Generalized Ratio-Product cum Regression Variance Estimator in Two-Phase Sampling

Isah Muhammad[1,2]

This study develops a flexible and efficient generalized ratio-product cum regression-type estimator of population variance utilizing auxiliary variable in two-phase sampling that incorporates the properties of ratio-type and product-type estimators. The properties of the estimator were derived using first order approximation. The theoretical conditions under which the precision and the flexibility of the estimator is better than some classical estimators are also provided. Empirical evidence from five real datasets suggests that the proposed estimator outperforms the classical variance, ratio variance, product, and exponential ratio type estimators in terms of precision and efficiency. The estimator can be utilized to provide better variance estimates for various phenomena such as inflation variation, exchange rate variation and standard of living variation for better policymaking.

Keywords: Variance; supplementary information; bias; mean square error; efficiency; double sampling.
JEL Classification: C02, C13, C83
DOI: 10.33429/Cjas.14223.4/5

1. Introduction

Estimation of population variance is an important issue which has been discussed by many experts engaged in sampling experiments. For instance, the variation of economic variables such as inflation and exchange rates over the past years is required to develop and implement a better policy. In agriculture, the production variation of crop is required for further planning while in manufacturing industries and pharmaceutical laboratories, the variability of their products is a necessity for their quality control (Muhammad *et al.*, 2022). In sampling experiments, the precision of the population characteristics estimation such as; mean, variance and sum, increases when the auxiliary information is utilized (Muhammad *et al.*, 2022; Oyeyemi *et al.,*2023; Audu *et al.*, 2023; Muhammad *et al.*, 2023; Zakari & Muhammad, 2022; Zakari *et*

[1]Department of Statistics, Binyaminu Usman Polytechnic, Hadejia, Nigeria.
[2]Corresponding author: isahsta@gmail.com

al., 2023). Various approaches of ratio, product and regression methods of estimation have been widely utilized where supplementary information is available. The strategy of linear transformation of the auxiliary variable has been utilized in the literature by various researchers in developing mixtures of ratio estimators of population mean, total and variance (Mishra *et al.,* 2019 and Muhammad *et al.,* 2021). Thus, exponential estimators of population variance have been extensively discussed in literature (Isaki, 1983; Singh *et al.,* 2009; Shabbir & Gupta, 2007). Similarly, the strategy of exponentiation has been considered by various researchers in developing new estimators under simple random sampling and stratified random sampling which proved to be more efficient than the ordinary estimators. In sample survey, when the population variance of the secondary character is not available, the two-phase sampling scheme can be used in obtaining an enhanced estimator rather than the single sampling method (Neyman, 1938).

The limitations of the usual variance ratio and variance regression estimators can be handled by ratio-regression-type estimators and some developments was made including Shabbir and Gupta (2007) that developed an estimator by combining the concept of Isaki (1983) classical regression estimator and ratio exponential type estimator. Yadav *et al.* (2015) proposed a regression-ratio-type exponential estimator by combining Isaki (1983), Singh *et al.* (2009) and Bahl and Tuteja (1991) estimators. Mishra *et al.* (2019) proposed a class of log-product type exponential estimators.

In sampling experiments, it is well known that ratio-type estimators cannot provide better estimates in a situation where the relationship between the variable of interest and auxiliary variable is negative, while product-type estimators cannot provide better estimates in a situation where the relationship is positive (Isaki, 1983). Also the bias of the estimators is still extreme which may lead to over estimation or under estimation of population characteristics. To address these problems, this study under two-phase sampling suggests a new generalized ratio-product cum regression-type estimator with some special classes that provide efficient and precise estimation of population characteristics. The study also derives the properties and theoretical efficiency conditions of the proposed estimator and compare its performance with those of the classical estimators using real datasets based on the criteria of mean square error and percentage relative efficiency.

The rest of the study is structured as follows: Section 2 presents notations and reviewed relevant literature, while the data used in this study and methodology are presented in Section 3. The empirical results are discussed in Section 4 and Section 5 presents the conclusion and policy recommendations.

2. Literature Review

2.1 Empirical Overview

Isaki (1983) proposed ratio and regression estimators of population variance in two-phase sampling, the mean square error of the proposed estimators up to the first order of approximation were obtained. The study used the criterion of mean square error and percentage relative efficiency in comparing the efficiency of the proposed and existing estimators. In the work of Singh *et al.* (1988), a two-phase sampling scheme was considered in developing difference-type estimator to provide precise and efficient population characteristics such as; variance and mean. The efficiency of the developed difference-type estimator was compared using real datasets based on the criteria of bias, minimum mean square error and percentage relative efficiency. Ahmed *et al.* (2003) utilized the supplementary information and proposed class of two-phase estimators. The performances of the suggested estimators and other existing estimators were studied using real datasets based on bias and mean square error.

Shabbir and Gupta (2007) considered the work of Singh *et al.* (1988) and proposed a new estimator of two-phase sampling by applying exponential transformation approach. They derived the bias and mean square error of the estimator and compared it with some available estimators in the literature. Mishra *et al.* (2019) employed the log-type transformation approach and proposed a class of product exponential-type estimators for the estimation of population variance under Double Sampling Scheme. They used first order of approximation and derived the bias and mean square error of their estimators. By using real datasets they found that their estimators have minimum biases and mean square errors. A two-phase generalized ratio-type variance estimator was also suggested by Sukhatme (1962), where the properties of the suggested estimator were extensively derived and the efficiency of the estimator was assessed using real datasets.

In a situation where there is non-response problem and the population has diverse features, Singh *et al.* (2020) proposed a class of new variance estimators and derived their properties using first order of approximation. The criteria of bias and mean square error were used to assess the efficiency of the proposed estimators and existing estimators. Cochran (1977) provided some basic information for two-phase sampling. The estimators for the population variance considering a supplementary auxiliary variable have been discussed in literature (Grover, 2010; Upadhaya & Singh, 1999; Ahmed *et al.,* 2000; Singh & Singh, 2001; Al-Jararha & Ahmed, 2002; Ahmed *et al.,* 2003; Kadilar & Cingi, 2006; Bahl & Tuteja, 1991; Upadhyaya *et al.,* 2004; Singh & Solanki, 2013; Swain, 2015; Subramani & Kumarapandiyan, 2015; Yakub & Shabbir, 2016; Shabbir & Gupta, 2007; Singh & Malik, 2014; Yadav *et al.,* 2015; Yadav & Kadilar, 2014; Zakari & Muhammad, 2023).

The bias of the aforementioned estimators are extreme which may lead to over estimation or under estimation. The generalized estimator in this study possesses the minimum bias and mean square error and provides better estimates when the relationship between the variable of interest and auxiliary variable is either positive or negative, and this could make the proposed estimator more flexible and better.

3.0 Data and Methodology

The performance of the developed generalized estimator over some existing estimators is measured empirically based on the criteria of bias, mean square error, and percentage relative efficiency. The analysis is carried out using R statistical package.

3.1 Data

Assessments based on the flexibility and efficiency of the classes of developed generalized estimator with some existing population variance estimators are carried out using five real datasets including those used by Cochran (1977) and Sarjinder (2003) and those sourced from CBN *Statistical Bulletin* (2010; 2014) and CBN *Annual Reports.*

Population I: Cochran (1977);

y: the number of persons in the city;

x: the number of rooms in the city.

Population II: CBN Statistical Bulletin(2010);

y: average annual exchange rate in Nigeria from 1960-2010;

x: average annual inflation rate in Nigeria from 1960-2010.

Population III: CBN Statistical Bulletin(2014);

y: Nigerian monthly average crude oil price from 2006-2010;

x: Nigerian monthly average crude oil production from 2006-2010.

Population IV: CBN Statistical Bulletin(2014);

y: Nigerian monthly average crude oil price from 2010-2014;

x: Nigerian monthly average crude oil production from 2010-2014.

Population V: Sarjinder (2003);

y: duration of sleep (in minutes);

x: age (in years) of the persons.

Hence, the summary statistics obtained from datasets are presented in Table 1.

Table 1: Summary Statistics of the Datasets

Parameters	Dataset I	Dataset II	Dataset III	Dataset IV	Dataset V
N	100	63	60	60	30
n'	85	35	37	40	25
N	10	12	15	10	10
S_y^2	214.690	12378.571	411.924	638.305	3582.582
S_x^2	56.761	225.912	0.034	0.017	85.237
ρ_{xy}	0.956	0.805	-0.096	0.465	-0.855
β_{2y}^*	2.239	4.507	4.421	2.765	2.579
β_{2x}^*	2.252	6.758	3.726	2.389	2.164
λ_{22}^*	1.543	0.349	1.180	0.915	1.930

where N is the size of the population; n' is the size of the preliminary sample; n is size of the second sample; S_x^2 and S_y^2 represent the variances of the study variable and the auxiliary character, respectively; ρ_{xy} is the correlation coefficient between the variables; β_{2y}^* and β_{2x}^* are the population coefficient of kurtosis of the study variable y and the auxiliary variable x, respectively; λ_{22}^* is the population coefficient of covariance between the study and auxiliary variables (Mishra *et al.* 2019).

3.2 Symbols, Nomenclature and Related Estimators

Consider $U = U_1, ,,,,, U_N$ to be a finite population of size N and let (y_i, x_i) be the value of the variable of interest, Y, and the auxiliary character, X, on ith unit, U_i, $(i = 1, \ldots N)$. Let $\overline{Y}$ and $\overline{X}$ be means obtained from population of the variable of interest, Y, and the secondary character, X, respectively. Neyman (1938) highlighted that when the population characteristics such as; mean $\overline{X}$ and variance S_x^2 of the auxiliary character are unknown, a two-phase sampling scheme is used to estimate population variance. In this study, a two-phase sampling scheme under simple random sampling is employed, where the first phase sample $S'(S' \subset U)$ of a fixed size n' is drawn to measure only on the auxiliary variable X in order to formulate a good estimate of a population mean, $\overline{X}$, and population variance, S_x^2, respectively. The second phase sample $S(S \subset S')$ of a fixed size n is drawn from the first sample to measure the variable of interest, y, and auxiliary variable character, x, respectively. Furthermore, the following two-phase notation and formulaes were described in Isaki (1983), Zakari *et al.* (2020), Muhammad *et al.* (2022), Muhammad *et al.* (2021) and Mishra *et al.* (2019) as;

$$\bar{y} = \sum_{i\varepsilon S} y_i/n, \quad \bar{x} = \sum_{i\varepsilon S} x_i/n \ and \ \bar{x}^I = \sum_{i\varepsilon S'} x_i/n^I$$

$$s_y^2 = \frac{1}{(n-1)} \sum_{i\varepsilon S} (y_i - \bar{y})^2, s_x^2 = \frac{1}{(n-1)} \sum_{i\varepsilon S} (x_i - \bar{x})^2 \ and \ s_x^{\prime 2} = \frac{1}{(n'-1)} \sum_{i\varepsilon S'} \left(x_i - \bar{x}'\right)^2$$

where the first sample of size n' is denoted by S'; the second sample of size n is represented by S; y denotes the variable of interest defined on ith units and $\overline{Y}$ represents the population average of the study variable; x denotes the auxiliary variable defined on ith units; $\bar{x}^I$ denotes the mean obtained from the primary sample of the auxiliary variable; $\bar{x}$ denotes the mean obtained from the sub-sample of the auxiliary variable; $\bar{y}$ denotes the mean obtained from the sub-sample of the study variable.

The following equations and notations are also defined:

$\lambda = \left(\frac{1}{n} - \frac{1}{N}\right)$, $\lambda' = \left(\frac{1}{n'} - \frac{1}{N}\right)$, $C_y = S_y/\overline{Y}$, $C_x = S_x/\overline{X}$, $\rho_{yx} = S_{yx}/(S_y S_x)$, $S_y^2 = \frac{\sum_{i=1}^N \left(y_i - \overline{Y}\right)^2}{N-1}$, $S_x^2 = \frac{\sum_{i=1}^N \left(x_i - \overline{X}\right)^2}{N-1}$, $S_{yx} = \frac{\sum_{i=1}^N \left(y_i - \overline{Y}\right)\left(x_i - \overline{X}\right)}{N-1}$, $\beta_{2y}^* = \lambda_{40} - 1$, $\beta_{2x}^* = \lambda_{04} - 1$ and $\lambda_{22}^* = \lambda_{22} - 1$, $\lambda_{rs} = \frac{\mu_{rs}}{\sqrt{\mu_{20}^r \mu_{02}^s}}$ and $\mu_{rs} = \frac{1}{N} \sum_{j=1}^N (Y_j - \bar{Y})^r (X_j - \bar{X})^s$

where n', n, N, $\overline{Y}$ $\overline{X}$, S_{yx}, S_x^2 and S_y^2 remained as earlier defined; λ and λ' denote the sampling fraction obtained from the respective units of the samples and populations; ρ_{yx} denote the coefficient of correlation between the variable of interest and the secondary character; C_x and C_y represent the coefficients of variation obtained from the study variable and the auxiliary variable, respectively; λ_{rs} is the population coefficient of kurtosis; μ_{rs} is the moments about the mean; r and s are the order of the moments (Mishra *et al.* 2019).

Note;

MSE = Mean Square Error;

PRE =Percentage Relative Efficiency.

The usual unbiased estimator of variance is given as:

$$t_0 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2 \tag{1}$$

The expression for the variance of conventional unbiased estimator of variance (t_0) is given as:

$$Var(t_0) = \lambda S_y^4 \beta_{2y}^* \tag{2}$$

Isaki (1983) proposed usual ratio estimator of the population variance using auxiliary information under two-phase sampling as:

$$t_1 = s_y^2\left(\frac{s_x^{\prime 2}}{s_x^2}\right) \tag{3}$$

Applying the first order of approximation, Isaki (1983) obtained the bias and mean square error of the usual variance ratio estimator as:

$$Bias(t_1) \cong S_y^2\left(\lambda - \lambda^{\prime}\right)(\beta_{2x}^* - \lambda_{22}^*) \tag{4}$$

$$MSE(t_1) \cong S_y^4\left[\lambda \beta_{2y}^* + \left(\lambda - \lambda^{\prime}\right)(\beta_{2x}^* - 2\lambda_{22}^*)\right] \tag{5}$$

Murthy (1964) proposed usual product estimator of the population variance using auxiliary information under two-phase sampling as:

$$t_2 = s_y^2\left(\frac{s_x^2}{s_x^{\prime 2}}\right) \tag{6}$$

Applying the first order of approximation, the bias and mean square error of the usual variance product estimator is obtained as:

$$Bias(t_2) \cong S_y^2\left(\lambda - \lambda^{\prime}\right)\lambda_{22}^* \tag{7}$$

$$MSE(t_2) \cong S_y^4 \left[\lambda \beta_{2y}^* + \left(\lambda - \lambda'\right) (\beta_{2x}^* + 2\lambda_{22}^*)\right] \tag{8}$$

Singh *et al*. (1988) proposed usual difference estimator of the population variance using auxiliary information under two-phase sampling as:

$$t_3 = k_1 s_y^2 + k_2 (s_x'^2 - s_x^2) \tag{9}$$

k_1 and k_2 denote the suitable chosen constants, whose values are to be obtained by minimizing the mean square error. The optimum values of k_1 and k_2 along with the minimum MSE equation of the estimator in (9), up to the first order of approximation are given, respectively as:

$$k_1^{opt} = \frac{\beta_{2x}^*}{\beta_{2x}^* + \beta_{2y}^*\beta_{2x}^* - \lambda_{22}^*}$$

$$k_2^{opt} = \frac{S_x^2 \lambda_{22}^*}{S_y^2 \left(\beta_{2x}^* + \beta_{2y}^*\beta_{2x}^* - \lambda_{22}^*\right)}$$

$$Bias(t_3) \cong S_y^2 (k_1 - 1) \tag{10}$$

$$MSE(t_3)_{\min} \cong \frac{MSE(t_{reg})_{\min}}{1 + \frac{MSE(t_{reg})_{\min}}{S_y^4}} \tag{11}$$

where $MSE(t_{reg})_{\min} \cong S_y^4 \lambda \beta_{2y}^* - \frac{S_y^4(\lambda-\lambda')\lambda_{22}^*}{\beta_{2x}^*}$, S_y^2, S_x^2, β_{2y}^*, and λ_{22}^* bear the same definition given earlier. Applying exponential transformation to the usual regression, Shabbir and Gupta (2007) proposed a regression cum exponential variance estimator in two-phase sampling as:

$$t_4 = \left[k_3 s_y^2 + k_4 \left(s_x'^2 - s_x^2\right)\right] \exp\left(\frac{s_x'^2 - s_x^2}{s_x'^2 + s_x^2}\right) \tag{12}$$

where k_3 and k_4 are unknown constants. The optimum values of k_3 and k_4 and the minimized MSE of the usual variance regression estimator are obtained and given, respectively as:

$$k_3^{opt} = \frac{\beta_{2x}^*}{8}\left(\frac{8-\beta_{2x}^*}{\beta_{2x}^* + \beta_{2y}^*\beta_{2x}^* - \lambda_{22}^*}\right)$$

$$k_4^{opt} = \frac{S_y^2}{8S_x^2}\left(\frac{-4\beta_{2y}^* + \beta_{2x}^* + 8\lambda_{22}^* - \lambda_{22}^*\beta_{2x} + 4\beta_{2y}^*\beta_{2x}^* - 4\lambda_{22}^{*2}}{(\beta_{2x}^* + \beta_{2y}^*\beta_{2x}^* - \lambda_{22}^*)}\right)$$

$$Bias\,(t_4) = S_y^2\,(k_3-1) + S_y^2 k_3\left(\lambda - \lambda'\right)\left[\frac{3}{8}\beta_{2x}^* - \frac{1}{2}\lambda_{22}^*\right] + \frac{1}{2}S_x^2 k_4\left(\lambda - \lambda'\right)\beta_{2x}^* \quad (13)$$

$$MSE(t_4)_{\min} \cong \frac{MSE(t_{reg})_{\min}}{1+\frac{MSE(t_{reg})_{\min}}{S_y^4}} - \frac{(\lambda-\lambda')\beta_{2x}^*\left[MSE(t_{reg})_{\min} + \frac{(\lambda-\lambda')S_y^4\beta_{2x}^*}{16}\right]}{4\left[1+\frac{MSE(t_{reg})_{\min}}{S_y^4}\right]} \quad (14)$$

where S_y^2, S_x^2, β_{2y}^*, β_{2x}^*, λ_{22}^* λ, $\lambda'$and $MSE(t_2)_{\min}$ bear the same definition given earlier.

Mishra *et al.* (2019) proposed four classes of estimators for estimating population variance under two-phase sampling scheme using log-type transformation as:

$$Pl_1 = s_y^2 + w_0 \log\left(\frac{s_x^2}{s_x'^2}\right) \quad (15)$$

$$Pl_2 = s_y^2\,(w_1+1) + w_2 \log\left(\frac{s_x^2}{s_x'^2}\right) \quad (16)$$

$$Pl_3 = \left[s_y^2\,(w_3+1) + w_4 \log\left(\frac{s_x^2}{s_x'^2}\right)\right]\exp\left\{\frac{s_x'^2 - s_x^2}{s_x'^2 + s_x^2}\right\} \quad (17)$$

$$Pl_4 = s_y^2\,(w_5+1) + w_6 \log\left(\frac{s_x^2}{s_x'^2}\right)\exp\left\{\frac{s_x'^2 - s_x^2}{s_x'^2 + s_x^2}\right\} \quad (18)$$

where w_0, w_1, w_2, w_3, w_4, w_5 and w_6 are unknown parameters whose values depend on the sample and population information. The bias and minimum mean square error of the estimators along with the optimum values of the parameters are obtained, respectively as:

$$w_0 = \left[\frac{-S_y^2\lambda_{22}^*}{\beta_{2x}^*}\right],$$

$$w_1^{opt} = \frac{(A-C)\,D}{D^2-AB}$$

$$w_2^{opt} = \frac{CB-D^2}{D^2-AB}$$

$$w_3^{opt} = \frac{C_1B_1-D_1E_1}{E_1^2-A_1B_1}$$

$$w_4^{opt} = \frac{A_1D_1-C_1E_1}{E_1^2-A_1B_1}$$

$$w_5^{opt} = \frac{C_3B_3-D_3E_3}{E_3^2-A_3B_3}$$

$$w_6^{opt} = \frac{A_3D_3-C_3E_3}{E_3^2-A_3B_3}$$

$$Bias(Pl_1) \cong \frac{-S_y^2}{2}\lambda_{22}^*\left(\lambda-\lambda'\right) \tag{19}$$

$$Bias\,(Pl_2) = S_y^2\left[\frac{(A-C)\,D}{D^2-AB}\right] \tag{20}$$

$$Bias\,(Pl_3) \cong (1+w_3)\,S_y^2\left[\left(\lambda-\lambda'\right)\left(\frac{3}{8}\beta_{2x}^*-\frac{1}{2}\lambda_{22}^*\right)\right]-\frac{1}{2}w_4\left[\left(\lambda-\lambda'\right)\beta_{2x}^*\right] \tag{21}$$

$$Bias\,(Pl_4) \cong S_y^2 w_5-\frac{1}{2}w_6\left[\left(\lambda-\lambda'\right)\beta_{2x}^*\right] \tag{22}$$

$$MSE(Pl_1)_{\min} \cong S_y^4\left[\lambda\beta_{2y}^*-\left(\lambda-\lambda'\right)\frac{\lambda_{22}^*}{\beta_{2x}^*}\right] \tag{23}$$

$$MSE(Pl_2)_{\min} \cong C+\frac{BC^2+(A-2C)D^2}{D^2-AB} \tag{24}$$

$$MSE(Pl_3)_{\min} \cong F_1+\frac{B_1C_1^2+A_1D_1^2-2C_1D_1^2}{E_1^2-A_1B_1} \tag{25}$$

$$MSE(Pl_4)_{\min} \cong C_3 + \frac{B_3C_3^2 + A_3D_3^2 - 2C_3D_3^2}{E_3^2 - A_3B_3} \tag{26}$$

where:
$A = A_1 = A_3 = S_y^4\left(1 + \lambda\beta_{2y}^*\right)$, $B = B_1 = B_3 = \left(\lambda - \lambda'\right)\beta_{2x}^*$, $C = C_1 = C_3 = S_y^4\lambda\beta_{2y}^*$, $D = D_3 = S_y^2\left(\lambda - \lambda'\right)\lambda_{22}^*$, $D_1 = S_y^2\left(\lambda - \lambda'\right)\left(\lambda_{22}^* - \frac{\beta_{2x}^*}{2}\right)$, $E_1 = S_y^2\left(\lambda - \lambda'\right)$ $\left(\lambda_{22}^* - \beta_{2x}^*\right)$, $F_1 = S_y^4\left[\lambda\beta_{2y}^* + (\lambda - \lambda')\left(\frac{\beta_{2x}^*}{4} - \lambda_{22}^*\right)\right]$ and $E_3 = S_y^2\left(\lambda - \lambda'\right)\left(\lambda_{22}^* - \frac{\beta_{2x}^*}{2}\right)$, S_y^2, S_x^2, β_{2y}^*, β_{2x}^*, λ_{22}^* λ, λ' and $MSE(t_2)_{\min}$ bear the same definition given earlier.

3.3 Suggested Estimator

The proposed generalized estimator comprises of two components; where in the first component, ratio and product are combined with a driving parameter, α_1, and regression with a driving parameter, α_2; as $\alpha_1 s_y^2\left[\frac{1}{2}\left(\frac{s_x'^2}{s_x^2} + \frac{s_x^2}{s_x'^2}\right)\right]^{\beta} + \alpha_2\left(s_x'^2 - s_x^2\right)$. A linear combination approach was adopted in combining the first component and the second component: $exp\left(\frac{s_x'^2 - s_x^2}{s_x'^2 + s_x^2}\right)$ to produce a generalized two-phase variance estimator as:

$$\widehat{S}_{d_i}^2 = \left\{\alpha_1 s_y^2\left[\frac{1}{2}\left(\frac{s_x'^2}{s_x^2} + \frac{s_x^2}{s_x'^2}\right)\right]^{\beta} + \alpha_2\left(s_x'^2 - s_x^2\right)\right\}\exp\left\{\frac{s_x'^2 - s_x^2}{s_x'^2 + s_x^2}\right\} \tag{27}$$

where, $\hat{S}_{d_i}^2$ is the notation for the new estimator such that, if $\alpha_1 = 1$ and $\alpha_2 = 0$; the estimator becomes a ratio-product cum exponential estimator and, if $\alpha_1 = 0$ and $\alpha_2 = 1$; the generalized new estimator becomes ratio-regression cum exponential estimator. The parameter α_1 and α_2 are real numbers which minimized the MSE of the estimator and β is a driving parameter suitably chosen. Thus, the following relative error terms are defined to obtain the properties of the estimator:
$s_y^2 = S_y^2(1 + e_0)$, $s_x^2 = S_x^2(1 + e_1)$ and $s_x'^2 = S_x^2(1 + e_2)$
where e_0 denotes the relative error term corresponding to the study variable in the second sample; e_1 denotes the relative error term corresponding to the auxiliary variable in the second sample; e_2 denotes the relative error term corresponding to the auxiliary variable in the first sample. Such that

$$E(e_0) = E(e_1) = E(e_2) = 0$$

$$E\left(e_0^2\right)=\lambda\left(\beta_{2(y)}-1\right), E\left(e_1^2\right)=\lambda\left(\beta_{2(x)}-1\right), E\left(e_2^2\right)=\lambda^{'}\left(\beta_{2(x)}-1\right),$$
$$E\left(e_0e_1\right)=\lambda\left(\lambda_{22}-1\right), E\left(e_0e_2\right)=\lambda^{'}\left(\lambda_{22}-1\right), E\left(e_1e_2\right)=\lambda^{'}\left(\beta_{2(x)}-1\right)$$

The new generalized estimator $\hat{S}^2_{d_i}$ can be expressed in terms of e_i $(i=0,\ 1,\ 2)$ as:

$$\widehat{S}^2_{d_i}=\left\{\begin{array}{l}\alpha_1S_y^2\left(1+e_0\right)\frac{1}{2^\beta}\left[\left(1+e_2\right)\left(1+e_1\right)^{-1}+\left(1+e_1\right)\left(1+e_2\right)^{-1}\right]^\beta\\ -\alpha_2\left[S_x^2\left(1+e_2\right)-S_x^2\left(1+e_1\right)\right]\end{array}\right\}$$
$$\exp\left\{\frac{S_x^2\left(1+e_2\right)-S_x^2\left(1+e_1\right)}{S_x^2\left(1+e_2\right)+S_x^2\left(1+e_1\right)}\right\} \tag{28}$$

By expanding $\left(1+e_1\right)^{-1}$ using Taylor series and ignoring the error terms that have power greater than two, the equation (28) gives:

$$\widehat{S}^2_{d_i}=\left\{\begin{array}{l}\alpha_1S_y^2\left(1+e_0\right)\frac{1}{2^\beta}\left[\left(1+e_2\right)\left(1-e_1+e_1^2\right)+\left(1+e_1\right)\left(1-e_2+e_2^2\right)\right]^\beta\\ -\alpha_2\left(e_2-e_1\right)S_x^2\end{array}\right\}$$
$$\exp\left\{\frac{S_x^2\left(e_2-e_1\right)}{2S_x^2+S_x^2\left(e_2+e_1\right)}\right\} \tag{29}$$

Simplifying the RHS, evaluating out and ignoring the error terms that have power greater than two, we get:

$$\widehat{S}^2_{d_i}=\left\{\alpha_1S_y^2\left(1+e_0\right)\frac{1}{2^\beta}\left[2+e_1^2+e_2^2-2e_1e_2\right]^\beta-\alpha_2S_x^2\left(e_2-e_1\right)\right\}$$
$$\exp\left\{\frac{\left(e_2-e_1\right)}{2}\left[1+\frac{\left(e_2+e_1\right)}{2}\right]^{-1}\right\} \tag{30}$$

Multiplying out the terms in the RHS of equation (30) and ignoring the error terms that have power greater than two, we get:

$$\widehat{S}^2_{d_i}=\left\{\alpha_1S_y^2\left(1+e_0\right)\left[1+\frac{\beta e_1^2}{2}+\frac{\beta e_2^2}{2}-\beta e_1e_2\right]-\alpha_2S_x^2\left(e_2-e_1\right)\right\}$$
$$\exp\left\{\frac{e_2}{2}-\frac{e_1}{2}+\frac{e_1^2}{4}-\frac{e_2^2}{4}\right\} \tag{31}$$

Apply the concept of exponential series to equation (31) and ignoring the error terms that have power greater than two, the exponential component of (31), gives:

$$\widehat{S}^2_{d_i} = \left\{\alpha_1 S_y^2 (1+e_0)\left[1+\frac{\beta e_1^2}{2}+\frac{\beta e_2^2}{2}-\beta e_1 e_2\right]-\alpha_2 S_x^2 (e_2-e_1)\right\}$$

$$\left\{1+\frac{e_2}{2}-\frac{e_1}{2}-\frac{e_2^2}{8}+\frac{3e_1^2}{8}-\frac{e_1 e_2}{4}\right\} \tag{32}$$

Multiplying out the RHS of (32), ignoring the error terms that have power greater than two and factorizing the similar terms, we get:

$$\widehat{S}^2_{d_1} - S_y^2 = \left\{ \begin{array}{l} (\alpha_1 - 1) S_y^2 + \alpha_1 S_y^2 \left[e_0 - \frac{e_1}{2} + \frac{e_2}{2} + A e_1^2 + B e_2^2 - \frac{e_0 e_1}{2} + \frac{e_0 e_2}{2} - C e_1 e_2\right] \\ +\alpha_2 S_x^2 \left[e_1 - e_2 - \frac{e_1^2}{2} - \frac{e_2^2}{2} + e_1 e_2\right] \end{array} \right\} \tag{33}$$

where $A_i = \left[\frac{4\beta+3}{8}\right]$, $B_i = \left[\frac{4\beta-1}{8}\right]$ and $C_i = \left[\frac{8\beta+1}{8}\right]$
Taking expectation of (33), the generalized new estimator's bias is obtained as:

$$Bias\left(\widehat{S}^2_{d_i}\right) = \left\{ \begin{array}{l} (\alpha_1 - 1) S_y^2 + \alpha_1 S_y^2 \left[\left(A_i \lambda + B_i \lambda' - C_i \lambda'\right)(\beta_{2x} - 1) + \frac{1}{2}\left(\lambda' - \lambda\right)(\lambda_{22} - 1)\right] \\ +\alpha_2 S_x^2 \frac{1}{2}\left(\lambda' - \lambda\right)(\beta_{2x} - 1) \end{array} \right\} \tag{34}$$

Squaring equation (33), ignoring the error terms that have power greater than two and factorizing the similar terms, we get:

$$\left(\widehat{S}^2_{d_i} - S_y^2\right)^2 = \left\{ \begin{array}{l} (\alpha_1 - 1)^2 S_y^4 + \alpha_1^2 S_y^4 \left[\begin{array}{l} 2e_0 - e_1 + e_2 + e_0^2 + \frac{(8A_i+1)e_1^2}{4} + \frac{(8B_i+1)e_2^2}{4} \\ -2e_0 e_1 + 2e_0 e_2 - \frac{(4C_i+1)e_1 e_2}{2} \end{array} \right] \\ +\alpha_2^2 S_x^4 \left(e_1^2 + e_2^2 - 2e_1 e_2\right) - 2\alpha_1 \alpha_2 S_y^2 S_x^2 \\ \left(e_1 - e_2 - e_1^2 - e_2^2 + e_0 e_1 - e_0 e_2 + 2e_1 e_2\right) \\ -\alpha_1 S_y^4 \left(e_0 - \frac{e_1}{2} + \frac{e_2}{2} + A_i e_1^2 + +B_i e_2^2 - \frac{e_0 e_1}{2} + \frac{e_0 e_2}{2} - C_i e_1 e_2\right) \\ +\alpha_2 S_y^2 S_x^2 \left(e_1 - e_2 - \frac{e_1^2}{2} - \frac{e_2^2}{2} + e_1 e_2\right) \end{array} \right\} \tag{35}$$

Taking expectation of (35), the MSE of the proposed estimator, up to the first order of approximation, is given as:

$$MSE\left(\widehat{S}^2_{d_i}\right)=\left\{\begin{array}{l}(\alpha_1-1)^2S_y^4+\alpha_1^2S_y^4\left[\begin{array}{l}\lambda\left(\beta_{2y}-1\right)+\frac{1}{4}\left([8B_i-8C_i-1]\lambda'+[8A_i+1]\lambda\right)\\ \left(\beta_{2x}-1\right)\\ +2\left(\lambda'-\lambda\right)\left(\lambda_{22}-1\right)\end{array}\right]\\ +\alpha_2^2S_x^4\left(\lambda'-\lambda\right)\left(\beta_{2x}-1\right)+2\alpha_1\alpha_2S_y^2S_x^2\left[\left(\lambda'-\lambda\right)\left[\left(\beta_{2x}-1\right)-\left(\lambda_{22}-1\right)\right]\right]\\ -2\alpha_1S_y^4\left[\left(A_i\lambda+(B_i-C_i)\lambda'\right)\left(\beta_{2x}-1\right)+\frac{1}{2}\left(\lambda'-\lambda\right)\left(\lambda_{22}-1\right)\right]\\ -2\alpha_2S_y^2S_x^2\frac{\left(\lambda'-\lambda\right)}{2}\left(\beta_{2x}-1\right)\end{array}\right\} \tag{36}$$

The new generalized estimator's MSE expression is subsequently reduced to:

$$MSE\left(\widehat{S}^2_{d_i}\right)=(\alpha_1-1)^2S_y^4+\alpha_1^2S_y^4D_i+\alpha_2^2S_x^4E_i+2\alpha_1\alpha_2S_y^2S_x^2F_i-2\alpha_1S_y^4G_i-2\alpha_2S_y^2S_x^2H_i \tag{37}$$

where: $D_i=[\lambda\left(\beta_{2y}-1\right)+\frac{1}{4}\left([8B_i-8C_i-1]\lambda'+[8A_i+1]\lambda\right)\left(\beta_{2x}-1\right)+2\left(\lambda'-\lambda\right)\left(\lambda_{22}-1\right)$, $E_i=\left(\lambda'-\lambda\right)\left(\beta_{2x}-1\right)$, $F_i=\left(\lambda'-\lambda\right)\left[\left(\beta_{2x}-1\right)-\left(\lambda_{22}-1\right)\right]$, $G_i=\left[\left(A_i\lambda+(B_i-C_i)\lambda'\right)\left(\beta_{2x}-1\right)+\frac{1}{2}\left(\lambda'-\lambda\right)\left(\lambda_{22}-1\right)\right.$, $H_i=\frac{1}{2}\left(\lambda'-\lambda\right)\left(\beta_{2x}-1\right)$

We obtain the optimum values of α_1 and α_2, by differentiating (37) partially with respect to α_1 and α_2, and equating to zero, respectively as:

$$\alpha_1^{(opt)}=\frac{(1+G_i)E_i-F_iH_i}{(1+D_i)E_i-F_i^2}$$

and

$$\alpha_2^{(opt)}=\frac{S_y^2}{S_x^2}\left[\frac{(1+D_i)H_i-(1+G_i)F_i}{(1+D_i)E_i-F_i^2}\right]$$

The minimized new generalized estimator's MSE is obtained by substituting the optimum values of α_1 and α_2 into equation (37), as:

$$MSE\left(\widehat{S}^2_{d_i}\right)_{\min} = S_y^4\left\{1-\frac{\begin{bmatrix}(1+D_i)\left((G_i+1)^2E_i^2-F_i^2H_i^2\right)+2(1+G_i)\\ \left(F_i^2-D_iE_i-E_i\right)F_iH_i\\ +(1+2D_i)E_iH_i^2-(1+G_i)^2E_iF_i^2\end{bmatrix}}{\left[(1+D_i)E_i-F_i^2\right]^2}\right\} \quad (38)$$

Special Cases: For $\beta = 1$, the proposed estimator in equation (27) becomes

$$\widehat{S}^2_{d_1} = \left\{\alpha_1 s_y^2\left[\frac{1}{2}\left(\frac{s_x'^2}{s_x^2}+\frac{s_x'^2}{s_x'^2}\right)\right]+\alpha_2\left(s_x'^2-s_x^2\right)\right\}\exp\left\{\frac{s_x'^2-s_x^2}{s_x'^2+s_x^2}\right\} \quad (39)$$

The optimum values of α_1 and α_2, of the proposed estimator in equation (39) are obtain as:

$$\alpha_1^{(opt)} = \frac{(1+G_1)E_1-F_1H_1}{(1+D_1)E_1-F_1^2}$$

and

$$\alpha_2^{(opt)} = \frac{S_y^2}{S_x^2}\left[\frac{(1+D_1)H_1-(1+G_1)F_1}{(1+D_1)E_1-F_1^2}\right]$$

The minimum MSE, up to the first order of approximation is obtained as:

$$MSE\left(\widehat{S}^2_{d_1}\right)_{\min} = S_y^4\left\{1-\frac{\begin{bmatrix}(1+D_1)\left((G_1+1)^2E_1^2-F_1^2H_1^2\right)+2(1+G_1)\\ \left(F_1^2-D_1E_1-E_1\right)F_1H_1\\ +(1+2D_1)E_1H_1^2-(1+G_1)^2E_1F_1^2\end{bmatrix}}{\left[(1+D_1)E_1-F_1^2\right]^2}\right\} \quad (40)$$

When $\beta = 2$, the proposed estimator in equation (27) becomes

$$\widehat{S}^2_{d_2} = \left\{\alpha_1 s_y^2\left[\frac{1}{2}\left(\frac{s_x'^2}{s_x^2}+\frac{s_x'^2}{s_x'^2}\right)\right]^2+\alpha_2\left(s_x'^2-s_x^2\right)\right\}\exp\left\{\frac{s_x'^2-s_x^2}{s_x'^2+s_x^2}\right\} \quad (41)$$

The optimum values of α_1 and α_2, of the proposed estimator in equation (41) are obtain as:

$$\alpha_1^{(opt)} = \frac{(1+G_2)E_2 - F_2H_2}{(1+D_2)E_2 - F_2^2}$$

and

$$\alpha_2^{(opt)} = \frac{S_y^2}{S_x^2}\left[\frac{(1+D_2)H_2 - (1+G_2)F_2}{(1+D_2)E_2 - F_2^2}\right]$$

The minimum mean squared error, up to the first order of approximation is obtained as:

$$MSE\left(\widehat{S}^2_{d_1}\right)_{\min} = S_y^4\left\{1 - \frac{\left[\begin{array}{l}(1+D_2)\left((G_2+1)^2E_2^2 - F_2^2H_2^2\right) + 2(1+G_2)\\ \left(F_2^2 - D_2E_2 - E_2\right)F_2H_2\\ +(1+2D_2)E_2H_2^2 - (1+G_2)^2E_2F_2^2\end{array}\right]}{\left[(1+D_2)E_2 - F_2^2\right]^2}\right\} \quad (42)$$

3.3 Efficiency Conditions of the Proposed Estimator

This section illustrates the performance of the proposed estimator over existing variance estimator, ratio estimator of variance, regression variance estimator and other population variance estimators considered. The section also presents the conditions under which the developed estimator is more precise, flexible and efficient than population variance estimators considered in this study.

Comparison between developed generalized estimator's MSE and that of estimator of sample variance:

$Var(\widehat{s}_y^2) - MSE(\widehat{S}^2_{d_i})_{\min} > 0$, if

$$S_y^4\lambda\left(\beta_{2y} - 1\right) - S_y^4\left(1 - \frac{K}{M}\right) > 0$$

$$\lambda\left(\beta_{2y}-1\right)-\left(1-\frac{K}{M}\right)>0 \tag{43}$$

where: $K=(1+D_2)\left((G_2+1)^2E_2^2-F_2^2H_2^2\right)+2(1+G_2)\left(F_2^2-D_2E_2-E_2\right)F_2H_2+(1+2D_2)E_2H_2^2-(1+G_2)^2E_2F_2^2M=\left[(1+D_2)E_2-F_2^2\right]^2$

If condition I, in Eq. 43, is satisfied, then the new generalized estimator is better and more efficient than the estimator of sample variance.

Comparison between the new generalized estimator's MSE and that of existing ratio variance estimator defined by Isaki (1983):

$MSE(t_1)-MSE(\widehat{S}^2_{d_i})_{\min}>0$, if

$$S_y^4\left[\lambda\beta^*_{2y}+\left(\lambda-\lambda'\right)(\beta^*_{2x}-2\lambda^*_{22})\right]-S_y^4\left(1-\frac{K}{M}\right)>0$$

$$\left[\lambda\beta^*_{2y}+\left(\lambda-\lambda'\right)(\beta^*_{2x}-2\lambda^*_{22})\right]-\left(1-\frac{K}{M}\right)>0 \tag{44}$$

If condition II, in Eq. 44, is satisfied, then the new generalized estimator is better and more efficient than the ratio estimator defined by Isaki (1983).

Comparison between the new generalized estimator's MSE and that of existing product variance estimator defined by Murthy (1964):

$MSE(t_2)-MSE(\widehat{S}^2_{d_i})_{\min}>0$, if

$$S_y^4\left[\lambda\beta^*_{2y}+\left(\lambda-\lambda'\right)(\beta^*_{2x}+2\lambda^*_{22})\right]-S_y^4\left(1-\frac{K}{M}\right)>0$$

$$\left[\lambda\beta^*_{2y}+\left(\lambda-\lambda'\right)(\beta^*_{2x}+2\lambda^*_{22})\right]-\left(1-\frac{K}{M}\right)>0 \tag{45}$$

If condition III, in Eq. 45, is satisfied, then the new generalized estimator is better and more efficient than the product estimator defined by Murthy (1964).

Comparison between the new generalized estimator's MSE and that of existing difference variance estimator defined by Singh *et al.* (1988):

$MSE(t_3)_{\min} - MSE(\widehat{S}^2_{d_i})_{\min} > 0$, if

$$S_y^4 \left[\frac{MSE(t_{reg})_{\min}}{1 + \frac{MSE(t_{reg})_{\min}}{S_y^4}} \right] - S_y^4 \left(1 - \frac{K}{M} \right) > 0$$

$$\left[\frac{MSE(t_{reg})_{\min}}{1 + \frac{MSE(t_{reg})_{\min}}{S_y^4}} \right] - \left(1 - \frac{K}{M} \right) > 0 \tag{46}$$

If condition IV, in Eq. 46, is satisfied, then the new generalized estimator is better and more efficient than the difference variance estimator defined by Singh *et al.* (1988).

Comparison between the new generalized estimator's MSE and that of modified variance estimator defined by Shabbir and Gupta (2007):

$MSE(t_3)_{\min} - MSE(\widehat{S}^2_{d_i})_{\min} > 0$, if

$$\left[\frac{MSE(t_{reg})_{\min}}{1 + \frac{MSE(t_{reg})_{\min}}{S_y^4}} - \frac{(\lambda - \lambda')\beta_{2x}^* \left[MSE(t_{reg})_{\min} + \frac{(\lambda - \lambda')S_y^4 \beta_{2x}^*}{16} \right]}{4 \left[1 + \frac{MSE(t_{reg})_{\min}}{S_y^4} \right]} \right] - \left(1 - \frac{K}{M} \right) > 0 \tag{47}$$

If condition V, in Eq. 47, is satisfied, then the new generalized estimator is better and more efficient than the modified variance estimator defined by Shabbir and Gupta (2007).

Comparing the proposed estimator's MSE with that of ratio exponential variance estimators defined by Mishra *et al* (2019), we have:

$MSE(Pl_1)_{\min} - MSE(\widehat{S}^2_{d_i})_{\min} > 0$, if

$$S_y^4 \left[\lambda \beta_{2y}^* - \left(\lambda - \lambda' \right) \frac{\lambda_{22}^*}{\beta_{2x}^*} \right] - S_y^4 \left(1 - \frac{K}{M} \right) > 0$$

$$\left[\lambda \beta_{2y}^* - \left(\lambda - \lambda' \right) \frac{\lambda_{22}^*}{\beta_{2x}^*} \right] - \left(1 - \frac{K}{M} \right) > 0 \tag{48}$$

If condition VI, in Eq. 48, is satisfied, then the new generalized estimator is better and more efficient than the first modified variance estimator defined by Mishra *et al.* (2019).

$MSE(Pl_2)_{\min} - MSE(\widehat{S}^2_{d_i})_{\min} > 0$, if

$$\left[C + \frac{BC^2 + (A-2C)D^2}{D^2 - AB}\right] - \left(1 - \frac{K}{M}\right) > 0 \tag{49}$$

If condition VII, in Eq. 49, is satisfied, then the new generalized estimator is better and more efficient than the second modified variance estimator defined by Mishra *et al.* (2019).

$MSE(Pl_3)_{\min} - MSE(\widehat{S}^2_{d_i})_{\min} > 0$, if

$$\left[F_1 + \frac{B_1C_1^2 + A_1D_1^2 - 2C_1D_1^2}{E_1^2 - A_1B_1}\right] - \left(1 - \frac{K}{M}\right) > 0 \tag{50}$$

If condition VIII, in Eq. 50, is satisfied, then the new generalized estimator is better and more efficient than the third modified variance estimator defined by Mishra *et al.* (2019).

$MSE(Pl_4)_{\min} - MSE(\widehat{S}^2_{d_i})_{\min} > 0$, if

$$\left[C_3 + \frac{B_3C_3^2 + A_3D_3^2 - 2C_3D_3^2}{E_3^2 - A_3B_3}\right] - \left(1 - \frac{K}{M}\right) > 0 \tag{51}$$

If condition IX, in Eq. 51, is satisfied, then the new generalized estimator is better and more efficient than the fourth modified variance estimator defined by Mishra *et al.* (2019).

4. Results and Discussion

This section presents the empirical result of the study. Five datasets were used to evaluate the efficiencies of the proposed estimator over other estimators considered in this paper. The MSE and percentage relative efficiencies (PREs) were obtained using dataset I, II, III, IV and V by substituting the parameters of the datasets into the MSE equation of the proposed and existing estimators. The PRE were obtained using relationship between the variance of sample mean and MSE values of the proposed and existing estimators.

Table 2: Estimators' Bias

Estimators	Dataset I	Dataset II	Dataset III	Dataset IV	Dataset V
Usual Variance (t_0)	-	-	-	-	-
Isaki (1983) Classical Ratio (t_1)	13.433	4344.903	41.574	70.586	250.428
Murthy (1964) Classical Product (t_2)	29.233	236.295	19.269	63.802	414.798
Singh *et al.* (1988) (t_3)	-130.614	-10109.400	-331.227	-449.594	-2249.425
Shabbir and Gupta (2007) (t_4)	-153.791	-11977.010	-66.629	505.011	-2618.417
Mishra *et al.* (2019) Estimator (Pl_1)	14.617	2180.476	59.646	121.901	207.399
Mishra *et al.* (2019) Estimator (Pl_2)	-49.560	-6065.220	-84.511	-129.361	-310.850
Mishra *et al.* (2019) Estimator (Pl_3)	-45.002	-2528.050	-81.370	-136.429	-610.107
Mishra *et al.* (2019) Estimator (Pl_4)	-45.711	-3526.030	-88.474	-147.163	-591.052
Proposed Estimator ($\hat{S}^2_{d_1}$)	-13.324	-158.087	-17.155	-63.637	-177.560
Proposed Estimator ($\hat{S}^2_{d_2}$)	-7.496	-155.171	-16.172	-62.451	-119.935

Table 2 presents the results of the bias of the existing and proposed generalized estimators using datasets I, II, III, IV and V. Based on the results obtained from Dataset I, it is observed that the proposed generalized ratio-product cum regression-type estimators; $\hat{S}^2_{d_1}$ and $\hat{S}^2_{d_2}$ (-13.324 and -7.496), respectively, minimizes the bias values compared to the Isaki (1983) classical ratio (13.433); Murthy (1964) classical product (29.233); Singh *et al.,* (1988) estimator (-130.614); Shabbir and Gupta (2007) ratio-regression-type estimator (-153.791) and Mishra *et al.,* (2019) ratio estimators (14.617, -49.560, -45.002 and -45.711), respectively. The results obtained from Dataset II revealed that the proposed generalized ratio-product cum regression-type estimators; $\hat{S}^2_{d_1}$ and $\hat{S}^2_{d_2}$ (-158.087 and -155.171) , respectively, minimizes the bias values compared to the Isaki (1983) classical ratio (4344.903); Murthy (1964) classical product (236.295); Singh *et al.,* (1988) estimator (-10109.400); Shabbir and Gupta (2007) ratio-regression-type estimator (-11977.010) and Mishra *et al.,* (2019) ratio estimators (2180.476, -6065.220, -2528.050 and -3526.030), respectively. Similarly, the results obtained from dataset III revealed that the proposed generalized ratio-product cum regression-type estimators; $\hat{S}^2_{d_1}$ and $\hat{S}^2_{d_2}$ (-17.155 and -16.172) , respectively, minimizes the bias values compared to the Isaki (1983) classical ratio (49.574); Murthy (1964) classical product (19.269); Singh *et al.,* (1988) estimator (-331.227); Shabbir and Gupta (2007) ratio-regression-type estimator (-66.629) and

Mishra *et al.,* (2019) ratio estimators (59.646, -84.511, -81.370 and -88.474), respectively. Also, the results obtained from dataset IV signifies that the proposed generalized ratio-product cum regression-type estimators; $\hat{S}^2_{d_1}$ and $\hat{S}^2_{d_2}$ (-63.637 and -62.451) , respectively, minimizes the bias values compared to the Isaki (1983) classical ratio (70.586); Murthy (1964) classical product (63.802); Singh *et al.,* (1988) estimator (-449.594); Shabbir and Gupta (2007) ratio-regression-type estimator (505.011) and Mishra *et al.,* (2019) ratio estimators (121.901, -129.361, -136.429 and -147.163), respectively. Further, the results obtained from dataset V revealed that the proposed generalized ratio-product cum regression-type estimators; $\hat{S}^2_{d_1}$ and $\hat{S}^2_{d_2}$ (-177.560 and -119.935) , respectively, minimize the bias values compared to the Isaki (1983) classical ratio (250.428); Murthy (1964) classical product (414.798); Singh *et al.,* (1988) estimator (-2249.425); Shabbir and Gupta (2007) ratio-regression-type estimator (-2618.417) and Mishra *et al.,* (2019) ratio estimators (207.399, -310.850, -610.107 and -591.052), respectively.

Therefore, the results indicate that the special classes of the proposed generalized estimator possessed the minimum bias in comparison with the existing estimators. Further, the results showed greater gains by the special classes of the proposed estimator over existing ones in situation where the study and auxiliary variables are negatively correlated (dataset III and V), and where they are positively correlated (dataset I, II and IV).

Table 3 presents the results of the MSE of some existing and proposed generalized estimators using datasets I, II, III, IV and V. Based on the results obtained from Dataset I, it is observed that the proposed generalized ratio-product cum regression-type estimators; $\hat{S}^2_{d_1}$ and $\hat{S}^2_{d_2}$ (3218.826 and 1888.781), respectively, have minimum MSE values compared to the usual variance (9286.713); Isaki (1983) classical ratio (5894.493); Murthy (1964) classical product (30998.810); Singh *et al.,* (1988) estimator (6500.175); Shabbir and Gupta (2007) ratio-regression-type estimator (6192.213) and Mishra *et al.,* (2019) ratio estimators (4986.558, 4499.742, 4431.230 and 3787.798), respectively.

Table 3: Mean Square Errors

Estimators	Dataset I	Dataset II	Dataset III	Dataset IV	Dataset V
Usual Variance (t_0)	9286.713	46584427.000	37511.482	93891.915	2206658.000
Isaki (1983) Classical Ratio (t_1)	5894.493	97443119.000	46699.414	110988.323	901273.800
Murthy (1964) Classical Product (t_2)	30998.810	109143104.000	78448.803	222823.806	6845466.000
Singh *et al.* (1988) (t_3)	6500.175	46151619.000	35381.320	82190.662	1520040.000
Shabbir and Gupta (2007) (t_4)	6192.213	42617938.000	34252.375	78956.410	1472894.000
Mishra *et al.* (2019) Estimator (Pl_1)	4986.558	46433557.000	34997.722	83185.784	881690.700
Mishra *et al.* (2019) Estimator (Pl_2)	4499.742	35634961.000	29013.426	69081.414	825016.305
Mishra *et al.* (2019) Estimator (Pl_3)	4431.230	34932834.000	28912.620	66300.387	787613.412
Mishra *et al.* (2019) Estimator (Pl_4)	3787.798	34257837.000	27639.360	64315.864	703163.134
Proposed Estimator ($\hat{S}^2_{d_1}$)	3218.826	24781026.000	24173.932	55359.893	648085.511
Proposed Estimator ($\hat{S}^2_{d_2}$)	1888.781	13368041.000	19942.831	42865.790	455801.917

The results obtained from dataset II also revealed that the proposed generalized ratio-product cum regression-type estimators; $\hat{S}^2_{d_1}$ and $\hat{S}^2_{d_2}$ (24781026.0 and 13368041.0), respectively, have minimum MSE values compared to the usual variance (46584427.0); Isaki (1983) classical ratio (97443119.0); Murthy (1964) classical product (109143104.0); Singh *et al.,* (1988) estimator (46151619.0); Shabbir and Gupta (2007) ratio-regression-type estimator (42617938.0) and Mishra *et al.,* (2019) ratio estimators (46433557.0, 35634961.0, 34932834.0 and 34257837.0), respectively. Also, the results obtained from Dataset III revealed that the proposed generalized ratio-product cum regression-type estimators; $\hat{S}^2_{d_1}$ and $\hat{S}^2_{d_2}$ (24173.932 and 19942.831), respectively, have minimum mean square error values compared to the usual variance (37511.482); Isaki (1983) classical ratio (46699.414); Murthy (1964) classical product (78448.803); Singh *et al.,* (1988) estimator (35381.320); Shabbir and Gupta (2007) ratio-regression-type estimator (34252.375) and Mishra *et al.,* (2019) ratio estimators (34997.722, 29013.426, 28912.620 and 27639.360), respectively. The results obtained from Dataset IV signifies that the proposed generalized ratio-product cum regression-type estimators; $\hat{S}^2_{d_1}$ and $\hat{S}^2_{d_2}$ (55359.893 and 42865.790), respectively, have minimum MSE values compared to the usual variance (93891.915); Isaki (1983) classical ratio (110988.323); Murthy (1964) classical product (222823.806); Singh *et al.,* (1988) estimator (82190.662); Shabbir and Gupta (2007) ratio-regression-type estimator (78956.410) and Mishra *et al.,* (2019) ratio estimators (83185.784, 69081.414, 66300.387 and 64315.864), respectively.

Further, the results obtained from Dataset V revealed that the proposed generalized ratio-product cum regression-type estimators; $\hat{S}^2_{d_1}$ and $\hat{S}^2_{d_2}$ (648085.511 and 455801.917), respectively, have minimum MSE values compared to the usual variance (2206658.0); Isaki (1983) classical ratio (901273.80); Murthy (1964) classical product (6845466.0); Singh *et al.,* (1988) estimator (1520040.0); Shabbir and Gupta (2007) ratio-regression-type estimator (1472894.0) and Mishra *et al.,* (2019) ratio estimators (881690.70, 825016.305, 787613.412 and 703163.134), respectively. Therefore, the results obtained from I, II, III, IV and V indicate that the special classes of the proposed generalized estimator possessed the minimum MSE in comparison with the existing estimators. The special classes of the proposed estimator perform better based on the MSE when the study and auxiliary variables are negatively correlated (dataset III and V), and when they are positively correlated (dataset I, II and IV).

The results of the PRE of the existing and proposed generalized estimators using datasets I, II, III, IV and V are presented in Table 4. Based on the results obtained from dataset I, it is observed that the proposed generalized ratio-product cum regression-type estimators; $\hat{S}^2_{d_1}$ and $\hat{S}^2_{d_2}$ (288.512 and 491.678), respectively, have the highest PRE values compared to the usual variance (100.0); Isaki (1983) classical ratio (157.549); Murthy (1964) classical product (29.958); Singh *et al.,* (1988) estimator (142.869); Shabbir and Gupta (2007) ratio-regression-type estimator (149.974) and Mishra *et al.,* (2019) ratio estimators (186.235, 206.383, 209.574 and 245.175), respectively. The results obtained from dataset II also revealed that the proposed generalized ratio-product cum regression-type estimators; $\hat{S}^2_{d_1}$ and $\hat{S}^2_{d_2}$ (187.984 and 348.476), respectively, have the highest PRE values compared to the usual variance (100.0); Isaki (1983) classical ratio (47.807); Murthy (1964) classical product (42.682); Singh *et al.,* (1988) estimator (100.938); Shabbir and Gupta (2007) ratio-regression-type estimator (109.307) and Mishra *et al.,* (2019) ratio estimators (100.325, 130.727, 133.354 and 135.982), respectively.

Table 4: Estimates of the Percentage Relative Efficiency

Estimators	Dataset I	Dataset II	Dataset III	Dataset IV	Dataset V
Usual Variance (t_0)	100.000	100.000	100.000	100.000	100.000
Isaki (1983) Classical Ratio (t_1)	157.549	47.807	80.325	84.596	244.838
Murthy (1964) Classical Product (t_2)	29.958	42.682	47.817	42.137	32.235
Singh *et al.* (1988) (t_3)	142.869	100.938	106.021	114.237	145.171
Shabbir and Gupta (2007) (t_4)	149.974	109.307	109.515	118.916	149.818
Mishra *et al.* (2019) Estimator (Pl_1)	186.235	100.325	107.183	112.870	250.276
Mishra *et al.* (2019) Estimator (Pl_2)	206.383	130.727	129.290	135.915	267.468
Mishra *et al.* (2019) Estimator (Pl_3)	209.574	133.354	129.741	141.616	280.170
Mishra *et al.* (2019) Estimator (Pl_4)	245.175	135.982	135.718	145.986	313.819
Proposed Estimator $(\hat{S}^2_{d_1})$	288.512	187.984	155.173	169.603	340.489
Proposed Estimator $(\hat{S}^2_{d_2})$	491.678	348.476	188.095	219.037	484.127

More so, the results obtained from Dataset III revealed that the proposed generalized ratio-product cum regression-type estimators; $\hat{S}^2_{d_1}$ and $\hat{S}^2_{d_2}$ (155.173 and 188.095), respectively, have the highest PRE values compared to the usual variance (100.0); Isaki (1983) classical ratio (80.325); Murthy (1964) classical product (47.817); Singh *et al.,* (1988) estimator (106.021); Shabbir and Gupta (2007) ratio-regression-type estimator (109.515) and Mishra *et al.,* (2019) ratio estimators (107.183, 129.290, 129.741 and 135.718), respectively. The results obtained from Dataset IV signifies that the proposed generalized ratio-product cum regression-type estimators; $\hat{S}^2_{d_1}$ and $\hat{S}^2_{d_2}$ (169.603 and 219.037), respectively, have the highest PRE values compared to the usual variance (100.0); Isaki (1983) classical ratio (84.596); Murthy (1964) classical product (42.137); Singh *et al.,* (1988) estimator (114.237); Shabbir and Gupta (2007) ratio-regression-type estimator (118.916) and Mishra *et al.,* (2019) ratio estimators (112.870, 135.915, 141.616 and 145.986), respectively. Further, the results obtained from Dataset V revealed that the proposed generalized ratio-product cum regression-type estimators; $\hat{S}^2_{d_1}$ and $\hat{S}^2_{d_2}$ (340.489 and 484.127), respectively, have the highest PRE values compared to the usual variance (100.0); Isaki (1983) classical ratio (244.838); Murthy (1964) classical product (32.235); Singh *et al.,* (1988) estimator (145.171); Shabbir and Gupta (2007) ratio-regression-type estimator (149.818) and Mishra *et al.,* (2019) ratio estimators (250.276, 267.468, 280.170 and 313.819), respectively.

Therefore, the results obtained from dataset I, II, III, IV and V show that the special classes of the proposed generalized estimator possessed higher PRE in comparison with the existing estimators. Thus, the special classes of the proposed estimator showed greater gains in efficiency when the study and auxiliary variables are negatively correlated (dataset III and V), and when they are positively correlated (dataset I, II and IV).

5. Conclusion and Policy Implication

5.1 Conclusion

Controlling the variation of phenomena such as oil and gas prices, inflation and exchange rates, unemployment rate and so on is difficult in application and this has received attentions of many researchers. This study under two-phase sampling suggests a new generalized ratio-product cum regression-type estimator with some special classes that provides efficient and precise estimation of phenomena variation for better policy formulation. The new generalized estimator is obtained by applying a linear combination approach, power transformation and exponentiation strategy to some existing estimators. The properties and theoretical efficiency conditions of the new generalized estimator are derived up to first degree of approximation. The MSE and PRE of the proposed and some existing estimators are compared empirically using five real datasets. Evidence suggests that the classes of the proposed estimator perform better in terms of bias, MSE and PRE compared to existing estimators in the literature.

5.2 Policy Implication

The findings of this study are relevant for policy makers and industry analysts. The findings could be helpful by providing better variation estimates for various phenomena such as inflation, exchange rate, standard of living among other macro fundamentals for better policy-making. Also, real-world problems like estimating how income varies across different society, or how the number of deaths from a particular disease varies over a decade, or how the contribution to pollution varies among different types of industries over time, and so on can be solved by the proposed estimator.

For future work, the estimator can be modified by considering multi-auxiliary variables to enhance its precision and efficiency, since supplementary information has proven to enhance the precision and flexibility of estimators.

References

Ahmed, M. S., Raman, M. S., & Hossain, M. I. (2000). Some competitive estimators of finite population variance using multivariate auxiliary information. *Information Management Science,* 11(1), 49–54.

Ahmed, M. S., Walid, A. D., & Ahmed, A. O. H. (2003). Some estimators for finite population variance under two-phase sampling. *Statistics Transition,* 6(1), 143-150.

Al-Jararha, J. & Ahmed, M. S. (2002). The class of chain estimators for a finite population variance using double sampling. *Information Management Science*, 13(2), 13-18.

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, 246-254.

Bahl, S. & Tuteja, R. K. (1991). Ratio and product type exponential estimators. *Information Optimization Science*, 12(1), 159-164.

Central Bank of Nigeria (2014). The Structure of the Nigerian Economy and Implications for Development, *Realm Communication Ltd,* Lagos.

Central Bank of Nigeria (2010). *Statistical Bulletin, 50 years Special Anniversary Edition, Abuja*

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

Grover, L. K. (2010). A correction note on improvement in variance estimation using auxiliary information. *Communication in Statistics-Theory and Methods*, 43, 101–105.

Isaki, C. T. (1983). Variance estimation using auxiliary information. *Journal of American Statistics Association*, 78, 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), 1047–1059.

Mishra, P., Singh, R. & Khare, S. (2019). Estimation of population variance in log-product type estimators under double sampling scheme. *Journal of Reliability and Statistical Studies,* 12(2), 11-20.

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

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), 14-30.

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 mean using auxiliary variable in simple random sampling. *NIPES Journal of Science and Technology Research*, 5(1): 242-252.

Murthy, M. N. (1964). Product method of estimation. *Sankhya A*, 26, 69-74.

Neyman, J. (1938). Contribution to the theory of sampling human populations. *Journal of the American Statistical Association*, 33(201), 101-116.

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), 44-58.

Shabbir, J. & Gupta, S. (2007). On improvement in variance estimation using auxiliary information. *Communication in Statistical-Theory and Methods*, 36, 2177–2185.

Singh, G. N., Bhattacharyya, D. & Bandyopadhyay, A. (2020). A general class of calibration estimators under stratified random sampling in presence of various kinds of non-sampling errors. *Communications in Statistics - Simulation and Computation*, 52, 20-32.

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

Singh, R. & Malik, S. (2014). Improved estimation of population variance using information on auxiliary attribute in simple random sampling. *Applied Mathematics and Computation*. 235, 43–49.

Singh, H. P. & Singh, R. (2001) Improved ratio type estimator for variance using auxiliary information. *Journal of Indian Society of Agricultural Statistics.* 54(3), 276–287.

Singh, H. P. & Solanki, R. S. (2013). Improved estimation of finite population variance using auxiliary information. *Communication in Statistics-Theory and Methods*, 42, 2718–2730.

Singh, H. P., Upadhyaya, L. N. & Namjoshi, U. D. (1988). Estimation of finite population variance. *Curr Science*, 57, 1331–1334.

Sukhatme, B. V. (1962). Some ratio-type estimators in two-phase sampling. *Journal of the American Statistical Association*, 57(299), 628–632.

Sukhatme, P. V. & Sukhatme, B. V. (1970). *Sampling Theory of Surveys with Applications*. Ames, IA: Iowa State University Press.

Sumramani, J. & Kumarapandiyan, G. (2015). Generalized modified ratio type estimator for estimation of population variance. *Sri-Lankan Journal of Applied Statistics*, 16(1), 69–90.

Swain, A. K. P. C. (2015). Generalized estimator of finite population variance. *Journal of Statistical Theory and Application*, 14(1), 45–51.

Upadhyaya, L. N. & Singh, H. P. (1999). An estimator of population variance that utilizes the kurtosis of an auxiliary variable in sample surveys. *Vikram Mathematics Journal*, 19, 14–17.

Upadhyaya, L. N., Singh, H. P. & Singh, S. (2004). A class of estimators for estimating the variance of the ratio estimator. *Journal of Japan Statistical Society*, 34(1), 47–63.

Yadav, S. K. & Kadilar, C. (2014). A two parameter variance estimator using auxiliary information. *Applied Mathematics and Computation*, 226, 117–122.

Yadav, S. K., Kadilar, C. & Shabbir, J. (2015). Improved family of estimators of population variance in simple random sampling. *Journal of Statistical Theory Practice*, 9(2), 219–226.

Yaqub, M. & Shabbir, J. (2016). An improved class of estimators for finite population variance. *Hacettepe Journal of Mathematical Statistics*. 45(5), 1641–1660.

Zakari, Y & Muhammad, I. (2022). Improved product estimator using measurable auxiliary character. Proceedings of the 58^{th} 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 of the 4^{th} International Electronic Conference on Applied Sciences of Science Forum. Published by *Multidisciplinary Digital Publishing Institute (MDPI).*

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., Rashida, A., 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 of the 4^{th} International Electronic Conference on Applied Sciences of Science Forum. Published by *Multidisciplinary Digital Publishing Institute (MDPI).*