Stochastic Models in Probability and Statistics
Vol. 3, No. 1, 2026, pp. 1–11. Research Paper

Some ratio type estimators for estimating the population mean in case of missing data

Ajeet Kumar Singh[†*] **and Himanshi Kumawat**[‡]

[†] *Department of Statistics, University of Rajasthan, Jaipur, India*
[‡] *Research Scholar, Department of Statistics, University of Rajasthan, Jaipur, India*

Email(s): ajeetvns.singh@gmail.com, kumawathimanshi301@gmail.com

Abstract. This paper investigates ratio-type estimators for estimating the population mean in the presence of missing data. To address the problem of missing observations, a class of generalized estimators is proposed. The bias and mean squared error of the proposed estimators are derived under optimal conditions using first-order approximations. The performance of the proposed estimators is then evaluated and compared with that of the classical mean estimator, the ratio estimator, and the compromised estimator using several empirical data sets. The results show that the proposed estimators outperform the existing estimators in terms of efficiency.

Keywords: Bias, Imputation, Mean squared error, Percent relative efficiency.

1 Introduction and preliminaries

In survey sampling, missing data may arise due to non-response, attrition, or errors in data collection. Such missingness poses a serious challenge, as it can lead to biased estimates, undermine the validity of statistical inference, and reduce the efficiency of estimators. Common approaches for handling missing data include imputation, weighting adjustments, and model-based methods. Among these approaches, ratio-type estimators have been widely studied for estimating the population mean because of their ability to efficiently incorporate auxiliary information. These estimators exploit the relationship between the study variable and an auxiliary variable, often yielding more precise estimates than those obtained under simple random sampling. However, in the presence of missing data, classical ratio estimators must be modified to account for different missingness mechanisms, namely missing completely at random (MCAR), missing at random (MAR), and missing not at random (MNAR). In this study, we focus on ratio-type estimators under missing data and investigate their performance across various missingness scenarios.

Let a finite population consist of N distinct units with associated values (x_i, y_i). The population means are defined as

$$\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} y_i \quad \text{and} \quad \bar{X} = \frac{1}{N}\sum_{i=1}^{N} x_i,$$

*Corresponding author
Received: 22 September 2025 / Accepted: 13 February 2026
DOI: 10.22067/SMPS.2025.95580.1052

where Y denotes the study variable and X is an auxiliary variable used for estimation. A sample S of size n is drawn from the finite population $P = 1, 2, \ldots, N$ using simple random sampling without replacement (SRSWOR) to estimate the population mean.

The sample S consists of r responding units ($r < n$), forming the response set R, and $n - r$ non-responding units, forming the complementary set R^c. The variable Y is the primary variable of interest, while X is an auxiliary variable correlated with Y. For each unit $i \in R$, y_i denotes the observed value of the study variable, whereas for each unit $i \in R^c$, the corresponding y_i values are missing and therefore require imputation.

Kalton et al. (1981) addressed issues of non-response and imputation in survey sampling, while Meeden (2000) proposed a theoretical framework for imputation. Lee et al. (1994) introduced a ratio method of imputation, and Singh and Horn (2000) proposed an improved technique known as the compromised method of imputation. Subsequently, Singh and Deo (2003) developed a modified estimator to handle missing data more efficiently.

A widely recognized mean imputation method, along with its sample mean and variance, can be expressed as $\bar{y}_m = \bar{y}_r$ and

$$V(\bar{y}_m) = \delta_1 C_Y^2 \bar{Y}^2, \tag{1}$$

respectively, where $\delta_1 = \frac{1}{r} - \frac{1}{N}$. It is also well recognized that, under the ratio method of imputation, the ratio estimator of the population mean is

$$\bar{y}_R = \bar{y}_r \frac{\bar{x}_n}{\bar{x}_r}.$$

The bias and mean squared error (MSE) of $\bar{y}_R$, in the presence of missing data, can be expressed as follows:

$$B(\bar{y}_R) = \delta_1 \bar{Y} \left(C_X^2 - \rho_{YX} C_Y C_X \right)$$

and

$$MSE(\bar{y}_R) = \bar{Y}^2 (\delta_1 C_Y^2 + \delta_4 (C_X^2 - 2\rho_{YX} C_Y C_X)), \tag{2}$$

respectively, $\delta_4 = \frac{1}{r} - \frac{1}{n}$, ρ_{XY} denotes the correlation coefficient between X and Y, and C_X and C_Y represent the coefficients of variation of X and Y, respectively. Singh and Horn (2000) proposed the compromised method of imputation, defined as

$$\bar{y}_{\text{comp}} = \alpha \bar{y}_r + (1 - \alpha) \bar{y}_r \frac{\bar{x}_n}{\bar{x}_r}, \tag{3}$$

where α is a constant. The value of α that minimizes the MSE is given by $\alpha = 1 - \rho_{XY} C_Y / C_X$. The corresponding bias and MSE are given by

$$B(\bar{y}_{\text{comp}}) = \delta_4 \bar{Y} \left(1 - \rho_{YX} \frac{C_Y}{C_X} \right) \rho_{YX} C_Y C_X$$

and

$$MSE(\bar{y}_{\text{comp}}) = \bar{Y}^2 (\delta_1 C_Y^2 - \delta_4 \rho_{YX}^2 C_Y^2), \tag{4}$$

respectively. Pandey et al. (2015) and Singh and Nath (2018) proposed the following estimators

$$\bar{y}_{Re} = \bar{y}_r \exp\left(\frac{\bar{x}_1 - \bar{x}_r}{\bar{x}_1 + \bar{x}_r} \right),$$

and

$$\bar{y}_{dr} = \bar{y}_r \left(\frac{n_1 \bar{x}_1 - n \bar{x}_r}{(n_1 - n) \bar{x}_r} \right)^{\beta}.$$

The bias and MSE of $\bar{y}_{Re}$ are given by

$$B(\bar{y}_{Re}) = \delta_5 \bar{Y} \left(\frac{3}{8} C_X^2 - \frac{1}{2} \rho_{YX} C_Y C_X \right),$$

$$\mathrm{MSE}(\bar{y}_{Re}) = \bar{Y}^2\left(\delta_1 C_Y^2 + \delta_5\left(\frac{C_X^2}{4} - \rho_{YX}C_YC_X\right)\right), \tag{5}$$

where $\delta_5 = \frac{1}{r} - \frac{1}{n_1}$. Similarly, for $\bar{y}_{dr}$, the bias and MSE are

$$B(\bar{y}_{dr}) = -\delta_5\frac{\bar{Y}}{2}\left(g_1C_X + \rho_{YX}C_Y\right)\rho_{YX}C_Y,$$

and

$$\mathrm{MSE}(\bar{y}_{dr}) = \bar{Y}^2\left(\delta_1 C_Y^2 - \delta_5\rho_{YX}^2C_Y^2\right), \tag{6}$$

respectively, where $g_1 = \frac{n}{n_1 - n}$.

In the remainder of this paper, new estimators are proposed in Section 2. Their statistical properties and performance in terms of MSE are investigated. Section 3 presents numerical results and graphical analysis, while concluding remarks are provided in Section 4.

2 Proposed estimators and their properties

Motivated by Bahl and Tuteja (1991) and Singh et al. (2014, 2015a,b), we propose a new class of ratio-type estimators for estimating the population mean under SRSWOR.

$$T_{1P} = \bar{y}_r\left(\frac{\bar{X}}{\bar{x}_r}\right)\exp\left[\alpha\left(\frac{\bar{X}-\bar{x}_r}{\bar{X}+\bar{x}_r}\right)\right], \tag{7}$$

$$T_{2P} = \bar{y}_r\left(\frac{\bar{X}}{\bar{x}_r}\right)^{\alpha}\exp\left[\frac{\bar{X}-\bar{x}_r}{\bar{X}+\bar{x}_r}\right], \tag{8}$$

$$T_{3P} = \bar{y}_r\left(\frac{\bar{X}}{\bar{x}_r}\right)^{\alpha}\exp\left[\beta\left(\frac{\bar{X}-\bar{x}_r}{\bar{X}+\bar{x}_r}\right)\right], \tag{9}$$

where α and β are constants chosen to minimize the MSE of the estimators T_{1P}, T_{2P}, and T_{3P}.
Using the transformations $\bar{y}_r = \bar{Y}(1+e_1)$ and $\bar{x}_r = \bar{X}(1+e_2)$, such that $|e_i| < 1, i = 1,2$. Expanding T_{1P}, T_{2P}, and T_{3P} in terms of e_1 and e_2, retaining terms up to the first order of approximation, the expressions in (7), (8) and (9) can be written as

$$T_{1P} = \bar{Y} + \bar{Y}\left(e_1 + e_2^2 - e_2\left(1+\frac{\alpha}{2}\right) - e_1e_2\left(1+\frac{\alpha}{2}\right) + \alpha\left(1+\frac{\alpha}{2}\right)\frac{e_2^2}{4} + \frac{\alpha}{2}e_2^2\right), \tag{10}$$

$$T_{2P} = \bar{Y} + \bar{Y}\left(e_1 + \frac{3}{8}e_2^2 - e_2\left(\alpha+\frac{1}{2}\right) - e_1e_2\left(\alpha+\frac{1}{2}\right) + \alpha\left(1+\frac{\alpha}{2}\right)e_2^2\right), \tag{11}$$

and

$$T_{3P} = \bar{Y} + \bar{Y}\left(e_1 - \left(\alpha+\frac{\beta}{2}\right)\left(e_2 + e_1e_2 - \frac{e_2^2}{2}\right) + \left(\alpha+\frac{\beta}{2}\right)^2\frac{e_2^2}{2}\right), \tag{12}$$

respectively. We recall that the following relations hold for e_1 and e_2:

$$E(e_1) = E(e_2) = 0,\ E(e_1^2) = \delta_1C_Y^2,\ E(e_2^2) = \delta_1C_X^2 \text{ and } E(e_1e_2) = \delta_1\rho_{YX}C_YC_X. \tag{13}$$

Using (10)–(13), we derive the bias and MSE of the estimators T_{1P}, T_{2P}, and T_{3P}. The results are presented below.

(i) Bias of proposed estimators. Using (10) and (13), we obtain the bias of T_{1P} up to the first-order approximation, which is given by

$$\begin{aligned}
Bias(T_{1P}) &= E(T_{1P}-\bar{Y}) \\
&= E\left(\bar{Y}\left[e_1+e_2^2-e_2\left(1+\frac{\alpha}{2}\right)-e_1e_2\left(1+\frac{\alpha}{2}\right)+\alpha\left(1+\frac{\alpha}{2}\right)\frac{e_2^2}{4}+\frac{\alpha}{2}e_2^2\right]\right) \\
&= \bar{Y}\left[E(e_2^2)-E(e_1e_2)\left(1+\frac{\alpha}{2}\right)+\alpha\left(1+\frac{\alpha}{2}\right)\frac{E(e_2^2)}{4}+\frac{\alpha}{2}E(e_2^2)\right] \\
&= \bar{Y}\left[\delta_1C_X^2-\delta_1\rho_{YX}C_XC_Y\left(1+\frac{\alpha}{2}\right)+\alpha\left(1+\frac{\alpha}{2}\right)\frac{\delta_1C_X^2}{4}+\frac{\alpha}{2}\delta_1C_X^2\right] \\
&= \bar{Y}\delta_1\left[(C_X^2-\rho_{YX}C_YC_X)+\frac{\alpha}{2}\left(\frac{3}{2}C_X^2-\rho_{YX}C_YC_X\right)+\frac{\alpha^2}{8}C_X^2\right].
\end{aligned}$$

Similarly, using (11), (12) and (13), we obtain the bias of T_{2P} and T_{3P} up to the first-order approximation, which are given by

$$Bias(T_{2P}) = \bar{Y}\delta_1\left(\frac{3}{8}C_X^2-\frac{1}{2}\rho_{YX}C_YC_X+\alpha(C_X^2-\rho_{YX}C_YC_X)+\frac{\alpha^2}{2}C_X^2\right)$$

and

$$Bias(T_{3P}) = \bar{Y}\delta_1\left[\left(\alpha+\frac{\beta}{2}\right)\left(\frac{C_X^2}{2}\left(\alpha+\frac{\beta}{2}+1\right)-\rho_{YX}C_YC_X\right)\right],$$

respectively.

(ii) MSE of proposed estimators. Using (10) and (13), we obtain the MSE of T_{1P} up to the first-order approximation, which is given by

$$\begin{aligned}
MSE(T_{1P};\alpha) &= E\left(T_{1P}-\bar{Y}\right)^2 \\
&= E\left(\bar{Y}^2\left[e_1+e_2^2-e_2\left(1+\frac{\alpha}{2}\right)-e_1e_2\left(1+\frac{\alpha}{2}\right)+\alpha\left(1+\frac{\alpha}{2}\right)\frac{e_2^2}{4}+\frac{\alpha}{2}e_2^2\right]^2\right) \\
&= E\left(\bar{Y}^2\left[e_1^2+e_2^2+\frac{\alpha^2}{4}-2e_1e_2+\alpha e_2^2-\alpha e_1e_2\right]\right) \\
&= \bar{Y}^2\delta_1\left(C_Y^2+C_X^2-\rho_{YX}C_YC_X+\alpha\left(C_X^2-\rho_{YX}C_YC_X\right)+\frac{\alpha^2}{4}C_X^2\right).
\end{aligned} \tag{14}$$

Similarly, using (11), (12) and (13), we obtain the MSE of T_{2P} and T_{3P} up to the first-order approximation, which are given by

$$MSE\left(T_{2P};\alpha\right) = \bar{Y}^2\delta_1\left(C_Y^2+\frac{C_X^2}{4}-\rho_{YX}C_YC_X+\alpha\left(C_X^2-2\rho_{YX}C_YC_X\right)+\alpha^2C_X^2\right), \tag{15}$$

and

$$MSE\left(T_{3P};\alpha,\beta\right) = \bar{Y}^2\delta_1\left(C_Y^2+\left(\alpha+\frac{\beta}{2}\right)^2C_X^2-2\left(\alpha+\frac{\beta}{2}\right)\rho_{YX}C_YC_X\right), \tag{16}$$

respectively.
The MSEs in (14), (15) and (16) depend on the parameters α and β. Since these expressions are differentiable with respect to α and β, the values of α and β that minimize the MSEs can be obtained

by setting the corresponding partial derivatives equal to zero.
From (14), we obtain

$$\frac{\partial}{\partial\alpha}\mathrm{MSE}(T_{1P};\alpha)=\bar{Y}^2\delta_1\left(C_X^2-\rho_{YX}C_YC_X+\frac{\alpha}{2}C_X^2\right).$$

Setting the above expression equal to zero yields

$$\alpha_{\mathrm{op}}=2\rho_{YX}\frac{C_Y}{C_X}-2.$$

Consequently,

$$\mathrm{MSE}(T_{1P};\alpha_{\mathrm{op}})=\bar{Y}^2\delta_1C_Y^2\left(1-\rho_{YX}^2\right). \tag{17}$$

Also, from (15), we have

$$\frac{\partial}{\partial\alpha}\mathrm{MSE}(T_{2P};\alpha)=\bar{Y}^2\delta_1\left(C_X^2-2\rho_{YX}C_YC_X+2\alpha C_X^2\right).$$

Equating this expression to zero gives $\alpha_{\mathrm{op}}=\rho_{YX}\frac{C_Y}{C_X}-1/2$, and hence

$$\mathrm{MSE}(T_{2P};\alpha_{\mathrm{op}})=\bar{Y}^2\delta_1C_Y^2\left(1-\rho_{YX}^2\right). \tag{18}$$

Similarly, from (16), to obtain the optimum condition of the proposed estimator T_{3p}, let $k=\alpha+\beta/2$. Then, we have

$$MSE(T_{3P};k)=\bar{Y}^2\delta_1\left(C_Y^2+k^2C_X^2-2k\rho_{YX}C_YC_X\right). \tag{19}$$

Differentiating (19) with respect to k and equating to zero, we get $k_{op}=\rho_{YX}C_Y/C_X$. Hence, the optimum condition is

$$\alpha+\beta/2=\rho_{YX}C_Y/C_X.$$

Thus, in this case, the minimum MSE is given by

$$\mathrm{MSE}\big(T_{3P};\alpha_{\mathrm{op}},\beta_{\mathrm{op}}\big)=\bar{Y}^2\delta_1C_Y^2\left(1-\rho_{YX}^2\right). \tag{20}$$

Remark 1. *As observed from (17), (18), and (20), the MSEs of the proposed estimators (7), (8), and (9) are identical when evaluated at the optimal values α_{op} and β_{op}. Henceforth, we denote the resulting optimal estimator by T_P, for example by $\alpha+\beta/2=\rho_{YX}C_Y/C_X$ the estimator in (12) reduced to*

$$T_P=\bar{Y}+\bar{Y}\left(e_1-(\rho_{YX}C_Y/C_X)\left(e_2+e_1e_2-e_2^2/2\right)+(\rho_{YX}C_Y/C_X)^2e_2^2/2\right),$$

and its MSE is given by

$$\mathrm{MSE}(T_P)=\bar{Y}^2\delta_1C_Y^2\left(1-\rho_{YX}^2\right).$$

The MSEs of all estimators are summarized in Table 1. From Table 1, the relative performance of the estimators can be readily compared in terms of their MSE.

3 Numerical results and graphical analysis

To examine the efficiency of the proposed estimators, seven finite population data sets (from Cochran (1977), Murthy (1967), Anderson (2003), Shukla et al. (2011), Mukhopadhyay (2000), Kadilar and Cingi (2008) and Diana and Francesco (2010)) are considered. The corresponding descriptive data N, n, r, $\bar{Y}$, $\bar{X}$, C_{Y}, C_{X} and ρ_{YX} are presented in Table 2.

Table 1: Performance of T_P relative to other estimators in terms of MSE.

Estimator	MSE	T_p is more efficient if
$\bar{y}_m$	(1)	$N-r>0$
$\bar{y}_R$	(2)	$\delta_1\rho_{YX}^2C_Y^2+\delta_4\left(C_X^2-2\rho_{YX}C_YC_X\right)>0$
$\bar{y}_{comp}$	(4)	$2nN-r(N-n)>0$
$\bar{y}_{Re}$	(5)	$\left(\delta_1\rho_{YX}^2C_Y^2+\delta_5\left(\frac{1}{4}C_X^2-\rho_{YX}C_YC_X\right)\right)>0$
$\bar{y}_{dr}$	(6)	$N-n_1>0$

Table 2: Descriptive data of seven populations.

Population	N	n	r	$\bar{Y}$	$\bar{X}$	$C^2{}_{\mathrm{Y}}$	$C^2{}_{\mathrm{X}}$	ρ_{YX}
I Cochran (1977)	34	11	9	4.92	2.59	1.0248	1.5175	0.7326
II Murthy (1967)	34	10	7	199.44	208.89	0.5673	0.5191	0.9801
III Anderson (2003)	25	9	7	183.84	185.72	0.0029	0.0027	0.7108
IV Shukla et al. (2011)	200	50	45	42.485	18.515	0.1102	0.1415	0.8652
V Mukhopadhyay (2000)	20	7	5	41.5	441.95	0.0555	0.0522	0.6521
VI Kadilar and Cingi (2008)	19	7	5	575	13573.68	2.2290	0.9101	0.88
VII Diana and Francesco (2010)	8011	400	250	28229.43	1.69	0.6193	0.2130	0.46

Table 3: Bias values of the estimators.

Pop.	$\bar{y}_m$	$\bar{y}_R$	$\bar{y}_{COMP}$	$\bar{y}_{Re}$	$\bar{y}_{dr}$	T_P
I	0	0.24275	0.03614	0.0362	-0.2363	0.0854
II	0	-0.2888	-0.1118	-1.3845	-9.5980	-0.1246
III	0	0.0136	0.0031	0.0002	-0.0340	0.0050
IV	0	0.0244	0.0024	-0.0005	-0.0481	0.0093
V	0	0.1066	0.0273	0.0092	-0.1927	0.0358
VI	0	-29.089	-15.534	-17.903	-122.934	-20.031
VII	0	5.0246	1.5255	-0.3899	-7.8680	1.9706

Numerical results for the bias, MSE, and percent relative efficiency (PRE) of the competing estimators are reported and discussed. In what follows, $\mathrm{PRE}(\hat{\theta}_1,\hat{\theta}_2)$ denotes the percent relative efficiency of $\hat{\theta}_2$ with respect to $\hat{\theta}_1$, defined as

$$\mathrm{PRE}(\hat{\theta}_1,\hat{\theta}_2)=100\,\frac{\mathrm{MSE}(\hat{\theta}_1)}{\mathrm{MSE}(\hat{\theta}_2)}.$$

Table 4: MSE values of the estimators.

Pop.	$\bar{y}_m$	$\bar{y}_R$	$\bar{y}_{COMP}$	$\bar{y}_{Re}$	$\bar{y}_{dr}$	T_P
I	2.02669	1.87525	1.75773	1.1776	1.1525	0.9389
II	2559.907	1631.473	1630.937	1002.079	448.613	100.8705
III	10.3697	8.9553	8.7527	7.1110	6.7314	5.1305
IV	3.4282	3.1288	3.0970	1.7046	1.4715	0.8619
V	14.3605	12.5869	12.0342	10.0211	9.91936	8.2539
VI	108607.5	78441.43	75995.09	71606.06	46347.47	24501.85
VII	1912690	1767868	1756029	1570428	1518632	1507965

The numerical results in Table 4 indicate that the proposed imputation strategy yields a lower MSE compared to the existing imputation strategies.

The MSE comparisons of the estimators $\bar{y}_m$, $\bar{y}_R$, $\bar{y}_{comp}$, $\bar{y}_{Re}$, $\bar{y}_{dr}$ and T_P for different non-response rates are presented in Table 5.

Table 5: MSE values of estimators for different non-response rates.

Rate	Pop.	$\bar{y}_m$	$\bar{y}_R$	$\bar{y}_{COMP}$	$\bar{y}_{Re}$	$\bar{y}_{dr}$	T_P
5%	I	1.7511	1.6829	1.6300	1.0457	1.0248	0.8113
	II	1592.831	1592.831	1592.831	720.4471	410.5064	62.7639
	III	7.1692	7.1692	7.1692	5.3588	5.1479	3.5471
	IV	3.1518	3.0395	3.0276	1.6105	1.4020	0.7924
	V	8.8898	8.8898	8.8898	6.9139	6.7750	5.1096
	VI	66494.38	66494.38	66494.38	48874.65	36846.75	15001.13
	VII	1237271	1224567	1223529	1019142	986131.3	975464.2
10%	I	1.7511	1.6829	1.6300	1.0457	1.0248	0.8113
	II	1843.554	1602.849	1602.710	793.4627	420.3859	72.6434
	III	8.5694	7.9506	7.8620	6.1254	5.8407	4.2398
	IV	3.4282	3.1288	3.0971	1.7047	1.4715	0.8620
	V	11.1693	10.4303	10.1999	8.2877	8.0852	6.4197
	VI	84041.50	71472.31	70453.01	58346.07	40805.38	18959.76
	VII	1309431	1282612	1280420	1078040	1043022	1032355
15%	I	2.0267	1.8753	1.7577	1.1777	1.1525	0.9390
	II	2156.958	1615.372	1615.06	884.7321	432.7352	84.9928
	III	8.5694	7.9506	7.8620	6.1254	5.8407	4.2398
	IV	3.7442	3.2309	3.1765	1.8123	1.5510	0.9414
	V	11.1693	10.4303	10.1999	8.2877	8.0852	6.4197
	VI	84041.5	71472.31	70453.01	58346.07	40805.38	18959.76
	VII	1390080	1347486	1344004	1143867	1106606	1095939

1. The PRE results presented in Tables 6, 7, and 8 clearly demonstrate that the proposed estimator T_p consistently outperforms the sample mean, ratio, and compromised estimators across all populations considered.

2. The uniformly higher PRE values confirm that the proposed estimator achieves a substantial reduction in MSE and maintains strong performance across diverse population structures.

3. For all populations, the MSE increases as the non-response rate rises from 5% to 15% for each

Table 6: PRE($\bar{y}_m$, .) values of estimator.

Pop.	$\bar{y}_R$	$\bar{y}_{COMP}$	$\bar{y}_{Re}$	$\bar{y}_{dr}$	T_P
I	108.0756	115.3019	172.0946	175.8439	215.8442
II	156.9077	156.9592	255.4597	570.6269	2537.8141
III	115.7939	118.4746	145.8264	154.0497	202.1168
IV	109.5687	110.6917	201.1054	232.9688	397.7267
V	114.0903	119.3309	140.6372	144.7726	173.984
VI	138.4568	142.9138	151.6736	234.3332	443.2624
VII	108.1919	108.9213	121.7942	125.9482	126.8392

Table 7: PRE($\bar{y}_R$, .) values of estimator.

Pop.	$\bar{y}_{comp}$	$\bar{y}_{Re}$	$\bar{y}_{dr}$	T_P
I	106.6864	159.2354	162.7045	199.7159
II	100.0328	162.8088	363.6704	1617.3927
III	102.3151	125.9362	133.0378	174.5488
IV	101.0249	183.5428	212.6236	362.993
V	104.5933	123.2683	126.8929	152.4967
VI	103.2191	109.5458	169.2464	320.145
VII	100.6742	112.5723	116.4119	117.2353

Table 8: PRE($\bar{y}_{comp}$, .) values of estimator.

Pop.	$\bar{y}_{Re}$	$\bar{y}_{dr}$	T_P
I	149.2556	152.5073	187.1991
II	162.7554	363.5511	1616.862
III	123.0867	130.0276	170.5993
IV	181.6807	210.4664	359.3103
V	117.8548	121.3203	145.7996
VI	106.1294	163.9682	310.1607
VII	111.8185	115.6323	116.4503

estimator under consideration. This pattern is expected, since higher non-response rates reduce the effective sample size and lead to increased variability.

4. However, the magnitude of the MSE increase differs notably among the estimators, highlighting variations in their robustness to non-response.

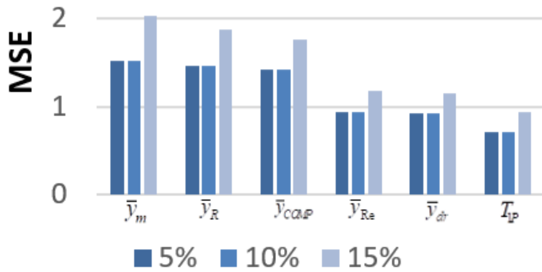

Figure 1: MSE comparison to non-response rate for population I.

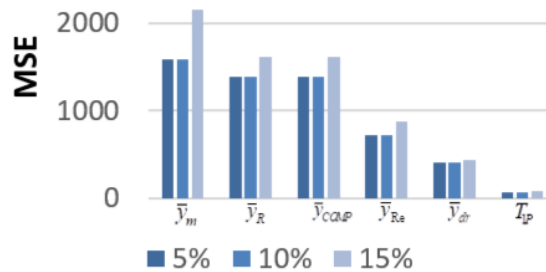

Figure 2: MSE comparison to non-response rate for population II.

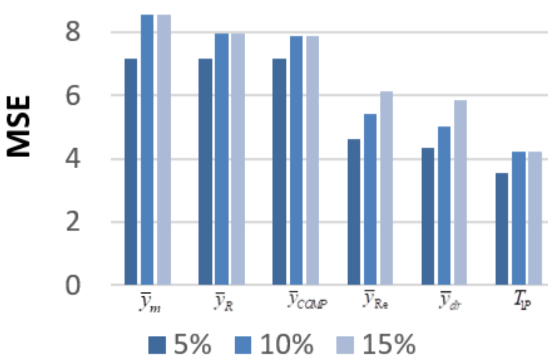

Figure 3: MSE comparison to non-response rate for population III.

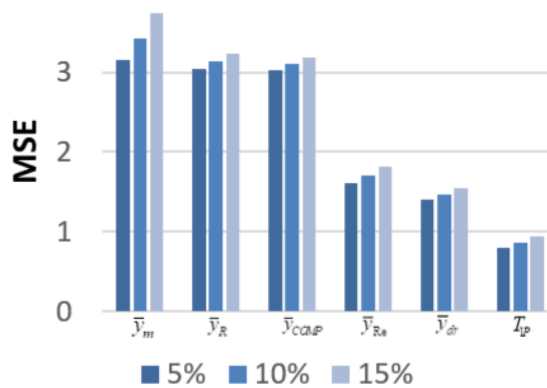

Figure 4: MSE comparison to non-response rate for population IV.

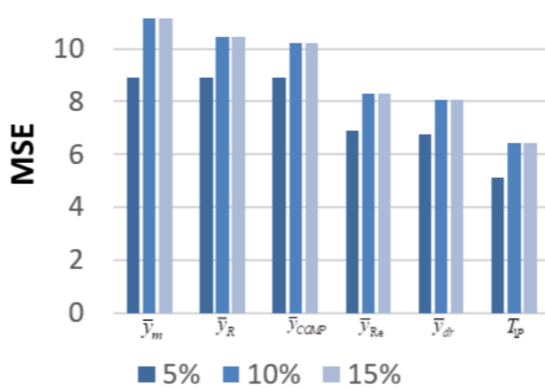

Figure 5: MSE comparison to non-response rate for population V.

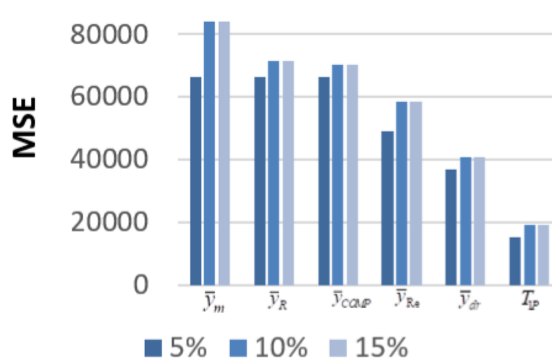

Figure 6: MSE comparison to non-response rate for population VI.

1. The numerical results and graphical analyses jointly confirm that the proposed estimator T_p consistently outperforms the mean, ratio, and compromised estimators in terms of efficiency.

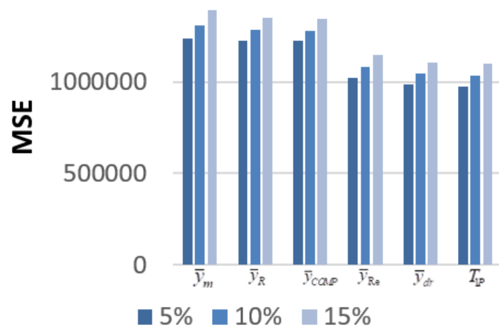

Figure 7: MSE comparison to non-response rate for population VII.

2. The estimator T_p not only achieves the minimum theoretical mean squared error (MSE) but also demonstrates superior empirical performance across varying non-response rates.

3. These findings provide strong evidence of the practical applicability of the proposed estimator in real-world survey settings, where non-response is inevitable.

4 Conclusions

This study proposes a generalized class of ratio-type estimators for estimating the population mean by effectively incorporating auxiliary information. The proposed class includes several existing estimators as special cases and allows for the determination of optimal parameter values based on the minimum MSE criterion. Under these optimal conditions, the proposed estimators attain identical MSEs.

Theoretical analysis establishes that the proposed estimator achieves a lower MSE than the traditional sample mean, ratio, and compromised estimators. These findings are further supported by an empirical investigation based on PRE measures, which consistently demonstrates the superior efficiency and robustness of the proposed estimator across different populations and varying non-response rates.

Overall, the results confirm the practical usefulness of the proposed methodology in survey sampling applications where auxiliary information is available and non-response is unavoidable. The analysis in this paper is restricted to first-order approximations under simple random sampling. Future research may extend the proposed framework to more complex sampling designs, such as stratified and multi-phase sampling, and examine its performance in the presence of non-response and measurement errors. In addition, the methodology may be adapted for estimating other population parameters, including the median, variance, proportion, and related characteristics.

5 Acknowledgements

We express our gratitude to the editors and reviewers for their thorough evaluation of the manuscript and for their valuable comments and suggestions.

References

Anderson, T. W. (2003). *An Introduction to the Multivariate Statistical Analysis.* 3rd Edition, Wiley, New York.

Bahl, S. Tuteja, R. K. (1991). Ratio and product type exponential estimator. *Information and Optimization Sciences*, Vol, XII, I, 159–163.

Cochran, W. (1977). *Sampling Techniques*. 3rd Edition. John Wiley and Sons, New York.

Diana, G. and Francesco, P. (2010). Improved estimators of the population mean for missing data. *Communications in Statistics-Theory and Methods*, **39** 3245–3251.

Kadilar, C. and Cingi, H. (2008). Estimators for the population mean in the case of missing data. *Communications in Statistics-Theory and Methods*, **37**, 2226–2236.

Kalton, G., Kasprzyk, D., and Santos, R. (1981). Issues of nonresponse and imputation in the survey of income and program participation. *In Current topics in survey sampling* (pp. 455-480). Academic Press.

Lee, H., Rancourt, E. and Sarndal, C. (1994). Experiments with variance estimation from survey data with imputed values. *Journal of Official Statistics*, **10**, 231–243.

Meeden, G. (2000). A decision theoretic approach to imputation in finite population sampling. *Journal of the American Statistical Association*, **95**, 586–595.

Mukhopadhyay, P. (2000). *Theory and Methods of Survey Sampling*, Prentice-Hall of India, New Delhi, India.

Murthy, M. N. (1967). *Sampling Theory and Methods*. 2nd Edition, Statistical Publishing Society, Calcutta.

Pandey, R., Thakur, N. S. and Yadav, K. (2015). Estimation of population mean using exponential ratio type imputation method under survey non-response. *Journal of the Indian Statistical Association*, **53**, 89–107.

Shukla, D., Thakur, N. S. and Thakur, D.S. (2011). Linear combination-based imputation method for missing data in sample. *International Journal of Modern Engineering Research*, **1**, 580–596.

Singh, B. K., and Nath, K. (2018). Some imputation methods in two-phase sampling scheme for estimation of population mean. *Research Reviews: Journal of Statistics (RRJoST)*, **7**, 1–16.

Singh, S. and Deo, B. (2003). Imputation by power transformation. *Statistical Papers*, **44**, 555–579.

Singh, S. and Horn, S. (2000). Compromised imputation in survey sampling. *Metrika*, **51**, 267–276.

Singh, A. K., Singh, P. and Singh, V. K. (2014). Exponential-type compromised imputation in survey sampling, *Journal of the Statistics Applications and Probability*, **3**, 211–217.

Singh, Priyanka, Singh, A. K. and Singh, V. K. (2015a). On the use of compromised imputation for missing data using factor-type estimators. *Journal of Statistics Applications and Probability Letters* 2(2), 1-9.

Singh, Priyanka, Singh, A. K. and Singh, V. K. (2015b). Estimation of population mean under non-response using various imputation methods for stratified population. *Journal of Advance Computing*, **4**, 112–122.