Caliphate Journal of Science & Technology (CaJoST)

ISSN: 2705-313X (**PRINT**); 2705-3121 (**ONLINE**) **Research Article**

Open Access Journal available at: https://cajost.com.ng/index.php/files and https://www.ajol.info/index.php/cajost/index

DOI: https://dx.doi.org/10.4314/cajost.v5i3.1

Article Info

Received: 10th August 2022

Revised: 22nd February 2023

Accepted: 24th February 2023

[1,3,5]Department of Statistics, Usmanu Danfodiyo University, Sokoto, Nigeria.

[2]Department of Statistics, Ahmadu Bello University, Zaria, Nigeria.

[4]Department of Statistics, Kano University of Sci. and Tech, Wudil, Nigeria

[6]Department of Statistics, Binyaminu Usman Polytechnic, Hadejia, Nigeria.

***Corresponding author's email:**

yahzaksta@gmail.com

Cite this: *CaJoST,* 2023, 3, 246-254

Regression-Cum-Ratio Mean Imputation Class of Estimators using Non-Conventional Robust Measures

Ahmed Audu[1], Yahaya Zakari[2*], Mojeed A. Yunusa[3], Ishaq O. Olawoyin[4], Faruk Manu[5], and Isah Muhammad[6]

Different imputation strategies have been developed by several authors to take care of missing observations during analyses. Nevertheless, the estimators involved in some of these schemes depend on known parameters of the auxiliary variable which outliers can easily influence. In this study, a new class of ratio-type imputation methods that utilize parameters that are free from outliers has been presented. The estimators of the schemes were obtained and their MSEs were derived up to first-order approximation using the Taylor series approach. Also, conditions for which the new estimators are more efficient than others considered in the study were also established. Numerical examples were conducted and the results revealed that the proposed class of estimators is more efficient.

Keywords: Imputation, Non-response, Estimator, Population Mean, Mean Squared Error (MSE).

1. Introduction

It is often assumed at the beginning of the survey that information on sampling units drawn from the population is completely available. This assumption is often violated due to non-response due to incomplete information or inaccessibility to respondents or refusal to answer questions, especially surveys in medical and social science, etc. which often involve sensitive questions. In such situations, responses of non-respondents after often imputed or estimated using imputation techniques. Imputation is the process of replacing missing data with substituted values. There are three main problems that missing data due to non-response causes. It can introduce a substantial amount of bias, make the handling and analysis of the data more arduous, and create reductions in efficiency. Missing data due to non-response creates problems of complications during data analysis. The imputation approach provides all cases by replacing missing data with an estimated value based on other available information and auxiliary variable. Once all missing values have been imputed, the data set can then be analyzed using standard techniques for complete data.

[1] were the first to consider the problem of non-response. Several authors also proposed imputation methods to deal with non-response or missing values. Among them are [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [15], [16] and [17]. Recently, [18] suggested a generalized class of imputation in which they compared the efficiency of the estimators obtained from the scheme with that of the estimators of the schemes by the previous authors and found that their estimators outperformed the estimators of the previous authors. Nonetheless, having studied the estimators by [18], it was observed that the estimators depend on the known parameters of the auxiliary variable which outliers can easily influence. In this study, new classes of ratio-type imputation methods which utilized parameters that are free from outliers have been presented.

Notations

The following notations have been used as described in **[19],** [20], [21], [22], [23], and [24].

Y: Study variable.
X: Auxiliary variable.
$\bar{X}, \bar{Y}$: Population mean of the variables X and Y respectively.
N: Population Size.
n: Size of the sample
r: Number of respondents.
R: Ratio of the population mean of study variable to the population mean of auxiliary variable.
$\bar{x}_n$: The sample mean for the sample of size n.

$\bar{x}_r$: The mean of the variable X for set Φ

$\bar{y}_r$: The mean of the variable Y for set Φ

S_Y^2, S_X^2 : Population variance of the variables X and Y

S_Y, S_X : Population standard deviation of Y and X.

β_1 : Population coefficient of skewness of X.

β_2 : Population coefficient of kurtosis of X.

ρ_{YX} : Population coefficient of correlation between Y and X.

β_{rg} : Population regression coefficient.

C_Y, C_X : Population coefficient of variation of Y and X.

$G = \frac{4}{N-1}\sum_{i=1}^{N}\left(\frac{2i-N-1}{2N}\right)X_{(i)}$: Gini's mean difference for X.

$D = \frac{2\sqrt{\pi}}{N-1}\sum_{i=1}^{N}\left(i-\frac{N+1}{2N}\right)X_{(i)}$: Downtown's method for X.

$S_{pw} = \frac{\sqrt{\pi}}{N^2}\sum_{i=1}^{N}(2i-N-1)X_{(i)}$: Probability weighted moments for X.

$\sum$: Population variance covariance matrix.

Sample Mean and [18] Imputation Schemes

Let Φ denotes the set of r units response and Φ^c denotes the se of $n-r$ units non-response or missing out of n units sampled without replacement from the N units population. For each $i \in \Phi$, the value of y_i is observed. However, for unit $i \in \Phi^c$, y_i is missing but calculated using different methods of imputation.

The mean method of imputation is defined as

$$y_{.i} = \begin{cases} y_i & \text{i} \in \Phi \\ \bar{y}_r & \text{i} \in \Phi^c \end{cases} \qquad (1.1)$$

The point estimator of scheme in (1.1) denoted by $\hat{\mu}_0$ is given as in (1.2)

$$\hat{\mu}_0 = r^{-1}\sum_{i\in R} y_i \qquad (1.2)$$

The variance of $\hat{\mu}_0$ is given by (1.3).

$$Var(\hat{\mu}_0) = \psi_{r,N} S_Y^2 \qquad (1.3)$$

where

$$\psi_{r,N} = r^{-1} - N^{-1},\ S_Y^2 = (N-1)^{-1}\sum_{i=1}^{N}\left(y_i - \bar{Y}\right)^2,\ \bar{Y} = N^{-1}\sum_{i=1}^{N} y_i$$

[18] proposed the following generalized class of imputation schemes;

$$y_{.i} = \begin{cases} y_i & \text{i} \in \Phi \\ \frac{\hat{\mu}_0 + \hat{\beta}_{rg}\left(\bar{X}-\bar{x}_r\right)}{\pi_1\bar{x}_r + \pi_2}\left(\pi_1\bar{X}+\pi_2\right)\exp\left(\frac{\varpi_1\left(\bar{X}-\bar{x}_r\right)}{\varpi_1\left(\bar{X}+\bar{x}_r\right)+2\varpi_2}\right) & \text{i} \in \Phi^c \end{cases} \qquad (1.4)$$

where π_1 and π_2 are known functions of auxiliary variables like coefficients of skewness $\beta_1(x)$, kurtosis $\beta_2(x)$, variation C_X, standard deviation S_X etc, and $\pi_1 \neq \pi_2, \varpi_1 = 1, \varpi_2 = 0$ and $\pi_1 \neq 0$. .

The point estimators of finite population mean under these methods of imputation are given by

$$\hat{\mu}_q^{(*)} = \frac{r}{n}\hat{\mu}_0 + \left(1-\frac{r}{n}\right)\frac{\hat{\mu}_0 + \hat{\beta}_{rg}\left(\bar{X}-\bar{x}_r\right)}{\pi_1\bar{x}_r + \pi_2}\left(\pi_1\bar{X}+\pi_2\right)\exp\left(\frac{\varpi_1\left(\bar{X}-\bar{x}_r\right)}{\varpi_1\left(\bar{X}+\bar{x}_r\right)+2\varpi_2}\right) \qquad (1.5)$$

$$MSE\left(\hat{\mu}_q^{(*)}\right) = \psi_{r,N}(S_Y^2 + \Upsilon^2 S_X^2 - 2\Upsilon S_{YX}) \qquad (1.6)$$

where, $\eta_1 = \frac{\pi_1\bar{X}}{\pi_1\bar{X}+\pi_2}$, $\eta_2 = \frac{\varpi_1\bar{X}}{2(\varpi_1\bar{X}+\varpi_2)}$,

and $\Upsilon = \left(1-\frac{r}{n}\right)(R(\eta_1+\eta_2)+\beta_{rg})$

2. Materials and Methods

Proposed Imputation Scheme

Having studied the scheme and estimators of [18], which utilize known functions of the auxiliary variable, that are sensitive to outliers, we proposed the following scheme to obtain estimators that are not sensitive to outliers by using nonconventional robust measures of the auxiliary variable defined as

$$y_{.i} = \begin{cases} y_i & i \in \Phi \\ \frac{\bar{y}_r + \beta_{rg}(\bar{X}-\bar{x}_r)}{\phi_j\bar{x}_r + \phi_k}(\phi_j\bar{X}+\phi_k) & i \in \Phi^c \end{cases} \qquad (2.1)$$

The associated class of point estimators for scheme (2.1) is obtained as

$$\hat{t}_m^1 = \frac{r}{n}\bar{y}_r + \left(1-\frac{r}{n}\right)\frac{(\bar{y}_r + \beta_{rg}(\bar{X}-\bar{x}_r))}{\phi_j\bar{x}_r+\phi_k}\left(\phi_j\bar{X}+\phi_k\right) \quad (2.2)$$

Remark 1: Note that $\phi_j \neq \phi_k$,

where,

$\phi_t = \left\{(G\times n),(D\times n),(S_{pw}\times n)\right\}, \left(\phi_j,\phi_k\right)\in\phi_t,\ t=j,k$

, $m = 1,2,3,4,5,6$,

Table 1: Some member of $\hat{t}_i^1$ for different values of ϕ_j and ϕ_k

i	Estimators	Values of Constants	
		ϕ_j	ϕ_k
1	$\hat{t}_1^1 = \frac{r}{n}\bar{y}_r + \left(1-\frac{r}{n}\right)\frac{(\bar{y}_r+\beta_{rg}(\bar{X}-\bar{x}_r))}{(G\times n)\bar{x}_r+(D\times n)}((G\times n)\bar{X}+(D\times n))$	$(G\times n)$	$(D\times n)$
2	$\hat{t}_2^1 = \frac{r}{n}\bar{y}_r + \left(1-\frac{r}{n}\right)\frac{(\bar{y}_r+\beta_{rg}(\bar{X}-\bar{x}_r))}{(G\times n)\bar{x}_r+(S_{pw}\times n)}((G\times n)\bar{X}+(S_{pw}\times n))$	$(G\times n)$	$(S_{pw}\times n)$
3	$\hat{t}_3^1 = \frac{r}{n}\bar{y}_r + \left(1-\frac{r}{n}\right)\frac{(\bar{y}_r+\beta_{rg}(\bar{X}-\bar{x}_r))}{(D\times n)\bar{x}_r+(G\times n)}((D\times n)\bar{X}+(G\times n))$	$(D\times n)$	$(G\times n)$
4	$\hat{t}_4^1 = \frac{r}{n}\bar{y}_r + \left(1-\frac{r}{n}\right)\frac{(\bar{y}_r+\beta_{rg}(\bar{X}-\bar{x}_r))}{(D\times n)\bar{x}_r+(S_{pw}\times n)}((D\times n)\bar{X}+(S_{pw}\times n))$	$(D\times n)$	$(S_{pw}\times n)$
5	$\hat{t}_5^1 = \frac{r}{n}\bar{y}_r + \left(1-\frac{r}{n}\right)\frac{(\bar{y}_r+\beta_{rg}(\bar{X}-\bar{x}_r))}{(S_{pw}\times n)\bar{x}_r+(G\times n)}((S_{pw}\times n)\bar{X}+(G\times n))$	$(S_{pw}\times n)$	$(G\times n)$
6	$\hat{t}_6^1 = \frac{r}{n}\bar{y}_r + \left(1-\frac{r}{n}\right)\frac{(\bar{y}_r+\beta_{rg}(\bar{X}-\bar{x}_r))}{(S_{pw}\times n)\bar{x}_r+(D\times n)}((S_{pw}\times n)\bar{X}+(D\times n))$	$(S_{pw}\times n)$	$(D\times n)$

Data for Empirical Study

For the empirical examples on the precision of proposed estimators, three sets of data obtained from [25] in Table 2 were used.

Table 2: Data used for empirical study

Parameters	Pop. 1	Pop. 2	Pop. 3
N	34	34	80
n	20	20	20
r (Assumed)	13	13	13
$\bar{Y}$	856.4117	856.4117	5182.637
$\bar{X}$	208.8823	199.4412	1126.463
C_Y	0.8561	0.8561	0.354193
C_X	0.7205	0.7531	0.7506772
$\beta_1(x)$	0.9782	1.1823	1.050002
$\beta_2(x)$	0.0978	1.0445	-0.063386
ρ_{YX}	0.4491	0.4453	0.9410
S_X	150.5059	150.2150	845.610

S_Y	733.1409	733.1407	1835.659
M_d	150	142.5	757.5
TM	162.25	165.562	931.562
MR	284.5	320	1795.5
HL	190	184	1040.5
QD	80.25	89.375	588.125
G	155.446	162.996	901.081
D	140.891	144.481	801.381
S_{pw}	199.961	206.944	791.364

Simulated Data for Empirical Study

In this section, Data of size 1000 units were generated for study populations using function defined in Table 3. Samples of size 100 units from which 60 units were selected as respondents were randomly chosen 10,000 times by method of simple random sampling without replacement (SRSWOR). The Biases and MSEs of the considered estimators were computed using (2.3) and (2.4) respectively

$$MSE\left(\hat{\theta}_d\right) = \frac{1}{10000}\sum_{d=1}^{10000}\left(\hat{\theta}_d - \bar{Y}\right)^2 , \hat{\theta}_d = \bar{y}_r, \hat{\mu}_i^{(*)}, i = 1,2,...,17, \hat{t}_m^1 \quad (2.3)$$

(2.4)

Table 3: Populations used for Simulation Study

Populations	Auxiliary variable (x)	Study variable (y)
I	$X \sim beta\left(1.1, 2.0\right)$	$Y = 50 + 10X + 20X^2 + e,$ where, $e \sim \left(0, 4\right)$
II	$X \sim gamma\left(10, 25\right)$	

3. Results and Discussion

Properties of the Estimators suggested

The MSE of estimators for suggested imputation schemes is obtained as

$$MSE\left(\hat{t}_m^1\right) = \Delta\Sigma\Delta^T$$

$$\Sigma = \begin{pmatrix} \psi_{r,N}S_Y^2 & \psi_{r,N}\rho S_Y S_X \\ \psi_{r,N}\rho S_Y S_X & \psi_{r,N}S_X^2 \end{pmatrix}$$

(3.1) where $\Delta = \left(\frac{\partial \hat{t}_m^1}{\partial \bar{y}_r}\right|$

On differentiating the estimator $\hat{t}_m^1$ with respect to $\bar{y}_r$ partially, we have,

$$\frac{\partial \hat{t}_m^1}{\partial \bar{y}_r} = \frac{r}{n} + \left(1 - \frac{r}{n}\right)\frac{(\phi_j \bar{X} + \phi_k)}{(\phi_j \bar{x}_r + \phi_k)} \quad (3.2)$$

On setting $\bar{y}_r = \bar{Y}, \bar{x}_r = \bar{X}, \hat{\beta}_{rg} = \beta_{rg}$, we have,

$$\left.\frac{\partial \hat{t}_m^1}{\partial \bar{y}_r}\right|_{\bar{y}_r = \bar{Y}, \bar{x}_r = \bar{X}, \hat{\beta}_{rg} = \beta_{rg}} = 1 \quad (3.3)$$

On differentiating the estimator $\hat{t}_m^1$ with respect to $\bar{x}_r$ partially, we have,

$$\frac{\partial \hat{t}_m^1}{\partial \bar{x}_r} = -\left(1 - \frac{r}{n}\right)(\phi_j \bar{X} + \phi_k)\left(\frac{(\phi_j \bar{x}_r + \phi_k)\hat{\beta}_{rg} + \left(\bar{y}_r + \hat{\beta}_{rg}(\bar{X} - \bar{x}_r)\right)\phi_j}{(\phi_j \bar{x}_r + \phi_k)^2}\right)$$

On setting $\bar{y}_r = \bar{Y}, \bar{x}_r = \bar{X}, \hat{\beta}_{rg} = \beta_{rg}$, we have,

$$\left.\frac{\partial \hat{t}_m^1}{\partial \bar{x}_r}\right|_{\bar{y}_r = \bar{Y}, \bar{x}_r = \bar{X}, \hat{\beta}_{rg} = \beta_{rg}} = -\left(1 - \frac{r}{n}\right)\left(\beta_{rg} + \bar{Y}\frac{\phi_j}{\phi_j \bar{X} + \phi_k}\right)$$

$$\left.\frac{\partial \hat{t}_m^1}{\partial \bar{x}_r}\right|_{\bar{y}_r = \bar{Y}, \bar{x}_r = \bar{X}, \hat{\beta}_{rg} = \beta_{rg}} = -M$$

(3.6)

where, $M = \left(1 - \frac{r}{n}\right)\left(\beta_{rg} + \bar{Y}\frac{\phi_j}{\phi_j \bar{X} + \phi_k}\right)$

By substituting (3.3) and (3.6) into (3.1), we obtain the mean square error of the estimators as

$$MSE(\hat{t}_m^1) = \begin{pmatrix} 1 & -M \end{pmatrix} \begin{pmatrix} \psi_{r,N} S_Y^2 & \psi_{r,N} \rho S_Y S_X \\ \psi_{r,N} \rho S_Y S_X & \psi_{r,N} S_X^2 \end{pmatrix} \begin{pmatrix} 1 \\ -M \end{pmatrix}$$

$$MSE(\hat{t}_m^1) = \psi_{r,N}\left(S_Y^2 - 2M\rho S_Y S_X + M^2 S_X^2\right) \quad (3.8)$$

Test for the Consistency of $\hat{t}_m^1$

Theorem 1: the estimators $\hat{t}_m^1$ are consistent.

Proof: Let f(x) and g(x) be continuous function, then

$$\lim_{x\to p}\left(f(x) \pm g(x)\right) = \lim_{x\to p} f(x) \pm \lim_{x\to p} g(x), \quad p \neq \infty \quad (3.9)$$

$$\lim_{x\to p}\left(f(x) \times g(x)\right) = \lim_{x\to p} f(x) \times \lim_{x\to p} g(x), \quad p \neq \infty \quad (3.10)$$

$$\lim_{x\to p} \frac{f(x)}{g(x)} = \frac{\lim_{x\to p} f(x)}{\lim_{x\to p} g(x)}, \quad p \neq \infty, \ \lim_{x\to p} g(x) \neq 0$$

(3.11)

As $r \to N, \ n = N$, using the results of (3.9), (3.10) and (3.11), we have

$$\lim_{r\to N}\left(\hat{t}_m^1\right) = \lim_{r\to N}\frac{r}{n}\lim_{r\to N}\left(\bar{y}_r\right) + \left(1 - \lim_{r\to N}\frac{r}{n}\right)\frac{(\lim_{r\to N}\bar{y}_r + \lim_{r\to N}\hat{\beta}_{rg}(\bar{X} - \lim_{r\to N}\bar{x}_r))}{\phi_j \lim_{r\to N}\bar{x}_r + \phi_k}(\phi_j \bar{X} + \phi_k)$$

(3.12)

Since $n = N$ if $r \to N$, then $\lim_{r\to N} \bar{y}_r = \bar{Y}$, $\lim_{r\to N} \bar{x}_r = \bar{X}$, $\lim_{r\to N} \frac{r}{n} = 1$ and $\lim_{r\to N} \hat{\beta}_{rg} = \beta_{rg}$.

Therefore,

$$\lim_{r\to N}\left(\hat{t}_m^1\right) = \bar{Y} \quad (3.7)$$

Hence, the proof.

Theoretical Efficiency Comparison

In this section, conditions for the efficiency of the new estimators over some existing related estimators were established.

Theorem 2: Estimator $\hat{t}_i^1$ is more efficient than $\hat{\mu}_0$ if (3.14) is satisfied.

$$M < 2\beta_{rg} \quad (3.14)$$

Proof: Minus (3.8) from (1.3), theorem 2 is proved.

Theorem 3: Estimator $\hat{t}_i^1$ is more efficient than $\hat{\mu}_i^{(*)}$ if (3.14) is satisfied.

$$\left(\Upsilon + M\right)\beta_{rg} + \left(M^2 - \Upsilon^2\right) < 0 \quad (3.15)$$

Proof: Subtract (3.8) from (1.6), theorem 3 is proved.

Empirical Study using Real life Data

Empirical examples on the precision of proposed estimators using real life data in Table 2 were considered in this section

Table 4: MSE of Estimators $\hat{\mu}_0, \hat{\mu}_i^{(*)}, i = 1,2,...,17, \ \hat{t}_k^1, k = 0,1,2,...,6$

Estimators	**Pop. 1**	**Pop. 2**	**Pop. 3**
Sample mean $\hat{\mu}_0$	25537.11	25537.11	217082.8
[18] Estimators			
$\hat{\mu}_1^{(*)}$	20960.84	21232.84	79355.19
$\hat{\mu}_2^{(*)}$	20953.11	21222.67	79247.78
$\hat{\mu}_3^{(*)}$	20950.37	21216.94	79205.03
$\hat{\mu}_4^{(*)}$	20959.78	21218.77	79364.27
$\hat{\mu}_5^{(*)}$	20404.57	20514.50	32126.76

$\hat{\mu}_6^{(*)}$	20946.36	21211.80	79155.26
$\hat{\mu}_7^{(*)}$	20959.38	21214.22	79367.39
$\hat{\mu}_8^{(*)}$	20386.69	20482.02	28536.62
$\hat{\mu}_9^{(*)}$	20952.94	21224.23	79252.89
$\hat{\mu}_{10}^{(*)}$	20959.76	21220.92	79363.84
$\hat{\mu}_{11}^{(*)}$	20402.49	20544.03	32845.29
$\hat{\mu}_{12}^{(*)}$	20886.54	21223.10	81083.16
$\hat{\mu}_{13}^{(*)}$	20862.36	21217.61	81790.12
$\hat{\mu}_{14}^{(*)}$	20692.22	20521.50	45609.92
$\hat{\mu}_{15}^{(*)}$	20960.79	21232.77	79355.06
$\hat{\mu}_{16}^{(*)}$	20960.77	21232.73	79355.01
$\hat{\mu}_{17}^{(*)}$	20960.83	21232.74	79355.20
Proposed Estimators			
$\hat{t}_1^1$	20386.56	20480.79	28503.63
$\hat{t}_2^1$	20386.53	20480.29	28504.04
$\hat{t}_3^1$	20386.54	20480.47	28494.95
$\hat{t}_4^1$	20386.52	20480.08	28500.00
$\hat{t}_5^1$	20386.57	20480.93	28494.42
$\hat{t}_6^1$	20386.58	20481.05	28499.07

Table 4 shows the numerical results of the MSE of estimators $\hat{\mu}_0$, $\hat{\mu}_i^{(*)},\ i=1,2,3,4,....,17$ and $\hat{t}_m^1,\ m=1,2,3,4,5,6$, using data sets in Table 2. Of all the subjects examined, the proposed two proposals have a minimum MSE for all data sets. This means that the proposed methods have shown a high level of efficiency on others considered in the study, and can produce a better estimate of the average population in the presence of an unresponsive or missing observation on average.

Empirical Study using Simulated Data

In this section, simulation studies were conducted to assess the performance of the estimators of the proposed schemes with respect to [18] estimators using study populations simulated by functions defined in Table 3 and the results are presented in table 5

Table 5: Biases and MSEs of Proposed and Other from Simulated Data

Estimators	Population I		Population II	
	Biases	MSEs	Biases	MSEs
Sample mean $\hat{\mu}_0$	0.005212425	0.08212077	-0.00238708	0.02055153
[18] Estimators				
$\hat{\mu}_1^{(*)}$	0.01881164	0.5988081	0.01224332	0.1290456
$\hat{\mu}_2^{(*)}$	0.006281752	0.1573342	0.006784724	0.05279314
$\hat{\mu}_3^{(*)}$	0.00771724	0.2157503	0.005456455	0.03706221
$\hat{\mu}_4^{(*)}$	0.01400318	0.120201	0.008066989	0.06924065
$\hat{\mu}_5^{(*)}$	0.009613986	0.2862123	0.00908595	0.08303842
$\hat{\mu}_6^{(*)}$	0.006427576	0.1636258	0.003340483	0.01579654
$\hat{\mu}_7^{(*)}$	0.007157499	0.0326743	0.005124688	0.03338117
$\hat{\mu}_8^{(*)}$	0.007912665	0.2232647	0.006160419	0.04521465
$\hat{\mu}_9^{(*)}$	0.004970697	0.0910103	0.005198207	0.03418736
$\hat{\mu}_{10}^{(*)}$	0.005403985	0.0271994	0.006411459	0.04822499
$\hat{\mu}_{11}^{(*)}$	0.006415286	0.1630998	0.007531888	0.06224325
$\hat{\mu}_{12}^{(*)}$	0.01825179	0.1904569	0.003381086	0.01614697
$\hat{\mu}_{13}^{(*)}$	0.3306025	21.33659	0.002833185	0.01166292
$\hat{\mu}_{14}^{(*)}$	0.8956953	1218.282	0.004979215	0.03180247
$\hat{\mu}_{15}^{(*)}$	0.004744487	0.0735999	0.002918198	0.01232231
$\hat{\mu}_{16}^{(*)}$	0.00497412	0.0912365	0.002534123	0.00946083
$\hat{\mu}_{17}^{(*)}$	0.004917712	0.0328937	0.003465998	0.01688858
Estimators of Proposed Scheme				
$\hat{t}_1^1$	0.002074542	0.02661696	4.807805e-06	0.001722871
$\hat{t}_2^1$	0.002074627	0.02662312	6.330344e-06	0.00172385
$\hat{t}_3^1$	0.002088068	0.02596925	-0.0003362034	0.001703694
$\hat{t}_4^1$	0.002073213	0.02609718	-0.0001708949	0.001661943
$\hat{t}_5^1$	0.00208825	0.02596969	-0.0003375063	0.001704425
$\hat{t}_6^1$	0.002073333	0.02609199	-0.0001737193	0.001661828

Table 5 shows the results of the biases and MSEs of the estimators of the proposed schemes, using the simulated data for population 1 in Table 3. The results revealed that the estimators of the proposed scheme have minimum biases and MSEs. This implies that the estimators of the proposed scheme are more efficient than Sample mean, [18] estimators and can produce a reliable estimate closer to the true population mean of the study variable.

4. Conclusion

From the results of the empirical studies, it was obtained that the proposed estimator is more efficient than other estimators considered in the study and, therefore, its use is recommended to estimate the population average when certain values of the variables of the study are missing in the study.

Conflict of Interest

The author declares that there is no conflict of interest.

Acknowledgements

The authors acknowledge the contributions of the reviewers in shaping the standard and quality of this manuscript.

References

[1] M. N. Hansen; W. N. Hurwitz (1946). The problem of non-response in sample surveys. Journal of the American Statistical Association. 41, 517–529. Available: https://doi.org/10.1080/01621459.1946.10501894.

[2] A. Audu; O. O. Ishaq; U, Isah; S. Muhammed; K. A. Akintola; A. Rashida; A. Abubakar (2020). On the Class of Exponential-Type Imputation Estimators of Population Mean with Known Population Mean of Auxiliary Variable. NIPES Journal of Science and Technology. *Research* 2(4), 1–11. https://doi.org/10.37933/nipes/2.4.2020.1

[3] A. Audu; O. O. Ishaq; J. O. Muili; Y. Zakari; A. M. Ndatsu; S. Muhammed (2020). On the Efficiency of Imputation Estimators using Auxiliary Attribute. Continental Journal of Applied Sciences. 15 (1), 1-13. https://doi.org/10.5281/zenodo.3721046

[4] A. Audu; O. O. Ishaq; Y. Zakari; D. D. Wisdom; J. O. MuiLI; A. M. Ndatsu (2020). Regression-cum-exponential ratio imputation class of estimators of population mean in the presence of non-response. Science Forum Journal of Pure and Applied Science. 20, 58-63. http://dx.doi.org/10.5455/sf.71109

[5] S. Singh; S. Horn (2000). Compromised imputation in survey sampling. Metrika. 51, 267-276. http://dx.doi.org/10.1007/s001840000054

[6] G. Diana; P. F. Perri (2010). Improved estimators of the population mean for missing data. Communications in Statistics- Theory and Methods. 39, 3245-3251.http://dx.doi.org/10.1080/03610920903009400

[7] Singh, S., (2009). A new method of imputation in survey sampling. Statistics. 43, 499-511. https://doi.org/10.1080/02331880802605114

[8] H. Toutenburg; V. K. Srivastava; A. Shalabh (2008). Imputation versus imputation of missing values through ratio method in sample surveys. Statistical Papers. 49, 237-247. https://doi.org/10.1007/s00362-006-0009-4

[9] A. I. Al-Omari; C. N. Bouza; C. Herrera (2013). Imputation methods of missing data for estimating the population mean using simple random sampling with known correlation coefficient. Quality and Quantity. 47, 353-365.https://doi.org/10.1007/s11135-011-9522-1

[10] A. K. Singh; P. Singh; V. K. Singh (2014). Exponential-type compromised imputation in survey sampling. Journal of Statistics and Application. 3 (2), 211-217. https://doi.org/10.12785/jsap/030211

[11] S. Singh; B. Deo (2003). Imputation by power transformation. Statistical Papers. 44, 555-579. https://doi.org/10.1007/BF02926010

[12] V. K. Singh; R. Singh (2014). Predictive estimation of finite population mean using generalized family of estimators. Istatistik. 7(2), 43-54.

[13] A. A. Gira (2015). Estimation of population mean with a new imputation methods. Applied Mathematical Sciences. 9(34), 1663-1672. https://doi.org/10.12988/ams.2015.5293

[14] S. Bhushan; A. P. Pandey (2016). Optimal imputation of missing data for estimation of population mean. Journal of Statistics and Management Systems. 19(6), 755-769. https://doi.org/10.1080/09720510.2016.1220099

[15] S. Prasad (2016). A study on new methods of ratio exponential type imputation in sample surveys. *Hacettepe Journal of Mathematics and Statistics*. 47(2), 1-23. https://doi.org/10.15672/HJMS.2016.392

[16] G. N. Singh; S. Maurya; M. Khetan; C. Kadilar (2016). Some imputation methods for missing data in sample surveys. Hacettepe Journal of Mathematics and Statistics. 45 (6), 1865-1880. https://doi.org/10.15672/HJMS.20159714095

[17] A. Audu; R. Singh; S. Khare (2021). New Regression-Type Compromised Imputation Class of Estimators with known Parameters of Auxiliary Variable. *Communication in Statistics-Simulation and Computation*,1-13. http://dx.doi.org/10.1080/03610918.2021.1970182

[18] A. Audu; R. V. K. Singh (2021). Exponential-type regression compromised imputation class of estimators. Journal of Statistics and Management System. 24(6), 1252-1266. http://dx.doi.org/10.1080/09720510.2020.1814501

[19] J. O. Muili; Y. Zakari*; A.* Audu (2019)**.** Modified estimator of finite population variance. *FUDMA Journal of Sciences***.** 3(4), 67-78.

[20] A. Audu; A. Danbaba; A. Abubakar; O. O. Ishaq; Y. Zakari (2020). On the efficiency of calibration ratio estimators of population mean in stratified random sampling. Proceedings of Royal Statistical Society Nigeria Local Group. 1, 247-261

[21] Y. Zakari; I. Muhammad; N. M. Sani (2020). Alternative ratio-product type estimator in simple random sampling. Communication in Physical Sciences. 5(4), 418-426.

[22] Y. Zakari; J. O. Muili; M. N. Tela; N. S. Danchadi; A. Audu (2020). Use of unknown weight to enhance ratio-type estimator in simple random sampling. Lapai Journal of Applied and Natural Sciences*.* 5(1), 74-81.

[23] I. Muhammad; Y. Zakari; A. Audu; (2021). An alternative class of ratio-regression-type estimator under two-phase sampling scheme. *CBN Journal of Applied Statistics*, 12(2),1-26. https://dx.doi.org/10.33429/Cjas.12221.1/5

[24] I. Muhammad; Y. Zakari; A. Audu (2022). Generalized estimators for finite population variance using measurable and affordable auxiliary character. Asian Research Journal of Mathematics. 18(1), 14-30. Available: https//doi.org/10.9734/ARJOM/2022/v18i130351

[25] D. Singh; F. S. Chaudhary (1986). Theory and analysis of sample survey designs. 1st edition, New Age International Publisher, India.

