Revista Colombiana de Estadística - Applied Statistics
January 2025, volume 48, issue 1, pp. 117–155
http://doi.org/10.15446/rce.v48n1.112391

Efficient Factor-Type Estimators of Population Mean in Case of Missing Data and Measurement Error

Estimadores factor tipo de eficientes del de la media de la población en caso de datos faltantes y error de medición

Lovleen Kumar Grover[1,a], Anchal Sharma[2,b]

[1]Department of Mathematics, Guru Nanak Dev University, Amritsar, India
[2]P. G. Department of Mathematics, Khalsa College, Amritsar, India

Abstract

In sample surveys, dealing with missingness in data is one of the most frequent problem that can be handled by replacing missing values with some imputed values. In addition to such missingness, oftenly data provided by respondents are under reported or over reported which results to "Measurement Error". In this paper, we have proposed three modified regression type estimators of population mean, using Factor-Type imputation strategy in two-phase sampling set up to deal with the problem of missing data and measurement error. While proposing our efficient estimators, we have considered two auxiliary variables which have chained correlation with the given study variable. The Bias and Mean Square Error of proposed estimators have been derived up to first order of approximation. The suitable conditions for the superiority of proposed estimators over some existing estimators have been derived. A simulation study is carried out using three artificial data sets to illustrate the supremacy of proposed estimators. Finally, real data set is used to demonstrate the efficiency of proposed estimators in practice.

***Key words*:** Auxiliary variable; Bias; Chain type estimators; Imputation; Correlated measurement error; Mean Square Error; Percent relative efficiency; Simple random sampling; Study variable; Two-phase sampling.

Resumen

En las encuestas por muestreo, lidiar con la falta de datos es uno de los problemas más frecuentes que se pueden manejar reemplazando los valores faltantes con algunos valores imputados. Además de tal falta, a menudo

[a]Ph.D. E-mail: lovleen_2@yahoo.co.in
[b]M.Sc. E-mail: anchalasr@gmail.com

los datos proporcionados por los encuestados no se informan o se informan en exceso, lo que resulta en un error de medición. En este artículo, hemos propuesto tres estimadores de tipo de regresión modificada de la media de la población, utilizando la estrategia de imputación de tipo factor en el muestreo de dos fases establecido para tratar el problema de los datos faltantes y el error de medición. Al proponer nuestros estimadores eficientes, hemos considerado dos variables auxiliares que tienen una correlación encadenada con la variable de estudio dada. El sesgo y el error cuadrático medio de los estimadores propuestos se han derivado hasta el primer orden de aproximación. Se han derivado las condiciones adecuadas para la superioridad de los estimadores propuestos sobre algunos estimadores existentes. Se realiza un estudio de simulación utilizando tres conjuntos de datos artificiales para ilustrar la supremacía de los estimadores propuestos. Finalmente, el conjunto de datos reales se utiliza para demostrar la eficiencia de los estimadores propuestos en la práctica.

***Palabras clave*:** Error cuadrático medio; Eficiencia relativa porcentual; Variable de estudio; Variable auxiliar; Muestreo aleatorio simple; Muestreo bifásico; Sesgo; Estimadores de tipo cadena; Imputación; Error de medición correlacionado.

1. Introduction

Besides sampling errors, in any sampling survey, non-sampling errors which arises due to either non-coverage of units, or unwillingness of respondents etc. poses a serious problem in the estimation of population parameters. The reason of this problem is the incompleteness of data due to which a sample cannot be the representative of population and so the estimators are biased. Missing data can dangerously affect the inferences from randomized clinical trials, if the problem of missing data is not handled properly. Hansen & Hurwitz (1946) were the first researchers to deal with the problem of incomplete samples in mail surveys. Rubin (1976) addressed the following concepts: Missing at Random (MAR), Observed at Random (OAR), Missing Completely at Random (MCAR) and Parameter Distinctness (PD). Sande (1979) and Kalton et al. (1981) suggested imputation methods that make an incomplete data set structurally complete. In the literature of survey sampling, some popular imputation techniques are: Mean imputation, ratio imputation, regression imputation, hot deck imputation, cold deck imputation, nearest neighbor imputation, etc. Researchers have made continuous efforts to devise improved estimators of population mean of the study variable by developing efficient imputation techniques and explore the effects of consequences of incomplete data. In this direction, some well known references are: Singh & Horn (2000), Ahmed et al. (2006), Diana & Francesco Perri (2010), Singh et al. (2010), Seaman et al. (2013), Bhaskaran & Smeeth (2014), Pandey et al. (2015), Singh et al. (2016), Kumar et al. (2017), Doretti et al. (2018), Sohail et al. (2019), Singh et al. (2021), Grover & Sharma (2023), Rehman et al. (2024), Yadav et al. (2024) and many more.

Measurement Errors (ME) arises when the observed value of the sampling units deviates from true values. ME can seriously alter the characteristics of the estimators used to estimate the population parameters of interest and can lead to incorrect conclusions unless some reasonable precautions are applied. Many authors demonstrated the effect of ME under different sampling schemes. The problem of ME on ratio and product estimators under simple random sampling technique was studied by Shalabh (1997). Manisha & Singh (2001) studied the effect of ME on regression type estimator and gave a family of estimators for population mean. Singh et al. (2014) studied class of difference type estimators for estimating population mean in presence of ME. When auxiliary variable is contaminated with ME, then Sahoo et al. (2006) have shown that regression estimator is more sensitive than ratio estimator. Several other authors like Singh & Karpe (2009), Diana & Giordan (2012), Singh & Sharma (2015), Singh & Singh (2017), Singh et al. (2018) and many more studied the impact of ME considering full response and non response from respondents.

In any survey sampling situation, it is tradition to utilize the prior known auxiliary information to improve the precision of an estimator of parameter of interest either at the stage of planning or at the stage of designing or at the estimation stage or combination of these stages. But, if such prior auxiliary information is unknown, then the concept of two phase sampling plan is utilized to obtain the requisite estimates of population parameters. Sometimes, information of another auxiliary variable which is highly correlated with the earlier auxiliary variable is easily available at lower cost. Under such circumstances, Chand (1975), introduced the concept of chain estimators. After this, Singh et al. (1994), Al-Jararha & Ahmed (2002), Kumar & Bahl (2006), Choudhury & Singh (2012), Kumar & Sharma (2020), Mehta & Tailor (2020) have judiciously used known functions of auxiliary variables using such chaining technique. Shukla et al. (2009) suggested Factor-Type (F-T) estimators of population mean when observations on study variable are missing. Singh et al. (2015) suggested one parameter family of F-T estimators using one auxiliary variable only. Authors like Pandey et al. (2016), Audu & Adewara (2017), Thakur & Shukla (2022) have made contributions in improving efficiency of estimator of population mean of research variable using F-T imputation strategies.

Recently, Bhushan et al. (2023), Tiwari et al. (2023), Kumar et al. (2024) and Sajjad & Ismail (2024) considered the problem of estimation of population mean in survey sampling under the situation of missing data along with the impact of ME. Most of the authors considered only the case of uncorrelated ME corresponding to various available variables. But, such ME corresponding to various available variables maybe correlated and neglection of their correction can lead to serious false inferences. Some of the authors like Shalabh & Tsai (2017), Kumar et al. (2023), and Vishwakarma et al. (2020) have considered the situation of correlated ME in their study. The motive of this paper is to propose efficient and bias-controllable F-T estimators for population mean of study variable using chaining technique as suggested by Thakur & Shukla (2022) for two auxiliary variables under the presumption that population mean of second auxiliary variable is known and the population mean of first auxiliary variable is unknown. While proposing these efficient F-T estimators, we have considered that some observation of study variable are missing which are imputed by our proposed strategy.

The work of this paper is organized as follows: In Section 2, the sampling strategy, methodology and notations are given, followed by Section 3 where the existing imputation methods are discussed considering correlated Measurement errors. Their biases and mean square errors have been expressed using the notations as mentioned in Section 2. Further, taking motivation from Thakur & Shukla (2022), we have proposed the algorithm for imputation technique and finally proposed three point estimators for population mean $\overline{Y}$ of study variable, followed by some Theorems on Biases and Mean Square Error (MSE) of these estimators in Section 4. An efficiency comparison is made between the proposed estimators and the existing ones in Section 5. In Section 6, the results of earlier sections have been derived for uncorrelated ME. The results obtained from simulation process and real data set are discussed in Section 7, followed by a discussion and findings in Section 8.

2. Methodology and Notations

Suppose a finite population $U=\{U_1, U_2, \ldots, U_N\}$ consists of N identifiable units and i^{th} unit U_i characterized by triplet (Y, X, Z), where Y, X, Z are study, first and second auxiliary variables respectively. We assume that variables Y and X are highly correlated while Y and Z are remotely correlated. Moreover, the correlation between two auxiliary variables X and Z is sufficiently significant. A first phase sample S' of size m is drawn using Simple Random Sampling Without Replacement (SRSWOR) from the population. The measurements of variables (X, Z) are taken on this first phase sample. Suppose a second phase sample S of size n is drawn from S' using again SRSWOR. The measurements of variables (Y, X, Z) are taken on this second phase sample. Let R be the subset of S consisting of r responding units and R^c be the subset of S consisting of $(n-r)$ non-responding units. Noting that $R \cup R^c = S$. So, assuming that information of variables (Y, X, Z) is available for R but information of only variables (X, Z) is available for R^c i.e. values of Y are missing on R^c. Let (y_i, x_i, z_i) be the observed values and (Y_i, X_i, Z_i) be the true values of the associated variables (Y, X, Z) on the i^{th} unit in the population. Suppose measurement errors are: $U_i = y_i - Y_i$, $V_i = x_i - X_i$, $W_i = z_i - Z_i$: $i = 1, 2, \ldots, N$. Note that, these measurement errors are assumed to be random in nature.

The following notations are used in this paper: $y_{.i}$: value of the variable Y on the i^{th} unit of second phase sample.

$\overline{Y}$, $\overline{X}$, $\overline{Z}$: population means of the variables (Y, X, Z).

$\overline{Y}_r$, $\overline{x}_r$, $\overline{z}_r$: sample means of the respective variables based on the responding units of second phase sample.

$\overline{x}_m$, $\overline{z}_m$: sample means of the respective variables based on the first phase sample.

ρ_{YX}, ρ_{YZ}, ρ_{XZ}, ρ_{UV}, ρ_{VW}, ρ_{UW}: population correlation coefficients between the variables as shown in subscripts.

β_{YX}, β_{YZ}, β_{XZ}: Population regression coefficients of the variables as shown in subscripts.

S_Y^2, S_X^2, S_Z^2, S_U^2, S_V^2, S_W^2: Population variances (with divisors $(N-1)$) of the variables as shown in subscripts.

S_{YX}, S_{YZ}, S_{XZ}: population covariances (with divisors $(N-1)$) between the variables as shown in subscripts.

C_Y, C_X, C_Z, C_U, C_V, C_W: population Coefficients of variations of the variables as shown in subscripts.

$K_{YX}=\rho_{YX}\frac{C_Y}{C_X}$; $K_{YZ}=\rho_{YZ}\frac{C_Y}{C_Z}$; $K_{XZ}=\rho_{XZ}\frac{C_X}{C_Z}$. Also define, $\lambda_1=\frac{1}{r}-\frac{1}{N}$, $\lambda_2=\frac{1}{m}-\frac{1}{N}$, $\lambda_3=\lambda_1-\lambda_2=\frac{1}{r}-\frac{1}{m}$, $f=\frac{m}{N}$. Again for the purpose of various required expectations used in the derivations, we are taking the following notations s.t. $|e_i|<1\ \forall\ i=1,2,\ldots,5$. $\overline{Y}_r=\overline{Y}(1+e_1)$, $\overline{x}_r=\overline{X}(1+e_2)$, $\overline{x}_m=\overline{X}(1+e_3)$, $\overline{z}_r=\overline{Z}(1+e_4)$, $\overline{z}_m=\overline{Z}(1+e_5)$.

Under the influence of Measurement Errors, corresponding to variables (Y,X,Z) and using SRSWOR, we have the following expectations:

$$E(e_i)=0 \ \text{ for } \ i=1,2,\ldots,5.$$

$$E(e_1^2)=\lambda_1 C_Y^2\left(1+\frac{S_U^2}{S_Y^2}\right)=\lambda_1 C_{YM}^2, E(e_2^2)=\lambda_1 C_X^2\left(1+\frac{S_V^2}{S_X^2}\right)=\lambda_1 C_{XM}^2,$$

$$E(e_3^2)=\lambda_2 C_X^2\left(1+\frac{S_V^2}{S_X^2}\right)=\lambda_2 C_{XM}^2, \qquad E(e_4^2)=\lambda_1 C_Z^2\left(1+\frac{S_W^2}{S_Z^2}\right)=\lambda_1 C_{ZM}^2,$$

$$E(e_5^2)=\lambda_2 C_Z^2\left(1+\frac{S_W^2}{S_Z^2}\right)=\lambda_2 C_{ZM}^2$$

where $C_{YM}^2=C_Y^2\left(1+\frac{S_U^2}{S_Y^2}\right)$, $C_{XM}^2=C_X^2\left(1+\frac{S_V^2}{S_X^2}\right)$, $C_{ZM}^2=C_Z^2\left(1+\frac{S_W^2}{S_Z^2}\right)$. Since the given variables (Y,X,Z) are correlated to each other, therefore the variables corresponding to the ME of these variables i.e. (U,V,W) may also be correlated. Thus, under the situation of correlated ME we have the following expectations.

$$E(e_1e_2)=\lambda_1\left(\rho_{YX}C_YC_X+\frac{\rho_{UV}S_US_V}{\overline{Y}\overline{X}}\right)=\lambda_1 C_{YXM}$$

$$E(e_1e_3)=\lambda_2\left(\rho_{YX}C_YC_X+\frac{\rho_{UV}S_US_V}{\overline{Y}\overline{X}}\right)=\lambda_2 C_{YXM}$$

$$E(e_1e_4)=\lambda_1\left(\rho_{YZ}C_YC_Z+\frac{\rho_{UW}S_US_W}{\overline{Y}\overline{Z}}\right)=\lambda_1 C_{YZM}$$

$$E(e_1e_5)=\lambda_2\left(\rho_{YZ}C_YC_Z+\frac{\rho_{UW}S_US_W}{\overline{Y}\overline{Z}}\right)=\lambda_2 C_{YZM}$$

$$E(e_2e_3) = \lambda_2 C_{XM}^2, \qquad E(e_2e_4) = \lambda_1 \left(\rho_{XZ} C_X C_Z + \frac{\rho_{VW} C_V C_W}{\overline{X}\overline{Z}}\right) = \lambda_1 C_{XZM}$$

$$E(e_2e_5) = E(e_3e_4) = E(e_3e_5) = \lambda_2 \left(\rho_{XZ} C_X C_Z + \frac{\rho_{VW} C_V C_W}{\overline{X}\overline{Z}}\right) = \lambda_2 C_{XZM}$$

$E(e_4e_5) = E(e_5^2)$ where $C_{YXM} = \rho_{YX} C_Y C_X + \frac{\rho_{UV} S_U S_V}{\overline{Y}\overline{X}}$, $C_{YZM} = \rho_{YZ} C_Y C_Z + \frac{\rho_{UW} S_U S_W}{\overline{Y}\overline{Z}}$, $C_{XZM} = \rho_{XZ} C_X C_Z + \frac{\rho_{VW} C_V C_W}{\overline{X}\overline{Z}}$. Further taking, $K_{YXM} = \frac{C_{YXM}}{C_{XM}^2}$, $K_{YZM} = \frac{C_{YZM}}{C_{ZM}^2}$, $K_{XZM} = \frac{C_{XZM}}{C_{ZM}^2}$.

To consider the general set up, we have yet considered the situation of correlated ME. In the later section i.e. Section 6, we will also consider the situation of uncorrelated ME as the special cases of these generalized situation of correlated ME.

3. Review of Relevant Existing Estimators in the Literature Under the Impact of ME

In the present section, a review of some existing imputation strategies for $\overline{Y}$ have been given which have direct relevance with missingness in data related to study variable. For comparison purposes, we have adapted these existing estimators and then derived their biases and MSE in presence of correlated ME up to first order of approximation.

3.1. Simple Mean Method of Imputation

The imputation scheme under mean method of imputation is

$$y_{.i} = \begin{cases} y_i & i \in R \\ \overline{Y}_r & i \in R^c \end{cases}$$

The corresponding point estimator for $\overline{Y}$ is:

$$\overline{Y}_{MN} = \overline{Y}_r$$

The Bias and Variance of $\overline{Y}_{MN}$, in presence of correlated ME, are respectively given as:

$$B(\overline{Y}_{MN}) = 0$$

$$Var(\overline{Y}_{MN}) = \lambda_1 \overline{Y}^2 C_{YM}^2$$

3.2. Ratio Method of Imputation

Now, on reformulating idea of Singh & Horn (2000), under the two phase sampling set up and in the presence of only first auxiliary variable X, assuming that the value of $\overline{X}$ is unknown, the imputation scheme is:

$$y_{.i} = \begin{cases} y_i & i \in R \\ \hat{b}x_i & i \in R^c \end{cases}$$

where $\hat{b} = \frac{\sum_{i \in R} y_i}{\sum_{i \in R} x_i}$. The corresponding point estimator for $\overline{Y}$ is:

$$\overline{Y}_R = \overline{Y}_r \frac{\overline{x}_m}{\overline{x}_r}$$

Up to first order of approximation, the Bias and MSE of this estimator, in presence of correlated ME, are respectively:

$$B(\overline{Y}_R) = \overline{Y}\lambda_3(C^2_{XM} - C_{YXM})$$

$$MSE(\overline{Y}_R) = \overline{Y}^2[\lambda_1 C^2_{YM} + \lambda_3(C^2_{XM} - 2C_{YXM})]$$

3.3. Compromised Method of Imputation

Under the two phase sampling set up and presence of only first auxiliary variable X, assuming that the value of $\overline{X}$ is unknown, now on following Singh & Horn (2000), the imputation scheme is:

$$y_{.i} = \begin{cases} y_i & i \in R \\ \frac{1}{n-r}\left(n\left(\alpha\overline{Y}_r + (1-\alpha)\overline{Y}_r\frac{\overline{x}_m}{\overline{x}_r}\right) - r\overline{Y}_r\right) & i \in R^c \end{cases}$$

where α is a suitable constant. The corresponding point estimator of $\overline{Y}$ is:

$$\overline{Y}_{CI} = \alpha\overline{Y}_r + (1-\alpha)\overline{Y}_r\frac{\overline{x}_m}{\overline{x}_r}$$

Up to first order of approximation, the $Bias$ and MSE of this estimator, in presence of correlated ME, are respectively:

$$B(\overline{Y}_{CI}) = (1-\alpha)\overline{Y}\lambda_3(C^2_{XM} - C_{YXM})$$

$$MSE_{\min}(\overline{Y}_{CI}) = \overline{Y}^2\left(\lambda_1 C^2_{YM} - \lambda_3\frac{C^2_{YXM}}{C^2_{XM}}\right)$$

3.4. Factor-Type Imputation

Under the two phase sampling set up and presence of only first auxiliary variable X, assuming that the value of $\overline{X}$ is unknown, now on following Shukla et al. (2009) Factor-type imputation techniques are:

$$y_{.i} = \begin{cases} y_i & i \in R \\ \frac{\overline{Y}_r}{n-r}[nd_1(k) - r] & i \in R^c \end{cases}$$

and

$$y_{.i} = \begin{cases} y_i & i \in R \\ \frac{\overline{Y}_r}{n-r}[nd_2(k) - r] & i \in R^c \end{cases}$$

where $d_1(k) = \frac{(A+C)\overline{x}_m+fB\overline{x}_n}{(A+fB)\overline{x}_m+C\overline{x}_n}$ and $d_2(k) = \frac{(A+C)\overline{x}_m+fB\overline{x}_r}{(A+fB)\overline{x}_m+C\overline{x}_r}$. Here, we have $A = (k-1)(k-2)$; $B = (k-1)(k-4)$; $C = (k-2)(k-3)(k-4)$ and k is any real constant such that $0 < k < \infty$.

The corresponding point estimators for $\overline{Y}$ are:

$$\overline{Y}_{D_1(k)} = \overline{Y}_r \left[\frac{(A+C)\overline{x}_m + fB\overline{x}_n}{(A+fB)\overline{x}_m + C\overline{x}_n}\right]$$

$$\overline{Y}_{D_2(k)} = \overline{Y}_r \left[\frac{(A+C)\overline{x}_m + fB\overline{x}_r}{(A+fB)\overline{x}_m + C\overline{x}_r}\right]$$

Up to first order of approximation, the respective Biases and MSE of these estimators, in the presence of correlated ME are:

$$B(\overline{Y}_{D_1(k)}) = \psi\overline{Y}\left(\frac{1}{n} - \frac{1}{m}\right)(\psi_2 C^2_{XM} - C_{YXM})$$

$$B(\overline{Y}_{D_2(k)}) = (\psi + \psi_2\psi_4)\lambda_2 C^2_{XM} - \psi_2\psi\lambda_1 C^2_{XM} - (\psi_4\lambda_2 + \psi_2\lambda_1)C_{YXM}$$

where $\psi_1 = \frac{fB}{A+fB+C}$, $\psi_2 = \frac{C}{A+fB+C}$, $\psi_3 = \frac{A+C}{A+fB+C}$, $\psi_4 = \frac{A+fB}{A+fB+C}$, $\psi_1 + \psi_3 = \psi_2 + \psi_4 = 1$, $\psi = \psi_1 - \psi_2$, and

$$MSE(\overline{Y}_{D_1(k)}) = \overline{Y}^2\left(\lambda_1 C^2_{YM} - \left(\frac{1}{n} - \frac{1}{m}\right)\frac{C^2_{YXM}}{C^2_{XM}}\right)$$

$$MSE(\overline{Y}_{D_2(k)}) = \overline{Y}^2\left(\lambda_1 C^2_{YM} - \lambda_3\frac{C^2_{YXM}}{C^2_{XM}}\right)$$

For the derivation of the above results, refer to Appendix A.1.

3.5. Factor Type Chained Imputation

Under the two phase sampling set up and in the presence of both first and second auxiliary variables X and Z, assuming that $\overline{X}$ is unknown but $\overline{Z}$ is known, now on following, Thakur & Shukla (2022) Factor-type chained imputation techniques are:

$$y_{.i} = \begin{cases} y_i & i \in R \\ \frac{\overline{Y}_r}{n-r}[nt_1(k) - r] & i \in R^c \end{cases}$$

$$y_{.i} = \begin{cases} y_i & i \in R \\ \frac{\overline{Y}_r}{n-r}[nt_2(k) - r] & i \in R^c \end{cases}$$

$$y_{.i} = \begin{cases} y_i & i \in R \\ \frac{\overline{Y}_r}{n-r}[nt_3(k) - r] & i \in R^c \end{cases}$$

where $t_1(k) = \frac{\overline{x}_m}{\overline{x}_r}\left[\frac{(A+C)\overline{Z}+fB\overline{z}_m}{(A+fB)\overline{Z}+C\overline{z}_m}\right]$, $t_2(k) = \frac{\overline{x}_m}{\overline{x}_r}\left[\frac{(A+C)\overline{z}_m+fB\overline{z}_r}{(A+fB)\overline{z}_r+C\overline{z}_r}\right]$ and $t_3(k) = \frac{\overline{x}_m}{\overline{x}_r}\left[\frac{(A+C)\overline{Z}+fB\overline{z}_r}{(A+fB)\overline{Z}+C\overline{z}_r}\right]$. The corresponding point estimators for $\overline{Y}$ are:

$$\overline{Y}_{T_1(k)} = \overline{Y}_r\frac{\overline{x}_m}{\overline{x}_r}\left[\frac{(A+C)\overline{Z} + fB\overline{z}_m}{(A+fB)\overline{Z} + C\overline{z}_m}\right]$$

$$\overline{Y}_{T_2(k)} = \overline{Y}_r \frac{\overline{x}_m}{\overline{x}_r} \left[\frac{(A+C)\overline{z}_m + fB\overline{z}_r}{(A+fB)\overline{z}_m + C\overline{z}_r} \right]$$

$$\overline{Y}_{T_3(k)} = \overline{Y}_r \frac{\overline{x}_m}{\overline{x}_r} \left[\frac{(A+C)\overline{Z} + fB\overline{z}_r}{(A+fB)\overline{Z} + C\overline{z}_r} \right]$$

Up to first order of approximation, the respective Biases and MSE of these estimators, in presence of correlated ME, are:

$$B(\overline{Y}_{T_1(k)}) = \overline{Y} \left[\lambda_3 C^2_{XM} \left(1 - K_{YXM}\right) - \psi\lambda_2 C^2_{ZM} \left(\psi_2 - K_{YZM}\right) \right]$$

$$B(\overline{Y}_{T_2(k)}) = \overline{Y}\lambda_3 \left[C^2_{XM} \left(1 - K_{YXM}\right) - \psi C^2_{ZM} \left(\psi_2 - K_{YZM} + K_{XZM}\right) \right]$$

$$B(\overline{Y}_{T_3(k)}) = \overline{Y} \left[\lambda_3 C^2_{XM} \left(1 - K_{YXM}\right) + \psi C^2_{ZM} \left(-\lambda_1\psi_2 + \lambda_1 K_{YZM} - \lambda_3 K_{XZM}\right) \right]$$

and

$$MSE(\overline{Y}_{T_1(k)}) = \overline{Y}^2 \left[\lambda_1 C^2_{YM} + \lambda_3 C^2_{XM} - 2\lambda_3 C_{YXM} - \lambda_2 \frac{C^2_{YZM}}{C^2_{ZM}} \right]$$

$$MSE(\overline{Y}_{T_2(k)}) = \overline{Y}^2 \left[\lambda_1 C^2_{YM} + \lambda_3 C^2_{XM} - 2\lambda_3 C_{YXM} - \lambda_3 \frac{(C_{XZM} - C_{YZM})^2}{C^2_{ZM}} \right]$$

$$MSE(\overline{Y}_{T_3(k)}) = \overline{Y}^2 \left[\lambda_1 C^2_{YM} + \lambda_3 C^2_{XM} - 2\lambda_3 C_{YXM} - \frac{(\lambda_3 C_{XZM} - \lambda_1 C_{YZM})^2}{\lambda_1 C^2_{ZM}} \right]$$

For the justification of the above expressions, we give the derivation of the biases and MSEs of the above estimators in Appendix A.2.

4. Proposed Imputation Techniques and Corresponding Proposed Estimators Along with Their Biases and MSEs

We consider the same situation of Thakur & Shukla (2022) i.e., under the two phase sampling set up and in the presence of both first and second auxiliary variables X and Z, assuming that $\overline{X}$ is unknown but $\overline{Z}$ is known. So, motivated with the recent work of Thakur & Shukla (2022), we have extended their idea and proposed the following three imputation set ups:

$$y_{.1i} = \begin{cases} y_i & i \in R \\ \frac{1}{n-r}[n\phi_1(k) - r\overline{Y}_r] & i \in R^c \end{cases}$$

$$y_{.2i} = \begin{cases} y_i & i \in R \\ \frac{1}{n-r}[n\phi_2(k) - r\overline{Y}_r] & i \in R^c \end{cases}$$

$$y_{.3i} = \begin{cases} y_i & i \in R \\ \frac{1}{n-r}[n\phi_3(k) - r\overline{Y}_r] & i \in R^c \end{cases}$$

where

$$\phi_1(k) = \overline{Y}_r \frac{\overline{x}_m}{\overline{x}_r}\left[\frac{(A+C)\overline{Z}+fB\overline{z}_m}{(A+fB)\overline{Z}+C\overline{z}_m}\right] + \alpha_1\left(1-\frac{\overline{x}_r}{\overline{x}_m}\right)$$

$$\phi_2(k) = \overline{Y}_r \frac{\overline{x}_m}{\overline{x}_r}\left[\frac{(A+C)\overline{z}_m+fB\overline{z}_r}{(A+fB)\overline{z}_m+C\overline{z}_r}\right] + \alpha_2\left(1-\frac{\overline{x}_r}{\overline{x}_m}\right)$$

$$\phi_3(k) = \overline{Y}_r \frac{\overline{x}_m}{\overline{x}_r}\left[\frac{(A+C)\overline{Z}+fB\overline{z}_r}{(A+fB)\overline{Z}+C\overline{z}_r}\right] + \alpha_3\left(1-\frac{\overline{x}_r}{\overline{x}_m}\right)$$

$\alpha_i; i = 1,2,3$ are suitable chosen constants respectively.

The corresponding proposed imputed point estimators of $\overline{Y}$ are:

$$\overline{Y}_{\phi_1(k)} = \overline{Y}_r \frac{\overline{x}_m}{\overline{x}_r}\left[\frac{(A+C)\overline{Z}+fB\overline{z}_m}{(A+fB)\overline{Z}+C\overline{z}_m}\right] + \alpha_1\left(1-\frac{\overline{x}_r}{\overline{x}_m}\right) \tag{1}$$

$$\overline{Y}_{\phi_2(k)} = \overline{Y}_r \frac{\overline{x}_m}{\overline{x}_r}\left[\frac{(A+C)\overline{z}_m+fB\overline{z}_r}{(A+fB)\overline{z}_m+C\overline{z}_r}\right] + \alpha_2\left(1-\frac{\overline{x}_r}{\overline{x}_m}\right) \tag{2}$$

$$\overline{Y}_{\phi_3(k)} = \overline{Y}_r \frac{\overline{x}_m}{\overline{x}_r}\left[\frac{(A+C)\overline{Z}+fB\overline{z}_r}{(A+fB)\overline{Z}+C\overline{z}_r}\right] + \alpha_3\left(1-\frac{\overline{x}_r}{\overline{x}_m}\right) \tag{3}$$

Theorem 1. *The Biases of proposed estimators* $\bar{y}_{\phi_i(k)}$ $(i = 1,2,3)$, *derived up to* $o(n^{-1})$ *are respectively:*

$$B(\overline{Y}_{\phi_1(k)}) = \overline{Y}[\lambda_3 C^2_{XM} - \lambda_3 C_{YXM} - \psi\psi_2\lambda_2 C^2_{ZM} + \psi\lambda_2 C_{YZM}] \tag{4}$$

$$B(\overline{Y}_{\phi_2(k)}) = \overline{Y}\lambda_3[C^2_{XM} - C_{YXM} - \psi\psi_2 C^2_{ZM} + \psi C_{YZM} - \psi C_{XZM}] \tag{5}$$

$$B(\overline{Y}_{\phi_3(k)}) = \overline{Y}[\lambda_3 C^2_{XM} - \lambda_3 C_{YXM} - \lambda_1\psi\psi_2 C^2_{ZM} + \lambda_1\psi C_{YZM} - \lambda_3\psi C_{XZM}] \tag{6}$$

where $\psi_1 = \frac{fB}{A+fB+C}$, $\psi_2 = \frac{C}{A+fB+C}$ *and* $\psi = \psi_1 - \psi_2$

Proof. For the derivation of above results, refer to Appendix A.3. □

Theorem 2. *The MSE of proposed estimators* $\bar{y}_{\phi_i(k)}$ $(i = 1,2,3)$, *derived up to* $o(n^{-1})$ *are respectively:*

$$\begin{gathered} MSE(\overline{Y}_{\phi_1(k)}) = \overline{Y}^2(\lambda_1 C^2_{YM} + \lambda_2\psi^2 C^2_{ZM} + \lambda_3 C^2_{XM} + 2\psi\lambda_2 C_{YZM} - 2\lambda_3 C_{YXM}) + \\ \alpha_1^2\lambda_3 C^2_{XM} + 2\alpha_1\overline{Y}(C^2_{XM} - C_{YXM}) \end{gathered} \tag{7}$$

$$\begin{gathered} MSE(\overline{Y}_{\phi_2(k)}) = \overline{Y}^2(\lambda_1 C^2_{YM} + \lambda_3\psi^2 C^2_{Z} + \lambda_3 C^2_{X} + 2\psi\lambda_3 C_{YZM} - 2\lambda_3 C_{YXM} \\ -2\psi\lambda_3 C_{XZM}) + \alpha_2^2\lambda_3 C^2_{XM} + 2\alpha_2\overline{Y}\lambda_3(C^2_{XM} - C_{YXM} - \psi C_{XZM}) \end{gathered} \tag{8}$$

$$\begin{gathered} MSE(\overline{Y}_{\phi_3(k)}) = \overline{Y}^2(\lambda_1 C^2_{YM} + \lambda_1\psi^2 C^2_{ZM} + \lambda_3 C^2_{XM} + 2\psi\lambda_1 C_{YZM} - 2\lambda_3 C_{YXM} \\ -2\psi\lambda_3 C_{XZM}) + \alpha_3^2\lambda_3 C^2_{XM} + 2\alpha_3\overline{Y}\lambda_3(C^2_{XM} - C_{YXM} - \psi C_{XZM}) \end{gathered} \tag{9}$$

Proof. For the derivation of above results, refer to Appendix A.4. □

Theorem 3. *The minimum MSE of proposed estimators $\bar{y}_{\phi_i(k)}$ $(i = 1, 2, 3)$, derived up to $o(n^{-1})$ are respectively:*

$$MSE_{\min}(\overline{Y}_{\phi_1(k)}) = \overline{Y}^2 \left[\lambda_1 C^2_{YM} - \lambda_2 \frac{C^2_{YZM}}{C^2_{ZM}} - \lambda_3 \frac{C^2_{YXM}}{C^2_{XM}}\right] \tag{10}$$

$$MSE_{\min}(\overline{Y}_{\phi_2(k)}) = \overline{Y}^2 \left[\lambda_1 C^2_{YM} - \lambda_3 \frac{C^2_{YXM}}{C^2_{XM}} - \lambda_3 \frac{(C_{XZM}C_{YXM} - C_{YZM}C^2_{XM})^2}{(C^2_{XM}C^2_{ZM} - C^2_{XZM})C^2_{XM}}\right] \tag{11}$$

$$MSE_{\min}(\overline{Y}_{\phi_3(k)}) = \overline{Y}^2 \left[\lambda_1 C^2_{YM} - \lambda_3 \frac{C^2_{YXM}}{C^2_{XM}} - \frac{(\lambda_3 C_{YXM}C_{XZM} - \lambda_1 C_{YZM}C^2_{XM})^2}{(\lambda_1 C^2_{XM}C^2_{ZM} - \lambda_3 C^2_{XZM})C^2_{XM}}\right] \tag{12}$$

Proof. For the derivation of above results, refer to Appendix A.5. □

Remark 1. For $\alpha_i = 0$, $\overline{Y}_{\phi_i(k)} = \overline{Y}_{T_i(k)} (i = 1, 2, 3)$ i.e. our proposed estimators become the same estimators as suggested by Thakur & Shukla (2022).

Remark 2. About choices of k.

For the three proposed estimators of $\overline{Y}$, the optimality conditions of ψ are

$$\psi_{opt} = C_i \ (i = 1, 2, 3) \text{ (see Appendix A.5)}.$$

Reddy (1978) has proved that the values of population parameters, like C_Y, C_X, C_Z, ρ_{YX}, ρ_{XZ}, ρ_{YZ} are stable over moderate period of time and can be obtained in advance from past data. Also, value of f can be known a priori. Above equations are polynomials (with real coefficients) in k of degree three and thus can be solved for values of k. Since, these equations are cubic in nature so it will have at least one real root. The choice of k should be done in such a way that bias of the corresponding proposed estimator become minimum.

Remark 3. About almost unbiased proposed estimators.

The Bias of $\overline{Y}_{\phi_i(k)} (i = 1, 2, 3)$ can be made zero, up to first order of approximation, using expressions (4), (5), (6) after putting optimum values of ψ in these expressions (for details see Thakur & Shukla, 2022). By doing so, we get three equations of degree 3 in k. These will provide multiple choices of k on which bias of corresponding proposed estimators is zero. The best choice of k among these values is that which will provide minimum MSE of the corresponding estimator. Here, $B(\overline{Y}_{\phi_i(k)}) = 0$ $(i = 1, 2, 3)$ will be approximately zero since biases have been obtained up to first order of approximation only. So, in this way the proposed estimators become almost unbiased in such situation.

5. Theoretical Comparison of Proposed estimators

In this section, MSE of the proposed estimators have been compared mathematically with all the existing estimators which have been considered in Section 3.

5.1. For Proposed Estimator $\overline{Y}_{\phi_1(k)}$

We considered the differences of MSE of the existing and proposed estimator $\overline{Y}_{\phi_1(k)}$ for the purpose of comparison as given below:

$$Var(\overline{Y}_{MN}) - MSE_{\min}(\overline{Y}_{\phi_1(k)}) = \overline{Y}^2\left[\lambda_3\frac{C^2_{YXM}}{C^2_{XM}} + \lambda_2\frac{C^2_{YZM}}{C^2_{ZM}}\right] > 0, \text{ always.}$$

$$MSE(\overline{Y}_R) - MSE_{\min}(\overline{Y}_{\phi_1(k)}) = \overline{Y}^2\left[\lambda_3\frac{(C^2_{XM} - C_{YXM})^2}{C^2_{XM}} + \lambda_2\frac{C^2_{YZM}}{C^2_{ZM}}\right] > 0, \text{ always.}$$

$$MSE(\overline{Y}_{CI}) - MSE_{\min}(\overline{Y}_{\phi_1(k)}) = \overline{Y}^2\left[\lambda_2\frac{C^2_{YZM}}{C^2_{ZM}}\right] > 0, \text{ always.}$$

$$MSE(\overline{Y}_{D_1(k)}) - MSE_{\min}(\overline{Y}_{\phi_1(k)}) = \overline{Y}^2\left[\left(\frac{1}{r} - \frac{1}{n}\right)\frac{C^2_{YXM}}{C^2_{XM}} + \lambda_2\frac{C^2_{YZM}}{C^2_{ZM}}\right] > 0, \text{ always.}$$

$$MSE(\overline{Y}_{D_2(k)}) - MSE_{\min}(\overline{Y}_{\phi_1(k)}) = \overline{Y}^2\left[\lambda_2\frac{C^2_{YZM}}{C^2_{ZM}}\right] > 0, \text{ always.}$$

$$MSE(\overline{Y}_{T_1(k)}) - MSE_{\min}(\overline{Y}_{\phi_1(k)}) = \overline{Y}^2\left[\lambda_3\frac{(C^2_{XM} - C_{YXM})^2}{C^2_{XM}}\right] > 0, \text{ always.}$$

$$MSE(\overline{Y}_{T_2(k)}) - MSE_{\min}(\overline{Y}_{\phi_1(k)}) = \overline{Y}^2\left[\lambda_3\frac{(C^2_{XM} - C_{YXM})^2}{C^2_{XM}} + \lambda_2\frac{C^2_{YZM}}{C^2_{ZM}} - \lambda_3\frac{(C_{XZM} - C_{YZM})^2}{C^2_{ZM}}\right] > 0,$$

if

$$\lambda_3\frac{(C^2_{XM} - C_{YXM})^2}{C^2_{XM}} + \lambda_2\frac{C^2_{YZM}}{C^2_{ZM}} - \lambda_3\frac{(C_{XZM} - C_{YZM})^2}{C^2_{ZM}} > 0.$$

$$MSE(\overline{Y}_{T_3(k)}) - MSE_{\min}(\overline{Y}_{\phi_1(k)}) = \overline{Y}^2\left[\lambda_3\frac{(C^2_{XM} - C_{YXM})^2}{C^2_{XM}} + \lambda_2\frac{C^2_{YZM}}{C^2_{ZM}} - \frac{(\lambda_3 C_{XZM} - \lambda_1 C_{YZM})^2}{\lambda_1 C^2_{ZM}}\right] > 0,$$

if

$$\lambda_3\frac{(C^2_{XM} - C_{YXM})^2}{C^2_{XM}} + \lambda_2\frac{C^2_{YZM}}{C^2_{ZM}} - \frac{(\lambda_3 C_{XZM} - \lambda_1 C_{YZM})^2}{\lambda_1 C^2_{ZM}} > 0.$$

5.2. For Proposed Estimator $\overline{Y}_{\phi_2(k)}$

We considered the differences of MSE of the existing and proposed estimator $\overline{Y}_{\phi_2(k)}$ for the purpose of comparison as given below:

$$Var(\overline{Y}_{MN}) - MSE_{\min}(\overline{Y}_{\phi_2(k)}) = \overline{Y}^2\left[\lambda_3\frac{C^2_{YXM}}{C^2_{XM}} + \lambda_3\frac{(C_{XZM}C_{YXM} - C_{YZM}C^2_{XM})^2}{(C^2_{XM}C^2_{ZM} - C^2_{XZM})C^2_{XM}}\right] > 0, \text{always.}$$

$$MSE(\overline{Y}_R) - MSE_{\min}(\overline{Y}_{\phi_2(k)}) = \overline{Y}^2\lambda_3\left[\frac{(C^2_{XM} - C_{YXM})^2}{C^2_{XM}} + \frac{(C_{XZM}C_{YXM} - C_{YZM}C^2_{XM})^2}{(C^2_{XM}C^2_{ZM} - C^2_{XZM})C^2_{XM}}\right] > 0, \text{ always.}$$

$$MSE(\overline{Y}_{CI}) - MSE_{\min}(\overline{Y}_{\phi_2(k)}) = \overline{Y}^2\left[\lambda_3\frac{(C_{XZM}C_{YXM} - C_{YZM}C^2_{XM})^2}{(C^2_{XM}C^2_{ZM} - C^2_{XZM})C^2_{XM}}\right] > 0, \text{always.}$$

$$MSE(\overline{Y}_{D_1(k)}) - MSE_{\min}(\overline{Y}_{\phi_2(k)}) = \overline{Y}^2\left[\left(\frac{1}{r} - \frac{1}{n}\right)\frac{C^2_{YXM}}{C^2_{XM}} + \lambda_3\frac{(C_{XZM}C_{YXM} - C_{YZM}C^2_{XM})^2}{(C^2_{XM}C^2_{ZM} - C^2_{XZM})C^2_{XM}}\right] > 0, \text{ always.}$$

$$MSE(\overline{Y}_{D_2(k)}) - MSE_{\min}(\overline{Y}_{\phi_2(k)}) = \overline{Y}^2\left[\lambda_3\frac{(C_{XZM}C_{YXM} - C_{YZM}C^2_{XM})^2}{(C^2_{XM}C^2_{ZM} - C^2_{XZM})C^2_{XM}}\right] > 0, \text{always.}$$

$$MSE(\overline{Y}_{T_1(k)}) - MSE_{\min}(\overline{Y}_{\phi_2(k)}) = \overline{Y}^2\left[\lambda_3\frac{(C^2_{XM} - C_{YXM})^2}{C^2_{XM}} + \lambda_3\frac{(C_{XZM}C_{YXM} - C_{YZM}C^2_{XM})^2}{(C^2_{XM}C^2_{ZM} - C^2_{XZM})C^2_{XM}} - \lambda_2\frac{C^2_{YZM}}{C^2_{ZM}}\right] > 0,$$

if

$$\lambda_3\frac{(C^2_{XM} - C_{YXM})^2}{C^2_{XM}} + \lambda_3\frac{(C_{XZM}C_{YXM} - C_{YZM}C^2_{XM})^2}{(C^2_{XM}C^2_{ZM} - C^2_{XZM})C^2_{XM}} - \lambda_2\frac{C^2_{YZM}}{C^2_{ZM}} > 0.$$

$$MSE(\overline{Y}_{T_2(k)}) - MSE_{\min}(\overline{Y}_{\phi_2(k)}) = \frac{\overline{Y}^2\lambda_3}{(C^2_{XM}C^2_{ZM} - C^2_{XZM})C^2_{XM}C^2_{ZM}}\times \left[(C^2_{XM}C^2_{ZM} - C^2_{XZM})C_{XM} - (\rho_{XZ}\rho_{YX}C^2_X - \rho_{YZ}C^2_{XM})C_YC_Z\right]^2 > 0, \text{ always.}$$

$$MSE(\overline{Y}_{T_3(k)}) - MSE_{\min}(\overline{Y}_{\phi_{2(k)}}) = \overline{Y}^2\left[\lambda_3\frac{(C^2_{XM}-C_{YXM})^2}{C^2_{XM}} + \lambda_3\frac{(C_{XZM}C_{YXM}-C_{YZM}C^2_{XM})^2}{(C^2_{XM}C^2_{ZM}-C^2_{XZM})C^2_{XM}} - \frac{(\lambda_3 C_{XZM}-\lambda_1 C_{YZM})^2}{\lambda_1 C^2_{ZM}}\right] > 0,$$

if

$$\lambda_3\frac{(C^2_{XM}-C_{YXM})^2}{C^2_{XM}} + \lambda_3\frac{(C_{XZM}C_{YXM}-C_{YZM}C^2_{XM})^2}{(C^2_{XM}C^2_{ZM}-C^2_{XZM})C^2_{XM}} - \frac{(\lambda_3 C_{XZM}-\lambda_1 C_{YZM})^2}{\lambda_1 C^2_{ZM}} > 0.$$

5.3. For Proposed Estimator $\overline{Y}_{\phi_3(k)}$

We considered the differences of MSE of the existing and proposed estimator $\overline{Y}_{\phi_3(k)}$ for the purpose of comparison as given below:

$$Var(\overline{Y}_{MN}) - MSE_{\min}(\overline{Y}_{\phi_3(k)}) = \overline{Y}^2\left[\lambda_3\frac{C^2_{YXM}}{C^2_{XM}} + \frac{(\lambda_3 C_{YXM}C_{XZM}-\lambda_1 C_{YZM}C^2_{XM})^2}{(\lambda_1 C^2_{XM}C^2_{ZM}-\lambda_3 C^2_{XZM})C^2_{XM}}\right] > 0, \text{ always.}$$

$$MSE(\overline{Y}_R) - MSE_{\min}(\overline{Y}_{\phi_3(k)}) = \overline{Y}^2\left[\lambda_3\frac{(C^2_{XM}-C_{YXM})^2}{C^2_{XM}} + \frac{(\lambda_3 C_{YXM}C_{XZM}-\lambda_1 C_{YZM}C^2_{XM})^2}{(\lambda_1 C^2_{XM}C^2_{ZM}-\lambda_3 C^2_{XZM})C^2_{XM}}\right] > 0, \text{ always.}$$

$$MSE(\overline{Y}_{CI}) - MSE_{\min}(\overline{Y}_{\phi_3(k)}) = \overline{Y}^2\left[\frac{(\lambda_3 C_{YXM}C_{XZM}-\lambda_1 C_{YZM}C^2_{XM})^2}{(\lambda_1 C^2_{XM}C^2_{ZM}-\lambda_3 C^2_{XZM})C^2_{XM}}\right] > 0, \text{ always.}$$

$$MSE(\overline{Y}_{D_1(k)}) - MSE_{\min}(\overline{Y}_{\phi_3(k)}) = \overline{Y}^2\left[\left(\frac{1}{r}-\frac{1}{n}\right)\frac{C^2_{YXM}}{C^2_{XM}} + \frac{(\lambda_3 C_{YXM}C_{XZM}-\lambda_1 C_{YZM}C^2_{XM})^2}{(\lambda_1 C^2_{XM}C^2_{ZM}-\lambda_3 C^2_{XZM})C^2_{XM}}\right] > 0 \text{ always.}$$

$$MSE(\overline{Y}_{D_2(k)}) - MSE_{\min}(\overline{Y}_{\phi_3(k)}) = \overline{Y}^2\left[\frac{(\lambda_3 C_{YXM}C_{XZM}-\lambda_1 C_{YZM}C^2_{XM})^2}{(\lambda_1 C^2_{XM}C^2_{ZM}-\lambda_3 C^2_{XZM})C^2_{XM}}\right] > 0, \text{ always.}$$

$$MSE(\overline{Y}_{T_1(k)}) - MSE_{\min}(\overline{Y}_{\phi_3(k)}) = \overline{Y}^2\left[\lambda_3\frac{(C^2_{XM}-C_{YXM})^2}{C^2_{XM}} + \frac{(\lambda_3 C_{YXM}C_{XZM}-\lambda_1 C_{YZM}C^2_{XM})^2}{(\lambda_1 C^2_{XM}C^2_{ZM}-\lambda_3 C^2_{XZM})C^2_{XM}} - \lambda_2\frac{C^2_{YZM}}{C^2_{ZM}}\right] > 0,$$

if

$$\lambda_3 \frac{(C^2_{XM} - C_{YXM})^2}{C^2_{XM}} + \frac{(\lambda_3 C_{YXM} C_{XZM} - \lambda_1 C_{YZM} C^2_{XM})^2}{(\lambda_1 C^2_{XM} C^2_{ZM} - \lambda_3 C^2_{XZM}) C^2_{XM}} - \lambda_2 \frac{C^2_{YZM}}{C^2_{ZM}} > 0.$$

$$MSE(\overline{Y}_{T_2(k)}) - MSE_{\min}(\overline{Y}_{\phi_3(k)}) =$$
$$\overline{Y}^2 \left[\lambda_3 \frac{(C^2_{XM} - C_{YXM})^2}{C^2_{XM}} + \frac{(\lambda_3 C_{YXM} C_{XZM} - \lambda_1 C_{YZM} C^2_{XM})^2}{(\lambda_1 C^2_{XM} C^2_{ZM} - \lambda_3 C^2_{XZM}) C^2_{XM}} - \lambda_3 \frac{(C_{XZM} - C_{YZM})^2}{C^2_{ZM}}\right] > 0,$$

$$\text{if } \lambda_3 \frac{(C^2_{XM} - C_{YXM})^2}{C^2_{XM}} + \frac{(\lambda_3 C_{YXM} C_{XZM} - \lambda_1 C_{YZM} C^2_{XM})^2}{(\lambda_1 C^2_{XM} C^2_{ZM} - \lambda_3 C^2_{XZM}) C^2_{XM}} - \frac{(\lambda_3 C_{XZM} - \lambda_1 C_{YZM})^2}{\lambda_1 C^2_{ZM}} > 0.$$

$$MSE(\overline{Y}_{T_3(k)}) - MSE_{\min}(\overline{Y}_{\phi_3(k)}) = \frac{\overline{Y}^2 \lambda_3}{(\lambda_1 C^2_{XM} C^2_{ZM} - \lambda_3 C^2_{XZM}) C^2_{XM} C^2_{ZM} \lambda_1} \times$$
$$\left[(\lambda_1 C^2_{XM} C^2_{ZM} - \lambda_3 C^2_{XZM}) C_{XM} - (\rho_{XZ} \rho_{YX} C^2_X - \rho_{YZ} C^2_{XM}) C_Y C_Z\right]^2 > 0, \text{always.}$$

5.4. Comparison Among Proposed Estimators

In the previous subsections 5.1-5.3, we have compared the MSEs of existing estimators with the proposed estimator but in this section, we want to compare the three proposed estimators among themselves by taking the difference of their MSE as given below:

(i) The estimator $\overline{Y}_{\phi_1(k)}$ is better than $\overline{Y}_{\phi_2(k)}$ if:

$$MSE_{\min}(\overline{Y}_{\phi_2(k)}) - MSE_{\min}(\overline{Y}_{\phi_1(k)}) > 0$$

provided that $\lambda_2 \frac{C^2_{YZM}}{C^2_{ZM}} - \lambda_3 \frac{(C_{XZM} C_{YXM} - C_{YZM} C^2_{XM})^2}{(C^2_{XM} C^2_{ZM} - C^2_{XZM}) C^2_{XM}} > 0.$

(ii) The estimator $\overline{Y}_{\phi_1(k)}$ is better than $\overline{Y}_{\phi_3(k)}$ if:

$$MSE_{\min}(\overline{Y}_{\phi_3(k)}) - MSE_{\min}(\overline{Y}_{\phi_1(k)}) > 0$$

provided that $\lambda_2 \frac{C^2_{YZM}}{C^2_{ZM}} - \frac{(\lambda_3 C_{YXM} C_{XZM} - \lambda_1 C_{YZM} C^2_{XM})^2}{(\lambda_1 C^2_{XM} C^2_{ZM} - \lambda_3 C^2_{XZM}) C^2_{XM}} > 0.$

(iii) The estimator $\overline{Y}_{\phi_3(k)}$ is better than $\overline{Y}_{\phi_2(k)}$ if:

$$MSE_{\min}(\overline{Y}_{\phi_2(k)}) - MSE_{\min}(\overline{Y}_{\phi_3(k)}) > 0$$

provided that

$$\lambda_3 \frac{(C_{XZM} C_{YXM} - C_{YZM} C^2_{XM})^2}{(C^2_{XM} C^2_{ZM} - C^2_{XZM}) C^2_{XM}} - \frac{(\lambda_3 C_{YXM} C_{XZM} - \lambda_1 C_{YZM} C^2_{XM})^2}{(\lambda_1 C^2_{XM} C^2_{ZM} - \lambda_3 C^2_{XZM}) C^2_{XM}} > 0.$$

6. A Special Case: Results of Previous Sections for Uncorrelated Measurement Error

As already mentioned in the last of Section 2, that we have considered the situation of correlated ME for all the three variables (Y, X, Z), that means the variables (U, V, W) are correlated. Under this general situation, we have taken up each and every fact of the proposed estimators till Section 5. Now in this section, we will consider the special case when ME are not correlated i.e. the variables (U, V, W) are uncorrelated. Here, in this section we are going to reproduce all the results of previous section of our study as special cases. Keeping in view the fact of uncorrelated variables (U, V, W), we should have $\rho_{UV} = 0$, $\rho_{UW} = 0$, $\rho_{VW} = 0$. So, following are expectation values which shall be used for the derivation of proofs in case of uncorrelated ME. $E(e_1e_2) = \lambda_1\rho_{YX}C_YC_X$, $E(e_1e_3) = \lambda_2\rho_{YX}C_YC_X$, $E(e_1e_4) = \lambda_1\rho_{YZ}C_YC_Z$, $E(e_1e_5) = \lambda_2\rho_{YZ}C_YC_Z$, $E(e_2e_3) = \lambda_2C^2_{XM}$, $E(e_2e_4) = \lambda_1\rho_{XZ}C_XC_Z$, $E(e_2e_5) = \lambda_2\rho_{XZ}C_XC_Z$, $E(e_3e_4) = \lambda_2\rho_{XZ}C_XC_Z$, $E(e_3e_5) = \lambda_2\rho_{XZ}C_XC_Z$.

6.1. Biases and *MSE* of Existing Estimators

In this sub-section, the Biases and MSEs respectively of the existing estimators discussed in Section 3 have been summarized below, taking in account the uncorrelated ME case.

(a) Ratio method of Imputation

$$B(\overline{Y}_R) = \overline{Y}\lambda_3(C^2_{XM} - \rho_{YX}C_YC_X)$$

$$MSE(\overline{Y}_R) = \overline{Y}^2(\lambda_1C^2_{YM} + \lambda_3C^2_{XM} - 2\lambda_3\rho_{YX}C_YC_X)$$

(b) Compromised method of Imputation

$$B(\overline{Y}_{CI}) = (1-\alpha)\overline{Y}\lambda_3(C^2_{XM} - \rho_{YX}C_YC_X)$$

$$MSE(\overline{Y}_{CI}) = \overline{Y}^2\left(\lambda_1C^2_{YM} - \lambda_3\frac{\rho^2_{YX}C^2_YC^2_X}{C^2_{XM}}\right)$$

(c) Factor-type Imputation

$$B(\overline{Y}_{D_1(k)}) = -\overline{Y}P\left(\frac{1}{n} - \frac{1}{m}\right)(\psi_2C^2_{XM} - \rho_{YX}C_YC_X)$$

$$B(\overline{Y}_{D_2(k)}) = -\overline{Y}P\left(\frac{1}{r} - \frac{1}{m}\right)(\psi_2C^2_{XM} - \rho_{YX}C_YC_X)$$

$$MSE(\overline{Y}_{D_1(k)}) = \overline{Y}^2\left(\lambda_1C^2_{YM} - \left(\frac{1}{n} - \frac{1}{m}\right)\frac{\rho^2_{YX}C^2_YC^2_X}{C^2_{XM}}\right)$$

$$MSE(\overline{Y}_{D_2(k)}) = \overline{Y}^2\left(\lambda_1C^2_{YM} - \lambda_3\frac{\rho^2_{YX}C^2_YC^2_X}{C^2_{XM}}\right)$$

Here, $P = -\rho_{YX}\frac{C_Y C_X}{C^2_{XM}}$

(d) Factor-type chained Imputation

$$B(\overline{Y}_{T_1(k)}) = \overline{Y}\left[\lambda_3 C^2_{XM}\left(1-\frac{\rho_{YX}C_YC_X}{C^2_{XM}}\right) - \psi\lambda_2 C^2_{ZM}\left(\psi_2 - \frac{\rho_{YZ}C_YC_Z}{C^2_{ZM}}\right)\right]$$

$$B(\overline{Y}_{T_2(k)}) = \overline{Y}\lambda_3\left[C^2_{XM}\left(1-\frac{\rho_{YX}C_YC_X}{C^2_{XM}}\right) - \psi C^2_{ZM}\left(\psi_2 - \frac{\rho_{YZ}C_YC_Z}{C^2_{ZM}} + \frac{\rho_{XZ}C_XC_Z}{C^2_{ZM}}\right)\right]$$

$$B(\overline{Y}_{T_3(k)}) = \overline{Y}\left[\lambda_3 C^2_{XM}\left(1-\frac{\rho_{YX}C_YC_X}{C^2_{XM}}\right) + \psi C^2_{ZM}\left(-\lambda_1\psi_2 + \lambda_1\frac{\rho_{YZ}C_YC_Z}{C^2_{ZM}} - \lambda_3\frac{\rho_{XZ}C_XC_Z}{C^2_{ZM}}\right)\right]$$

$$MSE(\overline{Y}_{T_1(k)}) = \overline{Y}^2\left[\lambda_1 C^2_{YM} + \lambda_3 C^2_{XM} - 2\lambda_3\rho_{YX}C_YC_X - \lambda_2\frac{\rho^2_{YZ}C^2_YC^2_Z}{C^2_{ZM}}\right]$$

$$MSE(\overline{Y}_{T_2(k)}) = \overline{Y}^2\left[\lambda_1 C^2_{YM} + \lambda_3 C^2_{XM} - 2\lambda_3\rho_{YX}C_YC_X - \lambda_3\frac{(\rho_{XZ}C_XC_Z - \rho_{YZ}C_YC_Z)^2}{C^2_{ZM}}\right]$$

$$MSE(\overline{Y}_{T_3(k)}) = \overline{Y}^2\left[\lambda_1 C^2_{YM} + \lambda_3 C^2_{XM} - 2\lambda_3\rho_{YX}C_YC_X - \frac{(\lambda_3\rho_{XZ}C_XC_Z - \lambda_1\rho_{YZ}C_YC_Z)^2}{\lambda_1 C^2_{ZM}}\right]$$

Following are the theorems based on Biases and MSE of the proposed estimators $\overline{Y}_{\phi_i(k)}$ $(i = 1, 2, 3)$ using the above expectation values when ME are uncorrelated.

Theorem 4. *The Biases of proposed estimators* $\bar{y}_{\phi_i(k)}$ $(i = 1, 2, 3)$, *derived up to* $o(n^{-1})$ *are respectively:*

$$B(\overline{Y}_{\phi_1(k)}) = \overline{Y}[\lambda_3(C^2_{XM} - \rho_{YX}C_YC_X) - \lambda_2(\psi\psi_2C^2_{ZM} - \psi\rho_{YZ}C_YC_Z)] \tag{13}$$

$$B(\overline{Y}_{\phi_2(k)}) = \overline{Y}\lambda_3[C^2_{XM} - \rho_{YX}C_YC_X - \psi\psi_2C^2_{ZM} + \psi\rho_{YZ}C_YC_Z - \psi\rho_{XZ}C_XC_Z] \tag{14}$$

$$B(\overline{Y}_{\phi_3(k)}) = \overline{Y}[\lambda_3C^2_{XM} - \lambda_3\rho_{YX}C_YC_X - \lambda_1\psi\psi_2C^2_{ZM} + \lambda_1\psi\rho_{YZ}C_YC_Z - \lambda_3\psi\rho_{XZ}C_XC_Z] \tag{15}$$

***Proof*.** For the derivation of above results, refer to Appendix A.6. □

Theorem 5. *The MSE of proposed estimators* $\bar{y}_{\phi_i(k)}$ $(i = 1, 2, 3)$*, derived up to* $o(n^{-1})$ *are respectively:*

$$\begin{aligned} MSE(\overline{Y}_{\phi_1(k)}) = \overline{Y}^2(\lambda_1 C^2_{YM} + \lambda_2\psi^2 C^2_{ZM} + \lambda_3 C^2_{XM} + 2\psi\lambda_2\rho_{YZ}C_Y C_Z \\ - 2\lambda_3\rho_{YX}C_Y C_X) + \alpha_1^2\lambda_3 C^2_{XM} + 2\alpha_1\overline{Y}(C^2_{XM} - \rho_{YX}C_Y C_X) \end{aligned} \quad (16)$$

$$\begin{aligned} MSE(\overline{Y}_{\phi_2(k)}) = \overline{Y}^2(\lambda_1 C^2_{YM} + \lambda_3\psi^2 C^2_{ZM} + \lambda_3 C^2_{XM} + 2\psi\lambda_3\rho_{YZ}C_Y C_Z \\ - 2\lambda_3\rho_{YX}C_Y C_X - 2\psi\lambda_3\rho_{XZ}C_X C_Z) \\ + \alpha_2^2\lambda_3 C^2_{XM} + 2\alpha_2\overline{Y}\lambda_3(C^2_{XM} - \rho_{YX}C_Y C_X - \psi\rho_{XZ}C_X C_Z) \end{aligned} \quad (17)$$

$$\begin{aligned} MSE(\overline{Y}_{\phi_3(k)}) = \overline{Y}^2(\lambda_1 C^2_{YM} + \lambda_1\psi^2 C^2_{ZM} + \lambda_3 C^2_{XM} + 2\psi\lambda_1\rho_{YZ}C_Y C_Z \\ - 2\lambda_3\rho_{YX}C_Y C_X - 2\psi\lambda_3\rho_{XZ}C_X C_Z) \\ + \alpha_3^2\lambda_3 C^2_{XM} + 2\alpha_3\overline{Y}\lambda_3(C^2_{XM} - \rho_{YX}C_Y C_X - \psi\rho_{XZ}C_X C_Z) \end{aligned} \quad (18)$$

***Proof*.** For the derivation of above results, refer to Appendix A.6. □

Theorem 6. *The minimum MSE of proposed estimators* $\bar{y}_{\phi_i(k)}$ $(i = 1, 2, 3)$*, derived up to* $o(n^{-1})$ *are respectively:*

$$MSE_{\min}(\overline{Y}_{\phi_1(k)}) = \overline{Y}^2\left[\lambda_1 C^2_{YM} - \lambda_2\frac{\rho^2_{YZ}C^2_Y C^2_Z}{C^2_{ZM}} - \lambda_3\frac{\rho^2_{YX}C^2_Y C^2_X}{C^2_{XM}}\right] \quad (19)$$

$$\begin{aligned} MSE_{\min}(\overline{Y}_{\phi_2(k)}) = \\ \overline{Y}^2\left[\lambda_1 C^2_{YM} - \lambda_3\frac{\rho^2_{YX}C^2_Y C^2_X}{C^2_{XM}} - \lambda_3\frac{(\rho_{YX}\rho_{XZ}C^2_X - \rho_{YZ}C^2_{XM})^2 C^2_Y C^2_Z}{(C^2_{XM}C^2_{ZM} - \rho^2_{XZ}C^2_X C^2_Z)C^2_{XM}}\right] \end{aligned} \quad (20)$$

$$\begin{aligned} MSE_{\min}(\overline{Y}_{\phi_3(k)}) = \\ \overline{Y}^2\left[\lambda_1 C^2_{YM} - \lambda_3\frac{\rho^2_{YX}C^2_Y C^2_X}{C^2_{XM}} - \frac{(\lambda_3\rho_{YX}\rho_{XZ}C^2_X - \lambda_1\rho_{YZ}C^2_{XM})^2 C^2_Y C^2_Z}{(\lambda_1 C^2_{XM}C^2_{ZM} - \lambda_3\rho^2_{XZ}C^2_X C^2_Z)C^2_{XM}}\right] \end{aligned} \quad (21)$$

***Proof*.** For the derivation of above results, refer to Appendix A.6. □

Remark 4. In case of no measurement error, minimum MSE of the proposed estimators become:
$MSE_{\min}(\overline{Y}_{\phi_1(k)}) = \overline{Y}^2 C^2_Y[\lambda_1 - \lambda_3\rho^2_{YX} - \lambda_2\rho^2_{YZ}]$
$MSE_{\min}(\overline{Y}_{\phi_2(k)}) = \overline{Y}^2 C^2_Y[(\lambda_1 - \lambda_3\rho^2_{YX}) - \lambda_3\frac{(\rho_{YX}\rho_{XZ} - \rho_{YZ})^2}{1-\rho^2_{XZ}}]$
$MSE_{\min}(\overline{Y}_{\phi_3(k)}) = \overline{Y}^2 C^2_Y[(\lambda_1 - \lambda_3\rho^2_{YX}) - \lambda_3\frac{(\lambda_3\rho_{YX}\rho_{XZ} - \lambda_1\rho_{YZ})^2}{\lambda_1 - \lambda_3\rho^2_{XZ}}]$.

Remark 5. It should be noted that when the measurement errors are not correlated, i.e. $\rho_{UV} = \rho_{UW} = \rho_{VW} = 0$, the MSEs mentioned in Section 4 reduces to corresponding MSEs mentioned in Section 6.

6.2. Comparison of Proposed Estimators with Existing Estimators

In this sub-section, MSE of the proposed estimators have been compared analytically with all the existing estimators discussed above, considering uncorrelated ME.

6.2.1. For Proposed Estimator $\overline{Y}_{\phi_1(k)}$

We considered the differences of MSE of the existing and proposed estimator $\overline{Y}_{\phi_1(k)}$ for the purpose of comparison as given below:

$$Var(\overline{Y}_{MN}) - MSE_{\min}(\overline{Y}_{\phi_1(k)}) = \overline{Y}^2\left[\lambda_3\frac{\rho^2_{YX}C^2_YC^2_X}{C^2_{XM}} + \lambda_2\frac{\rho^2_{YZ}C^2_YC^2_Z}{C^2_{ZM}}\right] > 0, \text{ always.}$$

$MSE(\overline{Y}_R) - MSE_{\min}(\overline{Y}_{\phi_1(k)}) = \overline{Y}^2\left[\lambda_3\frac{(C^2_{XM}-\rho_{YX}C_YC_X)^2}{C^2_{XM}} + \lambda_2\frac{\rho^2_{YZ}C^2_YC^2_Z}{C^2_{ZM}}\right] > 0$, always.

$MSE(\overline{Y}_{CI}) - MSE_{\min}(\overline{Y}_{\phi_1(k)}) = \overline{Y}^2\left[\lambda_2\frac{\rho^2_{YZ}C^2_YC^2_Z}{C^2_{ZM}}\right] > 0$, always.

$$MSE(\overline{Y}_{D_1(k)}) - MSE_{\min}(\overline{Y}_{\phi_1(k)}) = \overline{Y}^2\left[\left(\frac{1}{r}-\frac{1}{n}\right)\frac{\rho^2_{YX}C^2_YC^2_X}{C^2_{XM}} + \lambda_2\frac{\rho^2_{YZ}C^2_YC^2_Z}{C^2_{ZM}}\right] > 0, \text{ always.}$$

$MSE(\overline{Y}_{D_2(k)}) - MSE_{\min}(\overline{Y}_{\phi_1(k)}) = \overline{Y}^2\left[\lambda_2\frac{\rho^2_{YZ}C^2_YC^2_Z}{C^2_{ZM}}\right] > 0$, always.

$MSE(\overline{Y}_{T_1(k)}) - MSE_{\min}(\overline{Y}_{\phi_1(k)}) = \overline{Y}^2\left[\lambda_3\frac{(C^2_{XM}-\rho_{YX}C_YC_X)^2}{C^2_{XM}}\right] > 0$, always.

$$MSE(\overline{Y}_{T_2(k)}) - MSE_{\min}(\overline{Y}_{\phi_1(k)}) = \overline{Y}^2\left[\lambda_3\frac{(C^2_{XM}-\rho_{YX}C_YC_X)^2}{C^2_{XM}} + \lambda_2\frac{\rho^2_{YZ}C^2_YC^2_Z}{C^2_{ZM}} - \lambda_3\frac{(\rho_{XZ}C_XC_Z-\rho_{YZ}C_YC_Z)^2}{C^2_{ZM}}\right] > 0,$$

if $\lambda_3\frac{(C^2_{XM}-\rho_{YX}C_YC_X)^2}{C^2_{XM}} + \lambda_2\frac{\rho^2_{YZ}C^2_YC^2_Z}{C^2_{ZM}} - \lambda_3\frac{(\rho_{XZ}C_XC_Z-\rho_{YZ}C_YC_Z)^2}{C^2_{ZM}} > 0$.

$$MSE(\overline{Y}_{T_3(k)}) - MSE_{\min}(\overline{Y}_{\phi_1(k)}) = \overline{Y}^2\left[\lambda_3\frac{(C^2_{XM}-\rho_{YX}C_YC_X)^2}{C^2_{XM}} + \lambda_2\frac{\rho^2_{YZ}C^2_YC^2_Z}{C^2_{ZM}} - \frac{(\lambda_3\rho_{XZ}C_XC_Z-\lambda_1\rho_{YZ}C_YC_Z)^2}{\lambda_1C^2_{ZM}}\right] > 0,$$

if

$$\lambda_3\frac{(C^2_{XM}-\rho_{YX}C_YC_X)^2}{C^2_{XM}} + \lambda_2\frac{\rho^2_{YZ}C^2_YC^2_Z}{C^2_{ZM}} - \frac{(\lambda_3\rho_{XZ}C_XC_Z-\lambda_1\rho_{YZ}C_YC_Z)^2}{\lambda_1C^2_{ZM}} > 0.$$

6.2.2. For Proposed Estimator $\overline{Y}_{\phi_2(k)}$

We considered the differences of MSE of the existing and proposed estimator $\overline{Y}_{\phi_2(k)}$ for the purpose of comparison as given below:

$$Var(\overline{Y}_{MN}) - MSE_{\min}(\overline{Y}_{\phi_2(k)}) = \overline{Y}^2\left[\lambda_3\frac{\rho_{YX}^2 C_Y^2 C_X^2}{C_{XM}^2} + \lambda_3\frac{(\rho_{YX}\rho_{XZ}C_X^2 - \rho_{YZ}C_{XM}^2)^2 C_Y^2 C_Z^2}{(C_{XM}^2 C_{ZM}^2 - \rho_{XZ}^2 C_X^2 C_Z^2)C_{XM}^2}\right] > 0, \text{ always.}$$

$$MSE(\overline{Y}_R) - MSE_{\min}(\overline{Y}_{\phi_2(k)}) = \overline{Y}^2\lambda_3\left[\frac{(C_{XM}^2 - \rho_{YX}C_Y C_X)^2}{C_{XM}^2} + \frac{(\rho_{YX}\rho_{XZ}C_X^2 - \rho_{YZ}C_{XM}^2)^2 C_Y^2 C_Z^2}{(C_{XM}^2 C_{ZM}^2 - \rho_{XZ}^2 C_X^2 C_Z^2)C_{XM}^2}\right] > 0, \text{ always.}$$

$$MSE(\overline{Y}_{CI}) - MSE_{\min}(\overline{Y}_{\phi_2(k)}) = \overline{Y}^2\left[\lambda_3\frac{(\rho_{YX}\rho_{XZ}C_X^2 - \rho_{YZ}C_{XM}^2)^2 C_Y^2 C_Z^2}{(C_{XM}^2 C_{ZM}^2 - \rho_{XZ}^2 C_X^2 C_Z^2)C_{XM}^2}\right] > 0, \text{ always.}$$

$$MSE(\overline{Y}_{D_1(k)}) - MSE_{\min}(\overline{Y}_{\phi_2(k)}) = \overline{Y}^2\left[\left(\frac{1}{r} - \frac{1}{n}\right)\frac{\rho_{YX}^2 C_Y^2 C_X^2}{C_{XM}^2} + \lambda_3\frac{(\rho_{YX}\rho_{XZ}C_X^2 - \rho_{YZ}C_{XM}^2)^2 C_Y^2 C_Z^2}{(C_{XM}^2 C_{ZM}^2 - \rho_{XZ}^2 C_X^2 C_Z^2)C_{XM}^2}\right] > 0, \text{ always.}$$

$$MSE(\overline{Y}_{D_2(k)}) - MSE_{\min}(\overline{Y}_{\phi_2(k)}) = \overline{Y}^2\left[\lambda_3\frac{(\rho_{YX}\rho_{XZ}C_X^2 - \rho_{YZ}C_{XM}^2)^2 C_Y^2 C_Z^2}{(C_{XM}^2 C_{ZM}^2 - \rho_{XZ}^2 C_X^2 C_Z^2)C_{XM}^2}\right] > 0, \text{ always.}$$

$$MSE(\overline{Y}_{T_1(k)}) - MSE_{\min}(\overline{Y}_{\phi_2(k)}) = \overline{Y}^2\left[\lambda_3\frac{(C_{XM}^2 - \rho_{YX}C_Y C_X)^2}{C_{XM}^2} + \lambda_3\frac{(\rho_{YX}\rho_{XZ}C_X^2 - \rho_{YZ}C_{XM}^2)^2 C_Y^2 C_Z^2}{(C_{XM}^2 C_{ZM}^2 - \rho_{XZ}^2 C_X^2 C_Z^2)C_{XM}^2} - \lambda_2\frac{\rho_{YZ}^2 C_Y^2 C_Z^2}{C_{ZM}^2}\right] > 0,$$

if

$$\lambda_3\frac{(C_{XM}^2 - \rho_{YX}C_Y C_X)^2}{C_{XM}^2} + \lambda_3\frac{(\rho_{YX}\rho_{XZ}C_X^2 - \rho_{YZ}C_{XM}^2)^2 C_Y^2 C_Z^2}{(C_{XM}^2 C_{ZM}^2 - \rho_{XZ}^2 C_X^2 C_Z^2)C_{XM}^2} - \lambda_2\frac{\rho_{YZ}^2 C_Y^2 C_Z^2}{C_{ZM}^2} > 0$$

$$MSE(\overline{Y}_{T_2(k)}) - MSE_{\min}(\overline{Y}_{\phi_2(k)}) = \frac{\overline{Y}^2\lambda_3}{(C_{XM}^2 C_{ZM}^2 - \rho_{XZ}^2 C_X^2 C_Z^2)C_{XM}^2 C_{ZM}^2} \times\left[(C_{XM}^2 C_{ZM}^2 - \rho_{XZ}^2 C_X^2 C_Z^2)C_{XM} - (\rho_{XZ}\rho_{YX}C_X^2 - \rho_{YZ}C_{XM}^2)C_Y C_Z\right]^2 > 0, \text{ always.}$$

$$MSE(\overline{Y}_{T_3(k)}) - MSE_{\min}(\overline{Y}_{\phi_2(k)}) = \overline{Y}^2\Bigg[\lambda_3\frac{(C_{XM}^2 - \rho_{YX}C_YC_X)^2}{C_{XM}^2} + \lambda_3\frac{(\rho_{YX}\rho_{XZ}C_X^2 - \rho_{YZ}C_{XM}^2)^2C_Y^2C_Z^2}{(C_{XM}^2C_{ZM}^2 - \rho_{XZ}^2C_X^2C_Z^2)C_{XM}^2} - \frac{(\lambda_3\rho_{XZ}C_XC_Z - \lambda_1\rho_{YZ}C_YC_Z)^2}{\lambda_1C_{ZM}^2}\Bigg] > 0,$$

if

$$\lambda_3\frac{(C_{XM}^2 - \rho_{YX}C_YC_X)^2}{C_{XM}^2} + \lambda_3\frac{(\rho_{YX}\rho_{XZ}C_X^2 - \rho_{YZ}C_{XM}^2)^2C_Y^2C_Z^2}{(C_{XM}^2C_{ZM}^2 - \rho_{XZ}^2C_X^2C_Z^2)C_{XM}^2} - \frac{(\lambda_3\rho_{XZ}C_XC_Z - \lambda_1\rho_{YZ}C_YC_Z)^2}{\lambda_1C_{ZM}^2} > 0$$

6.2.3. For Proposed Estimator $\overline{Y}_{\phi_3(k)}$

We considered the differences of MSE of the existing and proposed estimator $\overline{Y}_{\phi_3(k)}$ for the purpose of comparison as given below:

$$Var(\overline{Y}_{MN}) - MSE_{\min}(\overline{Y}_{\phi_3(k)}) = \overline{Y}^2\left[\lambda_3\frac{\rho_{YX}^2C_Y^2C_X^2}{C_{XM}^2} + \frac{(\lambda_3\rho_{YX}\rho_{XZ}C_X^2 - \lambda_1\rho_{YZ}C_{XM}^2)^2C_Y^2C_Z^2}{(\lambda_1C_{XM}^2C_{ZM}^2 - \lambda_3\rho_{XZ}^2C_X^2C_Z^2)C_{XM}^2}\right] > 0, \text{ always.}$$

$$MSE(\overline{Y}_R) - MSE_{\min}(\overline{Y}_{\phi_3(k)}) = \overline{Y}^2\left[\lambda_3\frac{(C_{XM}^2 - \rho_{YX}C_YC_X)^2}{C_{XM}^2} + \frac{(\lambda_3\rho_{YX}\rho_{XZ}C_X^2 - \lambda_1\rho_{YZ}C_{XM}^2)^2C_Y^2C_Z^2}{(\lambda_1C_{XM}^2C_{ZM}^2 - \lambda_3\rho_{XZ}^2C_X^2C_Z^2)C_{XM}^2}\right] > 0, \text{ always.}$$

$$MSE(\overline{Y}_{CI}) - MSE_{\min}(\overline{Y}_{\phi_3(k)}) = \overline{Y}^2\left[\frac{(\lambda_3\rho_{YX}\rho_{XZ}C_X^2 - \lambda_1\rho_{YZ}C_{XM}^2)^2C_Y^2C_Z^2}{(\lambda_1C_{XM}^2C_{ZM}^2 - \lambda_3\rho_{XZ}^2C_X^2C_Z^2)C_{XM}^2}\right] > 0, \text{ always.}$$

$$MSE(\overline{Y}_{D_1(k)}) - MSE_{\min}(\overline{Y}_{\phi_3(k)}) = \overline{Y}^2\left[\left(\frac{1}{r} - \frac{1}{n}\right)\frac{\rho_{YX}^2C_Y^2C_X^2}{C_{XM}^2} + \frac{(\lambda_3\rho_{YX}\rho_{XZ}C_X^2 - \lambda_1\rho_{YZ}C_{XM}^2)^2C_Y^2C_Z^2}{(\lambda_1C_{XM}^2C_{ZM}^2 - \lambda_3\rho_{XZ}^2C_X^2C_Z^2)C_{XM}^2}\right] > 0, \text{ always.}$$

$MSE(\overline{Y}_{D_2(k)}) - MSE_{\min}(\overline{Y}_{\phi_3(k)}) = \overline{Y}^2\left[\frac{(\lambda_3\rho_{YX}\rho_{XZ}C_X^2 - \lambda_1\rho_{YZ}C_{XM}^2)^2C_Y^2C_Z^2}{(\lambda_1C_{XM}^2C_{ZM}^2 - \lambda_3\rho_{XZ}^2C_X^2C_Z^2)C_{XM}^2}\right] > 0$, always.

$$MSE(\overline{Y}_{T_1(k)}) - MSE_{\min}(\overline{Y}_{\phi_3(k)}) = \overline{Y}^2\left[\lambda_3\frac{(C_{XM}^2 - \rho_{YX}C_YC_X)^2}{C_{XM}^2} + \frac{(\lambda_3\rho_{YX}\rho_{XZ}C_X^2 - \lambda_1\rho_{YZ}C_{XM}^2)^2C_Y^2C_Z^2}{(\lambda_1C_{XM}^2C_{ZM}^2 - \lambda_3\rho_{XZ}^2C_X^2C_Z^2)C_{XM}^2} - \lambda_2\frac{\rho_{YZ}^2C_Y^2C_Z^2}{C_{ZM}^2}\right] > 0,$$

if

$$\lambda_3\frac{(C_{XM}^2-\rho_{YX}C_YC_X)^2}{C_{XM}^2}+\frac{(\lambda_3\rho_{YX}\rho_{XZ}C_X^2-\lambda_1\rho_{YZ}C_{XM}^2)^2C_Y^2C_Z^2}{(\lambda_1C_{XM}^2C_{ZM}^2-\lambda_3\rho_{XZ}^2C_X^2C_Z^2)C_{XM}^2}-\lambda_2\frac{\rho_{YZ}^2C_Y^2C_Z^2}{C_{ZM}^2}>0$$

$$MSE(\overline{Y}_{T_2(k)})-MSE_{\min}(\overline{Y}_{\phi_3(k)})=\overline{Y}^2\Big[\lambda_3\frac{(C_{XM}^2-\rho_{YX}C_YC_X)^2}{C_{XM}^2}+\frac{(\lambda_3\rho_{YX}\rho_{XZ}C_X^2-\lambda_1\rho_{YZ}C_{XM}^2)^2C_Y^2C_Z^2}{(\lambda_1C_{XM}^2C_{ZM}^2-\lambda_3\rho_{XZ}^2C_X^2C_Z^2)C_{XM}^2}-\lambda_3\frac{(\rho_{XZ}C_XC_Z-\rho_{YZ}C_YC_Z)^2}{C_{ZM}^2}\Big]>0,$$

if

$$\lambda_3\frac{(C_{XM}^2-\rho_{YX}C_YC_X)^2}{C_{XM}^2}+\frac{(\lambda_3\rho_{YX}\rho_{XZ}C_X^2-\lambda_1\rho_{YZ}C_{XM}^2)^2C_Y^2C_Z^2}{(\lambda_1C_{XM}^2C_{ZM}^2-\lambda_3\rho_{XZ}^2C_X^2C_Z^2)C_{XM}^2}-\frac{(\lambda_3\rho_{XZ}C_XC_Z-\lambda_1\rho_{YZ}C_YC_Z)^2}{\lambda_1C_{ZM}^2}>0.$$

$$MSE(\overline{Y}_{T_3(k)})-MSE_{\min}(\phi_{3(k)})=\frac{\overline{Y}^2\lambda_3}{(\lambda_1C_{XM}^2C_{ZM}^2-\lambda_3\rho_{XZ}^2C_X^2C_Z^2)C_{XM}^2C_{ZM}^2\lambda_1}\times\left[(\lambda_1C_{XM}^2C_{ZM}^2-\lambda_3\rho_{XZ}^2C_X^2C_Z^2)C_{XM}-(\rho_{XZ}\rho_{YX}C_X^2-\rho_{YZ}C_{XM}^2)C_YC_Z\right]^2>0,$$

always.

6.2.4. Comparison Among Proposed Estimators

In this sub-section, we want to compare the three proposed estimators among themselves by taking the difference of their MSEs as given below:

(i) The estimator $\overline{Y}_{\phi_1(k)}$ is better than $\overline{Y}_{\phi_2(k)}$ if:

$$MSE_{\min}(\overline{Y}_{\phi_2(k)})-MSE_{\min}(\overline{Y}_{\phi_1(k)})>0$$

provided that

$$\lambda_2\frac{\rho_{YZ}^2C_Y^2C_Z^2}{C_{ZM}^2}-\lambda_3\frac{(\rho_{YX}\rho_{XZ}C_X^2-\rho_{YZ}C_{XM}^2)^2C_Y^2C_Z^2}{(C_{XM}^2C_{ZM}^2-\rho_{XZ}^2C_X^2C_Z^2)C_{XM}^2}>0$$

(ii) The estimator $\overline{Y}_{\phi_1(k)}$ is better than $\overline{Y}_{\phi_3(k)}$ if:

$$MSE_{\min}(\overline{Y}_{\phi_3(k)})-MSE_{\min}(\overline{Y}_{\phi_1(k)})>0$$

provided that $\lambda_2\frac{\rho_{YZ}^2C_Y^2C_Z^2}{C_{ZM}^2}-\frac{(\lambda_3\rho_{YX}\rho_{XZ}C_X^2-\lambda_1\rho_{YZ}C_{XM}^2)^2C_Y^2C_Z^2}{(\lambda_1C_{XM}^2C_{ZM}^2-\lambda_3\rho_{XZ}^2C_X^2C_Z^2)C_{XM}^2}>0$

(iii) The estimator $\overline{Y}_{\phi_3(k)}$ is better than $\overline{Y}_{\phi_2(k)}$ if:

$$MSE_{\min}(\overline{Y}_{\phi_2(k)})-MSE_{\min}(\overline{Y}_{\phi_3(k)})>0$$

provided that

$$\lambda_3\frac{(\rho_{YX}\rho_{XZ}C_X^2-\rho_{YZ}C_{XM}^2)^2C_Y^2C_Z^2}{(C_{XM}^2C_{ZM}^2-\rho_{XZ}^2C_X^2C_Z^2)C_{XM}^2}-\frac{(\lambda_3\rho_{YX}\rho_{XZ}C_X^2-\lambda_1\rho_{YZ}C_{XM}^2)^2C_Y^2C_Z^2}{(\lambda_1C_{XM}^2C_{ZM}^2-\lambda_3\rho_{XZ}^2C_X^2C_Z^2)C_{XM}^2}>0$$

7. Efficiency Comparison of Proposed Estimators

To have a precise idea about the performance of the various considered estimators, the conventional unbiased estimator $\overline{Y}_{MN}$ has been considered as base estimator for the comparison purposes. The Percent Relative Efficiency (PRE) of the estimators $\overline{Y}_l$ with respect to $\overline{Y}_{MN}$ is defined as:

$$PRE(\overline{Y}_l) = \frac{Var(\overline{Y}_{MN})}{MSE(\overline{Y}_l)} \times 100$$

where $\overline{Y}_l = \overline{Y}_{MN},\ \overline{Y}_R,\ \overline{Y}_{CI},\ \overline{Y}_{D_1(k)},\ \overline{Y}_{D_2(k)},\ \overline{Y}_{T_1(k)},\ \overline{Y}_{T_2(k)},\ \overline{Y}_{T_3(k)},\ \overline{Y}_{\phi_1(k)},$ $\overline{Y}_{\phi_2(k)},\ \overline{Y}_{\phi_3(k)}$.

7.1. Simulation Study

We have used `R` software for generating artificial values of a six-variate normal distributions using MVRNORM package i.e. (Y, X, Z, U, V, W) is assumed to follow a multivariate normal distribution with the population mean vector $[\overline{Y},\ \overline{X},\ \overline{Z},$ $0,\ 0,\ 0]$ and population covariance matrix:

$$\begin{pmatrix} S_Y^2 & \rho_{YX}S_YS_X & \rho_{YZ}S_YS_Z & 0 & 0 & 0 \\ \rho_{YX}S_YS_X & S_X^2 & \rho_{XZ}S_XS_Z & 0 & 0 & 0 \\ \rho_{YZ}S_YS_Z & \rho_{XZ}S_XS_Z & S_Z^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & S_U^2 & \rho_{UV}S_US_V & \rho_{UW}S_US_W \\ 0 & 0 & 0 & \rho_{UV}S_US_V & S_V^2 & \rho_{VW}S_VS_W \\ 0 & 0 & 0 & \rho_{UW}S_US_W & \rho_{VW}S_VS_W & S_W^2 \end{pmatrix}$$

To study the trend of PREs of various estimators w.r.t. variances (S_U^2, S_V^2, S_W^2) and correlation coefficients $(\rho_{UV}, \rho_{UW}, \rho_{VW})$ corresponding to variables of ME, we have taken different values of these variances and correlation coefficients. By doing this, we actually studied the impact of MEs on the PREs of different estimators. Keeping in view the above, we have taken two different cases in the following three artificial populations generated in `R`. Also, it is assumed that Y and X are highly correlated and Y and Z are moderately correlated. This has been done so that a better analysis of performance of estimators could be done.

The descriptions of three artificial populations are as follows:

Population 1: $N = 100$, $m = 80$, $n = 50$, $r = 30$, $\overline{Y} = 1.9819$, $\overline{X} = 7.7102$, $\overline{Z} = 10.0109$, $C_Y = 4.145$, $C_X = 2.223$, $C_Z = 1.338$, $\rho_{YX} = 0.991$, $\rho_{XZ} = 0.979$, $\rho_{YZ} = 0.944$.
Case(i). $(S_U^2, S_V^2, S_W^2) = (15, 20, 25)$ and $(\rho_{UV}, \rho_{VW}, \rho_{UW}) = (0.8, 0.7, 0.6)$.
Case(ii). $(S_U^2, S_V^2, S_W^2) = (20, 30, 35)$ and $(\rho_{UV}, \rho_{VW}, \rho_{UW}) = (0.96, 0.85, 0.78)$.

Population 2: $N = 200$, $m = 150$, $n = 120$, $r = 50$, $\overline{Y} = 4.7963$, $\overline{X} = 12.1738$, $\overline{Z} = 2.2473$, $C_Y = 1.3195$, $C_X = 1.3098$, $C_Z = 1.2471$, $\rho_{YX} = 0.9983$, $\rho_{XZ} = 0.7365$, $\rho_{YZ} = 0.6967$.

Case(i). $(S_U^2, S_V^2, S_W^2) = (22, 28, 34)$ and $(\rho_{UV}, \rho_{VW}, \rho_{UW}) = (0.91, 0.88, 0.62)$.
Case(ii). $(S_U^2, S_V^2, S_W^2) = (30, 36, 42)$ and $(\rho_{UV}, \rho_{VW}, \rho_{UW}) = (0.85, 0.78, 0.65)$.

Population 3: $N = 500$, $m = 450$, $n = 300$, $r = 250$, $\overline{Y} = 10.8614$, $\overline{X} = 16.9128$, $\overline{Z} = 38.6142$, $C_Y = 3.1562$, $C_X = 1.9824$, $C_Z = 2.6143$, $\rho_{YX} = 0.75$, $\rho_{XZ} = 0.68$, $\rho_{YZ} = 0.54$.
Case(i). $(S_U^2, S_V^2, S_W^2) = (12, 15, 18)$ and $(\rho_{UV}, \rho_{VW}, \rho_{UW}) = (0.80, 0.75, 0.60)$
Case(ii). $(S_U^2, S_V^2, S_W^2) = (17, 24, 32)$ and $(\rho_{UV}, \rho_{VW}, \rho_{UW}) = (0.71, 0.68, 0.54)$.

The following steps have been carried out in this simulation study:

(i) Using SRSWOR, a first phase random sample S' of size m is drawn from a population of size N.

(ii) Then, we draw a random sample S of size n from S' which is the second phase sample.

(iii) We have deleted $(n-r)$ sample units from sample S corresponding to variable Y.

(iv) Then, the deleted units are replaced with the corresponding imputed values as proposed in this study.

(v) We repeated the above steps 10,000 times so that we have got 10,000 values of $\widehat{\overline{Y}}$ i.e. $\widehat{\overline{Y}}$; $i = 1, 2, 3, \ldots, 10,000$.

(vi) The MSE, PRE and PCME (Percentage Contribution of the Measurement Error) are obtained by using the following formulae:

$$MSE(\widehat{\overline{Y}}_i) = \frac{1}{10000} \Sigma_{i=1}^{10000} (\widehat{\overline{Y}} - \overline{Y})^2$$

$$PRE\widehat{\overline{Y}} = \frac{MSE(\overline{Y}_{MN})}{MSE\widehat{\overline{Y}}} \times 100$$

$$PCME\widehat{\overline{Y}} = \left(\frac{MSE\widehat{\overline{Y}}_M - MSE\widehat{\overline{Y}}_0}{MSE\widehat{\overline{Y}}_M} \right) \times 100$$

where $MSE\widehat{\overline{Y}}_M$ is MSE of $\widehat{\overline{Y}}$ with the impact of measurement error and $MSE\widehat{\overline{Y}}_0$ is MSE of $\widehat{\overline{Y}}$ without the impact of measurement error.

The analysis of PREs of various considered estimators under different situations have been given in Tables 1-7.

Table 1: MSE, PRE and PCME of the existing and proposed estimators with respect to $\overline{Y}_{MN}$ when there are no ME.

Estimators	Population 1			Population 2			Population 3		
	MSE	PRE	PCME	MSE	PRE	PCME	MSE	PRE	PCME
$\overline{Y}_{MN}$	1.5746	100	0	0.6007	100	0	2.839	100	0
$\overline{Y}_{R}$	0.4845	324.9948	0	0.0685	876.9343	0	2.1169	134.1112	0
$\overline{Y}_{CI}$	0.1939	812.0681	0	0.0683	879.5022	0	0.986	287.9310	0
$\overline{Y}_{D_1(k)}$	0.7431	211.8961	0	0.4009	149.8379	0	1.106	256.6907	0
$\overline{Y}_{D_2(k)}$	0.1939	812.0681	0	0.0683	879.5022	0	0.986	287.9310	0
$\overline{Y}_{T_1(k)}$	0.6349	248.0076	0	0.1009	595.3419	0	0.854	332.4355	0
$\overline{Y}_{T_2(k)}$	0.2378	662.1531	0	0.0679	884.0324	0	0.654	434.0978	0
$\overline{Y}_{T_3(k)}$	0.1289	1221.5670	0	0.0672	893.1014	0	0.5180	548.0694	0
$\overline{Y}_{\phi_1(k)}$	**0.0435**	3619.7700	0	**0.0361**	1661.2280	0	**0.3286**	863.9683	0
$\overline{Y}_{\phi_2(k)}$	0.1707	922.437	0	0.0668	899.2515	0	0.4156	683.1087	0
$\overline{Y}_{\phi_3(k)}$	0.1279	1231.118	0	0.0664	904.6687	0	0.3854	736.6372	0

Table 2: MSE, PRE and PCME of the existing and proposed estimators with respect to $\overline{Y}_{MN}$ in case of ME for Population 1 (Case (i)).

	$S_U^2 = 15, S_V^2 = 20, S_W^2 = 25$					
Estimators	MSE	PRE	PCME	MSE	PRE	PCME
	$\rho_{UV} = 0, \rho_{VW} = 0, \rho_{UW} = 0$			$\rho_{UV} = 0.8, \rho_{VW} = 0.7, \rho_{UW} = 0.6$		
$\overline{Y}_{MN}$	1.8319	100	14.3773	2.0194	100	22.0263
$\overline{Y}_{R}$	0.7614	240.5962	36.3672	0.9128	221.2314	46.9215
$\overline{Y}_{CI}$	0.4384	417.8604	55.7709	0.7310	276.2517	73.4746
$\overline{Y}_{D_1(k)}$	0.9099	201.3298	18.3316	1.128	179.0248	34.1223
$\overline{Y}_{D_2(k)}$	0.4384	417.8604	55.7709	0.7310	276.2517	73.4746
$\overline{Y}_{T_1(k)}$	0.9146	200.2952	30.5816	1.038	194.5472	38.4343
$\overline{Y}_{T_2(k)}$	0.5187	353.1713	54.1546	0.7168	281.7243	66.8247
$\overline{Y}_{T_3(k)}$	0.4183	437.939	69.1848	0.6319	319.5759	79.6012
$\overline{Y}_{\phi_1(k)}$	**0.0763**	2400.917	42.9882	**0.0917**	2202.1816	52.5627
$\overline{Y}_{\phi_2(k)}$	0.3219	569.0897	46.9711	0.5912	342.2016	71.1265
$\overline{Y}_{\phi_3(k)}$	0.3187	574.8038	59.8682	0.5706	353.9082	77.5850

Table 3: MSE, PRE and PCME of the existing and proposed estimators with respect to $\overline{Y}_{MN}$ in case of ME for Population 1 (Case (ii)).

Estimators	$S_U^2 = 20, S_V^2 = 30, S_W^2 = 35$					
	MSE	PRE	PCME	MSE	PRE	PCME
	$\rho_{UV} = 0, \rho_{VW} = 0, \rho_{UW} = 0$			$\rho_{UV} = 0.96, \rho_{VW} = 0.85, \rho_{UW} = 0.78$		
$\overline{Y}_{MN}$	2.124	100	25.8662	2.5108	100	37.2869
$\overline{Y}_R$	0.9184	231.1271	47.2452	1.4031	178.9466	65.4693
$\overline{Y}_{CI}$	0.6312	336.5019	69.2807	1.0896	230.4331	82.20443
$\overline{Y}_{D_1(k)}$	1.2310	172.5426	39.6344	1.6190	155.0833	54.1013
$\overline{Y}_{D_2(k)}$	0.6312	336.5019	69.2807	1.0896	230.4331	82.2044
$\overline{Y}_{T_1(k)}$	1.128	188.2978	43.7145	1.5784	159.0724	59.7757
$\overline{Y}_{T_2(k)}$	0.7615	278.9231	68.7721	1.1064	226.9342	78.5068
$\overline{Y}_{T_3(k)}$	0.6137	346.0974	61.2514	1.0413	241.1216	77.1631
$\overline{Y}_{\phi_1(k)}$	**0.0912**	2328.9473	52.6659	**0.1616**	1553.7128	73.0816
$\overline{Y}_{\phi_2(k)}$	0.5018	423.2762	65.9824	0.8159	307.7337	79.0783
$\overline{Y}_{\phi_3(k)}$	0.4129	514.4102	69.0239	0.7462	336.4781	82.8598

Table 4: MSE, PRE and PCME of the existing and proposed estimators with respect to $\overline{Y}_{MN}$ in case of ME for Population 2 (Case (i)).

Estimators	$S_U^2 = 22, S_V^2 = 28, S_W^2 = 34$					
	MSE	PRE	PCME	MSE	PRE	PCME
	$\rho_{UV} = 0, \rho_{VW} = 0, \rho_{UW} = 0$			$\rho_{UV} = 0.91, \rho_{VW} = 0.88, \rho_{UW} = 0.62$		
$\overline{Y}_{MN}$	0.7867	100	23.6430	1.1067	100	45.7215
$\overline{Y}_R$	0.2316	338.2201	70.4231	0.6134	163.0257	88.8327
$\overline{Y}_{CI}$	0.2210	355.9728	69.0950	0.6012	184.0818	88.6393
$\overline{Y}_{D_1(k)}$	0.5109	153.9831	21.5306	0.9238	119.7986	56.6031
$\overline{Y}_{D_2(k)}$	0.2210	355.9728	69.0950	0.6012	184.0818	88.6393
$\overline{Y}_{T_1(k)}$	0.2813	279.6658	64.1308	0.6345	174.4208	84.0977
$\overline{Y}_{T_2(k)}$	0.1914	411.0240	63.3658	0.5176	213.8137	86.8817
$\overline{Y}_{T_3(k)}$	0.1615	487.1207	58.3900	0.4913	225.2595	86.3220
$\overline{Y}_{\phi_1(k)}$	**0.1037**	758.6306	65.1880	**0.3624**	305.3807	90.0386
$\overline{Y}_{\phi_2(k)}$	0.1546	508.8615	56.7917	0.4018	275.4355	83.3748
$\overline{Y}_{\phi_3(k)}$	0.1431	549.7554	53.5988	0.3921	282.2494	83.0655

TABLE 5: MSE, PRE and PCME of the existing and proposed estimators with respect to $\overline{Y}_{MN}$ in case of ME for Population 2 (Case (ii)).

	$S_U^2 = 30, S_V^2 = 36, S_W^2 = 42$					
Estimators	MSE	PRE	PCME	MSE	PRE	PCME
	$\rho_{UV} = 0, \rho_{VW} = 0, \rho_{UW} = 0$			$\rho_{UV} = 0.85, \rho_{VW} = 0.78, \rho_{UW} = 0.65$		
$\overline{Y}_{MN}$	0.8192	100	26.673	1.3168	100	54.3818
$\overline{Y}_{R}$	0.3149	260.1460	78.2470	0.8163	161.3132	91.6084
$\overline{Y}_{CI}$	0.3062	267.5375	77.6943	0.7994	91.4560	546.23
$\overline{Y}_{D_1(k)}$	0.6128	133.6814	34.5789	1.0213	128.9337	60.7461
$\overline{Y}_{D_2(k)}$	0.3062	267.5375	77.6943	0.7994	164.7235	91.4560
$\overline{Y}_{T_1(k)}$	0.3147	260.3114	67.9377	0.8004	164.5777	87.3938
$\overline{Y}_{T_2(k)}$	0.2893	283.1662	76.5295	0.6146	214.2531	88.9521
$\overline{Y}_{T_3(k)}$	0.2642	310.0681	74.5647	0.5981	220.1638	88.7644
$\overline{Y}_{\phi_1(k)}$	**0.1938**	422.7038	81.3725	**0.4264**	308.8180	91.5337
$\overline{Y}_{\phi_2(k)}$	0.2114	387.5118	68.4011	0.4996	263.5708	86.6293
$\overline{Y}_{\phi_3(k)}$	0.2006	408.3748	66.8993	0.4328	304.2513	84.6580

TABLE 6: MSE, PRE and PCME of the existing and proposed estimators with respect to $\overline{Y}_{MN}$ in case of ME for Population 3 (Case (i)).

	$S_U^2 = 12, S_V^2 = 15, S_W^2 = 18$					
Estimators	MSE	PRE	PCME	MSE	PRE	PCME
	$\rho_{UV} = 0, \rho_{VW} = 0, \rho_{UW} = 0$			$\rho_{UV} = 0.80, \rho_{VW} = 0.75, \rho_{UW} = 0.60$		
$\overline{Y}_{MN}$	3.5691	100	20.4561	3.7918	100	25.1279
$\overline{Y}_{R}$	3.1012	115.0877	31.7393	3.3219	114.1455	36.2744
$\overline{Y}_{CI}$	2.6345	135.4754	62.5735	2.8106	134.9106	64.9185
$\overline{Y}_{D_1(k)}$	2.8192	126.5997	60.7690	3.0013	126.3259	63.1493
$\overline{Y}_{D_2(k)}$	2.6345	135.4754	62.5735	2.8106	134.9106	64.9185
$\overline{Y}_{T_1(k)}$	2.0148	177.1441	57.6136	2.3478	161.5043	63.6255
$\overline{Y}_{T_2(k)}$	1.9237	185.5330	66.0030	2.1619	175.3920	69.7488
$\overline{Y}_{T_3(k)}$	1.7103	208.6826	69.7129	1.9314	196.3239	73.1800
$\overline{Y}_{\phi_1(k)}$	**0.7912**	451.0995	58.4681	**1.0048**	377.3686	67.2969
$\overline{Y}_{\phi_2(k)}$	1.1463	311.3582	63.7442	1.5781	240.2762	73.6645
$\overline{Y}_{\phi_3(k)}$	0.9218	387.1881	58.1905	1.3294	285.2264	71.0094

Table 7: MSE, PRE and PCME of the existing and proposed estimators with respect to $\overline{Y}_{MN}$ in case of ME for Population 3 (Case (ii)).

	$S_U^2 = 17, S_V^2 = 24, S_W^2 = 32$					
Estimators	MSE	PRE	PCME	MSE	PRE	PCME
	$\rho_{UV} = 0, \rho_{VW} = 0, \rho_{UW} = 0$			$\rho_{UV} = 0.71, \rho_{VW} = 0.68, \rho_{UW} = 0.54$		
$\overline{Y}_{MN}$	4.0236	100	29.4413	4.4197	100	35.7648
$\overline{Y}_R$	3.6198	111.1553	41.5188	4.0324	109.6047	47.5027
$\overline{Y}_{CI}$	3.1256	128.7304	68.4540	3.8236	115.5900	74.2127
$\overline{Y}_{D_1(k)}$	3.4618	116.2285	68.0513	3.9913	110.6862	72.2897
$\overline{Y}_{D_2(k)}$	3.1256	128.7304	68.4540	3.8236	115.5900	74.2127
$\overline{Y}_{T_1(k)}$	2.9147	138.0450	70.7002	3.6143	122.2837	76.3716
$\overline{Y}_{T_2(k)}$	2.6041	154.5101	74.8857	3.4916	126.5809	81.2693
$\overline{Y}_{T_3(k)}$	2.4713	162.8130	79.0393	3.1085	142.1811	83.3360
$\overline{Y}_{\phi_1(k)}$	**2.0615**	195.1782	84.0601	**2.5436**	173.7576	87.0813
$\overline{Y}_{\phi_2(k)}$	2.2983	175.0685	81.9170	2.8912	152.8673	87.8121
$\overline{Y}_{\phi_3(k)}$	2.1491	187.2225	82.0669	2.6317	167.9408	85.3554

7.2. Empirical Study Using Real Data Set

For proving the validity of our results using real data set, we have taken Population 4. which is a built-in data set in R called "mtcars"(Motor Trend Car Road Tests), which is retrieved from the 1974 Motor Trend US Magazine and comprises fuel consumption and 10 aspects of automobile design and performance for 32 automobiles (1973-74 models) (see Henderson & Velleman, 1981). We have selected the study variable Y as weight of automobile, and the auxiliary variables as Displacement (X) and Gross horsepower (Z). Suppose average of weights of automobiles which are used for analysis of fuel consumption is required. A normal variate $U \sim N(0,1)$ is generated in `R` and then added to Y to generate measurement error in Y. Then, we intentionally deleted some values of weights to make data comprehensible with our methodology. Since, no measurement errors are taken in variables X and Z, so this is case of uncorrelated Measurement Error. The description of all the necessary statistics is given below.
$N = 32$, $m = 17$, $n = 12$, $\overline{Y} = 3.217$, $\overline{X} = 230.7$, $\overline{Z} = 146.7$, $S_Y^2 = 0.9274$, $S_X^2 = 14880.77$, $S_Z^2 = 4553.965$, $S_U^2 = 0.7528$, $S_V^2 = 0$, $S_W^2 = 0$, $C_Y = 0.2993$, $C_X = 0.5287$, $C_Z = 0.4600$, $\rho_{YX} = 0.8762$, $\rho_{XZ} = 0.8433$, $\rho_{YZ} = 0.7038$, $\rho_{UV} = 0$, $\rho_{VW} = 0$, $\rho_{UW} = 0$.

The number of observed values (responding units) have been varied and comparison of proposed estimators with the existing ones have been shown in Table 8.

Table 8: MSE of the existing and proposed estimators in case of ME for Population 4.

Estimators	$r=10$		$r=9$		$r=8$		$r=7$	
	MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE
$\overline{Y}_{MN}$	0.1194	100	0.1381	100	0.1596	100	0.1899	100
$\overline{Y}_{R}$	0.1164	102.5773	0.1353	102.0692	0.1589	100.4405	0.1893	100.3169
$\overline{Y}_{CI}$	0.0861	138.5794	0.09692	142.5181	0.1103	144.6963	0.1276	148.8245
$\overline{Y}_{D_1(k)}$	0.0980	121.8367	0.1166	118.4391	0.1400	114.0000	0.1700	111.7059
$\overline{Y}_{D_2(k)}$	0.0861	138.5794	0.09692	142.5181	0.1103	144.6963	0.1276	148.8245
$\overline{Y}_{T_1(k)}$	0.1037	115.1398	0.1226	112.6427	0.1463	109.0909	0.1767	105.735
$\overline{Y}_{T_2(k)}$	0.0928	128.6638	0.1054	131.0247	0.1210	131.9008	0.1412	134.4901
$\overline{Y}_{T_3(k)}$	0.1141	104.6450	0.1298	106.3945	0.1484	107.5472	0.1713	110.8581
$\overline{Y}_{\phi_1(k)}$	**0.0345**	346.0870	**0.0347**	396.9531	**0.0350**	456.0000	**0.0354**	536.4407
$\overline{Y}_{\phi_2(k)}$	0.0861	138.6760	0.0968	142.6653	0.1102	144.8276	0.1275	148.9412
$\overline{Y}_{\phi_3(k)}$	0.0840	142.1429	0.0950	145.3684	0.1087	146.8261	0.1262	150.4754

8. Discussion

From Tables 1-8, we observed the following points:

(i) From Table 1, in case of no measurement error, PREs of all the proposed estimators are higher as compared to the existing estimators in all the Populations 1-3. PCME values are zero in each case.

(ii) From Tables 2-8, considering both the situation of measurement error (i.e. correlated ME and uncorrelated ME), all the proposed estimators perform efficiently as compared to the existing estimators in all the populations.

(iii) From Tables 2-7, for a given estimator, when one moves from uncorrelated ME to correlated ME, PCME values increases which shows when errors are correlated, there is significant contribution in MSE which cannot be ignored.

(iv) From Tables 2-7, for a specific estimator, when variance values are increased, PRE's value decreases while PCME values increases which shows effect of ME increases and efficiency of estimator decreases with increase in variance values.

(v) Among proposed estimators, $\overline{Y}_{\phi_1(k)}$ is superior to both $\overline{Y}_{\phi_2(k)}$ and $\overline{Y}_{\phi_3(k)}$ but $\overline{Y}_{\phi_3(k)}$ is more efficient than $\overline{Y}_{\phi_2(k)}$ due to the efficiency conditions discussed in previous Sections.

(vi) It should be noted that in all the populations, $\overline{Y}_{\phi_1(k)}$ is superior to all the estimators as its PRE is maximum in both situations i.e. in the absence and in the presence of ME.

(vii) From Table 8, on introducing measurement error in study variable, all the proposed estimators perform efficiently as indicated by decreased MSE values. Also, MSE of $\overline{Y}_{\phi_1(k)}$ is minimum of all which confirms it's superiority over all other considered estimators.

(viii) From Table 8, as the number of responding units decreases (i.e. number of missing observations increases), MSE values of various considered estimators increases which shows that efficiency of an estimator decreases as the number of missing units increases.

9. Conclusion

It is clear that all the proposed estimators perform efficiently as compared to the existing imputation strategies. The results reveal that $\overline{Y}_{\phi_i(k)}(i = 1, 2, 3)$ have high PRE in both the absence and presence of measurement error. In this paper, the proposed factor type imputation strategies have given certain equations, which on solving will provide multiple values of k that gives optimal MSE. Also, the value of k may be chosen, that yields minimum bias. So, all the F-T imputation strategies are bias-controllable at optimum level of MSE. Conclusively, on the basis of results obtained from analytical comparison of proposed chaining imputation techniques, we recommend these estimators for the practical applications in engineering and other relevant fields as these imputation techniques provide multiple choices of k which helps in curbing the negative effect of non-response in the data.

Acknowledgement

The Authors would like to thank the learned reviewers and the editor for their valuable suggestions that leads to improving the quality of the research paper in the present version.

[Received: February 2024 — Accepted: June 2024]

References

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

Al-Jararha, J. & Ahmed, M. (2002), 'The class of chain estimators for a finite population variance using double sampling', *International Journal of Information and Management Sciences* **13**(2), 13–18.

Audu, A. & Adewara, A. (2017), 'Modified factor-type estimators with two auxiliary variables under two-phase sampling', *Anale. Seria Informatica* **15**(1), 63–76.

Bhaskaran, K. & Smeeth, L. (2014), 'What is the difference between missing completely at random and missing at random?', *International Journal of Epidemiology* **43**(4), 1336–1339.

Bhushan, S., Kumar, A., Shukla, S., Bakr, M., Tashkandy, Y. A. & Hossain, M. M. (2023), 'New logarithmic type imputation techniques in presence of measurement errors', *Alexandria Engineering Journal* **71**, 707–730.

Chand, L. (1975), Some ratio-type estimators based on two or more auxiliary variables, Ph.D. Thesis, Iowa State University, Iowa.

Choudhury, S. & Singh, B. (2012), 'A class of chain ratio-cum-dual to ratio type estimator with two auxiliary characters under double sampling in sample surveys', *Statistics in Transition New Series* **13**(3), 519–536.

Diana, G. & Francesco Perri, P. (2010), 'Improved estimators of the population mean for missing data', *Communications in Statistics—Theory and Methods* **39**(18), 3245–3251.

Diana, G. & Giordan, M. (2012), 'Finite population variance estimation in presence of measurement errors', *Communications in Statistics-Theory and Methods* **41**(23), 4302–4314.

Doretti, M., Geneletti, S. & Stanghellini, E. (2018), 'Missing data: a unified taxonomy guided by conditional independence', *International Statistical Review* **86**(2), 189–204.

Grover, L. K. & Sharma, A. (2023), 'Predictive estimation of finite population mean in case of missing data under two-phase sampling', *Journal of Statistical Theory and Applications* **22**(4), 283–308.

Hansen, M. H. & Hurwitz, W. N. (1946), 'The problem of non-response in sample surveys', *Journal of the American Statistical Association* **41**(236), 517–529.

Henderson, H. V. & Velleman, P. F. (1981), 'Building multiple regression models interactively', *Biometrics* **37**(2), 391–411.

Kalton, G., Kasprzyk, D. & Santos, R. (1981), Issues of nonresponse and imputation in the survey of income and program participation, *in* 'Current topics in survey sampling', Academic Press, pp. 455–480. https://doi.org/10.1016/B978-0-12-426280-5.50032-0

Kumar, A., Bhushan, S., Shukla, S., Bakr, M., Alshangiti, A. M. & Balogun, O. S. (2024), 'Efficient imputation methods in case of measurement errors', *Heliyon* **10**(6), 1–15. https://doi.org/10.1016/j.heliyon.2024.e26864

Kumar, A., Bhushan, S., Shukla, S., Emam, W., Tashkandy, Y. & Gupta, R. (2023), 'Impact of correlated measurement errors on some efficient classes of estimators', *Journal of Mathematics* **2023**(1), 8140831.

Kumar, A., Singh, A. K., Singh, P. & Singh, V. (2017), 'A class of exponential chain type estimator for population mean with imputation of missing data under double sampling scheme', *Journal of Statistics Applications & Probability* **6**(3), 479–485.

Kumar, M. & Bahl, S. (2006), 'Class of dual to ratio estimators for double sampling', *Statistical Papers* **47**(2), 319.

Kumar, S. & Sharma, V. (2020), 'Improved chain ratio-product type estimators under double sampling scheme', *Journal of Statistics Applications and Probability Letters* **7**(2), 87–96.

Manisha & Singh (2001), 'An estimation of population mean in the presence of measurement errors', *Journal of the Indian Society of Agricultural Statistics* **54**(1), 13–18.

Mehta, P. & Tailor, R. (2020), 'Chain ratio type estimators using known parameters of auxiliary variates in double sampling', *Journal of Reliability and Statistical Studies* **13**(2), 243–252.

Pandey, R., Thakur, N. S. & Yadav, K. (2015), 'Estimation of population mean using exponential ratio type imputation method under survey non-response', *Journal of the Indian Statistical Association* **53**(1), 89–107.

Pandey, R., Yadav, K. & Thakur, N. (2016), 'Adapted factor-type imputation strategies', *Journal of Scientific Research* **8**(3), 321–339.

Reddy, V. (1978), 'A study on the use of prior knowledge on certain population parameters in estimation', *Sankhya C* **40**, 29–37.

Rehman, S. A., Shabbir, J. & Al-essa, L. A. (2024), 'On the development of survey methods for novel mean imputation and its application to abalone data', *Heliyon* **10**(11), 1–13. https://doi.org/10.1016/j.heliyon.2024.e31423

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

Sahoo, L., Sahoo, R. & Senapati, S. (2006), 'An empirical study on the accuracy of ratio and regression estimators in the presence of measurement errors', *Monte Carlo Methods & Applications* **12**, 495–501.

Sajjad, M. & Ismail, M. (2024), 'Efficient generalized estimators of population mean in the presence of non-response and measurement error', *Kuwait Journal of Science* **51**(3). https://doi.org/10.1016/j.kjs.2024.100224

Sande, I. (1979), 'A personal view of hot-deck imputation procedures', *Survey Methodology* **5**(2), 238–258.

Seaman, S., Galati, J., Jackson, D. & Carlin, J. (2013), 'What is meant by missing at random?', *Statistical Science* **28**(2), 257–268. https://doi.org/10.1214/13-STS415

Shalabh, A. & Tsai, J. R. (2017), 'Ratio and product methods of estimation of population mean in the presence of correlated measurement errors', *Communications in Statistics-Simulation and Computation* **46**(7), 5566–5593.

Shalabh, S. (1997), 'Ratio method of estimation in the presence of measurement errors', *Journal of the Indian Society of Agricultural Statistics* **52**, 150–155.

Shukla, D., Thakur, N. S., Pathak, S. & Rajput, D. S. (2009), 'Estimation of mean under imputation of missing data using factor-type estimator in two-phase sampling', *Statistics in Transition* **10**(3), 397–414.

Singh, G., Maurya, S., Khetan, M. & Kadilar, C. (2016), 'Some imputation methods for missing data in sample surveys', *Hacettepe Journal of Mathematics and Statistics* **45**(6), 1865–1880.

Singh, G., Priyanka, K., Kim, J.-M. & Singh, S. (2010), 'Estimation of population mean using imputation techniques in sample surveys', *Journal of the Korean Statistical Society* **39**(1), 67–74.

Singh, H. P., Gupta, A. & Tailor, R. (2021), 'Estimation of population mean using a difference-type exponential imputation method', *Journal of Statistical Theory and Practice* **15**, 1–43.

Singh, H. P. & Karpe, N. (2009), 'Ratio-product estimator for population mean in presence of measurement errors', *Journal of Applied Statistical Science* **16**(4), 437.

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

Singh, P. & Singh, R. (2017), 'Exponential ratio type estimator of population mean in presence of measurement error and non response', *International Journal of Statistics and Economics* **18**(3), 102–121.

Singh, P., Singh, R. & Bouza, C. N. (2018), 'Effect of measurement error and non-response on estimation of population mean', *Investigación Operacional* **39**(1), 108–120.

Singh, R. S. & Sharma, P. (2015), 'Method of estimation in the presence of non-response and measurement errors simultaneously', *Journal of Modern Applied Statistical Methods* **14**, 107–121.

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

Singh, V. K., Singh, R. & Smarandache, F. (2014), 'Difference-type estimators for estimation of mean in the presence of measurement error', *arXiv preprint arXiv:1410.0279* **1410**(0279). https://arxiv.org/abs/1410.0279

Singh, V., Singh, G. & Shukla, D. (1994), 'A class of chain ratio type estimators with two auxiliary variables under double sampling scheme', *Sankhyā: The Indian Journal of Statistics, Series B* **56**(2), 209–221.

Sohail, M. U., Shabbir, J. & Sohil, F. (2019), 'Imputation of missing values by using raw moments', *Statistics in Transition New Series* **20**(1), 21–40.

Thakur, N. S. & Shukla, D. (2022), 'Missing data estimation based on the chaining technique in survey sampling', *Statistics in Transition New Series* **23**(4), 91–111.

Tiwari, K. K., Bhougal, S. & Kumar, S. (2023), 'A general class of estimators in the presence of non-response and measurement error', *Journal of Statistics Application & Probability Letters* **10**(1), 13–33.

Vishwakarma, G. K., Singh, A. & Singh, N. (2020), 'Calibration under measurement errors', *Journal of King Saud University-Science* **32**(7), 2950–2961.

Yadav, S. K., Vishwakarma, G. K. & Sharma, D. K. (2024), 'A computational strategy for estimation of mean using optimal imputation in presence of missing observation', *Scientific Reports* **14**(1), 1–13. https://doi.org/10.1038/s41598-024-57264-y

Appendix A.

Appendix A.1.

Here, we are giving detailed derivation of the Bias and minimum MSE of existing estimator $\overline{Y}_{D_2(k)}$ as stated in subsection 3.4. The estimator $\overline{Y}_{D_2(k)}$ in terms of e_is is expressed as below:

$$\overline{Y}_{D_2(k)} = \overline{Y}(1+e_1)\left[\frac{(A+C)(1+e_3)+fB(1+e_2)}{(A+fB)(1+e_3)+C(1+e_2)}\right]$$

Using the following notations: $\psi_1 = \frac{fB}{A+fB+C}$, $\psi_2 = \frac{C}{A+fB+C}$, $\psi_3 = \frac{A+C}{A+fB+C}$, $\psi_4 = \frac{A+fB}{A+fB+C}$, $\psi_1+\psi_3=\psi_2+\psi_4=1$, $\overline{Y}_{D_2(k)}$ can be expressed as:

$$\overline{Y}_{D_2(k)} = \overline{Y}(1+e_1)(1+\psi_3 e_3+\psi_1 e_2)(1+\psi_4 e_3+\psi_2 e_2)^{-1}.$$

On further simplification $\overline{Y}_{D_2(k)}$ becomes:

$$\begin{aligned}\overline{Y}_{D_2(k)} = \overline{Y}(1 &- \psi_4 e_3 - \psi_2 e_2 + \psi_4^2 e_3^2 + \psi_2^2 e_2^2 + 2\psi_2\psi_4 e_2 e_3 + \psi_3 e_3 - \psi_3\psi_4 e_3^2 \\ &- \psi_2\psi_3 e_2 e_3 + \psi_1 e_2 - \psi_1\psi_4 e_2 e_3 - \psi_1\psi_2 e_2^2 + e_1 - \psi_4 e_1 e_3 - \psi_2 e_1 e_2) \qquad \text{(A1)}\end{aligned}$$

On subtracting $\overline{Y}$ from both sides of Equation (A1), we get:

$$\begin{aligned}\overline{Y}_{D_2(k)} - \overline{Y} = \overline{Y}(&-\psi_4 e_3 - \psi_2 e_2 + \psi_4^2 e_3^2 + \psi_2^2 e_2^2 + 2\psi_2\psi_4 e_2 e_3 + \psi_3 e_3 - \psi_3\psi_4 e_3^2 \\ &- \psi_2\psi_3 e_2 e_3 + \psi_1 e_2 - \psi_1\psi_4 e_2 e_3 - \psi_1\psi_2 e_2^2 + e_1 - \psi_4 e_1 e_3 - \psi_2 e_1 e_2) \qquad \text{(A2)}\end{aligned}$$

Using the expectation values from section 2, the Bias of $\overline{Y}_{D_2(k)}$ is given as: $B(\overline{Y}_{D_2(k)}) = (\psi + \psi_2\psi_4)\lambda_2 C^2_{XM} - \psi_2\psi\lambda_1 C^2_{XM} - (\psi_4\lambda_2 + \psi_2\lambda_1)C_{YXM}$.

For deriving MSE of $\overline{Y}_{D_2(k)}$, squaring both sides of Equation (A2), we get:

$$
\begin{aligned}
(\overline{Y}_{D_2(k)} - \overline{Y})^2 = \overline{Y}^2(\psi_4^2 e_3^2 + \psi_2^2 e_2^2 + \psi_3^2 e_3^2 + \psi_1^2 e_2^2 + e_1^2 + 2\psi_2\psi_4 e_2 e_3 - 2\psi_3\psi_4 e_3^2 \\
-2\psi_1\psi_4 e_2 e_3 - 2\psi_4 e_1 e_3 - 2\psi_2\psi_3 e_2 e_3 - 2\psi_1\psi_2 e_2^2 - 2\psi_2 e_1 e_2 \\
+2\psi_1\psi_3 e_2 e_3 + 2\psi_3 e_1 e_3 + 2\psi_1 e_1 e_2)
\end{aligned}
$$

Using $\psi_1 + \psi_3 = \psi_2 + \psi_4 = 1$, $\psi = \psi_1 - \psi_2 = -(\psi_3 - \psi_4)$, the above result becomes:

$$
(\overline{Y}_{D_{2(k)}} - \overline{Y})^2 = \overline{Y}^2(e_1^2 + \psi^2 e_2^2 + \psi^2 e_3^2 - 2\psi^2 e_2 e_3 + 2\psi e_1 e_2 - 2\psi e_1 e_3)
$$

Using expectations values given in Section 2, the MSE of $\overline{Y}_{D_2(k)}$ becomes:

$$
\begin{aligned}
MSE(\overline{Y}_{D_2(k)}) = \overline{Y}^2(\lambda_1 C^2_{YM} + \psi^2\lambda_1 C^2_{XM} + \lambda_2 C^2_{XM} \\
- 2\lambda_2 C^2_{XM} + 2\psi\lambda_1 C_{YXM} - 2\psi\lambda_2 C_{YXM})
\end{aligned}
$$

which can be rewritten as:

$$
MSE(\overline{Y}_{D_2(k)}) = \overline{Y}^2(\lambda_1 C^2_{YM} + \psi^2\lambda_3 C^2_{XM} + 2\psi\lambda_3 C_{YXM}) \tag{A3}
$$

For optimal value of ψ, we minimize $MSE(\overline{Y}_{D_2(k)})$ by partially differentiating w.r.t. ψ on both sides of Equation (A3), we get the optimal values of ψ as: $\psi_{opt} = -\frac{C_{YXM}}{C^2_{XM}}$. Now, replacing the optimum values of ψ in Equation (A3), the minimum MSE of $\overline{Y}_{D_2(k)}$ is given as:

$$
MSE(\overline{Y}_{D_2(k)}) = \overline{Y}^2 \left(\lambda_1 C^2_{YM} - \lambda_3 \frac{C^2_{YXM}}{C^2_{XM}}\right)
$$

Similarly, we can derive the expressions for bias and minimum MSE of the estimator $\overline{Y}_{D_1(k)}$.

Appendix A.2.

Here, we are giving detailed derivation of the results of existing estimator $\overline{Y}_{T_1(k)}$ as stated in Subsection 3.5. The estimator $\overline{Y}_{T_1(k)}$ in terms of e_is is expressed as below:

$$
\begin{aligned}
\overline{Y}_{T_1(k)} &= \overline{Y}(1+e_1)(1+e_3)(1+e_2)^{-1}(1+\psi_1 e_5)(1+\psi_2 e_5)^{-1} \\
&= \overline{Y}(1+e_1)(1+e_3)(1+e_2)^{-1}(1+\psi e_5 + a e_5^2),
\end{aligned}
$$

where $a = \psi_2^2 - \psi_1\psi_2$.

On further simplification $\overline{Y}_{T_1(k)}$ becomes:

$$
\begin{aligned}
\overline{Y}_{T_1(k)} = \overline{Y}(1 + \psi e_5 + a e_5^2 - e_2 - \psi e_2 e_5 + e_2^2 + e_3 + \psi e_3 e_5 - e_2 e_3 \\
+ e_1 + \psi e_1 e_5 - e_1 e_2 + e_1 e_3)
\end{aligned} \tag{A4}
$$

On subtracting $\overline{Y}$ from both sides of Equation (A4), we get:

$$\overline{Y}_{T_1(k)} - \overline{Y} = (\psi e_5 + a e_5^2 - e_2 - \psi e_2 e_5 + e_2^2 + e_3 + \psi e_3 e_5 - e_2 e_3 + e_1 + \psi e_1 e_5 - e_1 e_2 + e_1 e_3) \quad \text{(A5)}$$

Using the expectation values from Section 2, the Bias of $\overline{Y}_{T_1(k)}$ is given as: $\overline{Y}(a\lambda_2 C_{ZM}^2 + (\lambda_2 - \lambda_1) C_{YXM} + (\lambda_1 - \lambda_2) C_{XM}^2 + \psi\lambda_2 C_{YZM})$. After substituting the value of a, we get the required expression of Bias. For deriving MSE of $\overline{Y}_{T_1(k)}$, squaring both sides of Equation (A5), we get:

$$(\overline{Y}_{T_1(k)} - \overline{Y})^2 = \overline{Y}^2(\psi^2 e_5^2 + e_2^2 + e_3^2 + e_1^2 - 2\psi e_2 e_5 + 2\psi e_3 e_5 + 2\psi e_1 e_5 - 2e_2 e_3 - 2e_1 e_2 + 2e_1 e_3)$$

The MSE of $\bar{y}_{T_1(k)}$ becomes:

$$MSE(\overline{Y}_{T_1(k)}) = \overline{Y}^2(\psi^2\lambda_2 C_{ZM}^2 + \lambda_1 C_{XM}^2 + \lambda_2 C_{XM}^2 + \lambda_1 C_{YM}^2 - 2\psi\lambda_2 C_{XZM} + 2\psi\lambda_2 C_{XZM} + 2\psi\lambda_2 C_{YZM} - 2\lambda_2 C_{XM}^2 - 2\lambda_1 C_{YXM} + 2\lambda_2 C_{YXM})$$

The expression of MSE of $\overline{Y}_{T_1(k)}$ can further be written as:

$$MSE(\overline{Y}_{T_1(k)}) = \overline{Y}^2(\psi^2\lambda_2 C_{ZM}^2 - \lambda_3 C_{XM}^2 + \lambda_1 C_{YM}^2 + 2\psi\lambda_2 C_{YZM} - 2\lambda_3 C_{YXM}) \quad \text{(A6)}$$

For optimal values of ψ, we minimize $MSE(\overline{Y}_{T_1(k)})$ by partially differentiating w.r.t. ψ on both sides of Equation (A6). The optimal value of ψ is obtained as: $\psi_{opt} = -\frac{C_{YZM}}{C_{ZM}^2}$. Now, replacing the optimum value of ψ in Equation (A6), the minimum MSE of $\overline{Y}_{T_1(k)}$ is given as:

$$MSE(\overline{Y}_{T_1(k)}) = \overline{Y}^2 \left[\lambda_1 C_{YM}^2 + \lambda_3 C_{XM}^2 - 2\lambda_3 C_{YXM} - \lambda_2 \frac{C_{YZM}^2}{C_{ZM}^2}\right] \quad \text{(A7)}$$

Similarly, using the above steps and expectation values for correlated ME for the estimators $\overline{Y}_{T_2(k)}$ and $\overline{Y}_{T_3(k)}$, we have obtained the expressions for bias and minimum MSE as provided in Subsection 3.5.

Appendix A.3.

Here, we are giving detailed proof of Theorem 1. The proposed estimator $\overline{Y}_{\phi_1(k)}$ in terms of e_is is expressed as below:

$$\begin{aligned}\overline{Y}_{\phi_1(k)} &= \overline{Y}(1+e_1)(1+e_3)(1+e_2)^{-1}(1+\psi_1 e_5)(1+\psi_2 e_5)^{-1} + \alpha_1(e_3 - e_3^2 - e_2 + e_2 e_3)\\ &= \overline{Y}(1+e_1)(1+e_3)(1+e_2)^{-1}(1+\psi e_5 + a e_5^2) + \alpha_1(e_3 - e_3^2 - e_2 + e_2 e_3)\end{aligned}$$

where $\psi = \psi_1 - \psi_2$, $a = \psi_2^2 - \psi_1\psi_2$.

On further simplification, $\overline{Y}_{\phi_1(k)}$ becomes:

$$\begin{aligned}\overline{Y}_{\phi_1(k)} &= \overline{Y}(1+e_1+e_3+e_1e_3)(1+\psi e_5 \\ &\quad + ae_5^2 - e_2 - \psi e_2e_5 + e_2^2) + \alpha_1(e_3 - e_3^2 - e_2 + e_2e_3) \\ &= \overline{Y}(1+\psi e_5 + ae_5^2 - e_2 - \psi e_2e_5 + e_2^2 + e_3 + \psi e_3e_5 - e_2e_3 + e_1 \\ &\quad + \psi e_1e_5 - e_1e_2 + e_1e_3) + \alpha_1(e_3 - e_3^2 - e_2 + e_2e_3)\end{aligned} \tag{A8}$$

On subtracting $\overline{Y}$ from both sides of Equation (A8), we get:

$$\begin{aligned}\overline{Y}_{\phi_1(k)} - \overline{Y} &= (\psi e_5 + ae_5^2 - e_2 - \psi e_2e_5 + e_2^2 + e_3 + \psi e_3e_5 - e_2e_3 \\ &\quad + e_1 + \psi e_1e_5 - e_1e_2 + e_1e_3) + \alpha_1(e_3 - e_3^2 - e_2 + e_2e_3)\end{aligned} \tag{A9}$$

Using the expectation values from Section 2, the Bias of $\overline{Y}_{\phi_1(k)}$ is given as: $\overline{Y}(a\lambda_2C_{ZM}^2 + (\lambda_2-\lambda_1)\,C_{YXM} + (\lambda_1-\lambda_2)\,C_{XM}^2 + \psi\lambda_2C_{YZM})$. After substituting the value of a, we get the required expression of Bias.

Similarly, biases of other two proposed estimators $\overline{Y}_{\phi_2(k)}$ and $\overline{Y}_{\phi_3(k)}$ can be obtained.

Appendix A.4.

Here, we are giving detailed proof of Theorem 2. For deriving MSE of $\overline{Y}_{\phi_1(k)}$, squaring both sides of Equation (A9), we get:

$$\begin{aligned}(\overline{Y}_{\phi_1(k)} - \overline{Y})^2 &= \overline{Y}^2(\psi^2e_5^2 + e_2^2 + e_3^2 + e_1^2 - 2\psi e_2e_5 + 2\psi e_3e_5 + 2\psi e_1e_5 \\ &\quad -2e_2e_3 - 2e_1e_2 + 2e_1e_3) + \alpha_1^2(e_2^2 + e_3^2 - 2e_2e_3) \\ &\quad +2\alpha_1\overline{Y}(\psi e_3e_5 - e_2e_3 + e_3^2 + e_1e_3 - \psi e_2e_5 + e_2^2 - e_2e_3 - e_1e_2)\end{aligned}$$

Using the expectation values as given in Section 2, the MSE of $\overline{Y}_{\phi_1(k)}$ becomes:

$$\begin{aligned}MSE(\overline{Y}_{\phi_1(k)}) &= \overline{Y}^2(\psi^2\lambda_2C_{ZM}^2 + \lambda_1C_{XM}^2 + \lambda_2C_{XM}^2 + \lambda_1C_{YM}^2 - 2\psi\lambda_2C_{XZM} \\ &\quad +2\psi\lambda_2C_{XZM} + 2\psi\lambda_2C_{YZM} - 2\lambda_2C_{XM}^2 - 2\lambda_1C_{YXM} + 2\lambda_2C_{YXM}) \\ &\quad +\alpha_1^2(\lambda_1C_{XM}^2 + \lambda_2C_{XM}^2 - 2\lambda_2C_{XM}^2) + 2\alpha_1\overline{Y}(\psi\lambda_2\lambda_1C_{XZM} - \lambda_2C_{XM}^2 \\ &\quad +\lambda_2C_{XM}^2 + \lambda_2C_{YXM} - \psi\lambda_2\lambda_2C_{XZM} + \lambda_1\lambda_2C_{XM}^2 - \lambda_2C_{XM}^2 - \lambda_1C_{YXM})\end{aligned}$$

The expression of MSE of $\overline{Y}_{\phi_1(k)}$ can further be written as:

$$\begin{aligned}MSE(\overline{Y}_{\phi_1(k)}) &= \overline{Y}^2(\psi^2\lambda_2C_{ZM}^2 + (\lambda_2-\lambda_1)C_{XM}^2 + \lambda_1C_{YM}^2 + 2\psi\lambda_2C_{YZM} \\ &\quad + 2(\lambda_2-\lambda_1)C_{YXM}) + \alpha_1^2(\lambda_1-\lambda_2)C_{XM}^2 + 2\alpha_1\overline{Y}((\lambda_1-\lambda_2)C_{XM}^2 \\ &\quad - (\lambda_1-\lambda_2)C_{YXM})\end{aligned} \tag{A10}$$

or

$$\begin{aligned}MSE(\overline{Y}_{\phi_1(k)}) &= \overline{Y}^2(\lambda_1C_{YM}^2 + \lambda_2\psi^2C_{ZM}^2 + \lambda_3C_{XM}^2 + 2\psi\lambda_2C_{YZM} - 2\lambda_3C_{YXM}) \\ &\quad + \alpha_1^2\lambda_3C_{XM}^2 + 2\alpha_1\overline{Y}(C_{XM}^2 - C_{YXM})\end{aligned} \tag{A11}$$

Similarly, MSEs of proposed estimators $\overline{Y}_{\phi_2(k)}$ and $\overline{Y}_{\phi_3(k)}$ can be obtained.

Appendix A.5.

Here, we are giving detailed proof of Theorem 3. We minimize MSE of $\overline{Y}_{\phi_1(k)}$ w.r.t ψ and α_1 in Equation (A11), using the principle of minimization. So, we get the optimal values of ψ and α_1 as: $\psi_{opt} = -\frac{C_{YZM}}{C^2_{ZM}} = C_1$ (say), $\alpha_{1(opt)} = -\frac{C^2_{XM} - C_{YXM}}{C^2_{XM}}$. Now, replacing the optimum values of ψ and α_1 in Equation (A11), the minimum MSE of $\overline{Y}_{\phi_1(k)}$ is given as:

$$MSE_{\min}(\overline{Y}_{\phi_1(k)}) = \overline{Y}^2 \left[\lambda_1 C^2_{YM} - \lambda_2 \frac{C^2_{YZM}}{C^2_{ZM}} - \lambda_3 \frac{C^2_{YXM}}{C^2_{XM}}\right] \tag{A12}$$

Similarly, using the above steps and expectation values for correlated ME for the estimators $\overline{Y}_{\phi_2(k)}$ and $\overline{Y}_{\phi_3(k)}$, we have obtained the optimum values of constants and expressions for minimum MSE, which are given as follows:

(i) **For proposed estimator $\overline{Y}_{\phi_2(k)}$.** The optimum values of ψ and α_2 are obtained (by solving respective normal equations) as:

$$\psi_{opt} = \frac{C_{XZM}C_{YXM} - C_{YZM}C^2_{XM}}{C^2_{XM}C^2_{ZM} - C^2_{XZM}} = C_2 \text{ (say)}$$

$$\alpha_{2(opt)} = \overline{Y}\frac{C_{YXM} + \psi C_{XZM} - C^2_{XM}}{C^2_{XM}}$$

The optimum MSE of $\overline{Y}_{\phi_2(k)}$ is:

$$MSE_{\min}(\overline{Y}_{\phi_2(k)}) = \overline{Y}^2 \left[\lambda_1 C^2_{YM} - \lambda_3 \frac{C^2_{YXM}}{C^2_{XM}} - \lambda_3 \frac{(C_{XZM}C_{YXM} - C_{YZM}C^2_{XM})^2}{(C^2_{XM}C^2_{ZM} - C^2_{XZM})C^2_{XM}}\right]$$

(ii) **For proposed estimator $\overline{Y}_{\phi_3(k)}$.** The optimum values of ψ and α_3 are as follows:

$$\psi_{opt} = \frac{\lambda_3 C_{XZM}C_{YXM} - \lambda_1 C_{YZM}C^2_{XM}}{\lambda_1 C^2_{XM}C^2_{ZM} - \lambda_3 C^2_{XZM}} = C_3 \text{ (say)}$$

$$\alpha_{3(opt)} = \overline{Y}\left(\frac{\rho_{YX}C_Y C_X + \psi \rho_{XZ} C_Z C_X - C^2_{XM}}{C^2_{XM}}\right)$$

The optimum MSE of $\overline{Y}_{\phi_3(k)}$ is:

$$MSE_{\min}(\overline{Y}_{\phi_3(k)}) = \overline{Y}^2 \left[\lambda_1 C^2_{YM} - \lambda_3 \frac{C^2_{YXM}}{C^2_{XM}} - \frac{(\lambda_3 C_{YXM}C_{XZM} - \lambda_1 C_{YZM}C^2_{XM})^2}{(\lambda_1 C^2_{XM}C^2_{ZM} - \lambda_3 C^2_{XZM})C^2_{XM}}\right]$$

Appendix A.6.

We have provided a brief proof of Theorems 4-6 by using the expectation values given in Section 6, in case of uncorrelated ME. For this purpose, we have taken the proposed estimator $\overline{Y}_{\phi_1(k)}$ only. For this proposed estimator, we are giving the proofs of the expressions for the bias, MSE and minimum MSE. These results are a part of the Theorems 4-6. For the other two proposed estimators i.e. $\overline{Y}_{\phi_2(k)}$ and $\overline{Y}_{\phi_3(k)}$, we can derive the expressions of their biases, MSE and minimum MSE on the same lines.

The bias and MSE of $\overline{Y}_{\phi_1(k)}$ in terms of ψ and α_1, up to first order of approximation, after using the values of expectations (neglecting powers of e_i's which are greater than two) are:

$$B(\overline{Y}_{\phi_1(k)}) = \overline{Y}[\lambda_3(C^2_{XM} - \rho_{YX}C_YC_X) - \lambda_2(\psi\psi_2C^2_{ZM} - \psi\rho_{YZ}C_YC_Z)]$$

$$\begin{aligned} MSE(\overline{Y}_{\phi_1(k)}) = \overline{Y}^2(\lambda_1C^2_{YM} + \lambda_2\psi^2C^2_{ZM} + \lambda_3C^2_{XM} + 2\psi\lambda_2\rho_{YZ}C_YC_Z - 2\lambda_3\rho_{YX}C_YC_X) \\ + \alpha_1^2\lambda_3C^2_{XM} + 2\alpha_1\overline{Y}(C^2_{XM} - \rho_{YX}C_YC_X) \end{aligned}$$

For obtaining optimum MSE of $\overline{Y}_{\phi_1(k)}$, we have obtained optimum value of ψ and α_1 by differentiating $MSE(\overline{Y}_{\phi_1(k)})$ w.r.t ψ and α_1 and then putting these equal to zero, Then after solving the normal equations, we get

$$\psi_{opt} = -\frac{\rho_{YZ}C_YC_Z}{C^2_{ZM}}$$

$$\alpha_{1(opt)} = -\overline{Y}\frac{C^2_{XM} - \rho_{YX}C_YC_X}{C^2_{XM}}$$

The minimum MSE of $\overline{Y}_{\phi_1(k)}$, after putting optimum values of ψ and $\alpha_1 = \alpha_{1(opt)}$ in $MSE(\overline{Y}_{\phi_1(k)})$ is:

$$MSE_{\min}(\overline{Y}_{\phi_1(k)}) = \overline{Y}^2\left[\lambda_1C^2_{YM} - \lambda_2\frac{\rho^2_{YZ}C^2_YC^2_Z}{C^2_{ZM}} - \lambda_3\frac{\rho^2_{YX}C^2_YC^2_X}{C^2_{XM}}\right]$$