Improving the efficiency of exponential ratio-type estimator for population median: A calibration weight adjustment approach

Mathew J. Iseh [1,2] and Kufre J. Bassey [3]

This paper modifies the Bahl and Tuteja *exponential ratio-type estimator for population median under simple random and stratified sampling schemes using calibration weight adjustment technique with supplementary information to vary the stratum weights. The bias and mean square error of the modified estimator were obtained up to the second-order approximation, which satisfies the necessary conditions for efficiency. The findings show that the new estimator surpasses existing estimators in efficiency gain. This suggests the appropriateness of calibration weight modification in boosting the efficiency of a population parameter estimator under stratified random sampling especially where the population parameter of the auxiliary variable is known and correlate with the variable of interest.*

Keywords: Bias, calibration estimation, mean square error, stratified sampling.
JEL Classification: C82, C83
DOI: 10.33429/Cjas.14123.1/5

1. Introduction

The median has been singled out as a good indicator of location when the researched variables come from a highly skewed distribution, obtainable in surveys that estimate income, expenditure, and scores. The computation of the median, on the other hand, is usually more difficult to deal with because it involves sorted data. As a result, it is commonly assumed that since the population median is known, like the population mean, it can be used as auxiliary information in estimator development. Several researchers in recent time have implemented a number of weights and weight adjustments using direct method of weighting to improve the performance of median estimators. Weighting and weight adjustments are gaining a lot of attention as a means of enhancing the precision and efficiency of estimators in survey sampling. Nonetheless, theoretical research has demonstrated that concentrating more on reducing the influence of sample selection bias rather than increasing statistical

[1]Department of Statistics, Akwa Ibom State University, Nigeria.
[2]Corresponding author: mathewiseh@aksu.edu.ng
[3]Statistics Department, Central Bank of Nigeria.

efficiency is a key source of worry in statistical research. In most cases, the results of such investigations invariably give rise to the same mean square error (MSE) on the median estimator. Among the leading contributors who have made significant strides in this direction of weight adjustments are Singh *et al.* (2003), Aladag and Cingi (2015), Singh and Solanki (2013), and Enang *et al.* (2016). There has been a gap in these studiesdue to a typical case of identical MSE despite variances between weights adjustment techniques. To circumvent a comparable loophole, Iseh (2020) employed a calibration estimate for the population median under distinct ratio estimators in which the weights of samples were altered to recreate known population totals to achieve greater efficiency.

The study conducted by Deville and Särndal (1992), serves as watershed moment in the history of calibration estimation, as it afforded researchers with complete description of the techniques involved thereof. If large amounts of statistical data are to be collected, calibration can be a useful methodological tool, and weights can be calculated using supplementary information from administrative registers, censuses, and other reliable sources, such as software developed by several national statistics organizations, as an example. Although several researchers such as Deville and Särndal (1992), and Andersson and Thorburn (2005), Kott (2006) have discussed the calibration process as a weighting mechanism. Särndal (2007) provided an exhaustive definition of calibration for finite populations and proposed that a calibration approach to estimation should entail computing weights that incorporate specified auxiliary information which are constrained by the calibration equations; and utilizing these weights to compute linearly weighted estimates of totals and other parameters.

As a contribution to the literature, this article aims to develop an efficient exponential ratio-type estimator for the population median through calibration weight adjustment rather than the direct method of weighting, in accordance with Särndal (2007), with additional supplementary information under simple random and stratified sampling.

the rest of the paper is structured as follows Section 2 is literature review. It addresses the theoretical foundation on which the development of the new estimator is based. Section 3 presents the methodological process, which includes an asymptotic behaviour of the bias and MSE under specified optimality criteria. Section 4 presents the empirical demonstration using existing data, and Section 5 presents the conclusion of the findings.

2. Literature Review

2.1 Theoretical Literature

Several estimators have been developed to estimate the median of a population in the context of a simple random sample without replacement, but which estimator to use is conditional on details about the population, the sample, and the assumptions being made. While the sample median estimator is a straightforward and objective method, it may be inadequate for studies with insufficient data. The recurring challenges in all the existing estimators for estimating population median is that they are based on the usual conventional measures of an auxiliary variable. This section reviews the usual and well-known estimators for estimating population median under simple random sample without replacement (SRSWOR).

It can be challenging to estimate the median of a finite population using a simple random sampling method due to issues like small sample sizes, skewed distributions, concerns about efficiency, nonresponse, missing data, and the varying nature of the population median. This problem was first discussed by Gross (1980). Years after, Kuk and Mark (1989) introduced the idea of auxiliary information in the estimation of population median using the classical ratio estimator. Bahl and Tuteja (1991) proposed the ratio-type median estimator which is shown to be the most robust estimator (Iseh, 2020). Having acknowledged the usefulness of the median estimator in skewed data, several authors have adopted different techniques under different sampling schemes either in modifying the Bahl and Tuteja (1991) estimator or formulating a new estimator in a bid to obtaining a more precise and efficient estimator(Bahl & Tuteja, 1991; Singh *et al.*, 2001; Aladag & Cingi, 2015; Solanki & Singh, 2013; Enang *et al.*, 2016; Gupta *et al.*, 2008; Biag *et al.*, 2020). To futher improve performance, Iseh (2020) calibrated on the separate product-ratio estimator to obtain a more efficient estimator better than the existing ones that adopted the conventional measures of the auxiliary variables under different schemes. However, of concern is the fact that many authors have modified the Bahl and Tuteja (1991) estimator either by incorporating weights or additional supplementary information, possibly because of the robust nature of the estimator. These attempts were seen to be insufficient to checkmate the instability of some of the existing estimators with respect to data from some skewed distributions. In the progression, this study seeks to modify the Bahl and Tuteja (1991) estimator by formulation of a new weight, and by adopting the concept of calibration to enhance stability and efficiency.

2.2 Analytical Review

There are several propositions in the literature with respect to estimators under simple random and stratified sampling within the class of the estimators considered in this study. Prominent among them are the unbiased per unit estimator, the classical ratio estimator, the ratio-type exponential estimator, the chain ratio-type estimator, the product-type estimator, exponential product-type estimator, and the alternative exponential ratio estimator. These estimators are also used for empirical demonstration in further sections.

(A1). the unbiased per unit estimator $\hat{M}_G = \hat{M}_y$ was proposed by Gross (1980) with variance

$$Var\left(\hat{M}_G\right) = \lambda M_Y^2 C_{M_Y}^2 \tag{1}$$

(A2). Kuk and Mak (1989) proposed the classical ratio estimator: $\hat{M}_C = \hat{M}_y \frac{M_X}{\hat{M}_x}$ with the bias and mean square error given as

$$B\left(\hat{M}_C\right) = M_Y C_{M_X}^2 (1-k), \text{ and}$$

$$MSE\left(\hat{M}_C\right) = M_Y^2\left[C_{M_Y}^2 + C_{M_X}^2(1-2k)\right] \tag{2}$$

(A3). Bahl and Tuteja (1991) proposed the ratio-type exponential estimator: $\hat{M}_{ER} = \hat{M}_y exp\left[\frac{M_X - \hat{M}_x}{M_X + \hat{M}_x}\right]$ with bias and mean square error given as:

$$B\left(\hat{M}_{ER}\right) = \frac{\lambda M_Y C_{M_X}^2 (3-4k)}{8}, \text{ and}$$

$$MSE\left(\hat{M}_{ER}\right) = \lambda M_Y^2\left[C_{M_Y}^2 + \frac{C_{M_X}^2}{4}(1-4k)\right] \tag{3}$$

(A4). The chain ratio-type estimator: $\hat{M}_{CR} = \hat{M}_y\left(\frac{M_X}{\hat{M}_x}\right)^2$ was proposed by Kadilar and Cingi (2004), with bias and mean square error obtained as

$$B\left(\hat{M}_{CR}\right) = \lambda M_Y C_{M_X}^2(1+2k), \text{ and } MSE\left(\hat{M}_{CR}\right) = \lambda M_Y^2\left[C_{M_Y}^2 + 4C_{M_X}^2(1+k)\right] \tag{4}$$

(A5). the product-type estimator $\hat{M}_P = \hat{M}_y \left(\frac{\hat{M}_x}{M_X}\right)$ (Robson, 1957) and Murthy, 1964) was proposed with a derived bias and mean square error given as

$B\left(\hat{M}_P\right) = \lambda M_Y C_{M_X}^2 k$, and

$$MSE\left(\hat{M}_P\right) = \lambda M_Y^2 \left[C_{M_Y}^2 + C_{M_X}^2 (1+2k)\right] \quad (5)$$

(A6). the exponential product-type estimator $\hat{M}_{EP}$=$\hat{M}_y exp\left[\frac{M_X-\hat{M}_x}{M_X+\hat{M}_x}\right]$ was recommended by Bahl and Tuteja (1991) with a bias and the mean square error of

$B\left(\hat{M}_{EP}\right) = \frac{\lambda M_Y C_{M_X}^2 (4k-1)}{8}$, and

$$MSE\left(\hat{M}_{EP}\right) = \lambda M_Y^2 \left[C_{M_Y}^2 + \frac{C_{M_X}^2}{4}(1+4k)\right] \quad (6)$$

(A7). While Enang *et al.* (2016) developed an alternative exponential ratio estimator

$$\hat{M}_{AE} = \alpha_1 \left[\hat{M}_y exp\left[\frac{M_X - \hat{M}_x}{M_X + \hat{M}_x}\right]\right] + \alpha_2 \left[\hat{M}_y exp\left[\frac{M_X - \hat{M}_x}{M_X + \hat{M}_x}\right]\right]$$

with bias and mean square error given as

$B\left(\hat{M}_{AE}\right) = \lambda M_Y C_{M_X}^2 (4k - 8k^2 + 1)$, and

$$\left(\hat{M}_{AE}\right) = \lambda M_Y^2 C_{M_Y}^2 (1 - \rho_{M_Y M_X}{}^2) \quad (7)$$

Equations (1) - (7) were all recommended under the simple random scheme. For the median estimators, biases and MSEs under Stratified Random Sampling, the literature showed that:

(B1). the per unit unbiased estimator under stratified sampling followed Gross (1980), and is given as $\hat{M}_{Gst} = \hat{M}_{yh}$ with variance

$$var\left(\hat{M}_{Gst}\right) = \sum_h^H W_h^2 M_{Yh}^2 \lambda_h C_{M_{Yh}}^2 \quad (8)$$

(B2). Kuk and Mak (1989) proposed the classical ratio estimator

$$\hat{M}_{Rst} = \sum_h^H W_h \hat{M}_{yh} \left(\frac{M_{Xh}}{\hat{M}_{xh}}\right)$$

with bias and mean square error (MSE)

$B\left(\hat{M}_{Rst}\right) = \sum_h^H W_h \lambda_h M_{Yh} C_{M_{Xh}}^2 (1-k_h)$, and $MSE\left(\hat{M}_{Rst}\right) =$

$$\sum_h^H W_h^2 \lambda_h M_{Yh}^2 \left[C_{M_{Yh}}^2 + C_{M_{Xh}}^2 (1-2k_h)\right] \quad (9)$$

(B3). the exponential ratio-type estimator $\hat{M}_{ERst} = \sum_h^H W_h \hat{M}_{yh} exp\left[\frac{M_{Xh}-\hat{M}_{xh}}{M_{Xh}+\hat{M}_{xh}}\right]$ followed Bahl and Tuteja (1991), with bias and MSE

$B\left(\hat{M}_{ERst}\right) = \frac{\sum_h^H W_h \lambda_h M_{Yh} C^2_{M_{Xh}}(3-4k_h)}{8}$, and

$$MSE\left(\hat{M}_{ERst}\right) = \sum_h^H W_h^2 \lambda_h M^2_{Yh}\left[C^2_{M_{Yh}} + \frac{C^2_{M_{Xh}}}{4}(1-4k_h)\right] \quad (10)$$

(B4). a Chain Ratio-type estimator $\hat{M}_{CRst} = \sum_h^H W_h \hat{M}_{yh}\left(\frac{M_{Xh}}{\hat{M}_{xh}}\right)^2$ by Kadilar and Cingi (2004), has a bias and MSE given as

$B\left(\hat{M}_{CRst}\right) = \sum_h^H W_h \lambda_h M_{Yh} C^2_{M_X}(1+2k_h)$, and $MSE\left(\hat{M}_{CRst}\right) =$

$$\sum_h^H W_h^2 \lambda_h M^2_{Yh}\left[C^2_{M_{Yh}} + 4C^2_{M_{Xh}}(1+k_h)\right] \quad (11)$$

(B5). Robson (1957) and Murthy (1964) product-type estimator $\hat{M}_{Pst} = \sum_h^H W_h \hat{M}_{yh}\left(\frac{\hat{M}_{xh}}{M_{Xh}}\right)$ also has bias and MSE under stratified sampling scheme as

$B\left(\hat{M}_{Pst}\right) = \sum_h^H W_h \lambda_h M_{Yh} C^2_{M_{Xh}} k_h$, and

$$MSE\left(\hat{M}_{Pst}\right) = \sum_h^H W_h^2 \lambda_h M^2_{Yh}\left[C^2_{M_{Yh}} + C^2_{M_{Xh}}(1+2k_h)\right] \quad (12)$$

(B6). Bahl and Tuteja (1991) exponential product-type estimator $\hat{M}_{EPst} = \sum_h^H W_h \hat{M}_{yh} exp\left[\frac{M_{Xh}-\hat{M}_{xh}}{M_{Xh}+\hat{M}_{xh}}\right]$ under stratified sampling was developed with bias and MSE obtained as

$B\left(\hat{M}_{EPst}\right) = \frac{\sum_h^H W_h \lambda_h M_{Yh} C^2_{M_{Xh}}(4k_h-1)}{8}$, and

$$MSE\left(\hat{M}_{EPst}\right) = \sum_h^H W_h^2 \lambda_h M^2_{Yh}\left[C^2_{M_{Yh}} + \frac{C^2_{M_{Xh}}}{4}(1+4k_h)\right] \quad (13)$$

(B7). While Enang *et al.* (2016) proposed alternative exponential ratio estimator under stratified sampling

$$\hat{M}_{AEst} = \sum_h^H W_h \left\{ \alpha_{1h}\left[\hat{M}_{yh} exp\left[\frac{M_{Xh}-\hat{M}_{xh}}{M_{Xh}+\hat{M}_{xh}}\right]\right] + \alpha_{2h}\left[\hat{M}_{yh} exp\left[\frac{M_{Xh}-\hat{M}_{xh}}{M_{Xh}+\hat{M}_{xh}}\right]\right]\right\}$$

has a bias and MSE as

$B\left(\hat{M}_{AEst}\right) = \sum_h^H W_h \lambda_h M_{Yh} C^2_{M_{Xh}}(4k_h - 8k_h^2 + 1)$, and

$$MSE\left(\hat{M}_{AEst}\right)=\sum_{h}^{H}W_{h}^{2}\lambda_{h}M_{Yh}^{2}C_{M_{Yh}}^{2}(1-\rho_{M_{Yh}M_{Xh}}{}^{2}) \tag{14}$$

A limitation in one estimator led to the proposition of another. Howbeit, the need to consider an exponential ratio-type estimator for the population median through calibration weight adjustment rather than using the direct method of weighting as observed in the existing estimators, in accordance with Särndal (2007) recommendation motivated this study.

3. Data and Methodology

3.1 Design

In designing a study of this nature, the first thing is to consider the conventional notations for a finite population $U=\{u_1,u_2,\ldots,u_N\}$ of size N. Let Y and X represent the study and auxiliary variables, respectively. Assume that the observations y_i on Y and x_i on X are obtained for each sampling units and that the random sample is drawn from the population by SRSWOR. Let $f_Y(M_Y)$, and $f_X(M_X)$ represent the density functions of Y and X with sample and population median $\hat{M}_y$, $\hat{M}_x$, M_Y and M_X, respectively, and the correlation coefficient taken as $\rho_{M_YM_X}=4(P_{11}-0.25)$, where $P_{11}=P(Y\le M_Y\cap X\le M_X)$ (Gross, 1980). It follows that when the research population is heterogeneous, the goal of creating homogeneity (of comparable qualities) by stratification comes into play, with the aim of reducing the estimation error for a set survey cost.

Suppose that the auxiliary variable x_{hi} and the attribute of interest y_{hi} are non-negative for every unit in the population for the i^{th} element in the h^{th} stratum. Under stratification, assume that the population of size N is stratified into H strata with h^{th} stratum comprising N_h units, where $h=1,2,\ldots,H$, such that $\sum_{h=1}^{H}N_h=N$ and stratum weight given as $W_h=\frac{N_h}{N}$, then a simple random sample of size n_h can be taken without replacement from the h^{th} stratum, such that $\sum_{h=1}^{H}n_h=n$.

Let M_{Yh} and M_{Xh} be the population median for the study and auxiliary variables in the h^{th} stratum, respectively, with $\hat{M}_{yh}$ and $\hat{M}_{xh}$ being their respective sample medians, and $y_{h(1),}y_{h(2),}\cdots,y_{h(n),}$ the y_{hi} values of the sample units in ascending order. Let r be an integer satisfying $Y_{hr}\le M_{Yh}\le M_{h(r+1)}$ such that $P_h=\frac{r}{n}$ is the proportion of y_{hi} values in the sample that are less than or equal to the median value M_{Yh} (which denotes the unknown population parameter). If $\varphi_y(r)$ is defined as the $r-quantile$ of Y_h and $\hat{M}_{yh} = \varphi_y(0.5)$, then the correlation coefficient in the h^{th} stratum between the sampling distributions of M_{Yh} and M_{Xh} will be given as $\rho_{M_{Yh}M_{Xh}} = 4(P_{11h} - 0.25)$, where P_{11h} is the ratio of units in the population in the h^{th} stratum with $Y_h \leq M_{Yh}$ and $X_h \leq M_{Xh}$. A matrix of the proportion P_{ij} by Kuk and Mak (1989) is given in Table 1.

Table 1: Matrix of Proportions for Stratified Sampling by Kuk and Mak (1989)

	$X_h \leq M_{Xh}$	$X_h > M_{Xh}$	$Total$
$Y_h \leq M_{Yh}$	P_{11}	P_{12}	P_1
$Y_h > M_{Yh}$	P_{21}	P_{22}	P_2
$Total$	P_1	P_2	**1**

For large sample approximations, the derived parameters under simple and stratified random sampling scheme are as summarized in Table 2.

Table 2: Simple and Stratified Random Sampling Parameters for Large Sample Approximations

Simple	**Stratified**
$\hat{M}_y = M_Y(1+e_0)$	$\hat{M}_{yh} = M_{Yh}(1+e_{0h})$
$\hat{M}_x = M_X(1+e_1)$	$\hat{M}_{xh} = M_{Xh}(1+e_{1h})$
$e_0 = \frac{\hat{M}_y - M_Y}{M_Y}$	$e_{0h} = \frac{\hat{M}_{yh} - M_{Yh}}{M_{Yh}}$
$e_1 = \frac{\hat{M}_x - M_X}{M_X}$	$e_{1h} = \frac{\hat{M}_{xh} - M_{Xh}}{M_{Xh}}$
$E(e_0) = E(e_1) = 0$	$E(e_{0h}) = E(e_{1h}) = 0$
$E\left(e_0^2\right) = \lambda C_{M_Y}^2$	$E\left(e_{0h}^2\right) = \lambda_h C_{M_{Yh}}^2$
$E\left(e_1^2\right) = \lambda C_{M_X}^2$	$E\left(e_{1h}^2\right) = \lambda_h C_{M_{Xh}}^2$
$\lambda = \frac{1-f}{4n}$	$\lambda_h = \frac{1-f_h}{4n_h}$
$E(e_0 e_1) = C_{M_Y} C_{M_X M_Y M_X}$	$E(e_{0h} e_{1h}) = \lambda_h C_{M_{Yh}} C_{M_{Xh}} \rho_{M_{YhM_{Xh}}}$
$\lambda C_{M_Y} = \{M_Y f_Y(M_Y)\}^{-1}$	$C_{M_{Yh}} = \{M_{Yh} f_Y(M_{Yh})\}^{-1}$
$**C_{M_X} = \{M_X f_X(M_X)\}^{-1}$	$C_{M_{Xh}} = \{M_{Xh} f_X(M_{Xh})\}^{-1}$
$k = \frac{C_{M_Y} \rho_{M_Y M_X}}{C_{M_X}}$	$k_h = \frac{C_{M_{Yh}} \rho_{M_{Yh} M_{Xh}}}{C_{M_{Xh}}}$

** and ** are the coefficients of variation of the population median of the study variables*

With the objective of increasing the efficiency of median estimators, existing estimators for predicting the population median in simple random sampling up to the first order approximation were further evaluated and relevant comparison made, with consideration of their biases and mean square errors (MSEs).

3.2 Data

In deriving new estimators, it is important, in the general sense, to validate the proposed estimators empirically and for statistical comparability using existing data of comparable studies for credence. In this study, we consider data from four distinct populations under simple random sampling and two different populations under stratified sampling to drive our findings. The strength of our estimators is derived from the Bias and MSE values compared to those of previously published works to evaluate the improvement in the new estimators as shown in further section of this paper. In particular, data from selected previous studies were used to compute the percent relative efficiency of the derived estimators as $PRE = \frac{MSE(\hat{M}_C)}{MSE(.)} \times 100$, where $MSE(\hat{M}_C)$ is the MSE of classical median estimator under simple random and stratified sampling which is the same as $var\left(\hat{M}_G\right)$ and $var\left(\hat{M}_{Gst}\right)$, respectively, and $MSE(.)$ denotes the MSE of estimators considered in this study.

3.3 Improved Exponential Ratio-type Estimator: A Proposition

The aim of this paper is to propose an improved exponential ratio-type estimator for population median through calibration weight adjustment. To achieve this objective, we first consider the expedition based on a comparative analysis of the biases and MSEs of existing median estimators. Thereafter, we focused on the modification of Bahl and Tuteja (1991) ”exponential type-ratio estimator,” which is generally considered to be the most robust estimator (see Iseh, 2020), by introducing a weight factor and applying a weight adjustment process to vary the stratum weights under stratified sampling scheme using calibration weight adjustment technique. These modifications are intended to improve the estimator’s performance under simple and stratified random sampling, respectively. Survey statisticians use asymptotic to generate formal statements about estimators that are centered on arbitrarily large samples that are as close as possible to actual estimators based on large but finite samples[4]. Thus, this paper has derived and obtained the asymptotic equations for the bias and MSE of the modified estimator up to the first order approximation with necessary conditions for efficiency.

[4] Note that, developing an asymptotic structure for a survey population and the sample taken for that population is a crucial idealization because sample surveys, by their very nature, are limited, even when the sample size is large (Kott, 2015).

3.4 Modification of the Estimator of Interest

3.4.1 Modification under Simple Random Sampling scheme

It is a common understanding in survey statistics that the model for the sample data in simple random sampling is the same as the model for the population before sampling. However, in practice, the complex sampling designs often used can make the two models very different and failure to account for the sample selection process might bias the inference (Pfeffermann, 1993). In what follows, the idea of incorporating weights into the analysis is one way of dealing with the design effects.

The proposed modification of Bahl and Tuteja (1991) in Eq. (3) is presented as

$$\hat{t}_m^* = \hat{M}_y exp\left[\theta\frac{(M_X-\hat{M}_x)}{(M_X+\hat{M}_x)}\right] \tag{15}$$

with an introduction of an initial weight[5] $\theta = \alpha(\alpha-1)^{-1}$ (which will be determined while minimizing the mean square error and $\alpha \neq 1$). then by large number approximation, Eq.(15) becomes:

$$\hat{t}_m^* = M_Y(1+e_0)exp\left\{\frac{-\theta}{2}e_1\left[1-\frac{e_1}{2}+\frac{e_1^2}{4}\right]\right\}$$

$$\hat{t}_m^* = M_Y\left[1+e_0-\frac{-\theta}{2}e_1-\theta e_0e_1+\theta\frac{(2+\theta)}{8}e_1^2\right]$$

$$\hat{t}_m^* - M_Y = M_Y\left[e_0-\tfrac{-\theta}{2}e_1-\tfrac{\theta}{2}e_0e_1+\theta\tfrac{(2+\theta)}{8}e_1^2\right] \tag{16}$$

$$Bias(\hat{t}_m^*) = M_Y\left[\theta\tfrac{(2+\theta)}{8}\lambda C_{M_X}^2-\tfrac{\theta}{2}\lambda C_{M_Y}C_{M_X}\rho_{M_YM_X}\right] \tag{17}$$

The MSE of $\hat{t}_m^*$ to the first order of approximation is obtained by squaring both sides of Eq.(16) and retaining terms to the second-degree while taking expectations as;

$$MSE(\hat{t}_m^*) = M_Y^2\left[\lambda C_{M_Y}^2+\lambda\tfrac{\theta^2}{4}C_{M_X}^2-\lambda\theta C_{M_Y}C_{M_X}\rho_{M_YM_X}\right] \tag{18}$$

[5] Note: the use of this Weight is to test and protect against nonignorable sampling designs which could cause selection bias and against misspecification of the model holding in the population. The approach is intended to yield design consistent estimators for corresponding descriptive population quantities of the model parameters.

Substituting for θ and minimizing Eq. (17) with respect to α gives

$$a = \frac{2C_{M_Y}C_{M_X}\rho_{M_YM_X}}{2C_{M_Y}C_{M_X}\rho_{M_YM_X} - C^2_{M_X}}$$

Substituting the value of α in Eq.(18) gives

$$MSE_{opt}(\hat{t}^*_m) = \lambda M^2_Y C^2_{M_Y}(1-\rho_{M_YM_X}{}^2) \tag{19}$$

Eq.(19) can also be written as

$$MSE_{opt}(\hat{t}^*_m) = \lambda\{f_Y(M_Y)\}^{-2}(1-\rho_{M_YM_X}{}^2) \tag{20}$$

From Eq. (17) the optimum bias is given as

$$Bias_{opt}(\hat{t}^*_m) = \lambda M_Y\left[\tfrac{1}{2}C_{M_Y}C_{M_X}\rho_{M_YM_X} - C^2_{M_Y}\rho_{M_YM_X}{}^2\right] \tag{21}$$

3.5 Efficiency Comparison under Simple Random Sampling

The efficiency of the modified estimator in Eq. (15) is determined by comparing its MSE at its minimum (Eq.19) with that of others (A1 - A7) under the following necessary (optimality) conditions:

	Parameter	**Necessary Conditions for efficiency**
1.	$\hat{t}^*_m$ with $\hat{M}_G$	$MSE_{opt}(\hat{t}^*_m) < var(\hat{M}_G)$ if: $\lambda M^2_Y C^2_{M_Y}(1-\rho_{M_YM_X}{}^2) < \lambda M^2_Y C^2_{M_Y}$ $\Rightarrow\left(1-\rho_{M_YM_X}{}^2\right) < 1$ which is always true.
2.	$\hat{t}^*_m$ with $\hat{M}_R$	$MSE_{opt}(\hat{t}^*_m) < MSE(\hat{M}_C)$ if: $\lambda M^2_Y C^2_{M_Y}(1-\rho_{M_YM_X}{}^2) < \lambda M^2_Y\left[C^2_{M_Y} + C^2_{M_X}(1-2k)\right]$ $\Rightarrow(C_{MX} - C_{MY}\rho_{M_YM_X})^2 > 0$ or $(1-k)^2 > 0$ will always be satisfied.
3.	$\hat{t}^*_m$ with $\hat{M}_{ER}$	$MSE_{opt}(\hat{t}^*_m) < MSE(\hat{M}_{ER})$ if: $\lambda M^2_Y C^2_{M_Y}(1-\rho_{M_YM_X}{}^2) < \lambda M^2_Y\left[C^2_{M_Y} + \frac{C^2_{M_X}}{4}(1-4k)\right]$ $\Rightarrow(C_{M_X} - 2C_{M_Y}\rho_{M_YM_X})^2 > 0$ or $(1-2k)^2 > 0$ or $(2C_{M_Y}\rho_{M_YM_X} - C_{M_X})^2 > 0$ or $(2k-1)^2 > 0$ will always be satisfied.
4.	$\hat{t}^*_m$ with $\hat{M}_{CR}$	$MSE_{opt}(\hat{t}^*_m) < MSE(\hat{M}_{CR})$ if: $\lambda M^2_Y C^2_{M_Y}(1-\rho_{M_YM_X}{}^2) < \lambda M^2_Y\left[C^2_{M_Y} + 4C^2_{M_X}(1+k)\right]$ $\Rightarrow (2C_{M_X} + C_{M_Y}\rho_{M_YM_X})^2 > 0$ or $(k+2)^2 > 0$ will always be satisfied.

	Parameter	Necessary Conditions for efficiency
5.	$\hat{t}_m^*$ with $\hat{M}_P$	$MSE_{opt}\left(\hat{t}_m^*\right) < MSE\left(\hat{M}_P\right)$ if: $\lambda M_Y^2 C_{M_Y}^2\left(1-\rho_{M_Y M_X}{}^2\right) < \lambda M_Y^2\left[C_{M_Y}^2 + C_{M_X}^2(1+2k)\right]$ $\Rightarrow\left(C_{MX} + C_{MY}\rho_{M_Y M_X}\right)^2 > 0$ or $(k+1)^2 > 0$, will always be satisfied.
6.	$\hat{t}_m^*$ with $\hat{M}_{EP}$	$MSE_{opt}\left(\hat{t}_m^*\right) < MSE\left(\hat{M}_{EP}\right)$ if: $\Rightarrow \lambda M_Y^2 C_{M_Y}^2\left(1-\rho_{M_Y M_X}{}^2\right) < \lambda M_Y^2\left[C_{M_Y}^2 + \frac{C_{M_X}^2}{4}(1+4k)\right]$ $\Longrightarrow\left(C_{MX} + 2C_{MY}\rho_{M_Y M_X}\right)^2 > 0$ or $(1+2k)^2 > 0$, will always be satisfied.

3.5.1 Modification under Stratified Random Sampling scheme

The modified form of Bahl and Tuteja (1991) exponential type-ratio estimator under stratified random sampling is as follows:

$$\hat{t}_{mh}^* = \sum_h^H W_h \hat{M}_{yh}\, exp\left[\theta_h \frac{\left(M_{Xh} - \hat{M}_{xh}\right)}{\left(M_{Xh} + \hat{M}_{xh}\right)}\right] \tag{22}$$

Let $\theta_h = \alpha_h(\alpha_h - 1)^{-1}$ where $\alpha_h \neq 1$, then by large number approximation, Eq. (22) also becomes

$$\hat{t}_{mh}^* = \sum_h^H W_h M_{Yh}(1+e_{0h})exp\left\{\frac{-\theta_h}{2}e_{1h}\left[1 - \frac{e_{1h}}{2} + \frac{e_{1h}^2}{4}\right]\right\}$$

$$= \sum_h^H W_h M_{Yh}\left[1 + e_{0h} - \frac{\theta_h}{2}e_{1h} - \theta e_{0h}e_{1h} + \theta_h\frac{(2+\theta_h)}{8}e_{1h}^2\right]$$

$$\hat{t}_{mh}^* - M_Y = \sum_h^H W_h M_{Yh}\left[e_{0h} - \frac{\theta_h}{2}e_{1h} - \frac{\theta_h}{2}e_{0h}e_{1h} + \theta_h\frac{(2+\theta_h)}{8}e_{1h}^2\right] \tag{23}$$

$$Bias(\hat{t}_{mh}^*) = \sum_h^H W_h M_{Yh}\left[\theta_h\frac{(2+\theta_h)}{8}\lambda_h C_{M_{Xh}}^2 - \frac{\theta_h}{2}\lambda_h C_{M_{Yh}}C_{M_{Xh}}\rho_{M_{Yh}M_{Xh}}\right] \tag{24}$$

The MSE of $\hat{t}_{mh}^*$ to the first order of approximation is obtained by squaring both sides of Eq. (23) and retaining terms to the second-degree while taking expectations as;

$$MSE(\hat{t}_{mh}^*) = \sum_h^H W_h^2 M_Y^2 \lambda_h\left[C_{M_{Yh}}^2 + \frac{\theta_h{}^2}{4}C_{M_{Xh}}^2 - \theta_h C_{M_{Yh}}C_{M_{Xh}}\rho_{M_{Yh}M_{Xh}}\right] \tag{25}$$

Substituting for θ_h and minimizing Eq. (25) with respect to α_h we obtained

$$\alpha_h = \frac{2C_{M_{Yh}}C_{M_{Xh}}\rho_{M_{Yh}M_{Xh}}}{2C_{M_{Yh}}C_{M_{Xh}}\rho_{M_{Yh}M_{Xh}} - C^2_{M_{Xh}}}$$

And substituting for α_h, Eq. (25) becomes

$$MSE_{opt}(\hat{t}^*_{mh}) = \sum_h^H W_h^2 M_Y^2 \lambda_h C^2_{M_Y}(1 - \rho_{M_{Yh}M_{Xh}}{}^2) \tag{26}$$

Hence, Eq. (25) can also be expressed as

$$MSE_{opt}(\hat{t}^*_{mh}) = \sum_h^H W_h^2 \{f_Y(M_{Yh})\}^{-2} \lambda_h (1 - \rho_{M_{Yh}M_{Xh}}{}^2)$$

Hence, from Eq. (14) the optimum bias is obtained as

$$Bias_{opt}(\hat{t}^*_{mh}) = \sum_h^H W_h M_{Yh} \lambda_h \left[\frac{1}{2} C_{M_{Yh}} C_{M_{Xh}} \rho_{M_{YhM_{Xh}}} - C^2_{M_{Yh}} \rho_{M_{Yh}M_{Xh}}{}^2\right] \tag{27}$$

3.6 Adjustment of Stratum Weights via Calibration Method

In recent times, calibration techniques have been adopted to enhance efficiency and further improve the precision of survey estimates of population parameters over the existing estimators. This was first demonstrated by Devile and Sarndal (1992). Several authors like Clement and Enang (2017), Iseh (2020) and Iseh and Enang (2021), among others, have established significant results in the literature with the use of calibration techniques over the existing estimators.The primary objective here is to calibrate the initial weight by adjusting it with auxiliary variables and then constructing final weights (calibration weights) that accurately estimate the totals of all auxiliary variables used in the calibration process, so that the final weights are as close to the initial weights as possible in terms of the distance function used. Consequently, as a result of calibration, the precision of the estimated values will improve. Thus, Eq. (22) can be written as

$$\hat{t}^*_{mhcal} = \sum_h^H \Omega_h \hat{M}_{yh} exp\left[\theta_h \frac{(M_{Xh} - \hat{M}_{xh})}{(M_{Xh} + \hat{M}_{xh})}\right] \tag{28}$$

where Ω_h is the chosen calibration weights such that the chi square distance measure

$$\sum_{h=1}^{H}\left(\frac{\Omega_h - W_h}{q_h W_h}\right)^2 \tag{29}$$

is minimized, subject to the constraints

$$\sum_{h=1}^{H}\Omega_h\{f_X(M_{Xh})\}^{-2} = V(M_{Xh}) \tag{30}$$

Thus, we obtained the calibration weights as

$$\Omega_h = W_h + \left[V(M_{Xh}) - \sum_{h=1}^{H} W_h\{f_X(M_{Xh})\}^{-2}\right]\frac{q_{hW_h}\{f_X(M_{Xh})\}^{-2}}{\sum_{h=1}^{H} q_{hW_h}\{f_X(M_{Xh})\}^{-4}} \tag{31}$$

and with the assumption that the tuning parameter $q_h = \{f_X(M_{Xh})\}^2$, then

$$\Omega_h = W_h\left[\frac{V(M_{Xh})}{\hat{V}(M_{Xh})}\right] \tag{32}$$

Squaring both sides of Eq. (32) we have

$$\Omega_h^2 = W_h^2\left[\frac{V(M_{Xh})}{\hat{V}(M_{Xh})}\right]^2 \tag{33}$$

where $V(M_{Xh}) = \sum_{h=1}^{H} W_h^2\lambda_h\{f_X(M_{Xh})\}^{-2}$ and $\hat{V}(M_{Xh}) = \sum_{h=1}^{H} W_h\{f_X(M_{Xh})\}^{-2}$ Thus, Eq. (27) can be written as

$$Bias_{opt}(\hat{t}^*_{mhcal}) = \sum_{h}^{H} W_h\left[\frac{V(M_{Xh})}{\hat{V}(M_{Xh})}\right] M_{Yh}\lambda_h\left[\frac{1}{2}C_{M_{Yh}}C_{M_{Xh}}\rho_{M_{Yh}M_{Xh}} - C^2_{M_{Yh}}\rho_{M_{Yh}M_{Xh}}{}^2\right] \tag{34}$$

and the mean square error in Eq. (26) obtained as

$$MSE_{opt}(\hat{t}^*_{mhcal}) = \sum_{h=1}^{H} W_h^2\left[\frac{V(M_{xh})}{\hat{V}(M_{xh})}\right]^2\{f_Y(M_{Yh})\}^{-2}\lambda_h(1-\rho_{M_{Yh}M_{Xh}}{}^2) \tag{35}$$

3.7 Efficiency Comparison under Stratified Sampling

The efficiency comparison of the modified estimator in Eq. (22) were obtained by comparing its MSE in Eq. (26) with that of the existing estimators in (B1 – B7), while that of the calibration estimator in Eq. (28) was obtained by comparing the MSE in Eq. (35) with that of Eq. (26) under the following necessary (optimality) conditions:

Case 1 : $\hat{t}^*_{mh}$ with $\hat{M}_{Gst}$

$MSE_{opt}\left(\hat{t}^*_{mh}\right) < \ var\left(\hat{M}_{Gst}\right)$ if:

$$\Longrightarrow \sum_{h}^{H} W_h^2 M_Y^2 \lambda_h C_{M_Y}^2 (1-\rho_{M_{Yh}M_{Xh}}{}^2) < \sum_{h}^{H} W_h^2 M_Y^2 \lambda_h C_{M_Y}^2$$

$\Rightarrow (1-\rho_{M_{Yh}M_{Xh}}{}^2) < 1$ which is always true if

Case 2 : $\hat{t}^*_{mh}$ with $\hat{M}_{Rst}$

$MSE_{opt}\left(\hat{t}^*_{mh}\right) < \ MSE\left(\hat{M}_{Rst}\right)$ if:

$$\sum_{h}^{H} W_h^2 M_{Yh}^2 \lambda_h C_{M_Y}^2 (1-\rho_{M_{Yh}M_{Xh}}{}^2) < \sum_{h}^{H} W_h^2 \lambda_h M_{Yh}^2 \left[C_{M_{Yh}}^2 + C_{M_{Xh}}^2 (1-2k_h)\right]$$

$\Rightarrow (C_{M_{Xh}} - C_{M_{Yh}}\rho_{M_{Yh}M_{Xh}})^2 > 0$ or $(1-k_h)^2 > 0$ will always be satisfied

Case 3 : $\hat{t}^*_{mh}$ with $\hat{M}_{ERst}$

$MSE_{opt}\left(\hat{t}^*_{mh}\right) < \ MSE\left(\hat{M}_{ERst}\right)$ if:

$$\sum_{h}^{H} W_h^2 M_{Yh}^2 \lambda_h C_{M_{Yh}}^2 (1-\rho_{M_{Yh}M_{Xh}}{}^2) < \sum_{h}^{H} W_h^2 \lambda_h M_{Yh}^2 \left[C_{M_{Yh}}^2 + \frac{C_{M_{Xh}}^2}{4}(1-4k_h)\right]$$

$\Rightarrow (C_{M_{Xh}} - 2C_{M_{Yh}}\rho_{M_{Yh}M_{Xh}})^2 > 0$ or $(1-2k_h)^2 > 0$ will always be satisfied

Case 4 : $\hat{t}^*_{mh}$ with $\hat{M}_{CRst}$

$MSE_{opt}\left(\hat{t}^*_{mh}\right) < \ MSE\left(\hat{M}_{CRst}\right)$ if:

$$\sum_{h}^{H} W_h^2 M_{Yh}^2 \lambda_h C_{M_Y}^2 (1-\rho_{M_{Yh}M_{Xh}}{}^2) < \sum_{h}^{H} W_h^2 \lambda_h M_{Yh}^2 \left[C_{M_{Yh}}^2 + 4C_{M_{Xh}}^2 (1+k_h)\right]$$

$\Rightarrow (2C_{M_{Xh}} + C_{M_{Yh}}\rho_{M_{Yh}M_{Xh}})^2 > 0$ or $(k_h+2)^2 > 0$ will always be satisfied

Case 5 : $\hat{t}^*_{mh}$ with $\hat{M}_{Pst}$

$MSE_{opt}\left(\hat{t}^{*}_{mh}\right) < MSE\left(\hat{M}_{Pst}\right)$ if:

$$\sum_{h}^{H} W_h^2 M_{Yh}^2 \lambda_h C_{M_{Yh}}^2 (1-\rho_{M_{Yh}M_{Xh}}{}^2) < \sum_{h}^{H} W_h^2 \lambda_h M_{Yh}^2 \left[C_{M_{Yh}}^2 + C_{M_{Xh}}^2 (1+2k_h)\right]$$

$\Rightarrow\left(C_{MXh} + C_{MYh} ?_{M_{Yh}M_{Xh}}\right)^2 > 0$ or $(k_h+1)^2 > 0$ will always be satisfied
Case 6 : $\hat{t}^{*}_{mh}$ with $\hat{M}_{EPst}$
$MSE_{opt}\left(\hat{t}^{*}_{mh}\right) < MSE\left(\hat{M}_{EPst}\right)$ if:

$$\sum_{h}^{H} W_h^2 M_{Yh}^2 \lambda_h C_{M_Y}^2 (1-\rho_{M_{Yh}M_{Xh}}{}^2) < \sum_{h}^{H} W_h^2 \lambda_h M_{Yh}^2 \left[C_{M_{Yh}}^2 + \frac{C_{M_{Xh}}^2}{4}(1+4k_h)\right]$$

$\Rightarrow\left(C_{MXh} + 2C_{MYh} ?_{M_{Yh}M_{Xh}}\right)^2 > 0$ or $(1+2k_h)^2 > 0$ will always be satisfied**.**
Case 7: $MSE_{opt}\left(\hat{t}^{*}_{mhcal}\right) < MSE_{opt}\left(\hat{t}^{*}_{mh}\right)$ if:

$$\sum_{h}^{H} W_h^2 \left[\frac{V(M_{Xh})}{\hat{V}(M_{Xh})}\right]^2 M_{Yh}^2 \lambda_h C_{M_Y}^2 (1-\rho_{M_{Yh}M_{Xh}}{}^2) < \sum_{h}^{H} W_h^2 M_{Yh}^2 \lambda_h C_{M_Y}^2 (1-\rho_{M_{Yh}M_{Xh}}{}^2)$$

$\Rightarrow\left[\frac{V(M_{Xh})}{\hat{V}(M_{Xh})}\right]^2 < 1$ will always be satisfied if $V(M_{Xh}) < \hat{V}(M_{Xh})$

4. Results and Discussion

The derivation of improved estimator under the two-sampling scheme of simple and stratified random sampling is further validated with empirical demonstration in this section. The data used for the empirical investigations and the results obtained are as presented in the Tables 3 - 8. In addition, the results are discussed with respect to the performance of both existing and proposed estimators.

4.1 Descriptive Statistics

Tables 3-5 show the data description from existing studies that will be used in the statistical analyses to evaluate the characteristics of the population parameters and validate the theoretical propositions.

Table 3: Empirical Evaluation of Some Population Statistics: Populations I - IV

Statistics	Population I	Population II	Population III	Population IV
N	69	340	396	67
n	17	150	65	23
M_Y	2068	178	30	4.8
M_X	2307	3513	14.6	7
$\rho_{M_Y M_X}$	0.3166	0.92	0.84	0.6624
$f_Y(M_Y)$	0.00014	0.00018	0.01178	0.0763
$f_X(M_X)$	0.00013	0.00008	0.02194	0.0526
C_{M_Y}	3.45399	31.17635	2.82869	2.73045
C_{M_X}	3.33433	3.41315	3.12184	2.71592
λ	0.01108	0.00093	0.00321	0.00714
k	0.32796	8.40345	0.77611	0.08493

Note: Population I- Singh (2003), Population II- Al and Gingi (2009); Population III- Chen *et.al.* (2004); Population IV- Aczel and Sounderpandian (2004).

Population I: Let y, and x respectively be the number of fish caught by the marine recreational fisherman in years 1995 and 1994 in USA.

Population II: Let y be the number of teachers and x be the number of students in elementaryschools for 340 medium-developed districts in Turkey in 2007.

Population III: Let y be the entire height of conifer trees in feet and x be the diameter of conifertrees in centimeters at breast height.

Population IV: Let y be the U.S. exports to Singapore in billions of Singapore dollars, and x, the money supply figures in billions of Singapore dollars.

Population V: Computations on the development index of all districts in Turkey about educational opportunities using the data gathered from schools by Ministry of National Education for 2006-2007 educational year. The development groups are obtained by clustering the districts with the same development level in the same group.

Table 4: Empirical Evaluation of Some Population Statistics: Population V

Statistics	Stratum I	Stratum II	Stratum III	Stratum IV	Stratum V	Stratum VI
N	91	129	204	145	184	170
n	18	26	41	29	29	34
M_Y	81	93	24	54	44	101
M_X	1265	1139	614	763	533	911
$\rho_{M_Y M_X}$	0.84	0.96	0.84	0.88	0.88	0.96
$f_Y(M_Y)$	0.00316	0.00318	0.01151	0.0003	0.00512	0.00002
$f_X(M_X)$	0.00019	0.00024	0.00049	0.00442	0.00052	0.00009
C_{M_Y}	3.90686	3.38135	3.62004	61.93485	4.43892	37.8874
C_{M_X}	4.1606	3.65818	3.48005	0.29652	3.58733	12.61718
λ	0.01114	0.00768	0.00487	0.0069	0.00726	0.00588
k	0.78877	0.88735	0.87379	183.80833	1.08890	2.88273

Data for number of Teachers is considered as study variable in elementary schools for 923 districts in Turkey in 2007 and the number of students as auxiliary variable at six regions (1. Mediterranean, 2. Aegean, 3. East and Southeast Anatolia, 4. Central Anatolia, 5. Black Sea, 6. Marmara). Proportional allocation was used in determining the sample size of each stratum.(Adalag & Cingi, 2015)

Table 5: Statistics for Population VI.

Statistics	Stratum I	Stratum II	Stratum III	Stratum IV
N	36	36	36	36
n	5	5	5	5
M_Y	38	3058	2033	2382
M_X	1480	127289	54559	71615
$\rho_{M_Y M_X}$	0.7776	0.8888	0.8888	0.8888
$f_Y(M_Y)$	0.00706	0.00032	0.00042	0.00030
$f_X(M_X)$	0.00016	0.00001	0.00001	0.00001
C_{M_Y}	3.72956	1.02118	1.16588	1.39381
C_{M_X}	4.11746	1.00372	1.29991	1.36097
λ	0.04306	0.04306	0.04306	0.04306
k	0.70434	0.90425	0.79715	0.91025

Population VI: Let y be the number of teaching staff and x be the number of students in 4 different types of schools under 36 districts in Punjab province of Pakistan (Panjab Development Statistics(PDS), 2012 Pages 114-116).

4.2. Estimation Results

The results of this study are as presented in Tables 6-8.

Table 6: Bias under Simple Random Sampling.

Estimator	Population I	Population II	Population III	Population IV
$\hat{M}_G$	0	0	0	0
$\hat{M}_C$	171.241	**-14.298**	0.211	0.231
$\hat{M}_{ER}$	53.77	**-7.391**	**-0.012**	0.084
$\hat{M}_{CR}$	421.942	**34.391**	2.399	0.296
$\hat{M}_P$	83.567	**16.23**	0.73	0.022
$\hat{M}_{EP}$	**9.932**	**7.874**	0.247	-0.021
$\hat{M}_{AE}$	369.823	-1024.237	-0.671	0.324
$\widehat{t}_m^*$	**14.377**	**-128.271**	**-0.201**	**0.009**

Table 7: MSE and PRE under Simple Random Sampling.

Estimat	Population I		Population II		Population III		Population IV	
	MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE
$\hat{M}_G$	565443.6	100	28682.29	100	23.151	100	1.226	100
$\hat{M}_C$	746753.2	75.72	23248.48	123.37	7.579	305.45	2.233	54.90
$\hat{M}_{ER}$	524362.2	107.83	25879.35	110.83	8.316	278.40	1.426	85.96
$\hat{M}_{CR}$	3364479	16.81	41612.98	68.93	223.485	10.36	6.491	18.89
$\hat{M}_P$	1438018	39.32	34803.87	82.41	95.12	24.34	2.645	46.35
$\hat{M}_{EP}$	869994.6	64.99	31657.14	90.60	52.086	44.45	1.633	75.11
$\hat{M}_{AE}$	**508766**	**111.14**	**4405.6**	**651.04**	**6.816**	**339.67**	**0.688**	**178.19**
$\widehat{t}_m^*$	**508766**	**111.14**	**4405.6**	**651.04**	**6.816**	**339.67**	**0.688**	**178.19**

Table 8: Results for Stratified Random Sampling

	Population V			Population VI		
Estimator	Bias	MSE	PRE	Bias	MSE	PRE
$\hat{M}_{Gst}$	0	5160.735	100	0	71075.12	100
$\hat{M}_{Rst}$	**-33.295**	3273.761	157.64	**16.989**	16177.7	439.34
$\hat{M}_{ERst}$	**-19.327**	3978.364	129.72	**-7.076**	25503.34	278.69
$\hat{M}_{CRst}$	**130.89**	10220.9	50.49	339.715	615823.6	11.54
$\hat{M}_{Pst}$	**54.728**	7179.982	71.88	107.575	270957.1	26.23
$\hat{M}_{EPst}$	**24.685**	5931.475	87.01	**38.217**	152893	46.49
$\hat{M}_{AEst}$	-2333.837	689.193	748.81	-192.343	14938.25	475.79
$\widehat{t}_{mh}^*$	**-294.409**	**689.193**	**748.81**	**-39.613**	**14938.25**	**475.79**
$\widehat{t}_{mcal}^*$	**-0.319**	**0.001**	**636106865.52**	**-0.426**	**1.731**	**4106489.5**

4.3. Discussion

Data from the four populations in Tables 3–5 were used to generate the findings of the empirical investigation, which are shown in Tables 6–8. The modified estimator $\hat{t}_m^*$ has a comparable small absolute bias across the four populations as shown in Table 4. However, $\hat{M}_{EP}$ was observed to have smaller bias than the proposed estimator under population I and II. Again, Table 7 shows that the proposed estimator outperforms existing estimators in terms of efficiency gains under simple random sampling. However, the alternative exponential estimator $\hat{M}_{AE}$ has the same PRE as the proposed estimator in all the four populations perhaps. Nevertheless, the proposed estimator becomes superior to $\hat{M}_{AE}$ in terms of small bias. This also lends support to several claims made in the literature about weight selection. It is of interest to know that in Table 8, when the calibration weight modification approach is utilized under stratified sampling, the result shows that the calibration estimator $\hat{t}_{mcal}^*$ outperforms other existing estimators while exhibiting low bias and larger efficiency gains. This validates the arguments made in the literature by Deville and Sarndal (1992) that the calibration estimator would perform at its best when an auxiliary variable is carefully chosen and is closely related to the study variable. Specifically, a constraint on the typical median estimate of the auxiliary variable's variance is followed by the remarkable results of the calibration estimator in this scenario.

5. Conclusion and Recommendation

This paper proposes an efficient exponential ratio-type estimator of population median for simple and stratified random sampling. The bias and mean square error of the proposed estimator, as well as the optimality criteria, were derived. When the proposed estimator's efficiency is compared to that of various existing estimators, theoretical findings supported by empirical demonstrations reveal that the new estimator outperforms existing estimators under simple random sampling. More precisely, using a calibration rather than direct weight adjustment when changing stratum weights in stratified random sampling, the proposed estimator outperforms existing estimators in terms of efficiency gain and bias. When the population median of the auxiliary variable is known and positively correlates with the interest variable, this result demonstrates that calibration weight modification is more appropriate for increasing the stability as well as efficiency of the population median estimator in stratified random sampling.

The principal concern is that the distribution of target variables is typically asymmetric in many practical situations, particularly in economic surveys, and that certain units may have high values in comparison to others (outliers). From one point of view, it is possible that removing these units will lead to skewed results. Hence, this approach could be extremely useful in surveys where the distribution of variables such as income or revenue is highly asymmetric. To be as close as possible to the original design, weights must be positive and should fall within specific desirable limits. Some distance functions require extremely large or negative calibration weights, which is extremely undesirable in terms of estimation. Taking into consideration an appropriate distance function that can exclude negative or large calibration weights while satisfying the given calibration equations should be sufficient to meet this requirement as seen in the proposed estimator. In addition, rather than using the conventional method of weight adjustments, it will be beneficial to use the calibration approach, because the proposed estimator has desirable properties that can be seen in the results, rather than the conventional method of weight adjustments.

References

Aczel, A. D. & Sounderpandian, J. (2004). *Complete business statistics*, (5^{th} ed.), McGraw Hill, New York.

Andersson, P. G. & Thorburn, D. (2005). An optimal calibration distance leading to the optimal regression estimator. *Survey Methodology*, 31, 95-99.

Al, S. & Cingi, H. (2009). New estimators for the population median in simple random sampling. Tenth Islamic countries conference on statistical sciences, held in New Cairo, Egypt. 375-383, 2009.

Aladag, S. & Cingi, H. (2015). Improvement in estimating the population median in simple random sampling and stratified random sampling using auxiliary information. *Communication in Statistics-Theory and Methods, 45(5)*, 1013-1032.

Bahl, S. & Tuteja, R. K. (1991). Ratio and product type exponential estimator. *Journal of Information and Optimization Sciences, 12(1)*, 159-164.

Baig, A., Masood, S. & Tarray, T. A. (2020). Improved class of difference-type estimators for population median in survey sampling. *Communications in Statistics: Theory and Methods*, 49(23), 5778–5793.

Chen, Z., Bai, Z. & Sinha, B. K. (2004). *Ranked set sampling: theory and applications*. New York, Springer-Verlag.

Clement, E. P. & Enang, E. I. (2017). On the Efficiency of Ratio Estimators over Regression Estimators, *Communication in Statistics- theory and Methods*, 46, 5357-5367.

Deville, J. C. & Särndal, C. E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association, 87*, 376–382.

Enang, E. I., Etuk, S. I., Ekpenyong, E. J. & Akpan, V. M. (2016). An alternative exponential estimator of population median. *International Journal of Statistics and Economics, 17(3)*, 85-97.

Gross, T. S. (1980). Median estimation in sample surveys. *In:* American Statistical Association Proceedings of survey research methodology section, 181-184.

Gupta, S., Shabbir, J. & Ahmad, S. (2008) Estimation of median in two-phase sampling using two auxiliary variables, *Communications in Statistics—Theory and Methods*, 37(11), 1815–1822.

Iseh, M. J. (2020). Enhancing efficiency of ratio estimator of population median by calibration techniques. *International Journal of Engineering Sciences & Research Technology, 9(8)*, 14-23.

Iseh, M. J. & Enang, E. I. (2021) A calibration synthetic estimator of population mean in small area under stratified sampling design. *Transition in Statistics new series*, 22(3), 15-30. DOI: 10.21307/stattrans-2021-025.

Kadilar, C. & Cingi, H. (2004). Ratio estimators in simple random sampling. *Applied Mathematical Computations, 151*, 893-902.

Kott, P. S. (2015). Calibration weighting in survey sampling. *WIREs Computational Statistics, 8(1)*, 39-53.

Kott, P. S. (2006). Using calibration weighting to adjust for nonresponse and coverage errors. *Survey Methodology*, 32,133-142.

Kuk, A. Y. C. & Mak, T. K. (1989). Median estimation in the presence of auxiliary variable. *Journal of Royal Statistical Society. Series B, 51*, 261-269.

Murthy, M. N. (1964). Product method of estimation. *Sankhya: Indian Journal of Statistics, Series A, 26*, 69-74.

Panjab Development Statistics (2012). Bureau of Statistics. Government of the Punjab, Lahore, Pakistan.

Särndal, C. E. (2007). The calibration approach in survey theory and practice. *Survey Methodology, 33*, 99–119.

Singh, H. P., Singh, S. & Puertas, S. M. (2003). Ratio-type estimators for the median of finite populations. *Allegemeines Statistisches Archiv, 87*, 369-382.

Singh, S. (2003). *Advanced sampling theory and applications: How Michael 'Selected' Amy*. Kluwer academics publishers, the Netherlands.

Singh, H. P. & Solanki, R. S. (2013). Some classes of estimators for the population median using auxiliary information. *Communication in Statistics, 42*, 4222-4238.

Robson, D. S. (1957). Application of multivariate polykays to the theory of unbiased ratio-type estimator. *Journal of the American Statistical Association, 52*, 511-522.