),s2} = [\overline{Y}_{(StEMRSS)} + \widehat{\beta}_{xy}(\mu_x - \overline{x}_{(StEMRSS)}) + \widehat{\beta}_{yz}(\mu_z - \overline{z}_{(StEMRSS)})](\mu_x/\overline{x}_{(StEMRSS)})$
0	1	$\widehat{\mu}_{y,(StEMRSS),s3} = [\overline{Y}_{(StEMRSS)} + \widehat{\beta}_{xy}(\mu_x - \overline{x}_{(StEMRSS)}) + \widehat{\beta}_{yz}(\mu_z - \overline{z}_{(StEMRSS)})](\mu_z/\overline{z}_{(StEMRSS)})$
1	1	$\widehat{\mu}_{y,(StEMRSS),G4} = [\overline{Y}_{(StEMRSS)} + \widehat{\beta}_{xy}(\mu_x - \overline{x}_{(StEMRSS)}) + \widehat{\beta}_{yz}(\mu_z - \overline{z}_{(StEMRSS)})](\mu_x/\overline{x}_{(StEMRSS)})(\mu_z/\overline{z}_{(StEMRSS)})$
-1	0	$\widehat{\mu}_{y,(StEMRSS),s5} = [\overline{Y}_{(StEMRSS)} + \widehat{\beta}_{xy}(\mu_x - \overline{x}_{(StEMRSS)}) + \widehat{\beta}_{yz}(\mu_z - \overline{z}_{(StEMRSS)})](\overline{x}_{(StEMRSS)}/\mu_x)$
0	-1	$\widehat{\mu}_{y,(StEMRSS),s6} = [\overline{Y}_{(StEMRSS)} + \widehat{\beta}_{xy}(\mu_x - \overline{x}_{(StEMRSS)}) + \widehat{\beta}_{yz}(\mu_z - \overline{z}_{(StEMRSS)})](\overline{z}_{(StEMRSS)}/\mu_z)$
-1	-1	$\widehat{\mu}_{y,(StEMRSS),s7} = [\overline{Y}_{(StEMRSS)} + \widehat{\beta}_{xy}(\mu_x - \overline{x}_{(StEMRSS)}) + \widehat{\beta}_{yz}(\mu_z - \overline{z}_{(StEMRSS)})](\overline{x}_{(StEMRSS)}/\mu_x)(\overline{z}_{(StEMRSS)}/\mu_z)$

TABLE 2: Parametric values for variables X and Z.

Strata	Parameter	Normal		Uniform		Gamma		Weibull	
		X	Z	X	Z	X	Z	X	Z
1	μ	500	600	0	5	4	12	5	10
	σ^2	480	475	9	20	3	7	1.5	4.5
2	μ	1500	2000	4	8	8	14	10	18
	σ^2	500	775	16	15	5	6	4.5	9.5

TABLE 3: PRE of proposed estimator and existing estimators with respect to SRS for $r=3$ and $\rho_{xy} = 0.5$.

Estimator	Normal	$m = 10$			$m = 15$			Weibull
		Uniform	Gamma	Weibull	Normal	Uniform	Gamma	
$\widehat{\mu}_{yKS,StRSS}$	180.329	208.256	192.827	186.596	219.443	211.922	218.445	213.16
$\widehat{\mu}_{y,StEMRSS,s1}$	195.838	212.777	198.664	210.714	246.84	217.715	231.998	220.36
$\widehat{\mu}_{y,StEMRSS,s2}$	208.152	275.307	242.639	217.001	256.551	250.857	240.882	246.853
$\widehat{\mu}_{y,StEMRSS,s3}$	272.41	282.139	258.124	246.803	227.032	232.795	256.551	279.65
$\widehat{\mu}_{y,StEMRSS,s4}$	295.838	282.777	272.664	250.714	276.84	267.715	241.998	290.36
$\widehat{\mu}_{y,StEMRSS,G3}$	324.718	311.495	315.748	310.399	352.862	341.206	334.984	354.558

TABLE 4: PRE of proposed estimator and existing estimators with respect to SRS for $r=3$ and $\rho_{xy} = 0.75$.

Estimator	Normal	$m = 10$			$m = 15$			Weibull
		Uniform	Gamma	Weibull	Normal	Uniform	Gamma	
$\widehat{\mu}_{yKS,StRSS}$	186.628	210.133	189.468	199.347	220.186	217.198	220.979	216.207
$\widehat{\mu}_{y,StEMRSS,s1}$	198.926	232.398	201.296	221.532	248.286	222.356	238.088	231.297
$\widehat{\mu}_{y,StEMRSS,s2}$	218.164	277.189	248.392	219.823	261.275	259.196	251.178	248.366
$\widehat{\mu}_{y,StEMRSS,s3}$	277.936	281.175	263.088	255.254	224.471	241.345	258.129	289.144
$\widehat{\mu}_{y,StEMRSS,s4}$	299.256	292.176	275.83	256.197	279.196	268.113	246.389	293.974
$\widehat{\mu}_{y,StEMRSS,G3}$	332.672	315.196	325.745	328.927	358.187	351.391	344.186	368.352

TABLE 5: PRE of proposed estimator and existing estimators with respect to SRS for $r=3$ and $\rho_{xy} = 0.99$.

Estimator	Normal	$m = 10$			$m = 15$			Weibull
		Uniform	Gamma	Weibull	Normal	Uniform	Gamma	
$\widehat{\mu}_{yKS,StRSS}$	202.83	211.154	192.196	207.142	225.175	229.263	222.275	219.735
$\widehat{\mu}_{y,StEMRSS,s1}$	217.816	242.185	212.478	230.758	258.765	231.681	241.851	233.852
$\widehat{\mu}_{y,StEMRSS,s2}$	222.951	265.234	253.382	222.751	269.344	262.197	257.385	257.968
$\widehat{\mu}_{y,StEMRSS,s3}$	285.154	296.858	276.987	261.254	235.374	248.254	269.861	295.186
$\widehat{\mu}_{y,StEMRSS,s4}$	315.256	299.186	291.142	262.645	285.295	270.783	296.196	299.85
$\widehat{\mu}_{y,StEMRSS,G3}$	345.813	321.745	331.867	332.927	361.978	359.785	362.535	372.672

proposed estimator and its special cases are more efficient than existing estimators, revealing that PRE of proposed estimator increases with the increase in sample size.

In Table 4, hypothetical data of study and auxiliary variables with $r=3$, correlation coefficient of 0.75, and $m = 10$ and 15 are generated from normal, uniform, Gamma, and

TABLE 6: PRE of proposed estimator and existing estimators with respect to SRS for $r = 5$ and $\rho_{xy} = 0.5$.

Estimator	Normal	$m = 10$			$m = 15$			
		Uniform	Gamma	Weibull	Normal	Uniform	Gamma	Weibull
$\widehat{\mu}_{yKS,StRSS}$	196.862	201.754	185.623	201.852	220.368	219.184	220.853	213.762
$\widehat{\mu}_{y,StEMRSS,s1}$	199.286	232.285	203.722	215.268	251.623	225.746	237.354	228.296
$\widehat{\mu}_{y,StEMRSS,s2}$	201.723	247.972	247.861	207.718	258.344	258.296	241.571	253.285
$\widehat{\mu}_{y,StEMRSS,s3}$	245.835	287.254	270.868	252.979	221.381	241.854	247.297	287.488
$\widehat{\mu}_{y,StEMRSS,s4}$	287.123	282.926	284.754	251.175	276.964	263.186	289.867	279.197
$\widehat{\mu}_{y,StEMRSS,G3}$	315.199	319.815	325.131	330.385	350.362	357.482	358.478	362.208

TABLE 7: PRE of proposed estimator and existing estimators with respect to SRS for $r = 5$ and $\rho_{xy} = 0.75$.

Estimator	Normal	$m = 10$			$m = 15$			
		Uniform	Gamma	Weibull	Normal	Uniform	Gamma	Weibull
$\widehat{\mu}_{yKS,StRSS}$	206.824	216.846	188.386	212.586	228.355	229.487	231.586	218.497
$\widehat{\mu}_{y,StEMRSS,s1}$	210.296	245.194	217.978	229.386	261.873	227.397	246.395	238.243
$\widehat{\mu}_{y,StEMRSS,s2}$	216.723	257.083	251.153	222.978	274.867	264.586	258.593	258.187
$\widehat{\mu}_{y,StEMRSS,s3}$	255.227	292.275	283.196	261.354	239.376	244.397	253.933	290.907
$\widehat{\mu}_{y,StEMRSS,s4}$	299.589	299.097	290.472	273.099	287.988	269.597	297.397	284.847
$\widehat{\mu}_{y,StEMRSS,G3}$	335.475	327.164	331.586	349.486	362.486	364.386	364.275	368.486

TABLE 8: PRE of proposed estimator and existing estimators with respect to SRS for $r = 5$ and $\rho_{xy} = 0.99$.

Estimator	Normal	$m = 10$			$m = 15$			
		Uniform	Gamma	Weibull	Normal	Uniform	Gamma	Weibull
$\widehat{\mu}_{yKS,StRSS}$	210.835	229.286	195.296	228.133	238.987	236.386	237.386	225.384
$\widehat{\mu}_{y,StEMRSS,s1}$	219.296	255.294	222.573	237.386	275.486	232.687	252.875	248.856
$\widehat{\mu}_{y,StEMRSS,s2}$	223.164	263.283	264.967	229.586	284.274	272.836	268.186	262.197
$\widehat{\mu}_{y,StEMRSS,s3}$	269.496	306.296	291.242	274.354	249.496	248.976	262.799	308.195
$\widehat{\mu}_{y,StEMRSS,s4}$	304.253	317.866	301.682	279.286	292.297	278.386	318.476	291.292
$\widehat{\mu}_{y,StEMRSS,G3}$	345.968	338.114	357.286	353.385	374.381	372.186	388.196	371.143

Weibull distributions. Results show that under $m = 10$ and 15 , proposed estimator under StEMRSS has the highest PREs (bold values) for all choices of probability distributions. Under normal distribution, PREs of proposed estimator for $m = 10$ and 15 are 332.672 and 358.187 , respectively. Under uniform distribution, PREs of proposed estimator for $m = 10$ and 15 are 315.196 and 351.391 , respectively. Under Gamma distribution, PREs of proposed estimator for $m = 10$ and 15 are 325.745 and 344.186 , respectively. Under Weibull distribution, PREs of proposed estimator for $m = 10$ and 15 are 328.927 and 368.352 , respectively. Results also show that proposed estimator and its special cases are more efficient than existing estimators, revealing that PRE of proposed estimator increases with the increase in sample size. Results also indicate that PRE of proposed estimator increases when correlation increases from 0.5 to 0.75 .

In Table 5, hypothetical data of study and auxiliary variables with $r = 3$, correlation coefficient of 0.99 , and $m = 10$ and 15 are generated from normal, uniform, Gamma, and Weibull distributions. Results show that under $m = 10$ and 15 , proposed estimator under StEMRSS has the highest PREs (bold values) for all choices of probability distributions. Under normal distribution, PREs of proposed estimator for $m = 10$ and 15 are 345.813 and 361.978 , respectively. Under uniform distribution, PREs of proposed estimator for $m = 10$

and 15 are 321.745 and 359.785 , respectively. Under Gamma distribution, PREs of proposed estimator for $m = 10$ and 15 are 331.867 and 362.535 , respectively. Under Weibull distribution, PREs of proposed estimator for $m = 10$ and 15 are 332.927 and 372.672 , respectively. Results in Table 5 also show that proposed estimator and its special cases are more efficient than existing estimators, revealing that PRE of proposed estimator increases with the increase in sample size. Results also indicate that PRE of proposed estimator increases when correlation increases from 0.75 to 0.99 .

In Table 6, hypothetical data of study and auxiliary variables with $r = 5$, correlation coefficient of 0.5 , and $m = 10$ and 15 are generated from normal, uniform, Gamma, and Weibull distributions. Results show that under $m = 10$ and 15 , proposed estimator under StEMRSS has the highest PREs (bold values) for all choices of probability distributions. Under normal distribution, PREs of proposed estimator for $m = 10$ and 15 are 315.199 and 350.362 , respectively. Under uniform distribution, PREs of proposed estimator for $m = 10$ and 15 are 319.815 and 357.482 , respectively. Under Gamma distribution, PREs of proposed estimator for $m = 10$ and 15 are 325.131 and 358.478 , respectively. Under Weibull distribution, PREs of proposed estimator for $m = 10$ and 15 are 330.385 and 362.208 , respectively. Results also show that proposed estimator and its special cases are more efficient

than existing estimators, revealing that PRE of proposed estimator increases with the increase in sample size.

In Table 7, hypothetical data of study and auxiliary variables with $r = 5$, correlation coefficient of 0.75, and $m = 10$ and 15 are generated from normal, uniform, Gamma, and Weibull distributions. Results show that under $m = 10$ and 15, proposed estimator under StEMRSS has the highest PREs (bold values) for all choices of probability distributions. Under normal distribution, PREs of proposed estimator for $m = 10$ and 15 are 335.475 and 362.486, respectively. Under uniform distribution, PREs of proposed estimator for $m = 10$ and 15 are 327.164 and 364.386, respectively. Under Gamma distribution, PREs of proposed estimator for $m = 10$ and 15 are 331.586 and 364.275, respectively. Under Weibull distribution, PREs of proposed estimator for $m = 10$ and 15 are 349.486 and 368.486, respectively. Results also show that proposed estimator and its special cases are more efficient than existing estimators, revealing that PRE of proposed estimator increases with the increase in sample size. Results also indicate that PRE of proposed estimator increases when correlation increases from 0.5 to 0.75.

In Table 8, hypothetical data of study and auxiliary variables with $r = 5$, correlation coefficient of 0.99, and $m = 10$ and 15 are generated from normal, uniform, Gamma, and Weibull distributions. Results show that under $m = 10$ and 15, proposed estimator under StEMRSS has the highest PREs (bold values) for all choices of probability distributions. Under normal distribution, PREs of proposed estimator for $m = 10$ and 15 are 345.968 and 374.381, respectively. Under uniform distribution, PREs of proposed estimator for $m = 10$ and 15 are 338.114 and 372.186, respectively. Under Gamma distribution, PREs of proposed estimator for $m = 10$ and 15 are 357.286 and 388.196, respectively. Under Weibull distribution, PREs of proposed estimator for $m = 10$ and 15 are 353.385 and 371.143, respectively. Results also show that proposed estimator and its special cases are more efficient than existing estimators, revealing that PRE of proposed estimator increases with the increase in sample size. Results also indicate that PRE of proposed estimator increases when correlation increases from 0.75 to 0.99.

Results in Tables 3–8 reveal that PRE of proposed estimator increases when set size m increases with fixed number of cycles r and coefficient of correlation. Results also show that PRE of proposed estimator is directly proportional to r and coefficient of correlation. From the above results, we conclude that our proposed estimator performs efficiently under symmetrical as well as asymmetrical distributions.

5. Real-Life Application of Proposed Estimator

To compare the percent relative efficiencies (PREs) of proposed estimator and existing estimator under StRSS, a real-life data set given by Bierens and Ginther [19] has been utilized. Data consists of 4 variables related to wages of employees in USA: wage (in dollars per week) is selected as study variable Y , number of years of education and number of years of work experience are selected as auxiliary variables (X and Z), and variable ‘does the individual work part time or not?’ has been used for stratification purpose. Stratum 1

consists of those individuals who are regular employees, and the rest are placed in stratum 2. We also utilized t -test for comparison of means in two strata. Summary statistics of data and results of t -test are presented in Table 9.

Results in Table 9 show that population of size 28155 is divided into two strata of sizes 25631 and 2524, respectively. Average wage (in dollars per week) in stratum 1 is 640.1625 whereas in stratum 2 its value is 233.7264. Variations in wage (in dollars per week) in strata 1 and 2 are 197379.9 and 139784.5, respectively. P value of t -test is 0.000, which shows that we reject our H_0 ($\mu_{y_1} = \mu_{y_2}$) and conclude that there is statistically significant difference between means of two strata.

The best fitted probability distributions of study and auxiliary variables are presented in Table 10. The study variable (wage (in dollars per week)) follows Gamma distribution with values of shape parameter, scale parameter, and location parameter computed as 3.332, 177.320, and -12.628 , respectively. The auxiliary variable X (number of years of education) also follows Gamma distribution with values of shape parameter, scale parameter, and location parameter computed as 210.170, 0.182, and -25.284 , respectively. Moreover, the auxiliary variable Z (number of years of work experience) follows Weibull distribution with values of shape parameter, scale parameter, and location parameter computed as 1.482, 21.873, and -1.499 , respectively. Results of Tables 9 and 10 reveal that this data has heterogeneity and outliers. Therefore, in this type of data, we suggest the use of our proposed generalized estimator under stratified extreme-cum-median ranked set sampling. The distributional graphs of all variables are presented in Figures 1–3. For simplicity of analysis, we select $r = 1, 2, 3$, and 4 and $m = 10$ and 15.

Table 11 shows percent relative efficiencies (PREs) of proposed estimator and existing estimators under stratified ranked set sampling scheme with respect to SRS for real-life data. Samples from real-life data set have been taken by using $r = 1, 2, 3$, and 4 and $m = 10$ and 15. Proposed estimator under StEMRSS has the highest PREs (bold values) for all choices of r and m . For $r = 1$ and $m = 10$ and 15, PREs of proposed estimator are 326.363 and 339.367, respectively. For $r = 2$ and $m = 10$ and 15, PREs of proposed estimator are 342.497 and 354.834, respectively. For $r = 3$ and $m = 10$ and 15, PREs of proposed estimator are 350.1887 and 357.264, respectively. For $r = 4$ and $m = 10$ and 15, PREs of proposed estimator are 356.194 and 379.278, respectively. Results also show that proposed estimator and its special cases are more efficient than existing estimators, revealing that PRE of proposed estimator increases with the increase in sample size.

6. Distribution of Proposed Estimator for Simulated Data

In this section, we identify the best fitted probability distribution of proposed estimator under StEMRSS. The auxiliary variables are generated by using normal distribution with parameters defined in Table 2. One million samples of size $n = 75$, with $r = 5$, $m = 15$, and $\rho_{xy} = 0.99$, are taken, and

TABLE 9: Summary statistics of wage (in dollars per week).

Strata	Size	Mean	Variance	Minimum	Maximum	t statistic	P value
1	25631	640.1625	197387.624	50.39	18777.2		
2	2524	233.7263	139839.896	50.05	9269.26	51.163	0.000

TABLE 10: Probability distributions of variables with parametric values.

Variable	Distribution	Parameters		
		Shape	Scale	Location
Wages	Gamma	3.332	177.320	-12.628
Education	Gamma	210.170	0.182	-25.284
Experience	Weibull	1.482	21.873	-1.499

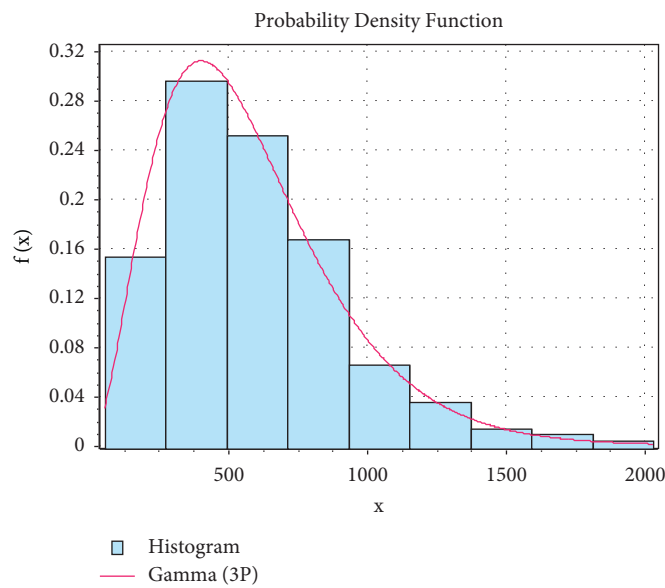


FIGURE 1: Probability distribution of variable “wage (in dollars per week).”.

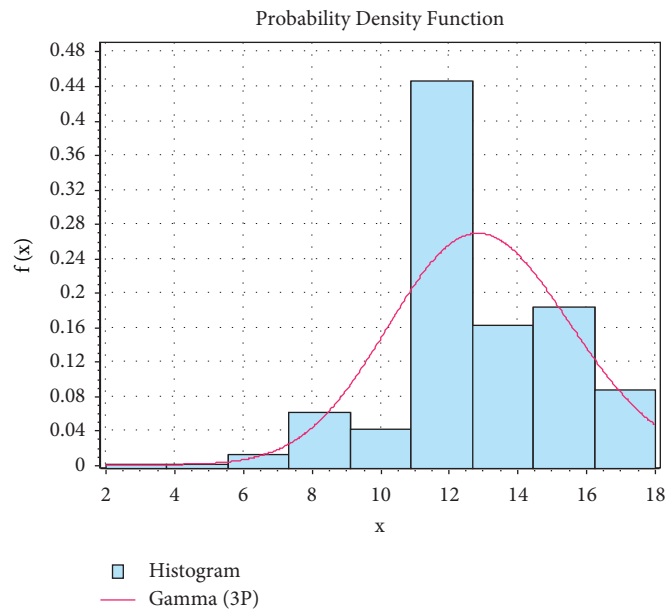


FIGURE 2: Probability distribution of variable “number of years of education.”.

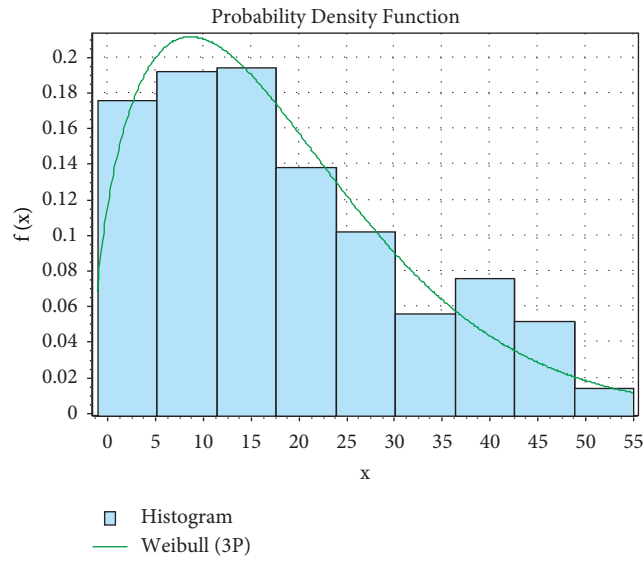


FIGURE 3: Probability distribution of variable “number of years of work experience.”.

TABLE 11: PRE of proposed estimator and existing estimators with respect to SRS for real-life data.

Estimator	$r=1$	$r=2$	$r=3$	$r=4$	$r=1$	$r=2$	$r=3$	$r=4$
	$m=10$				$m=15$			
$\hat{\mu}_{y,KS,StRSS}$	225.753	225.974	228.836	238.286	229.268	228.848	242.164	248.267
$\hat{\mu}_{y,StEMRSS,s1}$	228.862	236.865	245.286	247.482	233.197	246.977	249.083	257.478
$\hat{\mu}_{y,StEMRSS,s2}$	230.197	260.274	265.284	273.297	245.138	263.682	282.861	292.276
$\hat{\mu}_{y,StEMRSS,s3}$	244.386	256.387	283.274	296.973	250.389	284.867	299.235	324.278
$\hat{\mu}_{y,StEMRSS,s4}$	290.287	293.486	295.903	299.284	91.297	297.397	329.489	332.476
$\hat{\mu}_{y,StEMRSS,G3}$	326.363	342.497	350.187	356.194	339.367	354.834	357.264	379.278

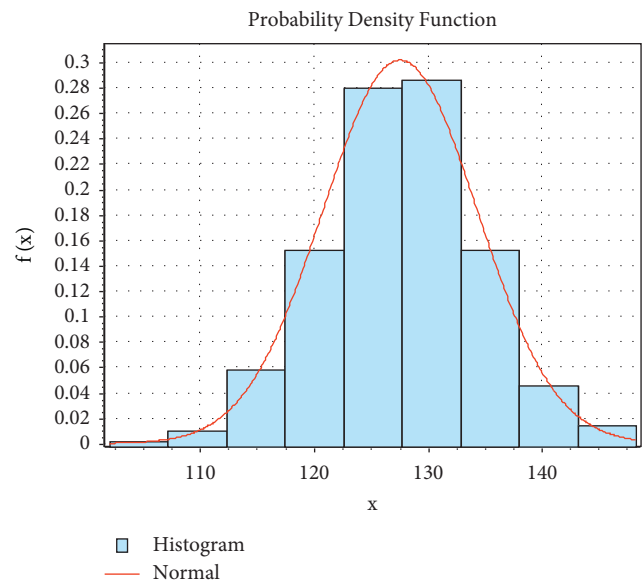


FIGURE 4: Probability distribution of proposed estimator under StEMRSS.

sampling distribution of proposed estimator is constructed. It is found out that proposed estimator under StEMRSS follows normal distribution with mean of 127.56 and standard deviation of 6.7945. Figure 4 shows the probability distribution of proposed estimator under StEMRSS.

7. Conclusion

The objective of this study is to propose a class of generalized estimators under stratified extreme-cum-median ranked set sampling (StEMRSS). To achieve this objective, a class of generalized chain regression-cum-chain ratio estimator for population mean is proposed using two auxiliary variables under StEMRSS. Monte Carlo simulation study has been conducted for comparison of proposed estimator under StEMRSS with some existing estimators using two auxiliary variables under StRSS. The results show that the proposed estimator under StEMRSS performs more efficiently as compared to existing estimators under both symmetrical and asymmetrical probability distributions. The use of the proposed estimator in real-life application includes an example to estimate average wages (in dollars per week) of employees in USA. The results of real-life application show that the proposed estimator efficiently estimates the average wages of employees in USA as compared to existing estimators using two auxiliary variables under RSS schemes. The probability distributions of study and auxiliary variables are explored. These results are also compatible with the results of Monte Carlo simulation, as both studies show the efficiency of the proposed estimator in StEMRSS under the specified distributions. The distribution of the proposed estimator is also identified, and it is found out that the proposed estimator follows normal distribution. Therefore, we can use any parametric test on the results obtained by the proposed estimator. In the light of the above results and discussions, it is recommended that the proposed estimator is used for estimating population mean of study variable when there are heterogeneity and outliers in the data sets.

Data Availability

The data used to support this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] W. G. Cochran, *Sampling Techniques*, Wiley, New York, NY, USA, 1977.
- [2] J. Neyman, "On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection," *Journal of the Royal Statistical Society*, vol. 97, no. 4, pp. 558–625, 1934.
- [3] J. Neyman, "Contribution to the theory of sampling human populations," *Journal of the American Statistical Association*, vol. 33, no. 201, pp. 101–116, 1938.
- [4] G. McIntyre, "A method for unbiased selective sampling, using ranked sets," *Australian Journal of Agricultural Research*, vol. 3, no. 4, pp. 385–390, 1952.
- [5] K. Takahasi and K. Wakimoto, "On unbiased estimates of the population mean based on the sample stratified by means of ordering," *Annals of the Institute of Statistical Mathematics*, vol. 20, no. 1, pp. 1–31, 1968.
- [6] T. R. Dell and J. L. Clutter, "Ranked set sampling theory with order statistics background," *Biometrics*, vol. 28, no. 2, pp. 545–555, 1972.
- [7] S. Lynne Stokes, "Ranked set sampling with concomitant variables," *Communications in Statistics - Theory and Methods*, vol. 6, no. 12, pp. 1207–1211, 1977.
- [8] H. M. Samawi, "Stratified ranked set sample," *Pakistan Journal of Statistics*, vol. 12, pp. 9–16, 1996.
- [9] J. Graunt, *Natural and Political Observations Mentioned in a Following Index, and Made upon the Bills of Mortality*, John Martin & James Allestry, London, UK, 2nd edition, 1662.
- [10] H. M. Samawai and L. J. Saeid, "Stratified extreme ranked set sample with application to ratio estimators," *Journal of Modern Applied Statistical Methods*, vol. 3, no. 1, pp. 117–133, 2004.
- [11] K. Ibrahim, M. Syam, and A. I. Al-Omari, "Estimating the population mean using stratified median ranked set sampling," *Applied Mathematical Sciences*, vol. 4, no. 47, pp. 2341–2354, 2010.
- [12] L. Khan, J. Shabbir, and S. Gupta, "Improved ratio-type estimators using stratified double-ranked set sampling," *Journal of Statistical Theory and Practice*, vol. 10, no. 4, pp. 755–767, 2016a.
- [13] A. Ali, M. M. Butt, M. D. Azad, Z. Ahmed, and M. Hanif, "Stratified extreme-cum-median ranked set sampling," *Pakistan Journal of Statistics*, vol. 37, no. 3, pp. 215–235, 2021.
- [14] K. Iqbal, M. Moeen, A. Ali, and A. Iqbal, "Mixture regression cum ratio estimators of population mean under stratified random sampling," *Journal of Statistical Computation and Simulation*, vol. 90, no. 5, pp. 854–868, 2020.
- [15] A. Ali, M. M. Butt, K. Iqbal, M. Hanif, and M. Zubair, "Estimation of population mean by using a generalized family of estimators under classical ranked set sampling," *RMS: Research in Mathematics & Statistics*, vol. 8, no. 1, Article ID 1948184, 2021.
- [16] I. Olkin, "Multivariate ratio estimation for finite populations," *Biometrika*, vol. 45, no. 1/2, pp. 154–165, 1958.
- [17] L. Khan and J. Shabbir, "Generalized exponential-type ratio-cum ratio estimators of population mean in ranked set and stratified ranked set sampling," *Journal of Statistics & Management Systems*, vol. 20, no. 1, pp. 133–151, 2017a.
- [18] M. Zubair and A. Ali, "A generalized regression-product estimator in presence of non-response," in *Proceedings of the 17th International Conference on Statistical Sciences*, pp. 183–198, Lahore, Pakistan, January 2019.
- [19] H. J. Bierens and D. K. Ginther, "Integrated conditional moment testing of quantile regression models," *Empirical Economics*, vol. 26, no. 1, pp. 307–324, 2001.