Robust calibration for estimating the population mean using stratified random sampling

Tolga Zaman[1]* , Hasan Bulut[2]

[1]Çankırı Karatekin University, Faculty of Science, Department of Statistics, 18100 Çankırı, Turkey

[2]Ondokuz Mayıs University, Faculty of Science, Department of Statistics, 55139 Samsun, Turkey

Email Corresponding Author: zamantolga@gmail.com (Zaman, T). Cell: 903762189540

Email: hasan.bulut@omu.edu.tr (Bulut, H).

Abstract

Estimating the population mean is of prime concern in many studies, and calibrations are popular choices. A robust calibration estimator estimates the mean using the minimum covariance determinant (MCD) and the minimum volume ellipsoid (MVE) estimations under stratified random sampling. Efficiency comparisons have been made between the robust calibration estimator and classical calibration estimator. Simulations and empirical results show that the proposed robust calibration estimator has a lower mean square error than the calibration estimators. When the relative efficiency and computation times are considered together, it is seen that the proposed robust calibration estimators based on MCD estimates are more efficient.

Keywords: Calibration estimator; Robust estimates; Mean square error; Efficiency; Stratified random sampling

2010 Mathematics Subject Classification: 62D05; 62G05

1 Introduction

The topic of ratio estimators has lost its practical appeal since the emergence of calibration techniques. In the sampling literature, the calibration technique is commonly used to improve the population parameter estimates. Deville and Sarndal [1] introduced the calibration method of estimation. Tracy et al. [2] provided the calibration estimator for the stratified random sampling estimator. Kim and Park [3] introduced the asymptotic properties of the class of calibration estimators. They found conditions for the design consistency of the calibration estimator. Koyuncu and Kadilar [4] provided novel calibration estimators under stratified random sampling. Climate [5] proposed a calibration estimator for estimating the population mean utilizing calibration weights under stratified double sampling. Ozgul [6] provided a calibration approach alternative to existing calibration estimators for estimating $\overline{Y}$ using an auxiliary variable in stratified sampling. Ozgul [7] proposed a calibration estimator for the population mean using two auxiliary variables in stratified sampling. The calibration approach has three main advantages in sampling theory. Ozgul [8] presented a novel calibration estimator for the $\bar{Y}$ with developing calibration weights based upon two auxiliary variables in stratified two-phase sampling. Shahzad et al. [9] provided a class of calibration estimators for the variance estimation in stratified random sampling. Shahzad et al. [10] presented some calibration estimators using L-Moments for double-stratified random sampling to estimate the population variance. First, the calibration approach leads to consistent estimates. Second, it ensures an important technical family for the efficient combination of datasets. Third, the calibration approach has the computing advantage to calculate estimates [11]. Garg and Pachori [12] introduced a new calibration estimator for estimating the population mean using the known coefficient of variation of the auxiliary variable in the stratified random sampling. However, the outlier problem, which is the presence of extreme values in data, generally decreases the efficiency since traditional estimators are sensitive to these extreme values. Keeping this fact in mind, many authors such as Kadilar et al. [13], Ali et al. [14], Zaman and Bulut [15], Zaman [16], Shahzad et al. [17], Abid et al. [18], Naz et al. [19], Zaman and Bulut [20], Zaman et al. [21] Grover and Kaur [22], and Zaman and Bulut [23] have used the robust regression methods, for minimizing the impact of the extreme values in ratio estimators of the mean. In sampling theory, information about the auxiliary variable used in rate estimators is used to increase the efficiency of the population mean of the study variable. However, in the case of outliers in the data structure, these estimators are negatively affected, and their efficiency decreases. To eliminate this negative effect caused by outliers, robust methods that are resistant to outliers are used [13]. Therefore, this article uses MCD and MVE estimates in calibration estimators to reduce the negative effect caused by an outlier data set. In addition, this study is the first study of calibration estimators using MCD and MVE robust estimates.

The main purpose of this study is to propose robust calibration estimators utilizing robust MCD and MVE covariance estimates to improve the precision of estimation of population mean in stratified sampling. In this study, we improve the robust calibration to decrease MSE in stratified random sampling. We express the MSE up to the first-order approximation and compare the efficiency of the robust calibration estimators with that of the existing estimator. Finally, we find a significantly lower MSE.

The remaining part of the study is organized as follows. Section 2 ensures a description of the classical calibration estimators. The structure of the robust calibration estimator based on MCD and MVE estimates is given in Section 3. The efficiency comparisons of the robust calibration estimators with the classical calibration estimators are presented in Section 4. Sections 5 and 6 consist of empirical and simulation studies of proposed robust calibration estimators. Finally, Section 7 summarizes the findings of this study.

2 Calibration Estimators

Kim et al. [11] extended Sisodia and Dwivedi's [24] estimator, Singh and Kakran's [25] estimator, and Upadhyaya and Singh's [26] estimator to several calibration ratio estimators for improving variance estimator with the aid of auxiliary information in stratified random sampling.

$$\bar{y}_{SD}^{*}=\left(\frac{\sum_{j=1}^{K} w_j \bar{y}_j}{\sum_{j=1}^{K} w_j\left(\bar{x}_j+C_{xj}\right)}\right)\sum_{j=1}^{K} w_j\left(\bar{X}_j+C_{xj}\right), \tag{1}$$

$$\bar{y}_{SK}^{*}=\left(\frac{\sum_{j=1}^{K} w_j \bar{y}_j}{\sum_{j=1}^{K} w_j\left(\bar{x}_j+\beta_{2j}(x)\right)}\right)\sum_{j=1}^{K} w_j\left(\bar{X}_j+\beta_{2j}(x)\right), \tag{2}$$

$$\bar{y}_{US1}^{*}=\left(\frac{\sum_{j=1}^{K} w_j \bar{y}_j}{\sum_{j=1}^{K} w_j\left(\bar{x}_j\beta_{2j}(x)+C_{xj}\right)}\right)\sum_{j=1}^{K} w_j\left(\bar{X}_j\beta_{2j}(x)+C_{xj}\right), \tag{3}$$

$$\bar{y}_{US2}^{*}=\left(\frac{\sum_{j=1}^{K} w_j \bar{y}_j}{\sum_{j=1}^{K} w_j\left(\bar{x}_jC_{xj}+\beta_{2j}(x)\right)}\right)\sum_{j=1}^{K} w_j\left(\bar{X}_jC_{xj}+\beta_{2j}(x)\right), \tag{4}$$

where $C_{xj}=\frac{S_{xj}}{\bar{X}_j}$ denotes the population coefficient of variation of the auxiliary variable in stratum j and $\beta_{2j}(x)=\frac{\sum_{i=1}^{N}\left(x_{ji}-\bar{X}_j\right)^4/N-1}{\left(\sum_{i=1}^{N}\left(x_{ji}-\bar{X}_j\right)^2/N-1\right)^2}$ denotes the population coefficient of the kurtosis of the auxiliary variable in stratum j; $\bar{y}_j$ and $\bar{x}_j$ denote the sample means of the study and auxiliary variable in stratum j, respectively, and it is assumed that the population mean $\bar{X}=\sum_{j=1}^{K} w_j\bar{X}_j$ of the auxiliary variable x is known.

The expression for MSE of the estimators Equations (1)-(4) computed utilizing a first-degree approximation of the Taylor series expansion and is as follows:

$$MSE_c\left(\bar{y}_{SD}^*\right)=\left(\frac{\bar{X}+C_x}{\bar{x}_{st}}\right)^2\sum_{j=1}^{K}\frac{w_j\left(1-f_j\right)}{n_j}s_{ejSD}^2, \tag{5}$$

where $s_{ejSD}^2=\left(n_j-1\right)^{-1}\sum_{i=1}^{n_j}\hat{e}_{jiSD}$ is the jth stratum sample variance, $\hat{e}_{jiSD}=y_{ji}-\bar{y}_j-b_{SD}\left(x_{ji}-\bar{x}_j\right)$ and $b_{SD}=\left(\sum_{j=1}^{K}w_jq_j\bar{y}_j\bar{x}_j/\sum_{j=1}^{K}w_jq_j\bar{x}_j^2\right)$, $q_{jSD}=\left(\bar{x}_j+C_{xj}\right)^{-1}$.

$$MSE_c\left(\bar{y}_{SK}^*\right)=\left(\frac{\bar{X}+\beta_2\left(x\right)}{\bar{x}_{st}}\right)^2\sum_{j=1}^{K}\frac{w_j\left(1-f_j\right)}{n_j}s_{ejSK}^2, \tag{6}$$

where $q_{jSK}=\left(\bar{x}_j+\beta_{2j}\left(x\right)\right)^{-1}$.

$$MSE_c\left(\bar{y}_{US1}^*\right)=\left(\frac{\bar{X}\beta_2\left(x\right)+C_x}{\bar{x}_{st}}\right)^2\sum_{j=1}^{K}\frac{w_j\left(1-f_j\right)}{n_j}s_{ejUS1}^2, \tag{7}$$

where $q_{jUS1}=\left(\bar{x}_j\beta_{2j}\left(x\right)+C_{xj}\right)^{-1}$.

$$MSE_c\left(\bar{y}_{US2}^*\right)=\left(\frac{\bar{X}C_x+\beta_2\left(x\right)}{\bar{x}_{st}}\right)^2\sum_{j=1}^{K}\frac{w_j\left(1-f_j\right)}{n_j}s_{ejUS2}^2, \tag{8}$$

where $q_{jUS2}=\left(\bar{x}_jC_{xj}+\beta_{2j}\left(x\right)\right)^{-1}$. For the $j\,th$ stratum, $w_j=N_j/N$ denotes the stratum weight and $f_j=n_j/N_j$ denotes the sample fraction.

3 MCD AND MVE Robust Covariance Estimates

Rousseeuw [27] introduced two estimators, namely the Minimum Volume Ellipsoid (MVE) and the Minimum Covariance Determinant (MCD) estimators. These estimators have a high breaking point. MCD and MVE robust estimators perform better than classical covariance estimators when there are outliers in the data. These estimators are based on iterative algorithms.

The MCD algorithm searches for a subset consisting of h observations such that the determinant of the covariance matrix is the minimum. Here, $h=(1-\alpha)*n$, and α is the trimming ratio. When this subset has the minimum determinant of the covariance matrix, the location and scatter parameters are calculated robustly as the sample mean vector and sample covariance matrix of this subset. Because of the definition, the algorithm's calculations may take a long time. So, Rousseeuw and Van Driessen [28] constructed the computationally efficient FastMCD algorithm. Finally, the breakdown point of MCD estimators is equal to the trimming ratio (α) [29, 30].

On the other hand, Rousseeuw [27] provided the MVE estimate to determine the unusual values in multivariate datasets. The MVE estimator for the location parameter of multivariate data was defined as the centre of an ellipsoid having a minimum volume spanned by the h points in $\boldsymbol{X}$ data, where $h = n/2 + 1$ and $.$ is a function that rounds the number to an integer [27]. The breakdown point of the MVE estimator is equal to $((\lfloor n/2 \rfloor - p + 1))/n$. And, when $n \to \infty$, the breakdown point of the MVE estimator is equal to 50% [27,31, 32].

There is a large body of literature on rate-type estimators for estimating the mean using MCD and MVE estimates. For example, Bulut and Zaman [33] extended Zaman and Bulut [15] using MCD estimates. Zaman and Bulut [32] provided ratio-type estimators for mean estimation using MCD and MVE estimates to stratified random sampling. Shahzad et al. [34] provided ratio estimators to estimate the mean using MCD and MVE estimates in case of missing data.

We use the CovMcd, and CovMve functions in the package rrcov in the R programming language for calculations belonging to MCD and MVE estimations, respectively [35].

4 Robust Calibration of Ratio Estimators

For the estimation of the population mean, we suggest the following eight robust calibration estimators using minimum covariance determinant (MCD) and minimum volume ellipsoid (MVE) estimators, instead of ratio estimators presented in Equations (1)-(4), to data which have outliers. We have considerable decline variance utilizing estimators $\bar{y}^{*}_{rSD(z)}$ and $\bar{y}^{*}_{rSK(z)}$, $\bar{y}^{*}_{rUS1(z)}$, and $\bar{y}^{*}_{rUS2(z)}$, $z = MCD$ and MVE, compare to the estimators $\bar{y}^{*}_{SD}$ and $\bar{y}^{*}_{SK}$, $\bar{y}^{*}_{US1}$, and $\bar{y}^{*}_{US2}$.

$$\bar{y}^{*}_{rSD(z)} = \left(\frac{\sum_{j=1}^{K} w_j \bar{y}_{j(z)}}{\sum_{j=1}^{K} w_j \left(\bar{x}_{j(z)} + C_{xj(z)} \right)} \right) \sum_{j=1}^{K} w_j \left(\bar{X}_{j(z)} + C_{xj(z)} \right), \tag{9}$$

$$\bar{y}^{*}_{rSK(z)} = \left(\frac{\sum_{j=1}^{K} w_j \bar{y}_{j(z)}}{\sum_{j=1}^{K} w_j \left(\bar{x}_{j(z)} + \beta_{2j(z)}(x) \right)} \right) \sum_{j=1}^{K} w_j \left(\bar{X}_{j(z)} + \beta_{2j(z)}(x) \right), \tag{10}$$

$$\bar{y}^{*}_{rUS1(z)} = \left(\frac{\sum_{j=1}^{K} w_j \bar{y}_{j(z)}}{\sum_{j=1}^{K} w_j \left(\bar{x}_j \beta_{2j(z)}(x) + C_{xj(z)} \right)} \right) \sum_{j=1}^{K} w_j \left(\bar{X}_{j(z)} \beta_{2j(z)}(x) + C_{xj(z)} \right), \tag{11}$$

$$\bar{y}^{*}_{rUS2(z)} = \left(\frac{\sum_{j=1}^{K} w_j \bar{y}_{j(z)}}{\sum_{j=1}^{K} w_j \left(\bar{x}_{j(z)} C_{xj(z)} + \beta_{2j(z)}(x) \right)} \right) \sum_{j=1}^{K} w_j \left(\bar{X}_{j(z)} C_{xj(z)} + \beta_{2j(z)}(x) \right), \tag{12}$$

where $z = MCD$ and MVE. $\bar{y}_{j(z)}$, $\bar{x}_{j(z)}$, $C_{xj(z)}$, $\beta_{2j(z)}(x)$ are obtained by utilizing MCD and MVE estimators in stratum j, respectively.

The MSE expression of the estimators given between Equations (9)-(12) are as follows:

$$MSE_c\left(\bar{y}^{*}_{rSD(z)}\right) = \left(\frac{\bar{X}_{(z)} + C_{x(z)}}{\bar{x}_{st(z)}} \right)^2 \sum_{j=1}^{K} \frac{w_j \left(1 - f_j\right)}{n_j} s^2_{ejDS(z)}, \tag{13}$$

where $s^2_{ejSD(z)} = \left(n_j - 1\right)^{-1} \sum_{i=1}^{n_j} \hat{e}_{jiSD(z)}$ is the jth stratum sample variance, $\hat{e}_{jiSD(z)} = y_{ji} - \bar{y}_{j(z)} - b_{SD(z)}\left(x_{ji} - \bar{x}_{j(z)}\right)$ and $b_{SD(z)} = \left(\sum_{j=1}^{K} w_j q_{j(z)} \bar{y}_{j(z)} \bar{x}_{j(z)} / \sum_{j=1}^{K} w_j q_{j(z)} \bar{x}^2_{j(z)} \right)$,

$q_{jSD(z)} = \left(\bar{x}_{j(z)} + C_{xj(z)} \right)^{-1}$.

$$MSE_c\left(\bar{y}^*_{r\text{SK}(z)}\right)=\left(\frac{\bar{X}_{(z)}+\beta_{2(z)}\left(x\right)}{\bar{x}_{st(z)}}\right)^2\sum_{j=1}^{K}\frac{w_j\left(1-f_j\right)}{n_j}s^2_{ejSK(z)}, \tag{14}$$

where $q_{jSK(z)}=\left(\bar{x}_{j(z)}+\beta_{2j(z)}\left(x\right)\right)^{-1}$.

$$MSE_c\left(\bar{y}^*_{rUS1(z)}\right)=\left(\frac{\bar{X}_{(z)}\beta_{2(z)}\left(x\right)+C_{x(z)}}{\bar{x}_{st(z)}}\right)^2\sum_{j=1}^{K}\frac{w_j\left(1-f_j\right)}{n_j}s^2_{ejUS1(z)}, \tag{15}$$

where $q_{jUS1(z)}=\left(\bar{x}_{j(z)}\beta_{2j(z)}\left(x\right)+C_{xj(z)}\right)^{-1}$.

$$MSE_c\left(\bar{y}^*_{rUS2(z)}\right)=\left(\frac{\bar{X}_{(z)}C_{x(z)}+\beta_{2(z)}\left(x\right)}{\bar{x}_{st(z)}}\right)^2\sum_{j=1}^{K}\frac{w_j\left(1-f_j\right)}{n_j}s^2_{ejUS2(z)}, \tag{16}$$

where $q_{jUS2(z)}=\left(\bar{x}_{j(z)}C_{xj(z)}+\beta_{2j(z)}\left(x\right)\right)^{-1}$.

The expression for MSE of the robust calibration estimators is in the same form as the MSE expression presented in Equations (5)-(8), but it is seen that, $\bar{X}$, C_x , $\beta_2\left(x\right)$, $\bar{x}_{st}$, s^2_{ej} , $\hat{e}_{ji}$, b and q_j in Equations (5)-(8) should be replaced by $\bar{X}_{(z)}$, $C_{x(z)}$, $\beta_{2(z)}\left(x\right)$, $\bar{x}_{st(z)}$, $s^2_{ejt(z)}$, $\hat{e}_{jit(z)}$, $b_{t(z)}$ and $q_{jt(z)}$, $\left(t=SD,SK,US1\, and\, US2\right)$ whose values as computed by MCD and MVE estimates $\left(z=MCD\, and\, MVE\right)$. The expressions of MSE two different robust covariance estimates for each value z will be calculated. Therefore, eight different the expressions for MSE will be obtained.

5 Efficiency Comparisons

In this section, we compare the expressions of MSE of the robust calibration estimators of Equations (13)-(16) with the expressions of MSE of the calibration ratio estimators of Equations (5)-(8).

(i) From Equations (5) and (13),

$$MSE_c\left(\bar{y}^*_{rSD(z)}\right)<MSE_c\left(\bar{y}^*_{SD}\right), \tag{17}$$

$$\left(\frac{\bar{X}_{(z)}+C_{x(z)}}{\bar{x}_{st(z)}}\right)^2 \sum_{j=1}^{K}\frac{w_j\left(1-f_j\right)}{n_j}s_{ejSD(z)}^2 < \left(\frac{\bar{X}+C_x}{\bar{x}_{st}}\right)^2 \sum_{j=1}^{K}\frac{w_j\left(1-f_j\right)}{n_j}s_{ejSD}^2. \quad (18)$$

Let

$A_{(z)}=\dfrac{\bar{X}_{(z)}+C_{x(z)}}{\bar{x}_{st(z)}}$ and $B_{(z)}=\sum_{j=1}^{K}\dfrac{w_j\left(1-f_j\right)}{n_j}s_{ejSD(z)}^2$; $\left(z=MCD\, and\, MVE\right)$. $A=\dfrac{\bar{X}+C_x}{\bar{x}_{st}}$ and

$$B=\sum_{j=1}^{K}\frac{w_j\left(1-f_j\right)}{n_j}s_{ejSD}^2.$$

Thus, Equation (18) becomes

$$A_{(z)}^2 B_{(z)} - A^2 B < 0. \quad (19)$$

When Equation (19) is satisfied, the estimator $\bar{y}_{rSD(z)}^{*}$, $z=MCD\, and\, MVE,$ performs better than the estimator $\bar{y}_{SD}^{*}$.

(ii) From Equations (6) and (14),

$$MSE_c\left(\bar{y}_{rSK(z)}^{*}\right) < MSE_c\left(\bar{y}_{SK}^{*}\right), \quad (20)$$

$$\left(\frac{\bar{X}_{(z)}+\beta_{2(z)}\left(x\right)}{\bar{x}_{st(z)}}\right)^2 \sum_{j=1}^{K}\frac{w_j\left(1-f_j\right)}{n_j}s_{ejSK(z)}^2 < \left(\frac{\bar{X}+\beta_2\left(x\right)}{\bar{x}_{st}}\right)^2 \sum_{j=1}^{K}\frac{w_j\left(1-f_j\right)}{n_j}s_{ejSK}^2. \quad (21)$$

$C_{(z)}=\dfrac{\bar{X}_{(z)}+\beta_{2(z)}\left(x\right)}{\bar{x}_{st(z)}}$ and $D_{(z)}=\sum_{j=1}^{K}\dfrac{w_j\left(1-f_j\right)}{n_j}s_{ejSK(z)}^2$; $\left(z=MCD\, and\, MVE\right)$. $C=\dfrac{\bar{X}+\beta_2\left(x\right)}{\bar{x}_{st}}$

and

$$D=\sum_{j=1}^{K}\frac{w_j\left(1-f_j\right)}{n_j}s_{ejSK}^2.$$

Thus, Equation (21) becomes

$$C_{(z)}^2 D_{(z)} - C^2 D < 0. \tag{22}$$

When Equation (22) is satisfied, the estimator $\bar{y}^*_{rSK(z)}$, $z = MCD\, and\, MVE,$ performs better than the estimator $\bar{y}^*_{SK}$.

(iii) From Equations (7) and (15),

$$MSE_c\left(\bar{y}^*_{rUS1(z)}\right) < MSE_c\left(\bar{y}^*_{US1}\right), \tag{23}$$

$$\left(\frac{\bar{X}_{(z)}\beta_{2(z)}(x) + C_{x(z)}}{\bar{x}_{st(z)}}\right)^2 \sum_{j=1}^{K} \frac{w_j\left(1-f_j\right)}{n_j} s^2_{ejUS1(z)} < \left(\frac{\bar{X}\beta_2(x) + C_x}{\bar{x}_{st}}\right)^2 \sum_{j=1}^{K} \frac{w_j\left(1-f_j\right)}{n_j} s^2_{ejUS1}. \tag{24}$$

$E_{(z)} = \dfrac{\bar{X}_{(z)}\beta_{2(z)}(x) + C_{x(z)}}{\bar{x}_{st(z)}}$ and $F_{(z)} = \sum_{j=1}^{K} \dfrac{w_j\left(1-f_j\right)}{n_j} s^2_{ejUS1(z)}$; $\left(z = MCD\, and\, MVE\right)$.

$E = \dfrac{\bar{X}\beta_2(x) + C_x}{\bar{x}_{st}}$ and $F = \sum_{j=1}^{K} \dfrac{w_j\left(1-f_j\right)}{n_j} s^2_{ejUS1}$.

Thus, Equation (24) becomes

$$E_{(z)}^2 F_{(z)} - E^2 F < 0. \tag{25}$$

When Equation (25) is satisfied, the estimator $\bar{y}^*_{rUS1(z)}$, $z = MCD\, and\, MVE,$ performs better than the estimator $\bar{y}^*_{US1}$.

(iv) From Equations (8) and (16),

$$MSE_c\left(\bar{y}^*_{rUS2(z)}\right) < MSE_c\left(\bar{y}^*_{US2}\right), \tag{26}$$

$$\left(\frac{\bar{X}_{(z)}C_{x(z)} + \beta_{2(z)}(x)}{\bar{x}_{st(z)}}\right)^2 \sum_{j=1}^{K} \frac{w_j\left(1-f_j\right)}{n_j} s^2_{ejUS2(z)} < \left(\frac{\bar{X}C_x + \beta_2(x)}{\bar{x}_{st}}\right)^2 \sum_{j=1}^{K} \frac{w_j\left(1-f_j\right)}{n_j} s^2_{ejUS2}. \tag{27}$$

$$G_{(z)} = \frac{\bar{X}_{(z)} C_{x(z)} + \beta_{2(z)}(x)}{\bar{x}_{st(z)}} \quad \text{and} \quad H_{(z)} = \sum_{j=1}^{K} \frac{w_j (1 - f_j)}{n_j} s^2_{ejUS2(z)}; \quad (z = MCD \, and \, MVE).$$

$$G = \frac{\bar{X} C_x + \beta_2(x)}{\bar{x}_{st}} \text{ and } H = \sum_{j=1}^{K} \frac{w_j (1 - f_j)}{n_j} s^2_{ejUS2}.$$

Thus, Equation (27) becomes

$$G^2_{(z)} H_{(z)} - G^2 H < 0. \quad (28)$$

When the Equation (28) is satisfied, the estimator $\bar{y}^*_{rUS2(z)}$, $z = MCD \, and \, MVE,$ performs better than the estimator $\bar{y}^*_{US2}$. The robust calibration estimators will perform better than the existing estimators if Conditions (i)- (iv) are satisfied.

6 Application

We use the data in Murthy [36] to compare the efficiencies between the robust calibration estimators and the classical calibration estimators under stratified random sampling. We have contaminated these datasets by multiplying with ten the Y value of the last five observations in each stratified. The contamination rates (ξ_i) have been given in Table 1 for population-1 and Table 2 for population-2. Therefore, datasets are contaminated in the direction of y. These datasets are presented as followings.

Population 1. y is factories in the region; x is the number of workers. The data from 80 factories have been classified arbitrarily into four strata based on x values Murthy [36]. The strata are $x<100$, $100 \le x < 200$, $200 \le x < 500$, and $x \ge 500$, respectively. We have randomly selected samples from each stratum by taking the proportional allocation, $n_h = n \frac{N_h}{N}$, using a total sample size of $n=45$

[Table 1]

Population 2. y is factories in the region; x is fixed capital. The data from 80 factories have been classified arbitrarily into four strata based on x values Murthy [36]. The strata are $x \le 500$,

$500 < x \leq 1000$, $1000 < x \leq 2000$, and $x > 2000$, respectively. Again, we use the same procedure of selecting the samples from each stratum as in Example 1. We use a total sample size of $n = 45$. Both data sets were used by Shabbir and Gupta [37].

[Table 2]

We compute the MSE values of the classical calibration and the robust calibration estimators as defined in Sections 2 and 3, respectively. And using these values obtain the relative efficiency (RE) for each robust calibration estimator in Equations (9)–(12) concerning the classical calibration estimators in Equations (1)–(4) by using Equation (29) as below:

$$RE\left(\bar{y}^*_{rt(z)}\right) = \frac{MSE\left(\bar{y}^*_t\right)}{MSE\left(\bar{y}^*_{rt(z)}\right)}, t = SD, SK, US1 \, and \, US2. \quad (29)$$

In Table 3, 24 relative efficiency values are computed. From Table 3, it is seen that the performances belonging to all of the robust calibration estimators concerning classical calibration estimators are greater than 1. The robust calibration estimators perform better than the calibration estimators. In addition, for both datasets, the most efficient estimator is the estimator $\bar{y}^*_{rUS2(MVE)}$ based on the MVE covariance estimate. This result is expected because the conditions (i)- (iv) are satisfied for all estimators in these datasets.

[Table 3]

7 Simulation

3 different simulation designs have been used to compare the performance of the robust calibration estimators and the classical calibration estimators. These designs are as follows;

Simulation Design-I

Here, the number of strata is taken as four, and each stratum contained $N_1 = N_2 = N_3 = N_4 = 20$ observations, respectively. The observations in each stratum are generated as follows:

$$(1) Y_i = 1 + 2X_1 + \varepsilon_i, \quad (30)$$

$$(2) Y_i = 5 + X_2 + \varepsilon_i, \quad (31)$$

$$(3) Y_i = 10 + 3X_3 + \varepsilon_i, \quad (32)$$

$$(4) Y_i = 20 + 7X_4 + \varepsilon_i. \quad (33)$$

Here, X variables are generated from $X_1 \sim N(0,1)$, $X_2 \sim N(5,1)$, $X_3 \sim N(0,9)$, $X_4 \sim N(5,9)$ distributions. Error terms $\varepsilon_i \sim N(0,1), \varepsilon_i \sim Exp(1), \varepsilon_i \sim U(0,1)$ are generated from 3 different distributions. Sample sizes are 20, 32 and outliers rates are 5%, 10% , and 25%.

Simulation Design-II

Here, the number of strata is taken as four, and each stratum contained $N_1 = N_2 = N_3 = N_4 = 20$ observations, respectively. The observations in each stratum are generated as follows:

$$(1) Y_i = 1 + 2X_1 + \varepsilon_i, \quad (34)$$

$$(2) Y_i = 5 + X_2 + \varepsilon_i, \quad (35)$$

$$(3) Y_i = 10 + 3X_3 + \varepsilon_i, \quad (36)$$

$$(4) Y_i = 20 + 7X_4 + \varepsilon_i. \quad (37)$$

Here, X variables are generated from $X_1 \sim N(0,1)$, $X_2 \sim N(5,1)$, $X_3 \sim N(0,9)$, $X_4 \sim N(5,9)$ distributions. Error terms are generated from $\varepsilon_i \sim N(0,1)$ distributions. Sample sizes are 20, 32 and 40 and outliers rates are 5%, 10%, and 25%.

Simulation Design-III

Here, the observations for all strata are generated from the distribution of $Y_i = 3X_{ki} + \varepsilon_i \ (k = 1,2,3,4)$, and each stratum contained $N_1 = N_2 = N_3 = N_4 = 20$ observations, respectively. The X and error values are obtained as follows, respectively.

$$(1) X_1 \sim N(0,1), \qquad \varepsilon_i \sim Exp(3), \quad (38)$$

$$(2) X_2 \sim N(5,1), \ \varepsilon_i \sim N(0,1), \quad (39)$$

$$(3)\, X_3 \sim N(5,9), \quad \varepsilon_i \sim Unif\,(0,1), \tag{40}$$

$$(4)\, X_4 \sim N(0,9),\ \varepsilon_i \sim t_{10}. \tag{41}$$

The sample sizes are 20, 32, and 40, and outliers rates were 5%, 10%, and 25%. The simulation steps are as follows;

Firstly, existing calibration estimators given in Section 2 are obtained for each sample size using SRSWOR (simple random sampling without replacement).

Then, for each sample taken, the existing calibration estimators, say $\bar{Y}_{\text{i}}$, such as $\bar{y}^{*}_{rt}$, given in Section 2 and the robust calibration estimators, $\bar{y}^{*}_{rt(z)}$, given in Section 3 are obtained.

The values of MSE for all cases are obtained with the help of Equation (42) as below:

$$MSE = \frac{1}{1000}\sum_{i=1}^{1000}\left(\bar{Y}_{\text{i}} - \bar{Y}\right)^2, \tag{42}$$

where $\bar{Y}$ is the population mean. In all cases, 1000 iterations are performed.

[Table 4]

[Table 5]

[Table 6]

The simulation study was performed with R programming language on an Intel(R) Core(TM) i3-4160 CPU with 3.60 GHz. The results obtained for the first, second, and third simulation designs are given in Tables 4, 5, and 6. We compute the computational times (CT) and the relative efficiency values of the robust and existing calibration estimators as given in Sections 2 and 3, respectively, using outliers data. In different simulation designs, the computational times and the relative efficiency values of the robust and existing calibration estimators are given in Tables 4, 5, and 6. These RE values are bigger than 1. It is seen that the robust calibration estimators are more efficient than the existing estimators, which indicates that the robust calibration estimators are more efficient in the presence of outliers. The relative efficiency of the robust calibration estimators concerning the existing estimators in Tables 4, 5, and 6 would increase dramatically, which shows that the performances of the robust calibration estimators would increase if there were more outliers in the dataset. Also, when the computation time and relative efficiency values are considered together, it is seen that estimators based on MCD are more efficient.

8 Conclusion

The study has proposed a set of calibration estimators for a finite population mean utilizing the information on some robust measures. The robust calibration estimator is very attractive and should be preferred in practice. It ensures consistent and more precise parameter estimates than the existing estimators under stratified random sampling, especially in the presence of unusual observations in the data. According to the computation times in the study, it is seen that the computational speeds of the classical estimators are faster than the proposed estimators. However, the MSE values of these estimators are large. That is, their efficiencies are small relative to the proposed estimators. Considering only the relative efficiency values, estimators based on MCD and MVE give more efficient results than classical estimators. However, there are no significant differences between the results of the proposed estimators based on MCD and MVE. Therefore, when the computation times and relative efficiency values are considered together, it is seen that the proposed estimators based on MCD are more efficient.

References

[1] Deville, J.C., Sarndal, C.E. “Calibration estimators in survey sampling”, *Journal of the American Statistical Association*, **87**, pp. 376-382, (1992)

[2] Tracy, D.S., Singh, S., Arnab, R. “Note on calibration in stratified and double sampling”, *Survey Methodology*, **29**, pp. 99–104, (2003).

[3] Kim, J.K., Park, M. “Calibration estimation in survey sampling”, *International Statistical Review*, **78**(1), pp. 21-29, (2010).

[4] Koyuncu, N., and Kadilar, C. “Calibration weighting in stratified random sampling”, *Communications in Statistics-Simulation and Computation*, **45**(7), pp. 2267-2275, (2016)

[5] Clement, E. P. “A note on calibration weightings for stratified double sampling with equal probability”, *Communications in Statistics-Theory and Methods*, **47**(12), pp. 2835-2847, (2018).

[6] Ozgul, N. “New calibration estimator in stratified sampling”, *Journal of Statistical Computation and Simulation*, **88**(13), 2561-2572, (2018).

[7] Ozgul, N. “New calibration estimator based on two auxiliary variables in stratified sampling”, *Communications in Statistics-Theory and Methods*, **48**(6), pp. 1481-1492, (2019).

[8] Ozgul, N. “New improved calibration estimator based on two auxiliary variables in stratified two-phase sampling”, *Journal of Statistical Computation and Simulation*, **In Press**, (2020).

[9] Shahzad, U., Ahmad, I., Almanjahie, I., et al. “Utilizing L-Moments and calibration method to estimate the variance based on COVID-19 data”, *Fresenius Environmental Bulletin*, **30**(7A), pp. 8988-8994, (2021).

[10] Shahzad, U., Ahmad, I., Almanjahie, I., et al. “L-Moments and calibration-based variance estimators under double stratified random sampling scheme: an application of covid-19 pandemic”, *Scientia Iranica*, **In Press**, (2021).

[11] Kim, J. M., Sungur, E. A., and Heo, T. Y. “Calibration approach estimators in stratified sampling”, *Statistics & Probability Letters*, **77**(1), pp. 99-103, (2007).

[12] Garg, N., and Pachori, M. “Use of coefficient of variation in calibration estimation of population mean in stratified sampling”, *Communications in Statistics-Theory and Methods*, **49**(23), pp. 5842-5852, (2020).

[13] Kadılar, C., M. Candan, M., and Çıngı, H. “Ratio estimators using robust regression”, *Hacettepe Journal of Mathematics and Statistics*, **36**(2), pp. 181–88, (2007).

[14] Ali, N., Ahmad, I., Hanif, M., and Shahzad, U. “Robust-regression-type estimators for improving mean estimation of sensitive variables by using auxiliary information”, *Communications in Statistics-Theory and Methods*, **50**(4), pp. 979-992, (2021).

[15] Zaman, T., and Bulut, H. “Modified ratio estimators using robust regression methods”, *Communications in Statistics-Theory and Methods*, **48**(8), pp. 2039-2048, (2019).

[16] Zaman, T. “Improvement of modified ratio estimators using robust regression methods”, *Applied Mathematics and Computation*, **348**, pp. 627-631, (2019).

[17] Shahzad, U., Al-Noor, N. H., Hanif, M., et al. “An exponential family of median based estimators for mean estimation with simple random sampling scheme”, *Communications in Statistics-Theory and Methods*, **In Press**, (2020).

[18] Abid, M., Ahmed, S., Tahir, M., et al. "Improved ratio estimators of variance based on robust measures", *Scientia Iranica*, **26**(4), pp. 2484-2494 (2019).

[19] Naz, F., Abid, M., Nawaz, T., et al. "Enhancing the efficiency of the ratio-type estimators of population variance with a blend of information on robust location measures", *Scientia Iranica*, **27**(4), pp. 2040-2056 (2019).

[20] Zaman, T. and Bulut, H. "An efficient family of robust-type estimators for the population variance in simple and stratified random sampling", *Communications in Statistics-Theory and Methods*, **In Press** (2021).

[21] Zaman, T., Dünder, E., Audu, A., et al. "Robust regression-ratio-type estimators of the mean utilizing two auxiliary variables: A simulation study", *Mathematical Problems in Engineering*, **2021,** pp. 1-9 (2021).

[22] Grover, L. K. and Kaur, A. "An improved regression type estimator of population mean with two auxiliary variables and its variant using robust regression method", *Journal of Computational and Applied Mathematics*, *382*, pp. 1-18 (2021).

[23] Zaman, T. and Bulut, H. "A simulation study: Robust ratio double sampling estimator of finite population mean in the presence of outliers", *Scientia Iranica*, **In Press** (2021).

[24] Sisodia, B. V. S. and Dwivedi, K. V. “A modified ratio estimator using coefficient of variation of auxiliary variable. Journal Indian Society of Agricultural Statistics”, **33**(2), pp. 13–18, (1981).

[25] Singh, H. P. and Kakran, M. S. “A modified ratio estimator using known coefficient of kurtosis of an auxiliary character”, *Journal of Indian Society of Agricultural Statistics,* (1993).

[26] Upadhyaya, L. N. and Singh, H. P. “Use of transformed auxiliary variable in estimating the finite population mean”, *Biometrical Journal*, **41**(5), pp. 627–636, (1999).

[27] Rousseeuw, P. J. "Multivariate estimation with high breakdown point", *Mathematical Statistics and Applications*, pp. 283-297, (1985).

[28] Rousseeuw, P., and Van Driessen, K. "A fast algorithm for the minimum covariance determinant estimator", *Technometrics*, **41**, pp. 212–223, (1999).

[29] Hubert, M., Debruyne, M., and Rousseeuw, P. J. "Minimum covariance determinant and extensions", Wiley Interdisciplinary Reviews: *Computational Statistics,* **10**(3), pp. 1-11, (2018).

[30] Bulut, H. "*Multivariate Statistical Methods with R Applications*", Nobel Academic Publishing, Ankara. ISBN: 978-605-7928-93-1, (2018).

[31] Bulut, H., and Öner, Y. "The evaluation of socio-economic development of development agency regions in Turkey using classical and robust principal component analyses", *Journal of Applied Statistics*, **44**(16), pp. 2936-2948, (2017).

[32] Zaman, T., and Bulut, H. "Modified regression estimators using robust regression methods and covariance matrices in stratified random sampling", *Communications in Statistics-Theory and Methods*, **49**(14), pp. 3407-3420, (2020).

[33] Bulut, H., and Zaman, T. "An improved class of robust ratio estimators by using the minimum covariance determinant estimation", *Communications in Statistics-Simulation and Computation*, **In Press**, (2019).

[34] Shahzad, U., Al-Noor, N. H., Hanif, M., et al. "Imputation based mean estimators in case of missing data utilizing robust regression and variance–covariance matrices", *Communications in Statistics-Simulation and Computation*, **In Press**, (2020).

[35] Todorov, V., and Filzmoser, P. (2009). An object-oriented framework for robust multivariate analysis. Journal of Statistical Software, **32**(13), 1-47, (2009).

[36] Murthy, M. N. "*Sampling theory and methods*", New Delhi, India: Statistical Publishing Society, (1967).

[37] Shabbir, J., and S. Gupta. "A new estimator of population mean in stratified sampling", *Communications in Statistics-Theory and Methods,* **35**(7), pp. 1201–1209, (2006).

Methods	Population	Stratum	1	2	3	4

List of table captions

Classical	C_x	1.98	C_x	1.471	1.442	1.246	1.249
	$\beta_2(x)$	9.20	$\beta_2(x)$	3.154	2.842	1.576	1.654
	$\bar{X}$	1243.41	$\bar{x}$	192.857	513.615	1806.000	3658.000
	$\bar{Y}$	18926.55	$\bar{y}$	7927.000	15926.380	27094.560	33807.890
			s^2_{ejSD}	2.38E+08	3.02E+08	6.41E+07	2.35E+09
			s^2_{ejSK}	2.38E+08	3.02E+08	6.38E+07	2.34E+09
			s^2_{ejUS1}	2.14E+08	3.01E+08	4.68E+07	1.52E+09
			s^2_{ejUS2}	2.30E+08	3.01E+08	5.58E+07	2.07E+09
MCD	C_x	0.94	C_x	0.231	0.209	0.325	0.236
	$\beta_2(x)$	10.47	$\beta_2(x)$	4.828	4.480	3.020	3.089
	$\bar{X}$	114.70	$\bar{x}$	73.625	127.222	329.000	627.000
	$\bar{Y}$	3999.42	$\bar{y}$	3477.500	4523.333	6276.833	7368.500
			s^2_{ejSD}	1.13E+08	8.55E+07	4.87E+08	8.49E+09
			s^2_{ejSK}	1.14E+08	8.89E+07	4.59E+08	8.22E+09
			s^2_{ejUS1}	1.26E+08	1.18E+08	2.71E+08	6.30E+09
			s^2_{ejUS2}	1.16E+08	9.29E+07	4.28E+08	7.92E+09
MVE	C_x	1.26	C_x	0.225	0.203	0.307	0.224
	$\beta_2(x)$	10.61	$\beta_2(x)$	4.828	4.480	3.020	3.089
	$\bar{X}$	191.84	$\bar{x}$	71.778	127.222	1806.000	2746.500
	$\bar{Y}$	4751.16	$\bar{y}$	3414.556	4523.333	27094.556	26471.375
			s^2_{ejSD}	2.14E+08	3.02E+08	2.18E+07	1.50E+09
			s^2_{ejSK}	2.14E+08	3.03E+08	2.24E+07	1.47E+09
			s^2_{ejUS1}	2.11E+08	3.09E+08	2.78E+07	1.23E+09
			s^2_{ejUS2}	2.12E+08	3.07E+08	2.62E+07	1.29E+09
	N	80	N_i	25	23	16	16
	n	45	n_i	14	13	9	9
			ξ_i	0.2	0.21	0.31	0.31
			f_i	0.56	0.5652	0.5625	0.5625
			w_j	0.31	0.29	0.2	0.2

Table 1: Data statistics for Population I

Methods	Population		Stratum	1	2	3	4
Classical	C_x	1.75	C_x	1.348	1.520	1.123	1.166
	$\beta_2(x)$	7.60	$\beta_2(x)$	2.055	4.616	1.300	1.523
	$\bar{X}$	4544.66	$\bar{x}$	1847.545	1067.5	7737.375	16264.38
	$\bar{Y}$	19026.79	$\bar{y}$	14506.46	7411.056	30038.38	46693.13
			s^2_{ejSD}	1.81E+08	2.09E+08	50281479	1.01E+09
			s^2_{ejSK}	1.8E+08	2.09E+08	50199765	1.01E+09
			s^2_{ejUS1}	1.63E+08	2.07E+08	32240548	6.86E+08
			s^2_{ejUS2}	1.73E+08	2.08E+08	41198684	8.73E+08
MCD	C_x	0.72	C_x	0.353	0.174	0.306	0.150
	$\beta_2(x)$	8.87	$\beta_2(x)$	3.602	6.359	2.634	2.907
	$\bar{X}$	608.22	$\bar{x}$	279	707.7692	1409.8	8873
	$\bar{Y}$	4082.87	$\bar{y}$	2550.143	4602.308	6333.4	24369
			s^2_{ejSD}	1.63E+08	2.13E+08	11269514	5.98E+08
			s^2_{ejSK}	1.63E+08	2.13E+08	11380985	5.92E+08
			s^2_{ejUS1}	1.68E+08	2.28E+08	25118909	3.53E+08
			s^2_{ejUS2}	1.68E+08	2.28E+08	25219536	3.52E+08
MVE	C_x	0.79	C_x	0.454	0.182	0.283	0.140
	$\beta_2(x)$	8.89	$\beta_2(x)$	3.603	6.359	2.634	2.907
	$\bar{X}$	646.06	$\bar{x}$	1847.546	707.7692	7737.375	14737.86
	$\bar{Y}$	4187.36	$\bar{y}$	14506.46	4602.308	30038.38	41677.86
			s^2_{ejSD}	1.74E+08	2.09E+08	38185742	8.91E+08
			s^2_{ejSK}	1.74E+08	2.09E+08	38055419	8.89E+08
			s^2_{ejUS1}	1.65E+08	2.09E+08	30207489	6.97E+08
			s^2_{ejUS2}	1.54E+08	2.1E+08	29008646	4.48E+08
	N	80	N_i	19	32	14	15
	n	45	n_i	11	18	8	8
			ξ_i	0.26	0.15	0.35	0.33
			f_i	0.58	0.56	0.57	0.53
			w_j	0.24	0.40	0.18	0.19

Table 2: Data statistics for Population II

Estimators	Population-I Classical	MCD	MVE	Population-II Classical	MCD	MVE
$\bar{y}^*_{rSD}$	1	1.4968	32.4457	1	14.2037	45.5237
$\bar{y}^*_{rSK}$	1	1.2924	29.8624	1	13.8667	44.5189
$\bar{y}^*_{rUS1}$	1	1.1686	24.6078	1	10.4495	33.3511
$\bar{y}^*_{rUS2}$	1	5.6305	74.4301	1	81.4782	218.905

Table 3: Theoretical results for the relative efficiencies of proposed robust calibration estimators with respect to existing estimator

ε	n	Estimators	5%			10%			25%		
			Classic	MCD	MVE	Classic	MCD	MVE	Classic	MCD	MVE
$N(0,1)$	20	$\bar{y}^{*}_{rSD}$	1.00	4803.37	52.71	1.00	21919.43	3011.74	1.00	3.48	19.18
		$\bar{y}^{*}_{rSK}$	1.00	196.78	196.21	1.00	100.71	100.49	1.00	12.42	12.48
		$\bar{y}^{*}_{rUS1}$	1.00	12.58	160.09	1.00	1327.69	516.52	1.00	5.73	3.59
		$\bar{y}^{*}_{rUS2}$	1.00	236.88	458.74	1.00	8.36	32.50	1.00	5147.77	28329.72
		CT:	0.35	7.33	7.83	0.38	7.45	8.7	0.41	7.67	8.64
	32	$\bar{y}^{*}_{rSD}$	1.00	5.38	51.25	1.00	147.17	5.04	1.00	187.51	10.57
		$\bar{y}^{*}_{rSK}$	1.00	236.76	236.67	1.00	82.90	83.09	1.00	9.16	9.20
		$\bar{y}^{*}_{rUS1}$	1.00	360.66	58.15	1.00	4.14	26.68	1.00	171.30	142.01
		$\bar{y}^{*}_{rUS2}$	1.00	16510.22	21345.92	1.00	39.11	209.65	1.00	2.25	448.84
		CT:	0.39	7.81	8.30	0.39	7.42	9.09	0.37	8.05	8.60
$Exp(1)$	20	$\bar{y}^{*}_{rSD}$	1.00	3697.07	284.54	1	2.06	8.54	1.00	448.20	490.65
		$\bar{y}^{*}_{rSK}$	1.00	272.53	272.96	1.00	122.59	122.54	1.00	13.30	13.35
		$\bar{y}^{*}_{rUS1}$	1.00	3035.88	862.00	1.00	7848.84	8736.25	1.00	121.93	161.00
		$\bar{y}^{*}_{rUS2}$	1.00	495.61	283.54	1.00	258.43	71.89	1.00	1839.04	1880.54
		CT:	0.38	8.04	8.68	0.44	7.29	8.94	0.44	7.75	8.22
	32	$\bar{y}^{*}_{rSD}$	1.00	15.00	15.23	1.00	159.58	40.11	1.00	627.98	1135.06
		$\bar{y}^{*}_{rSK}$	1.00	278.28	277.79	1.00	89.01	88.96	1.00	9.62	9.65
		$\bar{y}^{*}_{rUS1}$	1.00	27294.92	34053.35	1.00	26081.36	67235.98	1.00	225.18	18.34
		$\bar{y}^{*}_{rUS2}$	1.00	3225.21	5823.17	1.00	715.16	697.11	1.00	1087.82	565.48
		CT:	0.44	7.70	8.86	0.41	7.96	8.40	0.44	7.34	8.65
$Unif(0,1)$	20	$\bar{y}^{*}_{rSD}$	1.00	43.49	34.94	1.00	59.72	55.41	1.00	355.97	16.33
		$\bar{y}^{*}_{rSK}$	1.00	250.05	249.46	1.00	108.80	108.70	1.00	12.86	12.86

n	Estimators	Classic	MCD	MVE	Classic	MCD	MVE	Classic	MCD	MVE
	$\overline{y}^*_{rUS1}$	1.00	230.96	7.59	1.00	160.90	19.97	1.00	2410.50	950.86
	$\overline{y}^*_{rUS2}$	1.00	381.02	404.19	1.00	399.74	52.88	1.00	17101.73	21455.50
	CT:	0.36	7.19	9.02	0.34	7.47	8.29	0.47	7.65	8.93
32	$\overline{y}^*_{rSD}$	1.00	79.53	4288.86	1.00	7153.35	461.10	1.00	2.86	14.14
	$\overline{y}^*_{rSK}$	1.00	256.13	255.28	1.00	95.07	94.99	1.00	9.34	9.34
	$\overline{y}^*_{rUS1}$	1.00	5723.62	4719.90	1.00	42800.65	6105.50	1.00	51.59	113.11
	$\overline{y}^*_{rUS2}$	1.00	48.93	95.03	1.00	456.23	897.75	1.00	954341.05	957612.30
	CT:	0.50	7.39	8.54	0.45	7.47	8.03	0.36	7.71	8.31

Tablo 4: The Relative Efficiencies (RE) and Computational Time (CT) results of Simulation Design-I

n	Estimators	5%			10%			25%		
		Classic	MCD	MVE	Classic	MCD	MVE	Classic	MCD	MVE
20	$\overline{y}^*_{rSD}$	1.00E+00	1.91E+02	2.72E+00	1.00E+00	7.50E+01	3.22E+02	1.00E+00	1.19E+02	1.03E+03
	$\overline{y}^*_{rSK}$	1.00E+00	6.78E+02	6.78E+02	1.00E+00	2.43E+02	2.43E+02	1.00E+00	8.95E+01	8.95E+01
	$\overline{y}^*_{rUS1}$	1.00E+00	1.06E+05	1.38E+05	1.00E+00	1.14E+02	3.46E+03	1.00E+00	6.92E+01	1.48E+03
	$\overline{y}^*_{rUS2}$	1.00E+00	1.99E+05	2.01E+05	1.00E+00	1.78E+02	2.56E+02	1.00E+00	1.98E+04	2.29E+04
	CT:	0.47	7.73	8.35	0.37	7.32	8.10	0.36	7.64	8.14
32	$\overline{y}^*_{rSD}$	1.00E+00	1.01E+04	2.72E+04	1.00E+00	1.71E+05	1.09E+06	1.00E+00	3.47E+03	3.36E+03
	$\overline{y}^*_{rSK}$	1.00E+00	6.76E+02	6.74E+02	1.00E+00	3.14E+02	3.13E+02	1.00E+00	6.00E+01	6.00E+01
	$\overline{y}^*_{rUS1}$	1.00E+00	7.97E+01	1.79E+02	1.00E+00	2.07E+03	2.68E+03	1.00E+00	3.92E+03	4.82E+03
	$\overline{y}^*_{rUS2}$	1.00E+00	3.60E+05	4.53E+05	1.00E+00	2.52E+02	1.26E+02	1.00E+00	4.57E+01	2.57E+00
	CT:	0.39	7.46	8.91	0.34	7.52	8.92	0.28	7.72	8.58
40	$\overline{y}^*_{rSD}$	1.00E+00	1.72E+02	3.56E+02	1.00E+00	8.16E+02	9.67E+01	1.00E+00	7.95E+01	8.68E+01
	$\overline{y}^*_{rSK}$	1.00E+00	6.53E+02	6.53E+02	1.00E+00	2.93E+02	2.92E+02	1.00E+00	4.48E+01	4.48E+01

$\bar{y}^*_{rUS1}$	1.00E+00	1.15E+02	1.87E+01	1.00E+00	1.61E+05	1.47E+05	1.00E+00	1.92E+04	1.85E+04
$\bar{y}^*_{rUS2}$	1.00E+00	1.09E+03	7.38E+02	1.00E+00	1.91E+03	2.51E+03	1.00E+00	3.12E+02	3.88E+02
CT:	0.41	7.81	8.73	0.39	7.91	8.49	0.42	7.36	8.41

Tablo 5: The Relative Efficiencies (RE) and Computational Time (CT) results of Simulation Design-II

n	Estimators	5%			10%			25%		
		Classic	MCD	MVE	Classic	MCD	MVE	Classic	MCD	MVE
20	$\bar{y}^*_{rSD}$	1.00E+00	3.78E+01	1.14E+02	1.00E+00	3.07E+00	5.52E+01	1.00E+00	5.60E+03	5.01E+03
	$\bar{y}^*_{rSK}$	1.00E+00	4.33E+01	4.33E+01	1.00E+00	1.54E+01	1.54E+01	1.00E+00	1.78E+00	1.78E+00
	$\bar{y}^*_{rUS1}$	1.00E+00	9.21E+00	9.03E+00	1.00E+00	5.88E+02	3.25E+00	1.00E+00	1.60E+01	1.71E+00
	$\bar{y}^*_{rUS2}$	1.00E+00	6.58E+02	4.70E+02	1.00E+00	5.62E+02	4.71E+02	1.00E+00	1.03E+02	2.13E+02
	CT:	0.39	7.75	8.32	0.42	8.00	8.00	0.41	7.45	8.37
32	$\bar{y}^*_{rSD}$	1.00E+00	1.18E+02	2.87E+00	1.00E+00	1.69E+02	1.45E+01	1.00E+00	1.40E+05	6.88E+04
	$\bar{y}^*_{rSK}$	1.00E+00	5.26E+01	5.26E+01	1.00E+00	1.65E+01	1.64E+01	1.00E+00	1.84E+00	1.84E+00
	$\bar{y}^*_{rUS1}$	1.00E+00	1.41E+02	2.82E+02	1.00E+00	2.83E+02	1.81E+01	1.00E+00	1.15E+02	1.96E+02
	$\bar{y}^*_{rUS2}$	1.00E+00	2.71E+01	1.32E+01	1.00E+00	2.00E+03	5.04E+03	1.00E+00	6.95E+00	1.23E+01
	CT:	0.39	8.20	9.09	0.39	7.99	8.57	0.36	8.09	8.17
40	$\bar{y}^*_{rSD}$	1.00E+00	2.28E+02	1.64E+03	1.00E+00	3.19E+01	3.10E+01	1.00E+00	1.14E+03	2.74E+04
	$\bar{y}^*_{rSK}$	1.00E+00	5.08E+01	5.07E+01	1.00E+00	1.55E+01	1.55E+01	1.00E+00	1.64E+00	1.64E+00
	$\bar{y}^*_{rUS1}$	1.00E+00	5.02E+00	2.16E+01	1.00E+00	2.19E+00	4.88E+01	1.00E+00	1.33E+01	4.46E+01
	$\bar{y}^*_{rUS2}$	1.00E+00	1.01E+02	9.70E+01	1.00E+00	3.79E+02	1.36E+01	1.00E+00	8.86E+01	1.08E+01
	CT:	0.41	8.02	8.63	0.45	7.88	8.56	0.36	7.84	8.75

Tablo 6: The Relative Efficiencies (RE) and Computational Time (CT) results of Simulation Design-III

Author's Biography

Tolga Zaman is an Associate Professor at the Department of Statistics in Cankiri Karatekin University, Cankiri, Turkey. He received his MS and PhD degrees in Statistics from Ondokuz Mayis University Samsun, Turkey in 2013 and 2017, respectively. His research interests are sampling theory, resampling methods, robust statistics, and statistical inference. He has published more than 60 research papers in international/national journals and conferences. He has papers published in journals like Applied Mathematics and Computations, Communication in Statistics: Simulation and Computation, Communication in Statistics: Theory and Methods, Mathematical Population Studies, RevStat-Statistical Jounal, Scientia Irenica, Currency and Computation Practice and Experience.

Hasan Bulut is working as Associate Professor in Department of Statistics at Ondokuz Mayıs University, where he received his Doctor degree in 2017 based on robust clustering and robust multivariate analyses. His main research interests have been the fields of socio- economic development, robust principal component analysis, robust clustering analysis, multivariate statistical methods, and applied statistics. He has papers published in journals like Socioeconomic Planning Sciences, the Journal of Applied Statistics, Communication in Statistics: Theory and Methods, and Communication in Statistics: Simulation and Computation, Scientia Irenica. Moreover, he has a book about multivariate statistical methods with R applications.