On the Applicability and Efficiency of Calibrated Variance Estimators Utilizing Auxiliary Information Under Stratified Random Sampling

MOJEED ABIODUN YUNUSA[1], AWWAL ADEJUMOBI[2], AHMED AUDU[1]
[1]Department of Statistics,
Usmanu Danfodiyo University, Sokoto,
NIGERIA

[2]Department of Mathematics,
Abdullahi Fodio University of Science and Technology, Aliero,
NIGERIA

Abstract: - Estimating variance is a common practice among statisticians particularly, survey statisticians. The estimate of variance gives information about the spread within the dataset, and allows the statisticians make vital decision about the event at hand at that moment. Existing variance estimators are function of unknown constants whose applicability is not possible in real life situation. In this paper, we intend to modify some existing variance estimators using coefficient of variation and kurtosis of auxiliary variable as auxiliary information that are free of unknown constants, and are applicable in real life situation. Optimization approach (Lagrange function) was used to obtain the calibration weights of newly proposed estimators and therefore the proposed calibration variance estimators were obtained. A simulation study was used to prove the better efficiency (minimum mean square error) of these proposed estimators.

Key-Words: - Calibration, Estimate, Estimator, Sampling, Weights, Sample.

Received: April 11, 2026. Revised: May 4, 2026. Accepted: June 5, 2026. Published: July 16, 2026.

1 Introduction

It is very important to understand the variability in a dataset, so as to make a remarkable decision based on the situation on ground, for example, a classroom teacher need to know on the average in the variability that exist in the performance of students in a certain subject to know how class his or her students and look for wats of helping the weak ones among them. There are several variability measures among them is the variance. The sample variance estimator is used when the cost of collecting information from all member of the population is expensive and when the time involved in collecting this information is much. The sample variance estimator without the use of auxiliary information, is an unbiased estimator of the population variance of the study variable. Auxiliary information is used to enhance the performance of estimators. The first mam to propose variance estimator with the use of auxiliary information is Isaki (1983). After him several authors have developed variance estimators for estimation of population variance of the study variable, they include, Upadhyaya and Singh (1998), Singh et al. (2011), Khan and Shabbir (2013), Shabbir and Gupta (2015), Audu et al. (2016), Yunusa et al. (2022), Adejumobi and Yunusa (2022), Adejumobi and Yunusa (2023), Audu et al. (2023), etc.

The variance estimator has significant application in real life situation, for example, in drink production industry, all the content in the bottles will never be the same. This difference is what we call variance. We want to know the average in order to draw conclusion. We estimate the variability of the content of the drinks to see whether it is within, below or above standard for the drinks. Variance is the difference that normally happens in an event. As it is known standard is set for all product, so it is important we know whether this average variability is within, below or above specification. Another example, verifying the average variability of level of parastamol in a drug. Is the average level of parastamol within, below or above the standard. If the standard set for the level is 2.9 to 3.2, if the average level of parastamol in the drug is 2.4 then the drug will not work for the body. Again, if the average level of parastamol in the drug is 3.4 then the drug will be harmful to the body. So, if our estimator produce estimate of 3.09, the level of parastamol in the drug is okay.

Several researchers have proposed variance estimators under the framework of stratified random sampling, because estimators under stratified random sampling are efficient than estimators under simple random sampling approach, among them are Prasad (1992), Qzel et al. (2014), Muili et al. (2018), etc., the weight in stratified sampling estimators were modified by authors to give better estimate of population variance through the use of calibration approaches.

Calibration is a common method used by survey statisticians to develop population mean, population variance, population coefficient of variation estimators to produce calibration weights which in turn are used to obtain calibrated estimators through the use of optimization method that consists of Chi-square distance function, calibration constraints in the presence of auxiliary information incorporation. Hence, the calibration weights to be obtained will be a function of auxiliary information which will make the estimator produce a stable estimate. Authors have used this technique for developing population parameter estimators, among them include, Deville and Sarndal (2002), Kim et al. (2007), Audu et al (2021), Abubakar et al. (2024), Audu et al. (2025), Yunusa et al. (2026), etc. The calibrated variance estimators are often preferred over the ordinary ratio estimators because it can incorporate auxiliary information to improve efficiency and reduce bias. By adjusting weights to match known population variance.

The problem with the existing work of Muili et al. (2018) is that their estimators are function of unknown constants such as k_{1h}, k_{2h}, , k_{3h}, k_{4h}, k_{5h}, k_{6h}, hence the applicability of their estimators in real life situation is not possible. As a result of this identified problem prompted the current study.

In this current study, we intend to modify the variance ratio estimators proposed by Muili et al. (2018) due to lack of applicability of their estimators in real life practice. We aimed at proposing calibrated variance ratio estimators with the use of auxiliary information that are free from unknown constants and which can be applied in real life practice. The essence of proposing the calibration variance estimators is that, it reduces variance, that is, by using auxiliary information, they can produce more precise estimate. Again they improve accuracy, calibration estimators ensures estimates align with known population characteristics and they handle non-response, that is, it can adjust for biases due to missing data.

1.1 Statistical notations used in the study

Let a population of N units be stratified into populations, N_h, $h = 1, 2, ..., L$, such that N_1 N_2 ….. N_L are subpopulations in each stratum then we define the following notations

Y_{hj}: is the value of the study variable in the h^{th} stratum of the j^{th} element

X_{hj}: is the value of the auxiliary variable in the h^{th} stratum of the j^{th} element

$\bar{y}_h = \frac{1}{n_h}\sum_{i=1}^{n_h} y_{hj}$: is the sample mean of the study variable in the h^{th} stratum of the j^{th} element

$\bar{x}_h = \frac{1}{n_h}\sum_{i=1}^{n_h} x_{hj}$: is the sample mean of the auxiliary variable in the h^{th} stratum of the j^{th} element

$\bar{Y}_h = \frac{1}{N_h}\sum_{i=1}^{N_h} Y_{hj}$: is the population mean of the study variable in the h^{th} stratum of the j^{th} element

$\bar{X}_h = \frac{1}{N_h}\sum_{i=1}^{N_h} X_{hj}$: is the population mean of the auxiliary variable in the h^{th} stratum of the j^{th} element

$s_{yh}^2 = \frac{1}{n_h - 1}\sum_{i=1}^{n_h}\left(y_{hj} - \bar{y}_h\right)^2$: is the sample variance of the study in the h^{th} stratum of the j^{th} element

$s_{xh}^2 = \frac{1}{n_h - 1}\sum_{i=1}^{n_h}\left(x_{hj} - \bar{x}_h\right)^2$: is the sample variance of the auxiliary variable in the h^{th} stratum of the j^{th} element

$S_{yh}^2 = \frac{1}{N_h - 1}\sum_{i=1}^{N_h}\left(Y_{hj} - \bar{Y}_h\right)^2$: is the population variance of the study variable in the h^{th} stratum of the j^{th} element

$S_{xh}^2 = \frac{1}{N_h - 1}\sum_{i=1}^{N_h}\left(X_{hx} - \bar{X}_h\right)^2$: is the population variance of the auxiliary variable in the h^{th} stratum of the j^{th} element

2 Literature Review on Estimators

2.1 Existing Variance Estimators under Stratified Random Sampling

The sample variance estimator of the study variable with no auxiliary information under stratified random sampling is given by

$$t_0 = \sum_{h=1}^{L} \frac{W_h^2}{n_h} s_{yh}^2$$

(2.1)

The variance of the estimator is given by

$$\operatorname{var}(t_0) = \sum_{h=1}^{L} \frac{W_h^4}{n_h^2} S_{yh}^4 \left(\beta_{2(y)h} - 1\right)$$

(2.2)

Prasad and Singh (1992) proposed variance estimator in stratified sampling as

$$t_1 = \sum_{h=1}^{L} \frac{W_h^2}{n_h} \left(s_{yh}^2 - \frac{s_{xh}^2}{S_{xh}^2} + 1 \right)$$

(2.3)

The mean square error of the estimator are given by

$$\operatorname{var}(t_1) = \sum_{h=1}^{L} \frac{W_h^4}{n_h^2} S_{yh}^4 \left\{ \left(\beta_{2(y)h} - 1\right) + \frac{\left(\beta_{2(x)h} - 1\right)}{S_{yh}^4} - \frac{\left(\lambda_{22h} - 1\right)}{S_{yh}^2} \right\}$$

(2.4)

Muili et al. (2018) proposed some finite population variance estimators in stratified random sampling taking motivation from Audu et al. (2016) as

$$t_{j1} = \sum_{h=1}^{L} \frac{W_h^2}{n_h} s_{yh}^2 k_{1h} \left(\frac{S_{xh}^2}{s_{xh}^2} \right)$$

(2.5)

$$t_{j2} = \sum_{h=1}^{L} \frac{W_h^2}{n_h} s_{yh}^2 k_{2h} \left(\frac{S_{xh}^2 - C_{xh}}{s_{xh}^2 - C_{xh}} \right)$$

(2.6)

$$t_{j3} = \sum_{h=1}^{L} \frac{W_h^2}{n_h} s_{yh}^2 k_{3h} \left(\frac{S_{xh}^2 - \beta_{2(x)h}}{s_{xh}^2 - \beta_{2(x)h}} \right)$$

(2.7)

$$t_{j4} = \sum_{h=1}^{L} \frac{W_h^2}{n_h} s_{yh}^2 k_{4h} \left(\frac{\beta_{2(x)h} S_{xh}^2 - C_{xh}}{\beta_{2(x)h} s_{xh}^2 - C_{xh}} \right)$$

(2.8)

$$t_{j5} = \sum_{h=1}^{L} \frac{W_h^2}{n_h} s_{yh}^2 k_{5h} \left(\frac{C_{xh} S_{xh}^2 - \beta_{2(x)h}}{C_{xh} s_{xh}^2 - \beta_{2(x)h}} \right)$$

(2.9)

$$t_{j6} = \sum_{h=1}^{L} \frac{W_h^2}{n_h} s_{yh}^2 k_{6h} \left(\frac{S_{xh}^2 + \beta_{2(x)h}}{s_{xh}^2 + \beta_{2(x)h}} \right)$$

(2.10)

The mean square error of the estimator are given by

$$MSE(t_{ji}) = \sum_{h=1}^{L} \frac{W_h^4}{n_h^2} S_{yh}^4 \left\{ 1 - \frac{\left(1 + \theta_h \left(t_{ih}^2 \left(\beta_{2(x)h} - 1\right) - t_{ih}\left(\lambda_{22h} - 1\right)\right)\right)^2}{1 + \theta_h \left(\left(\beta_{2(x)h} - 1\right) + 3t_{ih}^2 \left(\beta_{2(x)h} - 1\right) - 4t_{ih}\left(\lambda_{22h} - 1\right)\right)} \right\}$$

(2.11)

Where,

$$k_{ih} = \frac{1 + \theta_h \left(t_{ih}^2 \left(\beta_{2(x)h} - 1\right) - t_{ih}\left(\lambda_{22h} - 1\right)\right)}{1 + \theta_h \left(\left(\beta_{2(x)h} - 1\right) + 3t_{ih}^2 \left(\beta_{2(x)h} - 1\right) - 4t_{ih}\left(\lambda_{22h} - 1\right)\right)}, \ i = 1,2,3,4,5,6$$

$$t_{1h} = 1, \qquad t_{2h} = \frac{S_{xh}^2}{S_{xh}^2 - C_{xh}}, \ t_{3h} = \frac{S_{xh}^2}{S_{xh}^2 - \beta_{2(x)h}},$$

$$t_{4h} = \frac{\beta_{2(x)h} S_{xh}^2}{\beta_{2(x)h} S_{xh}^2 - C_{xh}}, \ t_{5h} = \frac{C_x S_{xh}^2}{C_x S_{xh}^2 - \beta_{2(x)h}},$$

$$t_{6h} = \frac{S_{xh}^2}{S_{xh}^2 + \beta_{2(x)h}},$$

3. Proposed Estimators

Having studied the estimators of Muili et al. (2018) and we have identified the gap in their study. Hence, we proposed the following estimators in the following sections

3.1 First Proposed Class of Calibration Scheme and Variance Estimators

The first class of calibration estimator are defined as (3.1)

$$t_{p1i} = \sum_{h=1}^{L} \frac{\varpi_{hi}^2}{n_h} s_{yh}^2, \ i = 1,2,3,4,5,6$$

(3.1)

Where ϖ_h^2 is the new calibration weights such that the chi-square distance Z_1 is defined as

$$\left.\begin{aligned} Z_1 &= \sum_{h=1}^{L} \frac{\left(\varpi_{hi}^2 - W_h^2\right)^2}{n_h W_h^2 a_h} \\ \sum_{h=1}^{L} \frac{\varpi_{hi}^2}{n_h}\left(u(x)_{hi} s_{xh}^2 + v(x)_{hi}\right) &= \sum_{h=1}^{L} \frac{W_h^2}{n_h}\left(u(x)_{hi} S_{xh}^2 + v(x)_{hi}\right) \end{aligned}\right\}$$

(3.2)

Where, $\left(u(x)_{h1}=1, v(x)_{h1}=0\right)$,
$\left(u(x)_{h2}=1, v(x)_{h2}=-C_{xh}\right)$,
$\left(u(x)_{h3}=1, v(x)_{h3}=-\beta_{2(x)h}\right)$,
$\left(u(x)_{h4}=\beta_{2(x)h}, v(x)_{h4}=-C_{xh}\right)$,
$\left(u(x)_{h5}=C_{xh}, v(x)_{h5}=-\beta_{2(x)h}\right)$,
$\left(u(x)_{h6}=1, v(x)_{h6}=\beta_{2(x)h}\right)$

In order to obtain the new calibration weights ϖ_{hi}^2, the lagrange function is used as in (3.3)

$$L_1=\sum_{h=1}^{L}\frac{\left(\varpi_{hi}^2-W_h^2\right)^2}{2n_hW_h^2a_h}-g_{1i}\left\{\sum_{h=1}^{L}\frac{\varpi_{hi}^2}{n_h}\left(u(x)_{hi}\,s_{xh}^2+v(x)_{hi}\right)-\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(u(x)_{hi}\,S_{xh}^2+v(x)_{hi}\right)\right\}$$
(3.3)

By differentiating (3.17) partially, with respect to ϖ_{hi}^2 and g_{1i}, equating to zero and solve, we obtained equations (3.4) and (3.5)

$$\varpi_{hi}^2=W_h^2+g_{1i}W_h^2a_h\left(u(x)_{hi}\,s_{xh}^2+v(x)_{hi}\right)$$
(3.4)

$$\sum_{h=1}^{L}\frac{\varpi_{hi}^2}{n_h}\left(u(x)_{hi}\,s_{xh}^2+v(x)_{hi}\right)-\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(u(x)_{hi}\,S_{xh}^2+v(x)_{hi}\right)=0$$
(3.5)

Substituting equation (3.4) into equation (3.5) and solve for g_{1i}, the lagrange multiplier, we get equation (3.6)

$$g_{1i}=\frac{\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(u(x)_{hi}\,S_{xh}^2+v(x)_{hi}\right)-\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(u(x)_{hi}\,s_{xh}^2+v(x)_{hi}\right)}{\sum_{h=1}^{L}\frac{W_h^2}{n_h}a_h\left(u(x)_{hi}\,s_{xh}^2+v(x)_{hi}\right)^2}$$
(3.6)

By substituting equation (3.6) into equation (3.4) and putting $a_h=\left(u(x)_{hi}\,s_{xh}^2+v(x)_{hi}\right)^{-1}$, we get equation (3.7)

$$\varpi_{hi}^2=W_h^2+W_h^2\left\{\frac{\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(u(x)_{hi}\,S_{xh}^2+v(x)_{hi}\right)-\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(u(x)_{hi}\,s_{xh}^2+v(x)_{hi}\right)}{\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(u(x)_{hi}\,s_{xh}^2+v(x)_{hi}\right)}\right\}$$
(3.7)

Now, by substituting equation (3.7) into equation (3.1), we get equation (3.8)

$$t_{p1i}=\sum_{h=1}^{L}\frac{W_h^2}{n_h}s_{yh}^2\left(\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(u(x)_{hi}\,S_{xh}^2+v(x)_{hi}\right)\right)\left(\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(u(x)_{hi}\,s_{xh}^2+v(x)_{hi}\right)\right)^{-1}$$
(3.8)

So, the six members of the proposed estimator in equation (3.8) are

$$t_{p11}=\sum_{h=1}^{L}\frac{W_h^2}{n_h}s_{yh}^2\left(\sum_{h=1}^{L}\frac{W_h^2}{n_h}S_{xh}^2\right)\left(\sum_{h=1}^{L}\frac{W_h^2}{n_h}s_{xh}^2\right)^{-1}$$
(3.9)

$$t_{p12}=\sum_{h=1}^{L}\frac{W_h^2}{n_h}s_{yh}^2\left(\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(S_{xh}^2-C_{xh}\right)\right)\left(\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(s_{xh}^2-C_{xh}\right)\right)^{-1}$$
(3.10)

$$t_{p13}=\sum_{h=1}^{L}\frac{W_h^2}{n_h}s_{yh}^2\left(\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(S_{xh}^2-\beta_{2(x)h}\right)\right)\left(\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(s_{xh}^2-\beta_{2(x)h}\right)\right)^{-1}$$
(3.11)

$$t_{p14}=\sum_{h=1}^{L}\frac{W_h^2}{n_h}s_{yh}^2\left(\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(\beta_{2(x)h}S_{xh}^2-C_{xh}\right)\right)\left(\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(\beta_{2(x)h}s_{xh}^2-C_{xh}\right)\right)^{-1}$$
(3.12)

$$t_{p15}=\sum_{h=1}^{L}\frac{W_h^2}{n_h}s_{yh}^2\left(\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(C_{xh}S_{xh}^2-\beta_{2(x)h}\right)\right)\left(\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(C_{xh}s_{xh}^2-\beta_{2(x)h}\right)\right)^{-1}$$
(3.13)

$$t_{p16}=\sum_{h=1}^{L}\frac{W_h^2}{n_h}s_{yh}^2\left(\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(S_{xh}^2+\beta_{2(x)h}\right)\right)\left(\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(s_{xh}^2+\beta_{2(x)h}\right)\right)^{-1}$$
(3.14)

3.2 Second Proposed Class of Calibration Scheme and Variance Estimators

The second class of calibration estimator are defined as (3.15)

$$t_{p2i}=\sum_{h=1}^{L}\frac{\pi_{hi}^2}{n_h}s_{yh}^2,\ i=1,2,3,4,5,6$$
(3.15)

Where π_h^2 is the new calibration weights such that the chi-square distance Z_1 is defined as

$$\left.\begin{aligned}&Z_2=\sum_{h=1}^{L}\frac{\left(\pi_{hi}^2-W_h^2\right)^2}{n_hW_h^2b_h}\\&\sum_{h=1}^{L}\frac{\pi_{hi}^2}{n_h}\left(1+u(x)_{hi}\,s_{xh}^2+v(x)_{hi}\right)=\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(1+u(x)_{hi}\,S_{xh}^2+v(x)_{hi}\right)\end{aligned}\right\}$$
(3.16)

Where, $\left(u(x)_{h1}=1, v(x)_{h1}=0\right)$,
$\left(u(x)_{h2}=1, v(x)_{h2}=-C_{xh}\right)$,
$\left(u(x)_{h3}=1, v(x)_{h3}=-\beta_{2(x)h}\right)$,
$\left(u(x)_{h4}=\beta_{2(x)h}, v(x)_{h4}=-C_{xh}\right)$,
$\left(u(x)_{h5}=C_{xh}, v(x)_{h5}=-\beta_{2(x)h}\right)$,
$\left(u(x)_{h6}=1, v(x)_{h6}=\beta_{2(x)h}\right)$

In order to obtain the new calibration weights π_{hi}^2, the lagrange function is used as in (3.17)

$$L_2=\sum_{h=1}^{L}\frac{\left(\pi_{hi}^2-W_h^2\right)^2}{2n_hW_h^2b_h}-g_{2i}\left\{\sum_{h=1}^{L}\frac{\pi_{hi}^2}{n_h}\left(1+u\left(x\right)_{hi}s_{xh}^2+v\left(x\right)_{hi}\right)-\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(1+u\left(x\right)_{hi}S_{xh}^2+v\left(x\right)_{hi}\right)\right\}$$

(3.17)

By differentiating (3.17) partially, with respect to π_{hi}^2 and g_{2i}, equating to zero and solve, we obtained equation (3.18) and (3.19)

$$\pi_{hi}^2=W_h^2+g_{2i}W_h^2b_h\left(1+u\left(x\right)_{hi}s_{xh}^2+v\left(x\right)_{hi}\right)$$

(3.18)

$$\sum_{h=1}^{L}\frac{\pi_{hi}^2}{n_h}\left(1+u\left(x\right)_{hi}s_{xh}^2+v\left(x\right)_{hi}\right)-\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(1+u\left(x\right)_{hi}S_{xh}^2+v\left(x\right)_{hi}\right)=0$$

(3.19)

Substituting equation (3.18) into equation (3.19) and solve for g_{2i}, the lagrange multiplier, we get equation (3.20)

$$g_{2i}=\frac{\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(1+u\left(x\right)_{hi}S_{xh}^2+v\left(x\right)_{hi}\right)-\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(1+u\left(x\right)_{hi}s_{xh}^2+v\left(x\right)_{hi}\right)}{\sum_{h=1}^{L}\frac{W_h^2}{n_h}b_h\left(1+u\left(x\right)_{hi}s_{xh}^2+v\left(x\right)_{hi}\right)^2}$$

(3.20)

By substituting equation (3.20) into equation (3.18) and putting $b_h=\left(1+u\left(x\right)_{hi}s_{xh}^2+v\left(x\right)_{hi}\right)^{-1}$, we get equation (3.21)

$$\pi_{hi}^2=W_h^2+W_h^2\left\{\frac{\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(1+u\left(x\right)_{hi}S_{xh}^2+v\left(x\right)_{hi}\right)-\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(1+u\left(x\right)_{hi}s_{xh}^2+v\left(x\right)_{hi}\right)}{\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(1+u\left(x\right)_{hi}s_{xh}^2+v\left(x\right)_{hi}\right)}\right\}$$

(3.21)

Now, by substituting equation (3.21) into equation (3.15), we get equation (3.22)

$$t_{p2i}=\sum_{h=1}^{L}\frac{W_h^2}{n_h}s_{yh}^2\left(\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(1+u\left(x\right)_{hi}S_{xh}^2+v\left(x\right)_{hi}\right)\right)\left(\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(1+u\left(x\right)_{hi}s_{xh}^2+v\left(x\right)_{hi}\right)\right)^{-1}$$

(3.22)

So, the six members of the proposed estimator in equation (3.22) are

$$t_{p21}=\sum_{h=1}^{L}\frac{W_h^2}{n_h}s_{yh}^2\left(\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(1+S_{xh}^2\right)\right)\left(\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(1+s_{xh}^2\right)\right)^{-1}$$

(3.23)

$$t_{p22}=\sum_{h=1}^{L}\frac{W_h^2}{n_h}s_{yh}^2\left(\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(1+S_{xh}^2-C_{xh}\right)\right)\left(\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(1+s_{xh}^2-C_{xh}\right)\right)^{-1}$$

(3.24)

$$t_{p23}=\sum_{h=1}^{L}\frac{W_h^2}{n_h}s_{yh}^2\left(\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(1+S_{xh}^2-\beta_{2(x)h}\right)\right)\left(\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(1+s_{xh}^2-\beta_{2(x)h}\right)\right)^{-1}$$

(3.25)

$$t_{p24}=\sum_{h=1}^{L}\frac{W_h^2}{n_h}s_{yh}^2\left(\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(1+\beta_{2(x)h}S_{xh}^2-C_{xh}\right)\right)\left(\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(1+\beta_{2(x)h}s_{xh}^2-C_{xh}\right)\right)^{-1}$$

(3.26)

$$t_{p25}=\sum_{h=1}^{L}\frac{W_h^2}{n_h}s_{yh}^2\left(\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(1+C_{xh}S_{xh}^2-\beta_{2(x)h}\right)\right)\left(\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(1+C_{xh}s_{xh}^2-\beta_{2(x)h}\right)\right)^{-1}$$

(3.27)

$$t_{p26}=\sum_{h=1}^{L}\frac{W_h^2}{n_h}s_{yh}^2\left(\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(1+S_{xh}^2+\beta_{2(x)h}\right)\right)\left(\sum_{h=1}^{L}\frac{W_h^2}{n_h}\left(1+s_{xh}^2+\beta_{2(x)h}\right)\right)^{-1}$$

(3.28)

4. Empirical Study

To investigate the efficiency of the proposed estimators over the existing ones empirically, simulation study was conducted

4.1 Simulation Study

In this section, simulation study was performed to examine the performance of the proposed estimators over Sample variance and Muili et al. 2018) variance estimators in this study. We generate a bivariate random population (N=1000) for the study population stratified into three groups that are non-overlapping of size $N_1=200, N_2=300, N_3=500,$ using function defined in Table 1, sample of sizes 20, 30 and 50 (n=100) were selected 10,000 times using simple random sampling without replacement from each stratum respectively.

Table 1. MSE and PRE of Existing

Populations	**Auxiliary Variable X**	**Study Variable Y**
1	$X\sim\exp\left(\gamma_h\right),\gamma_1=0.3,\gamma_2=0.1,\gamma_3=0.1$	$Y=0.5X+X^2+e,e\sim N\left(0,1\right)$
2	$X\sim gamma\left(\alpha_h,\beta_h\right),\alpha_1=10,\beta_1=25,$ $\alpha_2=10,\beta_2=26,\alpha_3=10,\beta_3=27,$	$Y=0.5X+X^2+e,e\sim N\left(0,1\right)$
3	$X\sim beta\left(l_h,z_h\right),l_1=10,z_1=25,$ $l_2=10,z_2=26,l_3=10,z_3=27,$	$Y=0.5X+X^2+e,e\sim N\left(0,1\right)$
4	$X\sim Chisq\left(q_h\right),q_1=9,q_2=12,q_3=13$	$Y=0.5X+X^2+e,e\sim N\left(0,1\right)$

Table 2. MSE of Existing and Proposed Estimators using populations 1 and 2

Estimators	**MSE (Population 1)**	**MSE (Population 2)**

t_0	24456796645	1.096142
t_{j1}	24283465680	1.036588
t_{j2}	24283405172	1.036021
t_{j3}	24282267300	1.035639
t_{j4}	24283798929	1.050472
t_{j5}	24284432653	1.036004
t_{j6}	24170705095	1.035637
Proposed Estimators		
t_{p11}	**24078486256**	**1.034264**
t_{p12}	**24078543502**	**1.034949**
t_{p13}	**24078593931**	**1.034944**
t_{p14}	**24077972342**	**1.034949**
t_{p15}	**24078654723**	**1.034944**
t_{p16}	**24077711145**	**1.034943**
t_{p21}	**24078420486**	**1.034942**
t_{p22}	**24078486207**	**1.034941**
t_{p23}	**24103778106**	**0.982552**
t_{p24}	**24077964835**	**1.034933**
t_{p25}	**24078708383**	**1.034944**
t_{p26}	**24077598021**	**1.034943**

Table 3. MSE of Existing and Proposed Estimators using populations 3 and 4

Estimators	MSE (Population 3)	MSE (Population 4)
t_0	1.078126	356492811
t_{j1}	1.024699	329237304
t_{j2}	1.024089	32940906
t_{j3}	1.0239761	329292966
t_{j4}	1.03936	329256684
t_{j5}	1.024089	330101872
t_{j6}	1.023761	328899833
Proposed Estimators		
t_{p11}	**1.022665**	**328032847**
t_{p12}	**1-023036**	**328032716**
t_{p13}	**1.23035**	**328025710**
t_{p14}	**1.23036**	**328031732**
t_{p15}	**1.023035**	**327935053**
t_{p16}	**1.023035**	**328030302**
t_{p21}	**1.023035**	**328032722**
t_{p22}	**1.023035**	**32832843**
t_{p23}	**0.9132592**	**328494954**
t_{p24}	**1.023035**	**32832007**
t_{p25}	**1.023035**	**327995077**
t_{p26}	**1.023035**	**328028849**

Tables 2 and 3, show the MSEs of the existing and proposed estimators using populations in Table 1. The results of the mean square errors (MSEs) show that the proposed estimators are more efficient than variance and Muili et al. (2018) estimators considered in this study with the evidence of the proposed estimators having less or minimum MSE as compared to the existing ones considered in this study. Muili et al. (2018) estimators performs better than sample variance estimator with the evidence of smaller MSE. In all the populations considered, the proposed estimators are the most robust and highly efficient in estimating population variance of the study variable. The implication of the proposed estimators having minimum MSE is that the estimates produced by the proposed estimators are closer to the true estimate on the average than the variance and Muili et al. (2018) estimators.

5 Results and Discussion

Table 2. and 3 show the mean square error of the existing and proposed estimators. It can be observed that the proposed estimators have smaller MSEs as compared to other estimators considered in this study. The proposed estimators have two merits over the existing ones. First, they can be applied in real life surveys as compared to Muili et al. (2018) and second, they more efficient (low mean square error), that is, they can provide more precise estimates more the variance and Muili et al. (2018) estimators. The two main of objectives had been achieved. The proposed calibrated variance estimators was obtained after studying the work of Muili et al. (2018) and other related estimators in the literature. The developed calibrated variance estimators are most suitable for

estimating the finite population variance than the sample variance or Muili et al. (2018) estimators. The reasons include, by using auxiliary information, they can produce more precise estimates and ensure estimates align known population characteristics and also can handle non-response, and that is, they can adjust for biases due to missing data

6 Conclusion and Future Scope

In this study, we have modified some existing variance estimators for the estimation of population variance of the study variable by incorporating auxiliary information, coefficient of variation and coefficient of kurtosis of the auxiliary variable, using calibration and optimization approach to obtain calibration variance estimators with gain in efficiency over the existing ones. The existing variance estimators are function of unknown constants that depend on parameters of the study variable, hence the applicability of these estimators is not possible, and again the efficiency of the estimators depend on these unknown constants, .hence, to overcome these challenges, calibrated variance estimators were derived whose applicability is possible and free from unknown constants in conclusion, the two aims of the study were achieved, that is, proposing variance that can be apply in real life practice and to obtain estimators with better efficiency. The limitation of the current study is that the proposed estimator utilized auxiliary information of the auxiliary variable such as coefficient of variation and coefficient of kurtosis of the auxiliary variable which are prone to the influence of extreme values or outliers which may lead the estimators to either overestimation or underestimation. Again the properties of the proposed calibrated variance estimators were not derived. With these identified gaps, the future study can be undertaken by researchers to look into these directions of improving the current study to fill these gaps.

Acknowledgement:

The authors express their gratitude to the knowledgeable reviewers for their insightful recommendations on how to make this paper better. Furthermore, we are appreciative to Allah Ta'ala for enabling this research to be successfully completed.

Funding: The authors received no financial support for the research, authorship, and publication of this article.

Contribution of the Authors: Awwal handles the Introduction, Literature review and arrangement of the references for this research project, while Mojeed oversees the computation of the data and derivation of the proposed estimators. Then, Ahmed conducted the initial review of the research report before it was submitted to the journal for publication.

References:

[1] Isaki, C. T. (1983). Variance estimation using auxiliary information. Journal of the American Statistical Association. 78(381), 117-123.

[2] Prasad, B. and Singh, H. P. (1992). Unbiased estimators of finite population variance using auxiliary information in sample surveys. Communications in Statistics, 21(5), 1367-1376.

[3] Upadhyaya, L. and Singh, H. (1999). An estimator for population variancethat utilizes the kurtosis of an auxiliary variable in sample surveys. Vikram Mathematical Journal, 19, 14-17.

[4] Singh, R., Chauhan, P., Sawan, N. and Smarandache, F. (2011). Improved exponential estimator for population variance using two auxiliary variables. Italian Journal of Pure and Applied Mathematics, 28, 101-108.

[5] Khan, M and Shabbir, J. (2013). A ratio type estimator for the estimation of population variance using quartiles of an auxiliary variable. Journal of Statistics Application and Probability, 2(3), 31-325.

[6] Shabbir, J. and Gupta, S. (2015). A note on generalized exponential type estimator for population variance in survey sampling. Revista Colombiana de Estadistica, 38(2), 385-397.

[7] Audu, A., Singh, R.V.K., and Adewara, A.A. (2016). Ratio and product type exponential estimators of population variance under transformed sample information of study and supplementary variables. Asian Journal of Mathematics and Technology, 11(3), 175-183.

[8] Yunusa, M. A., Adejumobi, A. and Audu, A. (2022). A modified ratio estimator in the presence of Tri-mean and interquartile range for the estimation of population variance. LC International Journal of STEM, 3(4), 40-50.

[9] Adejumobi, A. and Yunusa, M. A. (2022). Some improved class of ratio estimators for finite population variance with the use of known parameters. LC International Journal of STEM, 3(3), 35-45

[10] Adejumobi, A. and Yunusa, M. A. (2023). An enhanced variance estimator with power transformation. Discovery, 59, e41d104.

[11] Audu, A., Muhammad, A., Kane, I. L., Mohammed, A., Ishaq, O. O. and Yunusa, M. A. (2023). Modified variance estimators in the presence of simultaneous effects of measurement error and non-response using auxiliary variable. Dutse Journal of Pure and Applied Sciences, 9(4b), 347-361

[12] Ozel, G., Cingi, H. and Oguz, M. (2014). Separate ratio estimators for the population variance in stratified random sampling. Communications in Statistics-Theory and Methods, 43(22), 4766-4779.

[13] Muili, J. O., Singh, R. V. K. and Audu, A.(2018). Study of Efficiency of some finite population variance estimators in stratified random sampling. Continental Journal of Applied Sciences, 13(2), 1-17.

[14] Devville, J. C. and Sarndal, C. E. 2002). Calibration estimators in survey sampling. Journal of American Statistical Association, 87418), 376-382.

[15] Kim, J. M., Sungur, E. A. and Heo, T. Y. 2007). Calibration approach estimators in stratified sampling. Statistics and Probability Letters, 771), 99-103.

[16] Audu, A., Singh, R. and Khare, S. (2021). Developing calibration estimators for population mean using robust measures of dispersion under stratified random sampling. Statistics in Transition New Series, 22(2), 125-142

[17] Abubakar, Z., Audu, A., Kane, L., Dogondaji, A. M. and Yunusa, M. A. (2024). Modified calibrated variance estimators in the presence of non-response and measurement error. Asian Journal of Probability and Statistics, 2610), 80-94

[18] Audu, A. Yunusa, M. A., Aphane, M. and Lekganyane, M. (2025). Calibrated optional randomized response techniques for efficient and robust estimation of quantitative variables. Scientific Reports, 15(1), 17138.

[19] Yunusa, M. A., Audu, A., Usman, U., Aremu, K. O. and Aphane, M. (2026). Calibrated-Two Optional Randomized Response Techniques (C-TORRT) for the estimation of quantitative variable information. PLOS one, 21(1), 1-24.

Declaration of Generative AI and AI - assisted technologies in the writing process

During the preparation of this work the author(s) used [MATHTYPE/QUILBOT] in order to write the equations correctly and also rephrasing of words. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

The authors contributed in the present research, at all stages from the formulation of the problem to the final findings and solution.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The authors have no conflicts of interest to declare that are relevant to the content of this article.