Communication in Physical Sciences, 2025, 12(2) 386-405

Generalized Variance Estimator using Two Auxiliary Variables under Stratified Random Sampling

Isah Muhammad and Gafar Matanmi Oyeyemi

Received: 28 November 2024/Accepted 01 February 2025/Published 03 February 2025
https://dx.doi.org/10.4314/cps.v12i2.15

Abstract: *Efficient and precise estimation in sample surveys often benefits from the incorporation of auxiliary information. This study addresses the challenge of improving variance estimation by developing a novel estimator for finite population variance that utilizes two auxiliary variables within the framework of stratified random sampling. The estimator's properties were derived using the approach of near-unbiasedness, ensuring theoretical rigor and robustness. Efficiency conditions that demonstrate the superiority of the suggested estimator over existing population variance estimators were established analytically. The performance of the proposed estimator was validated using four real datasets. From Dataset III, the estimator showed minimum bias (-6.937e-21), a mean square error of 7.991911, and a relative efficiency of 100.01%. Similarly, for Dataset IV, the proposed estimator achieved a bias of 9.249550e-22, a mean square error of 8.154949e-11, and a relative efficiency of 654.15%. In all cases, the proposed estimator outperformed the existing estimators based on the criteria of bias, mean square error, and percentage relative efficiency. These findings highlight the estimator's practical utility in delivering more accurate and reliable variance estimates across different applications. Consequently, the suggested estimator offers a significant contribution to the field, with potential for wide-ranging use in improving the quality of survey-based studies.*

Keywords: *Variance, auxiliary variable, bias, mean square error, efficiency.*

Isah Muhammad
Department of Statistics, Binyaminu Usman Polytechnic, Hadejia, Nigeria
Email: isahsta@gmail.com
Orcid id:

Gafar Matanmi Oyeyemi
Department of Statistics, University of Ilorin, Ilorin, Nigeria
Email: matanmi@unilorin.edu.ng

1.0 Introduction

The estimation of the population variance of the study variable is required in many application contexts. In several disciplines such as agriculture, medicine, biology, and business, where sometimes the populations are biased, variance estimate for the population is significant. Variance estimation plays a crucial role in various real-world applications where understanding variability is essential for decision-making and predictive modelling. For instance, in agricultural research, variance estimation is critical for assessing crop yield performance under different conditions such as soil types, irrigation methods, or fertilizer applications. This helps optimize resource allocation and improve agricultural productivity (Arata et al., 2020). Similarly, in clinical trials, variance estimation is fundamental for evaluating the effectiveness of new treatments or drugs. By quantifying the variability in patient responses, researchers can ensure that observed effects represent true treatment efficacy rather than random fluctuations (Miller, 2005).

Various authors including Isaki (1983), Shabir and Gupta (2010), Ahmad et al. (2022), and Zakari & Muhammad (2023) have worked on this area. Auxiliary information is frequently utilized in sample surveys to obtain accurate and reliable estimates of the population parameters such as proportions, averages, totals, and variance of the variable of interest. Cochran (1940) first introduced the use of auxiliary information in the ratio estimation method when there exists a positive correlation between the study and auxiliary variables. This method became inefficient when the correlation was negative, and thus the product estimation methods performed better. Different strategies have been employed by researchers in utilizing this information to estimate the population characteristics under different sampling techniques (Zakari et al., 2020; Muhammad et al., 2021; Yasmeen & Noor, 2021; Zakari & Muhammad, 2022; Muhammad et al., 2023; Zakari et al., 2023; Oyeyemi et al., 2023; Muhammad, 2023; Audu et al., 2023). In addition, estimation strategies with two or more auxiliary variables have been suggested by Singh, Choudhury & Kalitac (2013), Tato & Singh (2017) and Ahmad et al. (2022) among others. However, the nature of the relationship between the study and auxiliary variables in which the existing stratified variance estimators produce better estimates is a matter of concern. Therefore, to address this lingering problem, this study develops a generalized population variance estimator using two auxiliary variables that is flexible and efficient under stratified random sampling.

2.0 Sample Structure, Notations and Some Existing Estimators

Consider $\vartheta = \{\vartheta_1, , , , , , \vartheta_N\}$ to be a finite population of size N which is divided into G non-overlapping groups, with each group containing $N_h (h = 1,2, \dots, G)$ units, such that $\sum_{h=1}^{G} N_h = N$. A sample using the technique of simple random sample of size n_h is drawn *(SRSWOR)* from the population strata N_h without replacement such that $\sum_{h=1}^{G} n_h = n$. Let (y_{hi}, x_{hi}, z_{hi}) be the observation of the variable under study Y and the auxiliary variables X and Z, respectively. Supposed $\bar{y}_h, \bar{x}_h$ and $\bar{z}_h$ are the sample means corresponding to the population means $\bar{Y}_h$, $\bar{X}_h$ and $\bar{Z}_h$, respectively, in each stratum. Further, the following notations were defined by Singh (2020) and Muhammad (2023) as;

$S_{Y_h}^2 = \frac{\sum_{i=1}^{N_h}(Y_{hi} - \bar{Y}_h)^2}{(N_h - 1)}$: The population variance of the study variable Y for the i^{th} stratum

$S_{X_h}^2 = \frac{\sum_{i=1}^{N_h}(X_{hi} - \bar{X}_h)^2}{(N_h - 1)}$: The population variance of the auxiliary variable X for the i^{th} stratum

$S_{Z_h}^2 = \frac{\sum_{i=1}^{N_h}(Z_{hi} - \bar{Z}_h)^2}{(N_h - 1)}$: The population variance of the auxiliary variable Z for the i^{th} stratum

$s_{yh}^2 = \frac{\sum_{i=1}^{n_h}(y_{hi} - \bar{y}_h)^2}{(n_h - 1)}$: The sample variance of the study variable X for the i^{th} stratum

$s_{xh}^2 = \frac{\sum_{i=1}^{n_h}(x_{hi} - \bar{x}_h)^2}{(n_h - 1)}$: The sample variance of the auxiliary variable X for the i^{th} stratum

$s_{zh}^2 = \frac{\sum_{i=1}^{n_h}(z_{hi} - \bar{z}_h)^2}{(n_h - 1)}$: The sample variance of the auxiliary variable Z for the i^{th} stratum

$W_h = \frac{N_h}{N}$: The original weight of the h^{th} strata, $h = 1,2, \dots, G$

$f_{1h} = \left(\frac{1}{n_h} - \frac{1}{N_h}\right)$: The sampling fraction corresponding to the h^{th} strata

$\mu_{rst(h)} = \frac{1}{N_h}\sum_{j=1}^{N_h}(Y_{hj} - \bar{Y}_h)^r (X_{hj} - \bar{X}_h)^s (Z_{hj} - \bar{Z}_h)^t$: The moments about the mean for the i^{th} stratum

$\lambda_{rst(h)} = \dfrac{\mu_{rst(h)}}{\sqrt{\mu_{200(h)}^{r}\mu_{020(h)}^{s}\mu_{002(h)}^{t}}}$: The population kurtosis for the i^{th} stratum

The usual unbiased population variance estimator under stratified random sampling and its variance are, respectively, given as

$$T_{0(st)} = \sum_{h=1}^{G} W_h^2 f_{1h} s_{y_h}^2 \tag{1}$$

$$Var\left(T_{0(st)}\right) = \sum_{h=1}^{G} W_h^4 f_{1h}^3 S_{Y_h}^4 \left(\lambda_{40(h)} - 1\right) \tag{2}$$

Isaki (1983) suggested a ratio and product estimators using auxiliary variable, respectively, as

$$T_{1(st)} = \sum_{h=1}^{G} W_h^2 f_{1h} \left(s_{y_h}^2 \frac{S_{X_h}^2}{s_{x_h}^2} \right) \tag{3}$$

$$T_{2(st)} = \sum_{h=1}^{G} W_h^2 f_{1h} \left(s_{y_h}^2 \frac{s_{x_h}^2}{S_{X_h}^2} \right) \tag{4}$$

$$Bias\left(T_{1(st)}\right) = \sum_{h=1}^{G} W_h^2 f_{1h}^2 S_{Y_h}^2 \left[\left(\lambda_{40(h)} - 1\right) - \left(\lambda_{22(h)} - 1\right) \right] \tag{5}$$

$$Bias\left(T_{2(st)}\right) = \sum_{h=1}^{G} W_h^2 f_{1h}^2 S_{Y_h}^2 \left[\lambda_{22(h)} - 1 \right] \tag{6}$$

$$MSE\left(T_{1(st)}\right) \cong \sum_{h=1}^{G} W_h^4 f_{1h}^3 S_{Y_h}^4 \left[\left(\lambda_{40(h)} - 1\right) + \left(\lambda_{04(h)} - 1\right) - 2\left(\lambda_{22(h)} - 1\right) \right] \tag{7}$$

$$MSE\left(T_{2(st)}\right) \cong \sum_{h=1}^{G} W_h^4 f_{1h}^3 S_{Y_h}^4 \left[\left(\lambda_{40(h)} - 1\right) + \left(\lambda_{04(h)} - 1\right) + 2\left(\lambda_{22(h)} - 1\right) \right] \tag{8}$$

Rao (1991) suggested a difference type of estimator of population variance as

$$T_{R(st)} = \sum_{h=1}^{G} W_h^2 f_{1h} \left[\omega_{1h} s_{y_h}^2 + \omega_{1h} \left(S_{X_h}^2 - s_{x_h}^2 \right) \right] \tag{9}$$

$$Bias\left(T_{R(st)}\right) = \sum_{h=1}^{G} W_h^2 f_{1h}^2 S_{Y_h}^2 \left[\frac{\left(\lambda_{04(h)} - 1\right)}{f_{1h}\left\{ \left(\lambda_{40(h)} - 1\right)\left(\lambda_{04(h)} - 1\right) - \left(\lambda_{22(h)} - 1\right)^2 \right\}} + \left(\lambda_{04(h)} - 1\right) - 1 \right] \tag{10}$$

$$MSE\left(T_{R(st)}\right) = \sum_{h=1}^{G} W_h^4 f_{1h}^3 S_{Y_h}^4 \left[\left(\lambda_{40(h)} - 1\right)\left(\lambda_{04(h)} - 1\right) - \frac{\left(\lambda_{22(h)} - 1\right)^2}{f_{1h}\left\{ \begin{array}{l} \left(\lambda_{40(h)} - 1\right)\left(\lambda_{04(h)} - 1\right) \\ -\left(\lambda_{22(h)} - 1\right)^2 \end{array} \right\}} + \left(\lambda_{04(h)} - 1\right) \right] \tag{11}$$

where $\omega_{1h}=\dfrac{\left(\lambda_{04(h)}-1\right)}{\gamma_h\left[\left(\lambda_{40(h)}-1\right)\left(\lambda_{04(h)}-1\right)-\left(\lambda_{22(h)}-1\right)^2\right]+\left(\lambda_{04(h)}-1\right)}$ and

$$\omega_{2h}=\frac{S_{y_h}^2\left(\lambda_{22(h)}-1\right)}{S_{x_h}^2\left\{\gamma_h\left[\left(\lambda_{40(h)}-1\right)\left(\lambda_{04(h)}-1\right)-\left(\lambda_{22(h)}-1\right)^2\right]+\left(\lambda_{04(h)}-1\right)\right\}}$$

Singh et al. (2009) proposed exponential ratio and product-type estimators as

$$T_{er(st)}=\sum_{h=1}^{G}W_h^2 f_{1h}s_{y_h}^2\exp\left(\frac{S_{X_h}^2-s_{x_h}^2}{S_{X_h}^2+s_{x_h}^2}\right) \tag{12}$$

$$T_{ep(st)}=\sum_{h=1}^{G}W_h^2 f_{1h}s_{y_h}^2\exp\left(\frac{s_{x_h}^2-S_{X_h}^2}{S_{X_h}^2+s_{x_h}^2}\right) \tag{13}$$

$$Bias(T_{er(st)})\cong\sum_{h=1}^{G}W_h^2 f_{1h}^2S_{Y_h}^2\left[\frac{3}{8}\left(\lambda_{04(h)}-1\right)-\frac{1}{2}\left(\lambda_{22(h)}-1\right)\right] \tag{14}$$

$$Bias(T_{ep(st)})\cong\sum_{h=1}^{G}W_h^2 f_{1h}^2S_{Y_h}^2\left[\frac{1}{2}\left(\lambda_{22(h)}-1\right)-\frac{1}{8}\left(\lambda_{04(h)}-1\right)\right] \tag{15}$$

$$MSE(T_{er(st)})\cong\sum_{h=1}^{G}W_h^4 f_{1h}^3S_{Y_h}^4\left[\left(\lambda_{40(h)}-1\right)+\frac{1}{4}\left(\lambda_{04(h)}-1\right)-\left(\lambda_{22(h)}-1\right)\right] \tag{16}$$

$$MSE(T_{ep(st)})\cong\sum_{h=1}^{G}W_h^4 f_{1h}^3S_{Y_h}^4\left[\left(\lambda_{40(h)}-1\right)+\frac{1}{4}\left(\lambda_{04(h)}-1\right)+\left(\lambda_{22(h)}-1\right)\right] \tag{17}$$

Shabbir and Gupta (2010) proposed regression-cum-ratio estimator as

$$T_{SG(st)}=\sum_{h=1}^{G}\left(\frac{W_h^2}{n_h}\right)\left[k_{1_h}s_{y_h}^2+k_{2_h}\left(S_{X_h}^2-s_{x_h}^2\right)\right]\exp\left(\frac{S_{X_h}^2-s_{x_h}^2}{S_{X_h}^2+s_{x_h}^2}\right) \tag{18}$$

$$Bias(T_{SG(st)})\cong\sum_{h=1}^{G}\left(\frac{W_h^2}{n_h}\right)\left[\left(\alpha_{1h}-1\right)S_{Y_h}^2-\frac{1}{2}\alpha_{1h}S_{Y_h}^2\frac{\left(\lambda_{22(h)}-1\right)}{n_h}+\left\{S_{X_h}^2\alpha_{2h}+\frac{3}{4}\alpha_{1h}S_{Y_h}^2\frac{\left(\lambda_{04(h)}-1\right)}{n_h}\right\}\right] \tag{19}$$

The minimum mean square error (MSE) expression of the estimator T_{SG} up to the first order of approximation at optimum values;

$$k_{1h}^{(opt)}=\frac{\left(\lambda_{04(h)}-1\right)\left\{2-\frac{1}{4n_h}\left(\lambda_{04(h)}-1\right)\right\}}{2\left[\left(\lambda_{04(h)}-1\right)\left(1+\frac{A}{n_h}\right)-\frac{B^2}{n_h}\right]} \text{ and } k_{2h}^{(opt)}=\left[\frac{S_{Y_h}^2}{2S_{X_h}^2}-\frac{S_{Y_h}^2\left\{2-\frac{1}{4n_h}\left(\lambda_{04(h)}-1\right)\right\}B}{2S_{X_h}^2\left[\left(\lambda_{04(h)}-1\right)\left(1+\frac{A}{n_h}\right)-\frac{B^2}{n_h}\right]}\right]$$

where $A=\left(\lambda_{40(h)}-1\right)+\left(\lambda_{04(h)}-1\right)-2\left(\lambda_{22(h)}-1\right)$ and $B=\left(\lambda_{04h}-1\right)-\left(\lambda_{22h}-1\right)$, is given as

$$MSE(T_{SG(st)})_{\min} \cong \sum_{h=1}^{G}\left(\frac{W_h^4}{n_h^2}\right)\left[\frac{\left(\frac{S_{Y_h}^4}{n_h}\right)\left(\lambda_{40(h)}-1\right)\left(1-\rho_h^2\right)}{\left[1+\frac{\left(\frac{S_{Y_h}^4}{n_h}\right)\left(\lambda_{40(h)}-1\right)\left(1-\rho_h^2\right)}{S_{Y_h}^4}\right]} - \frac{\frac{\left(\lambda_{04(h)}-1\right)}{n_h}\left\{\left(\frac{S_{Y_h}^4}{n_h}\right)\left(\lambda_{40(h)}-1\right)\left(1-\rho_h^2\right)+\frac{S_{Y_h}^4\left(\lambda_{04(h)}-1\right)}{16n_h}\right\}}{4\left[1+\frac{\left(\frac{S_{Y_h}^4}{n_h}\right)\left(\lambda_{40(h)}-1\right)\left(1-\rho_h^2\right)}{S_{Y_h}^4}\right]}\right] \tag{20}$$

Koyuncu (2013) proposed improved estimators of finite population variance in stratified random sampling as

$$T_{Ni(st)} = \sum_{h=1}^{G}\left(\frac{W_h^2}{n_h}\right)\left[w_{1_h}s_{y_h}^2 + w_{2_h}\left(\frac{s_{x_h}^2}{S_{X_h}^2}\right)^{\gamma h}\right]\exp\left(\frac{\eta_h\left(S_{X_h}^2 - s_{x_h}^2\right)}{\eta_h\left(S_{X_h}^2 - s_{x_h}^2\right)+2\varpi_h}\right),\ (i=1,2,3,4,5) \tag{21}$$

$$Bias\left(T_{Ni(st)}\right) \cong \sum_{h=1}^{G}\left(\frac{W_h^2}{n_h}\right)\left[\begin{array}{l}\left(w_{1h}-1\right)S_{Y_h}^2 + w_{2h} - \frac{\tau_{ih}}{2}w_{1h}S_{Y_h}^2\frac{\left(\lambda_{22(h)}-1\right)}{n_h} \\ +\left\{w_{2h}\left[\frac{\gamma_h\left(\gamma_h-1\right)}{2}-\frac{\gamma_h}{2}\tau_{ih}+\frac{3}{4}\tau_{ih}^2\right]+\frac{3}{4}\tau_{ih}^2 w_{1h}S_{Y_h}^2\right\}\frac{\left(\lambda_{04(h)}-1\right)}{n_h}\end{array}\right] \tag{22}$$

The minimum mean square error of the estimator $T_{Ni(st)}$ up to the first order of approximation at the optimum values;

$$w_{1h}^{(opt)} = \frac{GF-2DB}{4AB-F^2} \quad \text{and} \quad w_{2h}^{(opt)} = \frac{S_{Y_h}^2\left(DF-2AG\right)}{4AB-F^2}$$

is given as

$$MSE\left(T_{Ni(st)}\right)_{min} = \sum_{h=1}^{G}W_h^4 S_{Y_h}^4\left[1-\frac{BD^2-DGF+AG^2}{\left(4AB-F^2\right)}\right] \tag{23}$$

where

$$A=\left(1+\frac{\left(\lambda_{40(h)}-1\right)}{n_h}+\tau_{ih}^2\frac{\left(\lambda_{04(h)}-1\right)}{n_h}-2\tau_{ih}\frac{\left(\lambda_{22(h)}-1\right)}{n_h}\right),$$

$$B=\left(1+\left\{\gamma_h^2+\tau_{ih}^2+\gamma_h(\gamma_h-1)-2\gamma_h\tau_{ih}\right\}\frac{\left(\lambda_{04(h)}-1\right)}{n_h}\right),\ D=\left(-2+\tau_{ih}\frac{\left(\lambda_{22(h)}-1\right)}{n_h}-\frac{3}{4}\tau_{ih}^2\frac{\left(\lambda_{04(h)}-1\right)}{n_h}\right),$$

$$G=\left(-2+\left\{\gamma_h\tau_{ih}-\gamma_h(\gamma_h-1)-\frac{3}{4}\tau_{ih}^2\right\}\frac{\left(\lambda_{04(h)}-1\right)}{n_h}\right),$$

$$F=\left(2+\left\{\gamma_h(\gamma_h-1)+\tau_{ih}^2-2\gamma_h\tau_{ih}\right\}\frac{\left(\lambda_{04(h)}-1\right)}{n_h}+2\left\{\gamma_h-\tau_{ih}\right\}\frac{\left(\lambda_{22(h)}-1\right)}{n_h}\right).$$

$$\tau_{i(h)}=\frac{\eta_h S_{X_h}^2}{\eta_h S_{X_h}^2+\varpi_h}, \qquad (i=1,2,3,4,5). \qquad \tau_{1(h)}=\frac{S_{X_h}^2}{S_{X_h}^2+\beta_{2(h)}(x)}, \qquad \tau_{2(h)}=\frac{S_{X_h}^2}{S_{X_h}^2+C_{x(h)}}$$

$$\tau_{3(h)}=\frac{C_{x(h)}S_{X_h}^2}{C_{x(h)}S_{X_h}^2+\beta_{2(h)}(x)}, \quad \tau_{4(h)}=\frac{\beta_{2(h)}(x)S_{X_h}^2}{\beta_{2(h)}(x)S_{X_h}^2+C_{x(h)}}, \quad \tau_{5(h)}=\frac{S_{X_h}^2}{S_{x(h)}^2+\theta_{22(h)}} \text{ and } \gamma_h=\frac{1}{n}$$

Yasmeen and Noor (2021) considered the adoption of multiple and transformed auxiliary information and suggested three ratio cum exponential type estimators as

$$T_{YN1(st)}=\sum_{h=1}^{G}\left(\frac{W_h^2}{n_h}\right)\left(s_{y_h}^2\,\frac{\alpha_h S_{X_h}^2-\chi_h}{\alpha_h s_{x_h}^2-\chi_h}\right) \tag{24}$$

$$T_{YN2(st)}=\sum_{h=1}^{G}\left(\frac{W_h^2}{n_h}\right)\left[s_{y_h}^2\left(\frac{S_{X_h}^2}{s_{x_h}^2}\right)\exp\left(\frac{S_{Z_h}^2-s_{z_h}^2}{s_{z_h}^2+S_{Z_h}^2}\right)\right] \tag{25}$$

$$T_{YN3(st)}=\sum_{h=1}^{G}\left(\frac{W_h^2}{n_h}\right)\left[s_{y_h}^2\left(\frac{s_{x_h}^{*2}}{S_{X_h}^2}\right)^{k_{1(h)}}\exp\left(k_{2(h)}\,\frac{s_{z_h}^{*2}-S_{Z_h}^2}{s_{z_h}^{*2}+S_{Z_h}^2}\right)\right] \tag{26}$$

The bias expression of the estimators up to the first order of approximation are, respectively, given as

$$Bias\left(T_{YN1(st)}\right)\cong\sum_{h=1}^{G}\left(\frac{W_h^2}{n_h^2}\right)S_{Y_h}^2\left[\phi_{i(h)}^2\left(\lambda_{04(h)}-1\right)-\phi_{i(h)}\left(\lambda_{22(h)}-1\right)\right] \tag{27}$$

$$Bias\left(T_{YN2(st)}\right)\cong\sum_{h=1}^{G}\left(\frac{W_h^2}{n_h^2}\right)S_{Y_h}^2\left[\frac{3}{8}\left(\lambda_{004(h)}-1\right)+\frac{1}{2}\left(\lambda_{022(h)}-1\right)-\frac{1}{2}\left(\lambda_{202(h)}-1\right)-\left(\lambda_{220(h)}-1\right)\right] \tag{28}$$

$$Bias\left(T_{YN3(st)}\right)\cong\sum_{h=1}^{G}\left(\frac{W_h^2}{n_h^2}\right)S_{Y_h}^2\left[\begin{array}{l}\frac{k_{1(h)}\left(k_{1(h)}-1\right)}{2}g_h^2\left(\lambda_{040(h)}-1\right)+\frac{k_{2(h)}\left(k_{2(h)}-1\right)}{8}g_h^2\left(\lambda_{004(h)}-1\right)\\+\frac{1}{2}k_{1(h)}k_{2(h)}g_h^2\left(\lambda_{022(h)}-1\right)-k_{1(h)}g_h f_{1h}\left(\lambda_{220(h)}-1\right)\\-\frac{1}{2}k_{2(h)}g_h\left(\lambda_{202(h)}-1\right)\end{array}\right] \tag{29}$$

where

$$\phi_{i(h)}=\frac{\alpha_h S_{X_h}^2}{\alpha_h S_{X_h}^2+\chi_h}, \ (i=1,2,3,4,5,\ldots,26). \ k_{1(h)}=A_{(h)}-\frac{\left(C_{(h)}-A_{(h)}D_{(h)}\right)B_{(h)}}{1-B_{(h)}D_{(h)}}, \ k_{2(h)}=\frac{C_{(h)}-A_{(h)}D_{(h)}}{1-B_{(h)}D_{(h)}},$$

$$A_{(h)}=\frac{\left(\lambda_{220(h)}-1\right)}{g_h\left(\lambda_{040(h)}-1\right)}, \ B_{(h)}=\frac{\left(\lambda_{022(h)}-1\right)}{2\left(\lambda_{040(h)}-1\right)}, \ C_{(h)}=\frac{2\left(\lambda_{202(h)}-1\right)}{g_h\left(\lambda_{040(h)}-1\right)}, \ D_{(h)}=\frac{2\left(\lambda_{022(h)}-1\right)}{\left(\lambda_{004(h)}-1\right)} \text{ and } g_h=\frac{n_h}{N_h-n_h}$$

The MSE expression of the estimators $T_{YN1(st)}$, $T_{YN2(st)}$ and $T_{YN3(st)}$ are, respectively, given as

$$MSE\left(T_{YN1(st)}\right) \cong \sum_{h=1}^{G}\left(\frac{W_h^4}{n_h^3}\right)S_{Y_h}^4\left[\left(\lambda_{40(h)}-1\right)+\phi_{i(h)}^2\left(\lambda_{04(h)}-1\right)-2\left(\lambda_{22(h)}-1\right)\right] \tag{30}$$

$$MSE\left(T_{YN2(st)}\right) \cong \sum_{h=1}^{G}\left(\frac{W_h^4}{n_h^3}\right)S_{Y_h}^4\left[\begin{array}{l}\left(\lambda_{400(h)}-1\right)+\left(\lambda_{040(h)}-1\right)+\frac{1}{2}\left(\lambda_{004(h)}-1\right)-2\left(\lambda_{220(h)}-1\right)\\ -\left(\lambda_{022(h)}-1\right)+\left(\lambda_{202(h)}-1\right)\end{array}\right] \tag{31}$$

$$MSE\left(T_{YN3(st)}\right)_{\min} \cong \sum_{h=1}^{G}\frac{W_h^4S_{Y_h}^4}{n_h^3\left(1-B_{(h)}D_{(h)}\right)}\left[\begin{array}{l}\left(\lambda_{400(h)}-1\right)+\dfrac{\left(A_{(h)}-B_{(h)}C_{(h)}\right)^2}{\left(1-B_{(h)}D_{(h)}\right)}g_h^2\left(\lambda_{040(h)}-1\right)\\ +\dfrac{\left(C_{(h)}-A_{(h)}D_{(h)}\right)^2}{4\left(1-B_{(h)}D_{(h)}\right)}g_h^2\left(\lambda_{004(h)}-1\right)\\ -2\left(A_{(h)}-B_{(h)}C_{(h)}\right)g_h\left(\lambda_{220(h)}-1\right)\\ +\dfrac{\left(\left(A_{(h)}-B_{(h)}C_{(h)}\right)\left(C_{(h)}-A_{(h)}D_{(h)}\right)\right)}{\left(1-B_{(h)}D_{(h)}\right)}g_h^2\left(\lambda_{022(h)}-1\right)\\ -\left(C_{(h)}-A_{(h)}D_{(h)}\right)g_h\left(\lambda_{202(h)}-1\right)\end{array}\right] \tag{32}$$

Ahmad et al. (2022) suggested an improved population variance estimator under stratified random sampling as

$$T_{a(st)} = \sum_{h=1}^{G}W_h^2 f_{1h}\left[k_{1h}s_{y_h}^2+k_{2h}\left(S_{X_h}^2-s_{x_h}^2\right)+k_{3h}\left(S_{rx_h}^2-s_{rx_h}^2\right)\right]\exp\left(\frac{S_{X_h}^2-s_{x_h}^2}{S_{X_h}^2+s_{x_h}^2}\right) \tag{33}$$

The bias and MSE expressions of the estimator T_a at optimum values;

$$k_{1h(opt)} = \frac{8-f_{1h}\lambda_{04(h)}}{\lambda_{04(h)}\left[8\left(1/\lambda_{40(h)}+f_{1h}(\gamma_h+1)\right)\right]},$$

$$k_{2h(opt)} = \frac{S_{Y_h}^2\left(\begin{array}{l}f_{1h}\lambda_{040(h)}\left[\lambda_{040(h)}\lambda_{004(h)}-\lambda_{220(h)}^2\right]+\left[\lambda_{004(h)}\lambda_{220(h)}-\lambda_{202(h)}\lambda_{022(h)}\right]\left(8-f_{1h}\lambda_{040(h)}\right)\\ +4\lambda_{400(h)}\left(\lambda_{040(h)}\lambda_{004(h)}-\lambda_{220(h)}^2\right)\left[\left(-1/\lambda_{400(h)}\right)+f_{1h}\left(\gamma_h+1\right)\right]\end{array}\right)}{8S_{xh}^2\lambda_{400(h)}\left\{\lambda_{040(h)}\lambda_{004(h)}-\lambda_{220(h)}^2\right\}\left[(-1/\lambda_{400(h)})+f_{1h}(\gamma_h+1)\right]} \text{ and}$$

$$k_{3h(opt)} = \frac{S_{Y_h}^2\left[8-f_{1h}\lambda_{040(h)}\right]\left(\lambda_{220(h)}\lambda_{202(h)}-\lambda_{040(h)}\lambda_{202(h)}\right)}{8S_{rx_h}^2\lambda_{400(h)}\left\{\lambda_{040(h)}\lambda_{004(h)}-\lambda_{220(h)}^2\right\}\left[(-1/\lambda_{400(h)})+f_{1h}(\gamma_h+1)\right]}$$

are, respectively, given as

$$Bias(T_{a(st)}) \cong \sum_{h=1}^{G}W_h^2 f_{1h}^2\left[\begin{array}{l}-S_{Y_h}^2+k_{1h}S_{Y_h}^2\left(1+\frac{3}{8}\lambda_{040(h)}-\frac{1}{2}\lambda_{220(h)}\right)\\ +\frac{1}{2}f_{1h}\left(k_{2h}S_{X_h}^2\lambda_{040(h)}+k_{3h}S_{rx_h}^2\lambda_{220(h)}\right)\end{array}\right] \tag{34}$$

$$MSE(T_{a(st)})_{\min} \cong \sum_{h=1}^{G} W_h^4 f_{1h}^3 S_{Y_h}^4 \left[\frac{64(\gamma_h + 1) - f_{1h}\left(\lambda_{040(h)}^2 / \lambda_{400(h)}\right) - 16 f_{1h}\lambda_{040(h)}(\gamma_h + 1)}{64\left[(1/\lambda_{400(h)}) + f_{1h}(\gamma_h + 1)\right]} \right] \tag{35}$$

where $\gamma_h = \dfrac{2\lambda_{220(h)}\lambda_{202(h)}\lambda_{022(h)} - \lambda_{040(h)}\lambda_{202(h)}^2 - \lambda_{004(h)}\lambda_{220(h)}^2}{\lambda_{400(h)}\left(\lambda_{040(h)}\lambda_{004(h)} - \lambda_{220(h)}^2\right)}$

3.0 Proposed Estimator

The proposed generalized population variance estimator using two auxiliary variables is given as;

$$t_{z(st)} = \beta_{1(h)} T_{1z(st)} + \beta_{2(h)} T_{2z(st)} + \beta_{3(h)} T_{3z(st)} \tag{36}$$

where,

$$T_{1z(st)} = \sum_{h=1}^{G} W_h^2 f_{1(h)} s_{y_h}^2, \quad T_{2z(st)} = \sum_{h=1}^{G} W_h^2 f_{1(h)} s_{y_h}^2 \left(\frac{S_{x_h}^2}{s_{x_h}^2} \frac{S_{z_h}^2}{s_{z_h}^2} \right) \text{ and } T_{3z(st)} = \sum_{h=1}^{G} W_h^2 f_{1(h)} s_{y_h}^2 \exp\left(\frac{s_{x_h}^2 \dfrac{s_{z_h}^2}{S_{z_h}^2} - S_{x_h}^2}{s_{x_h}^2 \dfrac{s_{z_h}^2}{S_{z_h}^2} + S_{x_h}^2} \right)$$

By definition, the class *P* will consist of all $T_{iz(st)}$ of the forms;

$$t_{z(st)} = \sum_{i=1}^{3} \beta_{i(h)} T_{iz(st)} \in \mathrm{P} \tag{37}$$

where *P* denotes the set of all possible estimators for estimating the population variance $S_{y_h}^2$ such that $\sum_{i=1}^{3} \beta_{i(h)} = 1$ and $\beta_{i(h)} \in \mathbb{R}$, where $\beta_{i(h)}$ (i =1,2,3) denotes the constants used for reducing the bias in the class of estimators $T_{1z(st)}, T_{2z(st)}, T_{3z(st)} \in \mathrm{P}$ and ℝ denotes the set of real numbers.

3.1 Bias and MSE of Proposed Estimator $(t_{z(st)})$

To derive, the bias and MSE of the proposed estimator $t_{z(st)}$ in equation (36), the following error terms are defined;

$e_{0(h)} = \left(s_{y_h}^2 - S_{y_h}^2\right) / S_{y_h}^2$ and $e_{1(h)} = \left(s_{x_h}^2 - S_{x_h}^2\right) / S_{x_h}^2$ $e_{2(h)} = \left(s_{z_h}^2 - S_{z_h}^2\right) / S_{z_h}^2$ such that $\left|e_{0(h)}\right| < 1$, $\left|e_{1(h)}\right| < 1$, $\left|e_{2(h)}\right| < 1$

$$\left.\begin{aligned}
&E\left(e_{0(h)}\right) = E\left(e_{1(h)}\right) = E\left(e_{2(h)}\right) = 0, E\left(e_{0(h)}^2\right) = f_{1h}\left(\lambda_{400(h)} - 1\right) = f_{1h} C_{0(h)}^2 \\
&E\left(e_{1(h)}^2\right) = f_{1h}\left(\lambda_{040(h)} - 1\right) = f_{1h} C_{1(h)}^2, E\left(e_{2(h)}^2\right) = f_{1h}\left(\lambda_{004(h)} - 1\right) = f_{1h} C_{2(h)}^2 \\
&E\left(e_{0(h)} e_{1(h)}\right) = f_{1h}\left(\lambda_{220(h)} - 1\right) = f_{1h}\rho_{01(h)}, E\left(e_{0(h)} e_{2(h)}\right) = f_{1h}\left(\lambda_{202(h)} - 1\right) = f_{1h}\rho_{02(h)} \\
&E\left(e_{1(h)} e_{2(h)}\right) = f_{1h}\left(\lambda_{022(h)} - 1\right) = f_{1h}\rho_{12(h)}
\end{aligned}\right\} \tag{38}$$

Expressing $t_{z(st)}$ in equation (36) in terms of error terms, we have

$$t_{z(st)}=\sum_{h=1}^{G}W_h^2 f_{1h}S_{y_h}^2\left(1+e_{0(h)}\right)\left\{\beta_{1(h)}+\beta_{2(h)}\left[\frac{S_{x_h}^2}{S_{x_h}^2\left(1+e_{1(h)}\right)}\frac{S_{z_h}^2}{S_{z_h}^2\left(1+e_{2(h)}\right)}\right]+\beta_{3(h)}\exp\left[\frac{\left(1+e_{1(h)}\right)S_{x_h}^2\dfrac{\left(1+e_{2(h)}\right)S_{z_h}^2}{S_{z_h}^2}-S_{x_h}^2}{\left(1+e_{1(h)}\right)S_{x_h}^2\dfrac{\left(1+e_{2(h)}\right)S_{z_h}^2}{S_{z_h}^2}+S_{x_h}^2}\right]\right\} \quad (39)$$

By some appropriate simplification, equation (39) becomes:

$$t_{z(st)}=\sum_{h=1}^{G}W_h^2 f_{1h}S_{y_h}^2\left(1+e_{0(h)}\right)\left\{\beta_{1(h)}+\beta_{2(h)}\left[\left(1+e_{1(h)}\right)^{-1}\left(1+e_{2(h)}\right)^{-1}\right]+\beta_{3(h)}\exp\left[\frac{\left(1+e_{1(h)}\right)\left(1+e_{2(h)}\right)S_{x_h}^2-S_{x_h}^2}{\left(1+e_{1(h)}\right)\left(1+e_{2(h)}\right)S_{x_h}^2+S_{x_h}^2}\right]\right\} \quad (40)$$

Thus it follows from equation (40) that

$$t_{z(st)}=\sum_{h=1}^{G}W_h^2 f_{1h}S_{y_h}^2\left(1+e_{0(h)}\right)\left\{\begin{array}{l}\beta_{1(h)}+\beta_{2(h)}\left[\left(1+e_{1(h)}\right)^{-1}\left(1+e_{2(h)}\right)^{-1}\right]\\+\beta_{3(h)}\exp\left[\dfrac{\left(e_{1(h)}+e_{2(h)}+e_{1(h)}e_{2(h)}\right)}{2}\left(1+\dfrac{\left(e_{1(h)}+e_{2(h)}+e_{1(h)}e_{2(h)}\right)}{2}\right)^{-1}\right]\end{array}\right\} \quad (41)$$

By expanding $\left(1+e_{1(h)}\right)^{-1}$, $\left(1+e_{2(h)}\right)^{-1}$ and $\left(1+\frac{\left(e_{1(h)}+e_{2(h)}+e_{1(h)}e_{2(h)}\right)}{2}\right)^{-1}$ to the first order approximation, and multiplying out the RHS, the equation (41) gives:

$$t_{z(st)}=\sum_{h=1}^{G}W_h^2 f_{1h}S_{y_h}^2\left(1+e_{0(h)}\right)\left\{\beta_{1(h)}+\beta_{2(h)}\left[\left(1-e_{1(h)}+e_{1(h)}^2\right)\left(1-e_{2(h)}+e_{2(h)}^2\right)\right]+\beta_{3(h)}\exp\left[\frac{e_{1(h)}}{2}+\frac{e_{2(h)}}{2}-\frac{e_{1(h)}^2}{4}-\frac{e_{2(h)}^2}{4}\right]\right\} \quad (42)$$

Applying the concept of exponential series to equation (42), we have

$$t_{z(st)}=\sum_{h=1}^{G}W_h^2 f_{1h}S_{y_h}^2\left(1+e_{0(h)}\right)\left\{\begin{array}{l}\beta_{1(h)}+\beta_{2(h)}\left(1-e_{1(h)}-e_{2(h)}+e_{1(h)}^2+e_{2(h)}^2+e_{1(h)}e_{2(h)}\right)\\+\beta_{3(h)}\left(1+\dfrac{e_{1(h)}}{2}+\dfrac{e_{2(h)}}{2}-\dfrac{e_{1(h)}^2}{8}-\dfrac{e_{2(h)}^2}{8}\right)\end{array}\right\} \quad (43)$$

Multiplying out the RHS of (43) up to the first-order (-second-degree) approximation, we have

$$t_{z(st)}=\sum_{h=1}^{G}W_h^2 f_{1h}S_{y_h}^2\left(1+e_{0(h)}\right)\left\{\begin{array}{l}1-\left(\beta_{2(h)}-\dfrac{\beta_{3(h)}}{2}\right)e_{1(h)}-\left(\beta_{2(h)}-\dfrac{\beta_{3(h)}}{2}\right)e_{2(h)}+\left(\beta_{2(h)}-\dfrac{\beta_{3(h)}}{8}\right)e_{1(h)}^2\\+\left(\beta_{2(h)}-\dfrac{\beta_{3(h)}}{8}\right)e_{2(h)}^2+\beta_{2(h)}e_{1(h)}e_{2(h)}\end{array}\right\}$$

(44)

where

$$\beta_{1(h)}+\beta_{2(h)}+\beta_{3(h)}=1 \quad (45)$$

Multiplying the RHS and factorizing the common terms, the equation (45) becomes;

$$\left(t_{z(st)}-\sum_{h=1}^{G}W_h^2 f_{1h}S_{y(h)}^2\right)=\sum_{h=1}^{G}W_h^2 f_{1h}S_{y_h}^2\left\{\begin{array}{l}e_{0(h)}-\left(\beta_{2(h)}-\frac{\beta_{3(h)}}{2}\right)e_{1(h)}-\left(\beta_{2(h)}-\frac{\beta_{3(h)}}{2}\right)e_{2(h)}+\left(\beta_{2(h)}-\frac{\beta_{3(h)}}{8}\right)e_{1(h)}^2\\+\left(\beta_{2(h)}-\frac{\beta_{3(h)}}{8}\right)e_{2(h)}^2+\beta_{2(h)}e_{1(h)}e_{2(h)}-\left(\beta_{2(h)}-\frac{\beta_{3(h)}}{2}\right)e_{0(h)}e_{1(h)}\\-\left(\beta_{2(h)}-\frac{\beta_{3(h)}}{2}\right)e_{0(h)}e_{2(h)}\end{array}\right\} \quad (46)$$

Taking expectation to both sides of equation (46) and applying the results of (38), the bias of the proposed estimator $t_{z(st)}$ is obtained as:

$$Bias\left(t_{z(st)}\right)=\sum_{h=1}^{G}W_h^2 f_{1(h)}^2 S_{y_h}^2\left\{\begin{array}{l}\left(\beta_{2(h)}-\frac{\beta_{3(h)}}{8}\right)C_{1(h)}^2+\left(\beta_{2(h)}-\frac{\beta_{3(h)}}{8}\right)C_{2(h)}^2-\left(\beta_{2(h)}-\frac{\beta_{3(h)}}{2}\right)\rho_{01(h)}\\-\left(\beta_{2(h)}-\frac{\beta_{3(h)}}{2}\right)\rho_{02(h)}-\beta_{2(h)}\rho_{12(h)}\end{array}\right\} \quad (47)$$

Squaring and taking the expectation of (46) and applying the results of (38), the mean square error of the proposed estimator $t_{z(st)}$ IIs obtained as:

$$MSE\left(t_{z(st)}\right)=\sum_{h=1}^{G}W_h^4 f_{1(h)}^3 S_{y(h)}^4\left\{\begin{array}{l}C_{0(h)}^2+\left(\beta_{2(h)}-\frac{\beta_{3(h)}}{2}\right)^2\left[C_{1(h)}^2+C_{2(h)}^2+2\rho_{12(h)}\right]\\-2\left(\beta_{2(h)}-\frac{\beta_{3(h)}}{2}\right)\left[\rho_{01(h)}+\rho_{02(h)}\right]\end{array}\right\} \quad (48)$$

The optimum values of the filtration parameters $\beta_{i(h)}$ $(i=1,2,3)$ are obtained by solving the three conditions using Echelon operation. Differentiating (48) partially with respect to $\left(\beta_{2(h)}-\frac{\beta_{3(h)}}{2}\right)$ and equating to zero, we have

$$\frac{\partial MSE\left(t_{z(st)}\right)}{\partial\left(\beta_{2(h)}-\frac{\beta_{3(h)}}{2}\right)}=\sum_{h=1}^{G}W_h^4 f_{1(h)}^3 S_{y_h}^4\left\{2\left(\beta_{2(h)}-\frac{\beta_{3(h)}}{2}\right)\left[C_{1(h)}^2+C_{2(h)}^2+2\rho_{12(h)}\right]-2\left[\rho_{01(h)}+\rho_{02(h)}\right]\right\}=0 \quad (49)$$

Simplifying equation (49), we obtain

$$\beta_{2(h)}-\frac{\beta_{3(h)}}{2}=\frac{H_h}{J_h} \quad (50)$$

where

$H_h=C_{1(h)}^2+C_{2(h)}^2+2\rho_{12(h)}$ and $J_h=\rho_{01(h)}+\rho_{02(h)}$

It is noted from (45) and (50) that three parameters are involved, and a unique solution may not to be obtained from the two equations. Therefore, the following additional linear restriction is imposed to obtain the unique values of $\beta_{i(h)}^{'s}$ $(i=1,2,3)$ as:

$$B\left(T_{1z(st)}\right)\beta_{1(h)}+B\left(T_{2z(st)}\right)\beta_{2(h)}+B\left(T_{3z(st)}\right)\beta_{3(h)}=0 \quad (51)$$

where $B\left(T_{1z(st)}\right)$, $B\left(T_{2z(st)}\right)$ and $B\left(T_{3z(st)}\right)$ denotes the bias expressions of estimators $T_{1z(st)}$, $T_{2z(st)}$ and $T_{3z(st)}$, respectively.

Expressing the three equations namely; (45), (51) and (50), respectively, in matrix form we have

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & B\left(T_{2z(st)}\right) & B\left(T_{3z(st)}\right) \\ 0 & 1 & -\frac{1}{2} \end{pmatrix}\begin{pmatrix} \beta_{1(h)} \\ \beta_{2(h)} \\ \beta_{3(h)} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \frac{H_h}{J_h} \end{pmatrix} \tag{52}$$

Applying the operation $B\left(T_{2z(st)}\right)R_3 - R_2 \to R_3^*$ into equation (52), we have

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & B\left(T_{2z(st)}\right) & B\left(T_{3z(st)}\right) \\ 0 & 0 & -\left[\frac{1}{2}B\left(T_{2z(st)}\right) + B\left(T_{3z(st)}\right)\right] \end{pmatrix}\begin{pmatrix} \beta_{1(h)} \\ \beta_{2(h)} \\ \beta_{3(h)} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \frac{H_h\left(T_{2z(st)}\right)}{J_h} \end{pmatrix} \tag{53}$$

From (53) we obtain the optimum values of the filtration parameters $\beta_{i(h)}^{'s}$ $(i = 1, 2, 3)$ as:

$$\beta_{1(h)} = \frac{B\left(T_{2z(st)}\right)\left[\frac{1}{2}J_h + H_h\right] + B\left(T_{3(st)}\right)\left[J_h - H_h\right]}{J_h\left[\frac{1}{2}B\left(T_{2(st)}\right) + B\left(T_{3(st)}\right)\right]} \tag{54}$$

$$\beta_{2(h)} = \frac{H_h B\left(T_{3z(st)}\right)}{J_h\left[\frac{1}{2}B\left(T_{2(st)}\right) + B\left(T_{3(st)}\right)\right]} \tag{55}$$

$$\beta_{3(h)} = -\frac{H_h B\left(T_{2z(st)}\right)}{J_h\left[\frac{1}{2}B\left(T_{2(st)}\right) + B\left(T_{3(st)}\right)\right]} \tag{56}$$

where

$B\left(T_{2z(st)}\right) = \sum_{h=1}^{G} W_h^2 f_{1(h)}^2 S_{y_h}^2 \left[C_{1(h)}^2 + C_{2(h)}^2 - \rho_{01(h)} - \rho_{02(h)} + \rho_{12(h)}\right]$ and

$B\left(T_{3z(st)}\right) = \sum_{h=1}^{G} W_h^2 f_{1(h)}^2 S_{y_h}^2 \left[\frac{1}{2}\left(\rho_{01(h)} + \rho_{02(h)}\right) - \frac{1}{8}\left(C_{1(h)}^2 + C_{2(h)}^2\right)\right]$

3.2 Efficiency Comparisons

In this section, the efficiency of the proposed generalized estimator $t_{z(st)}$ Is compared with the efficiency of some related estimators. Thus, conditions under which the proposed estimator is more efficient are given below:

(i) The proposed estimator $t_{z(st)}$ is more efficient than the sample variance estimator $T_{1(st)}$ if

$$MSE\left(T_1\right) - MSE\left(t_{z(st)}\right) > 0$$

$$\sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 C_{0(h)}^2 - \sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \left\{C_{0(h)}^2 + \Psi_h\right\} > 0$$

$$\sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y(h)}^4 \Psi_h < 0 \qquad (57)$$

where $\Psi_h = \left(\beta_{2(h)} - \frac{\beta_{3(h)}}{2}\right)^2 \left[C_{1(h)}^2 + C_{2(h)}^2 + 2\rho_{12(h)}\right] - 2\left(\beta_{2(h)} - \frac{\beta_{3(h)}}{2}\right)\left[\rho_{01(h)} + \rho_{02(h)}\right]$

If condition (57) is satisfied, the proposed estimator $t_{z(st)}$ is more efficient than$T_{1(st)}$.

(ii) The proposed estimator $t_{z(st)}$ is more efficient than the usual ratio variance estimator $T_{2(st)}$ if

$$MSE\left(T_{2(st)}\right) - MSE\left(t_{z(st)}\right) > 0$$

$$\sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \left[C_{0(h)}^2 + C_{1(h)}^2 - 2\rho_{01(h)}\right] - \sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y(h)}^4 \left\{C_{0(h)}^2 + \Psi_h\right\} > 0$$

$$\sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y(h)}^4 \Psi_h < \sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \left[C_{1(h)}^2 - 2\rho_{01(h)}\right] \qquad (58)$$

If condition (58) is satisfied, the proposed estimator $t_{z(st)}$ is more efficient than $T_{2(st)}$.

(iii) The proposed estimator $t_{z(st)}$ is more efficient than the usual product variance estimator $T_{3(st)}$ if

$$MSE\left(T_{3(st)}\right) - MSE\left(t_{z(st)}\right) > 0$$

$$\sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \left[C_{0(h)}^2 + C_{1(h)}^2 + 2\rho_{01(h)}\right] - \sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y(h)}^4 \left\{C_{0(h)}^2 + \Psi_h\right\} > 0$$

$$\sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y(h)}^4 \Psi_h < \sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \left[C_{1(h)}^2 + 2\rho_{01(h)}\right] \qquad (59)$$

If condition (59) is satisfied, the proposed estimator $t_{z(st)}$ is more efficient than $T_{3(st)}$.

(iv) The proposed estimator $t_{z(st)}$ is more efficient than the Rao variance estimator $T_{Rao(st)}$ if

$$MSE\left(T_{Rao(st)}\right)_{\min} - MSE\left(t_{z(st)}\right) > 0$$

$$\sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \left[C_{0(h)}^2 C_{1(h)}^2 + C_{1(h)}^2 - M_h\right] - \sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y(h)}^4 \left\{C_{0(h)}^2 + \Psi_h\right\} > 0$$

$$\sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y(h)}^4 \left\{C_{0(h)}^2 + \Psi_h\right\} < \sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \left[C_{0(h)}^2 C_{1(h)}^2 + C_{1(h)}^2 - M_h\right] \qquad (60)$$

where $M_h = \frac{\rho_{01(h)}^2}{f_{1(h)}\left\{C_{0(h)}^2 C_{1(h)}^2 - \rho_{01(h)}^2\right\}}$

If condition (60) is satisfied, the proposed estimator $t_{z(st)}$ is more efficient than $T_{Rao(st)}$.

(v) The proposed estimator $t_{z(st)}$ is more efficient than Singh et al. (2009) ratio estimator $T_{er(st)}$ if

$$MSE\left(T_{er(st)}\right) - MSE\left(t_{z(st)}\right) > 0$$

$$\sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \left[C_{0(h)}^2 + \frac{1}{4} C_{1(h)}^2 - \rho_{01(h)} \right] - \sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y(h)}^4 \left\{ C_{0(h)}^2 + \Psi_h \right\} > 0$$

$$\sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \Psi_h < \sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \left[\frac{1}{4} C_{1(h)}^2 - \rho_{01(h)} \right] \quad (61)$$

If condition (61) is satisfied, the proposed estimator $t_{z(st)}$ is more efficient than $T_{er(st)}$.

(vi) The proposed estimator $t_{z(st)}$ is more efficient than Singh et al. (2009) product estimator $T_{ep(st)}$ if

$$MSE\left(T_{ep(st)}\right) - MSE\left(t_{z(st)}\right) > 0$$

$$\sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \left[C_{0(h)}^2 + \frac{1}{4} C_{1(h)}^2 + \rho_{01(h)} \right] - \sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \left\{ C_{0(h)}^2 + \Psi_h \right\} > 0$$

$$\sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \Psi_h < \sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \left[\frac{1}{4} C_{1(h)}^2 + \rho_{01(h)} \right] \quad (62)$$

If condition (62) is satisfied, the proposed estimator $t_{z(st)}$ is more efficient than $T_{ep(st)}$.

(vii) The proposed estimator $t_{z(st)}$ is more efficient than Shabbir and Gupta (2010) estimator $T_{SG(st)}$ if

$$MSE\left(T_{SG(st)}\right)_{\min} - MSE\left(t_{z(st)}\right) > 0$$

$$\sum_{h=1}^{G} W_h^4 f_{1(h)}^2 S_{y_h}^4 \left[\frac{K_h Q_h}{K_h + Q_h} \right] - \sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \left\{ C_{0(h)}^2 + \Psi_h \right\} > 0$$

$$\sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \left\{ C_{0(h)}^2 + \Psi_h \right\} < \sum_{h=1}^{G} W_h^4 f_{1(h)}^2 S_{y_h}^4 \left[\frac{K_h Q_h}{K_h + Q_h} \right] \quad (63)$$

where $K_h = \left(1 - f_{1(h)} \tau_{i(h)}^2 C_{1(h)}^2\right)$ and $Q_h = \left(f_{1(h)} C_{0(h)}^2 \left(1 - \rho_h^2\right)\right)$

If condition (63) is satisfied, the proposed estimator $t_{z(st)}$ is more efficient than $T_{SG(st)}$.

(viii) The proposed estimator $t_{z(st)}$ is more efficient than Koyuncu (2013) estimator $T_{Ni(st)}$ if

$$MSE\left(T_{Ni(st)}\right)_{min} - MSE\left(t_{z(st)}\right) > 0$$

$$\sum_{h=1}^{G} W_h^4 S_{y_h}^4 \left[1 - \frac{BD^2 - DGF + AG^2}{\left(4AB - F^2\right)} \right] - \sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \left\{ C_{0(h)}^2 + \Psi_h \right\} > 0$$

$$\sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \left\{ C_{0(h)}^2 + \Psi_h \right\} < \sum_{h=1}^{G} W_h^4 S_{y_h}^4 \left[1 - \frac{BD^2 - DGF + AG^2}{\left(4AB - F^2\right)} \right] \quad (64)$$

If condition (64) is satisfied, the proposed estimator $t_{z(st)}$ is more efficient than $T_{Ni(st)}$.

(ix) The proposed estimator $t_{z(st)}$ is more efficient than Yasmeen and Noor (2021) ratio estimator $T_{YN1(st)}$ if

$$MSE\left(T_{YN1(st)}\right) - MSE\left(t_{z(st)}\right) > 0$$

$$\sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \left[C_{0(h)}^2 + \phi_{i(h)}^2 C_{1(h)}^2 - 2\rho_{01(h)} \right] - \sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \left\{ C_{0(h)}^2 + \Psi_h \right\} > 0$$

$$\sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \Psi_h < \sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \left[\phi_{i(h)}^2 C_{1(h)}^2 - 2\rho_{01(h)}\right] \quad (65)$$

If condition (65) is satisfied, the proposed estimator $t_{z(st)}$ is more efficient than $T_{YN1(st)}$.

(x) The proposed estimator $t_{z(st)}$ is more efficient than Yasmeen and Noor (2021) exponential ratio estimator $T_{YN2(st)}$ if

$$MSE\left(T_{YN2(st)}\right) - MSE\left(t_{z(st)}\right) > 0$$

$$\sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \left[C_{0(h)}^2 + C_{1(h)}^2 + \frac{1}{2}C_{2(h)}^2 - 2\rho_{01(h)} - \rho_{12(h)} + \rho_{02(h)}\right] - \sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \left\{C_{0(h)}^2 + \Psi_h\right\} > 0$$

$$\sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \Psi_h < \sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \left[C_{1(h)}^2 + \frac{1}{2}C_{2(h)}^2 - 2\rho_{01(h)} - \rho_{12(h)} + \rho_{02(h)}\right] \quad (66)$$

If condition (66) is satisfied, the proposed estimator $t_{z(st)}$ is more efficient than $T_{YN2(st)}$.

(xi) The proposed estimator $t_{z(st)}$ is more efficient than Yasmeen and Noor's (2021) exponential product estimator $T_{YN3(st)}$ if

$$MSE\left(T_{YN3(st)}\right)_{\min} - MSE\left(t_{z(st)}\right) > 0$$

$$\sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \frac{L_h}{\left(1 - B_{(h)}D_{(h)}\right)} - \sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \left\{C_{0(h)}^2 + \Psi_h\right\} > 0$$

$$\sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \left\{C_{0(h)}^2 + \Psi_h\right\} < \sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \frac{L_h}{\left(1 - B_{(h)}D_{(h)}\right)} \quad (67)$$

If condition (67) is satisfied, the proposed estimator $t_{z(st)}$ is more efficient than $T_{YN3(st)}$.

(xii) The proposed estimator $t_{z(st)}$ is more efficient than Ahmad et al. (2022) estimator $T_{\alpha(st)}$ if

$$MSE\left(T_{a(st)}\right)_{\min} - MSE\left(t_{z(st)}\right) > 0$$

$$\sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 R_h - \sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \left\{C_{0(h)}^2 + \Psi_h\right\} > 0$$

$$\sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 \left\{C_{0(h)}^2 + \Psi_h\right\} < \sum_{h=1}^{G} W_h^4 f_{1(h)}^3 S_{y_h}^4 R_h \quad (68)$$

where $R_h = \dfrac{64\left(\gamma_h + 1\right) - f_{1(h)}\left(C_{1(h)}^4 / C_{0(h)}^2\right) - 16 f_{1(h)} C_{1(h)}^2 \left(\gamma_h + 1\right)}{64\left[(1/C_{0(h)}^2) + f_{1(h)}(\gamma_h + 1)\right]}$

If condition (68) is satisfied, the proposed estimator $t_{z(st)}$ is more efficient than $T_{\alpha(st)}$.

4.0 Empirical study

In this section, real-life datasets were used to assess the performance of the proposed generalized estimator of finite population variance over some existing estimators base on the criteria of bias, mean square error and percentage relative efficiency.

Dataset I [Source: Agricultural Statistics (1999), Washington, D.C.]

Tobacco: Area (hectares), yield and production (metric tons) in specified countries during 1998.

Y = Yield (Metric tons), *X* = Area (Hectare), *Z* = Production (Metric tons), Strata = Continents, N=103, n=75 and m=16 (sub-sample size)

Dataset II [Source: Statistical Abstracts of the United States (1999)]

Y = Number of immigrants admitted in the USA in 1996, X = Number of immigrants admitted in the USA in 1995, Z = Number of immigrants admitted in the USA in 1994, Strata = Regions, N=50, n=32 and m=12.

Table 1: Parameters of the Datasets I and II

Parameters	Dataset I				Dataset II			
	Stratum 1	Stratum 2	Stratum 3	Stratum 4	Stratum 1	Stratum 2	Stratum 3	Stratum 4
N_h	12	18	34	39	9	16	12	13
n_h	9	13	25	28	6	10	8	8
m_h	2	3	5	6	2	4	3	3
$\bar{Y}$	1.680833	2.045294	1.156765	1.607179	27994.56	15437.31	10081.92	21471.23
$\bar{X}$	8927.25	17041.76	10714.18	37919.85	21999.22	11234.12	8438.667	17462.69
$\bar{Z}$	10030.75	33018.59	15425.88	111161.9	24436.67	11309	9267.5	21106.08
$S_{y_h}^2$	0.2046629	0.5203515	0.3295741	0.8479734	2639749017	707626160	123187113	2948783276
$S_{x_h}^2$	306708635	392241621	675871615	7566352438	1754057959	336303714	149721398	2021436188
$S_{z_h}^2$	21441542	1799038367	1817497001	169241418740	2220517339	343840938	121798863	3191962013
$\rho_{xy(h)}$	-0.6844773	-0.138175	0.2102041	-0.1597427	0.9948531	0.9878827	0.9960449	0.9998627
$\rho_{xz(h)}$	0.8728871	0.8966936	0.9490304	0.2680932	0.999965	0.9940247	0.9925293	0.9998764
$\rho_{yz(h)}$	-0.4232355	0.1140993	0.3273912	0.000882502	0.9946205	0.9986502	0.992753	0.9998665
$\lambda_{400(h)}$	3.083168	3.003996	2.733141	10.79979	4.611014	5.118251	4.582649	10.08216
$\lambda_{040(h)}$	6.21362	3.524234	13.23582	14.96476	5.252527	5.298082	4.09224	10.06581
$\lambda_{004(h)}$	2.715187	3.898346	16.09757	31.23819	5.284834	5.030387	5.114095	10.094
$\lambda_{220(h)}$	4.101859	0.5206339	1.032232	0.3680147	4.914612	4.948398	4.319464	10.07394
$\lambda_{202(h)}$	1.901612	0.6434228	1.676001	0.08727824	4.929345	5.048059	4.834475	10.08803
$\lambda_{022(h)}$	3.009616	3.312781	12.61424	0.9804824	5.268636	5.043032	4.559991	10.07986

Dataset III [Source: CBN Statistical Bulletin (2020)]

Y = Crude Oil Price from 2006-2020, X = Crude Oil Production from 2006-2020, Z = Crude Oil Exportation from 2006-2020, Strata = Five years interval, N=180, n=110 and m=30.

Dataset IV [Source: Agricultural Statistics (1998) Washington, US.]

Y = Season average price of Apples (in $) per pound, by States in 1996, X = Season average price of Apples (in $) per pound, by States in 1995, Z = Season average price of Apples (in $) per pound, by States in 1994, Strata: Regions, N=36, n=20 and m=11.

Table 2: Parameters of the Dataset III and IV

Parameters	Dataset III			Dataset IV			
	Stratum 1	Stratum 2	Stratum 3	Stratum 1	Stratum 2	Stratum 3	Stratum 4
N_h	60	60	60	8	10	9	9
n_h	37	46	27	4	5	7	4
m_h	10	15	5	2	2	5	2
$\bar{Y}$	77.4905	98.3045	55.658	0.198875	0.1702	0.2617778	0.1853333
$\bar{X}$	2.252333	2.243	1.877833	0.189375	0.1452	0.2406667	0.1722222
$\bar{Z}$	1.802333	1.793	1.427833	0.191	0.1493	0.2126667	0.1391111
$S_{y_h}^2$	411.92	638.305	207.565	0.004875554	0.004225067	0.01001044	0.00362525
$S_{x_h}^2$	0.03417073	0.0171739	0.03086133	0.004975982	0.0028204	0.00750175	0.004039444
$S_{z_h}^2$	0.03417073	0.0171739	0.03086133	0.005474286	0.002096678	0.00412075	0.002094361
$\rho_{xy(h)}$	-0.0958291	0.4648092	0.2640836	0.9760661	0.9627432	0.7857673	0.7567893
$\rho_{xz(h)}$	-0.5119518	-0.3334791	0.6405571	0.9918342	0.8683413	0.9074752	0.7033315

$\rho_{yz(h)}$	-0.3956622	0.5633014	0.3250112	0.9732945	0.8941405	0.7514282	0.9345371
$\lambda_{400(h)}$	4.421488	2.765368	2.683712	1.806926	3.200179	2.443142	2.382737
$\lambda_{040(h)}$	3.726242	2.389401	2.591385	1.809635	2.649353	2.527346	2.828628
$\lambda_{004(h)}$	3.726242	2.389401	2.591385	1.717057	1.524835	3.4034	1.997404
$\lambda_{220(h)}$	1.180103	0.9149597	0.8880379	1.729715	2.819442	2.24035	2.433403
$\lambda_{202(h)}$	1.14228	0.9138719	0.8880379	1.726914	1.777395	2.494846	1.977447
$\lambda_{022(h)}$	1345.015	482.8629	4771.051	1.74932	1.68456	2.844426	2.117876

Table 3: Bias, MSE and PRE of the Proposed and Existing Estimators using Dataset I & II

	Dataset I				Dataset II			
Estimator	**Bias**	**MSE**	**PRE**	**Rank**	**Bias**	**MSE**	**PRE**	**Rank**
$T_{1(st)}$	-	1.5598e-07	100.00	3rd	-	4.578572e+13	100.00	14th
$T_{2(st)}$	0.000255519	4.1309e-07	37.76	16th	118288.3	187424506292	24428.89	2nd
$T_{3(st)}$	-4.4957e-06	3.5264e-07	44.23	15th	5657379.0	1.846211e+14	24.80	16th
$T_{Rao(st)}$	0.001072130	7.9586e-07	19.60	17th	11653263.7	1.746581e+14	26.21	17th
$T_{er(st)}$	9.6382e-05	2.2534e-07	69.22	14th	-662814.3	1.134503e+13	403.58	8th
$T_{ep(st)}$	-3.3626e-05	2.0501e-07	76.08	12th	2106731.0	1.035095e+14	44.23	15th
$T_{SG(st)}$	0.176261600	2.1906e-07	71.21	13th	-309389.2	5.635698e+11	8124.23	3rd
$T_{N1(st)}$	0.001006607	1.6416e-07	95.02	6th	30583599.0	3.511147e+13	130.40	11.5th
$T_{N2(st)}$	0.001073261	1.6617e-07	93.87	8th	30500565.0	3.241526e+13	142.91	10th
$T_{N3(st)}$	0.001073219	1.6403e-07	95.09	5th	30534013.0	2.335528e+13	196.04	9th
$T_{N4(st)}$	0.001073214	1.6544e-07	94.29	7th	31066211.0	3.511147e+13	130.40	11.5th
$T_{N5(st)}$	0.001073223	1.6219e-07	96.18	4th	31533567.0	3.556147e+13	130.11	13th
$T_{YN1(st)}$	0.003319547	1.8641e-07	83.68	10th	1005680.0	3.900567e+12	1173.82	4th
$T_{YN2(st)}$	0.002770780	1.7797e-07	87.65	9th	-24712808.0	4.293213e+12	1066.47	5th
$T_{YN3(st)}$	0.000306262	1.9772e-07	78.89	11th	-104685.0	1.126826e+12	1016.65	6th
$T_{\alpha(st)}$	4.2621e-05	1.5579e-07	100.13	2nd	-566972.7	8.537835e+12	536.27	8th
$t_{z(st)}$	1.9883e-21	1.5474e-07	100.80	1st	4.2191e-10	7.150700e+10	64029.71	1st

Table 3 presents the results of the bias, mean square error and percentage relative efficiency of proposed generalized estimator and some existing estimators using datasets I and II. The fourth column of Table 3 indicates the rank of the estimators. Based on the results obtained from Dataset I, it is observed that the proposed generalized estimator ($t_{z(st)}$) have minimum bias and mean square error, and higher percentage relative efficiency values (1.9883e-21, 1.5474e-07 and 100.80, respectively) compared to the usual variance ($T_{1(st)}$); Isaki (1983) classical ratio and product estimators ($T_{2(st)}$ and $T_{3(st)}$); Rao (1991) regression type estimator ($T_{Rao(st)}$); Singh *et al.*, (1988) exponential ratio and product type estimators ($T_{er(st)}$ and $T_{ep(st)}$); Shabbir and Gupta (2007) ratio-regression-type estimator ($T_{SG(st)}$); Koyuncu (2013) estimators ($T_{N1(st)}$, $T_{N2(st)}$, $T_{N3(st)}$, $T_{N4(st)}$ and $T_{N5(st)}$); Yasmeen and Noor (2021) estimators ($T_{YN1(st)}$, $T_{YN2(st)}$ and $T_{YN3(st)}$); Ahmad et al. (2022) estimator ($T_{\alpha(st)}$). The results obtained from Dataset II also revealed that the proposed generalized estimator; ($t_{z(st)}$) have minimum bias and mean square error, and higher percentage relative efficiency values (4.2191e-10, 7.150700e+10 and 64029.71, respectively) compared to some existing estimators considered. Therefore, the proposed generalized estimator performs better based on the criteria of bias, mean square error and percentage relative efficiency values when the study and auxiliary variables are negatively correlated (dataset I), and when the study and auxiliary variables are positively correlated (dataset II).

Table 4: Bias, MSE and PRE of the Proposed and Existing Estimators using Dataset III & IV

	Dataset III				Dataset IV			
Estimator	**Bias**	**MSE**	**PRE**	**Rank**	**Bias**	**MSE**	**PRE**	**Rank**
$T_{1(st)}$	-	7.992012	100.00	2nd	-	5.334586e-10	100.00	12th
$T_{2(st)}$	1.440524	13.89579	57.51	13th	1.654535e-06	9.370419e-11	569.30	3rd
$T_{3(st)}$	0.050222	14.78973	54.04	15th	1.572381e-05	2.084210e-09	25.60	17th
$T_{Rao(st)}$	4.829258	19.56236	40.85	16th	0.0005428436	1.174455e-09	45.42	16th
$T_{er(st)}$	0.533929	9.356213	85.42	9th	-1.345026e-06	1.791798e-10	297.72	10th
$T_{ep(st)}$	-0.161232	9.803185	81.52	11th	5.689613e-06	1.156541e-09	46.13	15th
$T_{SG(st)}$	27.98110	11.57255	69.06	12th	4.157174e-05	1.309428e-10	407.40	8th
$T_{N1(st)}$	0.019113	8.069006	99.05	3rd	-2.315270e-06	1.118705e-10	476.85	6th
$T_{N2(st)}$	0.463060	8.258617	96.77	6th	-6.087897e-07	9.678856e-11	551.16	4th
$T_{N3(st)}$	-0.021755	8.078534	98.93	5th	-7.847469e-07	1.344898e-10	396.65	9th
$T_{N4(st)}$	0.564476	8.267766	96.67	7th	-5.162108e-07	9.123151e-11	584.73	2nd
$T_{N5(st)}$	0.006718	8.076830	98.95	4th	-7.665384e-07	1.115043e-10	478.42	5th
$T_{YN1(st)}$	2.120311	24.36189	32.81	17th	8.245939e-06	1.156776e-10	461.16	7th
$T_{YN2(st)}$	2.571675	9.545700	83.72	10th	-4.463605e-05	3.023131e-10	176.46	11th
$T_{YN3(st)}$	-2.644595	8.529883	93.69	8th	2.112711e-05	5.437352e-10	98.11	13th
$T_{\alpha(st)}$	-6.358264	14.08540	56.74	14th	-4.021615e-06	9.213379e-10	57.90	14th
$t_{z(st)}$	-6.937e-21	7.991911	100.01	1st	9.249550e-22	8.154949e-11	654.15	1st

Table 4 presents the results of the bias, mean square error and percentage relative efficiency of the proposed generalized estimator and some existing estimators using datasets III and IV. The fourth column of Table 4 indicates the rank of the estimators. Based on the results obtained from Dataset III, it is observed that the proposed generalized estimator $(t_{z(st)})$ have minimum bias and mean square error, and higher percentage relative efficiency values (-6.937e-21, 7.991911 and 100.01, respectively) compared to the usual variance $(T_{1(st)})$; Isaki (1983) classical ratio and product estimators ($T_{2(st)}$ and $T_{3(st)}$); Rao (1991) regression type estimator $(T_{Rao(st)})$; Singh *et al.*, (1988) exponential ratio and product type estimators ($T_{er(st)}$ and $T_{ep(st)}$); Shabbir and Gupta (2007) ratio-regression-type estimator $(T_{SG(st)})$; Koyuncu (2013) estimators ($T_{N1(st)}$, $T_{N2(st)}$, $T_{N3(st)}$, $T_{N4(st)}$ and $T_{N5(st)}$); Yasmeen and Noor (2021) estimators ($T_{YN1(st)}$, $T_{YN2(st)}$ and $T_{YN3(st)}$); Ahmad et al. (2022) estimator $(T_{\alpha(st)})$.. The results obtained from Dataset IV also revealed that the proposed generalized estimator; $(t_{z(st)})$ have minimum bias and mean square error, and higher percentage relative efficiency values (9.249550e-22, 8.154949e-11 and 654.15, respectively) compared to some existing estimators considered. Therefore, the proposed generalized estimator performs better based on the criteria of bias, mean square error and percentage relative efficiency values when the study and auxiliary variables are negatively correlated (dataset III), and when the study and auxiliary variables are positively correlated (dataset IV).

5.0 Conclusion

The study evaluated the performance of a proposed generalized estimator in comparison with several existing estimators using datasets with varying correlation structures between study and auxiliary variables. The results from Dataset III, where the variables were negatively correlated, showed that the proposed estimator had the least bias and mean square error, alongside the highest percentage relative efficiency. Similarly, when Dataset IV, characterized by positively correlated variables, was analyzed, the proposed estimator also demonstrated superior performance in terms of bias, mean square error, and efficiency. These findings underscore the robustness and reliability of the proposed estimator in diverse correlation scenarios.

In conclusion, the proposed generalized estimator outperforms the existing methods across multiple criteria, making it a more effective and reliable choice for applications requiring accurate parameter estimation. Its superior performance in both negatively and positively correlated datasets suggests its versatility and potential for broader applicability in statistical analysis.

6.0 Based on these findings, it is recommended that the proposed generalized estimator be adopted for use in studies where minimizing bias and error and maximizing efficiency are critical. Future research could further explore its performance across a wider range of datasets and real-world applications to validate its utility in practical scenarios.

6.0 References

Ahmad, S., Hussain, S., Shabbir, J., Zahid, E., Aamir, M. & Onyango, R. (2022). Improve Estimation of Finite Population Variance Using Dual Supplementary Information under Stratified Random Sampling. *Mathematical Problems in Engineering*, 3, 2, pp. 12-19.

Arata, L., Fabrizi, E., & Sckokai, P. (2020). A worldwide analysis of trends in crop yields and yield variability: Evidence from FAO data. *Economic Modelling, 90*, 190–208. https://doi.org/10.1016/j.econmod.2020.05.006.

Audu, A., Zakari, Y., Yunusa, M. A., Ishaq, O. O., Manu, F., & Muhammad, I. (2023). Regression-Cum-Ratio mean imputation class of estimators using non-conventional robust measures. *Caliphate Journal of Science & Technology (CaJoST)*, 3, pp. 246-254.

Cochran, W. G. (1977). *Sampling techniques*. John Wiley and Sons.

Das, A. K. & Tripathi, T. P. (1978). Use of auxiliary information in estimating the finite population variance. *Sankhya C,* 40, 2, pp. 139–148.

Isaki, C. T. (1983). Variance estimation using auxiliary information. *Journal of American Statistics Association*, 78, pp.117–123.

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

Koyuncu, N. (2013). Improved Estimators of Finite Population Variance in Stratified Random Sampling. *World Applied Sciences Journal*, 23, 12, pp. 130-137.

Miller F. (2005) Variance estimation in clinical studies with interim sample size re-estimation. *Biometrics*. 61, 2, pp. 355-61. doi: 10.1111/j.1541-0420.2005.00315.x.

Muhammad, I. (2023). Generalized Ratio-Product cum Regression Variance Estimator in Two-Phase Sampling. *CBN*

Journal of Applied Statistics, 14, 2, pp. 73-101. DOI: 10.33429/Cjas.14223.4/5

Muhammad, I., Zakari, Y., & Audu, A. (2022). Generalized estimators for finite population variance using measurable and affordable auxiliary character. *Asian Research Journal of Mathematics*, 18, 1, pp. 14-30.

Muhammad, I., Zakari, Y., & Audu, A. (2021). An alternative class of ratio-regression type estimator under a two-phase sampling scheme. *CBN Journal of Applied Statistics*, 12, 2, pp. 1-26.

Muhammad, I., Zakari, Y., Abdu, M., Iliyasu, R., Suleiman, M., Manzo, S., Muhammad, A. & Zakar, A. A. (2023). Enhanced ratio-type estimator for finite population means using auxiliary variable in simple random sampling. *NIPES Journal of Science and Technology Research*, 5, 1, pp. 242-252.

Muili, J. O., Audu, A., Singh, R.V.K. & Yunusa, I. A. (2019). Improved estimators of finite population variance using unknown weight of auxiliary variable. *Annals. Computer Science Series*, 17, 1, pp. 148-153.

Oyeyemi, G. M., Muhammad, I. & Kareem, A. O. (2023). Combined exponential-type estimators for finite population mean in two-phase sampling. *Asian Journal of Probability and Statistics*, 21, 2, pp. 44-58.

Rao, T. J. (1991). On certail methods of improving ratio and regression estimators. *Communication in Statistics-Theory and Methods*, 20, 10, pp. 3325–3340.

Shabbir, J. & Gupta, S. (2010). Some estimators of finite population variance of stratified sample mean. *Communications in Statistics - Theory and Methods*, 39, 16, pp. 001–3008.

Singh, G. N., Bhattacharyya, D. & Bandyopadhyay, A. (2020). A general class of calibration estimators under stratified random sampling in the presence of various kinds of non-sampling errors, *Communications in Statistics - Simulation and Computation*, DOI: 10.1080/03610918.2020.1855447

Singh, R., Chauhan, P., Sawan, N. & Smarandache, F. (2009). Improvement in estimating the population mean using exponential estimator in stratified random sampling. *International Journal of Statistics & Economics*, 3, 9, pp. 13–18.

Singh, B. K., Choudhury, S. & Kalitac, D. (2013). A Class of Exponential Chain Ratio-Product Type Estimator with Two Auxiliary Variables under Double Sampling Scheme. *Electronic Journal of Applied Statistical Analysis*, 6, 2, pp. 166-174.

Tato, Y. & Singh, B. K. (2017). Exponential Chain Dual to Ratio cum Dual to Product Estimator for Finite Population Mean in Double Sampling Scheme. *Applications and Applied Mathematics*, 12, 1, pp. 46-69.

Yasmeen, U. & Noor, M. (2021). Estimation of finite population variance under stratified sampling technique. *Journal of Reliability and Statistical Studies*, 14, 2, pp. 565-584.

Zakari, Y & Muhammad, I. (2022). *Improved product estimator using measurable auxiliary character*. Proceedings of the 58th Annual National Conference of Mathematical Association of Nigeria. Published by the *Mathematical Association of Nigeria (MAN)*.

Zakari, Y., & Muhammad, I. (2023). Modified Estimator of Finite Population Variance under Stratified Random Sampling. *Engineering Proceedings*, *56*(1), 177. https://doi.org/10.3390/ASEC2023-16308

Zakari, Y., Muhammad, I., & Sani, N. M. (2020). Alternative ratio-product type estimator in simple random sampling. *Communication in Physical Sciences*, 5, 4, 418-426.

Zakari, Y., Yunusa, M. A., Abdulrahman, R., Muhammad, I., Mohammed, A. S., &

Audu, A. (2023). The Efficiency of a Ratio Product Estimator in the Estimation of the Finite Population Coefficient of Variation. *Engineering Proceedings*, *5*, *1*, 305. https://doi.org/10.3390/ASEC2023-16263

Compliance with Ethical Standards Declaration

Ethical Approval

Not Applicable

Competing interests

The authors declare that they have no known competing financial interests

Funding

The authors declared no source of funding

Authors' Contribution

All the authors were involved in the design, experimental, analysis and reporting of the work.