Improved Estimation Procedure for Population Coefficient of Variation Using Calibrated Weights under Stratified Successive Sampling in the Presence of Non-Response and Measurement Errors

M. K. Pandey[a]*, G. N. Singh[a] and A. Bandyopadhyay[b]

[a]Department of Mathematics and Computing, Indian Institute of Technology (Indian School of Mines), Dhanbad- 826004, India;
[b]Department of Mathematics, Asansol Engineering College, Asansol, West Bengal, India

*Corresponding Author Email ID: *maheshbabu3797@gmail.com (Pandey, M. K.).*
Cell:+919794455683

Email: gnsingh@iitism.ac.in (Singh, G. N.), arnab.bandyopadhyay4@gmail.com (Bandyopadhyay, A.)

Abstract

In this paper, we introduce estimators for the population coefficient of variation within a two-occasion, stratified, successive sampling framework, aiming to mitigate the impact of non-response and measurement errors. We derive calibrated weights for the strata and thoroughly examine the properties of the proposed estimator through comprehensive numerical and simulation studies. Furthermore, we provide valuable recommendations for survey statisticians, guiding them on effective applications in real-world survey scenarios. By addressing the challenges of non-response and measurement errors within a stratified sampling approach, our proposed estimators aim to enhance the accuracy and precision of coefficient of variation estimates, ensuring more precise and accurate results.

Keywords: Two-occasion stratified successive sampling, coefficient of variation, calibration technique, bias, mean square error, random non-response, measurement error
2010 Mathematics Subject Classification: 62D05

1 Introduction

In the field of socio-economic research, accurate estimation of population parameters is essential for drawing meaningful inferences and making informed decisions. The coefficient of variation (CV) plays a pivotal role in this regard, measuring the relative variability of a population's characteristics. Facilitating the comparison of variability across different units, the CV, when expressed as a percentage, quickly illustrates the extent of variability present in the data. Its utility extends beyond socio-economic research, being common in various applied probability fields such as renewal theory, queuing theory, and reliability theory. Despite its significance, recent estimation techniques for the coefficient of variation often rely on complete information from sampled units, as observed in the work of Tripathi et al. [1], Archana and Rao [2]. However, such assumptions disregard the reality of data in real-life surveys, which are susceptible to non-sampling errors.

Non-response, the absence of data from certain respondents regarding specific variables, poses a significant challenge in data collection. Factors such as unavailability of respondents, reluctance to answer sensitive questions, or simply a lack of information contribute to non-response. For instance, in surveys targeting human populations, obtaining information from all selected units can be challenging, especially in mail surveys where respondents are asked to return completed questionnaires by a deadline. Non-response can manifest in various patterns and stem from diverse causes, affecting the representativeness of the sample and, subsequently, the accuracy of CV estimation. In agricultural production surveys, non-response may occur due to crop loss or damage from natural disasters, leading to missing data for certain seasons.

Measurement errors, discrepancies between recorded values and true values of variables under study, further complicate accurate estimation. Errors may arise from over-reporting, under-reporting, memory failures, interviewer biases, or defective measurement mechanisms. For instance, in surveys on household consumption, respondents may struggle to recall expenditure details accurately, leading to distorted data. Inaccurate measurements can skew variability and mean values, resulting in biased CV estimates, particularly problematic when comparing variability across different groups or over time.

Estimating the CV faces additional challenges when dealing with non-response and measurement errors. Several authors Sisodia and Dwivedi [3], Das and Tripathi [4, 5], Patel and Rina [6], Singh et al. [7], Muneer et al. [8], Yunusa et al. [9], Audu et al. [10], Shahzad et al. [11, 12], Yadav et al. [13], Rajyaguru and Gupta [14] and others have proposed estimators based on the simple random sampling scheme, which assumes accessibility to all sampling units and complete information without measurement errors. However, these assumptions are often unrealistic in real-life situations.

In situations where populations undergo continuous change, a single survey provides insights into the characteristics of the surveyed population for that specific occasion only. However, this singular approach fails to offer information about the rate of change over different occasions or the average value of characteristics across all occasions. To address these limitations, successive sampling is employed, as seen in scenarios like monthly data collection on goods prices to determine the consumer price index or periodic political opinion surveys to gauge voter preferences. Despite the widespread use of successive sampling in scientific and socio-economic surveys, existing research has primarily concentrated on developing estimators for population mean or variance, overlooking the crucial aspect of the population coefficient of variation (CV). Our proposed work seeks to bridge this gap by suggesting suitable estimation procedures in successive sampling, specifically addressing the challenges posed by non-response and measurement errors. In this context, where samples are taken on two occasions—match samples and fresh samples—non-response and measurement errors can independently or simultaneously affect either sample, presenting a significant challenge in estimating population parameters, including the often-neglected CV. This research builds upon the foundational work by Jessen [15] and subsequent expansions by researchers such as Yates et al. [16], Eckler [17], Sen [18], Feng and Zou [19], Singh and Homa [20], Naz et al. [21], Younis and Shabbir [22], Abid et al. [23], Irfan et al. [24], Bhushan and Pandey [25], Sen [26] and others who have done recent work in this field, has primarily focused on developing estimators for population mean or variance in the presence of non-response or measurement errors. Notably, the population coefficient of variation has been largely overlooked in this body of work Allen et al. [27], Kumar et al. [28], Ahmed and Shabbir [29], Audu et al. [30], Shahzad et al. [31].

Therefore, the aim of this research paper is to propose an improved estimation procedure for the population coefficient of variation (CV) under successive sampling, considering the presence of non-response and measurement errors, by utilizing calibrated weights. This introduces a fresh perspective on estimating the CV within a stratified successive sampling framework. While existing techniques primarily focus on mean or variance, our approach uniquely targets the CV, addressing a critical gap in the literature. The coefficient of variation (CV) is preferable to mean alone because it is unit-free, more stable in comparison to mean and variance, and facilitates comparisons between different populations, making it a valuable measure of dispersion across populations regardless of their scales or units of measurement. Incorporating CV estimation into stratified successive sampling provides a comprehensive understanding of variability within and between populations, offering valuable insights for decision-making and policy formulation. A significant aspect in the context of stratified successive sampling under non-response is the limited research conducted on the estimation of the coefficient of variation, which is particularly important given the challenges posed by non-response and measurement errors during data collection. Our proposed procedure combines existing methods for estimating the population CV with a model-based approach that accounts for non-response and measurement errors.

Deming and Stephan [32] introduced a calibration approach using least squares adjustment, which was later adopted by statistical authorities in various organizations. The main goal of the calibration approach is to formulate unbiased estimation procedures with the least amount of dispersion using the information on auxiliary variables. In follow-up, Deville and Särndal [33] proposed a calibration estimation procedure that decreases the distance between the initial and final weights while still respecting the calibration equations and constraints. Subsequently, Farrell and Singh [34], Särndal [35], Kim et al. [36], Kim and Park [37], Sud et al. [38], Singh et al. [39], Koyuncu and Kadilar [40], Nidhi et al. [41], Özgül [42], Shahzad et al. [43, 44], Pandey et al. [45-47], Clement [48] and others have produced notable calibrated estimation procedures.

In the estimation of population variance using stratified successive sampling, calibration is a technique used to improve the accuracy of the estimates. By incorporating calibration into the estimation process, resulting estimates are more representative of the population and have reduced bias. This is especially important in stratified successive sampling, where the goal is to ensure that each stratum is well-represented in the final estimate.

To demonstrate the practical relevance of our proposed method, we apply it to a real-life socio-economic example. Specifically, we consider the case of estimating the CV of household income in a developing country where non-response and measurement errors are prevalent due to the absence of reliable income data. Our proposed method provides a more accurate estimation of the population CV and helps to better understand the income distribution in the population, which has important implications for socio-economic policies and decision-making.

In the subsequent sections of this research paper, we will present the theoretical background and methodology of our proposed estimation procedure. We will then provide a detailed description of the application of our method to the real-life socio-economic example. Finally, we will discuss the results and implications of our research, highlighting the advantages and limitations of the proposed approach.

2 Sample structure and notations

Consider a finite population of size N divided into G non-overlapping strata, each containing N_k (k=1, 2,...,G) units. Let us use X and Y to represent the study character on the first and second occasions. It is assumed that information regarding an auxiliary variable Z is accessible on both occasions, and the population variance of Z is known.

Let us consider the kth strata, where k ranges from 1 to G., To begin with, we use simple random sampling without replacement (SRSWOR) to draw a preliminary sample of size n_k from the population for the first occasion, where r_{1k} units do not respond. From the responding part of this sample, we draw a second stage SRSWOR sample of size $m_k = n_k \lambda_k^{''}$, where $\lambda_k^{''}$ is the fraction of matched samples, and r_{2k} units do not respond. We use this sample for the second occasion and collect information on the study variable Y. Additionally, we draw a fresh sample of size $u_k = n_k - m_k = n_k u_k^{''}$ from the population using SRSWOR on Y again. Here, r_{3k} units do not respond. The fractions of matched and fresh samples on the current (second) occasion are represented by $\lambda_k^{''}$ and $\mu_k^{''}$, respectively, where $\lambda_k^{''} + \mu_k^{''} = 1$.

From now on, we will use the following notations:

$\overline{X}_k$, $\overline{Y}_k$: The population mean of study variables X and Y respectively in the kth strata.

$\overline{Z}_k$: The population mean of the auxiliary variable Z in the kth stratum.

$\bar{y}_{n_k}$, $\bar{y}_{m_k}$, $\bar{y}_{u_k}$, $\bar{x}_{n_k}$, $\bar{x}_{m_k}$, $\bar{x}_{u_k}$: The sample means of the variables Y and X respectively based on the respective sample sizes shown in suffice.

$\bar{z}_{n_k} = \frac{1}{n_k}\sum_{l=1}^{n_k} z_{kl}$, $\bar{z}_{m_k} = \frac{1}{m_k}\sum_{l=1}^{m_k} z_{kl}$ and $\bar{z}_{u_k} = \frac{1}{u_k}\sum_{l=1}^{u_k} z_{kl}$: The sample means of the auxiliary variable in the kth stratum are determined based on a sample size of n_k, m_k and u_k, respectively.

$S_{Y_{N_k}}^2 = \frac{1}{N_k - 1}\sum_{j=1}^{N_k} (Y_{k_j} - \overline{Y}_{N_k})^2$, $S_{X_{N_k}}^2 = \frac{1}{N_k - 1}\sum_{j=1}^{N_k} (X_{k_j} - \overline{X}_{N_k})^2$: The population mean squares of the kth stratum of the study variables Y and X, respectively.

$S_{Z_{N_k}}^2 = \frac{1}{N_k - 1}\sum_{j=1}^{N_k}(Z_{k_j} - \overline{Z}_{N_k})^2$: The population mean squares of the kth stratum for the auxiliary variable Z.

$s_{x_{n_k-r_{1k}}}^{*2} = \frac{1}{n_k - r_{1k} - 1}\sum_{l=1}^{n_k-r_{1k}}(x_{k_l} - \overline{x}_{n_k-r_{1k}})^2$: Depending on the responding part of sample of size n_k, the sample mean square of study variable X for the kth stratum.

$s_{x_{m_k-r_{2k}}}^{*2} = \frac{1}{m_k - r_{2k} - 1}\sum_{l=1}^{m_k-r_{2k}}(x_{k_l} - \overline{x}_{m_k-r_{2k}})^2$: Depending on the responding part of sample of size m_k, the sample mean square of study variable X for the kth stratum.

$s_{y_{n_k-r_{1k}}}^{*2} = \frac{1}{n_k - r_{1k} - 1}\sum_{l=1}^{n_k-r_{1k}}(y_{k_l} - \overline{y}_{n_k-r_{1k}})^2$: Depending on the responding part of sample of size n_k, the sample mean square of study variable Y for the kth stratum.

$s_{y_{u_k-r_{3k}}}^{*2} = \frac{1}{u_k - r_{3k} - 1}\sum_{l=1}^{u_k-r_{3k}}(y_{k_l} - \overline{y}_{u_k-r_{3k}})^2$: Depending on the responding part of sample of size u_k, the sample mean square of study variable Y for the kth stratum.

$s_{z_{n_k}}^{*2} = \frac{1}{n_k - 1}\sum_{l=1}^{n_k}(z_{k_l} - \overline{z}_{n_k})^2$: Depending on the sample of size n_k, the sample mean square of auxiliary variable Z for the kth stratum.

$s_{z_{m_k}}^{*2} = \frac{1}{m_k - 1}\sum_{l=1}^{m_k}(z_{k_l} - \overline{z}_{m_k})^2$: Depending on the sample of size m_k, the sample mean square of auxiliary variable Z for the kth stratum. $s_{z_{u_k}}^{*2} = \frac{1}{u_k - 1}\sum_{l=1}^{u_k}(z_{k_l} - \overline{z}_{u_k})^2$: Depending on the sample of size u_k, the sample mean square of auxiliary variable Z for the kth stratum.

$c_{y_{u_k-r_{3k}}} = \frac{s_{y_{u_k-r_{3k}}}}{\overline{y}_{u_k-r_{3k}}}$, $c_{x_{m_k-r_{2k}}} = \frac{s_{x_{m_k-r_{2k}}}}{\overline{x}_{m_k-r_{2k}}}$, $c_{y_{n_k-r_{1k}}} = \frac{s_{y_{n_k-r_{1k}}}}{\overline{y}_{n_k-r_{1k}}}$, $c_{x_{n_k-r_{1k}}} = \frac{s_{x_{n_k-r_{1k}}}}{\overline{x}_{n_k-r_{1k}}}$, $c_{z_{m_k}} = \frac{s_{z_{m_k}}}{\overline{z}_{m_k}}$, $c_{z_{n_k}} = \frac{s_{z_{n_k}}}{\overline{z}_{n_k}}$,

$c_{z_{u_k}} = \frac{s_{z_{u_k}}}{\overline{z}_{u_k}}$: The sample coefficients of variation for the variables Y, X and Z respectively based on the respective sample sizes shown in suffice.

$\rho_{XY_{N_k}}$, $\rho_{YZ_{N_k}}$, $\rho_{ZX_{N_k}}$: The population correlation coefficients between the variables shown suffice for the kth stratum.

$S_{X_{N_k}}^2$, $S_{Y_{N_k}}^2$, $S_{Z_{N_k}}^2$: The population mean squares of the variables X, Y, and Z for the kth stratum respectively.

$C_{X_{N_k}}$, $C_{Y_{N_k}}$, $C_{Z_{N_k}}$, $C_{XY_{N_k}}$, $C_{YZ_{N_k}}$, $C_{ZX_{N_k}}$: The coefficient of variation based on the variables in the suffices.

3 Non-response probability model

The kth stratum is considered using random non-response model Singh and Joarder [49]. Consider a sample S_{n_k} of size n_k for which some data on X could not be collected due to random non-response. Let r_{1k}, where r_{1k} ranges from

0, 1, 2,...,(n_k - 2), represent the number of such cases in S_{n_k} . Similarly, for a sample S_{m_k} of size m_k , let r_{2k} represent the number of units for which Y information on the second occasion could not be acquired due to random non-response, and r_{3k} represent the same for a sample of size u_k . It is presumed that r_{1k} , r_{2k} , and r_{3k} fall within their respective bounds. We assume that $0 \le r_{1k} \le (n_k - 2)$, $0 \le r_{2k} \le (m_k - 2)$ and $0 \le r_{3k} \le (u_k - 2)$. If p_1 , p_2 , and p_3 probabilities of non-response among the $(n_k - 2)$, $(m_k - 2)$, and $(u_k - 2)$ possible values of non-responses respectively, the discrete probability distributions for r_{1k} , r_{2k} , and r_{3k} are represented by

$$P(r_{1k}) = \frac{n_k - r_{1k}}{n_k q_1 + 2p_1} \binom{n_k - 2}{r_{1k}} p_1^{r_{1k}} q_1^{n_k - r_{1k} - 2} \;\; ; r_{1k} = 0, 1, 2, ..., n_k - 2$$

$$P(r_{2k}) = \frac{m_k - r_{2k}}{n_k q_2 + 2p_2} \binom{m_k - 2}{r_{2k}} p_2^{r_{2k}} q_2^{m_k - r_{2k} - 2} \;\; ; r_{2k} = 0, 1, 2, ..., m_k - 2$$

and $P(r_{3k}) = \dfrac{u_k - r_{3k}}{n_k q_3 + 2p_3} \dbinom{u_k - 2}{r_{3k}} p_3^{r_{3k}} q_3^{u_k - r_{3k} - 2} \;\; ; r_{3k} = 0, 1, 2, ..., u_k - 2$, respectively.

The number of ways to obtain r_{lk} (l = 1, 2, 3) non-responses from all potential non-response values for the three samples are represented by $\binom{n_k - 2}{r_{1k}}$, $\binom{m_k - 2}{r_{2k}}$, and $\binom{u_k - 2}{r_{3k}}$.

4 Proposed estimator

After considering the aforementioned discussions and building upon the research conducted by Bahl and Tuteja [50], we propose the estimator T_{m_k} for the population coefficient of variation, derived from the sample S_m of size *m* that is shared between both occasions.

$$T_{m_k} = c_{y_{n_k - r_{1k}}} + a_k (c^*_{x_{n_k}} - c^*_{x_{m_k}}) \tag{1}$$

The scalar quantities a_k may be determined by minimizing the mean square error of the estimators T_{m_k} . c $c^*_{x_{m_k}} = c_{x_{m_k - r_{2k}}} + (c_{z_{m_k}} - C_{Z_{N_k}})$ and $c^*_{x_{n_k}} = c_{x_{n_k - r_{1k}}} + (c_{z_{n_k}} - C_{Z_{N_k}})$. We also propose the estimators T_{u_k} for the population coefficient of variation. These estimators are based on a fresh sample S_u of size *u*, drawn on the current occasion.

$$T_{u_k} = c_{y_{u_k - r_{3k}}} + b_k (c_{z_{u_k}} - C_{Z_{N_k}}) \tag{2}$$

Where the scalar quantities b_k may be determined by minimizing the mean square error of the estimators T_{u_k} .

We use the estimators T_{m_k} and T_{u_k} , defined in Equations (1) and (2), to estimate the coefficient of variance. Additionally, we suggest using the estimators T_m and T_u to estimate the coefficient of variance for the matched sample and the fresh sample, respectively.

$$T_m = \sum_{k=1}^{G} \Omega^*_k T_{m_k} \tag{3}$$

and

$$T_u = \sum_{k=1}^{G} \Omega^{**}_k T_{u_k} \tag{4}$$

Remark 1. We have observed that the structure remains consistent in large sample approximations. Specifically, if m_k, u_k, and $n_k \to N_k$, then $c_{z_{n_k}} \to C_{Z_{N_k}}$, $c_{z_{m_k}} \to C_{Z_{N_k}}$, $c_{z_{u_k}} \to C_{Z_{N_k}}$, $c_{y_{n_k-r_{1k}}} \to C_{Y_{N_k}}$, and $c_{y_{u_k-r_{3k}}} \to C_{Y_{N_k}}$. By utilizing the identity, we may conclude that $c_{x_{n_k}} \to C_{X_{N_k}}$, and also $c_{x_{m_k}} \to C_{X_{N_k}}$. Therefore, it may be inferred that $T_{m_k} \to C_{Y_{N_k}}$ and $T_{u_k} \to C_{Y_{N_k}}$.

5 Suggested calibration technique

The new calibrated weights for the estimator of the population variance $T_m = \sum_{k=1}^{G} \Omega_k^* T_{m_k}$ under stratified sampling are provided by the values acquired through the minimization of the chi-square distance function $\sum_{k=1}^{G} \frac{(\Omega_k^* - W_k)^2}{Q_k W_k}$, considering the following calibration constraints.

1. $\sum_{k=1}^{G} \Omega_k^* = 1$
2. $\sum_{k=1}^{G} \Omega_k^* \log(\bar{z}_{n_k}) = \sum_{k=1}^{G} W_k \log(\bar{Z}_k)$
3. $\sum_{k=1}^{G} \Omega_k^* c_{x_{m_k-r_{2k}}} = \sum_{k=1}^{G} W_k c_{x_{n_k-r_{1k}}}$

where $c_{x_{m_k-r_{2k}}} = \frac{s_{x_{m_k-r_{2k}}}}{\bar{x}_{m_k-r_{2k}}}$ and $c_{x_{n_k-r_{1k}}} = \frac{s_{x_{n_k-r_{1k}}}}{\bar{x}_{n_k-r_{1k}}}$.

The Lagrange function may be expressed by utilizing the chi-square distance measure and the calibration constraints mentioned earlier as follows:

$$L_m = \sum_{k=1}^{G} \frac{(\Omega_k^* - W_k)^2}{Q_k W_k} - 2\lambda_{1m}\left(\sum_{k=1}^{G} \Omega_k^* - 1\right) - 2\lambda_{2m}\left(\sum_{k=1}^{G} \Omega_k^* \log(\bar{z}_{n_k}) - \sum_{k=1}^{G} W_k \log(\bar{Z}_k)\right)$$
$$-2\lambda_{3m}\left(\sum_{k=1}^{G} \Omega_k^* c_{x_{m_k-r_{2k}}} - \sum_{k=1}^{G} W_k c_{x_{n_k-r_{1k}}}\right) \tag{5}$$

where λ_{1m}, λ_{2m} and λ_{3m} are Lagrange multipliers.

Upon differentiation of Equation (5) with respect to Ω_k^*, we may determine the calibrated weights by setting the resulting expression equal to zero, as shown.

$$\Omega_k^* = W_k + (\lambda_{1m} + \lambda_{2m} \log(\bar{z}_{n_k}) + \lambda_{3m} c_{x_{m_k-r_{2k}}}) W_k Q_k \tag{6}$$

Put the value Ω_k^* in the above constraints, we get the matrix form as...

$$\begin{bmatrix} cal_{am} & cal_{bm} & cal_{cm} \\ cal_{bm} & cal_{em} & cal_{fm} \\ cal_{cm} & cal_{fm} & cal_{hm} \end{bmatrix} \begin{bmatrix} \lambda_{1m} \\ \lambda_{2m} \\ \lambda_{3m} \end{bmatrix} = \begin{bmatrix} cal_{dm} \\ cal_{gm} \\ cal_{im} \end{bmatrix}$$

The solution of the above matrix provides the values of the Lagrange multipliers, as stated below:

$$\lambda_{1m}=\frac{\det_{\alpha m}}{\det_m},\lambda_{2m}=\frac{\det_{\beta m}}{\det_m},\&\lambda_{3m}=\frac{\det_{\gamma m}}{\det_m} \tag{7}$$

where

$$\det_m = cal_{am}cal_{em}cal_{hm}-cal_{am}cal_{fm}^2-cal_{bm}^2cal_{hm}+2cal_{bm}cal_{cm}cal_{fm}-cal_{em}cal_{cm}^2 \tag{8}$$

$$\begin{aligned}\det_{\alpha m} &= cal_{dm}cal_{em}cal_{hm}-cal_{dm}cal_{fm}^2-cal_{bm}cal_{gm}cal_{hm}+cal_{bm}cal_{im}cal_{fm}\\ &\quad +cal_{cm}cal_{gm}cal_{fm}-cal_{cm}cal_{im}cal_{em}\end{aligned} \tag{9}$$

$$\begin{aligned}\det_{\beta m} &= cal_{am}cal_{gm}cal_{hm}-cal_{am}cal_{im}cal_{fm}-cal_{bm}cal_{dm}cal_{hm}+cal_{cm}cal_{dm}cal_{fm}\\ &\quad +cal_{bm}cal_{cm}cal_{im}-cal_{cm}^2cal_{gm}\end{aligned} \tag{10}$$

$$\begin{aligned}\det_{\gamma m} &= cal_{am}cal_{em}cal_{im}-cal_{am}cal_{gm}cal_{fm}-cal_{bm}^2cal_{im}+cal_{bm}cal_{cm}cal_{gm}\\ &\quad +cal_{bm}cal_{dm}cal_{fm}-cal_{cm}cal_{dm}cal_{em}\end{aligned} \tag{11}$$

Now, let us define the term cal_{am}, cal_{bm}, cal_{cm}, cal_{dm}, cal_{em}, cal_{fm}, cal_{gm}, cal_{hm}, cal_{im} as follows:

$cal_{am}=\sum_{k=1}^{G}W_kQ_k$	$cal_{bm}=\sum_{k=1}^{G}W_kQ_k\log(\bar{z}_{n_k})$
$cal_{cm}=\sum_{k=1}^{G}W_kQ_kc_{x_{m_k-r_2k}}$	$cal_{dm}=1-\sum_{k=1}^{G}W_kQ_k$
$cal_{em}=\sum_{k=1}^{G}W_kQ_k(\log(\bar{z}_{n_k}))^2$	$cal_{fm}=\sum_{k=1}^{G}W_kQ_k\log(\bar{z}_{n_k})c_{x_{m_k-r_2k}}$
$cal_{gm}=\sum_{k=1}^{G}W_kQ_k(\log(\bar{Z}_k)-\log(\bar{z}_{n_k}))$	$cal_{hm}=\sum_{k=1}^{G}W_kQ_kc^2_{x_{m_k-r_2k}}$
$cal_{im}=\sum_{k=1}^{G}W_kc_{x_{n_k-r_1k}}-\sum_{k=1}^{G}W_kc_{x_{m_k-r_2k}}$	

Now, the new calibrated weights for the estimator of the population variance, $T_u=\sum_{k=1}^{G}\Omega_k^{**}T_{u_k}$, under stratified sampling are provided. The values are acquired through the minimization of the chi-square distance function $\sum_{k=1}^{G}\frac{(\Omega_k^{**}-W_k)^2}{Q_kW_k}$, taking into account the following calibration constraints.

1. $\sum_{k=1}^{G}\Omega_k^{**}=1$
2. $\sum_{k=1}^{G}\Omega_k^{**}c_{\bar{z}_{u_k}}=C_Z$

where $c_{\bar{z}_{u_k}}=\frac{s_{z_{u_k}}}{\bar{z}_{u_k}}$ and $C_Z=\frac{S_Z}{\bar{Z}}$.

The Lagrange function may be expressed by utilizing the chi-square distance measure and the calibration constraints mentioned earlier as follows:

$$L_u=\sum_{k=1}^{G}\frac{(\Omega_k^{**}-W_k)^2}{Q_kW_k}-2\lambda_{1u}(\sum_{k=1}^{G}\Omega_k^{**}-1)-2\lambda_{2u}(\sum_{k=1}^{G}\Omega_k^{**}c_{\bar{z}_{u_k}}-C_Z) \tag{12}$$

where λ_{1u} and λ_{2u} are Lagrange multipliers.

Upon differentiation of Equation (12) with respect to Ω_k^{**}, we may determine the calibrated weights by setting the resulting expression equal to zero, as shown.

$$\Omega_k^{**}=W_k+(\lambda_{1u}+\lambda_{2u}c_{\bar{z}_{u_k}})W_kQ_k \tag{13}$$

Put the value Ω_k^{**} in the above constraints, we get the matrix form as...

$$\begin{bmatrix} cal_{au} & cal_{bu} \\ cal_{bu} & cal_{du} \end{bmatrix}\begin{bmatrix} \lambda_{1u} \\ \lambda_{2u} \end{bmatrix}=\begin{bmatrix} cal_{cu} \\ cal_{eu} \end{bmatrix}$$

The solution of the above matrix provides the values of the Lagrange multipliers, as stated below:

$$\lambda_{1u}=\frac{\det_{\alpha u}}{\det_u},\&\lambda_{2u}=\frac{\det_{\beta u}}{\det_u} \tag{14}$$

where

$$\det_u=cal_{au}cal_{du}-cal_{bu}^2 \tag{15}$$

$$\det_{\alpha u}=cal_{cu}cal_{du}-cal_{bu}cal_{eu} \tag{16}$$

$$\det_{\beta u}=cal_{au}cal_{eu}-cal_{bu}cal_{cu} \tag{17}$$

Now, let us define the term cal_{au}, cal_{bu}, cal_{cu}, cal_{du}, cal_{eu} as follows:

$cal_{au}=\sum_{k=1}^{G}W_kQ_k$	$cal_{bu}=\sum_{k=1}^{G}W_kQ_kc_{\bar{z}_{u_k}}$
$cal_{cu}=1-\sum_{k=1}^{G}W_k$	$cal_{du}=\sum_{k=1}^{G}W_kQ_kc_{\bar{z}_{u_k}}^2$
$cal_{eu}=C_Z-\sum_{k=1}^{G}W_kc_{\bar{z}_{u_k}}$	

5.1 Properties of the proposed estimators

We have derived the bias and mean square errors of the proposed estimators using the following transformations, under the assumptions of a large sample size and up to the first order of approximations:

$\bar{y}_{n_k-r_{1k}}=\bar{Y}_{N_k}(1+\varepsilon_0)$	$\bar{x}_{m_k-r_{2k}}=\bar{X}_{N_k}(1+\varepsilon_1)$	$\bar{x}_{n_k-r_{1k}}=\bar{X}_{N_k}(1+\varepsilon_2)$
$\bar{y}_{u_k-r_{3k}}=\bar{Y}_{N_k}(1+\varepsilon_3)$	$\bar{z}_{m_k}=\bar{Z}_{N_k}(1+\varepsilon_4)$	$\bar{z}_{n_k}=\bar{Z}_{N_k}(1+\varepsilon_5)$
$\bar{z}_{u_k}=\bar{Z}_{N_k}(1+\varepsilon_6)$	$S_{y_{n_k-r_{1k}}}^2=S_{Y_{N_k}}^2(1+\varepsilon_7)$	$s_{z_{m_k}}^2=S_{Z_{N_k}}^2(1+\varepsilon_8)$

$s^2_{z_{n_k}} = S^2_{Z_{N_k}}(1+\varepsilon_9)$	$S^2_{x_{n_k - r_{1k}}} = S^2_{X_{N_k}}(1+\varepsilon_{10})$	$S^2_{y_{u_k - r_{3k}}} = S^2_{Y_{N_k}}(1+\varepsilon_{11})$
$s^2_{z_{u_k}} = S^2_{Z_{N_k}}(1+\varepsilon_{12})$	$S^2_{x_{m_k - r_{2k}}} = S^2_{X_{N_k}}(1+\varepsilon_{13})$	

Such that $E(\varepsilon_i) = 0$ and $|\varepsilon_i| \le 1, \forall i = 0,1,2,...,13$. Using these conditions, we may obtain the following expectations:

$E(\varepsilon_0^2) = f_{1k}C^2_{Y_{N_k}}$	$E(\varepsilon_1^2) = f_{6k}C^2_{X_{N_k}}$	$E(\varepsilon_2^2) = f_{1k}C^2_{X_{N_k}}$
$E(\varepsilon_3^2) = f_{2k}C^2_{Y_{N_k}}$	$E(\varepsilon_4^2) = f_{5k}C^2_{Z_{N_k}}$	$E(\varepsilon_6^2) = f_{3k}C^2_{Z_{N_k}}$
$E(\varepsilon_7^2) = f_{1k}(\lambda_{400k} - 1)$	$E(\varepsilon_8^2) = f_{5k}(\lambda_{004k} - 1)$	$E(\varepsilon_9^2) = f_{4k}(\lambda_{004k} - 1)$
$E(\varepsilon_{10}^2) = f_{1k}(\lambda_{040k} - 1)$	$E(\varepsilon_{11}^2) = f_{2k}(\lambda_{400k} - 1)$	$E(\varepsilon_{12}^2) = f_{3k}(\lambda_{004k} - 1)$
$E(\varepsilon_{13}^2) = f_{6k}(\lambda_{040k} - 1)$	$E(\varepsilon_0\varepsilon_1) = f_{1k}C_{XY_{N_k}}$	$E(\varepsilon_0\varepsilon_2) = f_{1k}C_{XY_{N_k}}$
$E(\varepsilon_0\varepsilon_4) = f_{1k}C_{YZ_{N_k}}$	$E(\varepsilon_0\varepsilon_5) = f_{4k}C_{XY_{N_k}}$	$E(\varepsilon_0\varepsilon_7) = f_{1k}C_{Y_{N_k}}\lambda_{300k}$
$E(\varepsilon_0\varepsilon_8) = f_{1k}C_{Y_{N_k}}\lambda_{003k}$	$E(\varepsilon_0\varepsilon_9) = f_{4k}C_{Y_{N_k}}\lambda_{003k}$	$E(\varepsilon_0\varepsilon_{10}) = f_{1k}C_{Y_{N_k}}\lambda_{030k}$
$E(\varepsilon_0\varepsilon_{13}) = f_{1k}C_{Y_{N_k}}\lambda_{030k}$	$E(\varepsilon_1\varepsilon_2) = f_{1k}C^2_{X_{N_k}}$	$E(\varepsilon_1\varepsilon_4) = f_{5k}C_{XZ_{N_k}}$
$E(\varepsilon_1\varepsilon_5) = f_{4k}C_{XZ_{N_k}}$	$E(\varepsilon_1\varepsilon_7) = f_{1k}C_{X_{N_k}}\lambda_{210k}$	$E(\varepsilon_1\varepsilon_8) = f_{5k}C_{X_{N_k}}\lambda_{012k}$
$E(\varepsilon_1\varepsilon_9) = f_{4k}C_{X_{N_k}}\lambda_{012k}$	$E(\varepsilon_1\varepsilon_{10}) = f_{1k}C_{X_{N_k}}\lambda_{030k}$	$E(\varepsilon_1\varepsilon_{13}) = f_{6k}C_{X_{N_k}}\lambda_{030k}$
$E(\varepsilon_2\varepsilon_4) = f_{1k}C_{XZ_{N_k}}$	$E(\varepsilon_2\varepsilon_5) = f_{4k}C_{XZ_{N_k}}$	$E(\varepsilon_2\varepsilon_7) = f_{1k}C_{X_{N_k}}\lambda_{210k}$
$E(\varepsilon_2\varepsilon_8) = f_{1k}C_{X_{N_k}}\lambda_{012k}$	$E(\varepsilon_2\varepsilon_9) = f_{4k}C_{X_{N_k}}\lambda_{012k}$	$E(\varepsilon_2\varepsilon_{10}) = f_{1k}C_{X_{N_k}}\lambda_{030k}$
$E(\varepsilon_2\varepsilon_{13}) = f_{1k}C_{X_{N_k}}\lambda_{030k}$	$E(\varepsilon_3\varepsilon_6) = f_{3k}C_{YZ_{N_k}}$	$E(\varepsilon_3\varepsilon_{11}) = f_{2k}C_{Y_{N_k}}\lambda_{300k}$
$E(\varepsilon_3\varepsilon_{12}) = f_{3k}C_{Y_{N_k}}\lambda_{102k}$	$E(\varepsilon_4\varepsilon_5) = f_{4k}C^2_{Z_{N_k}}$	$E(\varepsilon_4\varepsilon_7) = f_{1k}C_{Z_{N_k}}\lambda_{201k}$
$E(\varepsilon_4\varepsilon_8) = f_{5k}C_{Z_{N_k}}\lambda_{003k}$	$E(\varepsilon_4\varepsilon_9) = f_{4k}C_{Z_{N_k}}\lambda_{003k}$	$E(\varepsilon_4\varepsilon_{10}) = f_{1k}C_{Z_{N_k}}\lambda_{021k}$
$E(\varepsilon_4\varepsilon_{13}) = f_{5k}C_{Z_{N_k}}\lambda_{021k}$	$E(\varepsilon_5\varepsilon_7) = f_{4k}C_{Z_{N_k}}\lambda_{201k}$	$E(\varepsilon_5\varepsilon_8) = f_{4k}C_{Z_{N_k}}\lambda_{003k}$
$E(\varepsilon_5\varepsilon_9) = f_{4k}C_{Z_{N_k}}\lambda_{003k}$	$E(\varepsilon_5\varepsilon_{10}) = f_{4k}C_{Z_{N_k}}\lambda_{021k}$	$E(\varepsilon_5\varepsilon_{13}) = f_{4k}C_{Z_{N_k}}\lambda_{021k}$
$E(\varepsilon_6\varepsilon_{11}) = f_{3k}C_{Z_{N_k}}\lambda_{201k}$	$E(\varepsilon_6\varepsilon_{12}) = f_{3k}C_{Z_{N_k}}\lambda_{003k}$	$E(\varepsilon_7\varepsilon_8) = f_{1k}(\lambda_{202k} - 1)$
$E(\varepsilon_7\varepsilon_9) = f_{4k}(\lambda_{202k} - 1)$	$E(\varepsilon_7\varepsilon_{10}) = f_{1k}(\lambda_{220k} - 1)$	$E(\varepsilon_7\varepsilon_{13}) = f_{1k}(\lambda_{220k} - 1)$
$E(\varepsilon_8\varepsilon_9) = f_{4k}(\lambda_{004k} - 1)$	$E(\varepsilon_8\varepsilon_{10}) = f_{1k}(\lambda_{022k} - 1)$	$E(\varepsilon_8\varepsilon_{13}) = f_{5k}(\lambda_{022k} - 1)$
$E(\varepsilon_9\varepsilon_{10}) = f_{4k}(\lambda_{022k} - 1)$	$E(\varepsilon_9\varepsilon_{13}) = f_{4k}(\lambda_{022k} - 1)$	$E(\varepsilon_{10}\varepsilon_{13}) = f_{1k}(\lambda_{040k} - 1)$
$E(\varepsilon_{11}\varepsilon_{12}) = f_{3k}(\lambda_{202k} - 1)$		

where

$$f_{1k} = (\frac{1}{n_k q_1 + 2p_1} - \frac{1}{N_k}), \qquad f_{2k} = (\frac{1}{u_k q_3 + 2p_3} - \frac{1}{N_k}), \qquad f_{3k} = (\frac{1}{u_k} - \frac{1}{N_k})$$

$$f_{4k} = (\frac{1}{n_k} - \frac{1}{N_k}), \qquad f_{5k} = (\frac{1}{m_k} - \frac{1}{N_k}) \qquad f_{6k} = (\frac{1}{m_k q_2 + 2p_2} - \frac{1}{N_k})$$

and

$$\lambda_{\alpha\beta\gamma k} = \frac{\mu_{\alpha\beta\gamma k}}{\sqrt{\mu_{200k}^{\alpha}\mu_{020k}^{\beta}\mu_{002k}^{\gamma}}}, \quad \mu_{\alpha\beta\gamma k} = \frac{1}{N_k - 1}\sum_{l=1}^{N_k}(Y_{k_l} - \overline{Y}_k)^{\alpha}(X_{k_l} - \overline{X}_k)^{\beta}(Z_{k_l} - \overline{Z}_k)^{\gamma}$$

Note. If there is no non-response, then p_1, p_2, and p_3 are all zero, and as a result, q_1, q_2, and q_3 are all one. In this case, the above values are identical to the standard results.

By applying these transformations to $c^*_{x_{n_k}}$ and $c^*_{x_{m_k}}$, we obtain:

$$c^*_{x_{n_k}} = C_{X_{N_k}}(1 - \varepsilon_2 + \frac{\varepsilon_{10}}{2} + \varepsilon_2^2 - \frac{\varepsilon_{10}^2}{8} - \frac{\varepsilon_2\varepsilon_{10}}{2} -) + (\frac{\varepsilon_9}{2} - \varepsilon_5 + \varepsilon_5^2 - \frac{\varepsilon_9^2}{8} - \frac{\varepsilon_5\varepsilon_9}{2} -) \tag{18}$$

$$c^*_{x_{m_k}} = C_{X_{N_k}}(1 - \varepsilon_1 + \frac{\varepsilon_{13}}{2} + \varepsilon_1^2 - \frac{\varepsilon_{13}^2}{8} - \frac{\varepsilon_1\varepsilon_{13}}{2} -) + (\frac{\varepsilon_8}{2} - \varepsilon_4 + \varepsilon_4^2 - \frac{\varepsilon_8^2}{8} - \frac{\varepsilon_4\varepsilon_8}{2} -) \tag{19}$$

By substituting Equations (18) and (19) into Equation (1), we get:

$$\begin{aligned} T_{m_k} &= C_{Y_{N_k}}(1 - \varepsilon_0 + \frac{\varepsilon_7}{2} + \varepsilon_0^2 - \frac{\varepsilon_7^2}{8} - \frac{\varepsilon_0\varepsilon_7}{2} -) \\ &+ a_k \left\{ \begin{array}{l} C_{X_{N_k}}(\varepsilon_1 - \varepsilon_2 + \frac{1}{2}(\varepsilon_{10} - \varepsilon_{13}) + \varepsilon_2^2 - \varepsilon_1^2 + \frac{1}{8}(\varepsilon_{13}^2 - \varepsilon_{10}^2) + \frac{1}{2}(\varepsilon_1\varepsilon_{13} - \varepsilon_2\varepsilon_{10})) \\ +(\frac{1}{2}(\varepsilon_9 - \varepsilon_8) + (\varepsilon_4 - \varepsilon_5) + \frac{1}{8}(\varepsilon_8^2 - \varepsilon_9^2) + (\varepsilon_5^2 - \varepsilon_4^2) + \frac{1}{2}(\varepsilon_4\varepsilon_8 - \varepsilon_5\varepsilon_9)) \end{array} \right\} \end{aligned} \tag{20}$$

Using Equations (20) and (3), we obtain:

$$\begin{aligned} T_m &= \sum_{k=1}^{G}\Omega_k^* C_{Y_{N_k}} + \sum_{k=1}^{G}\Omega_k^* C_{Y_{N_k}}(\frac{\varepsilon_7}{2} - \varepsilon_0 + \varepsilon_0^2 - \frac{\varepsilon_7^2}{8} - \frac{\varepsilon_0\varepsilon_7}{2} -) \\ &+ \sum_{k=1}^{G} a_k \Omega_k^* \left\{ \begin{array}{l} C_{X_{N_k}}(\varepsilon_1 - \varepsilon_2 + \frac{1}{2}(\varepsilon_{10} - \varepsilon_{13}) + \varepsilon_2^2 - \varepsilon_1^2 + \frac{1}{8}(\varepsilon_{13}^2 - \varepsilon_{10}^2) + \frac{1}{2}(\varepsilon_1\varepsilon_{13} - \varepsilon_2\varepsilon_{10})) \\ +(\frac{1}{2}(\varepsilon_9 - \varepsilon_8) + (\varepsilon_4 - \varepsilon_5) + \frac{1}{8}(\varepsilon_8^2 - \varepsilon_9^2) + (\varepsilon_5^2 - \varepsilon_4^2) + \frac{1}{2}(\varepsilon_4\varepsilon_8 - \varepsilon_5\varepsilon_9)) \end{array} \right\} \end{aligned} \tag{21}$$

Considering the close proximity of the calibrated weight Ω_k^* to the strata weight W_k, we may reasonably assume that:

$$\sum_{k=1}^{G}\Omega_k^* C_{Y_{N_k}} \approx \sum_{k=1}^{G} W_k C_{Y_{N_k}} = C_Y$$

To transform Equation (21) accordingly,

$$T_m = C_Y + \sum_{k=1}^{G} \Omega_k^* C_{Y_{N_k}} \left(\frac{\varepsilon_7}{2} - \varepsilon_0 + \varepsilon_0^2 - \frac{\varepsilon_7^2}{8} - \frac{\varepsilon_0\varepsilon_7}{2} - \right)$$

$$+ \sum_{k=1}^{G} a_k \Omega_k^* \left\{ \begin{array}{l} C_{X_{N_k}} (\varepsilon_1 - \varepsilon_2 + \frac{1}{2}(\varepsilon_{10} - \varepsilon_{13}) + \varepsilon_2^2 - \varepsilon_1^2 + \frac{1}{8}(\varepsilon_{13}^2 - \varepsilon_{10}^2) + \frac{1}{2}(\varepsilon_1\varepsilon_{13} - \varepsilon_2\varepsilon_{10})) \\ +(\frac{1}{2}(\varepsilon_9 - \varepsilon_8) + (\varepsilon_4 - \varepsilon_5) + \frac{1}{8}(\varepsilon_8^2 - \varepsilon_9^2) + (\varepsilon_5^2 - \varepsilon_4^2) + \frac{1}{2}(\varepsilon_4\varepsilon_8 - \varepsilon_5\varepsilon_9)) \end{array} \right\} \quad (22)$$

After performing some rearrangement and taking expectations on both sides of Equation (22), the Bias and MSE of T_m, considering the first-order approximation, may be expressed as follows:

$$Bias(T_m) = \sum_{k=1}^{G} \Omega_k^* f_{1k} C_{X_{N_k}} \left(C_{Y_{N_k}}^2 - \frac{(\lambda_{400k} - 1)}{8} - \frac{C_{Y_{N_k}} \lambda_{300k}}{2}\right)$$

$$+ \sum_{k=1}^{G} a_k \Omega_k^* \left\{ \begin{array}{l} C_{X_{N_k}} (f_{1k} - f_{6k})(C_{X_{N_k}}^2 - \frac{(\lambda_{040k} - 1)}{8} - \frac{C_{X_{N_k}} \lambda_{030k}}{2}) \\ +(f_{4k} - f_{5k})(C_{Z_{N_k}}^2 - \frac{(\lambda_{004k} - 1)}{8} - \frac{C_{Z_{N_k}} \lambda_{003k}}{2}) \end{array} \right\} \quad (23)$$

and

$$MSE(T_m) = \sum_{k=1}^{G} \Omega_k^{*2} f_{1k} C_{Y_{N_k}}^2 \left(C_{X_{N_k}}^2 + \frac{(\lambda_{400k} - 1)}{4} - C_{X_{N_k}} \lambda_{300k}\right)$$

$$+ \sum_{k=1}^{G} a_k^2 \Omega_k^{*2} \left\{ \begin{array}{l} C_{X_{N_k}}^2 (f_{1k} + f_{6k})(\frac{(\lambda_{040k} - 1)}{4} - \frac{f_{1k}(\lambda_{040k} - 1)}{2}) \\ +C_{X_{N_k}}^2 (f_{1k} - f_{6k})(C_{X_{N_k}} \lambda_{030k} - C_{X_{N_k}}^2) \\ +(f_{4k} + f_{5k})(C_{Z_{N_k}}^2 + \frac{(\lambda_{004k} - 1)}{4}) - \frac{(\lambda_{004k} - 1)}{2} \\ +(f_{4k} - f_{5k}) C_{Z_{N_k}} \lambda_{003k} - 2 f_{4k} C_{Z_{N_k}}^2) \\ \quad +2C_{X_{N_k}} (f_{4k} - f_{5k})(\frac{C_{X_{N_k}} \lambda_{012k}}{2} - C_{XZ_{N_k}} \\ \quad -\frac{(\lambda_{022k} - 1)}{4} + \frac{C_{Z_{N_k}} \lambda_{021k}}{2}) \end{array} \right\}$$

$$+ 2\sum_{k=1}^{G} a_k \Omega_k^{*2} C_{Y_{N_k}} \left\{ \begin{array}{l} (f_{1k} - f_{4k})(\frac{C_{Z_{N_k}} \lambda_{201k}}{2} + \frac{C_{Y_{N_k}} \lambda_{003k}}{2} - \frac{(\lambda_{202k} - 1)}{4}) \\ - f_{1k} C_{YZ_{N_k}} + f_{4k} C_{XY_{N_k}} \end{array} \right\} \quad (24)$$

By applying the same transformations to Equation (2), we obtain:

$$T_{u_k} = C_{Y_{N_k}} \left(1 - \varepsilon_3 + \frac{\varepsilon_{11}}{2} - \frac{\varepsilon_{11}^2}{8} + \varepsilon_3^2 - \frac{\varepsilon_3\varepsilon_{11}}{2} - \right) + b_k \left(\frac{\varepsilon_{12}}{2} - \varepsilon_6 - \frac{\varepsilon_{12}^2}{8} + \varepsilon_6^2 - \frac{\varepsilon_6\varepsilon_{12}}{2} - \right) \quad (25)$$

Using Equations (4) and (25), we obtain:

$$T_u = \sum_{k=1}^{G}\Omega_k^{**}C_{Y_{N_k}} + \sum_{k=1}^{G}\Omega_k^{**}C_{Y_{N_k}}\left(\frac{\varepsilon_{11}}{2}-\varepsilon_3+\varepsilon_3^2-\frac{\varepsilon_{11}^2}{8}-\frac{\varepsilon_3\varepsilon_{11}}{2}--\right) + \sum_{k=1}^{G}b_k\Omega_k^{**}\left(\frac{\varepsilon_{12}}{2}-\varepsilon_6+\varepsilon_6^2-\frac{\varepsilon_{12}^2}{8}-\frac{\varepsilon_6\varepsilon_{12}}{2}--\right) \tag{26}$$

Considering the close proximity of the calibrated weight Ω_k^{**} to the strata weight W_k, we may reasonably assume that:

$$\sum_{k=1}^{G}\Omega_k^{**}C_{Y_{N_k}} \approx \sum_{k=1}^{G}W_kC_{Y_{N_k}} = C_Y$$

To transform Equation (26) accordingly,

$$T_u = C_Y + \sum_{k=1}^{G}\Omega_k^{**}C_{Y_{N_k}}\left(\frac{\varepsilon_{11}}{2}-\varepsilon_3+\varepsilon_3^2-\frac{\varepsilon_{11}^2}{8}-\frac{\varepsilon_3\varepsilon_{11}}{2}--\right) + \sum_{k=1}^{G}b_k\Omega_k^{**}\left(\frac{\varepsilon_{12}}{2}-\varepsilon_6+\varepsilon_6^2-\frac{\varepsilon_{12}^2}{8}-\frac{\varepsilon_6\varepsilon_{12}}{2}--\right) \tag{27}$$

After performing some rearrangement and taking expectations on both sides of Equation (27), the Bias and MSE of T_u, considering the first-order approximation, may be expressed as follows:

$$Bias(T_u) = \sum_{k=1}^{G}\Omega_k^{**}f_{2k}C_{Y_{N_k}}\left(C_{Y_{N_k}}^2-\frac{(\lambda_{400k}-1)}{8}-\frac{C_{Y_{N_k}}\lambda_{300k}}{2}\right) + \sum_{k=1}^{G}b_k\Omega_k^{**}f_{3k}\left(\left(C_{Z_{N_k}}^2-\frac{(\lambda_{004k}-1)}{8}-\frac{C_{Z_{N_k}}\lambda_{003k}}{2}\right)\right) \tag{28}$$

and

$$\begin{aligned} MSE(T_u) &= \sum_{k=1}^{G}\Omega_k^{**2}f_{2k}C_{Y_{N_k}}^2\left(C_{Y_{N_k}}^2+\frac{(\lambda_{400k}-1)}{4}-C_{Y_{N_k}}\lambda_{300k}\right) \\ &+ \sum_{k=1}^{G}2b_k\Omega_k^{**2}f_{3k}C_{Y_{N_k}}\left(\frac{(\lambda_{202k}-1)}{4}-\frac{C_{Z_{N_k}}\lambda_{201k}}{2}-\frac{C_{Y_{N_k}}\lambda_{102k}}{2}+C_{YZ_{N_k}}\right) \\ &+ \sum_{k=1}^{G}b_k^2\Omega_k^{**2}f_{3k}\left(C_{Z_{N_k}}^2+\frac{(\lambda_{004k}-1)}{4}-C_{Z_{N_k}}\lambda_{003k}\right) \end{aligned} \tag{29}$$

Thus, we may obtain the bias and MSE of estimator T by combining the biases and MSEs of two non-overlapping samples T_u and T_m, as shown below:

$$Bias(T) = \phi Bias(T_u) + (1-\phi)Bias(T_m) \tag{30}$$

and

$$MSE(T) = \phi^2 MSE(T_u) + (1-\phi)^2 MSE(T_m) \tag{31}$$

Where Equations (23), (28), (24), and (29) provide the expressions for $Bias(T_m)$, $Bias(T_u)$, $MSE(T_m)$, and $MSE(T_u)$, respectively. As T_u and T_m are based on non-overlapping samples of sizes *u* and *m*, respectively, the covariance term may be ignored as it is of order N^{-1}, and $c(T_u, T_m) = 0$.

5.2 The estimator's minimum mean squared error (MSE)

MSEs of T_m and T_u depend on a_k and b_k. To find the best a_k and b_k, we minimize the MSEs in Equations (24) and (29). We get optimal a_k and b_k values.

$$a_{k_{opt}} = \frac{C_{Y_{N_k}}\left\{(f_{1k}-f_{4k})\left(\frac{\lambda_{202k}-1}{4}-\frac{C_{Z_{N_k}}\lambda_{201k}}{2}-\frac{C_{Y_{N_k}}\lambda_{003k}}{2}\right)+f_{1k}C_{YZ_{N_k}}-f_{4k}C_{XY_{N_k}}\right\}}{\left\{\begin{array}{l} C^2_{X_{N_k}}\left(\frac{(f_{1k}+f_{6k})(\lambda_{040k}-1)}{4}-\frac{f_{1k}(\lambda_{040k}-1)}{2}+(f_{1k}-f_{6k})(C_{X_{N_k}}\lambda_{030k}-C^2_{X_{N_k}})\right) \\ +2C_{X_{N_k}}(f_{1k}-f_{5k})\left(\frac{C_{X_{N_k}}\lambda_{012k}}{2}-C_{XZ_{N_k}}-\frac{\lambda_{022k}-1}{4}+\frac{C_{Z_{N_k}}\lambda_{021k}}{2}\right)-\frac{f_{4k}(\lambda_{004k}-1)}{2} \\ +(f_{4k}+f_{5k})\left(C^2_{Z_{N_k}}+\frac{\lambda_{004k}-1}{4}\right)+(f_{4k}-f_{5k})C_{Z_{N_k}}\lambda_{003k}-2f_{4k}C^2_{Z_{N_k}} \end{array}\right\}} \tag{32}$$

and

$$b_{k_{opt}} = \frac{C_{Y_{N_k}}\left\{\frac{C_{Z_{N_k}}\lambda_{201k}}{2}-\frac{\lambda_{202k}-1}{4}-C_{YZ_{N_k}}+\frac{C_{Y_{N_k}}\lambda_{102k}}{2}\right\}}{C^2_{Z_{N_k}}+\frac{\lambda_{004k}-1}{4}-C_{Z_{N_k}}\lambda_{003k}} \tag{33}$$

Using $a_{k_{opt}}$ from Equation (32) in Equation (24) and $b_{k_{opt}}$ from Equation (33) in Equation (29) gives the minimum MSE of T_m and T_u.

$$MinMSE(T_m) = \sum_{k=1}^{G}\Omega_k^{*2} f_{1k}C^2_{Y_{N_k}}\left(C^2_{X_{N_k}}+\frac{\lambda_{400k}-1}{4}-C_{X_{N_k}}\lambda_{210k}\right) - \sum_{k=1}^{G}\Omega_k^{*2}\frac{C^2_{Y_{N_k}}\left\{\begin{array}{l}(f_{1k}-f_{4k})\left(\frac{\lambda_{202k}-1}{4}-\frac{C_{Z_{N_k}}\lambda_{201k}}{2}-\frac{C_{Y_{N_k}}\lambda_{003k}}{2}\right) \\ +f_{1k}C_{YZ_{N_k}}-f_{4k}C_{XY_{N_k}}\end{array}\right\}^2}{\left\{\begin{array}{l} C^2_{X_{N_k}}\left(\frac{(f_{1k}+f_{6k})(\lambda_{040k}-1)}{4}-\frac{f_{1k}(\lambda_{040k}-1)}{2}\right. \\ \left.+(f_{1k}-f_{6k})(C_{X_{N_k}}\lambda_{030k}-C^2_{X_{N_k}})\right) \\ +2C_{X_{N_k}}(f_{1k}-f_{5k})\left(\frac{C_{X_{N_k}}\lambda_{012k}}{2}-C_{XZ_{N_k}}-\frac{\lambda_{022k}-1}{4}+\frac{C_{Z_{N_k}}\lambda_{021k}}{2}\right) \\ -\frac{f_{4k}(\lambda_{004k}-1)}{2}+(f_{4k}+f_{5k})\left(C^2_{Z_{N_k}}+\frac{\lambda_{004k}-1}{4}\right) \\ +(f_{4k}-f_{5k})C_{Z_{N_k}}\lambda_{003k}-2f_{4k}C^2_{Z_{N_k}} \end{array}\right\}} \tag{34}$$

and

$$MinMSE(T_u)=\sum_{k=1}^{G}\Omega_k^{**2}C_{Y_{N_k}}^2\left\{\begin{array}{l} f_{2k}\left(C_{Y_{N_k}}^2+\dfrac{\lambda_{004k}-1}{4}-C_{Y_{N_k}}\lambda_{300k}\right) \\ -\dfrac{f_{3k}\left\{\dfrac{C_{Z_{N_k}}\lambda_{201k}}{2}-\dfrac{\lambda_{202k}-1}{4}-C_{YZ_{N_k}}+\dfrac{C_{Y_{N_k}}\lambda_{102k}}{2}\right\}^2}{C_{Z_{N_k}}^2+\dfrac{\lambda_{004k}-1}{4}-C_{Z_{N_k}}\lambda_{003k}}\end{array}\right\} \tag{35}$$

MSE of T depends on ϕ. To find the best ϕ, we minimize the MSE. The optimal ϕ is

$$\phi_{opt}=\frac{MinMSE(T_m)}{MinMSE(T_m)+MinMSE(T_u)} \tag{36}$$

We may use this ϕ value to get the best MSE of T, which is

$$MSE(T)_{opt}=\frac{MinMSE(T_m)MinMSE(T_u)}{MinMSE(T_m)+MinMSE(T_u)} \tag{37}$$

Equations (34) and (35) give the expressions for $MinMSE(T_m)$ and $MinMSE(T_u)$, respectively.

6 Effects of measurement error

Let x_k, y_k, and z_k represent the true values of variables X, Y, and Z, respectively. The corresponding observed values are denoted as x_k^*, y_k^*, and z_k^*. We define the measurement errors for X, Y, and Z as $u_k=x_k-x_k^*$, $v_k=y_k-y_k^*$, and $w_k=z_k-z_k^*$. These errors are assumed to follow normal distributions: $u_k\in U\sim N(0,\ S_u^2)$, $v_k\in V\sim N(0,\ S_v^2)$, and $w_k\in W\sim N(0,\ S_w^2)$. The variables X, Y, Z, U, V, and W are pairwise uncorrelated. We consider two cases in this analysis.

Case I: Assuming Z as a variable without measurement errors
In this case, we assume that the auxiliary variable Z is free from measurement errors, meaning that $w_k=0$. Consequently, the joint moment about the mean is expressed as:

$$\varpi_{pqsk}=\frac{1}{N_k-1}\sum_{k=1}^{G}u_k^p v_k^q (z_k-\overline{Z})^s \tag{38}$$

Considering the influence of measurement errors U and V, we derive the expression for the minimum mean square error (MSE) of the proposed estimators ϕ as follows:

$$MSE(T)_{opt_I}=\frac{MinMSE(T_m)_I\,MinMSE(T_u)_I}{MinMSE(T_m)_I+MinMSE(T_u)_I} \tag{39}$$

where

$$MinMSE(T_m)_I=\sum_{k=1}^{G}\Omega_k^{*2}\left[A_{o.m(I)}-\frac{A_{1.m(I)}^2}{A_{2.m(I)}}\right]$$

$$A_{o.m(I)}=f_{1k}\left(C_{Y_{N_k}}^2+\frac{S_{U_{N_k}}^2}{\overline{Y}_{N_k}^2}\right)\left(C_{Y_{N_k}}^2+\frac{S_{U_{N_k}}^2}{\overline{Y}_{N_k}^2}+\frac{\varpi_{400k}-1}{4}-\sqrt{C_{Y_{N_k}}^2+\frac{S_{U_{N_k}}^2}{\overline{Y}_{N_k}^2}}\varpi_{300k}\right)$$

$$A_{1.m(I)} = \left[\begin{array}{l} \sqrt{C^2_{Y_{N_k}} + \dfrac{S^2_{U_{N_k}}}{\bar{Y}^2_{N_k}}}\,(f_{1k} - f_{4k}) \left(\dfrac{\varpi_{202k} - 1}{4} - \dfrac{C_{Z_{N_k}} \varpi_{201k}}{2} - \sqrt{C^2_{Y_{N_k}} + \dfrac{S^2_{U_{N_k}}}{\bar{Y}^2_{N_k}}}\, \dfrac{\varpi_{003k}}{2} \right) \\ \qquad - f_{1k} C_{YZ_{N_k}} - f_{4k} C_{XY_{N_k}} \end{array} \right]$$

$$A_{2.m(I)} = \left[\begin{array}{l} 2\sqrt{C^2_{X_{N_k}} + \dfrac{S^2_{V_{N_k}}}{\bar{X}^2_{N_k}}}\,(f_{1k} - f_{5k}) \left(\sqrt{C^2_{X_{N_k}} + \dfrac{S^2_{V_{N_k}}}{\bar{X}^2_{N_k}}}\, \dfrac{\varpi_{012k}}{2} - C_{XZ_{N_k}} - \dfrac{\varpi_{022k} - 1}{4} + \dfrac{C_{Z_{N_k}} \varpi_{021k}}{2} \right) \\ +(f_{4k} + f_{5k})\left(C^2_{Z_{N_k}} + \dfrac{\varpi_{004k} - 1}{4}\right) + (f_{4k} - f_{5k}) C_{Z_{N_k}} \varpi_{003k} - f_{4k}\left(\dfrac{\varpi_{004k} - 1}{2} - 2C^2_{Z_{N_k}}\right) \\ +\sqrt{C^2_{X_{N_k}} + \dfrac{S^2_{V_{N_k}}}{\bar{X}^2_{N_k}}}\,(f_{1k} - f_{6k}) \left(\sqrt{C^2_{X_{N_k}} + \dfrac{S^2_{V_{N_k}}}{\bar{X}^2_{N_k}}}\, \varpi_{030k} - C^2_{X_{N_k}} - \dfrac{S^2_{V_{N_k}}}{\bar{X}^2_{N_k}} - \dfrac{f_{1k}(\varpi_{040k} - 1)}{2} \right) \\ +\left(C^2_{X_{N_k}} + \dfrac{S^2_{V_{N_k}}}{\bar{X}^2_{N_k}} \right) \left((f_{1k} + f_{6k}) \dfrac{\varpi_{040k} - 1}{4} \right) \end{array} \right]$$

$$MinMSE(T_u)_I = \sum_{k=1}^{G} \Omega_k^{**2} \left(C^2_{Y_{N_k}} + \frac{S^2_{U_{N_k}}}{\bar{Y}^2_{N_k}} \right)^2 \left\{ \begin{array}{l} f_{2k}\left(C^2_{Y_{N_k}} + \dfrac{S^2_{U_{N_k}}}{\bar{Y}^2_{N_k}} + \dfrac{\varpi_{400k} - 1}{4} - \sqrt{C^2_{Y_{N_k}} + \dfrac{S^2_{U_{N_k}}}{\bar{Y}^2_{N_k}}}\, \varpi_{300k}\right) \\ - \dfrac{ f_{3k} \left\{ \begin{array}{l} \dfrac{C_{Z_{N_k}} \varpi_{201k}}{2} - \dfrac{\varpi_{202k} - 1}{4} - C_{YZ_{N_k}} \\ + \sqrt{C^2_{Y_{N_k}} + \dfrac{S^2_{U_{N_k}}}{\bar{Y}^2_{N_k}}}\, \dfrac{\varpi_{102k}}{2} \end{array} \right\}^2 }{ C^2_{Z_{N_k}} + \dfrac{\varpi_{004k} - 1}{4} - C_{Z_{N_k}} \varpi_{003k} } \end{array} \right\}$$

Case II: Assuming Z as a variable with measurement errors

In this case, we consider that the auxiliary variable Z is characterized by measurement errors. Thus, the joint moment about the mean is given by:

$$\kappa_{pqsk} = \frac{1}{N_k - 1} \sum_{k=1}^{G} u_k^p v_k^q w_k^s \tag{40}$$

Taking into account the effects of measurement errors U, V, and W, we obtain the expression for the minimum mean square error (MSE) of the proposed estimators ϕ as follows:

$$MSE(T)_{opt_{II}} = \frac{MinMSE(T_m)_{II}\, MinMSE(T_u)_{II}}{MinMSE(T_m)_{II} + MinMSE(T_u)_{II}} \tag{41}$$

where

$$MinMSE(T_m)_{II} = \sum_{k=1}^{G} \Omega_k^{*2} \left[A_{o.m(II)} - \frac{A^2_{1.m(II)}}{A_{2.m(II)}} \right]$$

$$A_{o.m(II)} = f_{1k}\left(C^2_{Y_{N_k}} + \frac{S^2_{U_{N_k}}}{\overline{Y}^2_{N_k}}\right)\left(C^2_{Y_{N_k}} + \frac{S^2_{U_{N_k}}}{\overline{Y}^2_{N_k}} + \frac{\kappa_{400k}-1}{4} - \sqrt{C^2_{Y_{N_k}} + \frac{S^2_{U_{N_k}}}{\overline{Y}^2_{N_k}}}\,\kappa_{300k}\right)$$

$$A_{1.m(II)} = \left[\begin{array}{l} \sqrt{C^2_{Y_{N_k}} + \frac{S^2_{U_{N_k}}}{\overline{Y}^2_{N_k}}}\,(f_{1k} - f_{4k})\left(\frac{\kappa_{202k}-1}{4} - \sqrt{C^2_{Z_{N_k}} + \frac{S^2_{W_{N_k}}}{\overline{Z}^2_{N_k}}}\,\frac{\kappa_{201k}}{2} - \sqrt{C^2_{Y_{N_k}} + \frac{S^2_{U_{N_k}}}{\overline{Y}^2_{N_k}}}\,\frac{\kappa_{003k}}{2}\right) \\ \qquad - f_{1k}C_{YZ_{N_k}} - f_{4k}C_{XY_{N_k}} \end{array}\right]$$

$$A_{2.m(II)} = \left[\begin{array}{l} 2\sqrt{C^2_{X_{N_k}} + \frac{S^2_{V_{N_k}}}{\overline{X}^2_{N_k}}}\,(f_{1k} - f_{5k})\left(\sqrt{C^2_{X_{N_k}} + \frac{S^2_{V_{N_k}}}{\overline{X}^2_{N_k}}}\,\frac{\kappa_{012k}}{2} - C_{XZ_{N_k}} - \frac{\kappa_{022k}-1}{4}\right) \\ +\left(C^2_{X_{N_k}} + \frac{S^2_{V_{N_k}}}{\overline{X}^2_{N_k}}\right)(f_{1k} - f_{6k})\left(\sqrt{C^2_{X_{N_k}} + \frac{S^2_{V_{N_k}}}{\overline{X}^2_{N_k}}}\,\kappa_{030k} - C^2_{X_{N_k}} - \frac{S^2_{V_{N_k}}}{\overline{X}^2_{N_k}} - \frac{f_{1k}(\kappa_{040k}-1)}{2}\right) \\ +(f_{4k} + f_{5k})(C^2_{Z_{N_k}} + \frac{S^2_{W_{N_k}}}{\overline{Z}^2_{N_k}} + \frac{\kappa_{004k}-1}{4}) + (f_{4k} - f_{5k})\sqrt{C^2_{Z_{N_k}} + \frac{S^2_{W_{N_k}}}{\overline{Z}^2_{N_k}}}\,\kappa_{003k} \\ +\left(C^2_{X_{N_k}} + \frac{S^2_{V_{N_k}}}{\overline{X}^2_{N_k}}\right)\left((f_{1k} + f_{6k})\frac{\kappa_{040k}-1}{4}\right) - f_{4k}\left(\frac{\kappa_{004k}-1}{2} - 2C^2_{Z_{N_k}} - \frac{2S^2_{W_{N_k}}}{\overline{Z}^2_{N_k}}\right) \\ +2\sqrt{C^2_{X_{N_k}} + \frac{S^2_{V_{N_k}}}{\overline{X}^2_{N_k}}}\,(f_{1k} - f_{5k})\left(\sqrt{C^2_{Z_{N_k}} + \frac{S^2_{W_{N_k}}}{\overline{Z}^2_{N_k}}}\,\frac{\kappa_{021k}}{2}\right) \end{array}\right]$$

$$MinMSE(T_u)_{II} = \sum_{k=1}^{G}\Omega_k^{**2}\left(C^2_{Y_{N_k}} + \frac{S^2_{U_{N_k}}}{\overline{Y}^2_{N_k}}\right)^2 \left\{\begin{array}{l} f_{2k}\left(C^2_{Y_{N_k}} + \frac{S^2_{U_{N_k}}}{\overline{Y}^2_{N_k}} + \frac{\kappa_{400k}-1}{4} - \sqrt{C^2_{Y_{N_k}} + \frac{S^2_{U_{N_k}}}{\overline{Y}^2_{N_k}}}\,\kappa_{300k}\right) \\ - \dfrac{f_{3k}\left\{\begin{array}{l}\sqrt{C^2_{Z_{N_k}} + \frac{S^2_{W_{N_k}}}{\overline{Z}^2_{N_k}}}\,\frac{\kappa_{201k}}{2} - \frac{\kappa_{202k}-1}{4} \\ -C_{YZ_{N_k}} + \sqrt{C^2_{Y_{N_k}} + \frac{S^2_{U_{N_k}}}{\overline{Y}^2_{N_k}}}\,\frac{\kappa_{102k}}{2}\end{array}\right\}^2}{C^2_{Z_{N_k}} + \frac{S^2_{W_{N_k}}}{\overline{Z}^2_{N_k}} + \frac{\kappa_{004k}-1}{4} - \sqrt{C^2_{Z_{N_k}} + \frac{S^2_{W_{N_k}}}{\overline{Z}^2_{N_k}}}\,\kappa_{003k}} \end{array}\right\}$$

7 Empirical study

Before employing an estimator in practical situations, it is crucial to assess its performance based on its inherent properties. In light of this, an empirical analysis has been carried out in this section utilizing both real and simulated data to evaluate the suggested estimator. To accomplish this, we will conduct a comparison between the suggested estimator T and an alternative estimator τ , which is also designed to handle random non-response and measurement errors and is defined in the same manner. The purpose of this comparison is to evaluate how well the suggested estimator T performs under conditions of random non-response and measurement errors.

$$\tau = \psi\tau_u + (1-\psi)\tau_m$$

The values of $\tau_u = \sum_{k=1}^{G}\left(\frac{N_k}{N}\right)c_{y_{u_k - r_3 k}}$, $\tau_m = \sum_{k=1}^{G}\left(\frac{N_k}{N}\right)c_{y_{n_k - r_1 k}}$ and $\psi\ (0 \le \psi \le 1)$ are unknown, and the constant ψ needs to be determined by minimizing the Mean Squared Error (MSE) of estimator τ .

The minimum Mean Squared Error (MSE) of estimator τ , up to the first order of approximations, may be expressed as:

$$MSE(\tau) = \frac{MSE(\tau_u)MSE(\tau_m)}{MSE(\tau_u) + MSE(\tau_m)} \tag{42}$$

where

$$MSE(\tau_u) = \sum_{k=1}^{G}\left(\frac{N_k}{N}\right)C_{Y_{N_k}}\left(\frac{1}{u_k q_3 + 2p_3} - \frac{1}{N_k}\right)\left(C_{Y_{N_k}}^2 + \frac{\lambda_{400k} - 1}{4} - C_{Y_{N_k}}\lambda_{300k}\right) \text{ and}$$

$$MSE(\tau_m) = \sum_{k=1}^{G}\left(\frac{N_k}{N}\right)C_{Y_{N_k}}\left(\frac{1}{n_k q_1 + 2p_1} - \frac{1}{N_k}\right)\left(C_{Y_{N_k}}^2 + \frac{\lambda_{400k} - 1}{4} - C_{Y_{N_k}}\lambda_{300k}\right)$$

The proposed estimator T may be evaluated in terms of its Absolute Relative Bias (ARB) and Percentage Relative Efficiency (PRE) with respect to the estimator τ respectively. This may be calculated using the following formula:

$$ARB = \frac{|E(T - C_Y)|}{C_Y}, \tag{43}$$

$$PRE = \frac{MSE(\tau)}{MSE(T)_{opt}} \times 100 \tag{44}$$

where $MSE(T)_{opt}$ and $MSE(\tau)$ are defined in Equations (37) and (42), respectively.

The following Q_k values have been considered:

Case I: $Q_k = 1.0$ This case assigns equal importance to each stratum, treating them equally in the calibration process.

Case II: $Q_k = \frac{1}{W_k}$ Here, the importance of each stratum is inversely proportional to its initial weight W_k , implying that strata with lower initial weights are given higher importance during calibration.

Case III: $Q_k = \frac{1}{\overline{Z}_k}$ In this case, the importance of each stratum is inversely proportional to its mean auxiliary variable value, suggesting that strata with higher mean auxiliary variable values are given more weight in the calibration process.

7.1 Simulation study

Using the statistical computing software *R*, we simulated data based on our theoretical findings. To generate data for both the study and auxiliary variables, following a normal distribution with specific parameters and correlation coefficients, we utilized the *mvrnorm* function from the *MASS* package. The population parameters for the generated data are presented in Table 1. We conducted simulations to analyze the impact of the controlling parameter Q_k under conditions of random non-response and measurement error.

7.2 Study based on real data

In this section, we examine the application of a proposed class of estimators to address a real-world problem concerning prostate cancer. Prostate cancer is a major cause of cancer-related deaths among men (Source: Ellsworth Pamela 2020). Developing effective screening tools for early detection of prostate cancer is crucial. The prostate-specific antigen (PSA) test, along with other diagnostic tests, is commonly employed for this purpose. Elevated PSA levels in the blood often indicate prostate cancer in patients. Individuals diagnosed with prostate cancer are advised to undergo regular PSA tests to monitor disease progression or detect its recurrence.

However, the PSA test suffers from a significant limitation - a high false-positive rate. Overdiagnosis may result in unnecessary invasive medical procedures, such as biopsies. One way to overcome this challenge is to establish age-specific cutoff values for PSA levels, which necessitates a study of age-related variations in PSA levels.

To explore this further, we consider data from a study involving men who were about to undergo a radical prostatectomy. The study analyzed the correlation between PSA levels and several clinical measures (Source: Stamey et al. 1989). The data utilized to illustrate the application of the proposed class of estimators is obtained from Section 4 and may be found under the name ‘prostate’ in the ‘faraway’ package in R (Faraway 2016).

The study incorporates the following variables:
Y: Logarithm of PSA
X: Logarithm of cancer volume
Z: Logarithm of prostate weight

To examine age-specific cutoff values for PSA range, the individuals participating in the study are divided into strata based on their ages:
Stratum 1: Individuals below the age of 60
Stratum 2: Individuals aged between 60 and 69
Stratum 3: Individuals aged 70 and above

Although complete data are available for this study, it is important to note that this is not always the case. In scenarios where some data on the study variable are missing, the goal is to estimate the coefficient of variance as accurately as possible. Table 2 provides statistical information about the population. Further tables present the PRE under non-response and measurement error for both cases, as well as the PRE in the absence of non-response and in the presence of measurement errors.

7.3 Discussion and results

Based on the simulation study and the study based on real data presented above:

1. Tables 3 and 4 show that calibrated strata weights are similar to original weights. Instead of calibration weights for both simulated and real data. Instead of calibration weights, we may also use $\frac{n_k}{\sum_{k=1}^{G} n_k}$ because it is an estimator of $\frac{N_k}{N}$. The estimator effectively reduces the negative impact of non-responses, and the bias is negligible for all Q_k choices.

2. Tables 5-7 reveal that the proposed estimator significantly outperforms the standard estimator, exhibiting higher Percent Relative Efficiency (PRE). This indicates that the proposed method is more effective in the presence of random non-response and assuming Z as a variable without measurement errors, as evidenced by its lower Mean Squared Error (MSE). The results validate the robustness of the proposed estimator, making it a more reliable choice when addressing non-response issues (simulated data).

3. Tables 8-10 demonstrate that the proposed estimator consistently outperforms the standard estimator, as indicated by higher PRE values. This performance advantage is particularly evident in scenarios with random non-response and assuming Z as a variable without measurement errors, underscoring the estimator's robustness and effectiveness in practical applications (real data).

4. The analysis in Tables 11-13 reveals that the Absolute Relative Bias (ARB) of the proposed estimator due to random non-response is minimal, approximately around 10^{-4}, suggesting that the proposed method effectively reduces non-response bias and enhances performance compared to the standard estimator. The results for the simulated data show that as p_3 increases, the bias rises while the PRE decreases. For fixed p_1 and p_3, increasing p_2 results in a constant bias but a decrease in PRE. Conversely, increasing p_1 while keeping p_2 and p_3 constant maintains a constant bias but increases the PRE. Additionally, as the non-response rate of p_1 increases, the PRE also increases, further highlighting the estimator's ability to maintain accuracy even in the presence of random non-response and measurement errors. These findings underscore significant patterns in the simulated data.

5. Tables 14-16 demonstrate that the Absolute Relative Bias (ARB) of the proposed estimator remains minimal (approximately 10^{-2} even in the presence of random non-response, highlighting its effectiveness in mitigating non-response bias and maintaining accuracy and reliability in real-world conditions. The findings further indicate that increasing p_3 results in a rise in bias and a reduction in PRE, while keeping p_1 and p_3 constant and increasing p_2 leads to a constant bias but reduced PRE. Conversely, increasing p_1 with p_2 and p_3 fixed results in a constant bias and an increase in PRE. Additionally, as the non-response rate of p_1 grows, the PRE also increases. These observations underscore the estimator's robustness and the key trends identified in the real data.

6. Tables 17 and 18 show that the proposed estimator has negligible ARB and a higher PRE than the standard estimator, even in the absence of non-response and measurement errors. This indicates that the proposed method is more effective, even without non-response and measurement errors for simulated as well as real data.

7. We may observe in a simulation study that as we increase the value of the correlation coefficient, the value of the PRE will increase and the bias will decrease. Conversely, as we decrease the value of the correlation coefficient, the PRE will decrease and the bias will increase.

The study assessed the feasibility of employing calibrated weights to combat non-response in stratified successive sampling, aiming to improve the accuracy of coefficient of variation estimation at the population level. Our evaluation included both simulation studies, as detailed in section *7.1*, and analyses of real data, outlined in section *7.2*. The developed estimator exhibited significant efficacy in mitigating random non-response in stratified two-occasion successive sampling, particularly when auxiliary information on positively correlated variables was available. This method not only reduced bias but also enhanced accuracy in CV estimation, demonstrating commendable performance across various scenarios of non-response rates, correlation coefficients, and error structures. Survey statisticians are encouraged to consider adopting this approach in similar contexts.

8 Conclusions

In summary, our research highlights the critical empirical outcomes of incorporating calibrated weights to address non-response in the context of stratified successive sampling. The proposed estimator offers substantial advantages, particularly when auxiliary information is available, enhancing its applicability in practical survey scenarios.

Validation through comprehensive numerical and simulation studies demonstrates the remarkable reduction in bias and improved precision in coefficient of variation (CV) estimation across diverse simulated and real-life scenarios. Notably, our approach proves highly feasible, showcasing its effectiveness in refining CV estimation at the population level, as evidenced by superior performance observed in both simulated data and its application to estimate the coefficient of variation for the logarithm of PSA in prostate cancer datasets. Survey statisticians are urged to integrate this estimator into their methodologies, given its established track record in improving precision and accuracy, especially in the face of challenges posed by non-response and measurement errors. Furthermore, the demonstrated effectiveness of our approach in mitigating random non-response in the context of stratified two-occasion successive sampling emphasizes its relevance and practical value for survey statisticians encountering similar challenges in real-world scenarios.

Acknowledgement

We are thankful to the IIT (ISM) Dhanbad for providing the financial and infrastructural support to accomplish the present work. Additionally, We would like to thank the esteemed editor and respected reviewers for their helpful and valuable suggestions, which greatly contributed to improving the manuscript's current state.

Disclosure statement

The authors have no conflicts of interest to disclose.

References

[1] Tripathi, T. P., Singh H. P., Upadhyaya, L. N. "A general method of estimation and its application to the estimation of co-efficient of variation", *Statistics in Transition*, **5**(6), pp. 887–909, (2002).

[2] Archana, V., Rao, K. A. "Improved estimators of coefficient of variation in a finite population", *Statistics in Transition new series*, **2**(12), pp. 357–380, (2011).

[3] Sisodia, B. V. S., Dwivedi, V. K. "Modified ratio estimator using coefficient of variation of auxiliary variable", *Journal-Indian Society of Agricultural Statistics*, (1981).

[4] Das, A. K., Tripathi, T. P. "A class of estimators for co-efficient of variation using knowledge on coefficient of variation of an auxiliary character", in *annual conference of Ind. Soc. Agricultural Statistics. Held at New Delhi, India*, (1981).

[5] Das, A. K., Tripathi, T. P. "Use of auxiliary information in estimating the coefficient of variation", *Alig. J. of. Statist*, **12**, pp. 51–58, (1992).

[6] Patel, P. A., Rina, S. "A Monte Carlo comparison of some suggested estimators of Co-efficient of variation in finite population", *Journal of Statistics sciences*, **1**(2), pp. 137–147, (2009).

[7] Singh, R., Mishra, M., Singh B. P., et al., "Improved estimators for population coefficient of variation using auxiliary variable", *Journal of Statistics and Management Systems*, **21**(7), pp. 1335–1355, (2018). DOI: *https://doi.org/10.1080/09720510.2018.1503405*

[8] Muneer, S., Khalil, A., Shabbir, J., et al., "Efficient estimation of population median using supplementary variable", *Scientia Iranica*, **29**(1), pp. 265–274, (2022). DOI: *10.24200/SCI.2020.52871.2924*

[9] Yunusa, M. A., *et al.*, "Logarithmic ratio-type estimator of population coefficient of variation", *Asian Journal of Probability and Statistics*, **14**(2), pp. 13–22, (2021). DOI: 10.9734/AJPAS/2021/v14i230323

[10] Audu, A., *et al.*, "Difference-cum-ratio estimators for estimating finite population coefficient of variation in simple random sampling", *Asian Journal of Probability and Statistics*, **13**(3), pp. 13–29, (2021). DOI: 10.9734/AJPAS/2021/v13i330308

[11] Shahzad, U., Ahmad, I., Hanif, M., et al., "Estimation of coefficient of variation using linear moments and calibration approach for nonsensitive and sensitive variables", *Concurr Comput*, **34**(18), pp. e7006, (2022). DOI: *https://doi.org/10.1002/cpe.7006*

[12] Shahzad, U., Ahmad, I., García-Luengo, A. V., et al., "Estimation of coefficient of variation using calibrated estimators in double stratified random sampling", *Mathematics*, **11**(1), pp. 252, (2023).
DOI: *https://doi.org/10.3390/math11010252*

[13] Yadav, S. K., Misra, S., Gupta, S. "Improved family of estimators of population coefficient of variation under simple random sampling", *Communications in Statistics-Theory and Methods*, **53**(2), pp. 727–747, (2024).
DOI: *https://doi.org/10.1080/03610926.2022.2091784*

[14] Rajyaguru, A., Gupta, P. C. "On the estimation of the coefficient of variation from finite population-II", *Model Assisted Statistics and Applications*, **1**(1), pp. 57–66, (2006).
DOI: 10.3233/MAS-2006-1110

[15] Jessen, R. J. *Statistical investigation of a sample survey for obtaining farm facts*, Iowa State University, (1943).

[16] Yates, F., et al., "Sampling methods for censuses and surveys", *Sampling methods for censuses and surveys.*, (1949).

[17] Eckler, A. R. "Rotation sampling", *The annals of mathematical statistics*, **26**(4), pp. 664–685, (1955).

[18] Sen, A. R. "346. Note: Theory and application of sampling on repeated occasions with several auxiliary variables", *Biometrics*, pp. 381–385, (1973).

[19] Feng, S., Zou, G. "Sample rotation method with auxiliary variable", *Communications in Statistics-Theory and Methods*, **26**(6), pp. 1497–1509, (1997).
DOI: *https://doi.org/10.1080/03610929708831996*

[20] Singh, G. N., Homa, F. "Effective rotation patterns in successive sampling over two occasions", *J Stat Theory Pract*, **7**, pp. 146–155, (2013).
DOI: https://doi.org/10.1080/15598608.2012.755484

[21] Naz, F., Abid, M., Nawaz, T., et al. "Enhancing efficiency of ratio-type estimators of population variance by a combination of information on robust location measures", *Scientia Iranica*, **27**(4), pp. 2040–2056, (2020).
DOI: *10.24200/SCI.2019.5633.1385*

[22] Younis F., Shabbir, J. "Estimation of general parameters under stratified adaptive cluster sampling based on dual use of auxiliary information", *Scientia Iranica*, **28**(3), pp. 1780–1801, (2021).
DOI: 10.24200/sci.2019.52515.2753

[23] Abid, M., Sherwani, R. A., Tahir, M., et al., "An improved and robust class of variance estimator", *Scientia Iranica*, **28**(6), pp. 3589–3601, (2021).
DOI: *10.24200/SCI.2020.52986.2986*

[24] Irfan, M., Javed, M., Bhatti, S. H., "Difference-type-exponential estimators based on dual auxiliary information under simple random sampling", *Scientia Iranica*, **29**(1), pp. 343–354, (2022).
DOI: *10.24200/SCI.2020.53592.3318*

[25] Bhushan S., Pandey, S. "An effective class of estimators for population mean estimation in successive sampling using simulation approach", *J Stat Comput Simul*, pp. 1–32, (2023).
DOI: *https://doi.org/10.1080/00949655.2023.2282741*

[26] Sen, A. R. "Some theory of sampling on successive occasions", *Australian Journal of Statistics*, pp. 105–110, (1973).
DOI: *https://doi.org/10.1111/j.1467-842X.1973.tb00014.x*

[27] Allen, J., Singh, H. P., Smarandache, F. "A family of estimators of population mean using multiauxiliary information in presence of measurement errors", *Int J Soc Econ*, **30**(7), pp. 837–848, (2003).
DOI: *https://doi.org/10.1108/03068290310478775*

[28] Kumar, S., Bhougal, S., Nataraja, N. S., et al., "Estimation of population mean in the presence of non-response and measurement error", *Rev Colomb Estad*, **38**(1), pp. 145–161, (2015).
DOI: http://dx.doi.org/10.15446/rce.v38n1.48807

[29] Ahmed, S., Shabbir, J. "A novel basis function approach to finite population parameter estimation", *Scientia Iranica*, (2021).
DOI: *10.24200/SCI.2021.56353.4682*

[30] Audu, A., Singh, R., Khare, S., et al. "Almost unbiased variance estimators under the simultaneous influence of non-response and measurement errors", *J Stat Theory Pract*, **16**(2), pp. 15, (2022).
DOI: https://doi.org/10.1007/s42519-021-00226-8

[31] Shahzad, U., Ahmad, I., Almanjahie, I. M., et al. "Mean estimation using robust quantile regression with two auxiliary variables", *Scientia Iranica*, (2022).
DOI: 10.24200/sci.2022.57170.5098
[32] Deming, W. E., Stephan, F. F. "On a least squares adjustment of a sampled frequency table when the expected marginal totals are known", *The Annals of Mathematical Statistics*, **11**(4), pp. 427–444, (1940).
DOI: https://www.jstor.org/stable/2235722
[33] Deville, J. C., Särndal, C. E. "Calibration estimators in survey sampling", *J Am Stat Assoc*, **87**(418), pp. 376–382, (1992).
DOI: https://doi.org/10.1080/01621459.1992.10475217
[34] Farrell, P. J., Singh, S. "*Model-Assisted higher-order calibration of estimators of variance*", *Aust N Z J Stat*, **47**(3), pp. 375–383, (2005).
DOI: *https://doi.org/10.1111/j.1467-842X.2005.00402.x*
[35] Särndal, C. E. "The calibration approach in survey theory and practice", *Surv Methodol*, 33(2), pp. 99–119, (2007).
[36] Kim, J. M., Sungur, E. A., Heo, T. Y. "Calibration approach estimators in stratified sampling", *Stat Probab Lett*, **77**(1), pp. 99–103, (2007).
[37] Kim, J. K., Park, M. "Calibration estimation in survey sampling", *International Statistical Review*, **78**(1), pp. 21–39, (2010).
DOI: https://doi.org/10.1111/j.1751-5823.2010.00099.x
[38] Sud, U. C., Chandra, H., Gupta, V. K. "Calibration approach-based regression-type estimator for inverse relationship between study and auxiliary variable", *J Stat Theory Pract*, **8**(4), pp. 707–721, (2014).
DOI: https://doi.org/10.1080/15598608.2013.832643
[39] Singh, G. K., Rao, D. K., Khan, M. G. M. "Calibration estimator of population mean in stratified random sampling", in *Asia-Pacific World Congress on Computer Science and Engineering*, pp. 1–5 (2014).
DOI: 10.1109/APWCCSE.2014.7053875
[40] Koyuncu, N., Kadilar, C. "Calibration weighting in stratified random sampling", *Communications in Statistics-Simulation and Computation*, 45(7), pp. 2267–2275, (2016).
DOI: *https://doi.org/10.1080/03610918.2014.901354*
[41] Nidhi, Sisodia, B. V. S., Singh, S., et al., "Calibration approach estimation of the mean in stratified sampling and stratified double sampling", *Communications in Statistics-Theory and Methods*, **46**(10), pp. 4932–4942, (2017).
DOI: *https://doi.org/10.1080/03610926.2015.1091083*
[42] Özgül, N. "New calibration estimator in stratified sampling", *J Stat Comput Simul*, **88**(13), pp. 2561–2572, (2018).
DOI: *https://doi.org/10.1080/00949655.2018.1478417*
[43] Shahzad, U., Ahmad, I., Alshahrani, F., et al., "Calibration-Based Mean Estimators under Stratified Median Ranked Set Sampling", *Mathematics*, **11**(8), pp. 1825, (2023).
DOI: https://doi.org/10.3390/math11081825
[44] Shahzad, U., Ahmad, I., Almanjahie, I. M., et al., "L-Moments and calibration-based variance estimators under double stratified random sampling scheme: Application of Covid-19 pandemic", *Scientia Iranica*, **30**(2), pp. 814–821, (2023). DOI: 10.24200/SCI.2021.56853.4942
[45] Pandey, M. K., Singh, G. N., Zaman, T., et al. "A general class of improved population variance estimators under non-sampling errors using calibrated weights in stratified sampling", *Sci Rep*, **14**(1), pp. 2948, (2024).
DOI: *https://doi.org/10.1038/s41598-023-47234-1*
[46] Pandey, M.K., Singh, G.N., Zaman, T., et al. "Improved estimation of population variance in stratified successive sampling using calibrated weights under non-response", *Heliyon*, ***10***(6), pp. e27738, (2024).
DOI: *10.1016/j.heliyon.2024.e27738*
[47] Pandey, M. K., Singh, G. N., Zaman, T. "Estimation of Population Mean Using Some Improved Imputation Methods for Missing Data in Sample Surveys", *Communications in Statistics-Theory and Methods*, pp. 1-15, (2024).
DOI: *https://doi.org/10.1080/03610926.2024.2369314*
[48] Clement, E. P. "A new ratio estimator of mean in survey sampling by calibration estimation", *Elixir International Journal: Statistics*, **106**, pp. 46461–46465, (2017).

[49] Singh S., Joarder, A. H. "Estimation of finite population variance using random non-response in survey sampling", *Metrika*, **47**, pp. 241–249, (1998).
DOI: https://doi.org/10.1007/BF02742876
[50] Bahl, S., Tuteja, R. K. "Ratio and product type exponential estimators", *Journal of information and optimization sciences*, **12**(1), pp. 159–164, (1991).
DOI: *https://doi.org/10.1080/02522667.1991.10699058*

List of table captions

Stratum	N_k	n_k	r_{1k}	m_k	r_{2k}	u_k	ρ_{XY}	ρ_{YZ}	ρ_{ZX}
Strata 1	10000	3000	500	2000	400	1000	0.9	0.9	0.9
Strata 2	15000	3000	500	2000	400	1000	0.8	0.8	0.8

Stratum	N_k	n_k	r_{1k}	m_k	r_{2k}	u_k	ρ_{XY}	ρ_{YZ}	ρ_{ZX}
Strata 3	5000	2000	400	1200	300	800	0.8	0.8	0.8
Strata 4	8000	800	50	500	50	300	0.7	0.7	0.7

Table 1: The statistical parameters corresponding to the simulated data.

Stratum	N_k	n_k	r_{1k}	m_k	r_{2k}	u_k	ρ_{XY}	ρ_{YZ}	ρ_{ZX}
Strata 1	20	11	2	6	1	5	0.85	0.47	0.48
Strata 2	60	45	12	25	5	20	0.64	0.35	0.15
Strata 3	17	10	2	5	2	4	0.69	0.61	0.31

Table 2: The statistical parameters corresponding to the real data.

Case	Stratum	W_k	Ω_k^*	Ω_k^{**}
I	1	0.2631579	0.1389900	0.10775276
	2	0.3947368	0.5512098	0.75216865
	3	0.1315789	0.1816670	0.05387638
	4	0.2105263	0.1281332	0.08620221
II	1	0.2631579	0.1195100	0.14401396
	2	0.3947368	0.5193507	0.75216865
	3	0.1315789	0.2121448	0.01243501
	4	0.2105263	0.1489945	0.09138238
III	1	0.2631579	0.1389244	0.10758672
	2	0.3947368	0.5511026	0.75216865
	3	0.1315789	0.1817695	0.05387086
	4	0.2105263	0.1282034	0.08637377

Table 3: Calibrated strata weights for simulated data.

Case	Stratum	W_k	Ω_k^*	Ω_k^{**}
I	1	0.2061856	0.26446760	0.3126980
	2	0.6185567	0.37455840	0.4659243
	3	0.1752577	0.36097410	0.2213777
II	1	0.2061856	0.34605130	0.1694016
	2	0.6185567	0.34189170	0.6420778
	3	0.1752577	0.31205710	0.1885206
III	1	0.2061856	0.28186790	0.3497671
	2	0.6185567	0.50458530	0.5310217
	3	0.1752577	0.21354680	0.1192113

Table 4: Calibrated strata weights for real data.

		p_3			
p_1	p_2	0.05	0.10	0.15	0.20
0.05	0.05	153.1851	141.6388	133.2162	126.8583
0.05	0.10	152.9561	141.4076	132.9828	126.6227
0.05	0.15	152.7422	141.1916	132.7647	126.4026
0.05	0.20	152.5418	140.9894	132.5606	126.1965
0.10	0.05	155.4660	143.2761	134.3790	127.6585
0.10	0.10	155.2714	143.0795	134.1804	127.4579
0.10	0.15	155.0926	142.8990	133.9980	127.2737
0.10	0.20	154.9279	142.7326	133.8299	127.1039
0.15	0.05	158.0574	145.1688	135.7562	128.6415

0.15	0.10	157.8970	145.0068	135.5924	128.4759
0.15	0.15	157.7532	144.8614	135.4455	128.3274
0.15	0.20	157.6234	144.7303	135.3129	128.1935
0.20	0.05	161.0059	147.3563	137.3818	129.8371
0.20	0.10	160.8801	147.2291	137.2531	129.7070
0.20	0.15	160.7711	147.1189	137.1416	129.5942
0.20	0.20	160.6757	147.0224	137.0441	129.4955

Table 5: For case I: $Q_1 = 1.0$, $Q_2 = 1.0$, $Q_3 = 1.0$, and $Q_4 = 1.0$, the PREs of T with respect to τ for simulated data, assuming Z as a variable without measurement errors.

		p_3			
p_1	p_2	0.05	0.10	0.15	0.20
0.05	0.05	153.8399	141.9968	133.4138	126.9665
0.05	0.10	153.6085	141.7633	133.1780	126.7285
0.05	0.15	153.3923	141.5450	132.9577	126.5061
0.05	0.20	153.1897	141.3405	132.7513	126.2977
0.10	0.05	156.1804	143.6734	134.6042	127.7872
0.10	0.10	155.9825	143.4735	134.4023	127.5833
0.10	0.15	155.8005	143.2897	134.2166	127.3958
0.10	0.20	155.6326	143.1202	134.0454	127.2229
0.15	0.05	158.8355	145.6074	136.0098	128.7908
0.15	0.10	158.6707	145.4409	135.8415	128.6208
0.15	0.15	158.5226	145.2913	135.6903	128.4680
0.15	0.20	158.3888	145.1561	135.5537	128.3299
0.20	0.05	161.8519	147.8381	137.6642	130.0066
0.20	0.10	161.7203	147.7050	137.5296	129.8705
0.20	0.15	161.6060	147.5894	137.4127	129.7523
0.20	0.20	161.5057	147.4880	137.3102	129.6486

Table 6: For case II: $Q_1 = 3.80$, $Q_2 = 2.53$, $Q_3 = 7.60$, and $Q_4 = 4.75$, the PREs of T with respect to τ for simulated data, assuming Z as a variable without measurement errors.

		p_3			
p_1	p_2	0.05	0.10	0.15	0.20
0.05	0.05	153.1566	141.6193	133.2024	126.8484
0.05	0.10	152.9276	141.3881	132.9690	126.6127
0.05	0.15	152.7136	141.1721	132.7510	126.3926
0.05	0.20	152.5133	140.9699	132.5468	126.1865
0.10	0.05	155.4360	143.2557	134.3645	127.6480
0.10	0.10	155.2414	143.0590	134.1659	127.4474
0.10	0.15	155.0626	142.8785	133.9835	127.2632
0.10	0.20	154.8978	142.7121	133.8154	127.0934
0.15	0.05	158.0257	145.1472	135.7409	128.6304
0.15	0.10	157.8653	144.9852	135.5771	128.4648
0.15	0.15	157.7215	144.8398	135.4302	128.3163
0.15	0.20	157.5917	144.7086	135.2976	128.1823
0.20	0.05	160.9724	147.3335	137.3656	129.8254
0.20	0.10	160.8466	147.2062	137.2369	129.6952
0.20	0.15	160.7375	147.0960	137.1254	129.5824

0.20	0.20	160.6421	146.9995	137.0278	129.4837

Table 7: For case III: $Q_1 = 0.3324468$, $Q_2 = 0.3315064$, $Q_3 = 0.3321156$, and $Q_4 = 0.3316337$, the PREs of T with respect to τ for simulated data, assuming Z as a variable without measurement errors.

		p_3			
p_1	p_2	0.05	0.10	0.15	0.20
0.05	0.05	124.2784	122.2180	120.3446	118.6380
0.05	0.10	124.1082	122.0459	120.1706	118.4620
0.05	0.15	123.9496	121.8855	120.0084	118.2979
0.05	0.20	123.8015	121.7357	119.8568	118.1447
0.10	0.05	125.4450	123.2509	121.2538	119.4326
0.10	0.10	125.2943	123.0983	121.0994	119.2763
0.10	0.15	125.1547	122.9570	120.9564	119.1316
0.10	0.20	125.0250	122.8257	120.8235	118.9971
0.15	0.05	126.7816	124.4531	122.3316	120.3951
0.15	0.10	126.6434	124.3130	122.1898	120.2514
0.15	0.15	126.5161	124.1841	122.0592	120.1191
0.15	0.20	126.3986	124.0650	121.9386	119.9969
0.20	0.05	128.2934	125.8302	123.5840	121.5317
0.20	0.10	128.1594	125.6943	123.4462	121.3920
0.20	0.15	128.0369	125.5701	123.3203	121.2643
0.20	0.20	127.9244	125.4560	123.2047	121.1471

Table 8: For case I: $Q_1 = 1.0$, $Q_2 = 1.0$, and $Q_3 = 1.0$, the PREs of T with respect to τ for real data, assuming Z as a variable without measurement errors.

		p_3			
p_1	p_2	0.05	0.10	0.15	0.20
0.05	0.05	133.5361	130.4289	127.6419	125.1352
0.05	0.10	133.2576	130.1458	127.3543	124.8429
0.05	0.15	132.9993	129.8832	127.0874	124.5717
0.05	0.20	132.7590	129.6390	126.8392	124.3194
0.10	0.05	134.7752	131.4922	128.5437	125.8880
0.10	0.10	134.5302	131.2430	128.2902	125.6301
0.10	0.15	134.3043	131.0131	128.0564	125.3922
0.10	0.20	134.0952	130.8005	127.8400	125.1720
0.15	0.05	136.1819	132.7252	129.6168	126.8133
0.15	0.10	135.9623	132.5015	129.3890	126.5813
0.15	0.15	135.7611	132.2966	129.1803	126.3688
0.15	0.20	135.5761	132.1082	128.9884	126.1733
0.20	0.05	137.7503	134.1221	130.8557	127.9062
0.20	0.10	137.5475	133.9153	130.6450	127.6913
0.20	0.15	137.3632	133.7274	130.4533	127.4959
0.20	0.20	137.1948	133.5557	130.2783	127.3174

Table 9: For case II: $Q_1 = 4.85$, $Q_2 = 1.62$, and $Q_3 = 5.71$, the PREs of T with respect to τ for real data, assuming Z as a variable without measurement errors.

		p_3			
p_1	p_2	0.05	0.10	0.15	0.20
0.05	0.05	110.5138	110.1408	109.7955	109.4762
0.05	0.10	110.1174	109.7405	109.3914	109.0682
0.05	0.15	109.7497	109.3692	109.0165	108.6897
0.05	0.20	109.4076	109.0237	108.6676	108.3376
0.10	0.05	110.1999	109.7928	109.4152	109.0651
0.10	0.10	109.8419	109.4311	109.0496	108.6958
0.10	0.15	109.5114	109.0971	108.7122	108.3548
0.10	0.20	109.2053	108.7877	108.3995	108.0389
0.15	0.05	110.0815	109.6407	109.2309	108.8503
0.15	0.10	109.7501	109.3055	108.8920	108.5076
0.15	0.15	109.4458	108.9978	108.5807	108.1929
0.15	0.20	109.1654	108.7142	108.2939	107.9028
0.20	0.05	110.1562	109.6823	109.2410	108.8305
0.20	0.10	109.8384	109.3606	108.9154	108.5009
0.20	0.15	109.5486	109.0673	108.6185	108.2004
0.20	0.20	109.2831	108.7985	108.3464	107.9250

Table 10: For case III: $Q_1 = 0.296$, 0.274, and $Q_3 = 0.261$, the PREs of T with respect to τ for real data, assuming Z as a variable without measurement errors.

		$p_3 = 0.05$		$p_3 = 0.10$		$p_3 = 0.15$		$p_3 = 0.20$	
p_1	p_2	ARB	PRE	ARB	PRE	ARB	PRE	ARB	PRE
0.05	0.05	0.0003814	132.1	0.0004163	126.9	0.0004486	122.6	0.0004787	119.1
0.05	0.10	0.0003814	131.9	0.0004163	126.6	0.0004486	122.4	0.0004787	118.9
0.05	0.15	0.0003814	131.7	0.0004163	126.4	0.0004486	122.1	0.0004787	118.6
0.05	0.20	0.0003814	131.5	0.0004163	126.2	0.0004486	121.9	0.0004787	118.4
0.10	0.05	0.0003814	133.2	0.0004163	127.7	0.0004486	123.1	0.0004787	119.5
0.10	0.10	0.0003814	133.0	0.0004163	127.5	0.0004486	122.9	0.0004787	119.2
0.10	0.15	0.0003814	132.8	0.0004163	127.3	0.0004486	122.8	0.0004787	119.1
0.10	0.20	0.0003814	132.7	0.0004163	127.1	0.0004486	122.6	0.0004787	118.9
0.15	0.05	0.0003814	134.5	0.0004163	128.6	0.0004486	123.9	0.0004787	119.9
0.15	0.10	0.0003814	134.3	0.0004163	128.5	0.0004486	123.7	0.0004787	119.8
0.15	0.15	0.0003814	134.2	0.0004163	128.3	0.0004486	123.5	0.0004787	119.6
0.15	0.20	0.0003814	134.1	0.0004163	128.2	0.0004486	123.4	0.0004787	119.5
0.20	0.05	0.0003814	136.0	0.0004163	129.8	0.0004486	124.7	0.0004787	120.6
0.20	0.10	0.0003814	135.9	0.0004163	129.7	0.0004486	124.6	0.0004787	120.5
0.20	0.15	0.0003814	135.8	0.0004163	129.5	0.0004486	124.5	0.0004787	120.3
0.20	0.20	0.0003814	135.7	0.0004163	129.4	0.0004486	124.4	0.0004787	120.2

Table 11: For case I: $Q_1 = 1.0$, $Q_2 = 1.0$, $Q_3 = 1.0$, and $Q_4 = 1.0$, the PREs of T with respect to τ for simulated data, assuming Z as a variable with measurement errors and ARB.

		$p_3 = 0.05$		$p_3 = 0.10$		$p_3 = 0.15$		$p_3 = 0.20$	
p_1	p_2	ARB	PRE	ARB	PRE	ARB	PRE	ARB	PRE

0.05	0.05	0.0003799	132.3	0.0004148	127.0	0.0004470	122.7	0.0004768	119.1
0.05	0.10	0.0003799	132.1	0.0004148	126.7	0.0004470	122.4	0.0004768	118.9
0.05	0.15	0.0003799	131.8	0.0004148	126.5	0.0004470	122.2	0.0004768	118.7
0.05	0.20	0.0003799	131.6	0.0004148	126.3	0.0004470	122.0	0.0004768	118.5
0.10	0.05	0.0003799	133.4	0.0004148	127.8	0.0004470	123.2	0.0004768	119.5
0.10	0.10	0.0003799	133.2	0.0004148	127.6	0.0004470	123.0	0.0004768	119.3
0.10	0.15	0.0003799	133.0	0.0004148	127.4	0.0004470	122.8	0.0004768	119.1
0.10	0.20	0.0003799	132.8	0.0004148	127.2	0.0004470	122.7	0.0004768	118.9
0.15	0.05	0.0003799	134.7	0.0004148	128.8	0.0004470	124.0	0.0004768	120.0
0.15	0.10	0.0003799	134.6	0.0004148	128.6	0.0004470	123.8	0.0004768	119.8
0.15	0.15	0.0003799	134.4	0.0004148	128.4	0.0004470	123.6	0.0004768	119.7
0.15	0.20	0.0003799	134.3	0.0004148	128.3	0.0004470	123.5	0.0004768	119.5
0.20	0.05	0.0003799	136.3	0.0004148	130.0	0.0004470	124.9	0.0004768	120.7
0.20	0.10	0.0003799	136.1	0.0004148	129.8	0.0004470	124.7	0.0004768	120.5
0.20	0.15	0.0003799	136.0	0.0004148	129.7	0.0004470	124.6	0.0004768	120.4
0.20	0.20	0.0003799	135.9	0.0004148	129.6	0.0004470	124.5	0.0004768	120.3

Table 12: For case II: $Q_1 = 3.80$, $Q_2 = 2.53$, $Q_3 = 7.60$, and $Q_4 = 4.75$, the PREs of T with respect to τ for simulated data, assuming Z as a variable with measurement errors and ARB.

		$p_3 = 0.05$		$p_3 = 0.10$		$p_3 = 0.15$		$p_3 = 0.20$	
p_1	p_2	ARB	PRE	ARB	PRE	ARB	PRE	ARB	PRE
0.05	0.05	0.0003814	132.1	0.0004163	126.8	0.0004486	122.6	0.0004787	119.1
0.05	0.10	0.0003814	131.9	0.0004163	126.6	0.0004486	122.3	0.0004787	118.9
0.05	0.15	0.0003814	131.7	0.0004163	126.4	0.0004486	122.1	0.0004787	118.6
0.05	0.20	0.0003814	131.5	0.0004163	126.2	0.0004486	121.9	0.0004787	118.4
0.10	0.05	0.0003814	133.2	0.0004163	127.6	0.0004486	123.1	0.0004787	119.4
0.10	0.10	0.0003814	133.0	0.0004163	127.4	0.0004486	122.9	0.0004787	119.2
0.10	0.15	0.0003814	132.8	0.0004163	127.3	0.0004486	122.8	0.0004787	119.1
0.10	0.20	0.0003814	132.6	0.0004163	127.1	0.0004486	122.6	0.0004787	118.9
0.15	0.05	0.0003814	134.5	0.0004163	128.6	0.0004486	123.8	0.0004787	119.9
0.15	0.10	0.0003814	134.3	0.0004163	128.4	0.0004486	123.7	0.0004787	119.8
0.15	0.15	0.0003814	134.2	0.0004163	128.3	0.0004486	123.5	0.0004787	119.6
0.15	0.20	0.0003814	134.0	0.0004163	128.2	0.0004486	123.4	0.0004787	119.5
0.20	0.05	0.0003814	136.0	0.0004163	129.8	0.0004486	124.7	0.0004787	120.6
0.20	0.10	0.0003814	135.9	0.0004163	129.6	0.0004486	124.6	0.0004787	120.4
0.20	0.15	0.0003814	135.8	0.0004163	129.5	0.0004486	124.5	0.0004787	120.3
0.20	0.20	0.0003814	135.7	0.0004163	129.4	0.0004486	124.4	0.0004787	120.2

Table 13: For case III: $Q_1 = 0.3324468$, $Q_2 = 0.3315064$, $Q_3 = 0.3321156$, and $Q_4 = 0.3316337$, the PREs of T with respect to τ for simulated data, assuming Z as a variable with measurement errors and ARB.

		$p_3 = 0.05$		$p_3 = 0.10$		$p_3 = 0.15$		$p_3 = 0.20$	
p_1	p_2	ARB	PRE	ARB	PRE	ARB	PRE	ARB	PRE
0.05	0.05	0.1244457	122.5	0.1244677	120.6	0.1244414	119.0	0.1243613	117.4
0.05	0.10	0.1244457	122.3	0.1244677	120.5	0.1244414	118.8	0.1243613	117.2
0.05	0.15	0.1244457	122.2	0.1244677	120.3	0.1244414	118.6	0.1243613	117.1
0.05	0.20	0.1244457	122.0	0.1244677	120.2	0.1244414	118.5	0.1243613	116.9
0.10	0.05	0.1244457	123.5	0.1244677	121.6	0.1244414	119.8	0.1243613	118.1
0.10	0.10	0.1244457	123.4	0.1244677	121.4	0.1244414	119.6	0.1243613	118.0
0.10	0.15	0.1244457	123.2	0.1244677	121.3	0.1244414	119.5	0.1243613	117.8

0.10	0.20	0.1244457	123.1	0.1244677	121.1	0.1244414	119.3	0.1243613	117.7
0.15	0.05	0.1244457	124.7	0.1244677	122.7	0.1244414	120.8	0.1243613	119.0
0.15	0.10	0.1244457	124.6	0.1244677	122.5	0.1244414	120.6	0.1243613	118.9
0.15	0.15	0.1244457	124.5	0.1244677	122.4	0.1244414	120.5	0.1243613	118.7
0.15	0.20	0.1244457	124.4	0.1244677	122.3	0.1244414	120.4	0.1243613	118.6
0.20	0.05	0.1244457	126.1	0.1244677	123.9	0.1244414	121.9	0.1243613	120.1
0.20	0.10	0.1244457	126.0	0.1244677	123.8	0.1244414	121.8	0.1243613	119.9
0.20	0.15	0.1244457	125.9	0.1244677	123.7	0.1244414	121.6	0.1243613	119.8
0.20	0.20	0.1244457	125.8	0.1244677	123.6	0.1244414	121.5	0.1243613	119.7

Table 14: For case I: $Q_1 = 1.0$, $Q_2 = 1.0$, and $Q_3 = 1.0$, the PREs of T with respect to τ for real data, assuming Z as a variable with measurement errors and ARB.

		$p_3 = 0.05$		$p_3 = 0.10$		$p_3 = 0.15$		$p_3 = 0.20$	
p_1	p_2	ARB	PRE	ARB	PRE	ARB	PRE	ARB	PRE
0.05	0.05	0.1256676	131.1	0.1259374	128.4	0.1261225	125.8	0.1262125	123.6
0.05	0.10	0.1256676	130.9	0.1259374	128.1	0.1261225	125.6	0.1262125	123.3
0.05	0.15	0.1256676	130.6	0.1259374	127.8	0.1261225	125.3	0.1262125	123.0
0.05	0.20	0.1256676	130.4	0.1259374	127.6	0.1261225	125.1	0.1262125	122.8
0.10	0.05	0.1256676	132.2	0.1259374	129.3	0.1261225	126.6	0.1262125	124.2
0.10	0.10	0.1256676	132.0	0.1259374	129.1	0.1261225	126.4	0.1262125	124.0
0.10	0.15	0.1256676	131.8	0.1259374	128.8	0.1261225	126.2	0.1262125	123.8
0.10	0.20	0.1256676	131.6	0.1259374	128.6	0.1261225	125.9	0.1262125	123.5
0.15	0.05	0.1256676	133.5	0.1259374	130.4	0.1261225	127.6	0.1262125	125.1
0.15	0.10	0.1256676	133.3	0.1259374	130.2	0.1261225	127.4	0.1262125	124.8
0.15	0.15	0.1256676	133.1	0.1259374	130.0	0.1261225	127.2	0.1262125	124.6
0.15	0.20	0.1256676	132.9	0.1259374	129.8	0.1261225	127.0	0.1262125	124.4
0.20	0.05	0.1256676	134.9	0.1259374	131.7	0.1261225	128.7	0.1262125	126.1
0.20	0.10	0.1256676	134.7	0.1259374	131.5	0.1261225	128.5	0.1262125	125.9
0.20	0.15	0.1256676	134.5	0.1259374	131.3	0.1261225	128.3	0.1262125	125.7
0.20	0.20	0.1256676	134.4	0.1259374	131.1	0.1261225	128.2	0.1262125	125.5

Table 15: For case II: $Q_1 = 4.85$, $Q_2 = 1.62$, and $Q_3 = 5.71$, the PREs of T with respect to τ for real data, assuming Z as a variable with measurement errors and ARB.

		$p_3 = 0.05$		$p_3 = 0.10$		$p_3 = 0.15$		$p_3 = 0.20$	
p_1	p_2	ARB	PRE	ARB	PRE	ARB	PRE	ARB	PRE
0.05	0.05	0.100534	110.1	0.1005059	109.8	0.1004329	109.5	0.1003104	109.2
0.05	0.10	0.100534	109.7	0.1005059	109.4	0.1004329	109.1	0.1003104	108.8
0.05	0.15	0.100534	109.4	0.1005059	109.0	0.1004329	108.7	0.1003104	108.4
0.05	0.20	0.100534	109.0	0.1005059	108.7	0.1004329	108.4	0.1003104	108.1
0.10	0.05	0.100534	109.8	0.1005059	109.4	0.1004329	109.1	0.1003104	108.8
0.10	0.10	0.100534	109.4	0.1005059	109.1	0.1004329	108.7	0.1003104	108.4
0.10	0.15	0.100534	109.1	0.1005059	108.7	0.1004329	108.4	0.1003104	108.1
0.10	0.20	0.100534	108.8	0.1005059	108.4	0.1004329	108.1	0.1003104	107.7
0.15	0.05	0.100534	109.7	0.1005059	109.3	0.1004329	108.9	0.1003104	108.5
0.15	0.10	0.100534	109.3	0.1005059	108.9	0.1004329	108.5	0.1003104	108.2
0.15	0.15	0.100534	109.0	0.1005059	108.6	0.1004329	108.2	0.1003104	107.9
0.15	0.20	0.100534	108.7	0.1005059	108.3	0.1004329	107.9	0.1003104	107.6
0.20	0.05	0.100534	109.7	0.1005059	109.3	0.1004329	108.9	0.1003104	108.5
0.20	0.10	0.100534	109.4	0.1005059	109.0	0.1004329	108.5	0.1003104	108.2
0.20	0.15	0.100534	109.1	0.1005059	108.7	0.1004329	108.3	0.1003104	107.9

0.20	0.20	0.100534	108.8	0.1005059	108.4	0.1004329	108.0	0.1003104	107.6

Table 16: For case III: $Q_1 = 0.296$, $Q_2 = 0.274$, and $Q_3 = 0.261$, the PREs of T with respect to τ for real data, assuming Z as a variable with measurement errors and ARB.

Stratum	ARB	PRE
Case I	0.00002797887	226.1336
Case II	0.00003177613	230.9878
Case III	0.00002800443	226.0122

Table 17: In the absence of non-response and measurement errors, ARB and PRE are observed from simulated data when $p_1 = p_2 = p_3 = 0$.

Stratum	ARB	PRE
Case I	0.05789603	125.3921
Case II	0.05533376	140.9515
Case III	0.03252067	119.7129

Table 18: In the absence of non-response and measurement errors, ARB and PRE are observed from real data when $p_1 = p_2 = p_3 = 0$.

Author's Biography

M. K. Pandey is a research scholar at the Indian Institute of Technology (Indian School of Mines), Dhanbad-826004, Jharkhand, India. His research, conducted under the guidance of Prof. G. N. Singh, mainly focuses on sample surveys.

G. N. Singh is a Higher Academic Grade Professor of Statistics at the Indian Institute of Technology (Indian School of Mines), Dhanbad-826004, Jharkhand, India. He received his PhD in Statistics from Banaras Hindu University, Varanasi, India. His research area is mainly in sample surveys. He has published more than 240 research papers in international and national journals and conferences. He has guided 24 Ph.D. and many M.Sc. and M.Tech students for their thesis work under his supervision. His work has appeared in journals such as Biometrical Journal, Journal of Applied Statistics, Statistics, Journal of Statistical Computation and Simulation, Communication in Statistics: Simulation and Computation, Communication in Statistics: Theory and Methods, Mathematical Population Studies, RevStat-Statistical Journal, and Korean Journal of Statistical Society, among others.

A. Bandyopadhyay was an Assistant Professor in the Department of Mathematics at Asansol Engineering College, Asansol, West Bengal, India. He is currently working as an Associate Professor in the Department of Mathematics, Basic Science, and Humanities at Dr. B. C. Roy Engineering College, Durgapur, West Bengal, India (Durgapur-713206). He has published more than 60 research papers in international and national journals and conferences. His research area is sample surveys, with publications in journals such as the Journal of Statistical Computation and Simulation, Communication in Statistics: Simulation and Computation, and Communication in Statistics: Theory and Methods.