VFAST Transactions on Mathematics
http://vfast.org/journals/index.php/VTM@ 2024 ISSN(e): 2309-0022, ISSN(p): 2411-6343
Volume 12, Number 1, January-June, 2024 pp: 264-279

Efficient Class of Variance Estimators for Population by using Supplementary Information in Stratified Random Sampling

Mujeeb Hussain[1*], **Lakhkar Khan**[2], **Qamruz Zaman** [3], **Abdurrahman Sabir**[4]

[1*]Department of Statistics, Government College Peshawar, Pakistan; [2]Department of Statistics, Government Post Graduate College Mardan, Pakistan; [3,4]Department of Statistics, University of Peshawar,Pakistan

Keywords: *Stratified, Efficient, Auxiliary, Variance, Estimators*
Subject Classification: *62D05*

Journal Info:
Submitted:
February 15, 2024
Accepted:
April 16, 2024
Published:
May 24, 2024

Abstract This paper addresses an efficient class of variance estimators for population using stratified random sampling. The suggested class of estimators using supplementary information has been studied in different circumstances. The expressions of bias and mean square error (MSE) of the proposed estimators are derived up to the first degree of approximation. The theoretical comparison of the proposed and considered estimators is also discussed, which shows that the proposed estimators are more efficient than the existing estimators. Theoretical findings are validated by three different types of real data sets and simulation studies. The numerical results of the proposed and existing estimators are compared in terms of mean square error, percentage relative efficiency (PRE) and diagrams. It is observed that all the proposed estimators outperform the existing estimators. For instance, the traditional unbiased estimator Ozel et.al [6] and other existing estimators. Lastly, appropriate recommendations have been provided for researchers to use these suggested estimators to solve real-world issues.

***Correspondence Author Email Address:**
mujeebhussain.stat@gmail.com
DOI: 10.21015/vtm.v12i1.1794

1 Introduction

In survey sampling, stratified sampling scheme reduces the heterogeneity of a population and increases the precision of the parameter of a population. In stratified sampling, the population of heterogenous units are divided into sub-populations called strata, such that each stratum is internally homogenous and externally heterogenous. A sample is independently selected from each subgroup of the population with the help of a suitable sampling scheme, mostly simple random sampling. Stratified random sampling is very helpful in administrative convenience and reduction of survey costs. In stratified sampling, supplementary information that are often accessible, also play a vital role in increasing the efficiency of the estimators. The traditional ratio, product and regression estimators are usually used to estimate the population parameters i.e. mean, variance and proportion [1].

In numerous real-world situations, estimation of population variance plays a crucial role and by excluding such information, the estimates of the parameters are very misleading. Consequently, the estimates will be overestimated or underestimated. In a variety of applications, variance estimation is important. Several examples are present i.e. blood pressure, temperature, customers' reaction regarding products, fuel prices, etc [2]. For the estimation of the variance of a finite population, many researchers used the sample variance in survey sampling. Powers [3] proposed variance estimators under different sampling techniques of the finite population.

In variance estimation case for a population using supplementary information under stratified sampling, insufficient consideration has been given in the literature. It is because of the traditional estimators of variance, that provide difficult and biased results. However, there are some authors who worked on the population variance estimation using stratified random sampling. In this connection, Kadilar and Cingi [4] suggested combine ratio estimator by utilizing the information of supplementary variable for increasing the precision of the estimator using stratified sampling. In addition, Sidelel et al. [5] prompted the study of Kadilar and Cingi [4], and proposed combine ratio estimator for population variance under stratified random sampling using two concomitant variables. Later Ozel et al. [6] and Cekim and Kadilar [7] proposed a separate biased ratio estimator for finite population variance. Similarly, Yasmeen et al. [8], Singh and Solanki [9], Qian [10], Cekim and Kadilar [11], Singh et al. [12], Mahanty and Mishra [13], Shahzad et al. [14], Singh et al. [15] and Ahmad et al. [16] also developed estimators for finite population variance.

Motivated by the aforesaid discussions, the current study suggests an efficient class of variance estimators for population using supplementary information under stratified random sampling. Different known parameters of the supplementary variable such as coefficient of variation, variance, standard deviation, mean and kurtosis have been used. The outcome of the proposed efficient class of variance estimators have been methodically investigated,i.e. theoretically numerically using real world data and simulation study.

The rest of the paper is organized as follows. Important notations that are used in this paper are introduced in Section 1.1. A review of existing estimators is provided in Section 2. In Section 3, the proposed class of efficient estimators is presented, while in Section 4, the proposed estimators are compared theoretically with the existing estimators and the efficiency of the suggested estimators is inspected. In the next Sections 5 and 6, the applications of proposed estimators are carried out using three different real-life data sets and simulation study respectively. The results and discussions are provided in Section 7, and in the last Section 8, the conclusion of the study is provided.

1.1 Notations

Let N observations of a finite population be divided into L strata. Suppose N_h denotes the size of the h^{th} stratum where $h = 1, 2, 3, \ldots, L$ and $\sum_{h=1}^{L} N_h = N$. A sample of size n_h is drawn randomly from the h^{th} stratum without replacement (WOR), such that $\sum_{h=1}^{L} n_h = n$. Let (y_{hi}, x_{hi}) denote the observations on the i^{th} unit of the h^{th} stratum of the study variable y and the supplementary variable x. Corresponding to the population means $(\bar{Y}_h, \bar{X}_h)$, let $(\bar{y}_h, \bar{x}_h)$ be the sample means of the h^{th} stratum. Let $\rho_{xh,yh}$ be the population correlation coefficient between the study variable and the auxiliary variable. Consider $W_h = \frac{N_h}{N}$ be the stratum weight and $f_h = \frac{n_h}{N_h}$ be the sample fraction. To obtain bias and MSE equations, we consider the following error terms:

$$s_{yh}^2 = S_{yh}^2(1 + e_{0h}),$$
$$s_{xh}^2 = S_{xh}^2(1 + e_{1h}),$$

where $E(e_{0h}) = E(e_{1h}) = 0$, and

$$E(e_{1h}^2) = \gamma_h(\lambda_{04h} - 1),$$
$$E(e_{0h}e_{1h}) = \gamma_h(\lambda_{22h} - 1),$$

$\gamma_h = \frac{1}{n_h} - \frac{1}{N_h}$ be the finite population correction. Also,

$$\lambda_{rsh} = \frac{\mu_{rsh}}{\mu_{20h}^{(r/2)}\mu_{h02h}^{(s/2)}},$$
$$\mu_{rsh} = \frac{\sum_{i=1}^{N_h}(Y_{hi} - \bar{Y}_h)^r (X_{hi} - \bar{X}_h)^s}{N_h - 1},$$

$$\lambda_{22h}^* = (\lambda_{22h} - 1),$$
$$\lambda_{04h}^* = (\lambda_{04h} - 1),$$
$$\lambda_{40h}^* = (\lambda_{40h} - 1),$$

where $\lambda_{40h} = \beta_2(y)_h$ and $\lambda_{04h} = \beta_2(x)_h$ be the coefficient of kurtosis for the population.

2 Literature Review

Ozel et al. [6] propose the following separate ratio variance estimators in stratified random sampling given as:

$$v_{c1(st)} = \sum_{h=1}^{L} W_h \left(\frac{s_{yh}^2}{s_{xh}^2}\right) S_{xh}^2 \quad (1)$$

$$v_{c2(st)} = \sum_{h=1}^{L} W_h \left(\frac{s_{yh}^2}{s_{xh}^2 + C_{xh}}\right) (S_{xh}^2 + C_{xh}) \quad (2)$$

$$v_{c3(st)} = \sum_{h=1}^{L} W_h \left(\frac{s_{yh}^2}{s_{xh}^2 + \beta_{xh}}\right) (S_{xh}^2 + \beta_{xh}) \quad (3)$$

$$v_{c4(st)} = \sum_{h=1}^{L} W_h \left(\frac{s_{yh}^2}{s_{xh}^2 C_{xh} + \beta_{xh}}\right) (S_{xh}^2 C_{xh} + \beta_{xh}) \quad (4)$$

$$v_{c5(st)} = \sum_{h=1}^{L} W_h \left(\frac{s_{yh}^2}{s_{xh}^2 \beta_{xh} + C_{xh}} \right) (S_{xh}^2 \beta_{xh} + C_{xh}) \tag{5}$$

The MSEs of the estimator v_{c1} is given as:

$$\text{MSE}(v_{c1(st)}) = \sum_{h=1}^{L} W_h^2 \gamma_h S_{yh}^4 (\beta_{yh} + \beta_{xh} - 2\lambda_{22h}) \tag{6}$$

The general MSE equation of the estimators v_{c2}, v_{c3}, v_{c4}, and v_{c5} are given as:

$$\text{MSE}(v_{cq(st)}) \approx \sum_{h=1}^{L} W_h^2 \gamma_h S_{yh}^4 (\lambda_{40h}^* - 2R_{qh} \lambda_{22h}^* + R_{qh}^2 \lambda_{04h}^*) \tag{7}$$

where $q = 2, 3, 4$ and 5

$$R_{2h} = \frac{S_{yh}^2}{S_{xh}^2 + C_{xh}},$$
$$R_{3h} = \frac{S_{yh}^2}{S_{xh}^2 + \beta_{xh}},$$
$$R_{4h} = \frac{S_{yh}^2 C_{xh}}{S_{xh}^2 C_{xh} + \beta_{xh}},$$
$$R_{5h} = \frac{S_{yh}^2 \beta_{xh}}{S_{xh}^2 \beta_{xh} + C_{xh}}.$$

Etebong

$$v_{Et(st)} = \sum_{h=1}^{L} W_h s_{yh}^2 \left[\tau_h - a_h \left(\frac{s_{xh}^2}{S_{xh}^2} \right)^{\eta_h} \exp \left(\frac{\omega_h (s_{xh}^2 - S_{xh}^2)}{(s_{xh}^2 - S_{xh}^2)} \right) \right] \tag{8}$$

where τ_h, a_h, and η_h are suitably chosen scalars such that a_h and τ_h satisfy the condition $\tau_h = 1 + \alpha_{(i)} - \infty \leq a_h \leq \infty$, also

$$a_h = \frac{2\lambda_{22h}^*}{(2\eta_k + \omega_h)\lambda_{04h}^*}, \quad \eta_h = \frac{(2\lambda_{22h}^* - \omega_h a_h \lambda_{04h}^*)}{(2a_h \lambda_{04h}^*)}, \quad \omega_h = \frac{(2\lambda_{22h}^* - \eta_h a_h \lambda_{04h}^*)}{(a_h \lambda_{04h}^*)}.$$

Etebong [17] suggested different estimators for population variance from the above equation (7), using $\tau_h = 1$, $\alpha_h = 0$, $\eta_k = \eta_k$, and $\omega_h = \omega_h$, and is given by:

$$v_{Et1(st)} = \sum_{h=1}^{L} W_h s_{yh}^2. \tag{9}$$

The MSE of $v_{Et1}(st)$ is given by:

$$\text{MSE}(v_{Et1(st)}) = \sum_{h=1}^{L} W_h^2 \gamma_h S_{yh}^4 \lambda_{40h}^*. \tag{10}$$

When $\tau_h = 0$, $\alpha_h = -1$, $\eta_k = 0$, and $\omega_h = -1$, the resulting family member of the suggested estimator is the Bahl and Tuteja [18] exponential ratio-type estimator given by:

$$v_{Et2(st)} = \sum_{h=1}^{L} W_h s_{yh}^2 \exp\left(\frac{S_{xh}^2 - s_{xh}^2}{S_{xh}^2 + s_{xh}^2}\right). \tag{11}$$

The MSE of the estimator v_{Et2} is given by:

$$\mathrm{MSE}(v_{Et2(st)}) = \sum_{h=1}^{L} W_h^2 \gamma_h S_{yh}^4 (\lambda_{40h}^* + \frac{1}{4}\lambda_{04h}^* - 4\lambda_{22h}^*). \tag{12}$$

When $\tau_h = 0$, $\alpha_h = -1$, $\eta_k = -2$, and $\omega_h = 0$, the resulting family member of the proposed estimator is the Kadilar and Cingi [19] chain ratio-type estimator of population variance in stratified random sampling given as:

$$v_{Et3(st)} = \sum_{h=1}^{L} W_h s_{hy}^2 \exp\left(\frac{S_{hx}^2}{s_{hx}^2}\right)^2. \tag{13}$$

The MSE of $v_{Et3(st)}$ is given by:

$$\mathrm{MSE}(v_{Et3(st)}) = \sum_{h=1}^{L} W_h^2 \gamma_h S_{hy}^4 (\lambda_{40h}^* + 4\lambda_{04h}^* - \lambda_{22h}^*). \tag{14}$$

3 Proposed Estimator

Motivated by [6], [17] and [19], an efficient class of variance estimators is proposed for a finite population under stratified random sampling, given as:

$$v_{Pi(st)} = \sum_{h=1}^{L} W_h s_{yh}^2 \left[k_{1h}\left(\frac{S_{xh}^2}{s_{xh}^2}\right) + k_{2h}\left(\frac{s_{xh}^2}{S_{xh}^2}\right)\right] \exp\left[\frac{d_1(M_h - N_h)}{d_1(M_h + N_h) + 2d_2}\right] \tag{15}$$

Where $i = 1, 2, \ldots, 5$ and k_{1h} and k_{2h} are unknown constants whose values are to be determined, such that the MSE is optimum, while d_1 and d_2 are quantities which take the values (-1, 0, 1) for designing different types of estimators.

Also $M_h = (E_{1h} + E_{2h})S_{xh}^2 + E_{3h}s_{xh}^2$ and $N_h = (E_{3h} + E_{2h})S_{xh}^2 + E_{1h}s_{xh}^2$ where E_{1h}, E_{2h}, and E_{3h} are different functions of parameters of the concomitant variable such as variances, mean, coefficient of variation, and kurtosis.

Rewriting equation (15) in terms of relative error and putting the values of M_h and N_h, we have:

$$v_{Pi(st)} = \sum_{h=1}^{L} W_h \left[S_{yh}^2(1 + e_{0h}) \left\{k_{1h}\left(\frac{S_{xh}^2}{S_{xh}^2(1 + e_{1h})}\right) + k_{2h}\left(\frac{S_{xh}^2(1 + e_{1h})}{S_{xh}^2}\right)\right\} \quad (16) \right. \\ \left. \times \exp\left[\frac{d_1\left((E_{1h}+E_{2h})S_{xh}^2 + E_{3h}s_{xh}^2 - (E_{3h}+E_{2h})S_{xh}^2 - s_{xh}^2 E_{1h}\right)}{d_1\left((E_{1h}+E_{2h})S_{xh}^2 + E_{3h}s_{xh}^2 + (E_{3h}+E_{2h})S_{xh}^2 + s_{xh}^2 E_{1h}\right) + 2d_2}\right]\right] \tag{6}$$

$$v_{Pi(st)} \approx \sum_{h=1}^{L} W_h \left[S_{yh}^2(1 + e_{0h}) \left\{k_{1h}(1 - e_{1h} + e_{1h}^2) + k_{2h}(1 + e_{1h})\right\} \exp\left(\frac{\theta_{1h}e_{1h}}{2(1 + \theta_{2h}/2e_{1h})}\right)\right] \tag{17}$$

Where

$$\theta_{1h} = \frac{d_1(E_{3h}-E_{1h})S_{xh}^2}{d_1(E_{3h}+E_{2h}+E_{1h})S_{xh}^2+d_2}$$

$$\theta_{2h} = \frac{d_1(E_{3h}+E_{1h})S_{xh}^2}{d_1(E_{3h}+E_{2h}+E_{1h})S_{xh}^2+d_2}$$

Table A: Members of the suggested class of estimators. $A = W_h s_{yh}^2 \left[k_{1h}\left(\frac{S_{xh}^2}{s_{xh}^2}\right) + k_{2h}\left(\frac{s_{xh}^2}{S_{xh}^2}\right)\right]$

d_1	d_2	E_{1h}	E_{2h}	E_{3h}	Estimators
–1	0	S_{xh}	S_{xh}^2	β_{2xh}	$v_{P1}(st) = \sum_{h=1}^{L} A \exp\left[\frac{(s_{xh}^2-S_{xh}^2)(\beta_{2xh}-S_{xh})}{(s_{xh}^2+S_{xh}^2)(\beta_{2xh}+S_{xh})+2S_{xh}^4}\right]$
1	–1	S_{xh}	β_{2_xh}	$\bar{X}_h$	$v_{P2}(st) = \sum_{h=1}^{L} A \exp\left[\frac{(s_{xh}^2-S_{xh}^2)(\bar{X}_h-S_{xh})}{[(s_{xh}^2+S_{xh}^2)(\bar{X}_h+S_{xh})+2\beta_2(x)hS_{xh}^2]-2}\right]$
1	–1	C_{xh}	β_{2xh}	C_{xh}	$v_{P3}(st) = \sum_{h=1}^{L} A \exp\left[\frac{(s_{xh}^2-S_{xh}^2)(C_{xh}-S_{xh})}{-2[C_{xh}(s_{xh}^2+S_{xh}^2)+\beta_{2xh}+1]}\right]$
1	–1	C_{xh}	S_{xh}	$\beta_2(x)h$	$v_{P4}(st) = \sum_{h=1}^{L} A \exp\left[\frac{(s_{xh}^2-S_{xh}^2)(\beta_2(x)h-C_{xh})}{[(s_{xh}^2+S_{xh}^2)(\beta_2(x)h+C_{xh})+2S_{xh}^3]-2}\right]$
1	0	C_{xh}	$\bar{X}_h$	S_{xh}	$v_{P5}(st) = \sum_{h=1}^{L} A \exp\left[\frac{(s_{xh}^2-S_{xh}^2)(S_{xh}-C_{xh})}{(s_{xh}^2+S_{xh}^2)(S_{xh}+C_{xh})+2\bar{X}_hS_{xh}^2}\right]$

$$v_{Pi(st)} = \sum_{h=1}^{L} W_h \left[S_{yh}^2(1+e_{0h})\left\{k_{1h}(1-e_{1h}+(e_{1h})^2)+k_{2h}(1+e_{1h})\right\}\exp\left\{\frac{\theta_{1h}e_{1h}}{2} - \frac{\theta_{1h}\theta_{2h}(e_{1h})^2}{4}\right\}\right] \tag{20}$$

$$v_{Pi(st)} = \sum_{h=1}^{L}\left[W_hS_{yh}^2\left\{\left[k_{1h}(1+e_{0h})(1-e_{1h}+(e_{1h})^2)+k_{2h}(1+e_{0h})(1+e_{1h})\right]\right.\right.$$
$$\left.\left.\times\left[1+\frac{\theta_{1h}}{2}e_{1h} - \frac{\theta_{1h}\theta_{2h}}{4}(e_{1h})^2 + \frac{(\theta_{1h})^2}{8}(e_{1h})^2\right]\right\}\right] \tag{21}$$

$$(v_{Pi(st)} - S_{yh}^2) = \sum_{h=1}^{L} W_hS_{yh}^2\left[\left[k_{1h}(1+e_{0h}+(\theta_{1h}/2-1)(e_{1h}+e_{1h}e_{1h}) + (1-\theta_{1h}/2-(\theta_{1h}\theta_{2h})/4+(\theta_{1h}^2)/8)e_{1h}^2)\right.\right.$$
$$\left.\left.+k_{2h}(1+e_{0h}+(\theta_{1h}/2+1)(e_{1h}+e_{0h}e_{1h})+\theta_{1h}/2(1-\theta_{2h}/2+\theta_{1h}/4)(e_{1h})^2)\right]-1\right] \tag{22}$$

$$\text{Bias}(v_{Pi(st)}) = \sum_{h=1}^{L} W_hS_{yh}^2\left[\left[k_{1h}\left(1+(\theta_{1h}/2-1)\gamma_h\lambda_{22h}^* + (1-\theta_{1h}/2-(\theta_{1h}\theta_{2h})/2+((\theta_{1h})^2)/8)\gamma_h\lambda_{04h}^*\right)\right.\right.$$
$$\left.\left.+k_2h\left(1+(\theta_{1h}/2+1)\gamma_h\lambda_{22h}^* + \theta_{1h}/2(1-\theta_{2h}+\theta_{1h}/4)\gamma_h\lambda_{04h}^*\right)\right]-1\right] \tag{23}$$

$$\text{Bias}(v_{Pi(st)}) = \sum_{h=1}^{L} W_h\frac{S_{yh}^2}{n_h}\left[\left[k_{1h}\left(n_h+(\theta_{1h}/2-1)\lambda_{22h}^* + (1-\theta_{1h}/2-(\theta_{1h}\theta_{2h})/2+((\theta_{1h})^2)/8)\lambda_{04h}^*\right)\right.\right.$$
$$\left.\left.+k_{2h}\left(n_h+(\theta_{1h}/2+1)\lambda_{22h}^* + \theta_{1h}/2(1-\theta_{2h}+\theta_{1h}/4)\lambda_{04h}^*\right)\right]-n_h\right] \tag{24}$$

To find MSE, taking eq: (22) and squaring on both sides, we get

$$
\begin{aligned}
&(v_{Pi(st)} - S_{yh}^2)^2 = \\
&\sum_{h=1}^{L} W_h^2 S_{yh}^4 \Bigg[\left(k_{1h}(1 + e_{0h} + (\theta_{1h}/2 - 1)(e_{1h} + e_{0h}e_{1h}) + (1 - \theta_{1h}/2 - (\theta_{1h}\theta_{2h})/4 + (\theta_{1h}^2)/8)e_{1h}^2)\right)^2 \\
&\quad + \left(k_{2h}(1 + e_{0h} + (\theta_{1h}/2 + 1)(e_{1h} + e_{0h}e_{1h}) + \theta_{1h}/2(1 - \theta_{2h}/2 + \theta_{1h}/4)e_{1h}^2)\right)^2 + (-1)^2 \\
&\quad +2k_{1h}k_{2h} \left(1 + e_{0h} + (\theta_{1h}/2 - 1)(e_{1h} + e_{0h}e_{1h}) + (1 - \theta_{1h}/2 - (\theta_{1h}\theta_{2h})/4 + (\theta_{1h}^2)/8)e_{1h}^2\right) \\
&\quad \times \left(1 + e_{0h} + (\theta_{1h}/2 + 1)(e_{1h} + e_{0h}e_{1h}) + \theta_{1h}/2(1 - \theta_{2h}/2 + \theta_{1h}/4)e_{1h}^2\right) \\
&\quad -2 \left(k_{1h}(1 + e_{0h} + (\theta_{1h}/2 - 1)(e_{1h} + e_{0h}e_{1h}) + (1 - \theta_{1h}/2 - (\theta_{1h}\theta_{2h})/4 + (\theta_{1h}^2)/8)e_{1h}^2)\right) \\
&\quad -2 \left(k_{2h}(1 + e_{0h} + (\theta_{1h}/2 + 1)(e_{1h} + e_{0h}e_{1h}) + \theta_{1h}/2(1 - \theta_{2h}/2 + \theta_{1h}/4)e_{1h}^2)\right) \Bigg]
\end{aligned}
\tag{25}
$$

Solving the above equation, we have

$$
\begin{aligned}
\text{MSE}(v_{Pi(st)}) &= \sum_{h=1}^{L} W_h^2 \frac{S_{yh}^4}{n_h} \Bigg[n_h + k_{1h}^2 \left(n_h + \lambda_{40}^* + \left\{(\theta_{1h}/2 - 1)^2 + 2\left(1 - \theta_{1h}/2 - (\theta_{1h}\theta_{2h})/4 + (\theta_{1h}^2)/8\right)\lambda_{04}^* + 4(\theta_{1h}/2 - 1)\lambda_{22}^*\right\}\right) \\
&\quad +k_{2h}^2 \left(n_h + \lambda_{40}^* + \left\{(\theta_{1h}/2 + 1)^2 + \theta_{1h}\left(1 - \theta_{2h}/2 + \theta_{1h}/4\right)\lambda_{04}^* + 4(\theta_{1h}/2 + 1)\lambda_{22}^*\right\}\right) \\
&\quad +2k_{1h}k_{2h} \left\{ \begin{vmatrix} n_h + \lambda_{40}^* + \left\{2(\theta_{1h}/2 - 1) + 2(\theta_{1h}/2 + 1)\right\}\lambda_{22}^* \\ + \left\{(1 - \theta_{1h}/2 - (\theta_{1h}\theta_{2h})/4 + (\theta_{1h}^2)/8) + ((\theta_{1h}^2)/4 - 1) + \theta_{1h}/2\left(1 - \theta_{2h}/2 + \theta_{1h}/4\right)\right\}\lambda_{04}^* \end{vmatrix} \right\} \\
&\quad -2k_{1h}\left(n_h + (\theta_{1h}/2 - 1)\lambda_{22}^* + \left(1 - \theta_{1h}/2 - (\theta_{1h}\theta_{2h})/4 + (\theta_{1h}^2)/8\right)\lambda_{04}^*\right) \\
&\quad -2k_{2h}\left(n_h + (\theta_{1h}/2 + 1)\lambda_{22}^* + \theta_{1h}/2\left(1 - \theta_{2h}/2 + \theta_{1h}/4\right)\lambda_{04}^*\right) \Bigg]
\end{aligned}
\tag{26}
$$

Minimizing the MSE($v_{Pi}(st)$) in equation (26) to obtain the optimum values of k_{1h} and k_{2h} given as:

$$\frac{\partial \text{MSE}(v_{Pi}(st))}{\partial k_{1h}} = 0 \quad \text{and} \quad \frac{\partial \text{MSE}(v_{Pi}(st))}{\partial k_{2h}} = 0$$

We obtain the optimum values of k_{1h} and k_{2h} as

$$k_{1h}^* = \frac{(A_{2h}A_{4h} - A_{5h}A_{3h})}{(A_{1h}A_{2h} - A_{3h}^2)},$$

$$k_{2h}^* = \frac{(A_{1h}A_{5h} - A_{3h}A_{4h})}{(A_{1h}A_{2h} - A_{3h}^2)}$$

where

$$
\begin{aligned}
A_{1h} &= [n_h + \lambda_{40h}^* + \{(\theta_{1h}/2 - 1)^2 + 2(1 - \theta_{1h}/2 - (\theta_{1h}\theta_{2h})/4 + (\theta_{1h}^2)/8)\lambda_{04h}^* + 4(\theta_{1h}/2 - 1)\lambda_{22h}^*\}] \\
A_{2h} &= [n_h + \lambda_{40h}^* + \{(\theta_{1h}/2 + 1)^2 + \theta_{1h}(1 - \theta_{2h}/2 + 1/4\theta_{1h})\lambda_{04h}^* + 4(\theta_{1h}/2 + 1)\lambda_{22h}^*\}] \\
A_{3h} &= [n_h + \lambda_{40h}^* + \{2(\theta_{1h}/2 - 1) + 2(\theta_{1h}/2 + 1)\}\lambda_{22h}^* + \\
&\qquad \{((1 - \theta_{1h}/2 - (\theta_{1h}\theta_{2h})/4 + (\theta_{1h}^2)/8) + ((\theta_{1h}^2)/4 - 1) + 1/2\theta_{1h}(1 - \theta_{2h}/2 + \theta_{1h}/4))\lambda_{04h}^*\}] \\
A_{4h} &= [n_h + (\theta_{1h}/2 - 1)\lambda_{22h}^* + (1 - \theta_{1h}/2 - (\theta_{1h}\theta_{2h})/4 + (\theta_{1h}^2)/8)\lambda_{04h}^*] \\
A_{5h} &= [n_h + (\theta_{1h}/2 + 1)\lambda_{22h}^* + \theta_{1h}/2(1 - \theta_{2h}/2 + \theta_{1h}/4)\lambda_{04h}^*]
\end{aligned}
$$

Replacing k_{1h} and k_{2h} with optimum values of k^*_{1h} and k^*_{2h} in equation (26), the optimum MSE of the suggested class of estimators $v_{pi}(st)$ is given as:

$$MSE(v_{Pi(st)})_{opt} = \sum_{h=1}^{L} W_h^2 \frac{S_{yh}^4}{n_h}(n_h + k^{*2}_{1h}A_{1h} + k^{*2}_{2h}A_{2h} + 2k^*_{1h}k^*_{2h}A_{3h} - 2k^*_{1h}A_{4h} - 2k^*_{2h}A_{5h}) \tag{27}$$

From Table A, it is observed that each combination of d_1, d_2, E_{1h}, E_{2h}, and E_{3h} are different. For this reason, the values of θ_{1h} and θ_{2h} in equation (18) and (19) are different from each other. Therefore, it is clear that optimum MSE values are different for each estimator $v_{p1}(st)$, $v_{p2}(st)$, $v_{p3}(st)$, $v_{p4}(st)$, and $v_{p5}(st)$, as A_{1h}, A_{2h}, A_{3h}, A_{4h}, and A_{5h} in equation (27) are computed using the values of θ_{1h} and θ_{2h}.

4 Theoretical Comparison

Condition (i): $MSE(v_{c1(st)}) - MSE(v_{Pi(st)})_{opt} > 0$ if

$$\sum_{h=1}^{L} \frac{W_h^2 S_{yh}^4}{n_h}(\beta_{yh} + \beta_{xh} - 2\lambda_{22h}) - \sum_{h=1}^{L} W_h^2 \frac{S_{yh}^4}{n_h}\left(n_h + k^{*2}_{1h}A_{1h} + k^{*2}_{2h}A_{2h} + 2k^*_{1h}k^*_{2h}A_{3h} - 2k^*_{1h}A_{4h} - 2k^*_{2h}A_{5h}\right) > 0$$

$$\sum_{h=1}^{L}\left[(\beta_{yh} + \beta_{xh} - 2\lambda_{22h}) - \left(n_h + k^{*2}_{1h}A_{1h} + k^{*2}_{2h}A_{2h} + 2k^*_{1h}k^*_{2h}A_{3h} - 2k^*_{1h}A_{4h} - 2k^*_{2h}A_{5h}\right)\right] > 0$$

Condition (ii): $MSE(v_{c2(st)}) - MSE(v_{Pi(st)})_{opt} > 0$ if

$$\sum_{h=1}^{L} \frac{W_h^2 S_{yh}^4}{n_h}(\lambda^*_{40h} - 2R_{h2}\lambda^*_{22h} + R^2_{h2}\lambda^*_{04h}) - \sum_{h=1}^{L} W_h^2 \frac{S_{yh}^4}{n_h}\left(n_h + k^{*2}_{1h}A_{1h} + k^{*2}_{2h}A_{2h} + 2k^*_{1h}k^*_{2h}A_{3h} - 2k^*_{1h}A_{4h} - 2k^*_{2h}A_{5h}\right) > 0$$

$$\sum_{h=1}^{L}\left[(\lambda^*_{40h} - 2R_{h2}\lambda^*_{22h} + R^2_{h2}\lambda^*_{04h}) - \left(n_h + k^{*2}_{1h}A_{1h} + k^{*2}_{2h}A_{2h} + 2k^*_{1h}k^*_{2h}A_{3h} - 2k^*_{1h}A_{4h} - 2k^*_{2h}A_{5h}\right)\right] > 0$$

Condition (iii): $MSE(v_{c3(st)}) - MSE(v_{Pi(st)})_{opt} > 0$ if

$$\sum_{h=1}^{L} \frac{W_h^2 S_{yh}^4}{n_h}(\lambda^*_{40h} - 2R_{h3}\lambda^*_{22h} + R^2_{h3}\lambda^*_{04h}) - \sum_{h=1}^{L} W_h^2 \frac{S_{yh}^4}{n_h}\left(n_h + k^{*2}_{1h}A_{1h} + k^{*2}_{2h}A_{2h} + 2k^*_{1h}k^*_{2h}A_{3h} - 2k^*_{1h}A_{4h} - 2k^*_{2h}A_{5h}\right) > 0$$

$$\sum_{h=1}^{L}\left[(\lambda^*_{40h} - 2R_{h3}\lambda^*_{22h} + R^2_{h3}\lambda^*_{04h}) - \left(n_h + k^{*2}_{1h}A_{1h} + k^{*2}_{2h}A_{2h} + 2k^*_{1h}k^*_{2h}A_{3h} - 2k^*_{1h}A_{4h} - 2k^*_{2h}A_{5h}\right)\right] > 0$$

Condition (iv): $MSE(v_{c4(st)}) - MSE(v_{Pi(st)})_{opt} > 0$ if

$$\sum_{h=1}^{L} \frac{W_h^2 S_{yh}^4}{n_h}(\lambda^*_{40h} - 2R_{h4}\lambda^*_{22h} + R^2_{h4}\lambda^*_{04h}) - \sum_{h=1}^{L} W_h^2 \frac{S_{yh}^4}{n_h}\left(n_h + k^{*2}_{1h}A_{1h} + k^{*2}_{2h}A_{2h} + 2k^*_{1h}k^*_{2h}A_{3h} - 2k^*_{1h}A_{4h} - 2k^*_{2h}A_{5h}\right) > 0$$

$$\sum_{h=1}^{L}\left[(\lambda^*_{40h} - 2R_{h4}\lambda^*_{22h} + R^2_{h4}\lambda^*_{04h}) - \left(n_h + k^{*2}_{1h}A_{1h} + k^{*2}_{2h}A_{2h} + 2k^*_{1h}k^*_{2h}A_{3h} - 2k^*_{1h}A_{4h} - 2k^*_{2h}A_{5h}\right)\right] > 0$$

Condition (v): $\mathrm{MSE}(v_{c5(st)}) - \mathrm{MSE}(v_{Pi(st)})_{\mathrm{opt}} > 0$ if

$$\sum_{h=1}^{L} \frac{W_h^2 S_{yh}^4}{n_h}(\lambda_{40h}^* - 2R_{h5}\lambda_{22h}^* + R_{h4}^2\lambda_{04h}^*) - \sum_{h=1}^{L} W_h^2 \frac{S_{yh}^4}{n_h}\left(n_h + k_{1h}^{*2}A_{1h} + k_{2h}^{*2}A_{2h} + 2k_{1h}^* k_{2h}^* A_{3h} - 2k_{1h}^* A_{4h} - 2k_{2h}^* A_{5h}\right) > 0$$

$$\sum_{h=1}^{L}\left[(\lambda_{40h}^* - 2R_{h5}\lambda_{22h}^* + R_{h4}^2\lambda_{04h}^*) - \left(n_h + k_{1h}^{*2}A_{1h} + k_{2h}^{*2}A_{2h} + 2k_{1h}^* k_{2h}^* A_{3h} - 2k_{1h}^* A_{4h} - 2k_{2h}^* A_{5h}\right)\right] > 0$$

Condition (vi): $\mathrm{MSE}(v_{Et1(st)}) - \mathrm{MSE}(v_{Pi(st)})_{\mathrm{opt}} > 0$ if

$$\sum_{h=1}^{L} \frac{W_h^2 S_{hy}^4 \lambda_{40h}^*}{n_h} - \sum_{h=1}^{L} W_h^2 \frac{S_{yh}^4}{n_h}\left(n_h + k_{1h}^{*2}A_{1h} + k_{2h}^{*2}A_{2h} + 2k_{1h}^* k_{2h}^* A_{3h} - 2k_{1h}^* A_{4h} - 2k_{2h}^* A_{5h}\right) > 0$$

$$\sum_{h=1}^{L}\left[\lambda_{40h}^* - \left(n_h + k_{1h}^{*2}A_{1h} + k_{2h}^{*2}A_{2h} + 2k_{1h}^* k_{2h}^* A_{3h} - 2k_{1h}^* A_{4h} - 2k_{2h}^* A_{5h}\right)\right] > 0$$

Condition (vii): $\mathrm{MSE}(v_{Et2(st)}) - \mathrm{MSE}(v_{Pi(st)})_{\mathrm{opt}} > 0$ if

$$\sum_{h=1}^{L} \frac{W_h^2 S_{hy}^4}{n_h}(\lambda_{40h}^* + \frac{1}{4}\lambda_{04h}^* - 4\lambda_{22h}^*) - \sum_{h=1}^{L} W_h^2 \frac{S_{yh}^4}{n_h}\left(n_h + k_{1h}^{*2}A_{1h} + k_{2h}^{*2}A_{2h} + 2k_{1h}^* k_{2h}^* A_{3h} - 2k_{1h}^* A_{4h} - 2k_{2h}^* A_{5h}\right) > 0$$

$$\sum_{h=1}^{L}\left[(\lambda_{40h}^* + \frac{1}{4}\lambda_{04h}^* - 4\lambda_{22h}^*) - \left(n_h + k_{1h}^{*2}A_{1h} + k_{2h}^{*2}A_{2h} + 2k_{1h}^* k_{2h}^* A_{3h} - 2k_{1h}^* A_{4h} - 2k_{2h}^* A_{5h}\right)\right] > 0$$

Condition (viii): $\mathrm{MSE}(v_{Et3(st)}) - \mathrm{MSE}(v_{Pi(st)})_{\mathrm{opt}} > 0$ if

$$\sum_{h=1}^{L} \frac{W_h^2 S_{hy}^4}{n_h}(\lambda_{40h}^* + 4\lambda_{04h}^* - \lambda_{22h}^*) - \sum_{h=1}^{L} W_h^2 \frac{S_{yh}^4}{n_h}\left(n_h + k_{1h}^{*2}A_{1h} + k_{2h}^{*2}A_{2h} + 2k_{1h}^* k_{2h}^* A_{3h} - 2k_{1h}^* A_{4h} - 2k_{2h}^* A_{5h}\right) > 0$$

$$\sum_{h=1}^{L}\left[(\lambda_{40h}^* + 4\lambda_{04h}^* - \lambda_{22h}^*) - \left(n_h + k_{1h}^{*2}A_{1h} + k_{2h}^{*2}A_{2h} + 2k_{1h}^* k_{2h}^* A_{3h} - 2k_{1h}^* A_{4h} - 2k_{2h}^* A_{5h}\right)\right] > 0$$

5 Numerical Investigation

To compare the efficiencies of traditional and suggested estimators in stratified sampling, we use the following data.

Data 1: [The data is publicly available in Etebong [17]

$$N = 92, \quad N_1 = 40, \quad N_2 = 20, \quad N_3 = 32, \quad n = 32,$$
$$n_1 = 14, \quad n_2 = 8, \quad n_3 = 10,$$
$$\bar{X}_1 = 24.62, \quad \bar{X}_2 = 48.63, \quad \bar{X}_3 = 32.85,$$
$$S_{y1}^2 = 125.26, \quad S_{y2}^2 = 82.08, \quad S_{y3}^2 = 78.480,$$
$$S_{x1}^2 = 90.25, \quad S_{x2}^2 = 64.83, \quad S_{x3}^2 = 105.38,$$
$$\lambda_{40(1)} = 21.28, \quad \lambda_{40(2)} = 35.32, \quad \lambda_{40(3)} = 18.450,$$
$$\lambda_{04(1)} = 10.25, \quad \lambda_{04(2)} = 16.82, \quad \lambda_{04(3)} = 22.42,$$
$$\lambda_{22(1)} = 8.86, \quad \lambda_{22(2)} = 18.56, \quad \lambda_{22(3)} = 16.200.$$

Data 2: [The data is publicly available in Singh and Mangat [20]] y shows quantity of juice and x denotes canes' weight.

$$N = 25,\quad N_1 = 6,\quad N_2 = 12,\quad N_3 = 7,\quad n = 10,$$
$$n_1 = 3,\quad n_2 = 4,\quad n_3 = 3,$$
$$C_{x1} = 0.1418,\quad C_{x2} = 0.1395,\quad C_{x3} = 0.1695,$$
$$S_{y1} = 8.1649,\quad S_{y2} = 14.409,\quad S_{y3} = 10.1519,$$
$$S_{x1} = 53.208,\quad S_{x2} = 43.371,\quad S_{x3} = 62.723,$$
$$\lambda_{40(1)} = 2.3437,\quad \lambda_{40(2)} = 3.7924,\quad \lambda_{40(3)} = 2.32942,$$
$$\lambda_{04(1)} = 2.2865,\quad \lambda_{04(2)} = 3.2689,\quad \lambda_{04(3)} = 3.1306,$$
$$\lambda_{22(1)} = 2.2641,\quad \lambda_{22(2)} = 3.3795,\quad \lambda_{22(3)} = 2.31177.$$

Data 3: [The data is publicly available in Singh and Mangat [17]]
y shows total number of milk cows in 1993 and x: total number of milk cows in 1990.

$$N = 24,\quad N_1 = 7,\quad N_2 = 12,\quad N_3 = 5,\quad n = 10,$$
$$n_1 = 3,\quad n_2 = 5,\quad n_3 = 2,$$
$$C_{x1} = 0.2991,\quad C_{x2} = 0.3186,\quad C_{x3} = 0.1837,$$
$$S_{y1} = 3.8861,\quad S_{y2} = 3.9042,\quad S_{y3} = 3.2619,$$
$$S_{x1} = 5.4695,\quad S_{x2} = 5.0338,\quad S_{x3} = 3.2710,$$
$$\lambda_{40(1)} = 1.6555,\quad \lambda_{40(2)} = 2.7509,\quad \lambda_{40(3)} = 1.5925,$$
$$\lambda_{04(1)} = 1.8497,\quad \lambda_{04(2)} = 2.3120,\quad \lambda_{04(3)} = 1.8333,$$
$$\lambda_{22(1)} = 1.1348,\quad \lambda_{22(2)} = 0.5695,\quad \lambda_{22(3)} = 1.3461.$$

Percentage Relative Efficiency (PRE):

$$\text{PRE} = \frac{\text{Var}(v_0(st))}{\text{MSE}(v_j) \text{ or } \text{MSE}(v_{Pj}(st))_{\text{opt}}} \times 100 \tag{28}$$

Where $j = c1, c2, c3, c4, c5, Et1, Et2,$ and $Et3$

6 Simulation Study

A simulation study is also conducted for the comparison of the suggested class of estimators and the estimators in the literature. The simulated dataset consists of information related to study variable y and supplementary variable x using the bivariate normal distribution $N(\mu, \Sigma)$, where μ denotes the mean vector while Σ showing a variance-covariance matrix. The dataset is divided into two strata using the following values of parameters:

$$\mu_1 = \begin{pmatrix} 18 \\ 18 \end{pmatrix} \quad \text{and} \quad \Sigma_1 = \begin{pmatrix} 3 & 2 \\ 2 & 2 \end{pmatrix}$$

$$\mu_2 = \begin{pmatrix} 22 \\ 22 \end{pmatrix} \quad \text{and} \quad \Sigma_2 = \begin{pmatrix} 4 & 2.5 \\ 2.5 & 3 \end{pmatrix}$$

Table 1. MSEs and PREs for suggested and existing estimators using three data sets

		Data 1		Data 2		Data 3	
Estimators		MSE	PRE	MSE	PRE	MSE	PRE
1	$v_{c1(st)}$	4155.058	100	993.0674	100	55.1434	100
2	$v_{c2(st)}$	4429.001	93.8148	6157.964	16.12655	26.13838	210.9675
3	$v_{c3(st)}$	4183.688	99.31569	6159.92	16.12143	25.47378	216.4715
4	$v_{c4(st)}$	4300.794	96.61142	6172.362	16.08894	24.4746	225.309
5	$v_{c5(st)}$	4441.126	93.55866	6157.903	16.12671	26.19847	210.4835
6	$v_{Et1(st)}$	6844.039	60.71061	7417.383	13.38838	25.9278	212.6808
7	$v_{Et2(st)}$	4393.278	94.57764	2621.339	37.88397	24.89248	221.5266
8	$v_{Et3(st)}$	20390	20.37792	27489.11	3.612585	119.8042	46.02798
9	$v_{P1(st)}$	1228.56	338.2056	792.0136	125.3851	13.25845	415.9119
10	$v_{P2(st)}$	676.4969	614.2022	772.1627	128.6086	10.70464	515.136
11	$v_{P3(st)}$	1309.135	317.3895	808.0062	122.9034	13.0568	422.3348
12	$v_{P4(st)}$	1869.649	222.2374	817.33	121.5014	13.866	397.686
13	$v_{P5(st)}$	1232.845	337.03	820.514	121.0299	13.8828	397.205

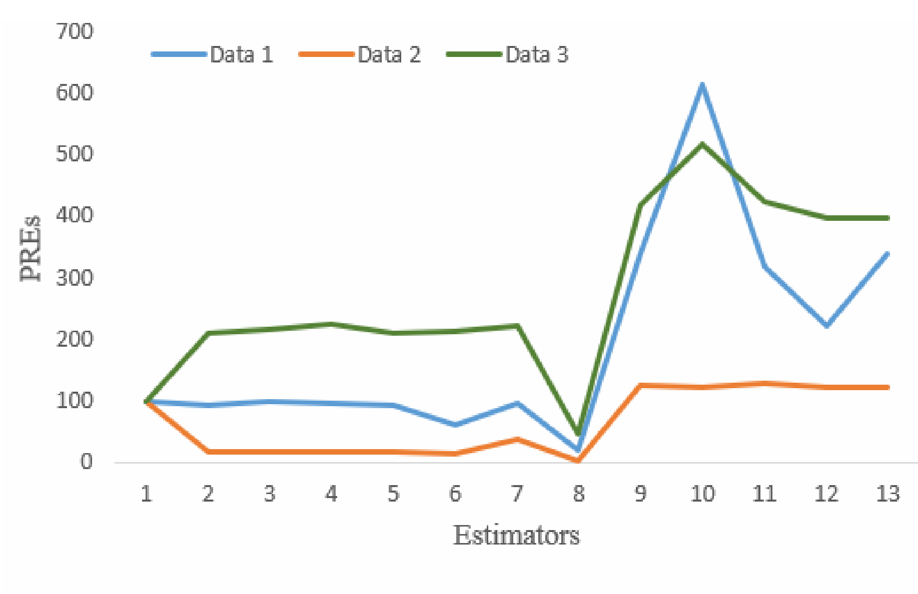

Figure 1. PREs of existing and proposed estimators for all three data sets

Table 2. Numerical Verification of Conditions (i)-(viii)

	Existing Estimators		Proposed Estimators				
Conditions		Data	$v_{P1(st)}$	$v_{P2(st)}$	$v_{P3(st)}$	$v_{P4(st)}$	$v_{P5(st)}$
		Data 1	2926.498	3478.561	2845.498	2285.498	2922.208
i	$v_{c1(st)}$	Data 2	201.0534	185.0614	220.9044	175.7374	172.5534
		Data 3	41.88497	44.43887	42.08667	41.27747	41.26067
		Data 1	3200.441	3752.504	3119.441	2559.441	3196.151
ii	$v_{c2(st)}$	Data 2	5365.95	5349.958	5385.801	5340.634	5337.45
		Data 3	12.87988	15.43378	13.08158	12.27238	12.25558
		Data 1	2955.128	3507.191	2874.128	2314.128	2950.838
iii	$v_{c3(st)}$	Data 2	5367.906	5351.914	5387.757	5342.59	5339.406
		Data 3	12.21528	14.76918	12.41698	11.60778	11.59098
		Data 1	3072.234	3624.297	2991.234	2431.234	3067.944
iv	$v_{c4(st)}$	Data 2	5380.348	5364.356	5400.199	5355.032	5351.848
		Data 3	11.2161	13.77	11.4178	10.6086	10.5918
		Data 1	3212.566	3764.629	3131.566	2571.566	3208.276
v	$v_{c4(st)}$	Data 2	5365.889	5349.897	5385.74	5340.573	5337.389
		Data 3	12.93997	15.49387	13.14167	12.33247	12.31567
		Data 1	5615.479	6167.542	5534.479	4974.479	5611.189
vi	$v_{Et1(st)}$	Data 2	6625.369	6609.377	6645.22	6600.053	6596.869
		Data 3	12.6693	15.2232	12.871	12.0618	12.045
		Data 1	3164.718	3716.781	3083.718	2523.718	3160.428
vii	$v_{Et2(st)}$	Data 2	1829.325	1813.333	1849.176	1804.009	1800.825
		Data 3	11.63398	14.18788	11.83568	11.02648	11.00968
		Data 1	19161.44	19713.5	19080.44	18520.44	19157.15
viii	$v_{Et3(st)}$	Data 2	26697.1	26681.1	26716.95	26671.78	26668.6
		Data 3	106.5457	109.0996	106.7474	105.9382	105.9214

Where μ_1 and Σ_1 corresponds to stratum 1 while μ_2 and Σ_2 are used for stratum 2. The population size is N = 450, which consists of N_1 = 250 and N_2 = 200 related to two strata. Using proportional allocation, samples of sizes 50, 80, and 100 are drawn from the population. MSEs and PREs are given in Table 3.

7 Results and Discussions

In this paper, an efficient class of variance estimators for population using supplementary information under stratified random sampling is proposed. To evaluate the performance of the suggested family of estimators, we used three real data sets and a simulation study. The criteria of MSE and PRE are used for comparison of different estimators. The MSEs and PREs of the suggested and existing estimators are provided in Table 2 for the real data sets and in Table 4, the values of the simulation study are given. MSEs and PREs of the proposed and existing estimators are also presented through different diagrams for both real data and simulation study. The results of the proposed family of estimators vary, based on different choices of d_1, d_2, E_{1h}, E_{2h}, and E_{3h}.

Some general findings are obtained as follows:

- Table 2 reveals that all the suggested estimators $v_{Pi(st)}$ have the optimum MSE than all the competitor estimators in the literature for all three data sets. Therefore, the recommended estimators $v_{Pi(st)}$ efficiently outperform all other considered estimators.
- The combinations of different choices of d_1, d_2, E_{1h}, E_{2h}, and E_{3h} impact the outcome of the suggested estimators.
- The proposed estimator $v_{P2(st)}$ shows the best performance and has the minimum MSE of 676.4969, 772.1627, and 10.70464 for all the three data sets when d_1 = 1, d_2 = –1, E_1 = S_{xh}, E_2 = β_{2xh}, and

Table 3. MSEs and PREs of different estimators using simulation study

		n=50		n=80		n=100	
Estimators		MSE	PRE	MSE	PRE	MSE	PRE
1	$v_{c1(st)}$	0.454859	100	0.236885	100	0.170496	100
2	$v_{c2(st)}$	0.434366	104.72	0.228477	103.68	0.16497	103.35
3	$v_{c3(st)}$	0.432782	105.1	0.225864	104.88	0.165724	102.88
4	$v_{c4(st)}$	0.434104	104.78	0.228038	103.88	0.16735	101.88
5	$v_{c5(st)}$	0.418812	108.6	0.232189	102.02	0.168789	101.01
6	$v_{Et1(st)}$	0.4247	107.1	0.228038	103.88	0.174724	97.58
7	$v_{Et2(st)}$	0.311383	146.08	0.172646	137.21	0.127961	133.24
8	$v_{Et3(st)}$	0.440276	103.31	0.23221	102.01	0.152392	111.88
9	$v_{P1(st)}$	0.270363	168.24	0.166805	142.01	0.120169	141.87
10	$v_{P2(st)}$	0.189669	239.82	0.102304	231.55	0.078857	216.2
11	$v_{P3(st)}$	0.265157	171.54	0.155832	152.01	0.115294	147.87
12	$v_{P4(st)}$	0.148848	305.58	0.082125	288.44	0.063681	267.73
13	$v_{P5(st)}$	0.12869	353.45	0.070421	336.38	0.054388	313.48

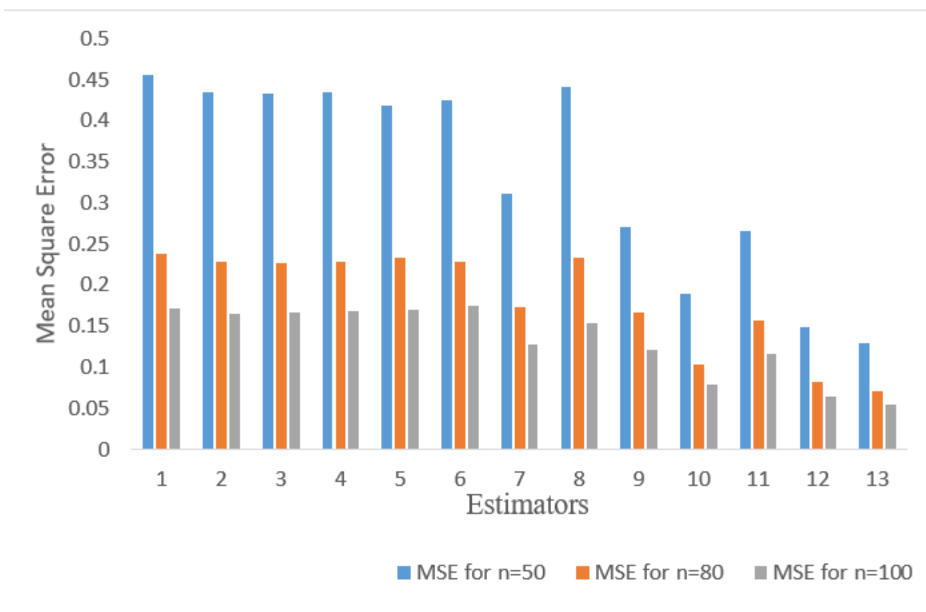

Figure 2. Comparison by MSEs of proposed and existing estimators using simulation study

$E_3 = \bar{X}_h$.

- Figure 1 of the three data sets reveal that all the recommended estimators have greater PRE than the estimators in the literature as the graph is going towards upward.
- Table 3 shows that all the efficiency conditions are satisfied as numerical differences of the MSEs between the proposed and existing estimators are positive, which is also described theoretically in Section 4.
- The results of the simulation study provided in Table 4 reveal that all the six suggested estimators have minimum MSE and maximum PRE than all the competitor estimators.
- In the simulation study, when n = 50, the variance of $v_{c1(st)}$ is 0.454859 and when n = 80, the value decreases to 0.236885. Similarly, comparing the value of the estimate Var($v_{c1(st)}$) when n = 100, its value further reduces to 0.170496. A similar finding is obtained for the remaining existing and proposed estimators. Therefore, it is clear that as the sample size increases, the variance of the estimators decreases.
- Figure 2 of the simulation study also verifies that all proposed estimators have smaller MSE than the existing estimators.

8 Conclusion

In this study, an efficient class of estimators is proposed for estimating the finite population variance using the information of an auxiliary variable under stratified random sampling. The suggested class of estimators is generated using different values of d_1, d_2, E_{1h}, E_{2h}, and E_{3h} given in Table A. Expressions for bias and MSE are derived up to the first degree of approximation. The proposed estimators are compared with the existing estimators both theoretically and numerically using PRE and MSE criteria. The numerical values of the MSE for the proposed and the mentioned estimators in the literature are computed from three populations, given in Table 2. For all the populations, the proposed estimators are quite efficient compared to the competitors' estimators. The estimator $v_{P2(st)}$ has the optimum MSE for all the data sets and shows the best performance among all the estimators considered in the study. From the results of real data sets given in Table 2 and the simulation study in Table 4, it is concluded that all the suggested estimators $v_{Pi(st)}$ have the maximum PRE compared to the competitor estimators. Overall, our proposed family of estimators are efficient performers, and we hope this family of estimators will be useful for practitioners.

Recommendations

It is recommended that this study can be extended by developing new estimators utilizing other sampling schemes.

Data Availability

The data used to support the findings of the study are available within this article.

Acknowledgment

The authors express appreciation to the editor and reviewers for their valuable and positive comments/suggestions which certainly have improved the presentation and quality of the paper.

Author Contributions

Mujeeb Hussain: Conceptualization, Supervision, analyzed and interpreted data. **Lakhkar Khan**: Investigation, software, removed all the grammatical mistakes. **Qamruz Zaman**: Data curation, Methodology. **Abdurrahman Sabir**: Improved language and Editing.

Compliance with Ethical Standards

The authors declare that they have no conflicts of interest. Additionally, this article does not involve studies with human participants or animals conducted by any of the authors. Furthermore, informed consent was taken from all individual participants involved in the study.

Author Information

ORCID:

Mujeeb Hussain: 0009-0007-5971-2685

Lakhkar Khan: 0009-0008-7160-2515

References

[1] Cochran, W. G. (2007). Sampling techniques (3rd ed.). New York: John Wiley & Sons.

[2] Pandey, M. K., Singh, G. N., Zaman, T., Mutairi, A. A., & Mustafa, M. S. (2024). A general class of improved population variance estimators under non-sampling errors using calibrated weights in stratified sampling. *Scientific Reports*, 14(1), 2948.

[3] Powers, T. (2016). Efficient estimator for population variance using auxiliary variable. *American Journal of Operational Research*, 6(1), 9-15.

[4] Kadilar, C., & Cingi, H. (2006). Ratio estimators for the population variance in simple and stratified random sampling. *Applied Mathematics and Computation*, 173(2), 1047-1059.

[5] Sidelel, E. B., Orwa, G. O., & Otieno, R. O. (2014). Variance estimation in stratified random sampling in the presence of two auxiliary random variables. *International Journal of Science and Research*, 3(9), 2453-9.

[6] Ozel, G., Cingi, H., & Oguz, M. (2014). Separate ratio estimators for the population variance in stratified random sampling. *Communications in Statistics—Theory and Methods*, 43(22), 4766-79.

[7] Cekim, H. O., & Kadilar, C. (2020). In-type estimators for the population variance in stratified random sampling. *Communications in Statistics—Simulation and Computation*, 49(7), 1665-77. doi:10.1080/03610918.2019.1577973

[8] Yasmeen, U., Noor-Ul Amin, M., & Hanif, M. (2019). Exponential estimators of finite population variance using transformed auxiliary variables. *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences*, 89(1), 185-91. doi:10.1007/s40010-017-0410-5

[9] Singh, H. P., & Solanki, R. S. (2013). A new procedure for variance estimation in simple random sampling using auxiliary information. *Statistical Papers*, 54(2), 479-97. doi:10.1007/s00362-012-0445-2

[10] Qian, J. (2020). Variance estimation with complex data and finite population correction-a paradigm for comparing jackknife and formula-based methods for variance estimation. *ETS Research Report Series*, 2020(1), 1-16. doi:10.1002/ets2.12294

[11] Cekim, H. O., & Kadilar, C. (2020). In-type variance estimators in simple random sampling. *Pakistan Journal of Statistics and Operation Research*, 16(4), 689-96. doi:10.18187/pjsor.v16i4.3072

[12] Singh, G. N., Bhattacharyya, D., & Bandyopadhyay, A. (2021). Calibration estimation of population variance under stratified successive sampling in presence of random non response. *Communications in Statistics—Theory and Methods*, 50(19), 4487-4509.

[13] Mahanty, B., & Mishra, G. (2020). An unbiased estimator of finite population mean using auxiliary information. *Journal of Statistical Theory and Applications*, 19(4), 534-9.

[14] Shahzad, U., Ahmad, I., Almanjahie, I. M., Al-Noor, N. H., & Hanif, M. (2021). A novel family of variance estimators based on L-moments and calibration approach under stratified random sampling. *Communications in Statistics—Simulation and Computation*, 1-14. doi:10.1080/03610918.2021.1945629

[15] Singh, N., Vishwakarma, G. K., & Gangele, R. K. (2021). Variance estimation in the presence of measurement errors under stratified random sampling. *REVSTAT*, 19(2), 275-90.

[16] Ahmad, S., et al. (2022). Improved estimation of finite population variance using dual supplementary information under stratified random sampling. *Math. Probl. Eng.*, 2022(2022).

[17] Etebong, P. C. (2018). Improved family of ratio estimators of finite population variance in stratified random sampling. *Biostatistics and Biometrics Open Access Journal*, 5(2), 55659. doi:10.19080/BBOAJ.2018.04.555659

[18] Bahl, S., & Tuteja, R. (1991). Ratio and product type exponential estimators. *Journal of Information and Optimization Sciences*, 12(1), 159-164.

[19] Kadilar, C., & Cingi, H. (2003). A study on the chain ratio-type estimator. *Hacettepe Journal of Mathematics and Statistics*, 32, 105-108.

[20] Singh, R., & Mangat, N. S. (1996). *Elements of Survey Sampling*. London: Kluwer Academic Publishers.