

Research Article

Särndal Approach and Separate Type Quantile Robust Regression Type Mean Estimators for Nonsensitive and Sensitive Variables in Stratified Random Sampling

Usman Shahzad ^{1,2}, Ishfaq Ahmad,¹ Nadia H. Al-Noor ,³ Soofia Iftikhar,⁴
A. H. Abd Ellah,⁵ and Troon J. Benedict ⁶

¹Department of Mathematics and Statistics, International Islamic University, Islamabad, Pakistan

²Department of Mathematics and Statistics, PMAS-Arid Agriculture University, Rawalpindi, Pakistan

³Department of Mathematics, College of Science, Mustansiriyah University, Baghdad, Iraq

⁴Department of Statistics, Shaheed Benazir Bhutto Women University, Peshawar, KP, Pakistan

⁵Department of Mathematics, College of Science, Al-Baha University, Al-Baha, Saudi Arabia

⁶Department of Economics, Maasai Mara University, Narok, Kenya

Correspondence should be addressed to Troon J. Benedict; troon@mmarau.ac.ke

Received 11 September 2022; Revised 17 November 2022; Accepted 18 November 2022; Published 6 December 2022

Academic Editor: Ali Sajid

Copyright © 2022 Usman Shahzad 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.

Surprising perceptions may happen in survey sampling. The arithmetic mean estimator is touchy to extremely enormous or potentially small observations, whenever selected in a sample. It can give one-sided (biased) results and eventually, enticed to erase from the selected sample. These extremely enormous or potentially small observations, whenever known, can be held in the sample and utilized as supplementary information to expand the exactness of estimates. Also, a supplementary variable is consistently a well-spring of progress in the exactness of estimates. A suitable conversion/transformation can be utilized for getting much more precise estimates. In the current study, regarding population mean, we proposed a robust class of separate type quantile regression estimators under stratified random sampling design. The proposed class is based on extremely enormous or potentially small observations and robust regression tools, under the framework of Särndal. The class is at first defined for the situation when the nature of the study variable is nonsensitive, implying that it bargains with subjects that do not create humiliation when respondents are straightforwardly interrogated regarding them. Further, the class is stretched out to the situation when the study variable has a sensitive nature or theme. Sensitive and stigmatizing themes are hard to explore by utilizing standard information assortment procedures since respondents are commonly hesitant to discharge data concerning their own circle. The issues of a population related to these themes (for example homeless and nonregular workers, heavy drinkers, assault and rape unfortunate casualties, and drug users) contain estimation errors ascribable to nonresponses as well as untruthful revealing. These issues might be diminished by upgrading respondent participation by scrambled response devices/techniques that cover the genuine value of the sensitive variable. Thus, three techniques (namely additive, mixed, and Bar-Lev) are incorporated for the purposes of the article. The productivity of the proposed class is also assessed in light of real-life dataset. Lastly, a simulation study is also done to determine the performance of estimators.

1. Introduction

For future development, each community needs careful planning to manage its affairs efficiently. Successful planning requires many types of data that are reasonably accurate.

Everything is changing rapidly in this modern environment, requiring the regular collection of up-to-date information. It is possible to collect data in two ways, which are a complete survey of the enumeration and a sample survey. Since data collection is subject to time and cost constraints, regular data

collection by full enumeration is typically not feasible. The only solution then is sample surveys. Through surveying part of a population as a sample, more effort can be made to gather more accurate data through hiring better-trained workers, better organization, better monitoring, etc.

Furthermore, in survey sampling, it is regular to make utilization of supplementary (auxiliary) information to acquire enhanced designs and more effective estimators. This information might be utilized at the planning phase of the study, in the estimation methodology, or at the two phases. The huge amount of sampling literature portrays an assortment of methods for using supplementary information, for example, see [1–4].

At the estimation stage, in many sampling situations of the survey, estimators of ratio and regression are commonly used when using supplementary information. It moves through the part of the origin, and the variance of the study variable is proportional to the auxiliary variable(s) when the relationship between the study and auxiliary variable(s) is a straight line, and the efficiencies of these estimators are practically equivalent. In reality, if this condition is not fulfilled, at that point usual ratio estimators are less productive than regression estimators. In order to overthrow this circumstance in this literature, huge research has been carried out to better the ratio estimators by providing different adjusted/modified ratio-type estimators. In addition, the ratio estimator is used quite effectively in case of a positive correlation between the study and the auxiliary variable(s). There are numerous practical circumstances (medical, biological, economical, and industrial sectors) when a positive correlation between the two or more variables (one is a study variable and the other are auxiliary variables) exists.ion are as follows: (i) the sale of a particular commodity rises/increases with the increase in the region's population and average per capita income; (ii) the productivity of the employee improves/increases with both his previous experience and his educational or intelligence level; (iii) the human body's immunity increases from the risk of certain diseases by following healthy diets and paying attention to fitness, etc.

A vast amount of literature is available on ratio-type estimators for mean estimation. Such literature studies [3–5] have developed some classes of estimators utilizing supplementary information under a simple random sampling scheme. However, for positive correlation, the traditional regression-ratio-type estimator is better for the estimation of population parameters [6]. It is noted that traditional regression-type-ratio estimators are based on conventional regression coefficient which becomes inappropriate when data are contaminated by outliers and hence mean estimation too. For solving this issue, there are some modifications available in the literature under a simple random sampling scheme (see, for example, [7–11]). In light of these developments [12], Zaman and Bulut first time introduced robust regression type mean estimators under stratified random sampling. To the best of our knowledge, utilization of quantile robust regression with Särndal approach [13] in the mean estimation does not discuss yet in mean estimation under stratified random sampling. Thus, drawing on these encounters and following a plan based methodology, in the

framework of [13], we propose in the current paper a new class of estimators for the mean of a study variable under stratified random sampling. The comparison is based on the real-life application and various simulation experiments which clearly indicates that the planned class outperforms the various estimators remembered for the highly sensitive and quickly investigated in the paper. At first, we accept that the examination/study variable is nonsensitive and, in this manner, a standard estimation setting is implemented. After that, we stretch out the outcomes to the instance of a sensitive, highly personal, stigmatizing, or in any event threatening variable, which is thought to be observed on sampled units by utilizing nonstandard survey methods so as to increase respondent collaboration. Current methods are preferable as randomized response theory proposed by [14] and well-demonstrated in the monographs via [15] and numerous other scholars. Furthermore, for the sensitive variable, mathematical design is conceded out to examine the behavior of the planned class on the origin of distressed values of the study variable attained by scrambling the factual values conferring to the directions given by a number of randomized response models.

The other main parts of this paper are structured into following sections. Section 2 documented some existing robust estimators of the population mean of the study variable under a stratified random sampling scheme. In Section 3, the proposed family of estimators has been determined and the expressions for its mean square error (MSE) have been attained to first order of approximation. We theoretically explained, in Section 4, the sensitive and nonsensitive responses of mean estimation. Section 5 enfold the numerical illustration of the existing and proposed family. Finally, in Section 6, the article leads up to concluding remarks.

2. Robust Regression and Mean Estimators

The most well-known traditional regression technique is ordinary least squares (OLS). It depends upon limiting (minimizing) the aggregate of the squares of the residuals (r_i). Due to computational effortlessness, OLS mostly utilized for parameter estimation. The OLS strategy gives the best estimation results of the straight-line regression under the perfect conditions portrayed by the OLS. It is noted that the estimates of parameters are based on OLS, influenced by outliers, and subsequently have not given significantly productive outcomes. According to [16], the threshold point of the OLS fitting is $1/n$ or 0%, which infers that it tends to be effectively influenced by even a single outlier. Hence, the mean estimation based on OLS in presence of outliers is also affected [9]. To overcome this issue in mean estimation, many robust regression tools and robust covariance matrices are utilized by authors. For the purposes of the article, the precise description of some of these tools and matrices is provided in upcoming lines.

The least absolute deviation (LAD) is based upon the minimization of the sum of absolute squared errors (SEs). The least median of squares (LMS) is based upon the minimization of the median of SE. The least trimmed squares (LTS) is another robust regression tool. In LTS, SEs

are sorted after that OLS is applied on initial Z observations. Hence, computations are not affected by extreme values in LTS. Another unique kind of robust regression tool is based on three steps, namely MM-estimation is also utilized in the presence of outliers, for details see [17]. The M -estimation is based upon minimization of the function $\varnothing(\cdot)$. Some of the designed formulae for the objective function are as follows:

Huber-M estimator [18] considered the objective function with $u = 4.685$ or 6 , as

$$\varnothing(r_i) = \begin{cases} \frac{1}{2}r_i^2, & \text{for } |r_i| \leq u, \\ v\left(|r_i| - \frac{1}{2}v\right), & \text{for } |r_i| > u. \end{cases} \quad (1)$$

Hampel-M estimator [19] considered the objective function with $k = 1.7$, $g = 3.4$, and $u = 8.5$, as

$$\varnothing(r_i) = \begin{cases} \frac{1}{2}r_i^2, & \text{for } 0 < |r_i| \leq k, \\ k|r_i| - \frac{1}{2}r_i^2, & \text{for } k < |r_i| \leq g, \\ \frac{-k}{2(u-g)}(u-r_i)^2 + \frac{k}{2}(g+u-k), & \text{for } g < |r_i| \leq u, \\ \frac{k}{2}(g+u-k), & \text{for } u < |r_i|. \end{cases} \quad (2)$$

Tukey-M estimator [20] considered the objective function with $u = 4.685$ or 6 , as

$$\varnothing(r_i) = \begin{cases} \frac{u^2}{6} \left[1 - \left(1 - \left(\frac{r_i}{u} \right)^2 \right)^3 \right], & \text{for } |r_i| \leq u, \\ \frac{u^2}{6}, & \text{for } |r_i| > u. \end{cases} \quad (3)$$

Furthermore, [21] constructed two robust covariance matrices. The first is the minimum covariance determinant (MCD) estimator, i.e. $T(x)$ = average of the g^{th} point of X where determinant of covariance matrix provides minimum value. It is noted that $g = n(1 - \alpha_t)$ where α_t is representing the ratio of trimming. Moreover, the threshold point of MCD is equivalent to α_t [22]. The second is the minimum volume ellipsoid (MVE) estimator, i.e. $T(x)$ = center/average of an ellipsoid with min-volume spanned by the g^{th} point X . It is noted that $g = [0.5] + 1$ and $[\cdot]$ used for rounding the number to an integer. For more details about robust regression tools, MCD and MVE, see [9, 12, 23].

In light of the abovementioned robust regression tools and covariance matrices, literature [12] defined some separate type robust regression estimators in stratified random sampling. We are providing their generalized form as follows:

$$\bar{y}_{\text{slr}(i)} = \sum_{h=1}^L \Omega_h \left[\bar{y}_h + \beta_{h,xy(\text{rob}_i)} (\bar{X}_h - \bar{x}_h) \right], \quad \text{for } i = 1, 2, \dots, 21. \quad (4)$$

MSE of separate robust regression estimator

$$\text{MSE}_{(\bar{y}_{\text{slr}(i)})} = \sum_{h=1}^L \Omega_h^2 \theta_h \left[S_{h,y}^2 + \beta_{h,xy(\text{rob}_i)}^2 S_{h,x}^2 - 2\beta_{h,xy(\text{rob}_i)} S_{h,xy} \right], \quad (5)$$

with

$(\bar{y}_h, S_{h,y}^2, \bar{x}_h, S_{h,x}^2, S_{h,xy})$ representing means variances and covariances of (Y, X) , computed from traditional method in h^{th} stratum for $i = 1, 2, \dots, 7$.

$(\bar{y}_h, S_{h,y}^2, \bar{x}_h, S_{h,x}^2, S_{h,xy})$ computed from MCD covariance matrix in h^{th} stratum for $i = 8, 9, \dots, 14$.

$(\bar{y}_h, S_{h,y}^2, \bar{x}_h, S_{h,x}^2, S_{h,xy})$ computed from MVE covariance matrix in h^{th} stratum for $i = 15, 16, \dots, 21$.

$(\beta_{h,xy(\text{rob}_i)})$ computed from robust regression tools (i.e. LAD, LMS, LTS, Huber-MM (HMM), Huber-M (HM), Hampel-M (HPL), and Tukey-M (TK)) for $i = 1, 2, \dots, 21$.

3. Quantile Robust-Regression-Type Estimators Using Särndal Approach

Much of the time, genuine information contains extremely enormous or potentially small observations. Different cross breed seed creation organizations present new assortments of seeds and furthermore indicate the scope of generation per section of land that rancher would profit by. Extremely enormous or potentially small observations can without much of a stretch be perused from the predetermined extents. For evaluating average pay of families, pay off the most extravagant people (greatest) in the general public is notable, and that of least fortunate (least) can without much of a stretch be surveyed. Correspondingly in different studies that are directed consistently after an explicit interim of time, data about extremely enormous or potentially small observations can undoubtedly be obtained. Mean per unit estimator for population mean is delicate to uncommon observations. In such circumstance, this estimator can deliver deceiving results if any of the extremely enormous or potentially small observations is chosen in the sample. As indicated by [24], when the outliers or extreme values are substantial, it can give new bits of knowledge about the nature of the data. Keeping this fact in mind, we shed light on [13] estimator for Y and X under stratified random sampling as

$$\bar{y}_{h,\text{sr}} = \begin{cases} \bar{y}_h + \eta_{1h}, & \text{if selected sample contain } Y_{h,m} \text{ but not } Y_{h,M}, \\ \bar{y}_h - \eta_{1h}, & \text{if selected sample contain } Y_{h,M} \text{ but not } Y_{h,m}, \\ \bar{y}_h, & \text{for rest of the samples.} \end{cases} \quad (6)$$

Similarly,

$$\bar{x}_{h, sr} = \begin{cases} \bar{x}_h + \eta_{2h}, & \text{if selected sample contain } Y_{h,m} \text{ but not } Y_{h,M}, \\ \bar{x}_h - \eta_{2h}, & \text{if selected sample contain } Y_{h,M} \text{ but not } Y_{h,m}, \\ \bar{y}_h, & \text{for rest of the samples,} \end{cases} \quad (7)$$

where $\eta_h, \{h = 1, 2, \dots, L\}$ are wisely chosen constants with respect to (x_h, y_h) . Further, $(X_{h,m}, Y_{h,m})$ representing the minimum values and $(X_{h,M}, Y_{h,M})$ representing the maximum values with respect to (x_h, y_h) . So, utilizing [13] technique and extending the idea of [12], we propose the following separate type robust regression estimators.

The variance of [13] estimator under stratified random sampling can be written as

$$\begin{aligned} \text{var}(\bar{y}_{c.st(sr)}) &= \text{var}\left(\sum_{h=1}^L \Omega_h \bar{y}_{h, sr}\right) \\ &= \sum_{h=1}^L \Omega_h^2 \text{var}(\bar{y}_{h, sr}) \\ &= \sum_{h=1}^L \Omega_h^2 \theta_h \left[S_{h,y}^2 - \frac{2n_h \eta_h}{N_h - 1} (Y_{h,M} - Y_{h,m} - n_h \eta_h) \right]. \end{aligned} \quad (8)$$

By substituting $\eta_h^{\text{opt}} = (Y_{h,M} - Y_{h,m}/2n_h)$ in the above expression, we get

$$\begin{aligned} \text{var}(\bar{y}_{c.st(sr)})_{\min} &= \sum_{h=1}^L \Omega_h^2 \theta_h \left[S_{h,y}^2 - \frac{1}{2(N_h - 1)} (Y_{h,M} - Y_{h,m})^2 \right]. \end{aligned} \quad (9)$$

In light of [13], we propose separate type robust quantile regression estimator:

$$\bar{y}_{slr, sr(pj)} = \sum_{h=1}^L \Omega_h \left[\bar{y}_{h, sr} + \beta_{h, xy}(q_j) (\bar{X}_h - \bar{x}_{h, sr}) \right], \quad (10)$$

for $i = 1, 2, \dots, 7$.

$\bar{y}_{h, sr}$ and $\bar{x}_{h, sr}$ are mean estimators based on Särndal technique. For theoretical MSE expressions, let us define the notations in light of [13], as follows:

$$\begin{aligned} \zeta_{y,h} &= \frac{\bar{y}_{h, sr}}{\bar{Y}_h} - 1, \\ \zeta_{x,h} &= \frac{\bar{x}_{h, sr}}{\bar{X}_h} - 1, \\ E(\zeta_{x,h}) &= (\zeta_{y,h}) = 0, \\ E(\zeta_{y,h}^2) &= \frac{\theta_h}{\bar{Y}_h^2} \left[S_{h,y}^2 - \frac{2n_h \eta_{1h}}{N_h - 1} (Y_{h,d} - n_h \eta_{1h}) \right], \\ E(\zeta_{x,h}^2) &= \frac{\theta_h}{\bar{X}_h^2} \left[S_{h,x}^2 - \frac{2n_h \eta_{2h}}{N_h - 1} (X_{h,d} - n_h \eta_{2h}) \right], \\ E(\zeta_{y,h} \zeta_{x,h}) &= \frac{\theta_h}{\bar{Y}_h \bar{X}_h} \left[S_{h,xy} - \frac{n_h}{N_h - 1} \right. \\ &\quad \left. (\eta_{2h} Y_{h,d} + \eta_{1h} X_{h,d} - 2n_h \eta_{1h} \eta_{2h}) \right]. \end{aligned} \quad (11)$$

Now, let we expand the right-hand side of equation (10), as follows:

$$\bar{y}_{slr, sr(pj)} - \bar{Y} = \sum_{h=1}^L \Omega_h \left[\bar{Y}_h (1 + \zeta_{y,h}) - \beta_{h, xy}(q_j) \bar{X}_h \zeta_{x,h} - \bar{Y}_h \right]. \quad (12)$$

Squaring both sides of equation (8) and disregarding terms of ζ s having a power higher than two, we get the MSE of $\bar{y}_{slr, sr(pj)}$ as

$$\text{MSE}(\bar{y}_{slr, sr(pj)}) = \text{MSE}(\bar{y}_{sqr}(j)) - y, \quad (13)$$

where

$$\begin{aligned} \text{MSE}(\bar{y}_{sqr}(j)) &= \sum_{h=1}^L \Omega_h^2 \theta_h \left(S_{h,y}^2 + \beta_{h, xy}^2(q_j) S_{h,x}^2 - 2\beta_{h, xy}(q_j) S_{h,xy} \right), \\ &\text{for } j = 1, 2, \dots, 7. \end{aligned} \quad (14)$$

and

$$\begin{aligned} Y &= \sum_{h=1}^L \frac{2\Omega_h^2 \theta_h n_h}{N_h - 1} \left(\eta_{1h} - \beta_{h, xy}(\text{rob}_j) \eta_{2h} \right) \left[Y_{h,d} \beta_{h, xy}(\text{rob}_j) X_{h,d} \right. \\ &\quad \left. - n_h \left(\eta_{1h} - \beta_{h, xy}(\text{rob}_j) \eta_{2h} \right) \right]. \end{aligned} \quad (15)$$

Each and every notation in $\text{MSE}(\bar{y}_{sqr, sr(j)})$ viably delineated in previous lines. Further, the proposed class can be planned in the structure of [13]. Nonetheless, we are executing their proposition in stratified random sampling with a robust quantile regression tool. Along these lines, exploiting of known results, with some direct numerical computations, keeping up a vital distance from dull or inconsequential algebra, we give the optimal values of (η_{1h}, η_{2h}) and subsequently, minimum $\text{MSE}(\bar{y}_{sqr, sr(j)})$ of the estimators $\bar{y}_{sqr, sr(j)}$ as pursues:

$$\begin{cases} \eta_{1h}^{\text{opt}} = \frac{Y_{h,M} - Y_{h,m}}{2n_h} = \frac{Y_{h,d}}{2n_h}, \\ \eta_{2h}^{\text{opt}} = \frac{X_{h,M} - X_{h,m}}{2n_h} = \frac{X_{h,d}}{2n_h}, \end{cases} \quad (16)$$

$$\begin{aligned} \text{MSE}_{\min}(\bar{y}_{sqr, sr(pj)}) &= \text{MSE}(\bar{y}_{sqr}(j)) - \sum_{h=1}^L \frac{\Omega_h^2 \theta_h}{2(N_h - 1)} \\ &\quad \left[Y_{h,d} - \beta_{h, xy}(\text{rob}_j) X_{h,d} \right]^2, \\ &\text{for } j = 1, 2, \dots, 7. \end{aligned}$$

Note that any quantile can be used here. However, some of them which are $(q^{15\text{th}}, q^{25\text{th}}, q^{35\text{th}}, q^{45\text{th}}, q^{55\text{th}}, q^{65\text{th}}, q^{75\text{th}})$ quantiles are considered in this article. In light of these seven referenced quantiles, the new class consists of seven members. For the sake of readability, let us provide seven

members of the new class with their min-MSE in compact form, as follows:

$$\begin{aligned}
 \bar{y}_{\text{sqr-sr}(p_j)} &= \left\{ \begin{aligned} &\sum_{h=1}^L \Omega_h [\bar{y}_{h\text{-sr}} + \beta_{h\text{-xy}(0.15)} (\bar{X}_h - \bar{X}_{h\text{-sr}})], & q_j = (0.15) \text{ for } j = 1, \\ &\sum_{h=1}^L \Omega_h [\bar{y}_{h\text{-sr}} + \beta_{h\text{-xy}(0.25)} (\bar{X}_h - \bar{X}_{h\text{-sr}})], & q_j = (0.25) \text{ for } j = 2, \\ &\sum_{h=1}^L \Omega_h [\bar{y}_{h\text{-sr}} + \beta_{h\text{-xy}(0.35)} (\bar{X}_h - \bar{X}_{h\text{-sr}})], & q_j = (0.35) \text{ for } j = 3, \\ &\sum_{h=1}^L \Omega_h [\bar{y}_{h\text{-sr}} + \beta_{h\text{-xy}(0.45)} (\bar{X}_h - \bar{X}_{h\text{-sr}})], & q_j = (0.45) \text{ for } j = 4, \\ &\sum_{h=1}^L \Omega_h [\bar{y}_{h\text{-sr}} + \beta_{h\text{-xy}(0.55)} (\bar{X}_h - \bar{X}_{h\text{-sr}})], & q_j = (0.55) \text{ for } j = 5, \\ &\sum_{h=1}^L \Omega_h [\bar{y}_{h\text{-sr}} + \beta_{h\text{-xy}(0.65)} (\bar{X}_h - \bar{X}_{h\text{-sr}})], & q_j = (0.65) \text{ for } j = 6, \\ &\sum_{h=1}^L \Omega_h [\bar{y}_{h\text{-sr}} + \beta_{h\text{-xy}(0.75)} (\bar{X}_h - \bar{X}_{h\text{-sr}})], & q_j = (0.75) \text{ for } j = 7, \end{aligned} \right. \\
 \\
 \text{MSE}_{\min}(\bar{y}_{\text{sqr-sr}(p_j)}) &= \left\{ \begin{aligned} &\text{MSE}(\bar{y}_{\text{sqr}(1)}) - \sum_{h=1}^L \frac{\Omega_h^2 \theta_h}{2(N_h - 1)} [Y_{h,d} - \beta_{h\text{-xy}(0.15)} X_{h,d}]^2, & q_j = (0.15) \text{ for } j = 1, \\ &\text{MSE}(\bar{y}_{\text{sqr}(2)}) - \sum_{h=1}^L \frac{\Omega_h^2 \theta_h}{2(N_h - 1)} [Y_{h,d} - \beta_{h\text{-xy}(0.25)} X_{h,d}]^2, & q_j = (0.25) \text{ for } j = 2, \\ &\text{MSE}(\bar{y}_{\text{sqr}(3)}) - \sum_{h=1}^L \frac{\Omega_h^2 \theta_h}{2(N_h - 1)} [Y_{h,d} - \beta_{h\text{-xy}(0.35)} X_{h,d}]^2, & q_j = (0.35) \text{ for } j = 3, \\ &\text{MSE}(\bar{y}_{\text{sqr}(4)}) - \sum_{h=1}^L \frac{\Omega_h^2 \theta_h}{2(N_h - 1)} [Y_{h,d} - \beta_{h\text{-xy}(0.45)} X_{h,d}]^2, & q_j = (0.45) \text{ for } j = 4, \\ &\text{MSE}(\bar{y}_{\text{sqr}(5)}) - \sum_{h=1}^L \frac{\Omega_h^2 \theta_h}{2(N_h - 1)} [Y_{h,d} - \beta_{h\text{-xy}(0.55)} X_{h,d}]^2, & q_j = (0.55) \text{ for } j = 5, \\ &\text{MSE}(\bar{y}_{\text{sqr}(6)}) - \sum_{h=1}^L \frac{\Omega_h^2 \theta_h}{2(N_h - 1)} [Y_{h,d} - \beta_{h\text{-xy}(0.65)} X_{h,d}]^2, & q_j = (0.65) \text{ for } j = 6, \\ &\text{MSE}(\bar{y}_{\text{sqr}(7)}) - \sum_{h=1}^L \frac{\Omega_h^2 \theta_h}{2(N_h - 1)} [Y_{h,d} - \beta_{h\text{-xy}(0.75)} X_{h,d}]^2, & q_j = (0.75) \text{ for } j = 7. \end{aligned} \right. \tag{17}
 \end{aligned}$$

4. Estimation in Sensitive Research

Frequently in biomedical studies and socioeconomic, normally the researcher has to collect information concerning embarrassing, threatening issues, or even highly sensitive. When exploiting sensitive inquiry, posing straight questions to the respondents by means of a customary collection of data approaches (for example, self-administered surveys with pencil and paper, computer-supported telephone interrogating, computer-supported self-interrogating, audio computer-supported self-interrogating, or by computer-supported Web-interrogating) may acquire refuse to respond or even untruthful responses as of social disgrace or distress about the threat of exposing. Such methodical nonsampling response errors lead to social allure inclination in the estimates of sensitive qualities. Social desirability bias happens when respondents will, in general, present themselves in a positive light, implying that they overreport socially adequate perspectives which adjust to accepted practices (for example providing for a noble cause, having confidence in God, voting, good dieting, accomplishing deliberate work) and underreport socially objected, bothersome practices, which digress from social standards (for example xenophobia, hostile to Semitism, gambling, utilization of alcohol, premature birth, sexual brutality, drug and upgrading substances, and tax avoidance). The impact is to imperfect the nature of the gathered information and produces an unreliable analysis of the sensitive behavior under scrutiny.

To restrict unsatisfactory paces of nonresponse and get more solid information, indirect questioning approaches (more details in [25]), for example, randomized response technique (RR) and (RRT), might be utilized. The RRT was begun by [14] who proposed an information assortment technique that permits researchers to acquire more dependable sensitive data by expanding respondents participation without endangering security insurance to ensure classification to the respondents, and a randomization gadget (decks of cards, dice, coins, colored numbered balls, spinners, irregular number generators, and so forth) is utilized to hide the appropriate responses as in the respondents answer to one of at least two chosen questions relying upon the result of the device. Private information (privacy) is endangered since respondents do not expose to anyone the question that has been nominated and unknown, excluding the respondent, knows the result generated by the randomization stratagem. Since privacy is completely endangered, the method should, as in our opinion, promote greater collaboration between respondents and diminish their incentive to incorrectly reveal their arrogance. Therefore, it is expected that study participants are acquiescent with the rules arranged by the assumed randomization instrument and are completely truthful in releasing their answers. The gadget for randomization produces a probabilistic connection between respondent's answers and the genuine sensitive status, which is utilized to make inferences about an obscure sensitive population, for example, the commonness of a slandering characteristic, the mean/complete of a quantitative sensitive variable, or its probability function.

Standard RR techniques have been essentially considered to be utilized in studies that require a binary response (i.e. yes or no) to a sensitive inquiry, and look to appraise the extent of individuals introducing a given sensitive property. However, experimental examinations may address circumstances in which the reaction to a sensitive inquiry brings about a quantitative variable and the enthusiasm of the researcher depends, in the least demanding case, on the estimation of the mean or the all out of the sensitive variable under investigation. To manage such circumstances, Warner's thought has been immediately stretched out to delicate quantitative factors by [26, 27]. After that many authors developed several RR devices. Furthermore, for the goals behind this paper, two scrambling factors, V_1 and V_2 , are considered. Likewise, the i respondent is approached to produce an incentive from V_1 , say v_{i1} , and a value from V_2 , say v_{i2} , and afterward to deliver the scrambled value $z_i = \varphi(y_i, v_{i1}, v_{i2})$, where φ presents a scrambling capacity that permits respondents to veil the genuine sensitive value. The φ closed form comes from the scrambled reaction device that might be one of the four models talked about underneath.

Since it is expected that the respondent does not uncover to anybody the created values v_{i1} and v_{i2} , the value estimation of y_i stays uncertain to the researcher, and, consequently, security is not imperiled. In any case, in spite of the fact that the individual values y_i cannot be determined, it is conceivable to get dependable estimates of specific attributes of Y (i.e. the sensitive variable) by choosing an example of n units and utilizing the scrambled reactions acquired from all the units having a sample, say z_1, \dots, z_n .

Here, we consider three scrambling models: (i) the additive model $Z = Y + V_1$ [28], (ii) the mixed model $Z = (Y + V_1)V_2$ [29], and (iii) the Bar-Lev model $Z = \phi(Y + V_1) + (1 - \phi)YV_2$. An unbiased estimator of the unknown sensitive mean in light of the mixed values z_i , through these three scrambling models, can be effortlessly determined in not many advances along with the variance of estimator. For example, assume that the additive model is utilized to bother the genuine responses and to create a gauge based on n units chosen from the population as indicated by SRSWOR (i.e. simple random sampling without replacement). Accept a scrambling variable $V_1 > 0$ with mean \bar{V}_1 and variance $S_{V_1}^2$, individually. Since the distribution of V_1 is known, likewise V_1 and $S_{V_1}^2$ will be known ahead of time.

Assume that the i^{th} sampled respondent, $i = 1, \dots, n$, is directed to create a number v_{i1} from V_1 , however, by utilizing a PC or a smartphone application, and to include the arbitrary number with his (her) actual worth y_i . The respondent, at a second subsequent step, is approached to deliver the scrambled response device $z_i = y_i + v_{i1}$ without uncovering to anybody the created value v_{i1} . After $z_i = y_i + v_{i1}$, it surveys that $\hat{y}_i = z_i - v_{i1}$, v_{i1} is an unbiased estimator of the unidentified y_i , where the unbiasedness is estimated with esteem to the scrambling stratagem. In other words, if E signifies the expectancy operator, we have

$$E(\hat{y}_i) = E(z_i) - \bar{V}_1 = (y_i + \bar{V}_1) - \bar{V}_1 = y_i. \quad (18)$$

Therefore, by exhausting the same representation presented in Section 2, it is at once deceptive that

$$\widehat{y}_{AM} = \bar{z} - \bar{V}_1. \tag{19}$$

with $\bar{z} = \sum_{i=1}^n z_i/n$, is a design-impartial estimator of Y with variance $\text{Var}(\widehat{y}_{AM}) = \gamma(S_y^2 - S_{v_1}^2)$.

Likewise, we can continue in the same way for the assorted model and for the Bar-Lev model [30]. However, for the sake of shortness, the details are mislaid and attentive readers are mentioned to [4, 11]. Improvements in the randomized response method allow supplementary information to be used to increase the competence of the

estimation method (see [4, 11]). In doing so, we present below reviewed and proposed classes of estimators when the study variable belongs to sensitive nature, and data are composed using the four beforehand stated scrambling approaches.

Let (z_1, \dots, z_n) denote the responses observed on a sample selected from the study population according to stratified random sampling, and let $\bar{z} = \sum_{i=1}^n z_i/n$ representing the mean of sensitive study variable in h^{th} stratum. Then, the reviewed and proposed classes of estimators are as follows:

$$\begin{aligned} \bar{z}_{\text{slr}(i)} &= \sum_{h=1}^L \Omega_h \left[\bar{z}_h + \beta_{h,xz}(\text{rob}_i) (\bar{X}_h - \bar{x}_h) \right], \quad \text{for } i = 1, 2, \dots, 21, \\ \bar{z}_{\text{slr.sr}(p)} &= \sum_{h=1}^L \Omega_h \left[\bar{z}_{h,\text{sr}} + \beta_{h,xz}(q_j) (\bar{X}_h - \bar{x}_{h,\text{sr}}) \right], \quad \text{for } i = 1, 2, \dots, 7. \end{aligned} \tag{20}$$

with the MSE

$$\begin{aligned} \text{MSE}(\bar{z}_{\text{slr}(i)}) &= \sum_{h=1}^L \Omega_h^2 \theta_h \left[S_{h,z}^2 + \beta_{h,xz}^2(\text{rob}_i) S_{h,x}^2 - 2\beta_{h,xz}(\text{rob}_i) S_{h,xz} \right], \quad \text{for } i = 1, 2, \dots, 21, \\ \text{MSE}_{\min}(\bar{z}_{\text{sqr.sr}(p)}) &= \text{MSE}(\bar{z}_{\text{sqr}(j)}) - \sum_{h=1}^L \frac{\Omega_h^2 \theta_h}{2(N_h - 1)} \left[Z_{h,d} - \beta_{h,xz}(\text{rob}_i) X_{h,d} \right]^2, \quad \text{for } j = 1, 2, \dots, 7. \end{aligned} \tag{21}$$

Note that these expressions promptly obtained by exchanging the population parameter for Y with the corresponding population parameters for Z . In order to make the paper slim, we intentionally avoid reporting detailed derivations of MSE because the modified formulas easily derived without any further efforts.

5. Numerical Illustration

5.1. *Real-Life Application (Population-1).* We consider the dataset of 80 factories, available in [31], where

- X = Data on the number of workers
- Y = Output for 80 factories in a region

The dataset is converted into four strata as follows:

- Stratum-I: $x < 100$, Stratum-II: $100 \leq x < 200$, Stratum-III: $200 \leq x < 500$, and
- Stratum-IV: $x \geq 500$.

Note that the dataset was free from outliers, so for the purposes of the article, we replace some values of (x, y) as an outlier. After that, we draw the scatter plot for each stratum, see Figures 1–4. These figures clearly show that each stratum contains outliers so suitable for robust regression tools. Some statistical descriptions of the dataset along with robust and quantile regression coefficients are provided in Tables 1

and 2. For more details about that data, see [12, 32]. The MSE of proposed and existing estimators is given in Table 3.

5.2. *Monte-Carlo Simulation (Population-2).* While evaluating the exhibition of new suggested estimators, it is customary to derive the MSE-based theoretical conditions under which an estimator is more productive than reviewed ones. Nonetheless, these MSE-based conditions are commonly difficult to affirm. Consequently, we skipped these conditions and move towards a Monte-Carlo simulation experiment, in the current section.

For Monte-Carlo simulation experiment, let we generate the stratified population from the following regression model:

$$y_i = m(u_i) + \epsilon_i, \quad \text{for } i = 1, 2, 3, 4. \tag{22}$$

where

- Stratum-I:
 $x_1 = m(u_1) = 2(x - 0.6) + \exp(-400(x - 0.5)^{2.5}) + 10$,
- Stratum-II:
 $x_2 = m(u_2) = 2(x - 0.5) + \exp(-200(x - 0.5)^2) + 1$,
- Stratum-III:
 $x_3 = m(u_3) = 2(x - 0.6) + \exp(-200(x - 0.5)^{3.5}) + 3$,

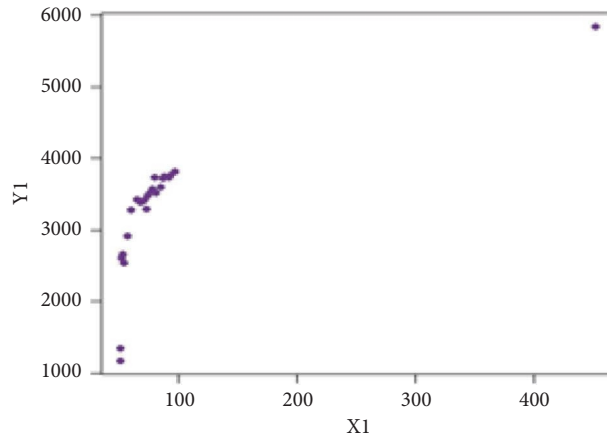


FIGURE 1: Stratum-I (population-1).

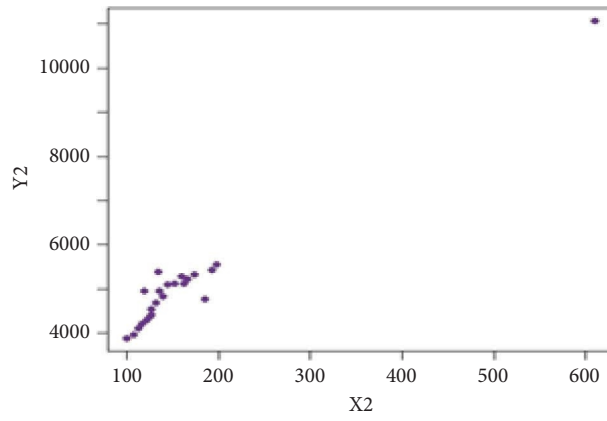


FIGURE 2: Stratum-II (population-1).

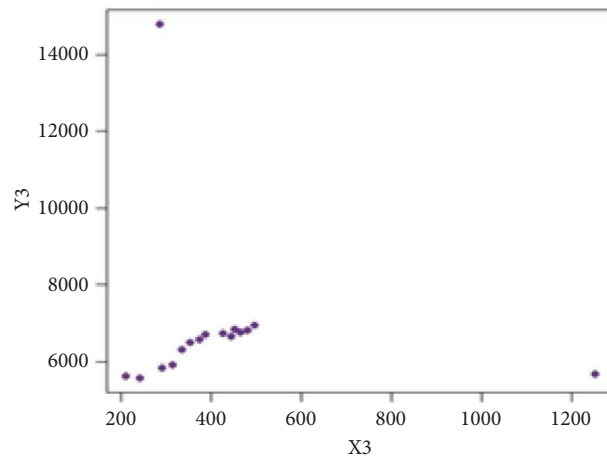


FIGURE 3: Stratum-III (population-1).

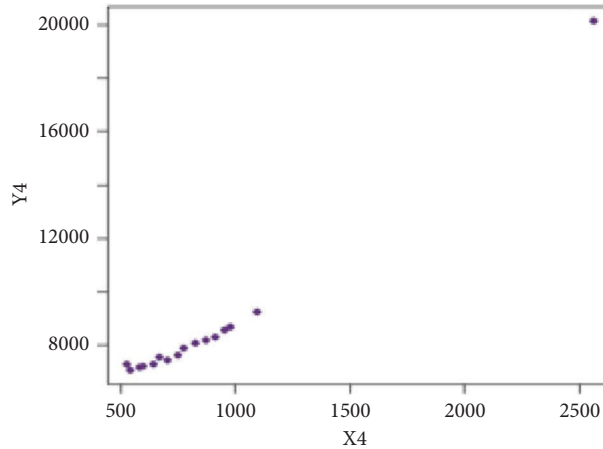


FIGURE 4: Stratum-IV (population-1).

TABLE 1: Population-1 characteristics.

	Stratum-I			Stratum-II		
	Classical	MCD	MVE	Classical	MCD	MVE
S_x	77.33154	164.1495	164.1400	101.2934	483.4461	483.4420
S_y	865.4893	11.15646	11.15599	1396.477	21.56164	21.56149
S_{xy}	47736.96	1784	1782	138369.9	10056.53	10056.47
\bar{Y}	3316.64	78.25	78.19	5070.565	135.1765	135.1761
\bar{X}	87	3550	3551	162.4348	4675.176	4675.169
	Stratum-III		Classical	MCD	Stratum-IV	
	Classical	MCD			Classical	MCD
S_x	237.0494	184.4691	184.4687	480.3166	585.9681	585.9677
S_y	2158.906	55.76827	55.76819	3142.285	158.2194	158.2171
S_{yx}	83703.39	9307.089	9307.080	1489313	92319.03	92319.01
\bar{Y}	6896.688	421.1	421.11	8607.812	759.1818	759.1815
\bar{X}	425.4375	6690.2	6690.0	874.5	7818.364	7818.357

TABLE 2: Population-1 regression coefficients.

Robust	Stratum-I		Robust	Stratum-II	
	Quantile			Quantile	
$\beta_{(LAD)} = 6.298699$	$\beta_{(0.15)} = 8.0875$		$\beta_{(LAD)} = 13.35679$	$\beta_{(0.15)} = 13.99598$	
$\beta_{(LMS)} = 15.22727$	$\beta_{(0.25)} = 7.417722$		$\beta_{(LMS)} = 25.90323$	$\beta_{(0.25)} = 13.83027$	
$\beta_{(LTS)} = 14.62162$	$\beta_{(0.35)} = 6.403141$		$\beta_{(LTS)} = 24.79245$	$\beta_{(0.35)} = 13.74534$	
$\beta_{(HMM)} = 15.88883$	$\beta_{(0.45)} = 6.362205$		$\beta_{(HMM)} = 16.95204$	$\beta_{(0.45)} = 13.50478$	
$\beta_{(HM)} = 7.022836$	$\beta_{(0.55)} = 6.256065$		$\beta_{(HM)} = 13.56174$	$\beta_{(0.55)} = 13.31106$	
$\beta_{(HPL)} = 7.275141$	$\beta_{(0.65)} = 6.172872$		$\beta_{(HPL)} = 13.50247$	$\beta_{(0.65)} = 13.14189$	
$\beta_{(TK)} = 6.596983$	$\beta_{(0.75)} = 11.2963$		$\beta_{(TK)} = 13.54156$	$\beta_{(0.75)} = 12.97162$	
Robust	Stratum-III		Robust	Stratum-IV	
	Quantile			Quantile	
$\beta_{(LAD)} = 1.367647$	$\beta_{(0.15)} = 0.05182342$		$\beta_{(LAD)} = 4.350333$	$\beta_{(0.15)} = 4.630435$	
$\beta_{(LMS)} = 2.192982$	$\beta_{(0.25)} = 0.2513312$		$\beta_{(LMS)} = 3.752212$	$\beta_{(0.25)} = 4.630303$	
$\beta_{(LTS)} = 2.433962$	$\beta_{(0.35)} = 0.6873638$		$\beta_{(LTS)} = 3.681193$	$\beta_{(0.35)} = 4.702564$	
$\beta_{(HMM)} = 5.043875$	$\beta_{(0.45)} = 1.005695$		$\beta_{(HMM)} = 3.787253$	$\beta_{(0.45)} = 4.380074$	
$\beta_{(HM)} = 0.2671355$	$\beta_{(0.55)} = 2.120879$		$\beta_{(HM)} = 6.463582$	$\beta_{(0.55)} = 4.625954$	
$\beta_{(HPL)} = 0.08898551$	$\beta_{(0.65)} = 0.5189873$		$\beta_{(HPL)} = 6.455515$	$\beta_{(0.65)} = 6.651899$	
$\beta_{(TK)} = 0.1068723$	$\beta_{(0.75)} = 1.12766$		$\beta_{(TK)} = 6.461783$	$\beta_{(0.75)} = 6.583715$	

TABLE 3: MSE for population-1 (nonsensitive).

$\hat{\theta}$	Robust	$\hat{\theta}$	MCD	$\hat{\theta}$	MVE	$\hat{\theta}$	Quantile
$\bar{y}_{slr(1)}$	13619.864	$\bar{y}_{slr(8)}$	128940.600	$\bar{y}_{slr(15)}$	128940.605	$\bar{y}_{slr.sr(p1)}$	7062.325
$\bar{y}_{slr(2)}$	20784.061	$\bar{y}_{slr(9)}$	459435.717	$\bar{y}_{slr(16)}$	459435.699	$\bar{y}_{slr.sr(p2)}$	4911.081
$\bar{y}_{slr(3)}$	20257.062	$\bar{y}_{slr(10)}$	421364.287	$\bar{y}_{slr(17)}$	421364.266	$\bar{y}_{slr.sr(p3)}$	5304.399
$\bar{y}_{slr(4)}$	10867.667	$\bar{y}_{slr(11)}$	147722.188	$\bar{y}_{slr(18)}$	147722.175	$\bar{y}_{slr.sr(p4)}$	6637.387
$\bar{y}_{slr(5)}$	10910.820	$\bar{y}_{slr(12)}$	146907.268	$\bar{y}_{slr(19)}$	146907.231	$\bar{y}_{slr.sr(p5)}$	7861.139
$\bar{y}_{slr(6)}$	10931.530	$\bar{y}_{slr(13)}$	146870.215	$\bar{y}_{slr(20)}$	146870.203	$\bar{y}_{slr.sr(p6)}$	5313.883
$\bar{y}_{slr(7)}$	20033.187	$\bar{y}_{slr(14)}$	215206.282	$\bar{y}_{slr(21)}$	215206.251	$\bar{y}_{slr.sr(p7)}$	8445.901

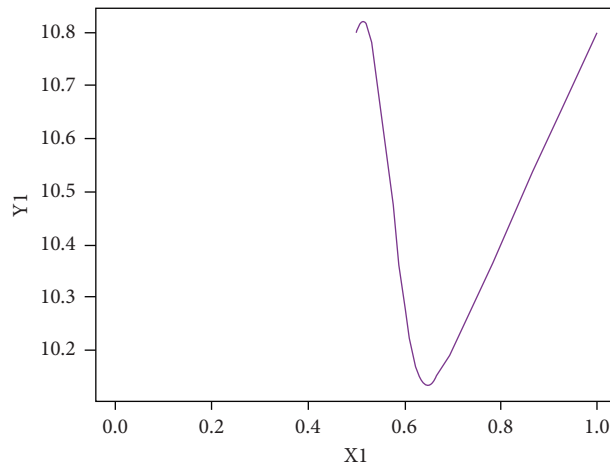


FIGURE 5: Stratum-I (population-2).

Stratum-IV: $x_4 = m(u_4) = \sin(2x\pi) + 2$.

Note that the errors (ϵ_i) are normally distributed with zero mean and 1 variance in all the strata (x_i, y_i) for $i = 1, \dots, 4$, (see [33]). The graphical representation is provided in Figures 5–8. Each stratum contains $N = 1000$ observations and 20% values selected as sample randomly. Graphical representations of the generated population are also available in Figures 5–8. We add noise in (X, Y) and guaranteed at least one outlier in each considered sample of the simulated population (see [10, 11]).

For the simulation design:

- (1) 20% sample is selected from each stratum and the mean of each estimator A_i (say) is calculated.
- (2) The above step is repeated $G' = 8000$ times and got G' values of each estimator
- (3) The empirical MSE is calculated for each up-to G' and then averaged as $MSE = (\sum_{i=1}^{G'} (MSE(A_i))/G')$. The MSE of proposed and existing estimators is provided in Table 4.

5.3. Assessment of Estimators regarding Scrambled Response.

In this section, we are assessing the performance of estimators regarding scrambled response. A similar strategy is repeated like the previous two sections for the sensitive setup. For this particular situation, the observations of the theme vitiate are nervous by the three scrambling strategies accessible in the previous section V_1 and V_2 re-expected to

be distributed normally with mean 0 and standard deviation 10 percent of the supplementary variate [4]. The MSE calculations regarding population-1 are given in Tables 5 and 6. The MSE calculations regarding simulation (population-2) are given in Tables 7 and 8.

5.4. Results Discussion. On the basis of the MSE values of the estimators, as shown in Tables 3–8, it has been observed as follows:

- (i) The MSE values associated with the estimators can be stated as follows: $MSE (Quantile) < MSE (Robust) < MSE (MCD, MVE)$ and this indicates the superior performance of the proposed class compared to the other relevant existing estimators.
- (ii) Both MCD and MVE performed similarly for most of the studied cases
- (iii) Regarding population-1, increasing the values of P from 0.3 up to 0.8, Table 6, has a good effect on the performance of all estimators. This is also true with all estimators in population-2, Table 8, except estimators that listed in the fifth and seventh rows.

Overall, under a stratified random sampling scheme, the results of MSE support the use of the proposed separate type quantile robust regression type mean estimators when supplementary information is available in addition to considering the nonsensitive and sensitive responses.

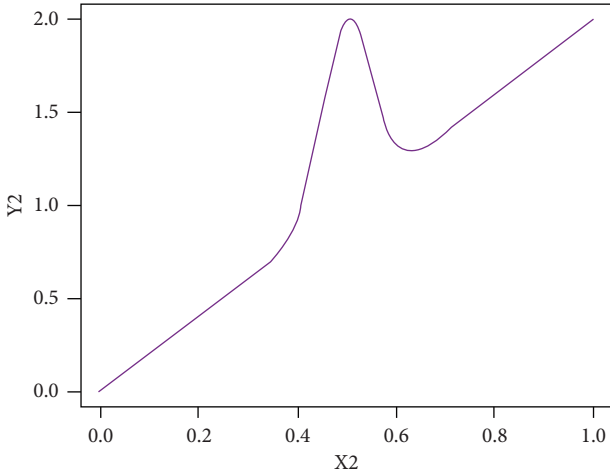


FIGURE 6: Stratum-II (population-2).

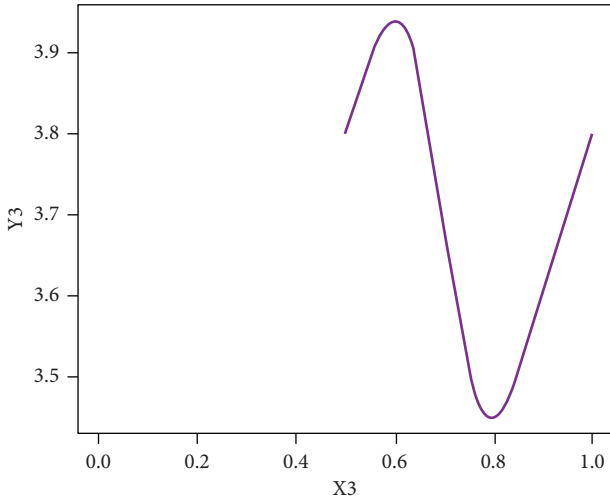


FIGURE 7: Stratum-III (population-2).

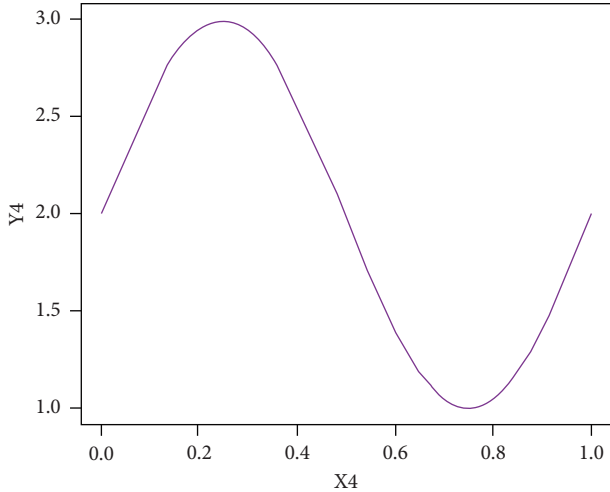


FIGURE 8: Stratum-IV (population-2).

TABLE 4: MSE for population-2 (nonsensitive).

$\hat{\theta}$	Robust	$\hat{\theta}$	MCD	$\hat{\theta}$	MVE	$\hat{\theta}$	Quantile
$\bar{y}_{slr(1)}$	567.3355	$\bar{y}_{slr(8)}$	707.1264	$\bar{y}_{slr(15)}$	707.1281	$\bar{y}_{slr-sr(p1)}$	466.8649
$\bar{y}_{slr(2)}$	567.3321	$\bar{y}_{slr(9)}$	707.1254	$\bar{y}_{slr(16)}$	707.1269	$\bar{y}_{slr-sr(p2)}$	466.8655
$\bar{y}_{slr(3)}$	567.3294	$\bar{y}_{slr(10)}$	707.1230	$\bar{y}_{slr(17)}$	707.1249	$\bar{y}_{slr-sr(p3)}$	466.8651
$\bar{y}_{slr(4)}$	567.3338	$\bar{y}_{slr(11)}$	707.1262	$\bar{y}_{slr(18)}$	707.1291	$\bar{y}_{slr-sr(p4)}$	466.8643
$\bar{y}_{slr(5)}$	567.3313	$\bar{y}_{slr(12)}$	707.1239	$\bar{y}_{slr(19)}$	707.1265	$\bar{y}_{slr-sr(p5)}$	466.8640
$\bar{y}_{slr(6)}$	567.3286	$\bar{y}_{slr(13)}$	707.1219	$\bar{y}_{slr(20)}$	707.1252	$\bar{y}_{slr-sr(p6)}$	466.8671
$\bar{y}_{slr(7)}$	567.3303	$\bar{y}_{slr(14)}$	707.1248	$\bar{y}_{slr(21)}$	707.1292	$\bar{y}_{slr-sr(p7)}$	466.8643

TABLE 5: Population-1 MSE for the mean estimators of the sensitive variable through using the additive and mixed models.

$\hat{\theta}$	Robust	$\hat{\theta}$	MCD	$\hat{\theta}$	MVE	$\hat{\theta}$	Quantile
<i>Additive model</i>							
$\bar{y}_{slr(1)}$	106.77184	$\bar{y}_{slr(8)}$	347.05782	$\bar{y}_{slr(15)}$	347.01028	$\bar{y}_{slr-sr(p1)}$	81.04381
$\bar{y}_{slr(2)}$	138.05336	$\bar{y}_{slr(9)}$	683.58398	$\bar{y}_{slr(16)}$	683.52247	$\bar{y}_{slr-sr(p2)}$	70.37243
$\bar{y}_{slr(3)}$	134.8567	$\bar{y}_{slr(10)}$	651.9963	$\bar{y}_{slr(17)}$	651.9319	$\bar{y}_{slr-sr(p3)}$	73.7262
$\bar{y}_{slr(4)}$	102.6549	$\bar{y}_{slr(11)}$	352.8103	$\bar{y}_{slr(18)}$	352.7633	$\bar{y}_{slr-sr(p4)}$	66.3055
$\bar{y}_{slr(5)}$	102.87412	$\bar{y}_{slr(12)}$	351.85693	$\bar{y}_{slr(19)}$	351.80910	$\bar{y}_{slr-sr(p5)}$	79.03844
$\bar{y}_{slr(6)}$	103.0261	$\bar{y}_{slr(13)}$	351.3355	$\bar{y}_{slr(20)}$	351.2872	$\bar{y}_{slr-sr(p6)}$	75.2371
$\bar{y}_{slr(7)}$	132.3506	$\bar{y}_{slr(14)}$	459.8484	$\bar{y}_{slr(21)}$	459.7547	$\bar{y}_{slr-sr(p7)}$	99.8949
<i>Mixed model</i>							
$\bar{y}_{slr(1)}$	22481.64	$\bar{y}_{slr(8)}$	4114956.29	$\bar{y}_{slr(15)}$	4111091.46	$\bar{y}_{slr-sr(p1)}$	15882.62
$\bar{y}_{slr(2)}$	22656.38	$\bar{y}_{slr(9)}$	2677711.85	$\bar{y}_{slr(16)}$	2672062.41	$\bar{y}_{slr-sr(p2)}$	16326.12
$\bar{y}_{slr(3)}$	22550.77	$\bar{y}_{slr(10)}$	3888537.46	$\bar{y}_{slr(17)}$	3886490.29	$\bar{y}_{slr-sr(p3)}$	14940.19
$\bar{y}_{slr(4)}$	22388.19	$\bar{y}_{slr(11)}$	5878210.18	$\bar{y}_{slr(18)}$	5874635.39	$\bar{y}_{slr-sr(p4)}$	13224.75
$\bar{y}_{slr(5)}$	22387.98	$\bar{y}_{slr(12)}$	5966204.91	$\bar{y}_{slr(19)}$	5962651.91	$\bar{y}_{slr-sr(p5)}$	12995.10
$\bar{y}_{slr(6)}$	22388.78	$\bar{y}_{slr(13)}$	5800679.41	$\bar{y}_{slr(20)}$	5797209.91	$\bar{y}_{slr-sr(p6)}$	13161.28
$\bar{y}_{slr(7)}$	22388.47	$\bar{y}_{slr(14)}$	5863638.13	$\bar{y}_{slr(21)}$	5860530.57	$\bar{y}_{slr-sr(p7)}$	13610.96

TABLE 6: Population-1 MSE for the mean estimators of the sensitive variable through using the Bar-Lev model.

$\hat{\theta}$	Robust	$\hat{\theta}$	MCD	$\hat{\theta}$	MVE	$\hat{\theta}$	Quantile
<i>P = 0.3</i>							
$\bar{y}_{slr(1)}$	12401.41	$\bar{y}_{slr(8)}$	1522958.14	$\bar{y}_{slr(15)}$	1271257.30	$\bar{y}_{slr-sr(p1)}$	7446.61
$\bar{y}_{slr(2)}$	23240.06	$\bar{y}_{slr(9)}$	9383508.77	$\bar{y}_{slr(16)}$	7783502.65	$\bar{y}_{slr-sr(p2)}$	10988.80
$\bar{y}_{slr(3)}$	12524.82	$\bar{y}_{slr(10)}$	969493.32	$\bar{y}_{slr(17)}$	825098.75	$\bar{y}_{slr-sr(p3)}$	8323.80
$\bar{y}_{slr(4)}$	12381.39	$\bar{y}_{slr(11)}$	1812632.22	$\bar{y}_{slr(18)}$	1508669.09	$\bar{y}_{slr-sr(p4)}$	6760.45
$\bar{y}_{slr(5)}$	12381.09	$\bar{y}_{slr(12)}$	1839191.58	$\bar{y}_{slr(19)}$	1529956.53	$\bar{y}_{slr-sr(p5)}$	6463.86
$\bar{y}_{slr(6)}$	12382.54	$\bar{y}_{slr(13)}$	1745061.82	$\bar{y}_{slr(20)}$	1453002.05	$\bar{y}_{slr-sr(p6)}$	6690.41
$\bar{y}_{slr(7)}$	12382.21	$\bar{y}_{slr(14)}$	1755391.42	$\bar{y}_{slr(21)}$	1461339.73	$\bar{y}_{slr-sr(p7)}$	7011.66
<i>P = 0.8</i>							
$\bar{y}_{slr(1)}$	4121.82	$\bar{y}_{slr(8)}$	232495.73	$\bar{y}_{slr(15)}$	232454.011	$\bar{y}_{slr-sr(p1)}$	3815.53
$\bar{y}_{slrs(2)}$	4605.74	$\bar{y}_{slr(9)}$	300460.47	$\bar{y}_{slr(16)}$	300412.58	$\bar{y}_{slr-sr(p2)}$	4006.40
$\bar{y}_{slr(3)}$	4390.20	$\bar{y}_{slr(10)}$	287290.96	$\bar{y}_{slr(17)}$	287274.49	$\bar{y}_{slr-sr(p3)}$	3961.16
$\bar{y}_{slr(4)}$	4089.90	$\bar{y}_{slr(11)}$	160720.80	$\bar{y}_{slr(18)}$	160672.10	$\bar{y}_{slr-sr(p4)}$	3511.96
$\bar{y}_{slr(5)}$	4088.11	$\bar{y}_{slr(12)}$	144396.23	$\bar{y}_{slr(19)}$	144341.39	$\bar{y}_{slr-sr(p5)}$	3396.34
$\bar{y}_{slr(6)}$	4089.55	$\bar{y}_{slr(13)}$	155541.14	$\bar{y}_{slr(20)}$	155489.83	$\bar{y}_{slr-sr(p6)}$	3468.02
$\bar{y}_{slr(7)}$	4088.86	$\bar{y}_{slr(14)}$	153434.14	$\bar{y}_{slr(21)}$	153382.08	$\bar{y}_{slr-sr(p7)}$	3490.03

TABLE 7: Population-2 MSE for the mean estimators of the sensitive variable through using the additive and mixed models.

$\hat{\theta}$	Robust	$\hat{\theta}$	MCD	$\hat{\theta}$	MVE	$\hat{\theta}$	Quantile
<i>Additive model</i>							
$\bar{y}_{slr(1)}$	572.9358	$\bar{y}_{slr(8)}$	712.7392	$\bar{y}_{slr(15)}$	712.7405	$\bar{y}_{slr.sr(p1)}$	472.4649
$\bar{y}_{slr(2)}$	571.7348	$\bar{y}_{slr(9)}$	711.5381	$\bar{y}_{slr(16)}$	711.5392	$\bar{y}_{slr.sr(p2)}$	471.2660
$\bar{y}_{slr(3)}$	573.6421	$\bar{y}_{slr(10)}$	713.4450	$\bar{y}_{slr(17)}$	713.4464	$\bar{y}_{slr.sr(p3)}$	473.1754
$\bar{y}_{slr(4)}$	569.4411	$\bar{y}_{slr(11)}$	709.2471	$\bar{y}_{slr(18)}$	709.2496	$\bar{y}_{slr.sr(p4)}$	468.9739
$\bar{y}_{slr(5)}$	563.6960	$\bar{y}_{slr(12)}$	703.5022	$\bar{y}_{slr(19)}$	703.5061	$\bar{y}_{slr.sr(p5)}$	463.2310
$\bar{y}_{slr(6)}$	572.6114	$\bar{y}_{slr(13)}$	712.4180	$\bar{y}_{slr(20)}$	712.4209	$\bar{y}_{slr.sr(p6)}$	472.1509
$\bar{y}_{slr(7)}$	564.3173	$\bar{y}_{slr(14)}$	704.1249	$\bar{y}_{slr(21)}$	704.1290	$\bar{y}_{slr.sr(p7)}$	463.8526
<i>Mixed model</i>							
$\bar{y}_{slr(1)}$	560.7090	$\bar{y}_{slr(8)}$	700.5124	$\bar{y}_{slr(15)}$	700.5137	$\bar{y}_{slr.sr(p1)}$	460.2381
$\bar{y}_{slr(2)}$	561.9062	$\bar{y}_{slr(9)}$	701.7095	$\bar{y}_{slr(16)}$	701.7106	$\bar{y}_{slr.sr(p2)}$	461.4374
$\bar{y}_{slr(3)}$	559.9956	$\bar{y}_{slr(10)}$	699.7985	$\bar{y}_{slr(17)}$	699.7999	$\bar{y}_{slr.sr(p3)}$	459.5289
$\bar{y}_{slr(4)}$	566.1994	$\bar{y}_{slr(11)}$	706.0054	$\bar{y}_{slr(18)}$	706.0079	$\bar{y}_{slr.sr(p4)}$	465.7322
$\bar{y}_{slr(5)}$	569.9402	$\bar{y}_{slr(12)}$	709.7464	$\bar{y}_{slr(19)}$	709.7503	$\bar{y}_{slr.sr(p5)}$	469.4752
$\bar{y}_{slr(6)}$	561.0187	$\bar{y}_{slr(13)}$	700.8253	$\bar{y}_{slr(20)}$	700.8282	$\bar{y}_{slr.sr(p6)}$	460.5582
$\bar{y}_{slr(7)}$	569.3081	$\bar{y}_{slr(14)}$	709.1157	$\bar{y}_{slr(21)}$	709.1198	$\bar{y}_{slr.sr(p7)}$	468.8434

TABLE 8: Population-2 MSE for the mean estimators of the sensitive variable through using the Bar-Lev model.

$\hat{\theta}$	Robust	$\hat{\theta}$	MCD	$\hat{\theta}$	MVE	$\hat{\theta}$	Quantile
<i>P = 0.3</i>							
$\bar{y}_{slr(1)}$	550.7090	$\bar{y}_{slr(8)}$	690.5124	$\bar{y}_{slr(15)}$	690.5137	$\bar{y}_{slr.sr(p1)}$	450.2381
$\bar{y}_{slr(2)}$	551.9062	$\bar{y}_{slr(9)}$	691.7095	$\bar{y}_{slr(16)}$	691.7106	$\bar{y}_{slr.sr(p2)}$	451.4374
$\bar{y}_{slr(3)}$	549.9956	$\bar{y}_{slr(10)}$	689.7985	$\bar{y}_{slr(17)}$	689.7999	$\bar{y}_{slr.sr(p3)}$	449.5289
$\bar{y}_{slr(4)}$	556.1994	$\bar{y}_{slr(11)}$	696.0054	$\bar{y}_{slr(18)}$	696.0079	$\bar{y}_{slr.sr(p4)}$	455.7322
$\bar{y}_{slr(5)}$	579.9402	$\bar{y}_{slr(12)}$	719.7464	$\bar{y}_{slr(19)}$	719.7503	$\bar{y}_{slr.sr(p5)}$	479.4752
$\bar{y}_{slr(6)}$	551.0187	$\bar{y}_{slr(13)}$	690.8253	$\bar{y}_{slr(20)}$	690.8282	$\bar{y}_{slr.sr(p6)}$	450.5582
$\bar{y}_{slr(7)}$	579.3081	$\bar{y}_{slr(14)}$	719.1157	$\bar{y}_{slr(21)}$	719.1198	$\bar{y}_{slr.sr(p7)}$	478.8434
<i>P = 0.8</i>							
$\bar{y}_{slr(1)}$	546.7090	$\bar{y}_{slr(8)}$	686.5124	$\bar{y}_{slr(15)}$	686.5137	$\bar{y}_{slr.sr(p1)}$	446.2381
$\bar{y}_{slr(2)}$	547.9062	$\bar{y}_{slr(9)}$	687.7095	$\bar{y}_{slr(16)}$	687.7106	$\bar{y}_{slr.sr(p2)}$	447.4374
$\bar{y}_{slr(3)}$	545.9956	$\bar{y}_{slr(10)}$	685.7985	$\bar{y}_{slr(17)}$	685.7999	$\bar{y}_{slr.sr(p3)}$	445.5289
$\bar{y}_{slr(4)}$	552.1994	$\bar{y}_{slr(11)}$	692.0054	$\bar{y}_{slr(18)}$	692.0079	$\bar{y}_{slr.sr(p4)}$	451.7322
$\bar{y}_{slr(5)}$	583.9402	$\bar{y}_{slr(12)}$	723.7464	$\bar{y}_{slr(19)}$	723.7503	$\bar{y}_{slr.sr(p5)}$	483.4752
$\bar{y}_{slr(6)}$	549.0187	$\bar{y}_{slr(13)}$	688.8253	$\bar{y}_{slr(20)}$	688.8282	$\bar{y}_{slr.sr(p6)}$	448.5582
$\bar{y}_{slr(7)}$	582.3081	$\bar{y}_{slr(14)}$	722.1157	$\bar{y}_{slr(21)}$	722.1198	$\bar{y}_{slr.sr(p7)}$	481.8434

6. Conclusion

In this paper taking motivation from [12], we propose separate type quantile robust regression type mean estimators. We also compare these estimators with reviewed estimators under stratified random sampling scheme when supplementary information is available alongside considering the nonsensitive and sensitive responses. We determine the MSE of the proposed class of estimators. In the article, two populations are presented. The outcomes obtained from the proposed class providing us reasonable suggestion that the quantile robust regression estimators perform better under randomized response as compared to the reviewed estimators based on regular strategy. To examine the MSE of the proposed classes, we consider three scrambling response models. Numerical results of proposed separate type quantile robust regression type and existing estimators of the mean (Tables 3–8) confirm the superiority of the proposed class of estimators, even in both presence or absence of nonresponses. This superiority reveals the utility

of the proposed class in practice and will perform very well in practical surveys. In future studies, the work can be extended in light of [34–36].

Data Availability

The data are included within the study for finding the results.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

- [1] D. T. Searls, “The utilization of a known coefficient of variation in the estimation procedure,” *Journal of the American Statistical Association*, vol. 59, no. 308, pp. 1225-1226, 1964.
- [2] C. E. Särndal, B. Swensson, and J. Wretman, *Model Assisted Survey Sampling*, Springer-Verlag, Berlin, Germany, 1992.
- [3] N. Koyuncu, “Efficient estimators of population mean using auxiliary attributes efficient estimators of population mean

- using auxiliary attributes,” *Applied Mathematics and Computation*, vol. 218, no. 22, pp. 10900–10905, 2012.
- [4] U. Shahzad, P. F. Perri, and M. Hanif, “A new class of ratio-type estimators for improving mean estimation of nonsensitive and sensitive variables by using supplementary information,” *Communications in Statistics - Simulation and Computation*, vol. 48, no. 9, pp. 2566–2585, 2019.
- [5] M. Hanif and U. Shahzad, “Estimation of population variance using kernel matrix,” *Journal of Statistics & Management Systems*, vol. 22, no. 3, pp. 563–586, 2019.
- [6] C. Kadilar and H. Cingi, “Ratio estimators in simple random sampling,” *Applied Mathematics and Computation*, vol. 151, no. 3, pp. 893–902, 2004.
- [7] C. Kadilar, M. Candan, and H. Cingi, “Ratio estimators using robust regression,” *Hacettepe Journal of Mathematics and Statistics*, vol. 36, no. 2, pp. 181–188, 2007.
- [8] T. Zaman, V. Saglam, M. Sagir, E. Yucesoy, and M. Zobu, “Investigation of some estimators via Taylor series approach and an application,” *American Journal of Theoretical and Applied Statistics*, vol. 3, no. 5, pp. 141–147, 2014.
- [9] T. Zaman and H. Bulut, “Modified ratio estimators using robust regression methods,” *Communications in Statistics - Theory and Methods*, vol. 48, no. 8, pp. 2039–2048, 2019.
- [10] H. Bulut and T. Zaman, “An improved class of robust ratio estimators by using the minimum covariance determinant estimation,” *Communications in Statistics - Simulation and Computation*, 2019.
- [11] N. Ali, I. Ahmad, M. Hanif, and U. Shahzad, “Robust-regression-type estimators for improving mean estimation of sensitive variables by using auxiliary information,” *Communications in Statistics - Theory and Methods*, vol. 50, no. 4, pp. 979–992, 2019.
- [12] T. Zaman and H. Bulut, “Modified regression estimators using robust regression methods and covariance matrices in stratified random sampling,” *Communications in Statistics - Theory and Methods*, vol. 49, no. 14, pp. 3407–3420, 2020.
- [13] C. E. Särndal, “Sampling survey theory vs. general statistical theory: Estimation of the population mean,” *International Statistical Institute*, vol. 40, pp. 1–12, 1972.
- [14] S. L. Warner, “Randomized response: a survey technique for eliminating evasive answer bias,” *Journal of the American Statistical Association*, vol. 60, no. 309, pp. 63–69, 1965.
- [15] J. A. Fox and P. E. Tracy, *Randomized Response: A Method for Sensitive Surveys*, Chapman & Hall, London, United Kingdom, 1986.
- [16] P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection: Wiley Series in Probability and Mathematical Statistics*, Wiley, Hoboken, New Jersey, 1987.
- [17] V. J. Yohai, “High breakdown-point and high efficiency robust estimates for regression,” *Annals of Statistics*, vol. 15, no. 2, pp. 642–656, 1987.
- [18] P. J. Huber, “Robust regression: asymptotics, conjectures and Monte Carlo,” *Annals of Statistics*, vol. 1, no. 5, pp. 799–821, 1973.
- [19] F. R. Hampel, “A general qualitative definition of robustness,” *The Annals of Mathematical Statistics*, vol. 42, no. 6, pp. 1887–1896, 1971.
- [20] J. W. Tukey, *Exploratory Data Analysis*, Addison-Wesley, Boston, 1977.
- [21] P. J. Rousseeuw, “Multivariate estimation with high breakdown point,” in *Mathematical Statistics and Applications*, W. Grossmann, G. C. Pflug, I. Vincze, and W. Wertz, Eds., Springer, Berlin, Germany, 1985.
- [22] H. Bulut and Y. Öner, “The evaluation of socio-economic development of development agency regions in Turkey using classical and robust principal component analyses,” *Journal of Applied Statistics*, vol. 44, no. 16, pp. 2936–2948, 2017.
- [23] N. H. Al-Noor and A. A. Mohammad, “Model of robust regression with parametric and non-parametric methods,” *Mathematical Theory and Modeling*, vol. 3, no. 5, pp. 27–39, 2013.
- [24] E. J. Pedhazur, *Multiple Regression in Behavioral Research: Explanation and Prediction*, Thompson Learning, Chicago, Illinois, 1997.
- [25] A. Chaudhuri and T. Christofides, *Indirect Questioning in Sample Surveys*, Springer, Berlin, Germany, 2013.
- [26] B. G. Greenberg, R. R. Kuebler, J. R. Abernathy, and D. G. Horvitz, “Application of the randomized response technique in obtaining quantitative data,” *Journal of the American Statistical Association*, vol. 66, no. 334, pp. 243–250, 1971.
- [27] S. A. Eriksson, “A new model for randomized response,” *International Statistical Review/Revue Internationale de Statistique*, vol. 41, no. 1, pp. 101–113, 1973.
- [28] K. H. Pollock and Y. Bek, “A comparison of three randomized response models for quantitative data,” *Journal of the American Statistical Association*, vol. 71, no. 356, pp. 884–886, 1976.
- [29] A. Saha, “A randomized response technique for quantitative data under unequal probability sampling,” *Journal of Statistical Theory and Practice*, vol. 2, no. 4, pp. 589–596, 2008.
- [30] S. K. Bar-Lev, E. Bobovitch, and B. Boukai, “A note on randomized response models for quantitative data,” *Metrika*, vol. 60, no. 3, pp. 255–260, 2004.
- [31] M. N. Murthy, *Sampling Theory and Methods*, Statistical Publishing Society, a, 1967.
- [32] U. Shahzad, M. Hanif, and N. Koyuncu, “A new estimator for mean under stratified random sampling,” *Mathematical Sciences*, vol. 12, no. 3, pp. 163–169, 2018.
- [33] S. Demir and Ö. Toktamiş, “On the adaptive Nadaraya-Watson kernel regression estimators,” *Hacettepe Journal of Mathematics and Statistics*, vol. 39, no. 3, pp. 429–437, 2010.
- [34] M. Khan and J. Shabbir, “Some improved ratio, product and regression estimators of finite population mean when using minimum and maximum values,” *The Scientific World Journal*, vol. 2013, no. 0, pp. 1–7, 2013.
- [35] S. Ali, M. Khan, and J. Shabbir, “Using extreme values and fractional raw moments for mean estimation in stratified sampling,” *Hacettepe Journal of Mathematics and Statistics*, vol. 47, no. 2, pp. 383–402, 2018.
- [36] E. Zahid, J. Shabbir, and O. A. Alamri, “A generalized class of estimators for sensitive variable in the presence of measurement error and non-response under stratified random sampling,” *Journal of King Saud University Science*, vol. 34, no. 2, pp. 101741–101751, 2022.