Estimation of Mean in Single-phase Sampling with High and Low Extreme Maximum values using auxiliary information

Peter I. Ogunyinka[1*], Gajendra K. Vishwakarma[2], Emmanuel F. Ologunleko[3], Ademola A. Sodipo[3]

[1:] Department of Mathematical Sciences, Olabisi Onabanjo University, Ago-Iwoye, Nigeria
ogunyinka.peter@gmail.com : +234706-758-4945, http://orcid.org/0000-0002-4811-7203
[2:] Department of Mathematics & Computing, Indian Institute of Technology (ISM) Dhanbad, Dhanbad-826004, India
[3:] Department of Statistics, University of Ibadan, Ibadan, Nigeria
**Corresponding author*

Abstract
Significant improvement has been introduced to regression-in-ratio estimators in simple random sampling. However, such improvement will be jeopardized when there is extreme maximum or minimum value in survey data. This study has proposed three improved regression-in-ratio estimators that would correct the over-estimation or under-estimation effect as a result of extreme maximum or minimum values in survey data, respectively. The bias and the mean square error expressions were established for comparison of the proposed estimators. Theoretical comparison and empirical comparison, through simulation for twenty six populations with high and low extreme maximum values, confirmed that the proposed estimators were, generally, efficient over the reviewed estimators. Though, the proposed estimators were less bias to the reviewed estimators, but they were confirmed to be asymptotically efficient. Suggestion for further study in the detection of significant extreme values in sample survey data was proposed.

Keywords/Phrase: Regression-in-ratio estimators, maximum values, minimum values, simple random sampling, efficiency

2010 AMS Classification: 62D05

Introduction

Auxiliary information has proved significant in the estimation of population parameters in sample survey theory. Simple random sampling estimators maximize on the advantages of auxiliary information. Similarly, ratio, regression and product estimators are efficient over the conventional Simple Random Sampling Without Replacement (SRSWOR) estimator when there is high correlation between the study and auxiliary variable(s) (Cochran, 1940; Robson, 1957 and Murthy, 1967). Mixed estimator combines two or more of ratio, regression or product estimator(s) into one estimator (Mohanthy, 1967 and Kiregrera, 1984). Ratio-in-regression, regression-in-ratio and ratio-cum-regression estimators are few examples of mixed estimators which had proved efficient over simple estimators. Similarly, few recent improved estimators in Survey Statistics that used auxiliary information included Singh *et al.* (2020), Sajjad *et al.* (2021) and Shabbir *et al.* (2021). However, such estimator would not be efficient in the presence of extreme values(s).

Abbas *et al.* (2018) argued in the direction of Sarndal (1972) that extreme values (either maximum or minimum value) will cause over or under estimation of the estimated parameter, respectively. However, the study was not primarily focused on the correction of this extreme value effect on the estimator because the methodology and conclusion of the study did not justify the aim of Sarndal

(1972). Sarndal (1972) had argued in the direction of Godambe (1955 and 1969) to describe the uniqueness of Sample Survey Theory to General Statistical (Statistical Inference) Theory. Let $X = (X_1, X_2, \dots, X_N)$ denotes the ordered population units while $x = (x_1, x_2, \dots, x_n)$ denotes the ordered sample value obtained from the N population.

$$\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n} \tag{1}$$

The conventional sample mean using SRSWOR as shown in equation (1) is a Uniformly Minimum Variance Unbiased Estimator (UMVUE) of the population mean $\bar{X}$. However, if a priori information has confirmed that X_N (and x_n in the sample) is exceedingly large (maximum value), then using equation (1) will yield over-estimated population mean. Similarly, if the prior knowledge confirms that X_1 (and x_1 in the sample) is exceedingly small (minimum value), then using equation (1) will yield under-estimated population mean. These maximum and minimum values are called extreme values.

Sarndal (1972) had proffered a unique solution to the correction of the extreme value effect in SRS in sample survey theory.-Khan and Shabbir (2013) seemed to be the first study that applied the method of Sarndal (1972) to ratio, regression and product estimators. Few authors have used the proposed method of Sarndal (1972) to correct for the effect of extreme value in both the study and one auxiliary variable. Al-Hossain and Khan (2014) minimized the extreme value effects in ratio, product and regression estimators with two auxiliary variables. Finally, Khan *et al.* (2015) improved on the ratio-type estimators with extreme value effect. This study aims to improve on the recent work of Abbas *et al.* (2018) ratio estimators by correcting the effect of the extreme values in the estimators using the method of Sarndal (1972). Similarly, this study will ascertain if the over-estimation or under-estimation of estimators has consequence on the bias and/or variance of the estimators.

Methodology

Reviewed ratio estimators

Abbas *et al.* (2018) had improved on the ratio estimators developed by Subramani and Kumarapandiyan (2012) by replacing the known median value of the auxiliary variable with the known maximum value of the auxiliary variable. The ratio estimators are presented as

$$\bar{y}_{P1} = \frac{\bar{y} + b(\bar{X} - \bar{x})}{(\bar{x} + M_x)}(\bar{X} + M_x), \tag{2}$$

$$\bar{y}_{P2} = \frac{\bar{y} + b(\bar{X} - \bar{x})}{(\bar{x}C_x + M_x)}(\bar{X}C_x + M_x), \tag{3}$$

$$\bar{y}_{P3} = \frac{\bar{y} + b(\bar{X} - \bar{x})}{(\bar{x}\rho_{yx} + M_x)}(\bar{X}\rho_{yx} + M_x), \tag{4}$$

where M_x and C_x are the maximum value and the coefficient of variation of the auxiliary variable, respectively. The b is the regression coefficient, $\overline{X}$ and $\overline{x}$ are the population and the sample means of the auxiliary variables, respectively. The $\overline{y}$ is the study variable and ρ_{yx} is the correlation coefficient of the study and auxiliary variables. The corresponding general form of bias and the Mean Square Errors (MSEs) were presented as:

$$B\left(\overline{y}_{Pi}\right) = \frac{\theta S_x^2}{\overline{Y}} R_{Pi}^2,$$

$$MSE\left(\overline{y}_{Pi}\right)_{min} \cong \theta\left(R_{Pi}^2 S_x^2 + S_y^2\left(1-\rho_{yx}^2\right)\right),$$

Where $i = 1, 2, and\ 3$; $\theta = \left(\frac{1}{n} - \frac{1}{N}\right)$; $R_{P1} = \left(\frac{\overline{Y}}{\overline{X}+M_x}\right)$; $R_{P2} = \left(\frac{\overline{Y}C_x}{\overline{X}C_x+M_x}\right)$ and $R_{P3} = \left(\frac{\overline{Y}\rho_{yx}}{\overline{X}\rho_{yx}+M_x}\right)$.

The S_x^2 and S_y^2 are the population variances of the auxiliary and study variables, respectively. These developed estimators of Abbas *et al.* (2018) proved efficient over the reviewed estimators of Subramani and Kumarapandiyan (2012).

Although, Abbas *et al.* (2018) referred to these estimators as ratio estimators but Muhanty (1967) had earlier referred to them as regression-in-ratio estimators. This study would refer to these estimators as regression-in-ratio estimator because the presence of the regression estimator in the referred estimator is very obvious and significant.

Proposed regression-in-ratio estimators

In equations (2, 3 and 4), this study assumes that there is extreme maximum or minimum value in both the sample of the study and auxiliary variables. The new estimators, based on the modification on equations (2, 3 and 4) are presented in equations (5, 6 and 7), respectively as:

$$\overline{y}_1 = \frac{\overline{y}_{c_0} + b\left(\overline{X} - \overline{x}_{c_1}\right)}{\left(\overline{x}_{c_1} + M_x\right)}\left(\overline{X} + M_x\right), \quad (5)$$

$$\overline{y}_2 = \frac{\overline{y}_{c_0} + b\left(\overline{X} - \overline{x}_{c_1}\right)}{\left(\overline{x}_{c_1}C_x + M_x\right)}\left(\overline{X}C_x + M_x\right), \quad (6)$$

$$\overline{y}_3 = \frac{\overline{y}_{c_0} + b\left(\overline{X} - \overline{x}_{c_1}\right)}{\left(\overline{x}_{c_1}\rho_{yx} + M_x\right)}\left(\overline{X}\rho_{yx} + M_x\right), \quad (7)$$

where $C_{0(opt)} = \frac{(y_{max}-y_{min})}{2n}$; and $C_{1(opt)} = \frac{(x_{max}-x_{min})}{2n}$.

To obtain the bias and the MSE for $\overline{y}_1$ in equation (5), substitute $\overline{Y}(1+\overline{e}_o)\ for\ \overline{y}_{co}$, $\overline{X}(1+\overline{e}_1)$ for $\overline{x}_{c1}$ and simplify, such that $E(\overline{e}_o) = E(\overline{e}_1) = 0$, then

$$E(e_0^2) = \frac{\theta}{\overline{Y}^2}\left(S_y^2 - \frac{2nc_0}{N-1}\left(y_{max} - y_{min} - nc_0\right)\right),$$

$$E(e_1^2) = \frac{\theta}{\overline{X}^2}\left(S_x^2 - \frac{2nc_1}{N-1}(x_{max} - x_{min} - nc_1)\right), \qquad and$$

$$E(e_0e_1) = \frac{\theta}{\overline{YX}}\left(S_{yx} - \frac{n}{N-1}(c_1(y_{max} - y_{min}) + c_0(x_{max} - x_{min}) - 2nc_0c_1)\right).$$

Hence, $\overline{y}_1 = (\overline{Y} + \overline{Y}\overline{e}_o - b\overline{X}\overline{e}_1)(1 + K_1\overline{e}_1)^{-1}$, such that $K_1 = \alpha_1\left(\frac{\overline{X}}{\overline{Y}}\right)$ and $\alpha_1 = \left(\frac{\overline{Y}}{\overline{X}+M_x}\right)$.

Apply Taylor series of expansion and the expectation, thereafter,

$$Bias(\overline{y}_1) = E(\overline{y}_1 - \overline{Y}) = E(\overline{Y}\overline{e}_o - \overline{Y}K_1\overline{e}_1 + \overline{Y}K_1^2\overline{e}_1^2 - \overline{Y}K_1\overline{e}_o\overline{e}_1 - b\overline{X}\overline{e}_1 + bK_1\overline{X}\overline{e}_1^2),$$

$$Bias(\overline{y}_1) = \frac{\theta S_x^2}{\overline{Y}}\alpha_{P1}^2 - \varepsilon_1 = Bias(\overline{y}_{P1}) - \varepsilon_1 \qquad (8)$$

where $\varepsilon_1 = [2c_1(\alpha_1 + \hat{\beta})(\Delta_x - nc_1) - (c_1\Delta_y + c_0\Delta_x - 2nc_0c_1)]\left[\frac{n\theta_1}{\overline{Y}(N-1)}\right]$; $\hat{\beta} = \left(\frac{S_{yx}}{S_x^2}\right)$;

$\Delta_y = (y_{max} - y_{min})$; *and* $\Delta_x = (x_{max} - x_{min})$.

To obtain the $MSE_{min}(\overline{y}_1)$

$$MSE\ (\overline{y}_1) = E\left[\overline{Y}^2\overline{e}_0^2 - 2\overline{Y}\overline{e}_o\overline{e}_1(\overline{Y}K_1 + b\overline{X}) + \left(\overline{Y}^2K_1^2 + 2\overline{Y}\overline{X}K_1b + b^2\overline{X}^2\right)\overline{e}_1^2\right]$$

Apply expectation and simplify further to obtain

$$MSE_{min}(\overline{y}_1) \cong MSE_{min}(\overline{y}_{P1}) - \gamma_1, \qquad (9)$$

where

$$\gamma_1 = [c_0(\Delta_y - nc_0) - (\alpha_1 + \hat{\beta})(c_1\Delta_y + c_0\Delta_x - 2nc_0c_1) + c_1(\Delta_x - nc_1)(\alpha_1^2 + 2\alpha_1\hat{\beta} + \hat{\beta}^2)]\left[\frac{2n\theta}{(N-1)}\right].$$

The general form of the obtained Bias and minimized MSE of the proposed $\overline{y}_1$, $\overline{y}_2$ and $\overline{y}_3$ are presented as

$$Bias(\overline{y}_i) = \frac{\theta S_x^2}{\overline{Y}}\alpha_{Pi}^2 - \varepsilon_i = Bias(\overline{y}_{Pi}) - \varepsilon_i \qquad (10)$$

and

$$MSE_{min}(\overline{y}_i) \cong MSE_{min}(\overline{y}_{Pi}) - \gamma_i, \qquad (11)$$

such that $\varepsilon_i = [2c_1(\alpha_i + \hat{\beta})(\Delta_x - nc_1) - (c_1\Delta_y + c_0\Delta_x - 2nc_0c_1)]\left[\frac{n\theta\alpha_i}{\overline{Y}(N-1)}\right]$, $\gamma_i = [c_0(\Delta_y - nc_0) - (\alpha_i + \hat{\beta})(c_1\Delta_y + c_0\Delta_x - 2nc_0c_1) + c_1(\Delta_x - nc_1)(\alpha_i^2 + 2\alpha_i\hat{\beta} + \hat{\beta}^2)]\left[\frac{2n\theta}{(N-1)}\right]$,
$\alpha_2 = \left(\frac{\overline{Y}C_x}{\overline{X}C_x + M_x}\right)$ *and* $\alpha_3 = \left(\frac{\overline{Y}\rho_{yx}}{\overline{X}\rho_{yx} + M_x}\right)$. The $i = 1, 2$ *and* 3, the c_0 *and* c_1 are obtained from $C_{0(opt)}$ and $C_{1(opt)}$ respectively.

Theoretical and Empirical Comparison

Theoretical comparison of the proposed estimators with the corresponding reviewed estimators of Abbas et al. (2018) considering the bias and the MSEs of the estimators

a. Comparison of the proposed and corresponding reviewed estimators based on the Biasness

i. $$Bias(\overline{y}_i) - Bias(\overline{y}_{Pi}) = -[2c_1(\alpha_i + \hat{\beta})(\Delta_x - nc_1) - (c_1\Delta_y + c_0\Delta_x - 2nc_0c_1)]\left[\frac{n\theta\alpha_i}{\overline{Y}(N-1)}\right].$$

The bias characteristic of the proposed estimator $\overline{y}_i$ would be determined in the empirical analysis.

b. This section compares the proposed estimators with the reviewed estimators based on the MSEs.

➢ $$(MSE_{min}(\overline{y}_i) - MSE_{min}(\overline{y}_{Pi}) = -\gamma_i) < 0$$

The estimators $\overline{y}_i$ will be efficient over $\overline{y}_{Pi}$ if $\gamma_i > 0$, such that $(-\gamma_i) < 0$. This would be subjected to empirical comparison.

Empirical comparison of estimators

This section compares the developed estimators and the reviewed estimators using numerical case. An algorithm and R code were developed for the study. Sixteen (16) simulated populations each with different population and sample sizes were developed. The R code is deposited on https://bit.ly/2GMG0Og as free and open source code. The algorithm conducted the simulation and analysis in accordance to the following procedure:

- ❑ Selection of different population and sample sizes for sixteen populations in asymptotic procedure.
- ❑ Simulation of data following normal distribution with pre-defined location and scalar values for each of the sixteen populations and for two population variables, Y and X with pre-defined high and positive correlation coefficient.
- ❑ An extreme maximum value was injected into each of Y and X such that these extreme values were also sampled into the sample variables y and x.
- ❑ The extreme value is structured into four classes (see Figure 1). The four classes are High Extreme Maximum Value (HEMaV), Low Extreme Maximum Value (LEMaV), High Extreme Minimum Value (HEMiV) and Low Extreme Minimum Value (LEMiV). However, this study focuses on HEMaV and LEMaV cases only.

- For each of HEMaV and LEMaV, other necessary parameters were computed, the bias, MSE and variance were, also, computed. Finally, the relative efficiencies were computed, all for the sixteen populations.

Tables 1 through 10 display the analysis results for the bias, MSEs and relative efficiencies.

Explanation on HEMaV and LEMaV cases

This study documented four cases at which extreme values can be present in a survey data. The conditions are High Extreme Maximum Values (HEMaV) case, High Extreme Minimum Values (HEMiV) case, Low Extreme Maximum Values (LEMaV) case and Low Extreme Minimum Values (LEMiV) case.

HEMaV distribution considered that both the study and the auxiliary variables had extreme values present in the data and it has positive (maximum) and very high (high) extreme values in the two variables, $Y \text{ and } X$. Such that these extreme values were sampled from the population data into the sample data. The HEMiV data considered the extreme values case such that there exists negative (minimum) and high negative (high) extreme values in both the population and sample data of the study and the auxiliary variables. The LEMaV data considered the extreme values case such that there is positive (maximum) but low positive (low) extreme values in both the population and sample data of the study and the auxiliary variables. Finally, the LEMiV case considered the extreme values case such that there is negative (minimum) and low negative (Low) extreme values in both the population and sample data of the study and the auxiliary variables. Figure 1 explains the nature of HEMaV, HEMiV, LEMaV and LEMiV cases on a number line.

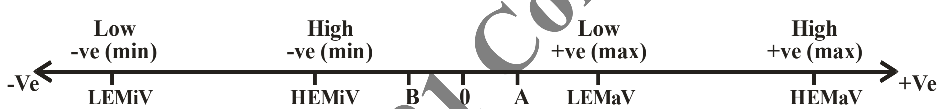

Figure 1: Calibrated number line for the explanation of HEMaV, HEMiV, LEMaV and LEMiV cases

Discussion

It had been mentioned that the efficiency of the proposed estimators would be ascertained empirically because the theoretical comparison could not confirm the efficiency. Hence, the empirical comparison was conducted in this section. Similarly, the asymptotic characteristic of the estimators was also confirmed empirically. The analyses were conducted for High Extreme Maximum Value (HEMaV) and Low Extreme Maximum Value (LEMaV) cases only.

Table 1 showed the distribution of the population and sample sizes using simulated data for sixteen (16) different populations. These populations were used in order to confirm the asymptotic characteristics of the estimators. Tables 2 through 5 showed the analyses for the bias, variance, Mean Square Error (MSE) and the Relative Efficiency (RE) analysis results for HEMaV case while Tables 6 through 9 showed the bias, variance, MSE and the relative efficiency analysis results for the LEMaV case. Finally, Table 10 showed the comparative analysis of the relative efficiency results for both HEMaV and LEMaV cases. The proposed estimators $(\overline{y}_1, \overline{y}_2 \text{ and } \overline{y}_3)$ would be compared with the corresponding reviewed estimators ($\overline{y}_{P1}$, $\overline{y}_{P2}$ and $\overline{y}_{P3}$) as developed by Abbas *et al.* (2018).

HEMaV Case Analysis

Table 2 showed that the bias of the reviewed (Abbas *et al.*, 2018) estimators $\overline{y}_{P1}$ and $\overline{y}_{P2}$had smaller overall bias values (with high rank) over the corresponding proposed estimators $\overline{y}_1$ and $\overline{y}_2$asymptotically. Contrarily, the reviewed estimator $\overline{y}_{P3}$ had large bias value over the corresponding proposed estimator $\overline{y}_3$. In general, among the six estimators, the proposed estimator $\overline{y}_3$ had the least bias (rank 1) while the $\overline{y}_2$ estimator had the highest bias (rank 6) (see Table 2). Hence, $\overline{y}_1$ and $\overline{y}_2$ estimators were asymptotically less efficient over the corresponding $\overline{y}_{P1}$ and $\overline{y}_{P2}$ estimators, respectively while $\overline{y}_3$ was asymptotically more efficient over the corresponding $\overline{y}_{P3}$.

Tables 3 and 4 revealed that the proposed estimators ($\overline{y}_3, \overline{y}_1$ and $\overline{y}_2$) asymptotically, had the least variance and MSE over the corresponding reviewed estimators ($\overline{y}_{P3}, \overline{y}_{P1}$ and $\overline{y}_{P2}$), respectively, as developed by Abbas *et al.* (2018). Tables 3 and 4 showed that $\overline{y}_3$ was the most efficient estimator while $\overline{y}_{P2}$ was the least efficient estimator among the six proposed and reviewed estimators. Hence, $\overline{y}_3, \overline{y}_1$ and $\overline{y}_2$ estimators were respectively and asymptotically efficient over the reviewed estimators $\overline{y}_{P3}$, $\overline{y}_{P1}$ and $\overline{y}_{P2}$, respectively, by Abbas *et al.* (2018). Table 5 showed that the proposed estimators $\overline{y}_1, \overline{y}_2$ and $\overline{y}_3$ were 120.16%, 118.93% and 120.39% asymptotically efficient over the corresponding reviewed estimators $\overline{y}_{P1}$, $\overline{y}_{P2}$ and $\overline{y}_{P3}$, respectively. Hence, the proposed estimators were placed in order of efficiency as $\overline{y}_3, \overline{y}_1$ and $\overline{y}_2$ when there is high extreme maximum value in the dataset (see Table 10).

LEMaV Case Analysis

Table 6 revealed that the developed estimators $\overline{y}_1, \overline{y}_2$ and $\overline{y}_3$were ranked lower than the reviewed estimators $\overline{y}_{P1}$, $\overline{y}_{P2}$ and $\overline{y}_{P3}$, respectively. The three reviewed estimators by Abbas *et al.* (2018) ($\overline{y}_{P1}$, $\overline{y}_{P2}$ and $\overline{y}_{P3}$) had smaller bias over the corresponding proposed estimators ($\overline{y}_1, \overline{y}_2$ and $\overline{y}_3$). In general, the revealed estimator $\overline{y}_{P3}$was the least bias estimator while the proposed estimator $\overline{y}_1$had the highest bias.

Tables 7 and 8 showed that the proposed estimators $\overline{y}_1, \overline{y}_2$ and $\overline{y}_3$, asymptotically, had smaller variances and MSEs compared to the corresponding reviewed estimators, $\overline{y}_{P1}$, $\overline{y}_{P2}$ and $\overline{y}_{P3}$, of Abbas *et al.* (2018). It was revealed that the proposed estimator $\overline{y}_3$ had the least variance and MSE while the reviewed estimator $\overline{y}_{P1}$had the highest variance and the MSE. Hence, the three proposed estimators $\overline{y}_1, \overline{y}_2$ and $\overline{y}_3$ proved to be asymptotically efficient over the corresponding three reviewed estimators ($\overline{y}_{P1}$, $\overline{y}_{P2}$ and $\overline{y}_{P3}$) by Abbas *et al.* (2018). Table 9 revealed that the proposed estimator $\overline{y}_1, \overline{y}_2$ and $\overline{y}_3$ were 105%, 119% and 120% relatively efficient over the corresponding reviewed estimators $\overline{y}_{P1}$, $\overline{y}_{P2}$ and $\overline{y}_{P3}$, respectively. Hence, the proposed estimators were ranked in order of efficiency as $\overline{y}_3, \overline{y}_2$ *and* $\overline{y}_1$ when there is low extreme maximum value(s) in the dataset (see Table 10).

Summary

This study had extended the ratio estimators developed by Abbas *et al.* (2018) with the method of Sarndal (1972) for the correction of the presence of extreme value in the sample survey data. Three improved regression-in-ratio estimators were proposed using coefficient of variation, correlation coefficient and extreme value correction factor in single-phase sampling. The proposed estimators were asymptotically tested under two types of extreme value in sample survey data. These extreme values were High Extreme Maximum Value (HEMaV) and Low Extreme Maximum Value (LEMaV) cases. The bias and the efficiency of the proposed estimators were confirmed using the empirical biasness, variance, MSEs and the relative efficiency under HEMaV and LEMaV cases. Results revealed that the three proposed estimators reacted differently under HEMaV and LEMaV cases. However, the three proposed estimators were asymptotically efficient over the reviewed estimators of Abbas *et al.* (2018).

Acknowledgement

The authors appreciate the professional contribution of Dr. Kayode Olurinola, the Chief Medical Director of Steksons and Sons Eye Clinic in Ijebu-Ode of Ogun State, Nigeria.

Table 1: Distribution of the population and sample sizes over the sixteen (16) simulated populations

Population	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
N	5000	4650	4300	3950	3600	3250	2900	2550	2200	1850	1500	1150	800	450	100	60
n	1667	1550	1433	1317	1200	1083	967	850	733	617	500	383	267	150	33	20

Bias, Variance and the Mean Square Error (MSE) Analyses on the High Extreme Maximum Value (HEMaV) case

Table 2: Rank of the Bias of the proposed estimators and the reviewed estimators for the sixteen (16) populations for HEMaV case

Population	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**Rank Average**
$Bias(\overline{y}_1)$	5	5	5	5	5	5	5	5	5	5	5	4	4	4	4	4	**5**
$Bias(\overline{y}_{P1})$	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	**3**
$Bias(\overline{y}_2)$	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	**6**
$Bias(\overline{y}_{P2})$	4	4	4	4	4	4	4	4	4	4	4	5	5	5	5	5	**4**
$Bias(\overline{y}_3)$	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**1**
$Bias(\overline{y}_{P3})$	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	**2**

Table 3: Rank of the variance of the proposed estimators and the reviewed estimators for the sixteen (16) populations for HEMaV case

Population	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**Rank Average**
$Var(\overline{y}_1)$	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	**2**
$Var(\overline{y}_{P1})$	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	**5**
$Var(\overline{y}_2)$	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	**3**
$Var(\overline{y}_{P2})$	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	**6**
$Var(\overline{y}_3)$	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**1**
$Var(\overline{y}_{P3})$	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	**4**

Table 4: Rank of the MSE of the proposed estimators and the reviewed estimators for the sixteen (16) populations for HEMaV case

Population	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Rank Average
$MSE(\bar{y}_1)$	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	**2**
$MSE(\bar{y}_{P1})$	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	**5**
$MSE(\bar{y}_2)$	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	**3**
$MSE(\bar{y}_{P2})$	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	**6**
$MSE(\bar{y}_3)$	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**1**
$MSE(\bar{y}_{P3})$	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	**4**

Table 5: Relative efficiency of the proposed estimators to the corresponding reviewed estimators for the 16 populations in HEMaV analysis

Population	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Average
$RE(\bar{y}_1/\bar{y}_{P1})$	119.91	119.93	119.92	119.93	119.93	119.89	119.94	119.90	119.95	119.95	119.94	119.97	120.09	120.14	120.98	122.22	**120.16**
$RE(\bar{y}_2/\bar{y}_{P2})$	119.41	119.38	119.36	119.34	119.32	119.26	119.25	119.20	119.14	119.10	118.98	118.83	118.75	118.39	117.40	117.71	**118.93**
$RE(\bar{y}_3/\bar{y}_{P3})$	120.04	120.04	120.03	120.05	120.05	120.05	120.06	120.06	120.07	120.11	120.12	120.13	120.25	120.38	121.58	123.21	**120.39**

Bias, Variance and the Mean Square Error (MSE) Analyses on the Low Extreme Maximum Value (LEMaV) case

Table 6: Rank of the Bias of the proposed estimators and the reviewed estimators for the sixteen (16) populations for LEMaV case

Population	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Rank Average
$Bias(\bar{y}_1)$	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	**6**
$Bias(\bar{y}_{P1})$	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	**5**
$Bias(\bar{y}_2)$	3	3	4	4	4	4	4	4	4	4	4	4	4	4	4	4	**4**
$Bias(\bar{y}_{P2})$	1	1	2	2	2	2	2	2	2	2	2	3	3	3	3	3	**2**
$Bias(\bar{y}_3)$	4	4	3	3	3	3	3	3	3	3	3	2	2	1	1	1	**3**
$Bias(\bar{y}_{P3})$	2	2	1	1	1	1	1	1	1	1	1	1	1	2	2	2	**1**

Table 7: Rank of the variance of the proposed estimators and the reviewed estimators for the sixteen (16) populations for LEMaV case

Population	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Rank Average
$Var(\bar{y}_1)$	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	**5**
$Var(\bar{y}_{P1})$	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	**6**
$Var(\bar{y}_2)$	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	**2**
$Var(\bar{y}_{P2})$	3	3	3	3	3	3	3	4	3	4	4	4	4	4	4	4	**4**
$Var(\bar{y}_3)$	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**1**
$Var(\bar{y}_{P3})$	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	**3**

Table 8: Rank of the MSE of the proposed estimators and the reviewed estimators for the sixteen (16) populations for LEMaV case

Population	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Rank Average
$MSE(\bar{y}_1)$	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	**5**
$MSE(\bar{y}_{P1})$	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	**6**
$MSE(\bar{y}_2)$	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	**2**
$MSE(\bar{y}_{P2})$	3	3	3	3	3	3	3	4	3	4	4	4	4	4	4	4	**4**
$MSE(\bar{y}_3)$	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	**1**
$MSE(\bar{y}_{P3})$	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	**3**

Table 9: Relative efficiency of the proposed estimators to the corresponding reviewed estimators for the16 populations in LEMaV analysis.

Population	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Average
$RE(\bar{y}_1/\bar{y}_{P1})$	105.38	104.68	104.78	105.48	105.32	104.88	105.04	105.29	104.58	105.30	105.03	104.86	105.15	105.28	105.66	105.42	**105.13**
$RE(\bar{y}_2/\bar{y}_{P2})$	119.76	119.79	119.83	119.88	119.46	119.96	119.59	119.46	119.08	119.14	119.45	119.30	119.07	119.14	118.25	118.78	**119.37**

$RE(\bar{y}_3/\bar{y}_{P3})$	119.76	119.79	119.83	119.88	119.59	120.08	119.94	119.85	119.45	119.81	120.16	120.26	120.23	120.86	121.91	124.00	**120.34**

Table 10: Relative Efficiency (RE) comparison of the proposed estimators with the corresponding reviewed estimators

	HEMaV case		**LEMaV case**	
Relative Efficiency	**Average**	**Average Rank**	**Average**	**Average Rank**
$RE(\bar{y}_1/\bar{y}_{P1})$	120.16	2	105.13	3
$RE(\bar{y}_2/\bar{y}_{P2})$	118.93	3	119.37	2
$RE(\bar{y}_3/\bar{y}_{P3})$	120.39	1	120.34	1

References

Abbas, N., Abid, M., Tahir, M., Abbas, N. & Hussain, Z. (2018). Enhancing ratio estimators for estimating population mean using maximum value of auxiliary variable. *J. Natn. Sci. Foundation Sri Lanka* 3, 453-463.

Al-Hossain, A. & Khan, M. (2014). Efficiency of ratio, product and regression estimators under maximu and minimum values, using two auxiliary variables. *The Scientific World Journal*, Article ID:693782, 1-6.

Cochran, W. G. (1940). Sampling Technique. 3rd Edition, Wiley Eastern Limited, India.

Godambe, V. P. (1955). A unified theory of sampling from finite populations. *J. R. Statist. Soc*. B, 17, 268-278.

Godambe, V. P. (1969). Some aspects of the theoretical developments in survey-sampling. New development in survey sampling. New York, Wiley, 27-58.

Khan, M. & Shabbir, J. (2013). Some improved ratio, product and regression estimators of finite population mean when using minimum and maximum values. *The Scientific World Journal*, Article ID:431868, 1-7.

Khan, M., Ullah, S., Al-Hossain, A. & Bashir, N. (2015). Improve ratio-typw estimators using maximum and minimum values under simple random sampling scheme. *Hacettepe Journal of Mathematics and Statistics*. 44(4), 923-931.

Kiregrera, B. (1984). Regression-type estimator using two auxiliary variables and model of double sampling from finite populations. *Metrika*, 31, 215-226.

Mohanty, S. (1967). Combination of Regression and Ratio Estimate. *Jour. Ind. Statist. Assoc., 5*, 16-19.

Murthy, M. N. (1967). Sampling Theory and Mathods. Statistical Publishing Society, Calcutta, India.

Robson, D. (1957). Application of Multivariate Polykays to the Theory of unbiade Ratio Type Estimators. *Journal of American Statistical Association, 52*, 511-522

Sajjad, I., Hanif, M., Koyuncu, N., Shahzad, U., & Al-Noor, N. H. (2021). A new family of robust regression estimators utilizing robust regression tools and supplementary attributes. *Statistics in Transition New Series*, *22*(1).

Sarndal, C. (1972). Sample Survey theory Vs. general statistical theory: Estimation of the population mean. *Int. Stat. Rev.*, 40(1), 1-12.

Shabbir, J., Gupta, S., & Onyango, R. (2021). On Using the Conventional and Nonconventional Measures of the Auxiliary Variable for Mean Estimation. *Mathematical Problems in Engineering*, *2021*.

Singh H.P., Patidar P. & Gorey S.M. (2020). A ratio-cum-ratio-type exponential estimator for estimating the mean of a sensitive variable using auxiliary information. *International Journal of Agricultural and Statistical Sciences*, 16(1), 479 - 486.

Subramani, J. & Kumarapandiyan, G. (2012). A class of almost unbiased modified ratio estimators for population mean with known population parameters. *Elixir Statistics*, 44, 7411-7415.

Zaman, T., Dünder, E., Audu, A., Alilah, D. A., Shahzad, U., & Hanif, M. (2021). Robust Regression-Ratio-Type Estimators of the Mean Utilizing Two Auxiliary Variables: A Simulation Study. *Mathematical Problems in Engineering*, *2021*.