

Robust mean estimation in stratified sampling: A quantile regression approach

A. SUBHANI, M. IRFAN, AND M. JAVED

Abstract

The presence of outliers can significantly reduce the reliability of conventional estimators, often resulting in biased estimation of the population mean. This challenge is particularly relevant in practical survey data, where irregular observations frequently occur. Although this issue is well-recognized, limited work has been done on population mean estimation under stratified random sampling when outliers are present, especially within a quantile regression framework. To address this gap, the present study introduces a new class of robust estimators based on quantile regression. The proposed estimators make use of non-conventional auxiliary information to improve estimation accuracy under a stratified random sampling scheme. By relying on quantile-based methods, the suggested approach provides greater resistance to the influence of outliers. The statistical properties of the proposed estimators are derived analytically, including measures of bias and mean squared error. In addition, their performance is examined through a real-life dataset. The findings indicate that the proposed estimators offer notable gains in efficiency and robustness compared to the adopted estimators, particularly in datasets affected by outliers.

Mathematics Subject Classification 2000: 62D05

Keywords: Efficiency, Mean Squared Error (MSE), Non-Conventional Measures, Outliers, Quantile Regression, Stratified Random Sampling.

1. INTRODUCTION

Sampling theory deals with the selection of a representative subset from a larger population. It finds applications in diverse fields, including science, education, bioengineering, healthcare, social sciences, and marketing. In practice, researchers typically rely on samples rather than complete populations, as samples are more practical to handle, easier to control, and more efficient to analyze. The selection of an appropriate sampling method depends on the nature of the population. For a homogeneous population, simple random sampling is generally suitable. In contrast, when the population is heterogeneous, stratified random sampling is more appropriate. Stratified random sampling is especially useful when the population consists of distinct subgroups or strata. By dividing a heterogeneous population into

¹ This research was supported by the Funding Agency.

homogeneous strata, this method improves the precision and reliability of estimates and allows more accurate inferences about the overall population, and it enables comparisons among different population groups. The stratified mean estimator outperforms the simple random sample mean when the variation between stratum means is substantial compared to the variation within strata. The precision gained from stratification depends on the method used to select units from each stratum. Key considerations for implementing stratified sampling include (1) determining the optimal number of strata (2) allocating the total sample size to different strata (3) selecting the appropriate strata boundaries. These factors are vital for maintaining the efficiency and accuracy of the stratified mean estimator. In sampling research, supplementary information is frequently utilized to improve the precision and performance of the estimators used to measure population characteristics such as the total, mean, median, proportion, and variance. This information is commonly incorporated into ratio, product, regression, exponential-, logarithmic- and difference-type estimators. For instance, (Kadilar et al., 2005) suggested a ratio estimator under stratified random sampling. Their suggested estimator is more efficient as compared to the traditional combined ratio estimators. (Shabbir et al., 2006) used an alternative estimator derived by (Bedi et al., 1996) under stratified random sampling. This was based on a transformation, which was proposed to be quite effective. Their comparison revealed that the methodology is more effective than the traditional combined ratio estimator and (Kadilar et al., 2005) estimator. (Hussain et al., 2020) constructed two new estimators under simple and stratified random sampling schemes. (Cekim et al., 2020) proposed a new ratio estimator of population variance in stratified random sampling using natural logarithm function. (Javed and Irfan, 2020) proposed new optimal estimators for population mean by using dual auxiliary information in stratified random sampling. Outliers can dramatically influence the performance of estimators, and it is therefore desirable to develop robust estimation methods that can mitigate their effects. Strong techniques are used to overcome this negative influence on the estimators e.g., (Kadilar et al., 2007, Subzar et al., 2019, Zaman et al., 2019, Zaman et al., 2020, Ali et al., 2021, Grover et al., 2021, Koc et al.,

2021, Koc and Koc, 2023, Zaman et al., 2022 and Zaman et al., 2024) and numerous others.

Quantile regression is one of the new techniques which are used to control the effect of extreme values/outliers. (Shahzad et al., 2021) applied the minimum covariance determinant-based measurements in a quantile regression framework to maximize the precision of the estimates, thus, increasing the resilience to outliers, created a new framework of regression ratio-type estimation. (Shahzad et al., 2022) suggested a family of quantile regression-ratio-type estimators of an unknown population mean by using traditional and non-traditional auxiliary measures. The authors used both complete auxiliary information (simple random sampling) and partial auxiliary information (two-phase sampling) for all the proposed estimators. (Anas et al., 2022) developed an efficient class of estimators for the finite population mean by utilizing a quantile regression framework in the presence of missing data. (Khalid et al., 2023) suggested new estimators for the population mean under simple random sampling based on a quantile regression approach. Their findings showed a significant reduction in the effects of extreme observations. (Shahzad et al., 2023) suggested efficient estimators for the estimation of population mean under systematic sampling through quantile regression technique. (Alomair and Iftikhar, 2024) introduced a new dual-type estimator for the population mean that incorporates quantile regression to handle outliers. Its efficiency was evaluated using asymmetric datasets and compared with existing estimators under both scrambled and non-scrambled response conditions. (Afzal et al., 2026) proposed novel estimators for the estimation of population mean under circular systematic sampling.

The present study fills an important gap in sampling theory by developing new estimators for the population mean under a stratified random sampling. Unlike conventional approaches, the proposed methodology incorporates quantile regression which is more robust in the presence of outliers and non-normal data structures.

2. SAMPLING METHOD, NOTATIONS AND ADOPTED ESTIMATORS

The population, comprising N units, is distributed into L distinct strata $N_1, N_2, N_3, \dots, N_L$ that collectively form the entire population without overlap. Within each stratum, individual units are represented by total of $n_1, n_2, n_3, \dots, n_L$ samples drawn from these strata, with each stratum treated as a separate population. In stratified random sampling framework, the mean estimators for the study variable Y and the auxiliary variable X are specifically denoted by $\bar{y}_{st}$ and $\bar{x}_{st}$, respectively. $\bar{y}_{st}$ and $\bar{x}_{st}$ are calculated as: $\bar{y}_{st} = \sum_{h=1}^L W_h \bar{y}_h$ and $\bar{x}_{st} = \sum_{h=1}^L W_h \bar{x}_h$ where $W_h = \frac{N_h}{N}$ is the stratum weight. The population mean of the study variable and auxiliary variable are calculated as: $\bar{Y}_{st} = \sum_{h=1}^L W_h \bar{Y}_h = \bar{Y}$ and $\bar{X}_{st} = \sum_{h=1}^L W_h \bar{X}_h = \bar{X}$. To derive the expressions for the bias, mean squared error (MSE) and minimum MSE of the suggested estimators, we define the relative error terms as:

$$\zeta_0 = \frac{\bar{y}_h - \bar{Y}_h}{\bar{Y}_h}, \zeta_1 = \frac{\bar{x}_h - \bar{X}_h}{\bar{X}_h} \text{ and } \zeta_2 = \frac{\hat{\beta}_{(hq)} - \beta_{(hq)}}{\beta_{(hq)}} \quad \text{such that} \quad E(\zeta_0) = E(\zeta_1) = E(\zeta_2) = 0,$$

$$E(\zeta_0^2) = \theta_h C_{yh}^2, E(\zeta_1^2) = \theta_h C_{xh}^2, E(\zeta_0 \zeta_1) = \theta_h C_{yxh}, E(\zeta_1 \zeta_2) = \frac{\text{Cov}(\bar{x}_h \hat{\beta}_{(hq)})}{\bar{X}_h \beta_{(hq)}} \text{ where}$$

$$\bar{y}_h = n_h^{-1} \sum_{i=1}^{n_h} y_{hi}, \bar{x}_h = n_h^{-1} \sum_{i=1}^{n_h} x_{hi}, \bar{Y}_h = N_h^{-1} \sum_{i=1}^{N_h} Y_{hi}, \bar{X}_h = N_h^{-1} \sum_{i=1}^{N_h} X_{hi},$$

$$C_{yh}^2 = (\bar{Y}_h^2)^{-1} S_{yh}^2, C_{xh}^2 = (\bar{X}_h^2)^{-1} S_{xh}^2, S_{yh}^2 = (N_h - 1)^{-1} \sum_{i=1}^N (y_{hi} - \bar{Y}_h)^2,$$

$$S_{xh}^2 = (N_h - 1)^{-1} \sum_{i=1}^N (x_{hi} - \bar{X}_h)^2, S_{yxh} = (N_h - 1)^{-1} \sum_{i=1}^N (y_{hi} - \bar{Y}_h)(x_{hi} - \bar{X}_h) \text{ and}$$

$$\theta_h = \left(\frac{1}{n_h} - \frac{1}{N_h} \right).$$

Remark 2.1: If $\bar{x}_h$ and $\hat{\beta}_{(hq)}$ are independent then

$$\text{Cov}(\bar{x}_h, \hat{\beta}_{(hq)}) = 0 \Rightarrow E(\zeta_1 \zeta_2) = 0.$$

(Kadilar and Cingi, 2004) introduced ratio-type estimators for the population mean estimation under simple random sampling. In this study, their approach is extended to the stratified random sampling framework by incorporating a quantile regression technique. Both conventional/traditional and non-conventional measures of auxiliary

variable are also utilized in the proposed methodology. Non-conventional measures in sampling theory are effective in reducing the impact of outliers. These measures provide robust alternatives to conventional measures that are highly influenced by outliers. Three reference quantiles are used: $q_{25th} = 0.25$, $q_{50th} = 0.50$ and $q_{75th} = 0.75$. Eq. (1) presents the separate quantile regression estimators adopted from (Kadilar and Cingi, 2004).

$$\hat{Y}_{Pi} = \sum_{h=1}^L W_h \left[\frac{\bar{y}_h + \hat{\beta}_{(hq)}(\bar{X}_h - \bar{x}_h)}{a\bar{x}_h + b} (a\bar{X}_h + b) \right] \quad (1)$$

where $\hat{\beta}_{(hq)}$ is the quantile regression coefficient and "a & b" are the unknown constants such as $C_{xh}, \rho_{yxh}, Q_{Ah}, T_{Mh}$ and H_{Lh} . Conventional/non-conventional parameters associated with the auxiliary variable are as below in Table 1. By replacing the constants "a and b" in Eq. (1), a variety of more refined and efficient estimators can be developed. Several such optimized estimators are listed and described in Table 2.

Expressing estimator " $\hat{Y}_{Pi}$ " in terms of relative error terms, we have

$$\hat{Y}_{Pi} = \sum_{h=1}^L W_h [\bar{Y}_h(1 + \zeta_0) - \beta_{(hq)}\bar{X}_h\zeta_1 - \beta_{(hq)}\bar{X}_h\zeta_1\zeta_2](1 + v_i\zeta_1)^{-1} \quad (2)$$

where $v_i = \frac{a\bar{x}_h}{a\bar{x}_h + b}, i = 1, 2, 3, \dots, 12$.

$$\begin{aligned} \hat{Y}_{Pi} = \sum_{h=1}^L W_h [\bar{Y}_h - v_i\bar{Y}_h\zeta_1 + v_i^2\bar{Y}_h\zeta_1^2 + \bar{Y}_h\zeta_0 - v_i\bar{Y}_h\zeta_0\zeta_1 - \beta_{(hq)}\bar{X}_h\zeta_1 + v_i\beta_{(hq)}\bar{X}_h\zeta_1^2 \\ - \beta_{(hq)}\bar{X}_h\zeta_1\zeta_2] \end{aligned} \quad (3)$$

Subtracting $\bar{Y}$ on both sides of Eq. (3), we get

$$\begin{aligned} (\hat{Y}_{Pi} - \bar{Y}) = \sum_{h=1}^L W_h [-v_i\bar{Y}_h\zeta_1 + v_i^2\bar{Y}_h\zeta_1^2 + \bar{Y}_h\zeta_0 - v_i\bar{Y}_h\zeta_0\zeta_1 - \beta_{(hq)}\bar{X}_h\zeta_1 + v_i\beta_{(hq)}\bar{X}_h\zeta_1^2 \\ - \beta_{(hq)}\bar{X}_h\zeta_1\zeta_2] \end{aligned} \quad (4)$$

Taking expectation on both sides of the Eq. (4), we get Bias($\hat{Y}_{Pi}$)

$$\begin{aligned} Bias(\hat{Y}_{Pi}) \cong \sum_{h=1}^L W_h \theta_h [v_i^2\bar{Y}_h C_{xh}^2 - v_i\bar{Y}_h C_{yxh} + v_i\beta_{(hq)}\bar{X}_h C_{xh}^2] \\ - \sum_{h=1}^L W_h Cov(\bar{x}_h, \hat{\beta}_{(hq)}) \end{aligned} \quad (5)$$

Squaring both sides of Eq. (4), we get

$$\begin{aligned} (\hat{Y}_{Pi} - \bar{Y})^2 = \sum_{h=1}^L W_h^2 [v_i^2 \bar{Y}_h^2 \zeta_1^2 + \bar{Y}_h^2 \zeta_0^2 + \beta_{(hq)}^2 \bar{X}_h^2 \zeta_1^2 - 2v_i \bar{Y}_h^2 \zeta_0 \zeta_1 + 2v_i \beta_{(hq)} \bar{X}_h \bar{Y}_h \zeta_1^2 \\ - 2\beta_{(hq)} \bar{X}_h \bar{Y}_h \zeta_0 \zeta_1] \end{aligned} \quad (6)$$

Taking expectation on both sides of the Eq. (6), we get $MSE(\hat{Y}_{Pi})$

$$\begin{aligned} MSE(\hat{Y}_{Pi}) \cong \sum_{h=1}^L W_h^2 \theta_h^2 [(v_i \bar{Y}_h + \beta_{(hq)} \bar{X}_h)^2 C_{xh}^2 + \bar{Y}_h^2 C_{yh}^2 \\ - 2(v_i \bar{Y}_h^2 + \beta_{(hq)} \bar{X}_h \bar{Y}_h) C_{yxh}] \end{aligned} \quad (7)$$

Table 1. Conventional and non-conventional measures

Conventional Measures	Non-conventional Measures
Coefficient of variation: C_{xh}	Quartile average: $Q_{Ah} = \frac{Q_3 + Q_1}{2}$
Coefficient of kurtosis: $\beta_{2(xh)}$	Tri-mean: $T_{Mh} = \frac{Q_1 + 2Q_2 + Q_3}{4}$
Coefficient of correlation: $\rho_{yxh} = (S_{yh} S_{xh})^{-1} S_{yxh}$	Hodges Lehmann estimator: $H_{Lh} = \text{Median}\left(\frac{x_j + x_k}{2}\right) \quad 1 \leq j \leq k \leq N$

Table 2. Family members of $\hat{Y}_{Pi}$

a	b	$\hat{Y}_{Pi}$
1	C_{xh}	$\hat{Y}_{P1} = \sum_{h=1}^L W_h \left[\frac{\bar{y}_h + \hat{\beta}_{(hq)}(\bar{X}_h - \bar{x}_h)}{\bar{x}_h + C_{xh}} (\bar{X}_h + C_{xh}) \right]$
1	$\beta_{2(xh)}$	$\hat{Y}_{P2} = \sum_{h=1}^L W_h \left[\frac{\bar{y}_h + \hat{\beta}_{(hq)}(\bar{X}_h - \bar{x}_h)}{\bar{x}_h + \beta_{2(xh)}} (\bar{X}_h + \beta_{2(xh)}) \right]$
$\beta_{2(xh)}$	C_{xh}	$\hat{Y}_{P3} = \sum_{h=1}^L W_h \left[\frac{\bar{y}_h + \hat{\beta}_{(hq)}(\bar{X}_h - \bar{x}_h)}{\beta_{2(xh)} \bar{x}_h + C_{xh}} (\beta_{2(xh)} \bar{X}_h + C_{xh}) \right]$
C_{xh}	$\beta_{2(xh)}$	$\hat{Y}_{P4} = \sum_{h=1}^L W_h \left[\frac{\bar{y}_h + \hat{\beta}_{(hq)}(\bar{X}_h - \bar{x}_h)}{C_{xh} \bar{x}_h + \beta_{2(xh)}} (C_{xh} \bar{X}_h + \beta_{2(xh)}) \right]$
ρ_{yxh}	C_{xh}	$\hat{Y}_{P5} = \sum_{h=1}^L W_h \left[\frac{\bar{y}_h + \hat{\beta}_{(hq)}(\bar{X}_h - \bar{x}_h)}{\rho_{yxh} \bar{x}_h + \beta_{2(xh)}} (\rho_{yxh} \bar{X}_h + C_{xh}) \right]$

C_{xh}	ρ_{yxh}	$\hat{Y}_{P6} = \sum_{h=1}^L W_h \left[\frac{\bar{y}_h + \hat{\beta}_{(hq)}(\bar{X}_h - \bar{x}_h)}{C_{xh}\bar{x}_h + \rho_{yxh}} (C_{xh}\bar{X}_h + \rho_{yxh}) \right]$
T_{Mh}	Q_{Ah}	$\hat{Y}_{P7} = \sum_{h=1}^L W_h \left[\frac{\bar{y}_h + \hat{\beta}_{(hq)}(\bar{X}_h - \bar{x}_h)}{T_{Mh}\bar{x}_h + Q_{Ah}} (T_{Mh}\bar{X}_h + Q_{Ah}) \right]$
Q_{Ah}	T_{Mh}	$\hat{Y}_{P8} = \sum_{h=1}^L W_h \left[\frac{\bar{y}_h + \hat{\beta}_{(hq)}(\bar{X}_h - \bar{x}_h)}{Q_{Ah}\bar{x}_h + T_{Mh}} (Q_{Ah}\bar{X}_h + T_{Mh}) \right]$
T_{Mh}	H_{Lh}	$\hat{Y}_{P9} = \sum_{h=1}^L W_h \left[\frac{\bar{y}_h + \hat{\beta}_{(hq)}(\bar{X}_h - \bar{x}_h)}{T_{Mh}\bar{x}_h + H_{Lh}} (T_{Mh}\bar{X}_h + H_{Lh}) \right]$
H_{Lh}	T_{Mh}	$\hat{Y}_{P10} = \sum_{h=1}^L W_h \left[\frac{\bar{y}_h + \hat{\beta}_{(hq)}(\bar{X}_h - \bar{x}_h)}{H_{Lh}\bar{x}_h + T_{Mh}} (H_{Lh}\bar{X}_h + T_{Mh}) \right]$
Q_{Ah}	H_{Lh}	$\hat{Y}_{P11} = \sum_{h=1}^L W_h \left[\frac{\bar{y}_h + \hat{\beta}_{(hq)}(\bar{X}_h - \bar{x}_h)}{Q_{Ah}\bar{x}_h + H_{Lh}} (Q_{Ah}\bar{X}_h + H_{Lh}) \right]$
H_{Lh}	Q_{Ah}	$\hat{Y}_{P12} = \sum_{h=1}^L W_h \left[\frac{\bar{y}_h + \hat{\beta}_{(hq)}(\bar{X}_h - \bar{x}_h)}{H_{Lh}\bar{x}_h + Q_{Ah}} (H_{Lh}\bar{X}_h + Q_{Ah}) \right]$

3. PROPOSED QUANTILE ROBUST ESTIMATORS

Within a stratified random sampling framework, outliers, whether arising within individual strata or across them, can seriously compromise the reliability of statistical analysis. Their presence can lead to biased estimation of the population mean and can adversely affect regression coefficients derived through the ordinary least squares (OLS) method, thereby undermining the overall accuracy of inference. We employ a quantile regression method to mitigate the influence of outliers, as it is well known for its robustness, particularly when the underlying data distribution deviates from normality (Shahzad et al., 2021, 2022, 2023). We propose a novel class of estimators for the estimation of population mean based on the quantile regression approach. The proposed estimators incorporate information from three key quantiles such as: $q_{25th} = 0.25$, $q_{50th} = 0.50$ and $q_{75th} = 0.75$. These three quantiles correspond to the first quartile, median, and third quartile, which together provide a balanced representation of the lower, central, and upper parts of the distribution. A generalized form of the

proposed estimators is presented in Eq. (8). This formulation provides a unified structure that encompasses various special classes of estimators under different choices of parameters associated with the auxiliary variable.

$$\hat{Y}_{pti} = \sum_{h=1}^L W_h \left[\alpha_h \bar{y}_h \left(\frac{a\bar{X}_h + b}{a\bar{x}_h + b} \right)^{\frac{a\bar{X}_h}{a\bar{X}_h + b}} + \hat{\beta}_{(hq)} (\bar{X}_h - \bar{x}_h) \right] \quad (8)$$

where α_h is the selected appropriate constant whose value is determined later, $\hat{\beta}_{(hq)}$ is the quantile regression coefficient and "a & b" are the unknown constants.

Remark 3.1. By assigning different unknown population parameters of the auxiliary variable, as listed in Table 1, to the constants "a and b" in Eq. (8), a wide variety of optimal estimators can be generated. Each specific choice of these parameters leads to a distinct form of the estimator, allowing flexibility in adapting the estimation procedure to different population characteristics. Consequently, this approach yields a broad class of optimal estimators. In the present study, several representative estimators derived through this framework are reported in Table 3.

Table 3. Family members of $\hat{Y}_{pti}$

a	b	$\hat{Y}_{pti}$
1	C_{xh}	$\hat{Y}_{pt1} = \sum_{h=1}^L W_h \left[\alpha_h \bar{y}_h \left(\frac{\bar{X}_h + C_{xh}}{\bar{x}_h + C_{xh}} \right)^{\frac{\bar{X}_h}{\bar{x}_h + C_{xh}}} + \hat{\beta}_{(hq)} (\bar{X}_h - \bar{x}_h) \right]$
1	$\beta_{2(xh)}$	$\hat{Y}_{pt2} = \sum_{h=1}^L W_h \left[\alpha_h \bar{y}_h \left(\frac{\bar{X}_h + \beta_{2(xh)}}{\bar{x}_h + \beta_{2(xh)}} \right)^{\frac{\bar{X}_h}{\bar{x}_h + \beta_{2(xh)}}} + \hat{\beta}_{(hq)} (\bar{X}_h - \bar{x}_h) \right]$
$\beta_{2(xh)}$	C_{xh}	$\hat{Y}_{pt3} = \sum_{h=1}^L W_h \left[\alpha_h \bar{y}_h \left(\frac{\beta_{2(xh)}\bar{X}_h + C_{xh}}{\beta_{2(xh)}\bar{x}_h + C_{xh}} \right)^{\frac{\beta_{2(xh)}\bar{X}_h}{\beta_{2(xh)}\bar{x}_h + C_{xh}}} + \hat{\beta}_{(hq)} (\bar{X}_h - \bar{x}_h) \right]$
C_{xh}	$\beta_{2(xh)}$	$\hat{Y}_{pt4} = \sum_{h=1}^L W_h \left[\alpha_h \bar{y}_h \left(\frac{C_{xh}\bar{X}_h + \beta_{2(xh)}}{C_{xh}\bar{x}_h + \beta_{2(xh)}} \right)^{\frac{C_{xh}\bar{X}_h}{C_{xh}\bar{x}_h + \beta_{2(xh)}}} + \hat{\beta}_{(hq)} (\bar{X}_h - \bar{x}_h) \right]$

ρ_{yxh}	C_{xh}	$\hat{Y}_{Pt5} = \sum_{h=1}^L W_h \left[\alpha_h \bar{y}_h \left(\frac{\rho_{yxh} \bar{X}_h + C_{xh}}{\rho_{yxh} \bar{x}_h + C_{xh}} \right)^{\frac{\rho_{yxh} \bar{X}_h}{\rho_{yxh} \bar{X}_h + C_{xh}}} + \hat{\beta}_{(hq)} (\bar{X}_h - \bar{x}_h) \right]$
C_{xh}	ρ_{yxh}	$\hat{Y}_{Pt6} = \sum_{h=1}^L W_h \left[\alpha_h \bar{y}_h \left(\frac{C_{xh} \bar{X}_h + \rho_{yxh}}{C_{xh} \bar{x}_h + \rho_{yxh}} \right)^{\frac{C_{xh} \bar{X}_h}{C_{xh} \bar{X}_h + \rho_{yxh}}} + \hat{\beta}_{(hq)} (\bar{X}_h - \bar{x}_h) \right]$
T_{Mh}	Q_{Ah}	$\hat{Y}_{Pt7} = \sum_{h=1}^L W_h \left[\alpha_h \bar{y}_h \left(\frac{T_{Mh} \bar{X}_h + Q_{Ah}}{T_{Mh} \bar{x}_h + Q_{Ah}} \right)^{\frac{T_{Mh} \bar{X}_h}{T_{Mh} \bar{X}_h + Q_{Ah}}} + \hat{\beta}_{(hq)} (\bar{X}_h - \bar{x}_h) \right]$
Q_{Ah}	T_{Mh}	$\hat{Y}_{Pt8} = \sum_{h=1}^L W_h \left[\alpha_h \bar{y}_h \left(\frac{Q_{Ah} \bar{X}_h + T_{Mh}}{Q_{Ah} \bar{x}_h + T_{Mh}} \right)^{\frac{Q_{Ah} \bar{X}_h}{Q_{Ah} \bar{X}_h + T_{Mh}}} + \hat{\beta}_{(hq)} (\bar{X}_h - \bar{x}_h) \right]$
T_{Mh}	H_{Lh}	$\hat{Y}_{Pt9} = \sum_{h=1}^L W_h \left[\alpha_h \bar{y}_h \left(\frac{T_{Mh} \bar{X}_h + H_{Lh}}{T_{Mh} \bar{x}_h + H_{Lh}} \right)^{\frac{T_{Mh} \bar{X}_h}{T_{Mh} \bar{X}_h + H_{Lh}}} + \hat{\beta}_{(hq)} (\bar{X}_h - \bar{x}_h) \right]$
H_{Lh}	T_{Mh}	$\hat{Y}_{Pt10} = \sum_{h=1}^L W_h \left[\alpha_h \bar{y}_h \left(\frac{H_{Lh} \bar{X}_h + T_{Mh}}{H_{Lh} \bar{x}_h + T_{Mh}} \right)^{\frac{H_{Lh} \bar{X}_h}{H_{Lh} \bar{X}_h + T_{Mh}}} + \hat{\beta}_{(hq)} (\bar{X}_h - \bar{x}_h) \right]$
Q_{Ah}	H_{Lh}	$\hat{Y}_{Pt11} = \sum_{h=1}^L W_h \left[\alpha_h \bar{y}_h \left(\frac{Q_{Ah} \bar{X}_h + H_{Lh}}{Q_{Ah} \bar{x}_h + H_{Lh}} \right)^{\frac{Q_{Ah} \bar{X}_h}{Q_{Ah} \bar{X}_h + H_{Lh}}} + \hat{\beta}_{(hq)} (\bar{X}_h - \bar{x}_h) \right]$
H_{Lh}	Q_{Ah}	$\hat{Y}_{Pt12} = \sum_{h=1}^L W_h \left[\alpha_h \bar{y}_h \left(\frac{H_{Lh} \bar{X}_h + Q_{Ah}}{H_{Lh} \bar{x}_h + Q_{Ah}} \right)^{\frac{H_{Lh} \bar{X}_h}{H_{Lh} \bar{X}_h + Q_{Ah}}} + \hat{\beta}_{(hq)} (\bar{X}_h - \bar{x}_h) \right]$

A generalized form of the proposed estimators, applicable to all three cases, such as q_{25th} , q_{50th} and q_{75th} , is given in Eq. (9) to Eq. (11).

$$\hat{Y}_{Pti} = \sum_{h=1}^L W_h \left[\alpha_h \bar{y}_h \left(\frac{a \bar{X}_h + b}{a \bar{x}_h + b} \right)^{\frac{a \bar{X}}{a \bar{x}_h + b}} + \hat{\beta}_{(h.25)} (\bar{X}_h - \bar{x}_h) \right] \quad (9)$$

$$\hat{Y}_{Pti} = \sum_{h=1}^L W_h \left[\alpha_h \bar{y}_h \left(\frac{a \bar{X}_h + b}{a \bar{x}_h + b} \right)^{\frac{a \bar{X}}{a \bar{x}_h + b}} + \hat{\beta}_{(h.50)} (\bar{X}_h - \bar{x}_h) \right] \quad (10)$$

$$\hat{Y}_{Pti} = \sum_{h=1}^L W_h \left[\alpha_h \bar{y}_h \left(\frac{a\bar{X}_h + b}{a\bar{x}_h + b} \right)^{\frac{a\bar{x}}{a\bar{x}_h + b}} + \hat{\beta}_{(h.75)} (\bar{X}_h - \bar{x}_h) \right] \quad (11)$$

Expressing estimator " $\hat{Y}_{Pti}$ " in terms of relative error terms, we have

$$\begin{aligned} \hat{Y}_{Pti} = \sum_{h=1}^L W_h \left[\alpha_h \bar{Y}_h - \alpha_h \bar{Y}_h v_i^2 \zeta_1 + \alpha_h \bar{Y}_h \zeta_o - \alpha_h \bar{Y}_h v_i^2 \zeta_o \zeta_1 - \beta_{(hq)} \bar{X}_h \zeta_1 - \beta_{(hq)} \bar{X}_h \zeta_1 \zeta_2 \right. \\ \left. + \frac{1}{2} \alpha_h \bar{Y}_h v_i^3 (v_i + 1) \zeta_1^2 \right] \end{aligned} \quad (12)$$

Subtracting $\bar{Y}$ on both sides of Eq. (12), we get

$$\begin{aligned} (\hat{Y}_{Pti} - \bar{Y}) = \sum_{h=1}^L W_h \left[\alpha_h \bar{Y}_h - \bar{Y}_h - \alpha_h \bar{Y}_h v_i^2 \zeta_1 + \alpha_h \bar{Y}_h \zeta_o - \alpha_h \bar{Y}_h v_i^2 \zeta_o \zeta_1 - \beta_{(hq)} \bar{X}_h \zeta_1 \right. \\ \left. - \beta_{(hq)} \bar{X}_h \zeta_1 \zeta_2 + \frac{1}{2} \alpha_h \bar{Y}_h v_i^3 (v_i + 1) \zeta_1^2 \right] \end{aligned} \quad (13)$$

Taking expectation on both side of Eq. (13), we get Bias($\hat{Y}_{Pti}$)

$$\begin{aligned} \text{Bias}(\hat{Y}_{Pti}) \cong \sum_{h=1}^L W_h \bar{Y}_h (\alpha_h - 1) + \sum_{h=1}^L W_h \theta_h \alpha_h v_i^2 \bar{Y}_h \left[\frac{v_i(v_i + 1)}{2} C_{xh}^2 - C_{yxh} \right] \\ - \sum_{h=1}^L W_h \text{Cov}(\bar{x}_h, \hat{\beta}_{(hq)}) \end{aligned}$$

Taking the square and expectation on both sides of Eq. (13), we get $MSE(\hat{Y}_{Pti})$

$$\begin{aligned} MSE(\hat{Y}_{Pti}) \cong \sum_{h=1}^L W_h^2 (\alpha_h \bar{Y}_h - \bar{Y}_h)^2 + \sum_{h=1}^L W_h^2 \theta_h \beta_{(hq)}^2 \bar{X}_h^2 C_{xh}^2 \\ + \alpha_h^2 \sum_{h=1}^L W_h^2 \theta_h \bar{Y}_h^2 (C_{yh}^2 + 2v_i^4 C_{xh}^2 + v_i^3 C_{xh}^2 - 4v_i^2 C_{yxh}) \\ - \alpha_h \sum_{h=1}^L W_h^2 \theta_h \bar{Y}_h (\bar{Y}_h v_i^4 C_{xh}^2 + \bar{Y}_h v_i^3 C_{xh}^2 - 2\beta_{(hq)} v_i^2 \bar{X}_h C_{xh}^2 - 2\bar{Y}_h v_i^2 C_{yxh} + 2\bar{X}_h \beta_{(hq)} C_{yxh}) \end{aligned}$$

After some simplification,

$$MSE(\hat{Y}_{Pti}) \cong \sum_{h=1}^L W_h^2 (\alpha_h \bar{Y}_h - \bar{Y}_h)^2 + \sum_{h=1}^L W_h^2 \theta_h \beta_{(hq)}^2 \bar{X}_h^2 C_{xh}^2 + \alpha_h^2 A_1 - \alpha_h A_2 \quad (14)$$

where

$$A_1 = \sum_{h=1}^L W_h^2 \theta_h \bar{Y}_h^2 (C_{yh}^2 + 2v_i^4 C_{xh}^2 + v_i^3 C_{xh}^2 - 4v_i^2 C_{yxh})$$

$$A_2 = \sum_{h=1}^L W_h^2 \theta_h \bar{Y}_h (\bar{Y}_h v_i^4 C_{xh}^2 + \bar{Y}_h v_i^3 C_{xh}^2 - 2\beta_{(hq)} v_i^2 \bar{X}_h C_{xh}^2 - 2\bar{Y}_h v_i^2 C_{yxh} + 2\bar{X}_h \beta_{(hq)} C_{yxh})$$

To determine the optimal value of α_h , we differentiate Eq. (14) with respect to α_h and impose the first-order optimality condition by setting the resulting derivative equal to zero. Solving this condition yields:

$$\alpha_h = \frac{\bar{Y}_h^2 + 0.5A_2}{\bar{Y}_h^2 + A_1}$$

By substituting the optimal value of α_h into Eq. (14) and carrying out the necessary simplifications, we derive the minimum mean squared error (MSE) of $\hat{Y}_{pti}$.

$$MSE_{min}(\hat{Y}_{pti}) \cong \left[\bar{Y}_h^2 - \frac{A_3^2}{2A_4} \right] \quad (15)$$

where

$$A_3 = 2\bar{Y}_h^2 + A_2$$

$$A_4 = 2\bar{Y}_h^2 + 2A_1$$

Theorem 3.1. When there is no information regarding unknown parameters of the auxiliary variable in advance then Eq. (8) can be modified as:

$$\hat{Y}_{pt} = \sum_{h=1}^L W_h \left[\alpha_h \bar{y}_h \frac{\bar{X}_h}{\bar{x}_h} + \hat{\beta}_{(hq)} (\bar{X}_h - \bar{x}_h) \right] \quad (16)$$

Using three different quantile regression coefficients such as q_{25th} , q_{50th} and q_{75th} three estimators have been constructed, as shown below.

$$\begin{cases} \hat{Y}_{pt} = \sum_{h=1}^L W_h \left[\alpha_h \bar{y}_h \frac{\bar{X}_h}{\bar{x}_h} + \hat{\beta}_{(h.25)} (\bar{X}_h - \bar{x}_h) \right] \\ \hat{Y}_{pt} = \sum_{h=1}^L W_h \left[\alpha_h \bar{y}_h \frac{\bar{X}_h}{\bar{x}_h} + \hat{\beta}_{(h.50)} (\bar{X}_h - \bar{x}_h) \right] \\ \hat{Y}_{pt} = \sum_{h=1}^L W_h \left[\alpha_h \bar{y}_h \frac{\bar{X}_h}{\bar{x}_h} + \hat{\beta}_{(h.75)} (\bar{X}_h - \bar{x}_h) \right] \end{cases} \quad (17)$$

We get bias, MSE and minimum MSE of $\hat{Y}_{pt}$ as:

$$\text{Bias}(\hat{Y}_{pt}) \cong \sum_{h=1}^L W_h \bar{Y}_h (\alpha_h - 1) + \sum_{h=1}^L W_h \theta_h \alpha_h \bar{Y}_h (C_{xh}^2 - C_{yxh}) - \sum_{h=1}^L W_h \text{Cov}(\bar{x}_h, \hat{\beta}_{(hq)})$$

$$MSE(\hat{\bar{Y}}_{Pt}) \cong \sum_{h=1}^L W_h^2 (\alpha_h \bar{Y}_h - \bar{Y}_h)^2 + \sum_{h=1}^L W_h^2 \theta_h \beta_{(hq)}^2 \bar{X}_h^2 C_{xh}^2 + \alpha_h^2 A_5 - \alpha_h A_6 \quad (18)$$

$$A_5 = \sum_{h=1}^L W_h^2 \theta_h \bar{Y}_h^2 (C_{yh}^2 + 3C_{xh}^2 - 4C_{yxh})$$

$$A_6 = \sum_{h=1}^L W_h^2 \theta_h \bar{Y}_h (2\bar{Y}_h C_{xh}^2 - 2\beta_{(hq)} \bar{X}_h C_{xh}^2 - 2\bar{Y}_h C_{yxh} + 2\bar{X}_h \beta_{(hq)} C_{yxh})$$

To obtain the optimal value of α_h , differentiating Eq. (18), with respect to α_h and equating to zero, we have:

$$\alpha_h = \frac{2\bar{Y}_h^2 + A_6}{2\bar{Y}_h^2 + 2A_5} = \frac{A_7}{A_8}$$

Putting the optimal value of α_h in Eq. (18), and simplifying, we have the minimum MSE

$$MSE_{min}(\hat{\bar{Y}}_{Pt}) \cong \left[\bar{Y}_h^2 - \frac{A_7^2}{2A_8} \right] \quad (19)$$

4. NUMERICAL ILLUSTRATION

In this section, we assess the performance of the proposed estimators in comparison with the adopted estimators using a real-life dataset. The primary objective of this empirical study is to estimate the average number of teachers in both primary and secondary schools in Turkey. For this purpose, we use the dataset used by (Javed et al., 2019). In this dataset, the total number of teachers in primary and secondary schools is taken as the study variable, while the corresponding student enrollment in these schools is used as the auxiliary variable. The use of an auxiliary variable is motivated by its strong association with the study variable which enhances the efficiency of the estimators. The study population consists of ($N = 923$) districts across Turkey, which are stratified into six geographical regions. To ensure adequate representation from each region and to improve the precision of the estimates, a stratified sampling design is adopted. A total sample of size $n = 180$ districts is selected from the six strata using Neyman allocation, which assigns sample sizes to strata in proportion to their population sizes and variability. The important descriptive

measures of the study and auxiliary variables for each stratum as well as for the overall population are presented in Table 4.

Table 4. Descriptive statistics for the population in h stratum

Symbol	Values of parameters					
	1	2	3	4	5	6
N_h	127	117	103	170	205	201
n_h	31	21	29	38	22	39
C_{xh}	1.465	1.647	1.925	1.922	1.525	1.776
C_{yh}	1.255	1.561	1.803	1.908	1.511	1.811
C_{yxh}	1.724	2.562	3.445	3.608	2.282	3.099
$\hat{\beta}_{(h.25)}$	0.027	0.041	0.036	0.040	0.043	0.025
$\hat{\beta}_{(h.50)}$	0.030	0.042	0.038	0.039	0.048	0.028
$\hat{\beta}_{(h.75)}$	0.035	0.046	0.041	0.048	0.051	0.031
W_h	0.138	0.127	0.112	0.184	0.222	0.212
$\bar{Y}_h$	703.740	413	573.17	424.66	267.03	393.86
$\bar{X}_h$	20804.59	9211.79	14309.30	9478.85	5569.95	12997.59

OLS coefficients and conventional estimators may not always provide efficient or reliable estimates of the population mean, particularly in the presence of outliers or departures from normality. To examine these issues in the dataset, diagnostic plots including histograms, box plots, and scatter plots were generated using R software. The results indicate the presence of possible outliers and departures from normality, suggesting that the assumptions required for OLS estimation may not hold. To overcome these limitations, quantile regression is used as a more robust alternative to OLS, particularly suitable for handling outliers and skewed distributions. Furthermore, non-conventional measures such as the quartile mean, tri-mean, and Hodges-Lehmann estimator are incorporated to enhance robustness. Together, these

methods reduce the impact of outliers and non-normality, resulting in more stable and reliable estimation of the population mean.

We evaluated the performance of the adopted and proposed estimators by calculating their mean squared error (MSE) values. The MSE expressions for the adopted estimators, as presented in Section 2, along with those derived for the proposed estimators up to the first-order approximation in Section 3, are considered for performance evaluation. The computed MSE values corresponding to all estimators are summarized in Table 5 for comparison and analysis. The MSE of the proposed class of estimators, $\hat{Y}_{pti}$, $i = 1, 2, 3, \dots, 12$, are defined as:

$$MSE(\hat{Y}_{pti}) \cong E(\hat{Y}_{pti} - \bar{Y})^2 \quad (20)$$

It is worth noting that when no auxiliary information regarding population parameters is available, the performance of the estimators can differ substantially. We observed that the MSE of the adopted estimators, denoted by $\hat{Y}_p$, is consistently higher than that of the proposed estimators $\hat{Y}_{pt}$ across all considered quantiles, specifically $q_{25th} = 0.25$, $q_{50th} = 0.50$, $q_{75th} = 0.75$. The MSE values for the estimators $\hat{Y}_p$ and $\hat{Y}_{pt}$ are presented as follows. The estimator $\hat{Y}_p$ yields MSE values of 2174.998, 2541.090 and 3050.304. In contrast, the corresponding MSE values for $\hat{Y}_{pt}$ are 1766.971, 2112.276 and 2555.012. A comparison of these results shows that the proposed estimators provide more accurate estimates of the population mean when auxiliary information is not available, highlighting their robustness and efficiency in practical scenarios. Table 5. demonstrates the superior performance of the proposed quantile robust estimators i.e., $\hat{Y}_{pti}$ over the adopted estimators $\hat{Y}_{pi}$. For all examined quantiles, the proposed estimators consistently achieve the lowest MSE. This improvement can be attributed to the ability of quantile-based methods to reduce the influence of outliers and to handle deviations from normality in the data, which are common challenges in real-world datasets. Moreover, it is observed that as the quantile regression coefficient increases, the MSE of both the adopted estimators $\hat{Y}_{pi}$ and the proposed estimators $\hat{Y}_{pti}$ also increases. This trend suggests that higher quantile regression coefficients, which give more weight to extreme observations, tend to amplify variability in the estimators. Overall, these results clearly indicate that the proposed quantile robust

estimators outperform the adopted estimators in terms of MSE, especially when the data contain outliers and exhibit non-normal characteristics.

Table 5. MSE's of adopted $(\hat{\bar{Y}}_{Pi})$ and proposed $(\hat{\bar{Y}}_{Pti})$ estimators

a	b	$\hat{\bar{Y}}_{Pi}$			$\hat{\bar{Y}}_{Pti}$		
		$q = 0.25$	$q = 0.50$	$q = 0.75$	$q = 0.25$	$q = 0.50$	$q = 0.75$
1	C_{xh}	2181.289	2549.319	3053.464	1773.553	2115.734	2563.332
1	$\beta_{2(xh)}$	2175.538	2543.009	3046.660	1769.086	2114.152	2558.763
$\beta_{2(xh)}$	C_{xh}	2181.939	2550.026	3054.233	1774.854	2117.148	2564.872
C_{xh}	$\beta_{2(xh)}$	2178.143	2545.859	3049.744	1767.278	2108.831	2555.909
ρ_{yxh}	C_{xh}	2181.265	2549.303	3053.446	1773.518	2115.703	2563.298
C_{xh}	ρ_{yxh}	2181.747	2549.827	3054.018	1774.482	2116.750	2564.441
T_{Mh}	Q_{Ah}	2181.498	2549.556	3053.722	1773.984	2116.209	2563.849
Q_{Ah}	T_{Mh}	2181.633	2549.703	3053.884	1774.255	2116.503	2564.172
T_{Mh}	H_{Lh}	2181.524	2549.585	3053.754	1774.036	2116.266	2563.913
H_{Lh}	T_{Mh}	2181.615	2549.684	3053.862	1774.219	2116.464	2564.128
Q_{Ah}	H_{Lh}	2181.592	2549.659	3053.835	1774.173	2116.414	2564.075
H_{Lh}	Q_{Ah}	2181.549	2549.611	3053.784	1774.087	2116.320	2563.973

5. CONCLUDING REMARKS

Outliers are an important concern in survey sampling because they can greatly influence the accuracy and consistency of estimates, especially when estimating the population mean. Their presence may lead to biased or inefficient results, particularly when traditional methods depend on assumptions such as normality. For this reason, careful detection and appropriate treatment of outliers are necessary to ensure that the analysis accurately represents the population. To address this issue, quantile regression is adopted as a more robust alternative, as it reduces the impact of outliers and yields more dependable estimates. In this study, a new approach is developed for estimating the population mean in stratified random sampling by employing quantile

regression, resulting in a class of robust quantile-based estimators. The theoretical properties of these estimators including bias, mean squared error (MSE), and optimum MSE, are derived. Although the findings are promising, the present study has some limitations that can be explored and refined in future work. The proposed estimators were evaluated under specific non-normal distributions and contamination scenarios, which may limit their generalizability to other data structures. The analysis was confined to a stratified random sampling scheme, and their performance under other sampling designs was not explored. Future extensions of this study could include:

- Development of quantile regression-based methods for estimating the population median under simple and stratified random sampling using non-conventional measures.
- Estimation of the population mean under various sampling designs by incorporating non-conventional auxiliary information while accounting for the presence of non-sampling errors.
- Estimation of population mean, median, and variance across different sampling schemes using incomplete or partially available auxiliary information.

REFERENCES

- AFZAL, S., IRFAN, M. AND JAVED, M. 2026. Quantile regression-based mean estimation in circular systematic sampling. *Maejo International Journal of Science and Technology* 20, 1, 67–82.
- ALOMAIR, A. M. AND IFTIKHAR, S. 2024. Dual-type quantile regression approach for mean estimation incorporating sampled and non-sampled data. *Heliyon* 10, Art. no. e31034.
- ALI, N., AHMAD, I., HANIF, M. AND SHAHZAD, U. 2021. Robust-regression-type estimators for improving mean estimation of sensitive variables by using auxiliary information. *Communications in Statistics-Theory and Methods* 50, 4, 979–992.
- ANAS, M. M., HUANG, Z., SHAHZAD, U., ZAMAN, T. AND SHAHZADI, S. 2022. Compromised imputation based mean estimators using robust quantile regression. *Communications in Statistics-Theory and Methods* 53, 1700–1715.
- BEDI, P. K. 1996. Efficient utilization of auxiliary information at estimation stage. *Biometrical Journal* 38, 8, 973–976.
- CEKIM, H. O. AND KADILAR, C. 2020. In-type estimators for the population variance in stratified random sampling *Communications in Statistics-Simulation and Computation* 49, 7, 1665–1677.

-
- GROVER, L. K. AND KAUR, A. 2021. 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, 113072.
- HUSSAIN, S., AHMAD, S., SALEEM, M. AND AKHTAR, S. 2020. Finite population distribution function estimation with dual use of auxiliary information under simple and stratified random sampling. *Plos One* 15, 9, e0239098.
- JAVED, M. AND IRFAN, M. 2020. A simulation study: new optimal estimators for population mean by using dual auxiliary information in stratified random sampling. *Journal of Taibah University for Science* 14, 1, 557–568.
- JAVED, M., IRFAN, M. AND PANG, T. 2019. Utilizing bivariate auxiliary information for enhanced estimation of population mean under simple and stratified random sampling schemes. *Journal of the National Science Foundation of Sri Lanka* 47, 2, 199–211.
- KADILAR, C. AND CINGI, H. 2007. Ratio estimators using robust regression. *Hacettepe Journal of Mathematics and Statistics* 36, 2, 181–188.
- KADILAR, C. AND CINGI, H. 2005. A new ratio estimator in stratified random sampling. *Communications in Statistics-Theory and Methods* 34, 3, 597–602.
- KADILAR, C. AND CINGI, H. 2004. Ratio estimators in simple random sampling. *Applied Mathematics and Computation* 151, 3, 893–902.
- KADILAR, C. AND CINGI, H. 2003. Ratio estimators in stratified random sampling. *Biometrical Journal* 45, 2, 218–225.
- KHALID, M., KHANA, H. AND SHABBIR, J. 2023. Exponentially quantile regression-ratio-type estimators for robust mean estimation. *Communications in Statistics-Theory and Methods*, 1–18.
- KOC, T. AND KOC, H. 2023. A new class of quantile regression ratio-type estimators for finite population mean in stratified random sampling. *Axioms* 12, 713.
- KOÇ, H. 2021. Ratio-type estimators for improving mean estimation using Poisson regression method. *Communications in Statistics-Theory and Methods* 50, 20, 4685–4691.
- PRASAD, B. 1989. Some improved ratio type estimators of population mean and ratio in finite population sample surveys. *Communications in Statistics-Theory and Methods* 18, 1, 379–392.
- SHABBIR, J. AND GUPTA, S. 2006. A new estimator of population mean in stratified sampling. *Communications in Statistics-Theory and Methods* 35, 7, 1201–1209.
- SHAHZAD, U., HANIF, M., SAJJAD, I. AND ANAS, M. M. 2022. Quantile regression-ratio-type estimators for mean estimation under complete and partial auxiliary information. *Scientia Iranica* 29, 3, 1705–1715.
- SHAHZAD, U., AHMAD, I., AL-NOOR, N. H., HANIF, M. AND ALMANJAHIE, I. M. 2023. Robust estimation of the population mean using quantile regression under systematic sampling. *Mathematical Population Studies* 30, 3, 195–207.
- SHAHZAD, U., AL-NOOR, N. H., AFSHAN, N., ALILAH, D. A., HANIF, M. AND ANAS, M. M. 2021. Minimum covariance determinant-based quantile robust regression-type estimators for mean parameter. *Mathematical Problems in Engineering*, 5255839.
- SINGH, R., KUMAR, M., SINGH, R. D. AND CHAUDHRY, M. K. 2013. Exponential ratio type estimators in stratified random sampling. *arXiv preprint arXiv:1301.5086*.
- SUBZAR, M., BOUZA, C. N. AND AL-OMARI, A. I. 2019. Utilization of different robust regression techniques for estimation of finite population mean in SRSWOR in case of presence of outliers through ratio method of estimation. *Investigación Operacional* 40, 5, 600–609.

-
- ZAMAN, T. AND BULUT, H. 2019. Modified ratio estimators using robust regression methods. *Communications in Statistics-Theory and Methods* 48, 8, 2039–2048.
- ZAMAN, T. AND BULUT, H. 2020. Modified regression estimators using robust regression methods and covariance matrices in stratified random sampling. *Communications in Statistics-Theory and Methods* 49, 14, 3407–3420.
- ZAMAN, T., SHAZAD, U. AND YADAV, K. V. 2024. An efficient Hartley-Ross type estimators of nonsensitive and sensitive variables using robust regression methods in sample surveys. *Journal of Computational and Applied Mathematics* 440, Art. no. 115645.
- ZAMAN, T., BULUT, H. AND YADAV, K. V. 2022. Robust ratio-type estimators for finite population mean in simple random sampling: A simulation study. *Concurrency and Computation: Practice and Experience* 34, Art. no. e7273.

Afzal Subhani

Department of Statistics,

Government College University Faisalabad, Pakistan.

E-mail:afzaalsubhani45@yahoo.com

Muhammad Irfan

Department of Statistics,

Government College University Faisalabad, Pakistan.

E-mail: mirfan@gcuf.edu.pk (Corresponding author)

Maria Javed

Department of Statistics,

Government College University Faisalabad, Pakistan.

E-mail: mariajaved@gcuf.edu.pk