http://communications.science.ankara.edu.tr

Commun.Fac.Sci.Univ.Ank.Ser. A1 Math. Stat.
Volume 73, Number 2, Pages 336–348 (2024)
DOI:10.31801/cfsuasmas.1373759
ISSN 1303-5991 E-ISSN 2618-6470

Research Article; Received: October 10, 2023; Accepted: December 13, 2023

ROBUST REGRESSION TYPE ESTIMATORS FOR BODY MASS INDEX UNDER EXTREME RANKED SET AND QUARTILE RANKED SET SAMPLING

Arzu Ece CETIN[1] and Nursel KOYUNCU[2]

[1]Department of Management, Gebze Technical University, Kocaeli, TÜRKİYE
[2]Department of Statistics, Hacettepe University, Ankara, TÜRKİYE

Abstract. Robust regression-type estimators of population mean that use auxiliary variable information are proposed by considering robust methods under extreme ranked set sampling (ERSS) and quartile ranked set sampling (QRSS). We have used the data concerning body mass index (BMI) for 800 people in Turkey in 2014. The real data example is applied to see efficiency of the estimators in ERSS and QRSS designs and it is found that the proposed estimators are better in these designs than the classical ranked set sampling (RSS) design. In addition, mean square error (MSE) and percent relative efficiency (PRE) are used to compare the performance of the adapted and proposed estimators.

1. Introduction

In sampling survey, the supplementary information is mostly used to enhance accuracy of the estimators due to the correlation between auxiliary and study variables. Auxiliary information has a major role according to the sampling theory. Because of improving the precision of estimates, making use of convenient auxiliary information such as mean, total population, skewness, attribute and correlation is pretty significant. Auxiliary information has been used in ratio, product and exponential type estimators to acquire effective estimators under distinct sampling designs.

RSS is an alternative sampling design to simple random sampling (SRS) for drawing a sample of observations from a population. It is intended for situations where the certain measurement of sample units is hard but they can be readily

2020 *Mathematics Subject Classification.* 62Dxx, 62D05.
Keywords. Robust regression estimators, extreme ranked set sampling, quartile ranked set sampling, percentage relative efficiency (PRE).
aecetin@gtu.edu.tr-Corresponding author; 0000-0001-7224-9698
nkoyuncu@hacettepe@edu.tr; 0000-0003-1065-3411.

ranked without real measurement. The ranking is done either through nominative judgment or via the use of an accompanying variable, and need not to be precise. This situation is named defective ranking. If the ranking process is correct, it will be referred to as excellent ranking structure. RSS design was first proposed by McIntyre [6]. Many authors such as Samawi and Muttlak [9], Bouza [2], Mehta and Mandowara [7] used judge mental RSS where ranking is done with respect to auxiliary variable. Later, the authors suggested new sampling designs based on ranking and used auxiliary information to get efficiency. Muttlak [8] proposed QRSS. Taking into account ranking error, Samawi et al. [10] suggested ERSS for estimating a population mean. Long et al. [5] suggested ratio estimators of population mean that used either the first or third quartiles of the auxiliary variable under RSS and ERSS. Koyuncu [3] studied regression type estimators (RTE) under different ranked set sampling. Shahzad et al. [11] suggested RTE for mean estimation under RSS besides the sensitivity issue.

Lately, robust tools are used in estimators under different sampling designs. Zaman and Bulut [14] are proposed new ratio type estimators using LTS, Huber MM, LMS, Tukey-M, LAD and Hampel M robust methods in SRS. Ali et al. [1] generalized estimators of Zaman and Bulut [14]. Subzar et al. [13] adapted the diverse robust regression methods to the ratio estimators. Shahzad et al. [12] identified the class of RTE utilizing robust regression tools. Recently, Koyuncu and Al-Omari [4] proposed generalized robust RTE under RSS and MRSS.

The target of this study is to suggest regression type estimators of the population mean using robust statistics under RSS, ERSS and QRSS. The article is composed as follows: In Section 2, RSS, ERSS and QRSS designs were explained. In Section 3, the recent robust literature were reviewed and adapted robust regression type estimators were given. The proposed exponential robust-RTE estimators in RSS, ERSS and QRSS were introduced in Section 4. In Section 5, a numerical study was conducted using a real data set on BMI. All results that were explained briefly and summarized also in Section 6.

2. RSS, ERSS and QRSS Designs

In this section, RSS, ERSS and QRSS designs are explained.

2.1. **RSS Design.**

The RSS procedure can be created by choosing r random samples of size r units from the population and order the units within each sample according to the variable of interest. Let $(X_1, Y_1), (X_2, Y_2), ..., (X_r, Y_r)$ be a SRS of r, then the measured RSS units are indicated by $\left(Y_{(i)j}, X_{[i]j}\right)$, $i = 1, 2, ..., r$, $j = 1, 2, ..., m$ where $\left(X_{[i]j}, Y_{(i)j}\right)$ is the i^{th} ranked unit from the j^{th} cycle of two auxiliary variables and study variable, respectively. [ ] and ( ) demonstrate the i^{th} perfect ordering in the i^{th} set for auxiliary variable X and the i^{th} judgment ordering in the i^{th} set for study variable Y. One of the most correlated auxiliary variables with study variable was choosed to rank the units. Further let $\bar{x}_{RSS} = \frac{1}{mr}\sum\limits_{j=1}^{m}\sum\limits_{i=1}^{r} X_{[i]j}$,

$\bar{y}_{RSS} = \frac{1}{mr} \sum_{j=1}^{m} \sum_{i=1}^{r} Y_{(i)j}$ are the sample means under RSS and $\bar{Y}, \bar{X}$ are population means, respectively for the study and auxiliary variables.

2.2. **ERSS Design.** ERSS investigated by Samawi et al. [10]. To predict the finite population mean $\left(\bar{Y}\right)$ using ERSS, the operation can be explained briefly as follows:

(1) The process includes drawing sets of each r units randomly from population for which the mean is to be predicted. The most important assumption is the smallest and the biggest units of the set can be fixed visually or with a little cost.
(2) The lowest ranked unit is determined from the first r unit set. Then, the largest ranked unit is determined from the second r unit set. And the lowest ranked unit is determined from the third set of r units and so on. Thus, the first $(r-1)$ determined units is obtained using the first $(r-1)$ sets. The event of choosing the $r-th$ unit from the $r-th$ (i.e very last) set depends on whether r is odd or even.
(3) When r is even, the measurement value of the largest unit ranked is measured.
(4) Two options exist when r is odd:
    (a) The average of the largest and lowest units in the $r-th$ set is measured for the measure of the $r-th$ unit.

    (b) The measure of the median for the measure of the $r-th$ unit is measured.
(5) This procedure complete one cycle of ERSS. The period may be repeated m times until n elements of desired to obtain.

$$\bar{x}_{ERSS_e} = \frac{1}{2}\left(\bar{X}_{[1]} + \bar{X}_{[r]}\right) \tag{1}$$

where $\bar{X}_{[1]} = \frac{2}{r} \sum_{i=1}^{r/2} X_{2i-1[1]}$. and $\bar{X}_{[r]} = \frac{2}{r} \sum_{i=1}^{r/2} X_{2i[r]}$.

To observe that $X_{1[1]}, X_{3[1]}, ..., X_{r-2[1]}$ and $X_{r[1]}$are identically distributed is easy and so are $X_{2[r]}, X_{4[r]}, ..., X_{r-1[r]}$ and $X_{r(r)}$.

$$\bar{x}_{ERSS_o} = \frac{X_{1[1]}, X_{2[r]}, X_{3[1]}, ..., X_{r-1[r]} + X_{r\left[\frac{r+1}{2}\right]}}{r} \tag{2}$$

2.3. **QRSS Design.** Muttlak [8] suggested QRSS to predict the population mean. The procedure of QRSS can be explained concisely as follows:

(1) Select randomly r^2 bivariate sample units of target population.
(2) If the sample size r is even, choose for measurement from the first $\frac{r}{2}$ samples the $q_1\left(r+1\right)$ th and from the second $\frac{r}{2}$ samples the $q_3\left(r+1\right)$th smallest ranked unit.

(3) If the sample size r is odd, choose for measurement from the first $\frac{(r-1)}{2}$ samples the $q_1\left(r+1\right)$th and from the last $\frac{(r-1)}{2}$samples the $q_3\left(r+1\right)$ th smallest ranked unit and from the remaining sample the median ranked unit.
(4) The nearest integer of $q_1\left(r+1\right)$th and $q_3\left(r+1\right)$th where $q_1 = 0.25$ and $q_3 = 0.75$ were always taken.
(5) This procedure complete one cycle of QRSS. The cycle may be repeated m times until $n = mr$ elements of desired to obtain.

Let $X_{i[q_1(r+1)]}$ and $X_{i[q_3(r+1)]}$ denote the $\left(q_1\left(r+1\right)\right)$th and $\left(q_3\left(r+1\right)\right)$th order statistics of the i^{th} sample respectively $\left(i = 1, 2, .., r\right)$.

The estimator of the population mean using QRSS with a cycle is given in equations 3 and 4, respectively, in the case of even and odd sample sizes.

$$\bar{x}_{QRSS_e} = \frac{1}{r}\left(\sum_{i=1}^{\frac{r}{2}} X_{i[q_1(r+1)]} + \sum_{i=\frac{r}{2}+1}^{r} X_{i[q_3(r+1)]}\right) \tag{3}$$

$$\bar{x}_{QRSS_o} = \frac{1}{r}\left(\sum_{i=1}^{\frac{r-1}{2}} X_{i[q_1(r+1)]} + \sum_{i=\frac{r+1}{2}}^{r} X_{i[q_3(r+1)]} + X_{i[(r+1)/2]}\right) \tag{4}$$

$X_{i[(r+1)/2]}$ is the median of sample $i = \left(r+1\right)/2$. To simplify the notations, let $X_{[i:q]}$ specify the $\left(q_1\left(r+1\right)\right)$th order statistic of i^{th} sample $\left(i = 1, 2, ..., \frac{r}{2}\right)$ and $\left(q_3\left(r+1\right)\right)$th order statistic of i^{th} sample $\left(i = \frac{r}{2}+1, \frac{r}{2}, ..., r\right)$ if the sample size n is even. Also specify the $\left(q_1\left(r+1\right)\right)$th order statistic of i^{th} sample $\left(i = 1, 2, ..., \frac{r-1}{2}\right)$, the median of the i^{th} sample $\left(i = \left(r+1\right)/2\right)$ and the $\left(q_3\left(r+1\right)\right)$th order statistic of i^{th} sample $\left(i = \frac{r-1}{2}+2, \frac{r-1}{2}+3, ..., r\right)$ if the sample size n is odd. Then the estimator of population mean using QRSS can be written as $\bar{x}_{QRSS} = \frac{1}{r}\sum_{i}^{r} X_{[i:q]}$.

3. Adapted Robust Regression Type Estimators

Koyuncu and Al-Omari [4] proposed generalized robust-RTE under SRS, RSS and median ranked set sampling (MRSS).

$$\bar{y}_{N(j)} = \left[\bar{y}_{(j)} + b_{i(j)}\left(\bar{X} - \bar{x}_{[j]}\right)\right]\left(\frac{F\bar{X}+G}{F\bar{x}_{[j]}+G}\right)^{\alpha} \tag{5}$$

where F may represent the coefficient of variation C_x, kurtosis $\beta_{2(x)}$, first and third quarters $q_{1(x)}, q_{3(x)}$ or any known population information of auxiliary variable. (j) represents the SRS, RSS and MRSS sampling designs. $b_{i(j)}$ is regression coefficient calculated from the i robust regression method under (j) design. i represents Huber M, LMS, Huber MM, S, LAD or LTS.

They showed that Zaman and Bulut [14] estimators are members of their generalized estimator. Putting suitable values as $\alpha = 1$, $F = 1$, $G = q_{1(x)}, q_{3(x)}$ and j=SRS in the $\bar{y}_{N(j)}$, we can get Zaman and Bulut [14] ratio-RTEs under SRS.

In the same manner, we can extend $\bar{y}_{N(j)}$ estimator to ERSS and QRSS designs putting j=ERSS, j=QRSS respectively. Zaman and Bulut [14] estimators and some members of $\bar{y}_{N(j)}$ can be given as

$$\bar{y}_{EN(j)1}=\frac{\left[\bar{y}_{(j)}+b_{i(j)}\left(\bar{X}-\bar{x}_{[j]}\right)\right]}{\bar{x}_{[j]}}\bar{X} \tag{6}$$

$$\bar{y}_{EN(j)2}=\frac{\left[\bar{y}_{(j)}+b_{i(j)}\left(\bar{X}-\bar{x}_{[j]}\right)\right]}{\bar{x}_{[j]}+C_x}\left(\bar{X}+C_x\right) \tag{7}$$

$$\bar{y}_{EN(j)3}=\frac{\left[\bar{y}_{(j)}+b_{i(j)}\left(\bar{X}-\bar{x}_{[j]}\right)\right]}{\bar{x}_{[j]}+\beta_{2(x)}}\left(\bar{X}+\beta_{2(x)}\right) \tag{8}$$

$$\bar{y}_{EN(j)4}=\frac{\left[\bar{y}_{(j)}+b_{i(j)}\left(\bar{X}-\bar{x}_{[j]}\right)\right]}{\bar{x}_{[j]}+q_{1(x)}}\left(\bar{X}+q_{1(x)}\right) \tag{9}$$

$$\bar{y}_{EN(j)5}=\frac{\left[\bar{y}_{(j)}+b_{i(j)}\left(\bar{X}-\bar{x}_{[j]}\right)\right]}{\bar{x}_{[j]}+q_{3(x)}}\left(\bar{X}+q_{3(x)}\right) \tag{10}$$

To obtain the specific MSE of adapted estimators in equation (5) under (j) design, let us define following notations

$$\vartheta_{o(j)}=\left(\bar{y}_{(j)}-\bar{Y}\right)/\bar{Y},\vartheta_{1(j)}=\left(\bar{x}_{[j]}-\bar{X}\right)/\bar{X}\quad \vartheta_{0(j)}\vartheta_{1(j)}=\left(\bar{x}_{[j]}-\bar{X}\right)\left(\bar{y}_{(j)}-\bar{Y}\right)/\bar{X}\bar{Y} \tag{11}$$

For the (j) design, expectaions of ϑ terms are given by

$$E\left(\vartheta_{0(j)}^2\right)=V\left(\bar{y}_{(j)}\right)/\bar{Y}^2,E\left(\vartheta_{1(j)}^2\right)=V\left(\bar{x}_{[j]}\right)/\bar{X}^2,E\left(\vartheta_{0(j)}\vartheta_{1(j)}\right)=\operatorname{cov}\left(\bar{x}_{[j]},\bar{y}_{(j)}\right)/\bar{Y}\bar{X}$$

If (j) design represents SRS, expectaions of ϑ terms are given by

$$E\left(\vartheta_{0(SRS)}^2\right)=\frac{S_y^2}{\bar{Y}^2},\quad E\left(\vartheta_{1(SRS)}^2\right)=\frac{S_x^2}{\bar{X}^2},\quad E\left(\vartheta_{0(SRS)}\vartheta_{1(SRS)}\right)=\frac{S_{xy}}{\bar{Y}\bar{X}}$$

If (j) design represents RSS, expectaions of ϑ terms are given by

$$E\left(\vartheta_{0(RSS)}^2\right)=\frac{1}{\bar{Y}^2}\left(\frac{S_y^2}{r}-\frac{1}{r^2}\sum_{i=1}^{r}\left(\mu_{y(j)}-\bar{Y}\right)^2\right)$$

$$E\left(\vartheta_{1(RSS)}^2\right)=\frac{1}{\bar{X}^2}\left(\frac{S_x^2}{r}-\frac{1}{r^2}\sum_{i=1}^{r}\left(\mu_{x[i]}-\bar{X}\right)^2\right),$$

$$E\left(\vartheta_{0(RSS)}\vartheta_{1(RSS)}\right) = \frac{1}{\bar{Y}\bar{X}}\left(\frac{S_{xy}}{r} - \frac{1}{r^2}\sum_{i=1}^{r}\left(\mu_{x[i]} - \bar{X}\right)\left(\mu_{y(i)} - \bar{Y}\right)\right)$$

If the sample size is odd and (j) represents the QRSS design, expectaions of ϑ terms are given by

$$E\left(\vartheta^2_{0(QRSS)_o}\right) = \frac{1}{\bar{Y}^2}\left[\frac{1}{r^2}\left(\frac{(r-1)}{2}\left(S^2_{y\left(\frac{r+1}{4}\right)} + S^2_{y\left(\frac{3(r+1)}{4}\right)}\right) + S^2_{y\left(\frac{r+1}{2}\right)}\right)\right],$$

$$E\left(\vartheta^2_{1(QRSS)_o}\right) = \frac{1}{\bar{X}^2}\left[\frac{1}{r^2}\left(\frac{(r-1)}{2}\left(S^2_{x\left[\frac{r+1}{4}\right]} + S^2_{x\left[\frac{3(r+1)}{4}\right]}\right) + S^2_{x\left[\frac{r+1}{2}\right]}\right)\right],$$

$$E\left(\vartheta_{0(QRSS)_o}\vartheta_{1(QRSS)_o}\right) = \frac{1}{\bar{Y}\bar{X}}\left[\frac{1}{r^2}\left(\frac{(r-1)}{2}\left(S_{xy\left(\frac{r+1}{4}\right)} + S_{xy\left(\frac{3(r+1)}{4}\right)}\right) + S_{xy\left(\frac{r+1}{2}\right)}\right)\right]$$

If the sample size is even and (j) represents the QRSS design, expectaions of ϑ terms are given by

$$E\left(\vartheta^2_{0(QRSS)_e}\right) = \frac{1}{\bar{Y}^2}\left[\frac{1}{2r}\left(S^2_{y\left(\frac{r+1}{4}\right)} + S^2_{y\left(\frac{3(r+1)}{4}\right)}\right)\right],$$

$$E\left(\vartheta^2_{1(QRSS)_e}\right) = \frac{1}{\bar{X}^2}\left[\frac{1}{2r}\left(S^2_{x\left[\frac{r+1}{4}\right]} + S^2_{x\left[\frac{3(r+1)}{4}\right]}\right)\right],$$

$$E\left(\vartheta_{0(QRSS)_e}\vartheta_{1(QRSS)_e}\right) = \frac{1}{\bar{Y}\bar{X}}\left[\frac{1}{2r}\left(S_{xy\left(\frac{r+1}{4}\right)} + S_{xy\left(\frac{3(r+1)}{4}\right)}\right)\right]$$

If the sample size is odd and (j) represents the ERSS design, expectaions of ϑ terms are given by

$$E\left(\vartheta^2_{0(ERSS)_o}\right) = \frac{1}{\bar{Y}^2}\left[\frac{1}{r^2}\left(\frac{(r-1)}{2}\left(S^2_{y(1)} + S^2_{y(r)}\right) + S^2_{y\left(\frac{r+1}{2}\right)}\right)\right],$$

$$E\left(\vartheta^2_{1(ERSS)_o}\right) = \frac{1}{\bar{X}^2}\left[\frac{1}{r^2}\left(\frac{(r-1)}{2}\left(S^2_{x[1]} + S^2_{x[r]}\right) + S^2_{x\left[\frac{r+1}{2}\right]}\right)\right],$$

$$E\left(\vartheta_{0(ERSS)_o}\vartheta_{1(ERSS)_o}\right) = \frac{S_{xy}}{\bar{Y}\bar{X}}\left[\frac{1}{r^2}\left(\frac{(r-1)}{2}\left(S_{xy(1)} + S_{xy(r)}\right) + S_{xy\left(\frac{r+1}{2}\right)}\right)\right]$$

If the sample size is even and (j) represents the ERSS design, expectaions of ϑ terms are given by

$$E\left(\vartheta^2_{0(ERSS)_e}\right) = \frac{1}{\bar{Y}^2}\left[\frac{1}{2r}\left(S^2_{y(1)} + S^2_{y(r)}\right)\right]$$

$$E\left(\vartheta^2_{1(ERSS)_e}\right) = \frac{1}{\bar{X}^2}\left[\frac{1}{2r}\left(S^2_{x[1]} + S^2_{x[r]}\right)\right]$$

$$E\left(\vartheta_{0(ERSS)_e}\vartheta_{1(ERSS)_e}\right) = \frac{1}{\bar{Y}\bar{X}}\left(\frac{1}{2r}\left(S_{xy(1)} + S_{xy(r)}\right)\right)$$

Writing $\bar{y}_{EN(j)}$ given in Equation 5 with ϑ terms, extracting $\bar{Y}$ and squaring both sides we get MSE of generalized estimator $\bar{y}_{EN(j)}$ under (j) design as

$$MSE\left(\bar{y}_{EN(j)i}\right) = E\Big(\bar{Y}^2\vartheta^2_{0(j)} + B_i^2\bar{X}^2\vartheta^2_{1(j)} + \alpha^2\psi^2\bar{Y}^2\vartheta^2_{1(j)} - 2B_i\bar{Y}\bar{X}\vartheta_{0(j)}\vartheta_{1(j)}$$
$$-2\alpha\psi\bar{Y}^2\vartheta_{0(j)}\vartheta_{1(j)} + 2\alpha\psi B_i\bar{Y}\bar{X}\vartheta^2_{1(j)}) \tag{12}$$

where $\psi = \dfrac{F\bar{X}}{F\bar{X}+G}$

$$MSE\left(\bar{y}_{EN(j)i}\right) = V\left(\bar{y}_{(j)}\right) + B_i^2 V\left(\bar{x}_{[j]}\right) + \alpha^2 R^2_{FG} V\left(\bar{x}_{[j]}\right) - 2B_i \operatorname{cov}\left(\bar{x}_{[j]}, \bar{y}_{(j)}\right)$$
$$-2\alpha R_{FG}\operatorname{cov}\left(\bar{x}_{[j]}, \bar{y}_{(j)}\right) + 2\alpha R_{FG} B_i V\left(\bar{x}_{[j]}\right) \tag{13}$$

where $R_{FG} = \dfrac{F\bar{Y}}{F\bar{X}+G}$ and B_i robust betas calculated with Huber M, LMS, Huber MM, S, LAD or LTS of population.

We can get MSEs of estimators given in Equation6-10 using Equation12 easily putting related expectations and suitable F and G values of each design. The R_{FG}s for the estimators in Equation6-10 can be given as $R_{FG1} = \frac{\bar{Y}}{\bar{X}}$, $R_{FG2} = \frac{\bar{Y}}{\bar{X}+C_x}$, $R_{FG3} = \frac{\bar{Y}}{\bar{X}+\beta_{2(x)}}$, $R_{FG4} = \frac{\bar{Y}}{\bar{X}+q_{1(x)}}$, $R_{FG5} = \frac{\bar{Y}}{\bar{X}+q_{3(x)}}$ respectively.

4. Proposed Robust Regression type Estimators in RSS, ERSS and QRSS

We can define the following estimators for the population mean of the study variable in RSS, ERSS and QRSS design as follows

$$\bar{y}_{E(j)} = \left[\bar{y}_{(j)} + b_{i(j)}\left(\bar{X} - \bar{x}_{[j]}\right)\right]\exp\left(\frac{\bar{X} - \bar{x}_{[j]}}{\bar{X} + 2F + \bar{x}_{[j]}}\right) \tag{14}$$

where F represents the coefficient of variation, kurtosis and quarters $C_x, \beta_{(x)}, q_{1(x)}$, $q_{3(x)}$ or any known population information of auxiliary variable. j represents the sampling design such as RSS, ERSS and QRSS and $b_{i(j)}$ is robust regression coefficient as defined in Section3. For particulars about all these robust regression methods, researchers are referred to Koyuncu and Al-Omari [4]. We have generated some members of $\bar{y}_{E(j)}$ as $\bar{y}_{E(j)1}$-$\bar{y}_{E(j)5}$ setting F=1, C_x,$\beta_{(x)}$, $q_{1(x)}$ and $q_{3(x)}$ respectively in Table2-Table4 under (j) design.

The MSE of $\bar{y}_{E(j)}$ is given by

$$MSE\left(\bar{y}_{E(j)}\right) = V\left(\bar{y}_{(j)}\right) + B_i^2 V\left(\bar{x}_{[j]}\right)$$
$$+\frac{1}{4}R_{Fi}^2 V\left(\bar{x}_{[j]}\right) - 2B_i \operatorname{cov}\left(\bar{x}_{[j]}, \bar{y}_{(j)}\right) - R_{Fi} \operatorname{cov}\left(\bar{x}_{[j]}, \bar{y}_{(j)}\right) + R_{Fi} B_i V\left(\bar{x}_{[j]}\right) \quad (15)$$

where $R_{Fi} = \frac{\bar{Y}}{\bar{X}+Fi}$, B_i is robust regression betas using i^{th} robust method, (j) represents RSS, ERSS and QRSS designs. One can easily obtain the spesific MSE from Eq.11-12 putting expectation terms belong to design.

5. Numerical Study

If a dataset contains outlying observations, classical methods can be affected by outliers. To obtain more reliable results in the estimation, different diagnostic methods and robust tools are used to determine the effect of these observations on the predictions. With robust methods, estimates that are insensitive to the effects of outliers and extreme values, can be obtained with little or no sensitivity. Moving in this direction, in this study, we considered robust methods for the estimation of population mean. To see the performance of robust regression type estimators of the population mean under RSS, ERSS and QRSS sampling designs, a numerical study is considered. A real data is used to observe the performances of the estimators concerning BMI as a study variable and the weight as an auxiliary variable for 800 people in Turkey in 2014. In Table 1, the summary of population information about BMI (Y) and weight (X) variables are given.

Table 1. Population information about Body Mass Index (Y) and Weight (X) variables

N = 800	$\bar{Y} = 23.776$
$\bar{X} = 67.558$	$C_x = 0.2047$
$\rho = 0.8674$	$C_y = 0.1763$
$q_{1(x)} = 56$	$q_{3(x)} = 78$
$\beta_{2(x)} = 0.2318$	$R = 0.3519$
$\S_x^2 = 191.295$	$\S_y^2 = 17.5804$

The scatter plot of BMI data is given in Figure1. As seen in Figure1, the data are not normally distributed and it is observed that some observations in the dataset are outliers. For this reason, the use of robust methods is found appropriate for this dataset. For application we have assumed that r=9 set, m=10 cycle, n=m*r=90 sample size and calculated theoretical MSE for each design using Equations 13 and 15.

The MSE and PRE of Koyuncu and Al-Omari [4] and the proposed estimators have been calculated under RSS and the results are given in Table 2. The MSE and PRE of Koyuncu and Al-Omari [4], Zaman and Bulut [14] adapted estimators and the proposed estimators for ERSS and QRSS designs are given in Table 3 and Table 4, respectively.

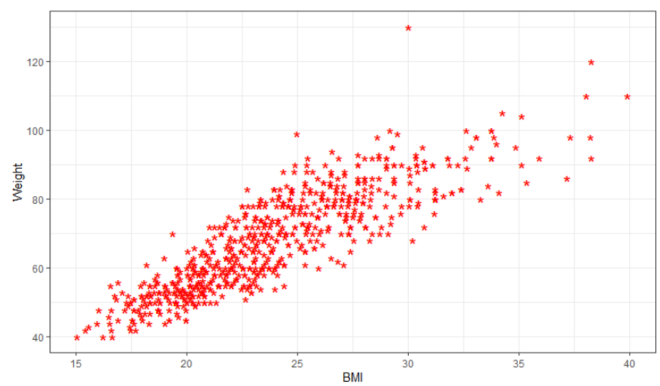

FIGURE 1. Scatter plot of BMI data

The numerical study can be summarized as follows:
The highest PRE values of Koyuncu and Al-Omari [4] and proposed estimators under RSS design are 132.55 and 316.81 respectively (see Table 2). From these values we can say that, for all estimators under RSS design, the best estimator is $\bar{y}_{(RSS)3}$ suggested estimator that used kurtosis of auxiliary variable and LMS robust beta. So, it is concluded that this proposed estimator is approximately three times more effective than other estimators.
The highest PRE values of adapted estimators of Zaman and Bulut [13] and Koyuncu and Al-Omari [4] and proposed estimators under ERSS design are 125.23; 132.47 and 316.70 respectively (see Table 3). From these values we can say that, for all estimators under ERSS design, the best estimator is $\bar{y}_{E(RSS)3}$ suggested estimator that used kurtosis of auxiliary variable and LMS robust beta. So, it is concluded that this proposed estimator is approximately three times more effective than other estimators.
The highest PRE values of adopted estimators of Zaman and Bulut [13] and Koyuncu and Al-Omari [4] and proposed estimators under QRSS design are 125.57; 133.15 and 322.25 respectively (see Table 4). From these values we can say that ,for all estimators under QRSS design, the best estimator is $\bar{y}_{Q(RSS)3}$ suggested estimator that used kurtosis of auxiliary variable and LMS robust beta. So, it was concluded

that this proposed estimator is approximately three times more effective than other estimators. In conclusion, QRSS have the best performance of all proposed estimators in other set sampling designs and LMS have the best performance of all robust methods.

6. Conclusion

We considered robust methods for robust-RTE for mean estimation in RSS, ERSS and QRSS. Firstly, recent proposed robust estimators have been examined. Then, theoretical results for different sampling designs RSS, ERSS and QRSS have been extended. A new exponential-robust- RTE of population mean is proposed and MSEs and PREs of the robust regression type estimators are also obtained for each designs. The existing estimators and proposed estimators have been compared. In conclusion, the suggested estimators perform better than present Zaman and Bulut [14] and Koyuncu and Al-Omari [4] estimators. Also, we demonstrated that the suggested estimator is more effective than adapted estimators of Zaman and Bulut [14] and Koyuncu and Al-Omari [4] in ERSS and QRSS. To see the performance of proposed estimators, we have carried out a numerical study applying on a real data set. When the results of the study are examined, the findings are summarized as follows. The estimators suggested based on the robust methods under RSS designs have better performance over SRS. Also, according to the results obtained from the numerical study, the best method among ranked set sampling methods is QRSS method and it is concluded that the best method among robust methods is LMS. In the light of these results, we desire to develop new estimators in other RSS methods in oncoming studies.

Author Contribution Statements All authors contributed equally and significantly in writing this article. All authors read and approved the final manuscript.

Declaration of Competing Interests All authors declare no conflicts of interest in this paper.

Acknowledgements The authors are thankful to the referees for making valuable suggestions leading to the better presentations of this paper.

References

[1] Ali, N., Ahmad, I., Hanif, M., Shahzad, U., Robust-regression-type estimators for improving mean estimation of sensitive variables by using auxiliary information, *Communications in Statistics-Theory and Methods,* 3 (2019), 385–390. https://doi.org/10.1080/03610926.2019.1645857

[2] Bouza, C. N., Ranked set sampling for the product estimator, *Rev. Invest. Oper,* 29(3) (2008), 201–206.

[3] Koyuncu, N., Regression estimators in ranked set, median ranked set and neoteric ranked set sampling, *Pakistan Journal of Statistics and Operation Research,* 14(1) (2018), 89–94. https://doi.org/10.18187/pjsor.v14i1.1825

[4] Koyuncu, N., Al-Omari, A. I., Generalized robust-regression-type estimators under different ranked set sampling, *Mathematical Sciences,* (2020). https://doi.org/10.1007/s40096-020-00360-7

[5] Long, C., Chen, W., Yang, R., Yao D., Ratio estimation of the population mean using auxiliary under the optimal sampling design, *Probability in the Engineering and Informational Sciences,* 3 (2022), 449–460. https://doi.org/10.1017/s0269964820000625

[6] Mclntyre, G. A., A method for unbiased selective sampling, using ranked sets, *Australian Journal of Agricultural Research,* 3 (1952), 385–390. https://doi.org/10.1071/ar9520385

[7] Mehta, N., Mandowara, V. A., A modified ratio-cum-product estimator of finite population mean using ranked set sampling, *Communication in Statistics- Theory and Methods,* 45(2) (2016), 267–276. https://doi.org/10.1080/03610926.2013.830748

[8] Muttlak, H. A., Investigating the use of quartile ranked set samples for estimating the population mean, *Applied Mathematics and Computation,* 146 (2003), 437–443. https://doi.org/10.1016/s0096-3003(02)00595-7

[9] Samawi, H. M., Muttlak, H. A., Estimation of ratio using rank set sampling, *Biometrical Journal,* 38 (1996), 753-764. https://doi.org/10.1002/bimj.4710380616

[10] Samawi, H. M., Ahmed, M. S.,Abu-Dayyeh, W., Estimating the population mean using extreme ranked set sampling, *Biometrical Journal,* 38(5) (1996), 577–586. https://doi.org/10.1002/bimj.4710380506

[11] Shahzad, U., Hanif, M., Koyuncu, N., Luengo, A. V. G., A regression type estimator for mean estimation under ranked set sampling alongside the sensitivity issue, *Communications Faculty of Sciences University of Ankara Series A1 Mathematics and Statistics*, 68(2) (2019), 2037–2049. https://doi.org/10.31801/cfsuasmas.586057

[12] Shahzad, U., Al-Noor, N., H., Hanif, M., Sajjad, I., Muhammad, A. M., Imputation based mean estimators in case of missing data utilizing robust regression and variance–covariance matrices, *Communications in Statistics-Simulation and Computation,* (2020), 1–20. https://doi.org/10.1080/03610918.2020.1740266

[13] Subzar, M., Bouza, C. N., Al-Omari, A. I.,Utilization of different robust regression techniques for estimation of finite population mean in srswor in case of presence of outliers through ratio method of estimation, *Communications Faculty of Sciences University of Ankara Series A1 Mathematics and Statistics*, 40(5) (2019), 600–609. https://doi.org/10.32604/cmc.2020.010230

[14] Zaman, T., Bulut, H., Modified ratio estimators using robust regression methods, *Communications in Statistics-Theory and Methods,* 48(8) (2019), 2019–2048. https://doi.org/10.1080/03610926.2018.1441419

Appendix

TABLE 2. MSE and PRE of Koyuncu and Al-Omari (2020) and proposed estimators under RSS

	MSE							PRE						
Estimators		Robust betas							Robust betas					
Koyuncu and Al-Omari (2020)	Linear	LTS	LAD	Huber	LMS	S	MM	Linear	LTS	LAD	Huber	LMS	S	MM
$\bar{y}_{N(RSS)1}$	2.454	2.676	2.490	2.502	1.960	2.5986	2.508	100	91.70	98.55	98.09	**125.18***	94.45	97.83
$\bar{y}_{N(RSS)2}$	2.440	2.662	2.476	2.488	1.949	2.584	2.495	100	91.69	98.55	98.08	**125.25***	94.46	97.83
$\bar{y}_{N(RSS)3}$	2.490	2.714	2.527	2.539	1.992	2.636	2.545	100	91.75	98.56	98.10	**125.02***	94.49	97.84
$\bar{y}_{N(RSS)4}$	0.920	1.033	0.938	0.944	0.694	0.9931	0.947	100	89.12	98.09	97.49	**132.55****	92.72	97.16
$\bar{y}_{N(RSS)5}$	0.753	0.845	0.768	0.772	0.576	0.8125	0.775	100	89.06	98.10	97.49	**130.78***	92.71	97.16
Proposed	**Linear**	**LTS**	**LAD**	**Huber**	**LMS**	**S**	**MM**	**Linear**	**LTS**	**LAD**	**Huber**	**LMS**	**S**	**MM**
$\bar{y}_{E(RSS)1}$	0.821	0.922	0.839	0.843	0.623	0.886	0.887	298.71	290.02	297.29	296.30	**314.58***	293.10	282.98
$\bar{y}_{E(RSS)2}$	0.818	0.919	0.834	0.839	0.621	0.883	0.842	298.14	289.53	296.74	296.30	**313.73***	292.55	296.05
$\bar{y}_{E(RSS)3}$	0.829	0.931	0.845	0.851	0.629	0.895	0.854	300.20	291.35	298.77	298.30	**316.81****	294.44	298.06
$\bar{y}_{E(RSS)4}$	0.499	0.545	0.506	0.508	0.434	0.528	0.509	184.53	**189.47***	185.55	185.85	159.92	188.09	186.01
$\bar{y}_{E(RSS)5}$	0.468	0.504	0.473	0.475	0.428	0.490	0.476	160.89	**167.65***	162.21	162.61	134.58	165.58	162.81

*demonstrates the most effective estimators with respect to robust methods

** demonstrates the most effective estimators with respect to all estimators and robust methods

TABLE 3. MSE and PRE of adapted estimators of Zaman and Bulut (2019), Koyuncu and Al-Omari (2020) and proposed estimators under ERSS

	MSE							PRE						
Estimators		Robust betas							Robust betas					
Adapted Estimators of Zaman and Bulut (2019)	**Linear**	**LTS**	**LAD**	**Huber**	**LMS**	**S**	**MM**	**Linear**	**LTS**	**LAD**	**Huber**	**LMS**	**S**	**MM**
$\bar{y}_{EN(ERSS)1}$	2.444	2.666	2.481	2.492	1.952	2.588	2.499	100	91.69	98.54	98.09	**125.23***	94.45	97.83
$\bar{y}_{EN(ERSS)2}$	2.431	2.652	2.467	2.478	1.940	2.574	2.485	100	91.67	98.54	98.08	**125.29***	94.44	97.83
$\bar{y}_{EN(ERSS)3}$	2.481	2.704	2.517	2.529	1.984	2.626	2.535	100	91.74	98.55	98.1	**125.06***	94.48	97.84
Adapted Estimators of Koyuncu and Al-Omari(2020)	**Linear**	**LTS**	**LAD**	**Huber**	**LMS**	**S**	**MM**	**Linear**	**LTS**	**LAD**	**Huber**	**LMS**	**S**	**MM**
$\bar{y}_{EN(ERSS)4}$	0.916	1.028	0.934	0.940	0.692	0.988	0.943	100	89.11	98.09	97.49	**132.47***	92.72	97.15
$\bar{y}_{EN(ERSS)5}$	0.750	0.842	0.764	0.769	0.574	0.809	0.772	100	89.09	98.12	97.5	**130.61***	92.73	97.17
Proposed	**Linear**	**LTS**	**LAD**	**Huber**	**LMS**	**S**	**MM**	**Linear**	**LTS**	**LAD**	**Huber**	**LMS**	**S**	**MM**
$\bar{y}_{E(ERSS)1}$	0.818	0.918	0.834	0.839	0.621	0.882	0.842	287.38	290.37	297.6	297.14	**314.48***	293.4	296.89
$\bar{y}_{E(ERSS)2}$	0.815	0.915	0.831	0.836	0.619	0.879	0.839	286.83	289.86	297.01	296.55	**313.63***	292.87	296.3
$\bar{y}_{an(ERSS)3}$	0.826	0.927	0.842	0.847	0.626	0.891	0.850	288.82	291.69	299.07	298.58	**316.70****	294.78	298.34
$\bar{y}_{E(ERSS)4}$	0.498	0.544	0.505	0.507	0.435	0.527	0.508	180.70	**189.09***	184.98	185.33	158.97	187.6	185.5
$\bar{y}_{E(ERSS)5}$	0.468	0.503	0.473	0.475	0.429	0.490	0.476	158.97	**167.16***	161.58	162.01	133.71	165.07	162.22

*demonstrates the most effective estimators with respect to robust methods

** demonstrates the most effective estimators with respect to all estimators and robust methods

TABLE 4. MSE and PRE of adapted estimators of Zaman and Bulut (2019), Koyuncu and Al-Omari (2020) and proposed estimators under QRSS

	MSE							PRE						
Estimators		**Robust betas**							**Robust betas**					
Adapted Estimators of Zaman and Bulut (2019)	**Linear**	**LTS**	**LAD**	**Huber**	**LMS**	**S**	**MM**	**Linear**	**LTS**	**LAD**	**Huber**	**LMS**	**S**	**MM**
$\bar{y}_{EN(QRSS)1}$	2.396	2.616	2.432	2.444	1.910	2.539	2.450	100	91.62	98.53	98.06	**125.50***	94.40	97.81
$\bar{y}_{EN(QRSS)2}$	2.383	2.601	2.419	2.430	1.898	2.525	2.436	100	91.6	98.53	98.07	**125.57***	94.39	97.81
$\bar{y}_{EN(QRSS)3}$	2.432	2.653	2.468	2.480	1.941	2.576	2.486	100	91.67	98.54	98.08	**125.33***	94.43	97.82
Adapted Estimators of Koyuncu and Al-Omari(2020)	**Linear**	**LTS**	**LAD**	**Huber**	**LMS**	**S**	**MM**	**Linear**	**LTS**	**LAD**	**Huber**	**LMS**	**S**	**MM**
$\bar{y}_{EN(QRSS)4}$	0.887	0.997	0.905	0.910	0.666	0.958	0.913	100	98.05	97.44	**133.15***	92.6	97.10	98.05
$\bar{y}_{EN(QRSS)5}$	0.723	0.814	0.738	0.742	0.551	0.781	0.745	100	98.07	97.45	**131.27***	92.6	97.13	98.07
Proposed	**Linear**	**LTS**	**LAD**	**Huber**	**LMS**	**S**	**MM**	**Linear**	**LTS**	**LAD**	**Huber**	**LMS**	**S**	**MM**
$\bar{y}_{E(QRSS)1}$	0.790	0.889	0.806	0.811	0.597	0.854	0.814	303.3	294.22	301.86	301.36	**319.96***	297.43	301.11
$\bar{y}_{E(QRSS)2}$	0.787	0.886	0.803	0.808	0.595	0.850	0.811	302.76	293.71	301.3	300.82	**319.12***	296.87	300.53
$\bar{y}_{E(QRSS)3}$	0.798	0.898	0.814	0.819	0.602	0.862	0.822	304.83	295.53	303.33	302.83	**322.25****	298.79	302.56
$\bar{y}_{E(QRSS)4}$	0.477	0.521	0.484	0.486	0.417	0.505	0.487	185.97	**191.24***	187.07	187.4	159.85	189.78	187.56
$\bar{y}_{E(QRSS)5}$	0.448	0.482	0.453	0.454	0.412	0.469	0.455	161.56	**168.73***	162.91	163.37	133.84	166.55	163.56

*demonstrates the most effective estimators with respect to robust methods

** demonstrates the most effective estimators with respect to all estimators and robust methods