



Research article

Symmetric hybrid transformations for robust mean estimation under stratified systematic sampling

Fatimah A. Almulhim¹, Umer Daraz^{2,*}, Hassan M. Aljohani³ and Javid Shabbir⁴

¹ Department of Mathematical Sciences, College of Science, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia

² Department of Management Sciences, College of Business Administration, Hunan University, Changsha 410082, China

³ Department of Mathematics and Statistics, College of Science, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia

⁴ Department of Statistics, University of Wah, Wah Cantt, Pakistan

* **Correspondence:** Email: umerdaraz099@gmail.com; Tel: +8613245090592.

Abstract: Challenges arise in stratified random sampling when the auxiliary variable shows nonlinear behavior, pronounced skewness, or heterogeneous patterns across strata, leading to the reduced efficiency of traditional linear estimators. To overcome these limitations, this study develops a new class of weighted symmetric hybrid estimators for the finite population mean under a stratified design. The proposed methodology is constructed by combining harmonic, geometric, and fractional power transformations of auxiliary information within a unified symmetric framework, allowing the estimator to adapt to complex structural relationships while maintaining a balance between extreme and moderate observations within and across strata. Analytical expressions for bias and mean squared error are obtained, and optimal weights are derived through mean squared error minimization. Comparative theoretical analysis demonstrates that the proposed estimators consistently outperform classical ratio, product, regression, and exponential-type estimators. Furthermore, extensive simulation experiments conducted under various population structures, correlation levels, and stratification schemes demonstrate substantial efficiency enhancements, particularly for skewed and nonlinear populations. The proposed class therefore offers a flexible, symmetric, and robust alternative for practical implementation in stratified random sampling surveys.

Keywords: stratified systematic sampling; symmetric hybrid estimators; finite population mean estimation; nonlinear auxiliary structure; bias and mean squared error; robust estimation; efficiency analysis

Mathematics Subject Classification: 62D05

Summary of notations

Notation	Description	Notation	Description
N	Population size	L	Number of strata
N_h	Stratum size	n	Total sample size
n_h	Sample size in stratum h	W_h	Stratum weight
T	Study variable	V	Auxiliary variable
$\bar{T}_h$	Stratum mean of T	$\bar{V}_h$	Stratum mean of V
$\bar{t}_h$	Sample mean of T in stratum h	$\bar{v}_h$	Sample mean of V in stratum h
$\bar{t}_{st, sy}$	Stratified systematic mean of T	$\bar{v}_{st, sy}$	Stratified systematic mean of V
S_{th}^2	Variance of T in stratum h	S_{vh}^2	Variance of V in stratum h
S_{tvh}	Covariance in stratum h	ρ_{tvh}	Correlation between T and V
C_{th}^2	Coefficient of variation of T	C_{vh}^2	Coefficient of variation of V
ρ_{th}	Intra-class correlation of T	ρ_{vh}	Intra-class correlation of V
θ_h	Sampling constant	D_{th}	Design effect for T
D_{vh}	Design effect for V	D_{tvh}	Ratio term
F_{tvh}	Standardized correlation term	ζ_{0h}	Relative error of $\bar{t}_h$
ζ_{1h}	Relative error of $\bar{v}_h$	$MSE(\cdot)$	Mean squared error
$Bias(\cdot)$	Bias of estimator	PRE	Percent relative efficiency

1. Introduction

Stratified random sampling represents a widely adopted probability sampling design in which the population is partitioned into internally homogeneous and mutually exclusive subgroups, referred to as strata, prior to sample selection. Independent random samples are then drawn from each stratum, allowing the design to explicitly account for population heterogeneity. This approach often yields improved precision over simple random sampling by reducing within-stratum variability and enabling flexible allocation strategies across strata. However, the effectiveness of stratified sampling is closely tied to the choice of an appropriate estimator, as differences in stratum characteristics and the nature of auxiliary information can substantially affect the estimation's performance. Consequently, careful estimator construction remains essential to fully exploit the advantages offered by stratified random sampling.

Systematic sampling emerged as a convenient alternative to simple random sampling for large and well-organized populations, particularly in situations where ease of implementation and an evenly

spread sample selection are important. The theoretical framework of this design was established by Cochran [1, 2], who provided a probabilistic justification for systematic sampling and discussed its usefulness in large-scale survey applications. Later, Kish [3] demonstrated that, unlike simple random sampling, the variability of the systematic sample mean is highly sensitive to the ordering of population units and the presence of underlying structural patterns. This characteristic sets systematic sampling apart from simple random sampling and underscores the importance of choosing an appropriate estimator under this design. Yates [4] established the foundations of systematic sampling by formalizing the selection of population units at regular intervals, providing an efficient alternative to simple random sampling. Additional theoretical developments and practical discussions on systematic sampling can be found in [5–7].

The auxiliary information has always been used in mean estimation to improve efficiency. Ratio and product estimators, initially developed under the assumption of simple random sampling by Hansen and Hurwitz [8], were subsequently generalized to systematic sampling by several researchers, such as [9–11]. Lahiri [12] examined the bias properties of systematic sampling and investigated the conditions under which systematic sampling can produce biased estimates. These estimators can be used to reduce variance when the auxiliary variable is strongly correlated with the study variable, but their performance deteriorates when these assumptions are violated. Later research showed that better performance can be achieved by regression-type estimators, which explicitly capture linear relationships, although at the expense of increased sensitivity to model misspecification. Generalized and regression-type estimators were also introduced by Kushwaha and Singh [13] to further enhance the methodology, and it is important to note that these estimator structures are adaptable. Further developments include nonlinear transformations such as difference-, product-, and ratio-type mean-based estimators, which have been proposed to enhance robustness and flexibility, including methods by Kushwaha et al. [14]. Several researchers have made important contributions to the development of efficient estimators under the systematic sampling scheme. For instance, Singh et al. [15] proposed modified ratio and product estimators to enhance efficiency when auxiliary information is available. In a later study, Singh and Jatwa [16] developed a broader and more efficient class of estimators for estimating the population mean in systematic sampling. Khan and Singh [17] introduced a chain ratio-type estimator to further refine population mean estimation within the systematic sampling framework. Other studies conducted by [18, 19] further suggest that weighted and transformed estimators can outperform traditional estimators under suitable conditions. Recent studies have focused on improving the efficiency of finite population mean estimation by developing advanced estimators under different sampling designs. In particular, Mishra et al. [20] proposed a ratio-cum-product exponential class of estimators for adaptive cluster sampling, Subzar et al. [21] developed generalized robust regression estimators for rare and clustered populations, and Nagy et al. [22] introduced memory-type estimators for systematic sampling in time-scaled surveys, demonstrating improved estimation efficiency through the effective use of auxiliary information. Improved ratio, product, and regression estimators using minimum and maximum values were developed [23] for finite population mean estimation. Under systematic sampling, a generalized semi-exponential estimator was proposed [24] to improve estimation efficiency. The use of maximum and minimum values was further explored [25] for finite population mean estimation under systematic sampling. Two auxiliary variables, together with maximum and minimum values, were utilized [26] to improve estimation under systematic sampling. Memory-type estimators were recently introduced [27] for finite population

mean estimation under systematic sampling. For estimation and monitoring purposes, an exponentially weighted moving average approach was developed [28] under systematic sampling. Difference-based estimation was further investigated [29] to obtain an efficient estimator of the finite population mean.

Research gap and motivation

Although considerable progress has been achieved in developing auxiliary information-based estimators, limited attention has been devoted to hybrid transformation-based estimators under stratified systematic sampling. Most existing ratio, product, regression, and exponential estimators rely on a single adjustment mechanism and are developed under the assumption of approximately linear relationships between the study and auxiliary variables. Consequently, these estimators may not fully utilize auxiliary information when the auxiliary variable exhibits nonlinear behavior, skewness, heterogeneous variability across strata, or the presence of influential observations. Such characteristics can affect the accuracy and reliability of estimation procedures, particularly when the underlying relationships depart substantially from simple structural assumptions. Recent methodological developments in related fields have also emphasized the importance of adaptive estimation strategies for maintaining accuracy and reliability under complex and highly nonlinear conditions [30].

To address this limitation, this paper proposes a new class of weighted hybrid estimators for estimating the finite population mean under stratified systematic sampling. The proposed methodology combines harmonic, geometric, and fractional power transformations of the auxiliary variable within a unified weighted framework. The harmonic transformation is used to reduce the influence of extreme auxiliary observations, the geometric transformation captures multiplicative relationships between the study and auxiliary variables, and the fractional power component provides additional flexibility for modeling moderate nonlinear departures. The unknown weighting constants are determined by minimizing the mean squared error (MSE), resulting in optimally calibrated estimators. Using first-order Taylor series approximations, analytical expressions for the bias and MSE of the proposed estimators are derived. Their theoretical properties are then compared with those of several existing estimators, and their performance is further evaluated through comprehensive simulation experiments and empirical studies based on real populations.

In this study, the term “symmetric” refers to the balanced use of complementary transformation mechanisms within the proposed hybrid framework rather than to a strict algebraic or distributional symmetry condition. The harmonic, geometric, and fractional power components provide different responses to auxiliary-variables’ deviations, while their contributions are determined through MSE-based optimization. This framework allows the proposed estimators to balance the effects of relatively large observations and moderate deviations according to the variability, association, and sampling-design characteristics within each stratum.

The remaining sections of this paper are organized as follows. Section 2 introduces the notation and sampling framework used throughout the paper. Section 3 reviews the existing estimators under stratified systematic sampling. Section 4 presents the proposed classes of hybrid estimators together with their theoretical properties, including their bias and MSE expressions. Section 5 presents a mathematical comparison between the existing estimators and the proposed estimators. Section 6 reports the simulation study conducted under various population structures to assess the estimators’ performance. Section 7 presents empirical comparisons using real data sets. Finally, Section 8 concludes the paper with a discussion of the main findings, practical implications, limitations, and possible directions for future research.

2. Notations

Consider a finite population $Z = (Z_1, Z_2, \dots, Z_N)$ consisting of N distinct units, which are divided into L non overlapping strata. Let N_h denote the number of units in stratum h so that $\sum_{h=1}^L N_h = N$. Let T be the study variable and V be an auxiliary variable available for all population units.

Within each stratum h , suppose we want to select a sample of size n_h using systematic sampling. For simplicity, assume that $N_h = n_h k_h$, where k_h is a positive integer. The population in stratum h is then partitioned into k_h systematic samples of size n_h , and one systematic sample is randomly selected from these k_h possible samples. Let (t_{hij}, v_{hij}) denote the observed values of the study and auxiliary variables for the j th unit in the i th systematic sample from stratum h , where $i = 1, 2, \dots, k_h$ and $j = 1, 2, \dots, n_h$.

The stratum-specific systematic sample means are given by

$$\bar{t}_h = \frac{1}{n_h} \sum_{j=1}^{n_h} t_{hij}$$

and

$$\bar{v}_h = \frac{1}{n_h} \sum_{j=1}^{n_h} v_{hij}.$$

The overall stratified systematic sample means are then computed as the weighted average across strata

$$\bar{t}_{stsy} = \sum_{h=1}^L W_h \bar{t}_h$$

and

$$\bar{v}_{stsy} = \sum_{h=1}^L W_h \bar{v}_h,$$

where

$$W_h = \frac{N_h}{N}$$

is the weight of stratum h . These estimators are unbiased for the population means $\bar{T}$ and $\bar{V}$, assuming that the systematic samples within each stratum are properly randomized.

Within stratum h , the population variance and covariance are defined based on the population units ($i = 1, 2, \dots, N_h$) as follows

$$S_{t_h}^2 = \frac{1}{N_h - 1} \sum_{i=1}^{N_h} (t_{hi} - \bar{T}_h)^2,$$

$$S_{v_h}^2 = \frac{1}{N_h - 1} \sum_{i=1}^{N_h} (v_{hi} - \bar{V}_h)^2,$$

$$S_{tv_h} = \frac{1}{N_h - 1} \sum_{i=1}^{N_h} (t_{hi} - \bar{T}_h)(v_{hi} - \bar{V}_h),$$

where $\bar{T}_h$ and $\bar{V}_h$ are the true stratum means of T and V , respectively. The population coefficients of variation for stratum h , are defined as

$$C_{t_h}^2 = \frac{S_{t_h}^2}{\bar{T}_h^2}$$

and

$$C_{v_h}^2 = \frac{S_{v_h}^2}{\bar{V}_h^2}.$$

Let ρ_{t_h} and ρ_{v_h} denote the intra-class correlations within the systematic sample in stratum h for the study and auxiliary variables, respectively. These are calculated as

$$\rho_{t_h} = \frac{E[(t_{hij} - \bar{T}_h)(t_{hij'} - \bar{T}_h)]}{E[(t_{hij} - \bar{T}_h)^2]}, \quad \rho_{v_h} = \frac{E[(v_{hij} - \bar{V}_h)(v_{hij'} - \bar{V}_h)]}{E[(v_{hij} - \bar{V}_h)^2]},$$

where $j \neq j'$ are distinct units in the selected systematic sample from stratum h . Similarly, the correlation coefficient between the study variable and auxiliary variable within stratum h is defined as

$$\rho_{tv_h} = \frac{E[(t_{hij} - \bar{T}_h)(v_{hij} - \bar{V}_h)]}{\sqrt{E[(t_{hij} - \bar{T}_h)^2] E[(v_{hij} - \bar{V}_h)^2]}}.$$

For analysis, define the relative errors in the stratum-specific systematic sample means as

$$\bar{t}_h = \bar{T}_h(1 + \zeta_{0h}), \quad \bar{v}_h = \bar{V}_h(1 + \zeta_{1h}),$$

where

$$\zeta_{0h} = \frac{\bar{t}_h - \bar{T}_h}{\bar{T}_h},$$

$$\zeta_{1h} = \frac{\bar{v}_h - \bar{V}_h}{\bar{V}_h},$$

such that

$$E(\zeta_{0h}) = E(\zeta_{1h}) = 0.$$

The expected squared and covariances of the relative errors across strata are given by

$$E(\zeta_{0h}^2) = \theta_h C_{t_h}^2 [1 + (n_h - 1)\rho_{t_h}] = \theta_h C_{t_h}^2 D_{t_h},$$

$$E(\zeta_{1h}^2) = \theta_h C_{v_h}^2 [1 + (n_h - 1)\rho_{v_h}] = \theta_h C_{v_h}^2 D_{v_h},$$

and

$$E(\zeta_{0h}\zeta_{1h}) = \theta_h \sqrt{D_{tv_h}} F_{tv_h} C_{v_h}^2,$$

where

$$D_{t_h} = [1 + (n_h - 1)\rho_{t_h}],$$

$$D_{v_h} = [1 + (n_h - 1)\rho_{v_h}],$$

$$D_{tv_h} = \frac{D_{t_h}}{D_{v_h}} = \frac{[1 + (n_h - 1)\rho_{t_h}]}{[1 + (n_h - 1)\rho_{v_h}]},$$

$$F_{tvh} = \rho_{tvh} \frac{C_{t_h}}{C_{v_h}},$$

$$\theta_h = \frac{N_h - 1}{n_h N_h}.$$

The intra-class correlations ρ_{th} and ρ_{vh} affect the estimator efficiency through the design effects D_{th} and D_{vh} , which determine the effective sampling variability of the study and auxiliary variables within each stratum. As the within-stratum dependence changes, these design effects modify the corresponding variance components in the MSE expressions. In addition, the association between the study and auxiliary variables is reflected through ρ_{tvh} and D_{tvh} , which determine the extent to which auxiliary information can reduce the MSE. Since these quantities also enter the optimal weight expressions and minimum MSE formulas of the proposed estimators, the correlation structure can affect both the magnitude of the MSE and the optimal balance between the study variable and transformed auxiliary components. Therefore, the efficiency of the proposed estimators is design-dependent and is jointly determined by within-stratum dependence, auxiliary variables' variability, and the study–auxiliary association.

These notations set the foundation for deriving the biases, MSE's, and efficiencies of estimators under stratified systematic random sampling. To improve readability and assist the reader, we summarize the main notations used throughout the manuscript. This table provides a quick reference for all symbols appearing in the development of the proposed estimators and theoretical results.

3. Existing estimators

(1) Stratified systematic sample mean estimator

The overall stratified systematic sample mean is defined as

$$\bar{t}_{stsy} = \sum_{h=1}^L W_h \bar{t}_h.$$

When no auxiliary information is used, the sampling variance (or MSE) of $\bar{t}_{stsy}$ is given by

$$\text{MSE}(\bar{t}_{stsy}) = \sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 D_{th} C_{t_h}^2. \quad (3.1)$$

(2) Stratified systematic ratio estimator

The stratified adaptation of the classical ratio estimator proposed by Swain [9] can be applied to estimate the population mean $\bar{T}$ by incorporating auxiliary information V within each stratum. In this approach, the ratio estimator is computed separately for each stratum and then combined using stratum weights, resulting in the overall stratified systematic estimator

$$\bar{t}_{Rstsy} = \sum_{h=1}^L W_h \left(\frac{\bar{t}_h}{\bar{v}_h} \right) \bar{V}_h. \quad (3.2)$$

The analytical expressions for the bias and the MSE of the estimator $\bar{t}_{Rstsy}$ are given by

$$\text{Bias}(\bar{t}_{Rstsy}) = \sum_{h=1}^L W_h \theta_h \bar{T}_h \left[D_{v_h} C_{v_h}^2 - \sqrt{D_{tv_h}} F_{tv_h} C_{v_h}^2 \right] \quad (3.3)$$

and

$$\text{MSE}(\bar{t}_{Rstsy}) = \sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[D_{t_h} C_{t_h}^2 + D_{v_h} C_{v_h}^2 \left(1 - 2F_{tv_h} \sqrt{D_{tv_h}} \right) \right]. \quad (3.4)$$

(3) Stratified systematic product estimator

The stratified adaptation of the classical product-type estimator introduced by Shukla [11] can be applied to estimate the population mean under stratified systematic sampling. In this approach, the product estimator is computed separately for each stratum and then combined using stratum weights, yielding the overall stratified systematic estimator

$$\bar{t}_{Pstsy} = \sum_{h=1}^L W_h \left(\frac{\bar{t}_h}{\bar{V}_h} \right) \bar{v}_h. \quad (3.5)$$

The expressions for approximate bias and MSE of $\bar{t}_{Pstsy}$ are, respectively, given by

$$\text{Bias}(\bar{t}_{Pstsy}) = \sum_{h=1}^L W_h \theta_h \bar{T}_h \sqrt{D_{tv_h}} F_{tv_h} C_{v_h}^2 \quad (3.6)$$

and

$$\text{MSE}(\bar{t}_{Pstsy}) = \sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[D_{t_h} C_{t_h}^2 + D_{v_h} C_{v_h}^2 \left(1 + 2F_{tv_h} \sqrt{D_{tv_h}} \right) \right]. \quad (3.7)$$

(4) Stratified systematic classical regression estimator

The stratified systematic form of classical regression estimator for the population mean can be expressed as

$$\bar{t}_{lstsy} = \sum_{h=1}^L W_h \left[\bar{t}_h + b_{tv_h} (\bar{V}_h - \bar{v}_h) \right], \quad (3.8)$$

where b_{tv_h} denotes the sample regression coefficient of t on v in the stratum h . Using a first-order approximation, the MSE of $\bar{t}_{lstsy}$ is given by

$$\text{MSE}(\bar{t}_{lstsy}) = \sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 C_{t_h}^2 D_{t_h} (1 - \rho_{tv_h}^2). \quad (3.9)$$

(5) Stratified systematic exponential ratio and product estimators

Singh et al. [15] proposed exponential-type estimators, including ratio and product variants, for estimating the finite population mean. In the context of stratified systematic sampling, these estimators can be computed within each stratum and then aggregated using the stratum weights, resulting in the overall stratified systematic exponential estimators

$$\bar{t}_{Restsy} = \sum_{h=1}^L W_h \bar{t}_h \exp \left(\frac{\bar{V}_h - \bar{v}_h}{\bar{V}_h + \bar{v}_h} \right) \quad (3.10)$$

and

$$\bar{t}_{Pesty} = \sum_{h=1}^L W_h \bar{t}_h \exp\left(\frac{\bar{v}_h - \bar{V}_h}{\bar{V}_h + \bar{v}_h}\right). \quad (3.11)$$

The formulas for the bias and MSE of the estimators $\bar{t}_{Restsy}$ and $\bar{t}_{Pesty}$ are expressed as

$$\text{Bias}(\bar{t}_{Restsy}) = \sum_{h=1}^L W_h \theta_h \bar{T}_h \left[\frac{3}{8} D_{v_h} C_{v_h}^2 - \frac{1}{2} F_{tv_h} \sqrt{D_{tv_h}} C_{v_h}^2 \right], \quad (3.12)$$

$$\text{Bias}(\bar{t}_{Pesty}) = \sum_{h=1}^L W_h \theta_h \bar{T}_h \left[\frac{1}{2} F_{tv_h} \sqrt{D_{tv_h}} C_{v_h}^2 - \frac{1}{8} D_{v_h} C_{v_h}^2 \right], \quad (3.13)$$

$$\text{MSE}(\bar{t}_{Restsy}) = \sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[D_{t_h} C_{t_h}^2 + \frac{1}{4} D_{v_h} C_{v_h}^2 - F_{tv_h} \sqrt{D_{tv_h}} C_{v_h}^2 \right], \quad (3.14)$$

and

$$\text{MSE}(\bar{t}_{Pesty}) = \sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[D_{t_h} C_{t_h}^2 + \frac{1}{4} D_{v_h} C_{v_h}^2 + F_{tv_h} \sqrt{D_{tv_h}} C_{v_h}^2 \right]. \quad (3.15)$$

(6) Improved stratified systematic exponential ratio estimator

The extended class of exponential estimators introduced by Singh and Jatwa [16] for estimating the population mean $\bar{T}$ under systematic sampling can be adapted to a stratified framework. In this setup, the exponential estimator is calculated within each stratum and then combined using stratum weights, leading to the overall stratified systematic exponential estimator

$$\bar{t}_{asty} = \sum_{h=1}^L W_h \bar{t}_h \exp\left[\frac{g_h (\bar{v}_h - \bar{V}_h)}{\bar{v}_h + \bar{V}_h}\right], \quad (3.16)$$

where g_h is a unknown constant.

The bias and MSE according of the estimator $\bar{t}_{asty}$ to first order of approximation are given by

$$\text{Bias}(\bar{t}_{asty})_{opt} \cong - \sum_{h=1}^L W_h \theta_h \frac{\bar{T}_h}{2} \left[\frac{D_{tv_h} C_{v_h}^2 F_{tv_h}^2}{D_v} \right] \quad (3.17)$$

and

$$\text{MSE}(\bar{t}_{asty})_{opt} \cong \sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[D_{t_h} C_{t_h}^2 - \frac{F_{tv_h}^2 D_{tv_h}}{D_{v_h}} C_{v_h}^2 \right], \quad (3.18)$$

where

$$g_{h(opt)} = \frac{2F_{tv_h} \sqrt{D_{tv_h}}}{D_{v_h}}.$$

4. Proposed mean estimators under stratified systematic sampling

In this section, we develop a novel type of estimators for the population mean under stratified systematic sampling. The proposed approach uses auxiliary information within each stratum through harmonic, geometric, and fractional power transformations, offering a flexible mechanism to capture complex relationships between the study variable T and auxiliary variable V . By combining these transformations, the estimators aim to provide more precise and stable estimates, reducing both bias and MSE, even when the correlation structures vary across different strata.

The proposed estimators are constructed as a weighted combination of the stratum-wise systematic sample mean of the study variable, $\bar{t}_h$, and a transformed component derived from the systematic sample mean of the auxiliary variable, $\bar{v}_h$, relative to the known stratum population mean $\bar{V}_h$. Each estimator applies different transformations, such as harmonic, geometric, or fractional power, to utilize auxiliary information effectively within each stratum. The specific forms of the proposed estimators are presented as follows.

4.1. Theoretical rationale for the hybrid transformations

The selection of harmonic, geometric, and fractional power transformations is motivated by their different responses to the structural characteristics of the auxiliary variable. The harmonic transformation provides a stabilizing adjustment by reducing the influence of relatively large auxiliary observations. This property is useful when the auxiliary variable has substantial dispersion or skewness, since extreme values can otherwise have a disproportionate effect on ratio-type adjustments. The geometric transformation is introduced to capture multiplicative relationships between the study and auxiliary variables. For positive auxiliary variables, this transformation provides an adjustment on a multiplicative scale and is therefore suitable when relative changes in the auxiliary variable are more informative than absolute differences. The fractional power transformation provides additional flexibility for moderate nonlinear departures between the study and auxiliary variables by compressing the scale of the adjustment and allowing a controlled response to deviations in the auxiliary information.

These transformations therefore play complementary rather than redundant roles. The harmonic component emphasizes stability, the geometric component represents multiplicative and skewed structures, and the fractional power component provides flexibility for nonlinear behavior. This combination is particularly relevant under stratified systematic sampling because the variability, skewness, and association between the study and auxiliary variables may differ across strata. The weighted formulation allows the contribution of each transformation to be determined according to the sampling characteristics of the individual strata. Rather than assigning fixed contributions to the individual transformations in advance, the proposed framework determines the corresponding weights through MSE minimization. Consequently, the hybrid structure provides a flexible mechanism for adapting the auxiliary variable adjustment to different structural patterns across strata while retaining the theoretical basis of MSE-based optimization.

Class 1. In the stratified systematic sampling framework, the first estimator combines the stratum-wise sample mean of the study variable with a harmonic-based correction derived from the auxiliary variable. Let $\hat{T}_{H_{1st}}$ denote this estimator, which is expressed as follows:

$$\hat{T}_{H_{1st}} = \sum_{h=1}^L W_h \left[P_{1h} \bar{t}_h + Q_{1h} \left(\frac{2\bar{V}_h \bar{v}_h}{\bar{V}_h + \bar{v}_h} \right) \right], \quad (4.1)$$

where (P_{1h}, Q_{1h}) are suitably chosen constants for stratum h ; $\bar{t}_h$ and $\bar{v}_h$ are the systematic sample means of the study variable T and auxiliary variable V in stratum h ; respectively, and $\bar{V}_h$ is the known stratum population mean of the auxiliary variable.

This formulation applies the harmonic mean between the stratum population mean and sample mean of the auxiliary variable to refine the estimate of $\bar{T}$. It is particularly advantageous when smaller auxiliary values exert more influence within a stratum, as it naturally reduces the effect of extreme values or outliers.

Theorem 1. Let $\hat{T}_{H_{1st}}$ denote the harmonic-adjusted estimator adapted for stratified systematic sampling. The optimal bias and MSE for this estimator within stratum h are determined at the stratum-specific constants P_{1h}^* and Q_{1h}^* as follows:

$$\text{Bias}(\hat{T}_{H_{1st}})_{opt} \cong - \sum_{h=1}^L W_h \theta_h \bar{T}_h \left[\frac{\theta_h D_{t_h} D_{v_h}^2 C_{t_h}^2 C_{v_h}^4 - 4D_{t_h} D_{v_h} C_{t_h}^2 C_{v_h}^2 + 4D_{tv_h} R_h^2}{\Delta_{1h}} \right] \quad (4.2)$$

and

$$\text{MSE}(\hat{T}_{H_{1st}})_{opt} \cong \sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[\frac{\theta_h D_{t_h} D_{v_h}^2 C_{t_h}^2 C_{v_h}^4 - 4D_{t_h} D_{v_h} C_{t_h}^2 C_{v_h}^2 + 4D_{tv_h} R_h^2}{\Delta_{1h}} \right], \quad (4.3)$$

where

$$\Delta_{1h} = \theta_h \left[\left(D_{v_h} C_{v_h}^2 - 2R_h \sqrt{D_{tv_h}} \right)^2 + 4D_{t_h} D_{v_h} C_{t_h}^2 C_{v_h}^2 \right] - 4 \left[D_{v_h} C_{v_h}^2 + 4D_{t_h} C_{t_h}^2 - 4R_h \sqrt{D_{tv_h}} \right],$$

$$R_h = \rho_{tv_h} C_{t_h} C_{v_h}.$$

Proof. The detailed derivation is given in Appendix A.

Remark

- The efficiency improves with an increasing correlation ρ_{tv} between the study and auxiliary variables.
- The harmonic-type adjustment contributes the factor of smaller than classical ratio estimators, improving stability.
- The optimal MSE is always less than the variance of the systematic sample mean $\bar{t}_{sy}$.
- The harmonic-type adjustment ensures stability when $\bar{v}_{sy}$ fluctuates around $\bar{V}$.

Stability remark. The quantity Δ_{1h} acts as the determinant of the coefficient system arising from the MSE minimization for the optimal constants P_{1h}^* and Q_{1h}^* . A value of Δ_{1h} close to zero indicates a nearly singular optimization system and may lead to unstable or excessively large optimal constants. Such a situation can occur under extreme combinations of the design effects, coefficients of variation, and study–auxiliary association within a stratum. Therefore, the closed-form optimal constants should be evaluated only when Δ_{1h} is sufficiently separated from zero. For practical

implementation, the magnitude of Δ_{1h} should be checked before computing the optimal constants; when near-singularity is detected, constrained or regularized MSE minimization can be used to obtain numerically stable weights.

Class 2. Under systematic stratified random sampling, this class of estimators uses auxiliary information through a geometric adjustment. The proposed approach enhances the usual systematic stratified sample mean of the study variable by adding a weighted geometric component based on the auxiliary variable. Specifically, the estimator combines the systematic stratified sample mean with a weighted geometric mean formed from the known population mean and the corresponding systematic stratified sample mean of the auxiliary variable. This adjustment is intended to improve estimation efficiency by effectively utilizing the auxiliary information available at the population level. The mathematical form of the estimator is given as follows:

$$\hat{T}_{H_{2st}} = \sum_{h=1}^L W_h \left[P_{2h} \bar{t}_h + Q_{2h} \sqrt{\bar{V}_h \cdot \bar{v}_h} \right], \quad (4.4)$$

where P_{2h} and Q_{2h} are unknown constants to be determined optimally. The geometric adjustment highlights multiplicative relationships and is particularly appropriate when the auxiliary variable exhibits positive skewness. This approach helps reduce bias in skewed distributions and results in a more stable estimator.

Theorem 2. Let $\hat{T}_{H_{2st}}$ be the weighted geometric-adjusted estimator under systematic random sampling, the optimum expressions for bias and MSE at P_{2h}^* and Q_{2h}^* are expressed as follows:

$$\text{Bias}(\hat{T}_{H_{2st}})_{\text{opt}} \cong - \sum_{h=1}^L W_h \theta_h \bar{T}_h \left[\frac{\theta_h D_{th} D_{vh}^2 C_{th}^2 C_{vh}^4 + 16 D_{tvh} R_h^2}{\Delta_{2h}} \right] \quad (4.5)$$

and

$$\text{MSE}(\hat{T}_{H_{2st}})_{\text{opt}} \cong \sum_{h=1}^L W_h^2 \theta_h^2 \bar{T}_h^2 \left[\frac{\theta_h D_{th} D_{vh}^2 C_{th}^2 C_{vh}^4 + 16 D_{tvh} R_h^2}{\Delta_{2h}} \right], \quad (4.6)$$

where

$$\Delta_{2h} = \theta_h \left\{ \left(D_{vh} C_{vh}^2 - 4 R_h \sqrt{D_{tvh}} \right)^2 + 16 D_{th} D_{vh} C_{th}^2 C_{vh}^2 \right\} - 64 \left(D_{th} C_{th}^2 - R_h \sqrt{D_{tvh}} \right).$$

Proof. The detailed derivation is given in Appendix B.

Class 3. Within the framework of stratified systematic sampling, this class introduces a hybrid estimator that makes use of auxiliary information through two complementary adjustments. The proposed estimator modifies the usual stratified systematic sample mean of the study variable by adding a harmonic-type component derived from the auxiliary variable, together with a fractional power term. The harmonic part is constructed using the known population mean of the auxiliary variable and its corresponding stratified systematic sample mean, while the fractional transformation is included to capture possible nonlinear discrepancies between these quantities. This combined structure allows the estimator to flexibly adapt to variations present in the auxiliary data across strata. The estimator is expressed as

$$\hat{T}_{H_{3st}} = \sum_{h=1}^L W_h \left[P_{3h} \bar{t}_h + P_{4h} \left(\frac{2 \bar{V}_h \bar{v}_h}{\bar{V}_h + \bar{v}_h} \right) + P_{5h} \left(\frac{\bar{v}_h}{\bar{V}_h} \right)^{\frac{1}{2}} \right], \quad (4.7)$$

where (P_{3h}, P_{4h}, P_{5h}) are unknown constants to be determined optimally. By combining a harmonic adjustment with a fractional power, this estimator provides additional flexibility for accommodating moderate nonlinear departures between the study and auxiliary variables, while the harmonic component helps stabilize the adjustment when the auxiliary variable exhibits substantial variability or skewness across strata.

Theorem 3. Let $\hat{T}_{H_{3st}}$ denote the proposed weighted hybrid harmonic and fractional power estimator of the population mean $\bar{T}_h$ under stratified systematic sampling. The minimum bias and MSE expressions at the optimal values P_{3h}^* , P_{4h}^* , and P_{5h}^* are defined as:

$$\text{Bias}(\hat{T}_{H_{3st}})_{\text{opt}} \cong - \sum_{h=1}^L \theta_h W_h \bar{T}_h \left[\frac{D_{t_h} D_{v_h} C_{t_h}^2 C_{v_h}^2 - D_{tvh} F_{tvh}^2 C_{v_h}^4}{\Omega_h} \right] \quad (4.8)$$

and

$$\text{MSE}(\hat{T}_{H_{3st}})_{\text{opt}} \cong \sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[\frac{D_{t_h} D_{v_h} C_{t_h}^2 C_{v_h}^2 - D_{tvh} F_{tvh}^2 C_{v_h}^4}{\Omega_h} \right], \quad (4.9)$$

where

$$\Omega_h = 4D_{t_h} C_{t_h}^2 + D_{v_h} C_{v_h}^2 - 4\sqrt{D_{tvh}} F_{tvh} C_{v_h}^2.$$

Proof. The detailed derivation is given in Appendix C.

5. Mathematical comparison

In this section, the efficiency conditions are established through a comparison of the MSE expressions of the proposed family of estimators presented in Section 4 with those of the existing estimators described in Section 3.

5.1. First proposed estimator

(i) A comparison between the MSE expressions given in (3.1) and (4.3) leads to the following efficiency condition:

$$\text{MSE}(\bar{t}_{stsy}) > \text{MSE}(\hat{T}_{H_{1st}})_{\text{opt}} \text{ if } \sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[D_{t_h} C_{t_h}^2 - \frac{\theta_h B_h}{\Delta_{1h}} \right] > 0,$$

where

$$B_h = D_{t_h} D_{v_h}^2 C_{t_h}^2 C_{v_h}^4 - 4D_{t_h} D_{v_h} C_{t_h}^2 C_{v_h}^2 + 4D_{tvh} R_h^2.$$

(ii) By comparing the expression obtained in (3.4) with that given in (4.3), we obtain the following condition:

$$\text{MSE}(\bar{t}_{Rstsy}) > \text{MSE}(\hat{T}_{H_{1st}})_{\text{opt}} \text{ if } \sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[D_{t_h} C_{t_h}^2 + D_{v_h} C_{v_h}^2 (1 - 2F_{tvh} \sqrt{D_{tvh}}) - \frac{\theta_h B_h}{\Delta_{1h}} \right] > 0.$$

(iii) The following condition is obtained by comparing the MSE in (3.7) with that provided in (4.3):

$\text{MSE}(\bar{t}_{Pstsy}) > \text{MSE}(\hat{\hat{T}}_{H1st})_{opt}$ if

$$\sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[D_{t_h} C_{t_h}^2 + D_{v_h} C_{v_h}^2 (1 + 2F_{tv_h} \sqrt{D_{tv_h}}) - \frac{\theta_h B_h}{\Delta_{1h}} \right] > 0.$$

(iv) The following condition is obtained by comparing the expression in (3.9) with its counterpart in (4.3):

$\text{MSE}(\bar{t}_{lstsy}) > \text{MSE}(\hat{\hat{T}}_{H1st})_{opt}$ if

$$\sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[C_{t_h}^2 D_{t_h} (1 - \rho_{tv_h}^2) - \frac{\theta_h B_h}{\Delta_{1h}} \right] > 0.$$

(v) By comparing the expression in (3.14) with that given in (4.3), the following condition is derived:

$\text{MSE}(\bar{t}_{Restsy}) > \text{MSE}(\hat{\hat{T}}_{H1st})_{opt}$ if

$$\sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[D_{t_h} C_{t_h}^2 + \frac{1}{4} D_{v_h} C_{v_h}^2 - F_{tv_h} \sqrt{D_{tv_h}} C_{v_h}^2 - \frac{\theta_h B_h}{\Delta_{1h}} \right] > 0.$$

(vi) A comparison between the expression derived in (3.15) and that given in (4.3) yields the following condition:

$\text{MSE}(\bar{t}_{Pestsy}) > \text{MSE}(\hat{\hat{T}}_{H1st})_{opt}$ if

$$\sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[D_{t_h} C_{t_h}^2 + \frac{1}{4} D_{v_h} C_{v_h}^2 + F_{tv_h} \sqrt{D_{tv_h}} C_{v_h}^2 - \frac{\theta_h B_h}{\Delta_{1h}} \right] > 0.$$

(vii) The following condition is obtained by comparing the expression in (3.18) with its counterpart in (4.3):

$\text{MSE}(\bar{t}_{aestsy})_{opt} > \text{MSE}(\hat{\hat{T}}_{H1st})_{opt}$ if

$$\sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[D_{t_h} C_{t_h}^2 - \frac{F_{tv_h}^2 D_{tv_h}}{D_{v_h}} C_{v_h}^2 - \frac{\theta_h B_h}{\Delta_{1h}} \right] > 0.$$

5.2. Second proposed estimator

(i) A comparison between the MSE expressions given in (3.1) and (4.6) leads to the following efficiency condition:

$$\sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[D_{t_h} C_{t_h}^2 - \frac{B'_h}{\Delta_{2h}} \right] > 0,$$

where

$$B'_h = \theta_h D_{t_h} D_{v_h}^2 C_{t_h}^2 C_{v_h}^4 + 16 D_{tv_h} R_h^2.$$

(ii) By comparing the expression obtained in (3.4) with that given in (4.6), we obtain the following condition:

$\text{MSE}(\bar{t}_{Rstsy}) > \text{MSE}(\hat{\hat{T}}_{H2st})_{opt}$ if

$$\sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[D_{t_h} C_{t_h}^2 + D_{v_h} C_{v_h}^2 (1 - 2F_{tv_h} \sqrt{D_{tv_h}}) - \frac{B'_h}{\Delta_{2h}} \right] > 0.$$

(iii) The following condition is obtained by comparing the MSE in (3.7) with that provided in (4.6):

$$\text{MSE}(\bar{t}_{Pstsy}) > \text{MSE}(\hat{\hat{T}}_{H_{2st}})_{opt} \text{ if}$$

$$\sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[D_{t_h} C_{t_h}^2 + D_{v_h} C_{v_h}^2 (1 + 2F_{tv_h} \sqrt{D_{tv_h}}) - \frac{B'_h}{\Delta_{2h}} \right] > 0.$$

(iv) The following condition is obtained by comparing the expression in (3.9) with its counterpart in (4.6):

$$\text{MSE}(\bar{t}_{lstsy}) > \text{MSE}(\hat{\hat{T}}_{H_{2st}})_{opt} \text{ if}$$

$$\sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[C_{t_h}^2 D_{t_h} (1 - \rho_{tv_h}^2) - \frac{B'_h}{\Delta_{2h}} \right] > 0.$$

(v) By comparing the expression in (3.14) with that given in (4.6), the following condition is derived:

$$\text{MSE}(\bar{t}_{Restsy}) > \text{MSE}(\hat{\hat{T}}_{H_{2st}})_{opt} \text{ if}$$

$$\sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[D_{t_h} C_{t_h}^2 + \frac{1}{4} D_{v_h} C_{v_h}^2 - F_{tv_h} \sqrt{D_{tv_h}} C_{v_h}^2 - \frac{B'_h}{\Delta_{2h}} \right] > 0.$$

(vi) A comparison between the expression derived in (3.15) and that given in (4.6) yields the following condition:

$$\text{MSE}(\bar{t}_{Pestsy}) > \text{MSE}(\hat{\hat{T}}_{H_{2st}})_{opt} \text{ if}$$

$$\sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[D_{t_h} C_{t_h}^2 + \frac{1}{4} D_{v_h} C_{v_h}^2 + F_{tv_h} \sqrt{D_{tv_h}} C_{v_h}^2 - \frac{B'_h}{\Delta_{2h}} \right] > 0.$$

(vii) The following condition is obtained by comparing the expression in (3.18) with its counterpart in (4.6):

$$\text{MSE}(\bar{t}_{aestsy})_{opt} > \text{MSE}(\hat{\hat{T}}_{H_{2st}})_{opt} \text{ if}$$

$$\sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[D_{t_h} C_{t_h}^2 - \frac{F_{tv_h}^2 D_{tv_h}}{D_{v_h}} C_{v_h}^2 - \frac{B'_h}{\Delta_{2h}} \right] > 0.$$

5.3. Third proposed estimator

(i) A comparison between the MSE expressions given in (3.1) and (4.9) leads to the following efficiency condition:

$$\text{MSE}(\bar{t}_{stsy}) > \text{MSE}(\hat{\hat{T}}_{H_{3st}})_{opt} \text{ if}$$

$$\sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[D_{t_h} C_{t_h}^2 - \left(\frac{D_{t_h} D_{v_h} C_{t_h}^2 C_{v_h}^2 - D_{tv_h} F_{tv_h}^2 C_{v_h}^4}{\Omega_h} \right) \right] > 0$$

(ii) By comparing the expression obtained in (3.4) with that given in (4.9), we obtain the following condition:

$\text{MSE}(\bar{t}_{Rstsy}) > \text{MSE}(\hat{\bar{T}}_{H_{3st}})_{opt}$ if

$$\sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[D_{t_h} C_{t_h}^2 + D_{v_h} C_{v_h}^2 (1 - 2F_{tv_h} \sqrt{D_{tv_h}}) - \left(\frac{D_{t_h} D_{v_h} C_{t_h}^2 C_{v_h}^2 - D_{tv_h} F_{tv_h}^2 C_{v_h}^4}{\Omega_h} \right) \right] > 0.$$

(iii) The following condition is obtained by comparing the MSE in (3.7) with that provided in (4.9):

$\text{MSE}(\bar{t}_{Pstsy}) > \text{MSE}(\hat{\bar{T}}_{H_{3st}})_{opt}$ if

$$\sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[D_{t_h} C_{t_h}^2 + D_{v_h} C_{v_h}^2 (1 + 2F_{tv_h} \sqrt{D_{tv_h}}) - \left(\frac{D_{t_h} D_{v_h} C_{t_h}^2 C_{v_h}^2 - D_{tv_h} F_{tv_h}^2 C_{v_h}^4}{\Omega_h} \right) \right] > 0.$$

(iv) The following condition is obtained by comparing the expression in (3.9) with its counterpart in (4.9):

$\text{MSE}(\bar{t}_{lstsy}) > \text{MSE}(\hat{\bar{T}}_{H_{3st}})_{opt}$ if

$$\sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[D_{t_h} C_{t_h}^2 (1 - \rho_{tv_h}^2) - \left(\frac{D_{t_h} D_{v_h} C_{t_h}^2 C_{v_h}^2 - D_{tv_h} F_{tv_h}^2 C_{v_h}^4}{\Omega_h} \right) \right] > 0.$$

(v) By comparing the expression in (3.14) with that given in (4.9), the following condition is derived:

$\text{MSE}(\bar{t}_{Restsy}) > \text{MSE}(\hat{\bar{T}}_{H_{3st}})_{opt}$ if

$$\sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[D_{t_h} C_{t_h}^2 + \frac{1}{4} D_{v_h} C_{v_h}^2 - F_{tv_h} \sqrt{D_{tv_h}} C_{v_h}^2 - \left(\frac{D_{t_h} D_{v_h} C_{t_h}^2 C_{v_h}^2 - D_{tv_h} F_{tv_h}^2 C_{v_h}^4}{\Omega_h} \right) \right] > 0.$$

(vi) A comparison between the expression derived in (3.15) and that given in (4.9) yields the following condition:

$\text{MSE}(\bar{t}_{Pestsy}) > \text{MSE}(\hat{\bar{T}}_{H_{3st}})_{opt}$ if

$$\sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[D_{t_h} C_{t_h}^2 + \frac{1}{4} D_{v_h} C_{v_h}^2 + F_{tv_h} \sqrt{D_{tv_h}} C_{v_h}^2 - \left(\frac{D_{t_h} D_{v_h} C_{t_h}^2 C_{v_h}^2 - D_{tv_h} F_{tv_h}^2 C_{v_h}^4}{\Omega_h} \right) \right] > 0.$$

(vii) The following condition is obtained by comparing the expression in (3.18) with its counterpart in (4.9):

$\text{MSE}(\bar{t}_{aestsy})_{opt} > \text{MSE}(\hat{\bar{T}}_{H_{1st}})_{opt}$ if

$$\sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[D_{t_h} C_{t_h}^2 - \frac{F_{tv_h}^2 D_{tv_h}}{D_{v_h}} C_{v_h}^2 - \left(\frac{D_{t_h} D_{v_h} C_{t_h}^2 C_{v_h}^2 - D_{tv_h} F_{tv_h}^2 C_{v_h}^4}{\Omega_h} \right) \right] > 0.$$

6. Simulation study

The present simulation experiments are designed to compare the performance of the suggested weighted hybrid estimators with the previous classical estimators and investigate the characteristics of the proposed estimators in the framework of systematic sampling. The simulation is specifically intended to investigate the impact of sample fraction, population structure, and correlation strength on MSEs and percentage relative efficiencies.

For every simulation scenario, a finite population of size $N = 2400$ is produced. A research variable T and an auxiliary variable V with a predefined correlation coefficient ρ_{tv} form each population. The population units are separated into $L = 4$ strata and indexed as $i = 1, 2, \dots, N$. For the purposes of systematic sampling, units within each stratum are arranged according to the auxiliary variable V . Three models of populations are examined.

Model I: Linear trend population

The population is generated according to the model

$$T_i = 100 + 0.3i + \varepsilon_i$$

and

$$V_i = 50 + 0.25i + \eta_i,$$

where $\varepsilon_i \sim N(0, 25^2)$ and $\eta_i \sim N(0, 16^2)$ are independent normal errors. The correlation coefficient ρ_{tv} is targeted at $\{0.65, 0.80, 0.90\}$.

Model II: Nonlinear and skewed population

The auxiliary variable is generated from a log-normal distribution,

$$V_i \sim \text{Lognormal}(2.8, 0.6^2)$$

and the study variable is generated as

$$T_i = 20 + 1.5V_i + 0.5\sqrt{V_i} + \varepsilon_i,$$

where $\varepsilon_i \sim N(0, 20^2)$.

Approximation under skewed populations

Since the theoretical bias and MSE expressions are derived using first-order Taylor approximations, some discrepancy between the theoretical and empirical MSE values may arise when the auxiliary variable is strongly skewed or contains influential observations. To assess the practical behavior of the proposed estimators under such conditions, Model II uses a log-normal auxiliary variable and a nonlinear relationship between the study and auxiliary variables. The simulation results therefore provide an empirical assessment of the proposed estimators under a skewed and nonlinear population structure. The theoretical expressions are interpreted as first-order approximations for comparative efficiency analysis rather than as exact finite-sample MSE values.

Model III: Periodic population

To introduce periodicity, the population is generated as

$$T_i = 80 + 15 \sin\left(\frac{2\pi i}{50}\right) + \varepsilon_i$$

and

$$V_i = 60 + 12 \sin\left(\frac{2\pi i}{50}\right) + \eta_i,$$

where $\varepsilon_i \sim N(0, 10^2)$ and $\eta_i \sim N(0, 8^2)$ are independent.

Three population models were considered to evaluate the performance of the proposed estimators under different structural relationships between the study and auxiliary variables. The linear model represents a simple association, whereas the nonlinear and periodic models provide more challenging scenarios involving complex and cyclic patterns.

6.1. Stratified systematic sampling design

Following population generation, units are separated into about equal-sized $L = 4$ strata ($N_h = 600$). From each stratum, a systematic sample of size n_h is taken; n_h is determined using proportionate allocation:

$$n_h = n \left(\frac{N_h}{N} \right), \quad n \in \{60, 120, 240\}.$$

For stratum h , the sampling interval is

$$k_h = \left\lfloor \frac{N_h}{n_h} \right\rfloor$$

and a random start r_h is selected uniformly from $\{1, 2, \dots, k_h\}$. The systematic sample from stratum h is

$$S_h = \{r_h, r_h + k_h, r_h + 2k_h, \dots, r_h + (n_h - 1)k_h\}.$$

The complete stratified systematic sample is

$$S = \bigcup_{h=1}^L S_h.$$

The stratum-level sample means are

$$\bar{t}_h = \frac{1}{n_h} \sum_{i \in S_h} T_i, \quad \bar{v}_h = \frac{1}{n_h} \sum_{i \in S_h} V_i$$

and the overall stratified systematic sample means are

$$\bar{t}_{sy} = \sum_{h=1}^L W_h \bar{t}_h, \quad \bar{v}_{sy} = \sum_{h=1}^L W_h \bar{v}_h.$$

6.2. Simulation algorithm

1. Generate a finite population of size $N = 2400$ based on one of the three models above, dividing it into $L = 4$ strata.
2. Compute the overall population mean $\bar{T}$ and auxiliary mean $\bar{V}$, as well as stratum means $\bar{T}_h$ and $\bar{V}_h$.

3. Select a total sample size $n \in \{60, 120, 240\}$ and calculate stratum sample sizes n_h using proportional allocation.
4. For each stratum, compute the sampling interval $k_h = \lfloor N_h/n_h \rfloor$ and choose a random start r_h .
5. Draw systematic samples S_h for each stratum and combine to form the full stratified systematic sample $S = \bigcup_{h=1}^L S_h$.
6. Compute the stratum-level sample means $\bar{t}_h$ and $\bar{v}_h$, then the overall stratified systematic sample means $\bar{t}_{sy}$ and $\bar{v}_{sy}$.
7. Calculate all proposed and comparative estimators using the stratified systematic sample.
8. Repeat Steps 4–7 for $R = 10,000$ Monte Carlo replications.
9. Store the estimated values for each replication and estimator.
10. Evaluate the performance of each estimator by computing:

$$\text{MSE}(\hat{T}) = \frac{1}{R} \sum_{r=1}^R (\hat{T}_r - \bar{T})^2$$

and

$$\text{PRE}(\hat{T}) = \frac{\text{MSE}(\bar{t}_{tsy})}{\text{MSE}(\hat{T})} \times 100.$$

11. Summarize the results in tables for different population models, correlation levels, and sample sizes are given in Tables 1–3.

Table 1. Simulation results for $n = 60$ under different population models using stratified systematic sampling.

Estimator	Model I (linear)		Model II (nonlinear)		Model III (periodic)	
	MSE	PRE	MSE	PRE	MSE	PRE
$\bar{t}_{stsy}$	318.762	100.000	389.541	100.000	298.374	100.000
$\bar{t}_{Rstsy}$	221.936	143.633	265.728	146.638	198.442	150.389
$\bar{t}_{Pstsy}$	238.574	133.585	291.263	133.713	216.829	137.602
$\bar{t}_{lstsy}$	207.845	153.389	255.417	152.442	192.316	155.159
$\bar{t}_{Restsy}$	172.638	184.528	210.436	185.097	154.327	193.369
$\bar{t}_{Pestsy}$	184.215	172.974	220.638	176.575	164.782	181.066
$\bar{t}_{aestsy}$	166.438	191.280	199.728	195.024	150.614	198.233
$\hat{T}_{H1st}$	124.736	255.542	168.927	230.628	134.516	221.812
$\hat{T}_{H2st}$	128.394	248.273	149.638	260.334	126.482	235.918
$\hat{T}_{H3st}$	145.817	218.453	177.326	219.591	139.284	214.233

Table 2. Simulation results for $n = 120$ under different population models using stratified systematic sampling.

Estimator	Model I (linear)		Model II (nonlinear)		Model III (periodic)	
	MSE	PRE	MSE	PRE	MSE	PRE
$\bar{t}_{stsy}$	220.18	100.00	267.42	100.00	205.31	100.00
$\bar{t}_{Rstsy}$	149.76	147.03	181.35	147.48	132.42	155.05
$\bar{t}_{Pstsy}$	158.84	138.63	194.26	137.66	142.18	144.40
$\bar{t}_{lstsy}$	139.25	158.12	171.44	155.96	126.91	161.81
$\bar{t}_{Restsy}$	108.34	203.24	138.72	192.77	91.86	223.52
$\bar{t}_{Pestsy}$	117.26	187.78	146.85	182.10	98.74	207.92
$\bar{t}_{aestsy}$	101.45	217.01	131.18	203.85	86.37	237.71
$\hat{\hat{T}}_{H_{1st}}$	72.18	305.03	98.52	271.44	65.43	313.76
$\hat{\hat{T}}_{H_{2st}}$	68.54	321.25	86.73	308.21	60.82	337.59
$\hat{\hat{T}}_{H_{3st}}$	82.37	267.29	109.64	243.90	74.51	275.55

Table 3. Simulation results for $n = 240$ under different population models using stratified systematic sampling.

Estimator	Model I (linear)		Model II (nonlinear)		Model III (periodic)	
	MSE	PRE	MSE	PRE	MSE	PRE
$\bar{t}_{stsy}$	158.64	100.00	193.18	100.00	148.26	100.00
$\bar{t}_{Rstsy}$	104.86	151.31	131.24	147.18	88.76	166.99
$\bar{t}_{Pstsy}$	112.73	140.82	141.62	136.40	96.38	153.83
$\bar{t}_{lstsy}$	98.41	161.20	125.37	154.10	90.24	164.25
$\bar{t}_{Restsy}$	72.83	217.81	96.45	200.35	62.74	236.25
$\bar{t}_{Pestsy}$	79.26	200.31	103.62	186.45	68.91	215.16
$\bar{t}_{aestsy}$	69.18	229.33	90.74	212.91	58.63	252.87
$\hat{\hat{T}}_{H_{1st}}$	45.72	347.14	66.83	288.99	39.86	371.86
$\hat{\hat{T}}_{H_{2st}}$	41.56	381.79	57.92	333.41	35.48	417.87
$\hat{\hat{T}}_{H_{3st}}$	52.84	300.30	72.46	266.63	45.17	328.15

6.3. Interpretation of simulation results

The empirical MSEs and percentage of relative efficiencies (PREs) of the competing estimators for three different population models and three sample sizes using the stratified systematic sampling strategy are shown in Tables 1–3. The stratified systematic sample mean, represented by $\bar{t}_{stsy}$, serves as the baseline estimator. With $n = 60, 120, 240$, as the sample size increases, the MSE of all estimators lowers. This is consistent with classical sampling theory and validates the efficiency improvements possible under stratified systematic sampling. Additionally, even at moderate sample sizes, the relative efficiency of the suggested estimators are significantly greater than the baseline estimator, demonstrating the benefits of using auxiliary data in a stratified design.

The estimators' performance exhibits a clear and consistent pattern across the three population models. All auxiliary variable-based estimators outperform the traditional stratified systematic estimator under the linear population (Model I). The proposed hybrid estimators provide the greatest improvement in efficiency, with PRE values increasing from approximately 218% to 256% for ($n = 60$), from 267% to 321% for ($n = 120$), from and 300% to 382% for ($n = 240$). Under the nonlinear population (Model II), where the conventional ratio- and product-type estimators provide comparatively smaller efficiency gains, the proposed hybrid estimators maintain substantially higher PRE values, demonstrating their robustness in handling nonlinear population structures. Similarly, for the periodic population (Model III), where systematic sampling is often influenced by population ordering, the proposed estimators continue to exhibit superior performance by producing the lowest MSE values and the highest PREs among all competing estimators. Overall, these results confirm that the proposed estimators remain highly efficient and stable under linear, nonlinear, and periodic population structures.

Across all population models and sample sizes, the proposed hybrid estimators $\hat{T}_{H1st}$, $\hat{T}_{H2st}$, and $\hat{T}_{H3st}$ consistently produce the smallest MSE values among all estimators considered. Among them, $\hat{T}_{H2st}$ is uniformly the most efficient estimator in every simulation scenario. Its PRE ranges from 236% to 260% for $n = 60$, from 308% to 338% for $n = 120$, and from 333% to 418% for $n = 240$ across the three population models. These results demonstrate that combining harmonic, geometric, and fractional power transformations of the auxiliary variable within the stratified systematic sampling framework yields substantial improvements in estimation efficiency over the existing estimators. Figure 1 provides a graphical comparison of the PRE values for all estimators under the three population models and different sample sizes. The plots clearly show that the proposed hybrid estimators (Estimators 8–10) consistently achieve higher PRE values than the existing estimators, with Estimator 9 ($\hat{T}_{H2st}$) exhibiting the highest efficiency across all population models and sample sizes.

According to the simulation results, there are significant efficiency advantages when auxiliary information is used in a stratified systematic design. With $\hat{T}_{H2st}$ consistently obtaining the smallest MSE, the suggested hybrid estimators are robust to linear, nonlinear, and periodic population patterns and are a good contender for real-world applications. All things considered, these results imply that the suggested family of estimators offers a very safe and efficient substitute for traditional stratified systematic sampling estimators, especially in situations where the population displays complicated trends, nonlinearity, or periodicity. In conclusion, the stratified systematic sampling simulations confirm that the proposed estimators considerably improve estimation efficiency over the classical mean, and the hybrid estimator $\hat{T}_{D2sy}$ remains the most efficient choice across all sample sizes, population structures, and correlation levels.

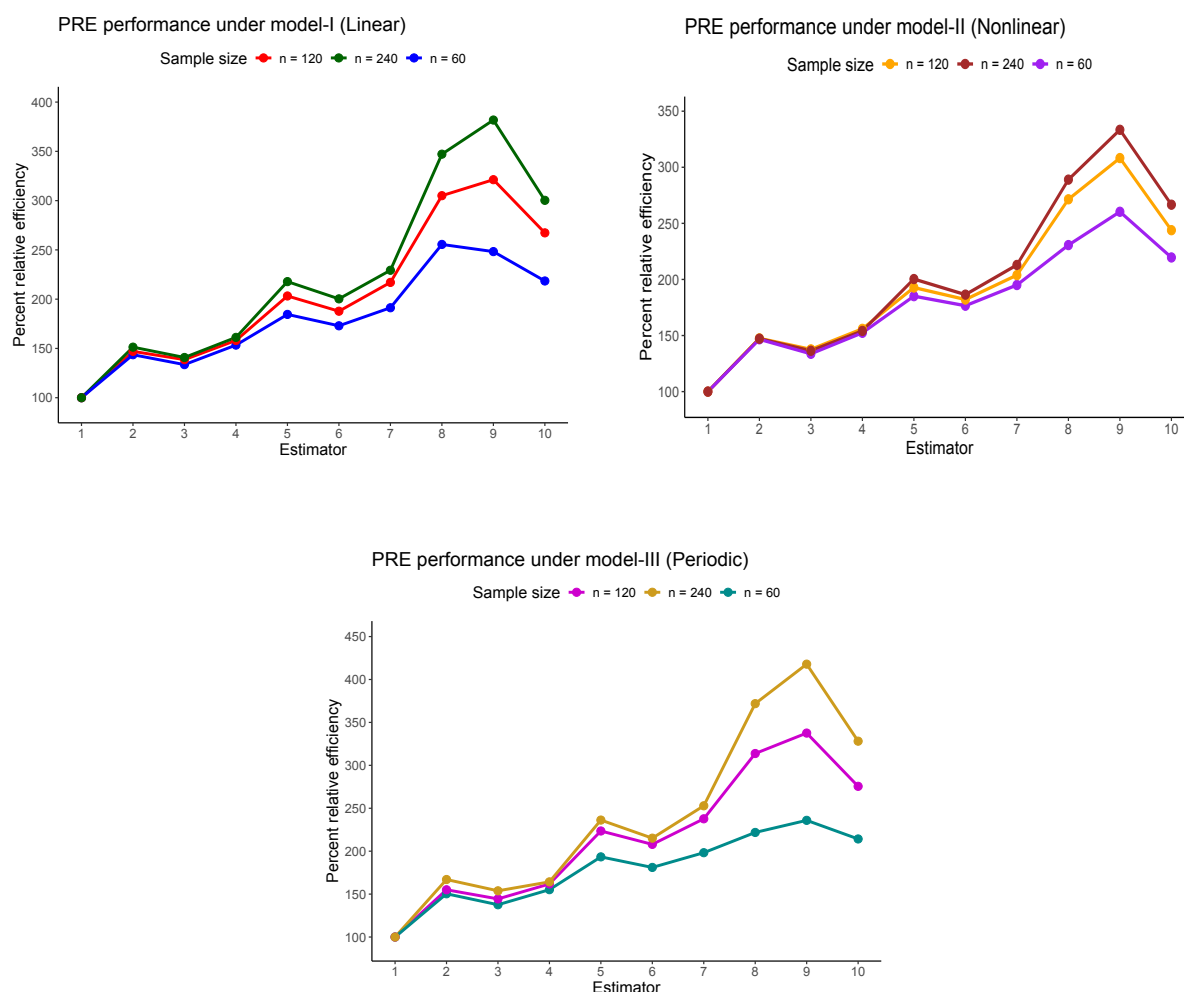


Figure 1. Graphical representation of the simulation results obtained under different population models. Estimators 1–7 represent the existing estimators, whereas Estimators 8–10 denote the proposed estimators.

7. Empirical study

To conduct empirical analysis, three independent data sets are used to evaluate and compare the numerical performance of the proposed methodology. The data sets are chosen in order to capture various population structures and variability patterns to enable a thorough and effectively investigated. Given the comparative analysis, it is necessary to state that the use of multiple data sources makes it strong and not tied to one experimental situation. The entire numerical findings in this paper are obtained on the basis of these three benchmark data sets obtain from [2, 31, 32]. The MSEs and PREs values are given in Table 4.

7.1. Empirical study based on real populations

The performance of the suggested estimators under a stratified systematic sampling design is further assessed by taking into account three real finite populations. The sample size, level of variability,

and degree of correlation between the study variable and the auxiliary variable vary among these groups. The range of possibilities provides a thorough analysis of the estimators' performance in many practical situations and guarantees that the empirical evaluation is not limited to a single population structure. Using appropriate auxiliary data, strata are generated for every population. Within each stratum, systematic samples are chosen independently. The competing estimators' MSE and PRE are calculated using the known population parameters. The PRE of the stratified systematic sample mean, which is used as the benchmark estimator, is set to 100 for all populations so that efficiency improvements between the estimators may be directly compared.

Table 4. MSE and PREs of various estimators for real populations under stratified systematic sampling.

Estimator	Population 1		Population 2		Population 3	
	MSE	PRE	MSE	PRE	MSE	PRE
$\bar{t}_{stsy}$	1320.45	100.00	2105.38	100.00	790200.62	100.00
$\bar{t}_{Rstsy}$	420.37	314.25	455.46	462.78	760500.84	103.90
$\bar{t}_{Pstsy}$	3010.56	43.86	5400.27	38.99	5523000.73	14.31
$\bar{t}_{lstsy}$	403.49	327.26	413.90	508.67	435822.08	181.31
$\bar{t}_{Restsy}$	1095.22	120.55	1680.64	125.28	301400.35	262.35
$\bar{t}_{Pestsy}$	1640.73	80.48	2860.51	73.60	1586000.48	49.82
$\bar{t}_{aestsy}$	1012.84	130.41	1595.36	131.99	602800.57	131.08
$\hat{\hat{T}}_{H_{1st}}$	347.68	380.59	386.29	545.08	252918.42	312.43
$\hat{\hat{T}}_{H_{2st}}$	330.54	399.48	378.73	555.91	241963.65	326.71
$\hat{\hat{T}}_{H_{3st}}$	389.43	339.07	395.22	532.71	277082.44	285.19

Results in Table 4 show that the classical ratio and product-type estimators are unstable in various populations and have a comparatively larger MSE value in certain populations. Conversely, the difference-type estimators, i.e., $\hat{\hat{T}}_{H_{1st}}$, $\hat{\hat{T}}_{H_{2st}}$, and $\hat{\hat{T}}_{H_{3st}}$, all have lower MSEs in all three populations when using a stratified systematic sampling scheme. This indicates their capability to use auxiliary information to the best and take within-stratum homogeneity. Out of all estimators to be evaluated, the minimum MSE and maximum PRE of the estimator of the population mean are found with the use of the estimator, $\hat{\hat{T}}_{H_{2st}}$, making it the most efficient estimator of the population mean. The effectiveness of this estimator is enhanced by the fact that the combination of stratification and systematic sampling, with the appropriately designed auxiliary-based adjustments, is beneficial. Figure 2 provides a graphical comparison of the PRE values for all estimators under the three population models and different sample sizes. The plots clearly show that the proposed hybrid estimators (Estimators 8–10) consistently achieve higher PRE values than the existing estimators, with Estimator 9 ($\hat{\hat{T}}_{H_{2st}}$) exhibiting the highest efficiency across all population models and sample sizes.

The theoretical outcomes are overall well supported by the empirical findings. The continuing decline in MSE in a variety of population contexts proves that the suggested estimators utilize effectively auxiliary information and provide sound and effective alternatives to the estimation of the population mean in the real-life applications of using survey directed at stratified systematic sampling.

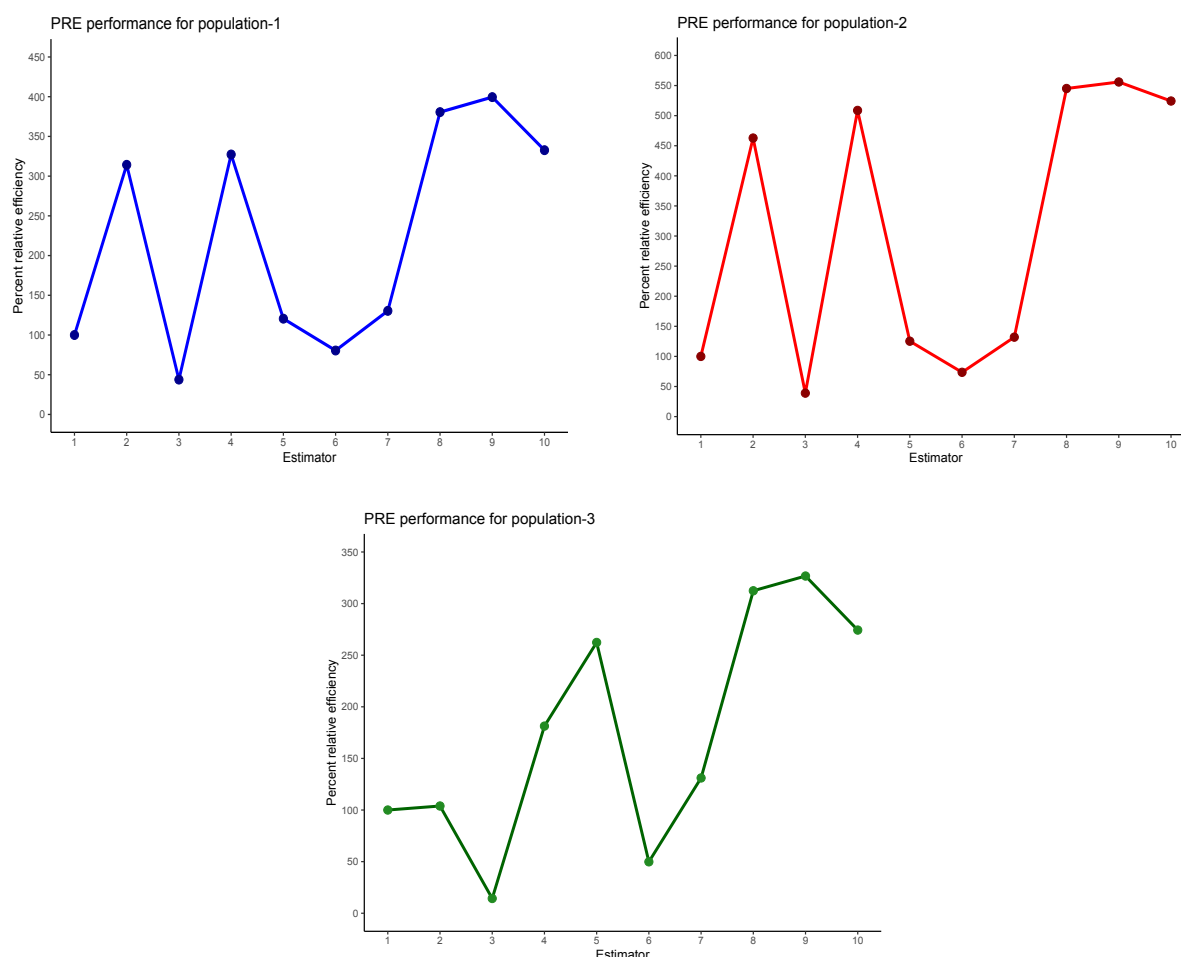


Figure 2. Graphical representation of the simulation results obtained under different population models. Estimators 1–7 represent the existing estimators, whereas Estimators 8–10 denote the proposed estimators.

7.2. Significance and benefits of the proposed estimators

In applied survey research, auxiliary information is available and often exhibits a strong association with the study variable. Under stratified random sampling, classical estimators generally use linear adjustments within or across strata, which may lead to sensitivity to extreme observations, skewed auxiliary distributions, or nonlinear relationships. Such limitations can compromise estimators' reliability, particularly when auxiliary characteristics vary substantially across strata or deviate from the ideal distributional assumptions.

- **Maximizing auxiliary data:** The proposed estimators efficiently incorporate available auxiliary data within the stratified framework, improving population mean estimation, especially when

strong correlations exist between study and auxiliary variables at the stratum level.

- **Minimization of bias and MSE:** These estimator classes are designed to reduce bias and MSE compared with traditional linear estimators, particularly under moderate to high correlation, through harmonic, geometric, and fractional power transformations.
- **Stability under skewness and outliers:** When auxiliary variables are skewed or contain outliers within strata, the estimators remain robust, as harmonic and geometric adjustments downweight extreme values.
- **Controlled adjustments via fractional powers:** Fractional power terms provide controlled corrections, allowing the estimators to handle moderate deviations between population and sample characteristics without overreacting to large differences.
- **Handling heterogeneous populations:** The proposed framework is suitable for various survey datasets, as it can effectively handle symmetric, skewed, and moderately heterogeneous auxiliary variables across strata.
- **Efficiency improvements from hybrid design:** Combining harmonic, geometric, and fractional components allows the hybrid estimators to leverage multiple adjustment strategies simultaneously, improving efficiency and stability in complex data scenarios.
- **Optimal parameter calibration:** Unknown constants in the estimators are determined by minimizing the MSE, enabling optimal calibration and improved performance compared with fixed-parameter forms.
- **Applicability in practical surveys:** These estimators provide reliable alternatives to conventional stratified estimators and are especially useful in large-scale surveys where auxiliary information may not follow ideal distributional assumptions.

7.3. Survey-based applications

The proposed estimators are applicable in a variety of survey contexts where stratified random sampling is used and auxiliary information is available. Key application areas include the following:

- **Agriculture and crop assessment:** Estimating average crop yield, total production, or farm income using auxiliary data such as cultivated land, fertilizer use, or satellite-derived vegetation indices within strata.
- **Environmental monitoring and conservation:** Estimating forest biomass, carbon reserves, pollution levels, or biodiversity using auxiliary information like land cover, canopy density, or remote sensing data across ecological strata.
- **Health and epidemiological surveys:** Estimating disease prevalence, vaccination coverage, or nutritional outcomes using auxiliary variables such as household size, demographics, or healthcare access within strata.
- **Economic and labor analysis:** Estimating mean income, employment rates, or expenditure levels using auxiliary variables including administrative records, business size classifications, or energy consumption across strata.
- **Industrial quality and performance assessment:** Assessing average production, defect rates, or machines' performance using auxiliary process indicators collected from stratified inspection units.

In these applications, populations often exhibit substantial heterogeneity, making stratified random

sampling an effective design for improving precision. The proposed harmonic, geometric, and fractional power-based estimators perform particularly well in the presence of skewed and nonlinear auxiliary relationships, which are frequently encountered in real survey datasets.

8. Final remarks and future directions

This study introduced a new class of weighted hybrid estimators for estimating the finite population mean under a stratified systematic sampling design. Explicit expressions for the bias and MSE of the proposed estimators were obtained using first-order Taylor series approximations under stratified systematic sampling. The MSE was minimized to derive the optimal values of the associated weights. Both the theoretical efficiency analysis and the comprehensive simulation experiments demonstrated that the proposed estimators achieve substantial efficiency improvements over the traditional ratio-, product-, regression-, and exponential-type estimators. Among them, the weighted geometric-adjusted estimator consistently produced the smallest MSE across all population models, stratification schemes, and sampling fractions, and was therefore identified as the most efficient and stable estimator.

The practical measurement based on actual finite populations also justified the efficiency of the suggested methodology. The estimator based on stratified systematic sampling, denoted $\hat{T}_{H_{2st}}$, was observed to perform better with a wide range of population structures and with high-correlation conditions, which supports the fact that it can be practically used in the real-world survey situation. These results suggest that stratification, in conjunction with systematic sampling and hybrid auxiliary-based correction, will result in significant improvements in estimation quality. To conclude, all three proposed estimators exhibit strong performance under stratified systematic sampling. Among them, the estimator $\hat{T}_{H_{2st}}$ emerges as the most suitable and reliable choice for surveys involving complex population characteristics and auxiliary information. The proposed framework provides a flexible and efficient alternative to classical estimators, particularly in situations where the usual assumptions are not satisfied.

Future studies can investigate higher-order Taylor approximations and robust versions of the suggested hybrid estimators for populations with severe skewness or influential observations. For practical surveys, the methodology can also be modified to account for small or extremely unequal strata sizes, providing dependable results. While maintaining design reliability, effectiveness can be further increased by combining the suggested framework with calibration or model-assisted estimation techniques. It would also be helpful to analyze the coverage properties of confidence intervals and derive consistent variance estimators. There are further opportunities to increase the applicability of the suggested methodology in complex survey settings by examining the sensitivity of the estimators to the number of strata and the choice of stratification variables, as well as by extending the approach to two-phase stratified systematic sampling or populations with spatial or chronological correlation.

Author contributions

Fatimah A. Almulhim: Conceptualization, methodology, formal analysis, investigation, validation, software, writing original draft, project administration and funding acquisition; Umer Daraz: Conceptualization, methodology, formal analysis, investigation, validation, software, data curation, resource, writing original draft, writing-review & editing, Supervision; Hassan M. Aljohani:

Methodology, formal analysis, investigation, validation, software, data curation, resource, writing-review & editing; Javid Shabbir: Data curation, resource, writing-review & editing. All authors have read and approved the final version of the manuscript.

Use of Generative-AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

The authors would like to thank Princess Nourah bint Abdulrahman University for supporting this work.

Conflict of interest

The authors have no competing interests.

Funding

This work was supported by the Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2026R515), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

References

1. W. G. Cochran, Relative efficiency of systematic and stratified random samples for a certain class of population, *Ann. Math. Stat.*, **17** (1946), 164–177. <https://doi.org/10.1214/aoms/1177730978>
2. W. G. Cochran, *Sampling techniques*, 3Eds., New York: Wiley, 1977.
3. L. Kish, *Survey sampling*, Wiley, New York, 1965.
4. F. Yates, Systematic sampling, *Philos. T. Roy. Soc. A*, **241** (1948), 345–377. <https://doi.org/10.1098/rsta.1948.0023>
5. W. Gautschi, Some remarks on systematic sampling, *Ann. Math. Stat.*, **28** (1957), 385–394. <https://doi.org/10.1214/aoms/1177706966>
6. J. Hájek, Optimum strategy and other problems in probability sampling, *Casopis Pěstov. Mat.*, **84** (1959), 387–423. <https://doi.org/10.21136/CPM.1959.117317>
7. W. G. Madow, On the theory of systematic sampling, II, *Ann. Math. Stat.*, **19** (1948), 333–354. <https://doi.org/10.1214/aoms/1177729988>
8. M. H. Hansen, W. N. Hurwitz, On the theory of sampling from finite populations, *Ann. Math. Stat.*, **14** (1943), 333–362. <https://doi.org/10.1214/aoms/1177731356>
9. A. Swain, The use of systematic sampling in ratio estimate, *J. Indian Stat. Assoc.*, **2** (1964), 160–164.

10. M. N. Murthy, *Sampling theory and methods*, Calcutta: Statistical Publishing Society, 1967.
11. N. Shukla, *Systematic sampling and product method of estimation*, In: Proceedings of the All India Seminar on Demography and Statistics, BHU, Varanasi, India, 1971.
12. D. Lahiri, *On the question of bias of systematic sampling*, Proceedings of world population conference, **6** (1954), 349–362.
13. K. Kushwaha, H. Singh, Class of almost unbiased ratio and product estimators in systematic sampling, *J. Indian Soc. Agric. Stat.*, **41** (1989), 193–205.
14. S. Kushwaha, S. Kushwaha, K. Kushwaha, A class of ratio, product and difference (RPD) estimators in systematic sampling, *Microelectron. Reliab.*, **33** (1993), 455–457. [https://doi.org/10.1016/0026-2714\(93\)90308-L](https://doi.org/10.1016/0026-2714(93)90308-L)
15. H. P. Singh, R. Tailor, N. K. Jatwa, Modified ratio and product estimators for population mean in systematic sampling, *J. Mod. Appl. Stat. Meth.*, **10** (2011), 424–435. <https://doi.org/10.22237/jmasm/1320120180>
16. H. P. Singh, N. K. Jatwa, A class of exponential type estimators in systematic sampling, *Econ. Qual. Control*, **27** (2013), 195–208. <https://doi.org/10.1515/eqc-2013-0005>
17. M. Khan, R. Singh, Estimation of population mean in chain ratio-type estimator under systematic sampling, *J. Probab. Stat.*, **2015** (2015), 248374. <https://doi.org/10.1155/2015/248374>
18. M. Noor-ul Amin, A. Javaid, M. Hanif, Estimation of population mean in systematic random sampling using auxiliary information, *J. Stat. Manag. Syst.*, **20** (2017), 1095–1106. <https://doi.org/10.1080/09720510.2017.1364500>
19. A. Javaid, M. Noor-ul Amin, M. Hanif, Modified ratio estimator in systematic random sampling under non-response, *P. Natl. A. Sci. India A*, **89** (2019), 817–825. <https://doi.org/10.1007/s40010-018-0509-3>
20. R. Mishra, R. Singh, N. K. Adichwal, A novel ratio cum product type exponential class of estimators of finite population mean in adaptive cluster sampling, *Braz. J. Biometr.*, **43** (2025). <https://doi.org/10.28951/bjb.v43i1.745>
21. M. Subzar, T. Alqurashi, D. Chandawat, S. Tamboli, T. A. Raja, A. K. Attri, et al., Generalized robust regression techniques and adaptive cluster sampling for efficient estimation of population mean in case of rare and clustered populations, *Sci. Rep.*, **15** (2025), 2069. <https://doi.org/10.1038/s41598-025-85328-0>
22. M. Nagy, M. N. Qureshi, N. Shaheen, M. Hanif, Mean estimation using memory-type estimators in systematic sampling for time-scaled surveys, *Mathematics*, **14** (2026), 1180. <https://doi.org/10.3390/math14071180>
23. M. Khan, J. Shabbir, Some improved ratio, product, and regression estimators of finite population mean when using minimum and maximum values, *Sci. World J.*, **2013** (2013), 158795. <https://doi.org/10.1155/2013/431868>
24. M. N. Qureshi, S. Khalil, M. Hanif, Generalized semi exponential type estimator under systematic sampling, *J. Stat. Theory Appl.*, **17** (2018), 283–290. <https://doi.org/10.2991/jsta.2018.17.2.8>

25. A. Iftikhar, H. Shi, S. Hussain, A. Qayyum, M. El-Morshedy, S. Al-Marzouki, Estimation of finite population mean in presence of maximum and minimum values under systematic sampling scheme, *AIMS Math.*, **7** (2022), 9825–9834. <https://doi.org/10.3934/math.2022547>
26. M. El-Morshedy, S. Hussain, K. Ullah, A. Khalil, J. Shabbir, W. Mansoor, Finite population mean estimation under systematic sampling scheme in presence of maximum and minimum values using two auxiliary variables, *Math. Probl. Eng.*, **2022** (2022), 2703178. <https://doi.org/10.1155/2022/2703178>
27. E. G. Koçyiğit, New memory-type estimators for systematic sampling, *J. Adv. Res. Nat. Appl. Sci.*, **11** (2025), 224–236. <https://doi.org/10.28979/jarnas.1671967>
28. A. Karim, H. Khan, Y. Mahmood, M. Riaz, S. Ahmad, On estimation and monitoring of population mean using systematic sampling under an exponentially weighted moving average scheme, *Pak. J. Stat. Oper. Res.*, 2024. <https://doi.org/10.18187/pjsor.v20i3.4429>
29. S. K. Pal, S. A. Mahmud, H. P. Singh, An efficient estimation of finite population mean through difference estimator in systematic sampling, *Afr. Mat.*, 2025. <https://doi.org/10.1007/s13370-024-01234-0>
30. S. Yang, D. Meng, H. Wang, C. Yang, A novel learning function for adaptive surrogate-model-based reliability evaluation, *Phil. T. Roy. Soc. A*, **382** (2024), 2264. <https://doi.org/10.1098/rsta.2022.0395>
31. Bureau of Statistics, *Punjab Development Statistics*, Government of the Punjab, Lahore, Pakistan, 2013.
32. Bureau of Statistics, *Punjab Development Statistics*, Government of the Punjab, Lahore, Pakistan, 2014.

Appendix A

A proof of Theorem 4.1, as detailed in Section 4

Proof. Let $\hat{T}_{H_{1st}}$ be the first proposed estimator

$$\hat{T}_{H_{1st}} = \sum_{h=1}^L W_h \left[P_{1h} \bar{t}_h + Q_{1h} \left(\frac{2\bar{V}_h \bar{v}_h}{\bar{V}_h + \bar{v}_h} \right) \right],$$

where $\bar{t}_h$ and $\bar{v}_h$ are the systematic sample means of the study and auxiliary variables in the h stratum, $\bar{V}_h$ is the known population mean of the auxiliary variable, and P_{1h} , Q_{2h} are unknown constants.

We can rewrite the first suggested estimator in Eq (4.1) in terms of errors to obtain the equations of bias and MSEs as

$$\begin{aligned} \hat{T}_{H_{1st}} &= \sum_{h=1}^L W_h \left[P_{1h} (\bar{t}_h - \bar{T}_h + \bar{T}_h) + Q_{1h} \left\{ \frac{2\bar{V}_h (\bar{v}_h - \bar{V}_h + \bar{V}_h)}{\bar{V}_h + (\bar{v}_h - \bar{V}_h + \bar{V}_h)} \right\} \right], \\ \hat{T}_{H_{1st}} &= \sum_{h=1}^L W_h \left[P_{1h} \bar{T}_h \left(1 + \frac{\bar{t}_h - \bar{T}_h}{\bar{T}_h} \right) + Q_{1h} \left\{ \frac{2\bar{V}_h^2 \left(1 + \frac{\bar{v}_h - \bar{V}_h}{\bar{V}_h} \right)}{\bar{V}_h + \bar{V}_h \left(1 + \frac{\bar{v}_h - \bar{V}_h}{\bar{V}_h} \right)} \right\} \right], \end{aligned}$$

$$\begin{aligned}\hat{T}_{H_{1st}} &= \sum_{h=1}^L W_h \left[P_{1_h} \bar{T}_h (1 + \zeta_{0_h}) + Q_{1_h} \left\{ \frac{2\bar{V}_h (1 + \zeta_{1_h})}{1 + (1 + \zeta_{1_h})} \right\} \right], \\ \hat{T}_{H_{1st}} &= \sum_{h=1}^L W_h \left[P_{1_h} \bar{T}_h (1 + \zeta_{0_h}) + Q_{1_h} \bar{V}_h (1 + \zeta_{1_h}) \left(1 + \frac{\zeta_{1_h}}{2} \right)^{-1} \right],\end{aligned}\quad (8.1)$$

where

$$\zeta_{0_h} = \frac{\bar{t}_h - \bar{T}_h}{\bar{T}_h}, \quad \zeta_{1_h} = \frac{\bar{v}_h - \bar{V}_h}{\bar{V}_h}.$$

Assume that $|\zeta_{1_h}| < 1$, so that the term $(1 + \zeta_{1_h})^{-1}$ can be expanded. Expanding the right-hand side of Eq (8.1) and simplifying the expression by neglecting terms involving powers of the error terms higher than two, we obtain

$$\begin{aligned}\hat{T}_{H_{1st}} &= \sum_{h=1}^L W_h \left[P_{1_h} \bar{T}_h (1 + \zeta_{0_h}) + Q_{1_h} \bar{V}_h (1 + \zeta_{1_h}) \left(1 - \frac{\zeta_{1_h}}{2} + \frac{\zeta_{1_h}^2}{4} \right) \right], \\ \hat{T}_{H_{1st}} &= \sum_{h=1}^L W_h \left[P_{1_h} \bar{T}_h + Q_{1_h} \bar{T}_h \zeta_{0_h} + Q_{1_h} \bar{V}_h \left(1 - \frac{\zeta_{1_h}}{2} + \frac{\zeta_{1_h}^2}{4} + \zeta_{1_h} - \frac{\zeta_{1_h}^2}{2} \right) \right], \\ \hat{T}_{H_{1st}} &= \sum_{h=1}^L W_h \left[P_{1_h} \bar{T}_h + Q_{1_h} \bar{T}_h \zeta_{0_h} + Q_{1_h} \bar{V}_h \left(1 + \frac{\zeta_{1_h}}{2} - \frac{\zeta_{1_h}^2}{4} \right) \right].\end{aligned}$$

Subtract $\sum_{h=1}^L W_h \bar{T}_h$ on both sides and after the simplifications, we obtain

$$\hat{T}_{H_{1st}} - \sum_{h=1}^L W_h \bar{T}_h = \sum_{h=1}^L W_h \left[(P_{1_h} - 1) \bar{T}_h + Q_{1_h} \bar{V}_h + P_{1_h} \bar{T}_h \zeta_{0_h} + \frac{Q_{1_h} \bar{V}_h}{2} \zeta_{1_h} - \frac{Q_{1_h} \bar{V}_h}{4} \zeta_{1_h}^2 \right]. \quad (8.2)$$

Taking the expectation on both sides of Eq (8.2) and using

$$E(\zeta_{0_h}) = E(\zeta_{1_h}) = 0, \quad E(\zeta_{1_h}^2) = \theta_h D_{v_h} C_{v_h}^2,$$

we obtain

$$\text{Bias}(\hat{T}_{H_{1st}}) = \sum_{h=1}^L W_h \left[(P_{1_h} - 1) \bar{T}_h + Q_{1_h} \bar{V}_h - \frac{1}{4} Q_{1_h} \theta_h \bar{V}_h D_{v_h} C_{v_h}^2 \right]. \quad (8.3)$$

Squaring both sides of Eq (8.2) and ignoring all terms of total degree ≥ 3 , we have

$$\begin{aligned}\left(\hat{T}_{H_{1st}} - \sum_{h=1}^L W_h \bar{T}_h \right)^2 &= \sum_{h=1}^L W_h^2 \left[\left\{ (P_{1_h} - 1) \bar{T}_h + Q_{1_h} \bar{V}_h \right\}^2 + P_{1_h}^2 \bar{T}_h^2 \zeta_{0_h}^2 \right. \\ &\quad + \frac{Q_{1_h}^2 \bar{V}_h^2}{4} \zeta_{1_h}^2 + P_{1_h} Q_{1_h} \bar{T}_h \bar{V}_h \zeta_{0_h} \zeta_{1_h} \\ &\quad \left. - \frac{Q_{1_h} \bar{V}_h}{2} \left\{ (P_{1_h} - 1) \bar{T}_h + Q_{1_h} \bar{V}_h \right\} \zeta_{1_h}^2 \right].\end{aligned}$$

Taking expectation on both sides and substituting

$$E(\zeta_{0h}^2) = \theta_h D_{t_h} C_{t_h}^2, \quad E(\zeta_{1h}^2) = \theta_h D_{v_h} C_{v_h}^2, \quad E(\zeta_{0h} \zeta_{1h}) = \theta_h \sqrt{D_{tv_h}} \rho_{tv_h} C_{t_h} C_{v_h},$$

we derive the MSE expressions as follows:

$$\begin{aligned} \text{MSE}(\hat{T}_{H_{1st}}) \cong \sum_{h=1}^L W_h^2 & \left[\left\{ (P_{1h} - 1) \bar{T}_h + Q_{1h} \bar{V}_h \right\}^2 + \theta_h P_{1h}^2 \bar{T}_h^2 D_{t_h} C_{t_h}^2 + \frac{\theta_h Q_{1h}^2 \bar{V}_h^2}{4} D_{v_h} C_{v_h}^2 \right. \\ & + \theta_h P_{1h} Q_{1h} \bar{T}_h \bar{V}_h \sqrt{D_{tv_h}} \rho_{tv_h} C_{t_h} C_{v_h} \\ & \left. - \frac{\theta_h Q_{1h} \bar{V}_h}{2} \left\{ (P_{1h} - 1) \bar{T}_h + Q_{1h} \bar{V}_h \right\} D_{v_h} C_{v_h}^2 \right]. \end{aligned} \quad (8.4)$$

To obtain the minimum bias and MSE expressions, we take the partial derivatives of Eq (8.4) with respect to P_{1h} and Q_{1h} as follows.

The partial derivative with respect to (w.r.t.) P_{1h} is

$$\begin{aligned} \frac{\partial \text{MSE}}{\partial P_{1h}} = & \left[2 \bar{T}_h \left\{ (P_{1h} - 1) \bar{T}_h + Q_{1h} \bar{V}_h \right\} + 2 \theta_h P_{1h} \bar{T}_h^2 D_{t_h} C_{t_h}^2 + \theta_h Q_{1h} \bar{T}_h \bar{V}_h \sqrt{D_{tv_h}} \rho_{tv_h} C_{t_h} C_{v_h} \right. \\ & \left. - \frac{\theta_h Q_{1h} \bar{T}_h \bar{V}_h}{2} D_{v_h} C_{v_h}^2 \right] = 0. \end{aligned}$$

The partial derivative w.r.t. Q_{1h} is

$$\begin{aligned} \frac{\partial \text{MSE}}{\partial Q_{1h}} = & \left[2 \bar{V}_h \left\{ (P_{1h} - 1) \bar{T}_h + Q_{1h} \bar{V}_h \right\} + \frac{\theta_h Q_{1h} \bar{V}_h^2}{2} D_{v_h} C_{v_h}^2 + \theta_h P_{1h} \bar{T}_h \bar{V}_h \sqrt{D_{tv_h}} \rho_{tv_h} C_{t_h} C_{v_h} \right. \\ & \left. - \frac{\theta_h \bar{V}_h}{2} \left\{ (P_{1h} - 1) \bar{T}_h + 2 Q_{1h} \bar{V}_h \right\} D_{v_h} C_{v_h}^2 \right] = 0. \end{aligned}$$

Solving the the above system of equations, which gives us the optimal values of P_{1h} and Q_{1h} , we have

$$P_{1h}^* = \frac{(D_{v_h} C_{v_h}^2 - 2 \sqrt{D_{tv_h}} \rho_{tv_h} C_{t_h} C_{v_h}) (\theta_h D_{v_h} C_{v_h}^2 - 4)}{\Delta_{1h}}$$

and

$$Q_{1h}^* = \frac{4 \bar{T}_h}{\bar{V}_h} \left[\frac{D_{t_h} C_{t_h}^2 (\theta_h D_{v_h} C_{v_h}^2 - 4) + 2 R_h \sqrt{D_{tv_h}}}{\Delta_{1h}} \right],$$

where

$$\begin{aligned} \Delta_{1h} = & \theta_h \left[(D_{v_h} C_{v_h}^2 - 2 R_h \sqrt{D_{tv_h}})^2 + 4 D_{t_h} D_{v_h} C_{t_h}^2 C_{v_h}^2 \right] - 4 \left[D_{v_h} C_{v_h}^2 + 4 D_{t_h} C_{t_h}^2 - 4 R_h \sqrt{D_{tv_h}} \right], \\ & R_h = \rho_{tv_h} C_{t_h} C_{v_h}. \end{aligned}$$

Substituting P_{1h}^* and Q_{1h}^* into the bias and MSE expression given in Eqs (8.3) and (8.4), we prove the theorem as follows:

$$\text{Bias}(\hat{T}_{H_{1st}})_{opt} \cong - \sum_{h=1}^L W_h \theta_h \bar{T}_h \left[\frac{\theta_h D_{t_h} D_{v_h}^2 C_{t_h}^2 C_{v_h}^4 - 4 D_{t_h} D_{v_h} C_{t_h}^2 C_{v_h}^2 + 4 D_{tv_h} R_h^2}{\Delta_{1h}} \right]$$

and

$$\text{MSE}(\hat{T}_{H_{1st}})_{opt} \cong \sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[\frac{\theta_h D_{t_h} D_{v_h}^2 C_{t_h}^2 C_{v_h}^4 - 4 D_{t_h} D_{v_h} C_{t_h}^2 C_{v_h}^2 + 4 D_{tv_h} R_h^2}{\Delta_{1h}} \right],$$

where

$$\Delta_{1h} = \theta_h \left[(D_{v_h} C_{v_h}^2 - 2 R_h \sqrt{D_{tv_h}})^2 + 4 D_{t_h} D_{v_h} C_{t_h}^2 C_{v_h}^2 \right] - 4 \left[D_{v_h} C_{v_h}^2 + 4 D_{t_h} C_{t_h}^2 - 4 R_h \sqrt{D_{tv_h}} \right],$$

$$R_h = \rho_{tv_h} C_{t_h} C_{v_h}.$$

Appendix B

A proof of Theorem 4.1, as detailed in Section 4

Proof. Let the second proposed estimator be

$$\hat{T}_{H_{2st}} = \sum_{h=1}^L W_h \left[P_{2h} \bar{t}_h + Q_{2h} \sqrt{\bar{V}_h \cdot \bar{v}_h} \right],$$

where $\bar{t}_h$ and $\bar{v}_h$ are the systematic sample means of the study and auxiliary variables in the h stratum, $\bar{V}_h$ is the known h stratum population mean of the auxiliary variable, and P_{2h}, Q_{2h} are unknown constants.

To derive the equations of bias and MSEs, rewrite the first proposed estimator in Eq (4.4) in terms of the error

$$\hat{T}_{H_{2st}} = \sum_{h=1}^L W_h \left[P_{2h} \bar{T}_h (1 + \zeta_{0h}) + Q_{2h} \bar{V}_h \sqrt{1 + \zeta_{1h}} \right].$$

If $|\zeta_{1h}| < 1$, then the square root can be expanded using the binomial (Taylor) series as follows:

$$\sqrt{1 + \zeta_{1h}} = (1 + \zeta_{1h})^{1/2} = 1 + \frac{1}{2}\zeta_{1h} - \frac{1}{8}\zeta_{1h}^2 + \frac{1}{16}\zeta_{1h}^3 - \dots.$$

Keeping terms up to the second order, we have

$$\sqrt{1 + \zeta_{1h}} \approx 1 + \frac{1}{2}\zeta_{1h} - \frac{1}{8}\zeta_{1h}^2.$$

Hence, the estimator becomes

$$\hat{T}_{H_{2st}} - \sum_{h=1}^L W_h \bar{T}_h = \sum_{h=1}^L W_h \left[-\bar{T}_h + P_{2h} \bar{T}_h (1 + \zeta_{0h}) + Q_{2h} \bar{V}_h \left(1 + \frac{\zeta_{1h}}{2} - \frac{\zeta_{1h}^2}{8} \right) \right]. \quad (8.5)$$

Taking the expectation on both sides of Eq (8.5) and substituting the values of the error terms, the expression for the bias is obtained as:

$$\text{Bias}(\hat{T}_{H_{2st}}) \cong \sum_{h=1}^L W_h \left[(P_{2h} - 1) \bar{T}_h + Q_{2h} \bar{V}_h - \frac{\theta_h}{8} Q_{2h} \bar{V}_h D_{v_h} C_{v_h}^2 \right]. \quad (8.6)$$

Squaring both sides of Eq (8.5), taking the expectation, and ignoring terms of degree ≥ 3 , the MSE expression is given by:

$$\text{MSE}(\hat{T}_{H_{2st}}) \cong \sum_{h=1}^L W_h^2 \left[\left\{ (P_{2h} - 1)\bar{T}_h + Q_{2h}\bar{V}_h \right\}^2 + \theta_h P_{2h}^2 \bar{T}_h^2 D_{t_h} C_{t_h}^2 + \theta_h P_{2h} Q_{2h} \bar{T}_h \bar{V}_h R_h \sqrt{D_{tv_h}} \right. \\ \left. - \frac{\theta_h Q_{2h} \bar{V}_h}{4} \left\{ (P_{2h} - 1)\bar{T}_h + Q_{2h}\bar{V}_h \right\} D_{v_h} C_{v_h}^2 \right]. \quad (8.7)$$

Minimizing MSE Eq (8.7) by solving

$$\frac{\partial \text{MSE}}{\partial P_{2h}} = 0, \quad \frac{\partial \text{MSE}}{\partial Q_{2h}} = 0,$$

we get the optimal values

$$P_{2h}^* = \frac{\theta_h D_{v_h} C_{v_h}^2 (D_{v_h} C_{v_h}^2 - 4R_h \sqrt{D_{tv_h}}) + 32R_h \sqrt{D_{tv_h}}}{\theta_h \left\{ (D_{v_h} C_{v_h}^2 - 4R_h \sqrt{D_{tv_h}})^2 + 16D_{t_h} D_{v_h} C_{t_h}^2 C_{v_h}^2 \right\} - 64 (D_{t_h} C_{t_h}^2 - R_h \sqrt{D_{tv_h}})}$$

and

$$Q_{2h}^* = \frac{8\bar{T}_h}{\bar{V}_h} \left[\frac{D_{t_h} C_{t_h}^2 (\theta_h D_{v_h} C_{v_h}^2 - 8) + 4R_h \sqrt{D_{tv_h}}}{\theta_h \left\{ (D_{v_h} C_{v_h}^2 - 4R_h \sqrt{D_{tv_h}})^2 + 16D_{t_h} D_{v_h} C_{t_h}^2 C_{v_h}^2 \right\} - 64 (D_{t_h} C_{t_h}^2 - R_h \sqrt{D_{tv_h}})} \right].$$

Substituting P_{2h}^* and Q_{2h}^* into the bias and MSE expressions, we obtain

$$\text{Bias}(\hat{T}_{H_{2st}})_{\text{opt}} \cong - \sum_{h=1}^L W_h \theta_h \bar{T}_h \left[\frac{\theta_h D_{t_h} D_{v_h}^2 C_{t_h}^2 C_{v_h}^4 + 16D_{tv_h} R_h^2}{\Delta_{2h}} \right]$$

and

$$\text{MSE}(\hat{T}_{H_{2st}})_{\text{opt}} \cong \sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[\frac{\theta_h D_{t_h} D_{v_h}^2 C_{t_h}^2 C_{v_h}^4 + 16D_{tv_h} R_h^2}{\Delta_{2h}} \right],$$

where

$$\Delta_{2h} = \theta_h \left\{ (D_{v_h} C_{v_h}^2 - 4R_h \sqrt{D_{tv_h}})^2 + 16D_{t_h} D_{v_h} C_{t_h}^2 C_{v_h}^2 \right\} - 64 (D_{t_h} C_{t_h}^2 - R_h \sqrt{D_{tv_h}}).$$

Remark. The proposed estimator becomes more efficient when ρ_{tv_h} is positive and high, and the geometric adjustment $\sqrt{\bar{V}_h \bar{V}_h}$ improves stability by reducing the contribution of $C_{v_h}^2$ in the MSE.

Appendix C

A proof of Theorem 4.1, as detailed in Section 4

Proof. To derive the minimum bias and MSE forms of the model $\hat{T}_{H_{3st}}$, we convert the form of the equation 4.7 in terms of error as shown below:

$$\hat{T}_{H_{3st}} = \sum_{h=1}^L W_h \left[P_{3h} (1 + \zeta_{0h}) \bar{T}_h + P_{4h} (1 + \zeta_{1h}) \left(1 + \frac{\zeta_{1h}}{2} \right)^{-1} \bar{V}_h + P_{5h} (1 + \zeta_{1h})^{1/2} \right].$$

Expanding this expression with the binomial expansion and dropping terms of order ≥ 3 , we have:

$$\hat{T}_{H_{3st}} = \sum_{h=1}^L W_h \left[P_{3h}(1 + \zeta_{0h})\bar{T}_h + P_{4h}(1 + \zeta_{1h}) \left(1 - \frac{\zeta_{1h}}{2} + \frac{\zeta_{1h}^2}{4} \right) \bar{V}_h + P_{5h} \left(1 + \frac{\zeta_{1h}}{2} - \frac{\zeta_{1h}^2}{8} \right) \right].$$

Subtracting $\sum_{h=1}^L W_h \bar{T}_h$ from both sides and carrying out the necessary simplifications, we obtain

$$\begin{aligned} \hat{T}_{H_{3st}} - \sum_{h=1}^L W_h \bar{T}_h &= \sum_{h=1}^L W_h \left[-\bar{T}_h + P_{3h}\bar{T}_h + P_{3h}\bar{T}_h\zeta_{0h} + P_{4h}\bar{V}_h \left(1 + \frac{\zeta_{1h}}{2} - \frac{\zeta_{1h}^2}{4} \right) \right. \\ &\quad \left. + P_{5h} \left(1 + \frac{\zeta_{1h}}{2} - \frac{\zeta_{1h}^2}{8} \right) \right]. \end{aligned} \quad (8.8)$$

To obtain the bias of the estimator $\hat{T}_{H_{3st}}$ to the first order of approximation, take the expectation on both sides of Eq (8.8).

$$\text{Bias}(\hat{T}_{H_{3st}}) \cong \sum_{h=1}^L W_h \left[-\bar{T}_h + P_{3h}\bar{T}_h + P_{4h}\bar{V}_h \left(1 - \frac{\theta_h}{4} D_{vh} C_{vh}^2 \right) - \frac{\theta_h}{8} P_{5h} D_{vh} C_{vh}^2 \right]. \quad (8.9)$$

By squaring both sides of Eq (8.8) and taking expectations, we obtain the MSE of $\hat{T}_{H_{3st}}$ under a first-order approximation.

$$\begin{aligned} \text{MSE}(\hat{T}_{H_{3st}}) &\cong \sum_{h=1}^L W_h^2 \left\{ \left(-\bar{T}_h + P_{3h}\bar{T}_h \right)^2 + P_{3h}^2 \bar{T}_h^2 \theta_h D_{th} C_{th}^2 \right. \\ &\quad + P_{4h}^2 \bar{V}_h^2 \left(1 - \frac{\theta_h}{4} D_{vh} C_{vh}^2 \right) + P_{5h}^2 \\ &\quad + 2P_{4h}(P_{3h} - 1)\bar{T}_h \bar{V}_h \left(1 - \frac{\theta_h}{4} D_{vh} C_{vh}^2 \right) \\ &\quad + 2P_{5h}(P_{3h} - 1)\bar{T}_h \left(1 - \frac{\theta_h}{8} D_{vh} C_{vh}^2 \right) \\ &\quad + 2P_{4h}P_{5h}\bar{V}_h \left(1 - \frac{\theta_h}{8} D_{vh} C_{vh}^2 \right) \\ &\quad \left. + \bar{T}_h \left(P_{3h}P_{4h}\bar{V}_h + P_{3h}P_{5h} \right) \theta_h \sqrt{D_{tvh}} F_{tvh} C_{vh}^2 \right\}. \end{aligned} \quad (8.10)$$

To obtain the optimum values of the unknown constants P_{3h} , P_{4h} , and P_{5h} , the MSE given in Eq (8.10) is minimized with respect to these constants. Thus, the required first-order conditions are obtained by setting the corresponding partial derivatives equal to zero as follows:

$$\frac{\partial \text{MSE}}{\partial P_{3h}} = 0, \quad \frac{\partial \text{MSE}}{\partial P_{4h}} = 0, \quad \frac{\partial \text{MSE}}{\partial P_{5h}} = 0,$$

we obtain the following three equations:

$$\begin{aligned} 0 &= 2\bar{T}_h^2 P_{3h} - 2\bar{T}_h^2 + 2\bar{T}_h^2 P_{3h} \theta_h D_{th} C_{th}^2 \\ &\quad + 2\bar{T}_h \bar{V}_h P_{4h} \left(1 - \frac{\theta_h}{4} D_{vh} C_{vh}^2 + \frac{\theta_h}{2} \sqrt{D_{tvh}} F_{tvh} C_{vh}^2 \right) \\ &\quad + 2\bar{T}_h P_{5h} \left(1 - \frac{\theta_h}{8} D_{vh} C_{vh}^2 + \frac{\theta_h}{2} \sqrt{D_{tvh}} F_{tvh} C_{vh}^2 \right), \end{aligned} \quad (8.11)$$

$$\begin{aligned}
0 &= 2\bar{T}_h \bar{V}_h P_{3h} \left(1 - \frac{\theta_h}{4} D_{vh} C_{vh}^2 + \frac{\theta_h}{2} \sqrt{D_{tvh}} F_{tvh} C_{vh}^2 \right) \\
&\quad - 2\bar{T}_h \bar{V}_h + \frac{1}{2} \bar{T}_h \bar{V}_h \theta_h D_{vh} C_{vh}^2 \\
&\quad + 2\bar{V}_h^2 P_{4h} \left(1 - \frac{\theta_h}{4} D_{vh} C_{vh}^2 \right) \\
&\quad + 2\bar{V}_h P_{5h} \left(1 - \frac{\theta_h}{8} D_{vh} C_{vh}^2 \right),
\end{aligned} \tag{8.12}$$

$$\begin{aligned}
0 &= 2\bar{T}_h P_{3h} \left(1 - \frac{\theta_h}{8} D_{vh} C_{vh}^2 + \frac{\theta_h}{2} \sqrt{D_{tvh}} F_{tvh} C_{vh}^2 \right) \\
&\quad - 2\bar{T}_h + \frac{1}{4} \bar{T}_h \theta_h D_{vh} C_{vh}^2 \\
&\quad + 2\bar{V}_h P_{4h} \left(1 - \frac{\theta_h}{8} D_{vh} C_{vh}^2 \right) + 2P_{5h}.
\end{aligned} \tag{8.13}$$

Solving the three simultaneous equations above with respect to P_{3h} , P_{4h} , and P_{5h} yields the corresponding optimum values of the constants, which are given below:

$$P_{3h}^* = \frac{D_{vh} C_{vh}^2 - 2\sqrt{D_{tvh}} F_{tvh} C_{vh}^2}{4D_{th} C_{th}^2 + D_{vh} C_{vh}^2 - 4\sqrt{D_{tvh}} F_{tvh} C_{vh}^2},$$

$$P_{4h}^* = \frac{2\bar{T}_h \left[2D_{th} D_{vh} C_{th}^2 C_{vh}^2 + D_{vh} C_{vh}^2 \sqrt{D_{tvh}} F_{tvh} - 4D_{tvh} F_{tvh}^2 C_{vh}^4 \right]}{\bar{V}_h D_{vh} C_{vh}^2 \left[4D_{th} C_{th}^2 + D_{vh} C_{vh}^2 - 4\sqrt{D_{tvh}} F_{tvh} C_{vh}^2 \right]},$$

and

$$P_{5h}^* = \frac{4\bar{T}_h \sqrt{D_{tvh}} F_{tvh} C_{vh}^2 \left[-D_{vh} C_{vh}^2 + 2\sqrt{D_{tvh}} F_{tvh} C_{vh}^2 \right]}{D_{vh} C_{vh}^2 \left[4D_{th} C_{th}^2 + D_{vh} C_{vh}^2 - 4\sqrt{D_{tvh}} F_{tvh} C_{vh}^2 \right]}.$$

Now, we can substitute these optimal values P_{3h}^* , P_{4h}^* , and P_{5h}^* into bias and MSE equations and simplify them, the final results are given below:

$$\text{Bias}(\hat{T}_{H_{3st}})_{\text{opt}} \cong - \sum_{h=1}^L \theta_h W_h \bar{T}_h \left[\frac{D_{th} D_{vh} C_{th}^2 C_{vh}^2 - D_{tvh} F_{tvh}^2 C_{vh}^4}{\Omega_h} \right]$$

and

$$\text{MSE}(\hat{T}_{H_{3st}})_{\text{opt}} \cong \sum_{h=1}^L W_h^2 \theta_h \bar{T}_h^2 \left[\frac{D_{th} D_{vh} C_{th}^2 C_{vh}^2 - D_{tvh} F_{tvh}^2 C_{vh}^4}{\Omega_h} \right],$$

where

$$\Omega_h = 4D_{th} C_{th}^2 + D_{vh} C_{vh}^2 - 4\sqrt{D_{tvh}} F_{tvh} C_{vh}^2.$$



AIMS Press

© 2026 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>)