REVSTAT – Statistical Journal
Volume 23, Number 1, January 2025, 53–77
https://doi.org/10.57805/revstat.v23i1.501

# Optimal Imputation Methods under Stratified Ranked Set Sampling

Authors: Shashi Bhushan
– Department of Statistics, University of Lucknow, Lucknow, India, 226007
shashi.py@gmail.com

Anoop Kumar ✉
– Department of Statistics, Central University of Haryana, Mahendergarh, India, 123031
anoop.asy@gmail.com

Received: September 2022 Revised: January 2023 Accepted: January 2023

Abstract:

- It is long familiar that the stratified ranked set sampling ($SRSS$) is more efficient than ranked set sampling ($RSS$) and stratified random sampling ($StRS$). The existence of missing values may alter the final inference of any study. This paper is a fundamental effort to suggest some combined and separate imputation methods in the presence of missing data under $SRSS$. The proposed imputation methods become superior than the mean imputation method, ratio imputation method, Diana and Perri (2010) type imputation method, and Sohail *et al.* (2018) type imputation methods. A simulation study is administered over two hypothetically drawn asymmetric populations.

Keywords:

- *missing values; imputation; stratified ranked set sampling.*

AMS Subject Classification:

- 62D05.

✉ Corresponding author

## 1. INTRODUCTION

The dilemma of missing value is very usual in a sample survey and its presence can spoil the traditional results. Therefore, it becomes essential to resolve the problem of missing values in a data set. The well-known imputation technique is used to replace the missing values. Three basic concepts on missing values were suggested by Rubin (1976), such as missing at random ($MAR$), observed at random ($OAR$), and parameter distribution ($PD$). Several renowned authors like Lee *et al.* (1994), Singh and Horn (2000), Singh and Deo (2003), Singh (2009), and Singh and Valdes (2009) introduced various imputation methods in the presence of missing values. Heitjan and Basu (1996) exhibited a difference between missing at random and missing completely at random ($MCAR$) approach. Thereafter, Ahmed *et al.* (2006), Kadilar and Cingi (2008), Diana and Perri (2010) and Bhushan and Pandey (2016, 2018), Mohamed *et al.* (2016), Prasad (2017), Bouza *et al.* (2020), Bouza-Herrera and Viada (2021), and Bhushan *et al.* (2018, 2022) utilized $MCAR$ strategy in their study for the imputation of missing values.

In real life, situations may emerge where it is either difficult to measure the study variable or indeed expensive but can be ranked either visually or by any cost free method. In such circumstances, McIntyre (1952) proposed the idea of ranked set sampling ($RSS$), which is superior to simple random sampling but did not furnish any mathematical support. Takahasi and Wakimoto (1968) extended the idea of McIntyre (1952) and provided the obligatory mathematical foundation to the theory of $RSS$. Samawi (1996) envisaged the idea of $SRSS$ superior to $StRS$. Samawi and Siam (2003) introduced combined and separate ratio estimators under $SRSS$. Mandowara and Mehta (2014) considered modified ratio estimators under $SRSS$. Linder *et al.* (2015) investigated the regression estimator under $SRSS$. Khan and Shabbir (2016) suggested Hartley-Ross type unbiased estimators under $RSS$ and $SRSS$. Recently, Saini and Kumar (2018) suggested the ratio estimator using quartile as an auxiliary information under $SRSS$.

In sample surveys, when each group contains very small observations, then each observation becomes essential to draw conclusions. Further use of such kind of data set consisting of missing values may vitiate the final conclusion and decrease the efficiency of the estimator as well. In order to tackle with such kind of problems, Bouza-Herrera and Al-Omari (2011) suggested mean imputation and ratio methods for the median estimator under $RSS$. Al-Omari and Bouza (2014) introduced ratio estimators of the population mean with missing values under $RSS$. Sohail *et al.* (2018) suggested ratio type imputation methods under $RSS$.

In this paper, we suggest some imputation methods in the presence of missing data under $SRSS$. The rest paper is arranged in subsequent sections. In the next section, we discuss the sampling methodology along with the notations used throughout the manuscript. In Section 3, the combined and separate imputation methods are reviewed. In Section 4, we have suggested combined and separate classes of imputation methods. The theoretical comparisons of combined and separate imputation methods are given in Section 5, whereas Section 6 deals with the simulation study conducted in favour of theoretical findings. Lastly, the conclusion is given in Section 7.

## 2. METHODOLOGY AND NOTATIONS

The procedure of ranked set sampling consists of drawing $m$ simple random samples of size $m$ from the population. These $m$ units are now ranked within each set with respect to the variable of interest, say $x$. The first smallest unit is quantified from the first set for the measurement of the auxiliary variable along with the associated study variables. The unit with the second smallest rank is quantified from the second ranked set for the measurement of the auxiliary variables along with the associated study variable and the process is carried on as far as the $m$-th smallest unit is quantified from the last set. The above process is known as a cycle. The repetition of this whole procedure up to $k$ times furnishes $n = mk$ ranked set samples.

The stratified ranked set sampling is a sampling procedure analogous to stratified random sampling, which is based on splitting a population into $L$ mutually exclusive and exhaustive strata and a ranked set sample of $n_h = m_h k$ units are measured within each stratum such that $h = 1, 2, ..., L$. The sampling is accomplished independently across the strata. Thus, $SRSS$ scheme can be supposed to a collection of $L$ separate ranked set samples.

Consider a finite population $U$ comprised of $N$ measurable units with values $y_i$, $i \in U$. Let a stratified ranked set sample of size $n = m_h k$ be chosen from $U$ to estimate the population mean of the study variable $y$. Let $r$ be the number of responding elements out of $n$ sampled elements. Let $P$ be the probability that $i$-th respondent associated with a responding class $A$ and $(1 - P)$ be the probability that $i$-th respondent associated with the non-responding class $\bar{A}$. Moreover, note that $s = A \cup \bar{A}$ and let the values $y_i$, $i \in A$ be observable for each characteristic, but for the characteristic $i \in \bar{A}$ the values are missing and require imputation in order to establish the complete frame of data to draw a reasonable inference. The auxiliary variable $x$ will be used to execute the imputation of missing values and let the ranking be performed over the auxiliary variables as well.

The succeeding notations would be used from the beginning to end in the case of combined estimators.

Let $\bar{y}_{r,srss} = \bar{Y}(1 + \epsilon_0)$, $\bar{x}_{r,srss} = \bar{X}(1 + \epsilon_1)$, $\bar{x}_{n,srss} = \bar{X}(1 + \epsilon_2)$ such that $E(\epsilon_0) = E(\epsilon_1) = E(\epsilon_2) = 0$ and

$$E(\epsilon_0^2) = \sum_{h=1}^{L} W_h^2 \left( \frac{C_{y_h}^2}{m_h k P} - \frac{1}{m_h^2 k P} \sum_{i=1}^{m_h} \frac{\tau_{y_h}^2}{\bar{Y}^2} \right) = \sum_{h=1}^{L} W_h^2 \left( \gamma^* C_{y_h}^2 - D_{y_h}^{2^*} \right) = I_0^*,$$

$$E(\epsilon_1^2) = \sum_{h=1}^{L} W_h^2 \left( \frac{C_{x_h}^2}{m_h k P} - \frac{1}{m_h^2 k P} \sum_{i=1}^{m_h} \frac{\tau_{x_h}^2}{\bar{X}^2} \right) = \sum_{h=1}^{L} W_h^2 \left( \gamma^* C_{x_h}^2 - D_{x_h}^{2^*} \right) = I_1^*,$$

$$E(\epsilon_2^2) = \sum_{h=1}^{L} W_h^2 \left( \frac{C_{x_h}^2}{m_h k} - \frac{1}{m_h^2 k} \sum_{i=1}^{m_h} \frac{\tau_{x_h}^2}{\bar{X}^2} \right) = \sum_{h=1}^{L} W_h^2 \left( \gamma C_{x_h}^2 - D_{x_h}^2 \right) = I_1,$$

$$E(\epsilon_0, \epsilon_1) = \sum_{h=1}^{L} W_h^2 \left( \frac{\rho_{x_h y_h} C_{x_h} C_{y_h}}{m_h k P} - \frac{1}{m_h^2 k P} \sum_{i=1}^{m_h} \frac{\tau_{x_h y_h}}{\bar{X}\bar{Y}} \right)$$
$$= \sum_{h=1}^{L} W_h^2 \left( \gamma^* \rho_{x_h y_h} C_{x_h} C_{y_h} - D_{x_h y_h}^* \right) = I_{01}^*,$$

$$
\begin{aligned}
E(\epsilon_0,\epsilon_2) &= \sum_{h=1}^{L} W_h^2\left(\frac{\rho_{x_h y_h} C_{x_h} C_{y_h}}{m_h k} - \frac{1}{m_h^2 k}\sum_{i=1}^{m_h}\frac{\tau_{x_h y_h}}{\bar{X}\bar{Y}}\right)\\
&= \sum_{h=1}^{L} W_h^2(\gamma\rho_{x_h y_h} C_{x_h} C_{y_h} - D_{x_h y_h}) = I_{01},\\
E(\epsilon_1,\epsilon_2) &= \sum_{h=1}^{L} W_h^2\left(\frac{C_{x_h}^2}{m_h k} - \frac{1}{m_h^2 k}\sum_{i=1}^{m_h}\frac{\tau_{x_h}^2}{\bar{X}^2}\right) = \sum_{h=1}^{L} W_h^2\left(\gamma C_{x_h}^2 - D_{x_h}^2\right) = I_1,
\end{aligned}
$$

where $\gamma^* = 1/m_h kP$, $\gamma = 1/m_h k$, $\tau_{y_h} = (\mu_{y_h} - \bar{Y}_h)$, $\tau_{x_h} = (\mu_{x_h} - \bar{X}_h)$ and $\tau_{x_h y_h} = (\mu_{x_h} - \bar{X}_h) \cdot$ $\cdot\, (\mu_{y_h} - \bar{Y}_h)$. Also, $C_{x_h} = S_{x_h}/\bar{X}$ and $C_{y_h} = S_{y_h}/\bar{Y}$ are the coefficients of variation of auxiliary variable $x$ and study variable $y$, respectively.

In the case of separate estimators, the following notations will be used throughout the paper.

Let $\bar{y}_{r,h[rss]} = \bar{Y}_h(1+e_{0_h})$, $\bar{x}_{r,h(rss)} = \bar{X}_h(1+e_{1_h})$, $\bar{x}_{n,h(rss)} = \bar{X}_h(1+e_{2_h})$ such that $E(e_{0_h}) = E(e_{1_h}) = E(e_{2_h}) = 0$ and

$$
\begin{aligned}
E(e_{0_h}^2) &= \left(\frac{C_{y_h}^2}{m_h kP} - \frac{1}{m_h^2 kP}\sum_{i=1}^{m_h}\frac{\tau_{y_h}^2}{\bar{Y}_h^2}\right) = \left(\gamma^* C_{y_h}^2 - M_{y_h}^{2^*}\right) = J_0^*,\\
E(e_{1_h}^2) &= \left(\frac{C_{x_h}^2}{m_h kP} - \frac{1}{m_h^2 kP}\sum_{i=1}^{m_h}\frac{\tau_{x_h}^2}{\bar{X}_h^2}\right) = \left(\gamma^* C_{x_h}^2 - M_{x_h}^{2^*}\right) = J_1^*,\\
E(e_{2_h}^2) &= \left(\frac{C_{x_h}^2}{m_h k} - \frac{1}{m_h^2 k}\sum_{i=1}^{m_h}\frac{\tau_{x_h}^2}{\bar{X}_h^2}\right) = \left(\gamma C_{x_h}^2 - M_{x_h}^2\right) = J_1,\\
E(e_{0_h}, e_{1_h}) &= \left(\frac{\rho_{x_h y_h} C_{x_h} C_{y_h}}{m_h kP} - \frac{1}{m_h^2 kP}\sum_{i=1}^{m_h}\frac{\tau_{x_h y_h}}{\bar{X}_h\bar{Y}_h}\right) = \left(\gamma^*\rho_{x_h y_h} C_{x_h} C_{y_h} - M_{x_h y_h}^*\right) = J_{01}^*,\\
E(e_{0_h}, e_{2_h}) &= \left(\frac{\rho_{x_h y_h} C_{x_h} C_{y_h}}{m_h k} - \frac{1}{m_h^2 k}\sum_{i=1}^{m_h}\frac{\tau_{x_h y_h}}{\bar{X}_h\bar{Y}_h}\right) = (\gamma\rho_{x_h y_h} C_{x_h} C_{y_h} - M_{x_h y_h}) = J_{01},\\
E(\epsilon_{1_h}, \epsilon_{2_h}) &= \left(\frac{C_{x_h}^2}{m_h k} - \frac{1}{m_h^2 k}\sum_{i=1}^{m_h}\frac{\tau_{x_h}^2}{\bar{X}_h^2}\right) = \left(\gamma C_{x_h}^2 - M_{x_h}^2\right) = J_1,
\end{aligned}
$$

where $\tau_{y_h} = (\mu_{y_h} - \bar{Y}_h)$, $\tau_{x_h} = (\mu_{x_h} - \bar{X}_h)$ and $\tau_{x_h y_h} = (\mu_{x_h} - \bar{X}_h)(\mu_{y_h} - \bar{Y}_h)$, $C_{x_h} = S_{x_h}/\bar{X}_h$ and $C_{y_h} = S_{y_h}/\bar{Y}_h$.

## 3. RECAP OF IMPUTATION METHODS

In this section, we consider a concise recap of existing prominent combined and separate imputation methods under $SRSS$.

### 3.1. Combined imputation methods

The mean method of imputation under $SRSS$ is given by

$$y^c_{.i_m} = \begin{cases} y_i & \text{for } i \in A, \\ \bar{y}_{r,srss} & \text{for } i \in \bar{A}. \end{cases}$$

The sequent estimator is given by

$$T^c_m = \bar{y}_{r,srss}$$

where $\bar{y}_{r,srss} = \sum_{h=1}^{L} W_h \bar{y}_{h[rss]}$ is the stratified ranked set sample mean of study variable $y$. Also, $W_h = N_h/N$ is the weight of stratum $h$ and $N_h$ and $N$ are the size of stratum $h$ and total population size, respectively.

The imputation methods are categorized into three situations under the availability of auxiliary informations:

*Situation I*: When $\bar{X}$ is known and $\bar{x}_{n,srss}$ is utilized.

*Situation II*: When $\bar{X}$ is known and $\bar{x}_{r,srss}$ is utilized.

*Situation III*: When $\bar{X}$ is unknown and $\bar{x}_{n,srss}$, $\bar{x}_{r,srss}$ are utilized.

The classical combined ratio type imputation methods are defined under $SRSS$ as:

*Situation I*

$$y^c_{.i_{R_1}} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{1}{n-r}\left[n\bar{y}_{r,rss}\left(\frac{\bar{X}}{\bar{x}_{n,srss}}\right) - r\bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}. \end{cases}$$

*Situation II*

$$y^c_{.i_{R_2}} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{1}{n-r}\left[n\bar{y}_{r,rss}\left(\frac{\bar{X}}{\bar{x}_{r,srss}}\right) - r\bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}. \end{cases}$$

*Situation III*

$$y^c_{.i_{R_3}} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{1}{n-r}\left[n\bar{y}_{r,rss}\left(\frac{\bar{x}_{n,srss}}{\bar{x}_{r,srss}}\right) - r\bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}. \end{cases}$$

The sequent estimators are

$$T^c_{R_1} = \bar{y}_{r,srss}\left(\frac{\bar{X}}{\bar{x}_{n,srss}}\right),$$

$$T^c_{R_2} = \bar{y}_{r,srss}\left(\frac{\bar{X}}{\bar{x}_{r,srss}}\right),$$

$$T^c_{R_3} = \bar{y}_{r,srss}\left(\frac{\bar{x}_{n,srss}}{\bar{x}_{r,srss}}\right),$$

where $\bar{x}_{n,srss} = \sum_{h=1}^{L} W_h \bar{x}_{h(rss)}$ is the stratified ranked set sample mean of auxiliary variable $x$.

Following Diana and Perri (2010), we define the regression imputation methods to impute the missing value under $SRSS$ as:

*Situation I*

$$y^c_{.i_{DP_1}} = \begin{cases} y_i & \text{for } i \in A, \\ \bar{y}_r + \frac{n}{n-r} b_1(\bar{X} - \bar{x}_{n,srss}) y_{r,srss} & \text{for } i \in \bar{A}. \end{cases}$$

*Situation II*

$$y^c_{.i_{DP_2}} = \begin{cases} y_i & \text{for } i \in A, \\ \bar{y}_r + \frac{n}{n-r} b_2(\bar{X} - \bar{x}_{r,srss}) y_{r,srss} & \text{for } i \in \bar{A}. \end{cases}$$

*Situation III*

$$y^c_{.i_{DP_3}} = \begin{cases} y_i & \text{for } i \in A, \\ \bar{y}_r + \frac{n}{n-r} b_3(\bar{x}_{n,srss} - \bar{x}_{r,srss}) y_{r,srss} & \text{for } i \in \bar{A}. \end{cases}$$

The sequent combined estimators under the above situations are given by

$$\begin{aligned} T^c_{DP_1} &= \bar{y}_{r,srss} + b_1(\bar{X} - \bar{x}_{n,srss}), \\ T^c_{DP_2} &= \bar{y}_{r,srss} + b_2(\bar{X} - \bar{x}_{r,srss}), \\ T^c_{DP_3} &= \bar{y}_{r,srss} + b_3(\bar{x}_{n,srss} - \bar{x}_{r,srss}). \end{aligned}$$

Following Sohail *et al.* (2018), one may envisage a combined class of ratio type imputation methods under $SRSS$ for the imputation of missing values as:

*Situation I*

$$y^c_{.iS_1} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{n}{n-r}\left[\bar{y}_{r,srss}\left(\frac{\bar{X}}{\bar{x}_{n,srss}}\right)^{\beta_1} - \bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}, \end{cases}$$

$$y^c_{.iS_4} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{n}{n-r}\left[\bar{y}_{r,srss}\left(\frac{\bar{X}}{\beta_4 \bar{x}_{n,srss} + (1-\beta_4)\bar{X}}\right) - \bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}. \end{cases}$$

*Situation II*

$$y^c_{.iS_2} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{n}{n-r}\left[\bar{y}_{r,srss}\left(\frac{\bar{X}}{\bar{x}_{r,srss}}\right)^{\beta_2} - \bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}, \end{cases}$$

$$y^c_{.iS_5} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{n}{n-r}\left[\bar{y}_{r,srss}\left(\frac{\bar{X}}{\beta_5 \bar{x}_{r,srss} + (1-\beta_5)\bar{X}}\right) - \bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}. \end{cases}$$

*Situation III*

$$y^c_{.iS_3} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{n}{n-r}\left[\bar{y}_{r,srss}\left(\frac{\bar{x}_{n,srss}}{\bar{x}_{r,srss}}\right)^{\beta_3} - \bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}, \end{cases}$$

$$y^c_{.iS_6} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{n}{n-r}\left[\bar{y}_{r,srss}\left(\frac{\bar{X}}{\beta_6 \bar{x}_{r,srss} + (1-\beta_6)\bar{x}_{n,srss}}\right) - \bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}. \end{cases}$$

The sequent estimators are given by

$$T_{S_1}^c = \bar{y}_{r,srss}\left(\frac{\bar{X}}{\bar{x}_{n,srss}}\right)^{\beta_1},$$
$$T_{S_2}^c = \bar{y}_{r,srss}\left(\frac{\bar{X}}{\bar{x}_{r,srss}}\right)^{\beta_2},$$
$$T_{S_3}^c = \bar{y}_{r,srss}\left(\frac{\bar{x}_{n,srss}}{\bar{x}_{r,srss}}\right)^{\beta_3},$$
$$T_{S_4}^c = \bar{y}_{r,srss}\left(\frac{\bar{X}}{\beta_4\bar{x}_{n,srss} + (1-\beta_4)\bar{X}}\right),$$
$$T_{S_5}^c = \bar{y}_{r,srss}\left(\frac{\bar{X}}{\beta_5\bar{x}_{r,srss} + (1-\beta_5)\bar{X}}\right),$$
$$T_{S_6}^c = \bar{y}_{r,srss}\left(\frac{\bar{X}}{\beta_6\bar{x}_{r,srss} + (1-\beta_6)\bar{x}_{n,srss}}\right),$$

where $\beta_i;\ \ i = 1, 2, ..., 6$ are suitably chosen optimizing scalars.

Appendix A of supplementary file contains the minimum mean square error (MSE) of the sequent estimators consisting of different imputation methods.

### 3.2. Separate imputation methods

The separate mean method of imputation under $SRSS$ is given by

$$y_{.i_m}^s = \begin{cases} y_i & \text{for } i \in A_h, \\ \bar{y}_{r,h[rss]} & \text{for } i \in \bar{A}_h. \end{cases}$$

The sequent estimator is given by

$$T_m^s = \sum_{h=1}^{L} W_h \bar{y}_{r,h[rss]},$$

where $\bar{y}_{r,h[rss]} = \frac{1}{m_h k}\sum_{i=1}^{m_h}\sum_{j=1}^{k} y_{h[i]j}$ is the ranked set sample mean of study variable in stratum $h$.

The separate imputation methods are categorized into three situations under the availability of auxiliary informations:

*Situation I*: When $\bar{X}$ is known and $\bar{x}_{n,h(rss)}$ is utilized.

*Situation II*: When $\bar{X}$ is known and $\bar{x}_{r,h(rss)}$ is utilized.

*Situation III*: When $\bar{X}$ is unknown and $\bar{x}_{n,h(rss)}$, $\bar{x}_{r,h(rss)}$ are utilized.

The classical separate ratio type imputation method is described under $SRSS$ as:

*Situation I*

$$y_{.i_{R_1}}^s = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\bar{x}_{n,h(rss)}}\right) - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h. \end{cases}$$

*Situation II*

$$y^s_{.i_{R_2}} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\bar{x}_{r,h(rss)}}\right) - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h. \end{cases}$$

*Situation III*

$$y^s_{.i_{R_3}} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\bar{y}_{r,h[rss]}\left(\frac{\bar{x}_{n,h(rss)}}{\bar{x}_{r,h(rss)}}\right) - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h. \end{cases}$$

The sequent estimators are given by

$$T^s_{R_1} = \sum_{h=1}^{L} W_h\left[\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\bar{x}_{n,h(rss)}}\right)\right],$$
$$T^s_{R_2} = \sum_{h=1}^{L} W_h\left[\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\bar{x}_{r,h(rss)}}\right)\right],$$
$$T^s_{R_3} = \sum_{h=1}^{L} W_h\left[\bar{y}_{r,h[rss]}\left(\frac{\bar{x}_{n,h(rss)}}{\bar{x}_{r,h(rss)}}\right)\right].$$

On the lines of Diana and Perri (2010), we define a separate regression imputation method under $SRSS$ as:

*Situation I*

$$y^s_{.i_{DP_1}} = \begin{cases} y_i & \text{for } i \in A_h, \\ y_{r,h[rss]} + \frac{n}{n-r}b_1(\bar{X} - \bar{x}_{n,h(rss)}) & \text{for } i \in \bar{A}_h. \end{cases}$$

*Situation II*

$$y^s_{.i_{DP_2}} = \begin{cases} y_i & \text{for } i \in A_h, \\ y_{r,h[rss]} + \frac{n}{n-r}b_2(\bar{X} - \bar{x}_{r,h(rss)}) & \text{for } i \in \bar{A}_h. \end{cases}$$

*Situation III*

$$y^s_{.i_{DP_3}} = \begin{cases} y_i & \text{for } i \in A_h, \\ y_{r,h[rss]} + \frac{n}{n-r}b_3(\bar{x}_{n,h(rss)} - \bar{x}_{r,h(rss)}) & \text{for } i \in \bar{A}_h. \end{cases}$$

The sequent separate estimators under the above situations are given by

$$T^s_{DP_1} = \sum_{h=1}^{L} W_h[\bar{y}_{r,h[rss]} + b_{1_h}(\bar{X}_h - \bar{x}_{n,h(rss)})],$$
$$T^s_{DP_2} = \sum_{h=1}^{L} W_h[\bar{y}_{r,h[rss]} + b_{2_h}(\bar{X}_h - \bar{x}_{r,h(rss)})],$$
$$T^s_{DP_3} = \sum_{h=1}^{L} W_h[\bar{y}_{r,h[rss]} + b_{3_h}(\bar{x}_{n,h(rss)} - \bar{x}_{r,h(rss)})].$$

Motivated by Sohail *et al.* (2018), we define a separate class of ratio type imputation methods under $SRSS$ as:

*Situation I*

$$y^{s}_{.is_1} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\bar{x}_{n,h(rss)}}\right)^{\beta_{1_h}} - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h. \end{cases}$$

$$y^{s}_{.is_4} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\beta_{4_h}\bar{x}_{n,h(rss)}+(1-\beta_{4_h})\bar{X}_h}\right) - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h. \end{cases}$$

*Situation II*

$$y^{s}_{.is_2} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\bar{x}_{r,h(rss)}}\right)^{\beta_{2_h}} - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h, \end{cases}$$

$$y^{s}_{.is_5} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\beta_{5_h}\bar{x}_{r,h(rss)}+(1-\beta_{5_h})\bar{X}_h}\right) - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h. \end{cases}$$

*Situation III*

$$y^{s}_{.is_3} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\bar{y}_{r,h[rss]}\left(\frac{\bar{x}_{n,h(rss)}}{\bar{x}_{r,h(rss)}}\right)^{\beta_{3_h}} - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h, \end{cases}$$

$$y^{s}_{.is_6} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\beta_{6_h}\bar{x}_{r,h(rss)}+(1-\beta_{6_h})\bar{x}_{n,h(rss)}}\right) - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h. \end{cases}$$

The sequent estimators are given by

$$T^{s}_{S_1} = \sum_{h=1}^{L} W_h\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\bar{x}_{n,h(rss)}}\right)^{\beta_{1_h}},$$

$$T^{s}_{S_2} = \sum_{h=1}^{L} W_h\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\bar{x}_{r,h(rss)}}\right)^{\beta_{2_h}},$$

$$T^{s}_{S_3} = \sum_{h=1}^{L} W_h\bar{y}_{r,h[rss]}\left(\frac{\bar{x}_{n,h(rss)}}{\bar{x}_{r,h(rss)}}\right)^{\beta_{3_h}},$$

$$T^{s}_{S_4} = \sum_{h=1}^{L} W_h\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\beta_{4_h}\bar{x}_{n,h(rss)} + (1-\beta_{4_h})\bar{X}_h}\right),$$

$$T^{s}_{S_5} = \sum_{h=1}^{L} W_h\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\beta_{5_h}\bar{x}_{r,h(rss)} + (1-\beta_{5_h})\bar{X}_h}\right),$$

$$T^{s}_{S_6} = \sum_{h=1}^{L} W_h\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\beta_{6_h}\bar{x}_{r,h(rss)} + (1-\beta_{6_h})\bar{x}_{n,h(rss)}}\right),$$

where $\beta_{i_h};\ i = 1, 2, ..., 6$ are suitably opted scalars.

Appendix B of supplementary file contains the minimum mean square error (MSE) of the sequent estimators consisting of different imputation methods.

## 4. PROPOSED IMPUTATION METHODS

The crux of this paper is binary:

1. To propose some efficient combined and separate imputation methods for the estimation of population mean $\bar{Y}$.
2. To determine the effect of the correlation coefficient, coefficient of skewness, and coefficient of kurtosis over the efficiency of the imputation procedures.

### 4.1. Combined imputation methods

Following Bhushan and Pandey (2016, 2018), we envisage nine new imputation methods under the three situations specified in the former section as:

*Situation I*

$$y^c_{.iSA_1} = \begin{cases} \alpha_1 y_i & \text{for } i \in A, \\ \alpha_1 \bar{y}_{r,srss} + \frac{n\theta_1}{n-r}(\bar{x}_{n,srss} - \bar{X}) & \text{for } i \in \bar{A}, \end{cases}$$

$$y^c_{.iSA_4} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{1}{n-r}\left[n\alpha_4 \bar{y}_{r,srss}\left(\frac{\bar{X}}{\bar{x}_{n,srss}}\right)^{\theta_4} - r\bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}, \end{cases}$$

$$y^c_{.iSA_7} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{1}{n-r}\left[n\alpha_7 \bar{y}_{r,srss}\left(\frac{\bar{X}}{\bar{X}+\theta_7(\bar{x}_{n,srss}-\bar{X})}\right) - r\bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}. \end{cases}$$

*Situation II*

$$y^c_{.iSA_2} = \begin{cases} \alpha_2 y_i & \text{for } i \in A, \\ \alpha_2 \bar{y}_{r,srss} + \frac{n\theta_2}{n-r}(\bar{x}_{r,srss} - \bar{X}) & \text{for } i \in \bar{A}, \end{cases}$$

$$y^c_{.iSA_5} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{1}{n-r}\left[n\alpha_5 \bar{y}_{r,srss}\left(\frac{\bar{X}}{\bar{x}_{r,srss}}\right)^{\theta_5} - r\bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}, \end{cases}$$

$$y^c_{.iSA_8} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{1}{n-r}\left[n\alpha_8 \bar{y}_{r,srss}\left(\frac{\bar{X}}{\bar{X}+\theta_8(\bar{x}_{r,srss}-\bar{X})}\right) - r\bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}. \end{cases}$$

*Situation III*

$$y^c_{.iSA_3} = \begin{cases} \alpha_3 y_i & \text{for } i \in A, \\ \alpha_3 \bar{y}_{r,srss} + \frac{n\theta_3}{n-r}(\bar{x}_{r,srss} - \bar{x}_{n,srss}) & \text{for } i \in \bar{A}, \end{cases}$$

$$y^c_{.iSA_6} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{1}{n-r}\left[n\alpha_6 \bar{y}_{r,srss}\left(\frac{\bar{x}_{n,srss}}{\bar{x}_{r,srss}}\right)^{\theta_6} - r\bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}, \end{cases}$$

$$y^c_{.iSA_9} = \begin{cases} y_i & \text{for } i \in A, \\ \frac{1}{n-r}\left[n\alpha_9 \bar{y}_{r,srss}\left(\frac{\bar{x}_{n,srss}}{\bar{x}_{n,srss}+\theta_9(\bar{x}_{n,srss}-\bar{x}_{r,srss})}\right) - r\bar{y}_{r,srss}\right] & \text{for } i \in \bar{A}. \end{cases}$$

Under the above situations, the sequent estimators are given by

$$
\begin{aligned}
T^c_{SA_1} &= \alpha_1 \bar{y}_{r,srss} + \theta_1(\bar{x}_{n,srss} - \bar{X}),\\
T^c_{SA_2} &= \alpha_2 \bar{y}_{r,srss} + \theta_2(\bar{x}_{r,srss} - \bar{X}),\\
T^c_{SA_3} &= \alpha_3 \bar{y}_{r,srss} + \theta_3(\bar{x}_{r,srss} - \bar{x}_{n,srss}),\\
T^c_{SA_4} &= \alpha_4 \bar{y}_{r,srss}\left(\frac{\bar{X}}{\bar{x}_{n,srss}}\right)^{\theta_4},\\
T^c_{SA_5} &= \alpha_5 \bar{y}_{r,srss}\left(\frac{\bar{X}}{\bar{x}_{r,srss}}\right)^{\theta_5},\\
T^c_{SA_6} &= \alpha_6 \bar{y}_{r,srss}\left(\frac{\bar{x}_{n,srss}}{\bar{x}_{r,srss}}\right)^{\theta_6},\\
T^c_{SA_7} &= \alpha_7 \bar{y}_{r,srss}\left[\frac{\bar{X}}{\bar{X} + \theta_7(\bar{x}_{n,srss} - \bar{X})}\right],\\
T^c_{SA_8} &= \alpha_8 \bar{y}_{r,srss}\left[\frac{\bar{X}}{\bar{X} + \theta_8(\bar{x}_{r,srss} - \bar{X})}\right],\\
T^c_{SA_9} &= \alpha_9 \bar{y}_{r,srss}\left[\frac{\bar{x}_{n,srss}}{\bar{x}_{n,srss} + \theta_9(\bar{x}_{r,srss} - \bar{x}_{n,srss})}\right],
\end{aligned}
$$

where $\alpha_1, \alpha_2, ..., \alpha_9$ and $\theta_1, \theta_2, ..., \theta_9$ are the suitably chosen scalars.

**Theorem 4.1.** *The $MSE$ of the sequent estimators consisting of the proposed imputation methods is given by*

$$
\begin{aligned}
MSE(T^c_{SA_1}) &= (\alpha_1 - 1)^2\bar{Y}^2 + \alpha_1^2\bar{Y}^2 I_0^* + \theta_1^2\bar{X}^2 I_1 + 2\alpha_1\theta_1\bar{X}\bar{Y}I_{01},\\
MSE(T^c_{SA_2}) &= (\alpha_2 - 1)^2\bar{Y}^2 + \alpha_2^2\bar{Y}^2 I_0^* + \theta_2^2\bar{X}^2 I_1^* + 2\alpha_2\theta_2\bar{X}\bar{Y}I_{01}^*,\\
MSE(T^c_{SA_3}) &= \begin{bmatrix} (\alpha_3 - 1)^2\bar{Y}^2 + \alpha_3^2\bar{Y}^2 I_0^* + \theta_3^2\bar{X}^2\{I_1^* - I_1\}\\ +2\alpha_3\theta_3\bar{X}\bar{Y}\{I_{01}^* - I_{01}\} \end{bmatrix},\\
MSE(T^c_{SA_4}) &= \bar{Y}^2\begin{bmatrix} 1 + \alpha_4^2\{\ 1 + I_0^* + \theta_4(2\theta_4 + 1)I_1 - 4\theta_4 I_{01}\ \}\\ -2\alpha_4\left\{1 - \theta_4 I_{01} + \frac{\theta_4(\theta_4+1)}{2} I_1\right\} \end{bmatrix},\\
MSE(T^c_{SA_5}) &= \bar{Y}^2\begin{bmatrix} 1 + \alpha_5^2\{\ 1 + I_0^* + \theta_5(2\theta_5 + 1)I_1^* - 4\theta_5 I_{01}^*\ \}\\ -2\alpha_5\left\{1 - \theta_5 I_{01}^* + \frac{\theta_5(\theta_5+1)}{2} I_1^*\right\} \end{bmatrix},\\
MSE(T^c_{SA_6}) &= \bar{Y}^2\begin{bmatrix} 1 + \alpha_6^2\{\ 1 + I_0^* + \theta_6(2\theta_6 + 1)(I_1^* - I_1) - 4\theta_6(I_{01}^* - I_{01})\ \}\\ -2\alpha_6\left\{\ 1 - \theta_6(I_{01}^* - I_{01}) + \frac{\theta_6(\theta_6+1)}{2}(I_1^* - I_1)\ \right\} \end{bmatrix},\\
MSE(T^c_{SA_7}) &= \bar{Y}^2\begin{bmatrix} 1 + \alpha_7^2\{\ 1 + I_0^* + 3\theta_7^2 I_1 - 4\theta_7 I_{01}\ \}\\ -2\alpha_7\{1 + \theta_7^2 I_1 - \theta_7 I_{01}\} \end{bmatrix},\\
MSE(T^c_{SA_8}) &= \bar{Y}^2\begin{bmatrix} 1 + \alpha_8^2\{\ 1 + I_0^* + 3\theta_8^2 I_1^* - 4\theta_8 I_{01}^*\ \}\\ -2\alpha_8\{1 + \theta_8^2 I_1^* - \theta_8 I_{01}^*\} \end{bmatrix},\\
MSE(T^c_{SA_9}) &= \bar{Y}^2\begin{bmatrix} 1 + \alpha_9^2\{\ 1 + I_0^* + 3\theta_9^2(\ I_1^* - I_1\ ) - 4\theta_9(I_{01}^* - I_{01})\ \}\\ -2\alpha_9\{\ 1 + \theta_9^2(I_1^* - I_1) - \theta_9(I_{01}^* - I_{01})\ \} \end{bmatrix}.
\end{aligned}
$$

**Proof:** Appendix C of supplementary file contains a summary of the derivations. The derivations can easily be done using Taylor series expansion. □

**Theorem 4.2.** *The minimum $MSE$ of the sequent estimators consisting of the proposed imputation methods are*

$$minMSE(T^c_{SA_i}) = \bar{Y}^2(1-\alpha_{i(opt)}) = \bar{Y}^2\left(1-\frac{A_i^2}{B_i}\right);\ i=1,2,3, \tag{4.1}$$

$$minMSE(T^c_{SA_j}) = \bar{Y}^2\left(1-\frac{A_j^2}{B_j}\right);\ j=4,5,6, \tag{4.2}$$

$$minMSE(T^c_{SA_k}) = \bar{Y}^2\left(1-\frac{A_k^2}{B_k}\right);\ k=7,8,9. \tag{4.3}$$

**Proof:** Appendix C of supplementary file contains a summary of the derivations. □

**Corollary 4.1.** *The proposed sequent estimators $T^c_{SA_i},\ i=1,2,3$ dominate the proposed sequent estimators $T^c_{SA_j},\ j=4,5,6$, iff*

$$\alpha_{i(opt)} > \frac{A_j^2}{B_j} \tag{4.4}$$

*and contrariwise. Otherwise, both are equally efficient when the equality holds in (4.4).*

**Proof:** By comparing the minimum MSEs of the proposed estimators from (4.1) and (4.2), we get (4.4). □

**Corollary 4.2.** *The proposed sequent estimators $T^c_{SA_i},\ i=1,2,3$ dominate the proposed sequent estimators $T^c_{SA_k},\ k=7,8,9$, iff*

$$\alpha_{i(opt)} > \frac{A_k^2}{B_k} \tag{4.5}$$

*and contrariwise. Otherwise, both are equally efficient when the equality holds in (4.5).*

**Proof:** On comparing the minimum MSEs of the proposed estimators from (4.1) and (4.3), we get (4.5). □

**Corollary 4.3.** *The proposed sequent estimators $T^c_{SA_j},\ i=4,5,6$ dominate the proposed sequent estimators $T^c_{SA_k},\ k=7,8,9$, iff*

$$\frac{A_j^2}{B_j} > \frac{A_k^2}{B_k} \tag{4.6}$$

*and contrariwise. Otherwise, both are equally efficient when the equality holds in (4.6).*

**Proof:** On comparing the minimum MSEs of the proposed estimators from (4.2) and (4.3), we get (4.6). □

The only way to determine if (4.4), (4.5), and (4.6) are true in practise is through the computational analysis done in Section 6.

## 4.2. Separate imputation methods

On the lines of Bhushan and Pandey (2016, 2018), we suggest nine new separate imputation methods under the three situations discussed in the earlier section as:

*Situation I*

$$y^s_{.i_{SA_1}} = \begin{cases} \alpha_{1_h} y_i & \text{for } i \in A_h, \\ \alpha_{1_h} \bar{y}_{r,h[rss]} + \frac{n\theta_{1_h}}{n-r}(\bar{x}_{n,h(rss)} - \bar{X}_h) & \text{for } i \in \bar{A}_h, \end{cases}$$

$$y^s_{.i_{SA_4}} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\alpha_{4_h}\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\bar{x}_{n,h(rss)}}\right)^{\theta_{4_h}} - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h, \end{cases}$$

$$y^s_{.i_{SA_7}} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\alpha_{7_h}\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\bar{X}_h+\theta_{7_h}(\bar{x}_{n,h(rss)}-\bar{X}_h)}\right) - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h. \end{cases}$$

*Situation II*

$$y^s_{.i_{SA_2}} = \begin{cases} \alpha_{2_h} y_i & \text{for } i \in A_h, \\ \alpha_{2_h} \bar{y}_{r,h[rss]} + \frac{n\theta_{2_h}}{n-r}(\bar{x}_{r,h(rss)} - \bar{X}_h) & \text{for } i \in \bar{A}_h, \end{cases}$$

$$y^s_{.i_{SA_5}} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\alpha_{5_h}\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\bar{x}_{r,h(rss)}}\right)^{\theta_{5_h}} - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h, \end{cases}$$

$$y^s_{.i_{SA_8}} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\alpha_{8_h}\bar{y}_{r,h[rss]}\left(\frac{\bar{X}_h}{\bar{X}+\theta_{8_h}(\bar{x}_{r,h(rss)}-\bar{X}_h)}\right) - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h. \end{cases}$$

*Situation III*

$$y^s_{.i_{SA_3}} = \begin{cases} \alpha_{3_h} y_i & \text{for } i \in A_h, \\ \alpha_{3_h} \bar{y}_{r,h[rss]} + \frac{n\theta_{3_h}}{n-r}(\bar{x}_{r,h(rss)} - \bar{x}_{n,h(rss)}) & \text{for } i \in \bar{A}_h, \end{cases}$$

$$y^s_{.i_{SA_6}} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\alpha_{6_h}\bar{y}_{r,h[rss]}\left(\frac{\bar{x}_{n,h(rss)}}{\bar{x}_{r,h(rss)}}\right)^{\theta_{6_h}} - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h, \end{cases}$$

$$y^s_{.i_{SA_9}} = \begin{cases} y_i & \text{for } i \in A_h, \\ \frac{1}{n-r}\left[n\alpha_{9_h}\bar{y}_{r,h[rss]}\left(\frac{\bar{x}_{n,h(rss)}}{\bar{x}_{n,h(rss)}+\theta_{9_h}(\bar{x}_{n,h(rss)}-\bar{x}_{r,h(rss)})}\right) - r\bar{y}_{r,h[rss]}\right] & \text{for } i \in \bar{A}_h. \end{cases}$$

The sequent estimators consisting of the above imputation methods are

$$T^s_{SA_1} = \sum_{h=1}^{L} W_h[\alpha_{1_h}\bar{y}_{r,h[rss]} + \theta_{1_h}(\bar{x}_{n,h(rss)} - \bar{X}_h)],$$

$$T^s_{SA_2} = \sum_{h=1}^{L} W_h[\alpha_{2_h}\bar{y}_{r,h[rss]} + \theta_{2_h}(\bar{x}_{r,h(rss)} - \bar{X}_h)],$$

$$T^s_{SA_3} = \sum_{h=1}^{L} W_h[\alpha_{3_h}\bar{y}_{r,h[rss]} + \theta_{3_h}(\bar{x}_{r,h(rss)} - \bar{x}_{n,h(rss)})],$$

$$T_{SA_4}^s = \sum_{h=1}^{L} W_h \alpha_{4_h} \bar{y}_{r,h[rss]} \left( \frac{\bar{X}_h}{\bar{x}_{n,h(rss)}} \right)^{\theta_{4_h}},$$

$$T_{SA_5}^s = \sum_{h=1}^{L} W_h \alpha_{5_h} \bar{y}_{r,h[rss]} \left( \frac{\bar{X}_h}{\bar{x}_{r,h(rss)}} \right)^{\theta_{5_h}},$$

$$T_{SA_6}^s = \sum_{h=1}^{L} W_h \alpha_{6_h} \bar{y}_{r,h[rss]} \left( \frac{\bar{x}_{n,h(rss)}}{\bar{x}_{r,h(rss)}} \right)^{\theta_{6_h}},$$

$$T_{SA_7}^s = \sum_{h=1}^{L} W_h \alpha_{7_h} \bar{y}_{r,h[rss]} \left[ \frac{\bar{X}_h}{\bar{X}_h + \theta_{7_h}(\bar{x}_{n,h(rss)} - \bar{X}_h)} \right],$$

$$T_{SA_8}^s = \sum_{h=1}^{L} W_h \alpha_{8_h} \bar{y}_{r,h[rss]} \left[ \frac{\bar{X}_h}{\bar{X}_h + \theta_{8_h}(\bar{x}_{r,h(rss)} - \bar{X}_h)} \right],$$

$$T_{SA_9}^s = \sum_{h=1}^{L} W_h \alpha_{9_h} \bar{y}_{r,h[rss]} \left[ \frac{\bar{x}_{n,h(rss)}}{\bar{x}_{n,h(rss)} + \theta_{9_h}(\bar{x}_{r,h(rss)} - \bar{x}_{n,h(rss)})} \right],$$

where $\alpha_{1_h}, \alpha_{2_h}, ..., \alpha_{9_h}$ and $\theta_{1_h}, \theta_{2_h}, ..., \theta_{9_h}$ are suitably chosen scalars.

**Theorem 4.3.** *The $MSE$ of the sequent estimators consisting of the proposed imputation methods is given by*

$$MSE(T_{SA_1}^s) = \sum_{h=1}^{L} W_h^2 \left[ (\alpha_{1_h} - 1)^2 \bar{Y}_h^2 + \alpha_{1_h}^2 \bar{Y}_h^2 J_0^* + \theta_{1_h}^2 \bar{X}_h^2 J_1 + 2\alpha_{1_h} \theta_{1_h} \bar{X}_h \bar{Y}_h J_{01} \right],$$

$$MSE(T_{SA_2}^s) = \sum_{h=1}^{L} W_h^2 \left[ (\alpha_{2_h} - 1)^2 \bar{Y}_h^2 + \alpha_{2_h}^2 \bar{Y}_h^2 J_0^* + \theta_{2_h}^2 \bar{X}_h^2 J_1^* + 2\alpha_{2_h} \theta_{2_h} \bar{X}_h \bar{Y}_h J_{01}^* \right],$$

$$MSE(T_{SA_3}^s) = \sum_{h=1}^{L} W_h^2 \begin{bmatrix} (\alpha_{3_h} - 1)^2 \bar{Y}_h^2 + \alpha_{3_h}^2 \bar{Y}_h^2 J_0^* + \theta_{3_h}^2 \bar{X}_h^2 \{J_1^* - J_1\} \\ +2\alpha_{3_h} \theta_{3_h} \bar{X}_h \bar{Y}_h \{J_{01}^* - J_{01}\} \end{bmatrix},$$

$$MSE(T_{SA_4}^s) = \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \begin{bmatrix} 1 + \alpha_{4_h}^2 \{ 1 + J_0^* + \theta_{4_h}(2\theta_{4_h} + 1) J_1 - 4\theta_{4_h} J_{01} \} \\ -2\alpha_{4_h} \left\{ 1 - \theta_{4_h} J_{01} + \frac{\theta_{4_h}(\theta_{4_h}+1)}{2} J_1 \right\} \end{bmatrix},$$

$$MSE(T_{SA_5}^s) = \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \begin{bmatrix} 1 + \alpha_{5_h}^2 \{ 1 + J_0^* + \theta_{5_h}(2\theta_{5_h} + 1) J_1^* - 4\theta_{5_h} J_{01}^* \} \\ -2\alpha_{5_h} \left\{ 1 - \theta_{5_h} J_{01}^* + \frac{\theta_{5_h}(\theta_{5_h}+1)}{2} J_1^* \right\} \end{bmatrix},$$

$$MSE(T_{SA_6}^s) = \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \begin{bmatrix} 1 + \alpha_{6_h}^2 \{ 1 + J_0^* + \theta_{6_h}(2\theta_{6_h} + 1)(J_1^* - J_1) - 4\theta_{6_h}(J_{01}^* - J_{01}) \} \\ -2\alpha_{6_h} \left\{ 1 - \theta_{6_h}(J_{01}^* - J_{01}) + \frac{\theta_{6_h}(\theta_{6_h}+1)}{2}(J_1^* - J_1) \right\} \end{bmatrix},$$

$$MSE(T_{SA_7}^s) = \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \begin{bmatrix} 1 + \alpha_{7_h}^2 \{ 1 + J_0^* + 3\theta_{7_h}^2 J_1 - 4\theta_{7_h} J_{01} \} \\ -2\alpha_{7_h} \{ 1 + \theta_{7_h}^2 J_1 - \theta_{7_h} J_{01} \} \end{bmatrix},$$

$$MSE(T_{SA_8}^s) = \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \begin{bmatrix} 1 + \alpha_{8_h}^2 \{ 1 + J_0^* + 3\theta_{8_h}^2 J_1^* - 4\theta_{8_h} J_{01}^* \} \\ -2\alpha_{8_h} \{ 1 + \theta_{8_h}^2 J_1^* - \theta_{8_h} J_{01}^* \} \end{bmatrix},$$

$$MSE(T_{SA_9}^s) = \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \begin{bmatrix} 1 + \alpha_{9_h}^2 \{ 1 + J_0^* + 3\theta_{9_h}^2 ( J_1^* - J_1 ) - 4\theta_{9_h}(J_{01}^* - J_{01}) \} \\ -2\alpha_{9_h} \{ 1 + \theta_{9_h}^2 (J_1^* - J_1) - \theta_{9_h}(J_{01}^* - J_{01}) \} \end{bmatrix}.$$

**Proof:** Appendix C of supplementary file contains a summary of the derivations. The derivations can easily be done using Taylor series expansion. □

**Theorem 4.4.** *The minimum $MSE$ of the sequent estimators consisting of the proposed imputation methods is given by*

$$minMSE(T^s_{SA_i}) = \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 (1 - \alpha_{i_h(opt)}) = \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left(1 - \frac{A^2_{i_h}}{B_{i_h}}\right);\ i = 1, 2, 3, \tag{4.7}$$

$$minMSE(T^s_{SA_j}) = \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left(1 - \frac{A^2_{j_h}}{B_{j_h}}\right);\ j = 4, 5, 6, \tag{4.8}$$

$$minMSE(T^s_{SA_k}) = \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left(1 - \frac{A^2_{k_h}}{B_{k_h}}\right);\ k = 7, 8, 9. \tag{4.9}$$

**Proof:** Appendix C of supplementary file contains a summary of the derivations. □

**Corollary 4.4.** *The proposed sequent estimators $T^s_{SA_i},\ i = 1, 2, 3$ dominate the proposed sequent estimators $T^s_{SA_j},\ j = 4, 5, 6$, iff*

$$\sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \alpha_{i_h(opt)} > \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left(\frac{A^2_{i_h}}{B_{i_h}}\right) \tag{4.10}$$

*and contrariwise. Otherwise, both are equally efficient when the equality holds in (4.10).*

**Proof:** On comparing the minimum MSEs of the proposed estimators from (4.7) and (4.8), we get (4.10). □

**Corollary 4.5.** *The proposed sequent estimators $T^s_{SA_i},\ i = 1, 2, 3$ dominate the proposed sequent estimators $T^s_{SA_k},\ k = 4, 5, 6$, iff*

$$\sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \alpha_{i_h(opt)} > \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left(\frac{A^2_{k_h}}{B_{k_h}}\right) \tag{4.11}$$

*and contrariwise. Otherwise, both are equally efficient when the equality holds in (4.11).*

**Proof:** By comparing the minimum MSEs of the proposed estimators from (4.7) and (4.9), we get (4.11). □

**Corollary 4.6.** *The proposed sequent estimators $T^s_{SA_j},\ j = 4, 5, 6$ dominate the proposed sequent estimators $T^s_{SA_k},\ k = 7, 8, 9$, iff*

$$\sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left(\frac{A^2_{j_h}}{B_{j_h}}\right) > \sum_{h=1}^{L} W_h^2 \bar{Y}_h^2 \left(\frac{A^2_{k_h}}{B_{k_h}}\right) \tag{4.12}$$

*and contrariwise. Otherwise, both are equally efficient when the equality holds in (4.12).*

**Proof:** By comparing the minimum MSEs of the proposed estimators from (4.8) and (4.9), we get (4.12). □

The only way to determine if (4.10), (4.11), and (4.12) are true in practise is through the computational analysis done in Section 6.

## 5. OPTIMALITY CONDITIONS

In this section, we obtain the optimality conditions under two heads, namely, optimality conditions for combined imputation methods and the optimality conditions for separate imputation methods.

### 5.1. Optimality conditions for the combined imputation methods

By comparing the minimum $MSE$ of the suggested combined imputation methods $y^{c}_{.i_{SA_i}}$, $i = 1, 2, ..., 9$ from (4.1) and (4.2) with the minimum $MSE$ of the other existing combined imputation methods from $(A.1)$, $(A.2)$, $(A.3)$, $(A.4)$, $(A.8)$, $(A.9)$, $(A.10)$, $(A.14)$, $(A.15)$, and $(A.16)$, respectively, given in Appendix A of supplementary file, we get the following optimality conditions:

$$MSE(T^c_m) > MSE(T^c_{SA_i}) \implies \frac{A_i^2}{B_i} > 1 - I_0^*,$$

$$MSE(T^c_{R_1}) > MSE(T^c_{SA_i}) \implies \frac{A_i^2}{B_i} > 1 - I_0^* - I_1 + 2I_{01},$$

$$MSE(T^c_{R_2}) > MSE(T^c_{SA_i}) \implies \frac{A_i^2}{B_i} > 1 - I_0^* - I_1^* + 2I_{01}^*,$$

$$MSE(T^c_{R_3}) > MSE(T^c_{SA_i}) \implies \frac{A_i^2}{B_i} > 1 - I_0^* - I_1 - I_1^* + 2I_{01}^*,$$

$$MSE(T^c_{DP_1}) > MSE(T^c_{SA_i}) \implies \frac{A_i^2}{B_i} > 1 - I_0^* + \frac{I_{01}^2}{I_1},$$

$$MSE(T^c_{DP_2}) > MSE(T^c_{SA_i}) \implies \frac{A_i^2}{B_i} > 1 - I_0^* + \frac{{I_{01}^*}^2}{I_1^*},$$

$$MSE(T^c_{DP_3}) > MSE(T^c_{SA_i}) \implies \frac{A_i^2}{B_i} > 1 - I_0^* + \frac{(I_{01}^* - I_{01})^2}{(I_1^* - I_1)},$$

$$MSE(T^c_{S_1}) > MSE(T^c_{SA_i}) \implies \frac{A_i^2}{B_i} > 1 - I_0^* + \frac{I_{01}^2}{I_1},$$

$$MSE(T^c_{S_2}) > MSE(T^c_{SA_i}) \implies \frac{A_i^2}{B_i} > 1 - I_0^* + \frac{{I_{01}^*}^2}{I_1^*},$$

$$MSE(T^c_{S_3}) > MSE(T^c_{SA_i}) \implies \frac{A_i^2}{B_i} > 1 - I_0^* + \frac{(I_{01}^* - I_{01})^2}{(I_1^* - I_1)}.$$

The optimality of the suggested combined imputation methods can be justified under the above conditions.

### 5.2. Optimality conditions for the separate imputation methods

By comparing the minimum $MSE$ of the proposed imputation methods $y^s_{.i_{SA_i}}$, $i = 1, 2, ..., 9$ given in (4.7) and (4.8) with the minimum $MSE$ of the other existing imputation methods given in $(B.17)$, $(B.18)$, $(B.19)$, $(B.20)$, $(B.24)$, $(B.25)$, $(B.26)$, $(B.30)$, $(B.31)$, and $(B.32)$, respectively, given in Appendix B of supplementary file, we get the following optimality conditions:

$$MSE(T^s_m) > MSE(T^s_{SA_i}) \implies \sum_{h=1}^{L} W_h^2\bar{Y}_h^2\left(1-\frac{A^2_{i_h}}{B_{i_h}}\right) < \sum_{h=1}^{L} W_h^2\bar{Y}_h^2 J_1^*,$$

$$MSE(T^s_{R_1}) > MSE(T^s_{SA_i}) \implies \sum_{h=1}^{L} W_h^2\bar{Y}_h^2\left(1-\frac{A^2_{i_h}}{B_{i_h}}\right) < \sum_{h=1}^{L} W_h^2\bar{Y}_h^2 [J_0^* + J_1 - 2J_{01}],$$

$$MSE(T^s_{R_2}) > MSE(T^s_{SA_i}) \implies \sum_{h=1}^{L} W_h^2\bar{Y}_h^2\left(1-\frac{A^2_{i_h}}{B_{i_h}}\right) < \sum_{h=1}^{L} W_h^2\bar{Y}_h^2 [J_0^* + J_1^* - 2J_{01}^*],$$

$$MSE(T^s_{R_3}) > MSE(T^s_{SA_i}) \implies \sum_{h=1}^{L} W_h^2\bar{Y}_h^2\left(1-\frac{A^2_{i_h}}{B_{i_h}}\right) < \sum_{h=1}^{L} W_h^2\bar{Y}_h^2 [J_0^* + J_1^* + J_1 - 2J_{01}^*],$$

$$MSE(T^s_{DP_1}) > MSE(T^s_{SA_i}) \implies \sum_{h=1}^{L} W_h^2\bar{Y}_h^2\left(1-\frac{A^2_{i_h}}{B_{i_h}}\right) < \sum_{h=1}^{L} W_h^2\bar{Y}_h^2 \left[J_0^* - \frac{J_{01}^2}{J_1}\right],$$

$$MSE(T^s_{DP_2}) > MSE(T^s_{SA_i}) \implies \sum_{h=1}^{L} W_h^2\bar{Y}_h^2\left(1-\frac{A^2_{i_h}}{B_{i_h}}\right) < \sum_{h=1}^{L} W_h^2\bar{Y}_h^2 \left[J_0^* - \frac{{J_{01}^*}^2}{J_1^*}\right],$$

$$MSE(T^s_{DP_3}) > MSE(T^s_{SA_i}) \implies \sum_{h=1}^{L} W_h^2\bar{Y}_h^2\left(1-\frac{A^2_{i_h}}{B_{i_h}}\right) < \sum_{h=1}^{L} W_h^2\bar{Y}_h^2 \left[J_0^* - \frac{\left(J_{01}^* - J_{01}\right)^2}{\left(J_1^* - J_1\right)}\right],$$

$$MSE(T^s_{S_1}) > MSE(T^s_{SA_i}) \implies \sum_{h=1}^{L} W_h^2\bar{Y}_h^2\left(1-\frac{A^2_{i_h}}{B_{i_h}}\right) < \sum_{h=1}^{L} W_h^2\bar{Y}_h^2 \left[J_0^* - \frac{J_{01}^2}{J_1}\right],$$

$$MSE(T^s_{S_2}) > MSE(T^s_{SA_i}) \implies \sum_{h=1}^{L} W_h^2\bar{Y}_h^2\left(1-\frac{A^2_{i_h}}{B_{i_h}}\right) < \sum_{h=1}^{L} W_h^2\bar{Y}_h^2 \left[J_0^* - \frac{{J_{01}^*}^2}{J_1^*}\right],$$

$$MSE(T^s_{S_3}) > MSE(T^s_{SA_i}) \implies \sum_{h=1}^{L} W_h^2\bar{Y}_h^2\left(1-\frac{A^2_{i_h}}{B_{i_h}}\right) < \sum_{h=1}^{L} W_h^2\bar{Y}_h^2 \left[J_0^* - \frac{\left(J_{01}^* - J_{01}\right)^2}{\left(J_1^* - J_1\right)}\right].$$

Under the above conditions, the optimality of the proposed separate imputation methods can be ascertained.

### 5.3. Comparison of proposed combined and separate imputation methods

By comparing the minimum $MSE$ of the proposed combined and separate classes of imputation methods from (4.1), (4.2) and (4.7), (4.8), we get

$$(5.1)\quad minMSE(T^c_{SA_i}) - minMSE(T^s_{SA_i}) = \sum_{h=1}^{L}\left[(\bar{Y}^2 - W_h^2\bar{Y}_h^2) - \left(\bar{Y}^2\frac{A_i^2}{B_i} - W_h^2\bar{Y}_h^2\frac{A^2_{i_h}}{B_{i_h}}\right)\right].$$

If the sequent estimators are conclusive and the relationship between auxiliary and study variables within each stratum is a straight line passing through the origine, then the last term of (5.1) is miniscule and it vanished.

In addition, except $R_h$ becomes invariant from stratum to stratum, the separate estimators perform better in each stratum provided the sample in each stratum is to be sufficiently large so that the approximate formula for $MSE(T^s_{SA_i}),\ i = 1, 2, ..., 9$ is valid and the cumulative bias that can affect the proposed estimators is negligible, whereas the proposed combined estimators are to be highly advocated with only a small sample in each stratum (see, Cochran, 1977).

## 6. SIMULATION STUDY

To highlight the properties and to access the performance of the proposed imputation methods, motivated by Singh and Horn (1998), simulations were carried out over two artificially generated asymmetric populations such as gamma and exponential of size $N = 2100$ units each with variables $X$ and $Y$ whose values are given by

$$y_i = 8.2 + \sqrt{(1-\rho^2_{xy})}\ y^*_i + \rho_{xy}\left(\frac{S_y}{S_x}\right)x^*_i,$$
$$x_i = 4.2 + x^*_i.$$

where $x^*_i$ and $y^*_i$ are independent variates of proportional distribution. Each population is divided into three equal strata and a stratified ranked set sample of size 9 with set size 3 and number of cycles 3 is drawn from each stratum with the help of the methodology described in Section 2. With 10000 iterations, the percent relative efficiency ($PRE$) of the sequent estimators with respect to the conventional mean estimator was obtained as

$$PRE = \frac{\frac{1}{10000}\sum_{i=1}^{10000}(T_m - \bar{Y})^2}{\frac{1}{10000}\sum_{i=1}^{10000}(T^* - \bar{Y})^2} \times 100,$$

where $T^*$ is the existing and proposed combined and separate class of estimators.

The findings of the simulation are disclosed from Table 1 to Table 4 through $PRE$ for reasonably chosen values of correlation coefficient $\rho_{xy} = 0.6, 0.7, 0.8, 0.9$ and fair choice of response probability $P = 0.4, 0.6$.

From Table 1 to Table 4, consisting of the simulation results of two asymmetric populations, namely, gamma and exponential, we have seen that the proposed combined and separate imputation methods $y^c_{.i_{SA_j}}$ and $y^s_{.i_{SA_j}}$, $j = 1, 2, ..., 9$ dominate the other existing imputation methods for reasonably chosen values of the correlation coefficient. We have also seen that the proposed combined and separate imputation methods $y^c_{.i_{SA_j}}$ and $y^s_{.i_{SA_j}}$, $j = 4, 5, 6$ perform better among the proposed class of imputation methods under situations I, II and III. Moreover, it is also seen that the $PRE$ of the proposed imputation methods under situations I, II and III in both populations decreases with the increase in asymmetry.

**Table 1**: $PRE$ of proposed combined estimators at $P = 0.4$.

| $\rho_{xy}$ | 0.6 | 0.7 | 0.8 | 0.9 |
|---|---|---|---|---|
| $T_m^c$ | 100 | 100 | 100 | 100 |
| $x^* \sim \Gamma(0.5, 1.5)$ | | | | |
| $y^* \sim \Gamma(1.5, 2)$ | | | | |
| Skewness of $y$ | 1.3292 | 1.4081 | 1.6083 | 1.9768 |
| Kurtosis of $y$ | 5.4000 | 5.6181 | 6.5418 | 8.4693 |
| Situation I | | | | |
| $T_{SA_i}^c,\ i = 1, 7$ | 106.0998 | 105.5777 | 105.4108 | 104.7131 |
| $\mathbf{T_{SA_4}^c}$ | **106.1067** | **105.5846** | **105.4176** | **104.7198** |
| $T_{R_1}^c$ | 100.2267 | 99.3709 | 98.3220 | 95.5744 |
| $T_{DP_1}^c/T_{S_i}^c,\ i = 1, 4$ | 105.983 | 105.46 | 105.3005 | 104.6141 |
| Situation II | | | | |
| $T_{SA_i}^c,\ i = 2, 8$ | 115.0805 | 113.7576 | 113.4881 | 111.7679 |
| $\mathbf{T_{SA_5}^c}$ | **115.1003** | **113.7771** | **113.5075** | **111.7865** |
| $T_{R_1}^c$ | 97.5580 | 95.7573 | 93.4999 | 87.5916 |
| $T_{DP_2}^c/T_{S_i}^c,\ i = 2, 5$ | 114.9637 | 113.6399 | 113.3778 | 111.6689 |
| Situation III | | | | |
| $T_{SA_i}^c,\ i = 3, 9$ | 108.1050 | 107.4685 | 107.3868 | 106.5397 |
| $\mathbf{T_{SA_6}^c}$ | **108.1161** | **107.4795** | **107.3977** | **106.5502** |
| $T_{R_1}^c$ | 97.3431 | 96.3414 | 95.0160 | 91.2945 |
| $T_{DP_3}^c/T_{S_i}^c,\ i = 3, 6$ | 107.9882 | 107.3508 | 107.2765 | 106.4406 |
| $x^* \sim Exp(3.0)$ | | | | |
| $y^* \sim Exp(2.0)$ | | | | |
| Skewness of $y$ | 1.4612 | 1.3814 | 1.3734 | 1.4769 |
| Kurtosis of $y$ | 5.9268 | 5.4885 | 5.4119 | 5.8395 |
| Situation I | | | | |
| $T_{SA_i}^c,\ i = 1, 7$ | 106.2035 | 106.5419 | 106.7350 | 105.8748 |
| $\mathbf{T_{SA_4}^c}$ | **106.2096** | **106.5481** | **106.7411** | **105.8808** |
| $T_{R_1}^c$ | 99.7239 | 100.2227 | 100.2960 | 98.5947 |
| $T_{DP_1}^c/T_{S_i}^c,\ i = 1, 4$ | 106.1075 | 106.4467 | 106.6419 | 105.7829 |
| Situation II | | | | |
| $T_{SA_i}^c,\ i = 2, 8$ | 112.9619 | 112.5542 | 112.8376 | 111.6621 |
| $\mathbf{T_{SA_5}^c}$ | **112.9805** | **112.5726** | **112.856** | **111.6799** |
| $T_{R_2}^c$ | 89.3047 | 88.4097 | 88.1260 | 85.9139 |
| $T_{DP_2}^c/T_{S_i}^c,\ i = 2, 5$ | 112.8659 | 112.4590 | 112.7445 | 111.5702 |
| Situation III | | | | |
| $T_{SA_i}^c,\ i = 3, 9$ | 105.3578 | 105.7988 | 105.8817 | 105.2694 |
| $\mathbf{T_{SA_6}^c}$ | **105.3684** | **105.8092** | **105.8920** | **105.2796** |
| $T_{R_3}^c$ | 87.5260 | 88.2363 | 87.8974 | 86.9790 |
| $T_{DP_3}^c/T_{S_i}^c,\ i = 3, 6$ | 105.2617 | 105.7036 | 105.7886 | 105.1775 |

**Table 2**: $PRE$ of proposed combined estimators at $P = 0.6$.

| $\rho_{xy}$ | 0.6 | 0.7 | 0.8 | 0.9 |
|---|---|---|---|---|
| $T_m^c$ | 100 | 100 | 100 | 100 |
| $x^* \sim \Gamma(0.5, 1.5)$ | | | | |
| $y^* \sim \Gamma(1.5, 2)$ | | | | |
| Skewness of $y$ | 1.3292 | 1.4081 | 1.6083 | 1.9768 |
| Kurtosis of $y$ | 5.4000 | 5.6181 | 6.5418 | 8.4693 |
| Situation I | | | | |
| $T_{SA_i}^c,\ i = 1, 7$ | 109.3365 | 108.5032 | 108.2434 | 107.1516 |
| $\mathbf{T_{SA_4}^c}$ | **109.3437** | **108.5104** | **108.2505** | **107.1584** |
| $T_{R_1}^c$ | 100.3407 | 99.0588 | 97.5032 | 93.5045 |
| $T_{DP_1}^c/T_{S_i}^c,\ i = 1, 4$ | 109.2587 | 108.4248 | 108.1698 | 107.0855 |
| Situation II | | | | |
| $T_{SA_i}^c,\ i = 2, 8$ | 115.3909 | 114.0088 | 113.6905 | 111.9040 |
| $\mathbf{T_{SA_5}^c}$ | **115.4041** | **114.0218** | **113.7034** | **111.9163** |
| $T_{R_2}^c$ | 98.2571 | 96.3813 | 94.0605 | 88.0659 |
| $T_{DP_2}^c/T_{S_i}^c,\ i = 2, 5$ | 115.3131 | 113.9304 | 113.6170 | 111.8379 |
| Situation III | | | | |
| $T_{SA_i}^c,\ i = 3, 9$ | 105.1641 | 104.7748 | 104.7360 | 104.2151 |
| $\mathbf{T_{SA_6}^c}$ | **105.1690** | **104.7796** | **104.7407** | **104.2197** |
| $T_{R_3}^c$ | 97.9303 | 97.2720 | 96.3820 | 93.8046 |
| $T_{DP_3}^c/T_{S_i}^c,\ i = 3, 6$ | 105.0863 | 104.6964 | 104.6625 | 104.1491 |
| $x^* \sim Exp(3.0)$ | | | | |
| $y^* \sim Exp(2.0)$ | | | | |
| Skewness of $y$ | 1.4612 | 1.3814 | 1.3734 | 1.4769 |
| Kurtosis of $y$ | 5.9268 | 5.4885 | 5.4119 | 5.8395 |
| Situation I | | | | |
| $T_{SA_i}^c,\ i = 1, 7$ | 109.5194 | 110.0755 | 110.3901 | 109.0041 |
| $\mathbf{T_{SA_4}^c}$ | **109.5258** | **110.0819** | **110.3965** | **109.0104** |
| $T_{R_1}^c$ | 99.5862 | 100.3351 | 100.4455 | 97.9046 |
| $T_{DP_1}^c/T_{S_i}^c,\ i = 1, 4$ | 109.4554 | 110.0121 | 110.3281 | 108.9429 |
| Situation II | | | | |
| $T_{SA_i}^c,\ i = 2, 8$ | 113.7547 | 113.6404 | 113.9884 | 112.5204 |
| $\mathbf{T_{SA_5}^c}$ | **113.7669** | **113.6525** | **114.0006** | **112.5322** |
| $T_{R_2}^c$ | 91.4958 | 91.0133 | 90.8178 | 88.2189 |
| $T_{DP_2}^c/T_{S_i}^c,\ i = 2, 5$ | 113.6907 | 113.577 | 113.9264 | 112.4592 |
| Situation III | | | | |
| $T_{SA_3}^c$ | 103.4189 | 103.5222 | 103.5633 | 103.3902 |
| $\mathbf{T_{SA_6}^c}$ | **103.4436** | **103.5267** | **103.5677** | **103.4246** |
| $T_{R_3}^c$ | 90.6449 | 90.7375 | 90.4534 | 89.9166 |
| $T_{DP_3}^c/T_{S_i}^c,\ i = 3, 6$ | 103.3749 | 103.4588 | 103.5013 | 103.3590 |

**Table 3**: $PRE$ of proposed separate estimators at $P = 0.4$.

| $\rho_{xy}$ | 0.6 | 0.7 | 0.8 | 0.9 |
|---|---|---|---|---|
| $T_m^s$ | 100 | 100 | 100 | 100 |
| $x^* \sim \Gamma(0.5, 1.5)$ | | | | |
| $y^* \sim \Gamma(1.5, 2)$ | | | | |
| Skewness of $y$ | 1.3292 | 1.4081 | 1.6083 | 1.9768 |
| Kurtosis of $y$ | 5.4000 | 5.6181 | 6.5418 | 8.4693 |
| Situation I | | | | |
| $T_{SA_i}^s,\ i = 1, 7$ | 106.0962 | 105.5742 | 105.4076 | 104.7104 |
| $\mathbf{T_{SA_4}^s}$ | **106.1027** | **105.5807** | **105.4141** | **104.7167** |
| $T_{R_1}^s$ | 100.6188 | 99.7795 | 98.7804 | 96.1227 |
| $T_{DP_1}^s/T_{S_i}^s,\ i = 1, 4$ | 105.983 | 105.4600 | 105.3005 | 104.6141 |
| Situation II | | | | |
| $T_{SA_i}^s,\ i = 2, 8$ | 115.0769 | 113.7541 | 113.4850 | 111.7652 |
| $\mathbf{T_{SA_5}^s}$ | **115.0954** | **113.7724** | **113.5031** | **111.7826** |
| $T_{R_2}^s$ | 98.5741 | 96.7889 | 94.6180 | 88.8216 |
| $T_{DP_2}^s/T_{S_i}^s,\ i = 2, 5$ | 114.9637 | 113.6399 | 113.3778 | 111.6689 |
| Situation III | | | | |
| $T_{SA_i}^s,\ i = 3, 9$ | 108.1014 | 107.4650 | 107.3836 | 106.5370 |
| $\mathbf{T_{SA_6}^s}$ | **108.1119** | **107.4753** | **107.3939** | **106.5469** |
| $T_{R_3}^s$ | 81.4466 | 80.3429 | 77.9883 | 72.6504 |
| $T_{DP_3}^s/T_{S_i}^s,\ i = 3, 6$ | 107.9882 | 107.3508 | 107.2765 | 106.4406 |
| $x^* \sim Exp(3.0)$ | | | | |
| $y^* \sim Exp(2.0)$ | | | | |
| Skewness of $y$ | 1.4612 | 1.3814 | 1.3734 | 1.4769 |
| Kurtosis of $y$ | 5.9268 | 5.4885 | 5.4119 | 5.8395 |
| Situation I | | | | |
| $T_{SA_i}^s,\ i = 1, 7$ | 106.2010 | 106.5393 | 106.7324 | 105.8724 |
| $\mathbf{T_{SA_4}^s}$ | **106.2067** | **106.5451** | **106.7382** | **105.8781** |
| $T_{R_1}^s$ | 100.0305 | 100.5150 | 100.5927 | 98.9217 |
| $T_{DP_1}^s/T_{S_i}^s,\ i = 1, 4$ | 106.1075 | 106.4467 | 106.6419 | 105.7829 |
| Situation II | | | | |
| $T_{SA_i}^s,\ i = 2, 8$ | 112.4594 | 112.5516 | 112.8350 | 111.6596 |
| $\mathbf{T_{SA_5}^s}$ | **112.4770** | **112.5691** | **112.8525** | **111.6766** |
| $T_{R_2}^s$ | 90.1295 | 89.1992 | 88.9243 | 86.7481 |
| $T_{DP_2}^s/T_{S_i}^s,\ i = 2, 5$ | 112.8659 | 112.4590 | 112.7445 | 111.5702 |
| Situation III | | | | |
| $T_{SA_i}^s,\ i = 3, 9$ | 105.3552 | 105.7962 | 105.8791 | 105.2670 |
| $\mathbf{T_{SA_6}^s}$ | **105.3654** | **105.8061** | **105.889** | **105.2767** |
| $T_{R_3}^s$ | 74.6746 | 73.9110 | 73.4349 | 71.8010 |
| $T_{DP_3}^s/T_{S_i}^s,\ i = 3, 6$ | 105.2617 | 105.7036 | 105.7886 | 105.1775 |

**Table 4**: $PRE$ of proposed separate estimators at $P = 0.6$.

| $\rho_{xy}$ | 0.6 | 0.7 | 0.8 | 0.9 |
|---|---|---|---|---|
| $T_m^s$ | 100 | 100 | 100 | 100 |
| $x^* \sim \Gamma(0.5, 1.5)$ | | | | |
| $y^* \sim \Gamma(1.5, 2)$ | | | | |
| Skewness of $y$ | 1.3292 | 1.4081 | 1.6083 | 1.9768 |
| Kurtosis of $y$ | 5.4000 | 5.6181 | 6.5418 | 8.4693 |
| Situation I | | | | |
| $T_{SA_i}^s,\ i = 1, 7$ | 109.3341 | 108.5009 | 108.2413 | 107.1498 |
| $\mathbf{T_{SA_4}^s}$ | **109.3409** | **108.5076** | **108.2479** | **107.1562** |
| $T_{R_1}^s$ | 100.9318 | 99.6695 | 98.1811 | 94.2940 |
| $T_{DP_1}^s/T_{S_i}^s,\ i = 1, 4$ | 109.2587 | 108.4248 | 108.1698 | 107.0855 |
| Situation II | | | | |
| $T_{SA_i}^s,\ i = 2, 8$ | 115.3885 | 114.0065 | 113.6884 | 111.9022 |
| $\mathbf{T_{SA_5}^s}$ | **115.4008** | **114.0187** | **113.7005** | **111.9138** |
| $T_{R_2}^s$ | 99.2688 | 97.4083 | 95.1744 | 89.2927 |
| $T_{DP_2}^s/T_{S_i}^s,\ i = 2, 5$ | 115.3131 | 113.9304 | 113.6170 | 111.8379 |
| Situation III | | | | |
| $T_{SA_i}^s,\ i = 3, 9$ | 105.1618 | 104.7725 | 104.7339 | 104.2133 |
| $\mathbf{T_{SA_6}^s}$ | **105.1663** | **104.7770** | **104.7383** | **104.2176** |
| $T_{R_3}^s$ | 75.3238 | 74.4041 | 72.0020 | 66.8509 |
| $T_{DP_3}^s/T_{S_i}^s,\ i = 3, 6$ | 105.0863 | 104.6964 | 104.6625 | 104.1491 |
| $x^* \sim Exp(3.0)$ | | | | |
| $y^* \sim Exp(2.0)$ | | | | |
| Skewness of $y$ | 1.4612 | 1.3814 | 1.3734 | 1.4769 |
| Kurtosis of $y$ | 5.9268 | 5.4885 | 5.4119 | 5.8395 |
| Situation I | | | | |
| $T_{SA_i}^s,\ i = 1, 7$ | 109.5177 | 110.0737 | 110.3884 | 109.0025 |
| $\mathbf{T_{SA_4}^s}$ | **109.5238** | **110.0798** | **110.3944** | **109.0084** |
| $T_{R_1}^s$ | 100.0457 | 100.7759 | 100.8935 | 98.3895 |
| $T_{DP_1}^s/T_{S_i}^s,\ i = 1, 4$ | 109.4554 | 110.0121 | 110.3281 | 108.9429 |
| Situation II | | | | |
| $T_{SA_i}^s,\ i = 2, 8$ | 113.7531 | 113.8386 | 113.9867 | 112.5188 |
| $\mathbf{T_{SA_5}^s}$ | **113.7646** | **113.8502** | **113.9983** | **112.5300** |
| $T_{R_2}^s$ | 92.3101 | 91.7963 | 91.6107 | 89.0465 |
| $T_{DP_2}^s/T_{S_i}^s,\ i = 2, 5$ | 113.6907 | 113.7770 | 113.9264 | 112.4592 |
| Situation III | | | | |
| $T_{SA_i}^s,\ i = 3, 9$ | 103.4372 | 103.5205 | 103.5615 | 103.4185 |
| $\mathbf{T_{SA_6}^s}$ | **103.4417** | **103.5247** | **103.5658** | **103.4228** |
| $T_{R_3}^s$ | 70.0312 | 69.5489 | 69.0562 | 67.4148 |
| $T_{DP_3}^s/T_{S_i}^s,\ i = 3, 6$ | 103.3749 | 103.4588 | 103.5013 | 103.3590 |

## 7. CONCLUSION

This paper is the outset to suggest some combined and separate classes of imputation methods along with their properties for the estimation of population mean in the presence of missing data using $SRSS$. The theoretical conditions are derived under which the proposed combined and separate classes of imputation methods are justified. In order to enhance the theoretical findings and to determine the effect of skewness and kurtosis over $PRE$, a simulation study is accomplished on two asymmetric populations viz. gamma and exponential with reasonable choice of correlation coefficient $\rho_{xy}$ and probability of non responding units $P$. It is noticed from the perusal of theoretical and simulation results that:

**1.** The proposed combined and separate of imputation methods $y^c_{.i_{SA_j}}$ and $y^s_{.i_{SA_j}}$, $j = 1, 2, ..., 9$ always perform better than the combined and separate mean imputation method $y^c_{.i_m}$ and $y^s_{.i_m}$, ratio imputation methods $y^c_{.i_{R_j}}$ and $y^s_{.i_{R_j}}$, $j = 1, 2, 3$ and their own conventional counterparts for different values of correlation coefficient $\rho_{xy}$, coefficient of skewness $\beta_1$ and coefficient of kurtosis $\beta_2$.

**2.** The proposed combined and separate imputation methods $y^c_{.i_{SA_j}}$ and $y^s_{.i_{SA_j}}$, $j = 4, 5, 6$ are best among the proposed classes of imputation methods under situations I, II and III.

**3.** The $PRE$ of the proposed combined and separate classes of imputation methods $y^c_{.i_{SA_j}}$, $y^s_{.i_{SA_j}}$, $j = 1, 2, ..., 9$ and their conventional counterparts under situations I, II and III are contrary to the asymmetry which is similar to the results of McIntyre (1952), Dell and Clutter (1972) and Bhushan and Kumar (2022) where they chose a wide range of skewed distributions and concluded that the asymmetry shows adverse effect over the efficiency of the estimators.

**4.** The suggested combined classes of imputation methods $y^c_{.i_{SA_j}}$, $j = 1, 2, ..., 9$ are superior than the suggested separate classes of imputation methods $y^s_{.i_{SA_j}}$, $j = 1, 2, ..., 9$ in situations I, II and III.

Therefore, due to the dominance of the proposed imputation methods over the existing imputation methods, we recommend them to survey persons for their real life problems.

## APPENDIXES A-B-C. Supplementary file

Supplementary data to this article can be found online.

## ACKNOWLEDGMENTS

The authors are extremely grateful to the learned referee for his valuable comments and Editor-in-Chief.

## REFERENCES


Ahmed, M. S., Al-Titi, O., Al-Rawi, Z., and Abu-Dayyeh, W. (2006). Estimation of a population mean using different imputation methods. *Statistics in Transition*, 7(6):1247–1264.

Al-Omari, A. and Bouza, C. (2014). Ratio estimators of the population mean with missing values using ranked set sampling. *Environmetrics*, 26(2):67–76.

Bhushan, S. and Kumar, A. (2022). Novel log type class of estimators under ranked set sampling. *Sankhya B*, 84:421–447.

Bhushan, S. and Pandey, A. P. (2016). Optimal imputation of missing data for estimation of population mean. *Journal of Statistics and Management System*, 19(6):755–769.

Bhushan, S. and Pandey, A. P. (2018). Optimality of ratio type estimation methods for population mean in presence of missing data. *Communication in Statistics – Theory and Methods*, 47(11):2576–2589.

Bhushan, S., Pandey, A. P., and Pandey, A. (2018). On optimality of imputation methods for estimation of population mean using higher order moments of an auxiliary variable. *Communication in Statistics – Simulation and Computation*, DOI: 10.1080/03610918.2018.1500595

Bhushan, S., Kumar, A., Pandey, A. P., and Singh, S. (2022). Estimation of population mean in presence of missing data under simple random sampling. *Communications in Statistics – Simulation and Computation*, `https://doi.org/10.1080/03610918.2021.2006713`.

Bouza-Herrera, C. N. and Al-Omari, A. I. (2011). Ranked set estimation with imputation of the missing observations: The median estimator. *Revista Investigation Operacional*, 32(1):30–37.

Bouza, C. N., Viada, C., and Vishwakarma, G. K. (2020). Studying the total under missingness by guessing the value of a superpopulation model for imputation. *Revista Investigation Operacional*, 41(7):979–989.

Bouza-Herrera, C. N. and Viada, C. E. (2021). Imputation of individual values of a variable using product predictors. *Revista Investigation Operacional*, 42(3):321–325.

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

Dell, T. R. and Clutter, J. L. (1972). Ranked set sampling theory with order statistics background. *Biometrics*, 28:545–555.

Diana, G. and Perri, P. F. (2010). Improved estimators of the population mean for missing data. *Communication in Statistics – Theory and Methods*, 39:3245–3251.

Heitjan, D. F. and Basu, S. (1996). Distinguishing 'Missing at Random' and 'Missing Completely at Random'. *The American Statistician*, 50:207–213.

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

Khan, L. and Shabbir, J. (2016). Hartley-Ross type unbiased estimators using ranked set sampling and stratified ranked set sampling. *North Carolina Journal of Mathematics and Statistics*, 2:10–22.

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

Linder, D. F., Samawi, H., Yu, L., Chatterjee, A., Huang, Y., and Vogel, R. (2015). On stratified ranked set sampling for regression estimators. *Journal of Applied Statistics*, 42(12):2571–2583.

Mandowara, V. L. and Mehta, N. (2014). Modified ratio estimators using stratified ranked set sampling. *Hacettepe Journal of Mathematics and Statistics*, 43(3):461–471.

McIntyre, G. A. (1952). A method of unbiased selective sampling using ranked set. *Australian Journal Agricultural Research*, 3:385–390.

Mohamed, C., Sedory, S. A., and Singh, S. (2016). Imputation using higher order moments of an auxiliary variable. *Communication in Statistics – Simulation and Computation*, 46:6588–6617.

Prasad, S. (2017). Ratio exponential type estimators with imputation for missing data in sample surveys. *Model Assisted Statistics and Applications*, 12(2):95–106.

Rubin, R. B. (1976). Inference and missing data. *Biometrika*, 63(3):581–592.

Saini, M. and Kumar, A. (2018). Ratio estimators using stratified random sampling and stratified ranked set sampling. *Life Cycle Reliability and Safety Engineering*, 84(5):931–945.

Samawi, H. M. (1996). Stratified ranked set sampling. *Pakistan Journal of Statistics*, 12(1):9–16.

Samawi, H. M. and Siam, M. I. (2003). Ratio estimation using stratified ranked set sampling. *Metron – International Journal of Statistics*, LXI(1):75–90.

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

Singh, S. (2009). A new method of imputation in survey sampling. *Statistics: A Journal of Theoretical and Applied Statistics*, 43(5):499–511.

Singh, S. and Horn, S. (1998). An alternative survey in multi-character survey. *Metrika*, 48:99–107.

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

Singh, S. and Valdes, S. R. (2009). Optimal method of imputation in survey sampling. *Applied Mathematical Sciences*, 3(35):1727–1737.

Sohail, M. U., Shabbir, J., and Ahmed, S. (2018). A class of ratio type estimators for imputing the missing values under rank set sampling. *Journal of Statistical Theory and Practice*, 1–31.

Takahasi, K. and Wakimoto, K. (1968). On unbiased estimates of the population mean based on the sample stratified by means of ordering. *Annals of the Institute of Statistical Mathematics*, 20:1–31.