A MODIFIED RATIO-PRODUCT ESTIMATOR OF POPULATION MEAN USING SOME KNOWN PARAMETERS OF THE AUXILIARY VARIABLE

Umar Kabir Abdullahi and Abubakar Yahaya
Department of Statistics, Ahmadu Bello University, Zaria, Nigeria.
umarkabir9@gmail.com, +2348109554170.
abubakaryahaya@abu.edu.ng, +2348069334466.

ABSTRACT
The estimation of population mean is one of the challenging aspects in sampling theory and population study and much effort has been vigorously employed to improve the precision of estimates. In this research work, a modified ratio-product estimator of population mean $\overline{Y}$ of the study variable Y using median and coefficient of variation of the auxiliary variable X, in simple random sampling scheme is proposed. The expression of bias and MSE of the proposed estimator have been obtained under large sample approximation, asymptotically optimum estimator (AOE) is identified with its approximate MSE formula. Estimator based on "estimated optimum values" was also investigated. Theoretical and empirical comparison of proposed estimator with some other ratio and product estimators justified the performance of the proposed estimators. A minimum of 20 percent reduction in the MSE were observed from each of the existing estimators considered. It is found that the proposed estimator were uniformly better than all other modified ratio and product estimators and thus most preferred over the existing estimators for the use in practical application.
Keywords: Finite population mean, bias, mean square error, auxiliary variable, optimum estimator, study variable.

INTRODUCTION
The use of auxiliary variable (supplementary variable) has been widely discussed in sampling theory. Auxiliary variables are in use in survey sampling to obtain improved sampling designs and to achieve more precision in the estimates of some population parameters such as the mean, or the variance of the variable under study. This information may be use at both the stages of designing a survey (design, execution and estimation stages). It is well established that ratio, linear regression, product are some good examples in this context.

The estimation of population parameters, particularly mean is one of the challenging aspects in sampling theory and much effort has been vigorously employed to improve the precision of estimates. In sample survey literature, great classes of techniques for using auxiliary information by means of ratio, regression and product methods have been discussed in the presence of single and/or multi-auxiliary variables. A wide variety of estimators were proposed following distinct notions; different possible linear combinations and mixing together ratio, product, and/or linear regression estimators, each one exploiting the variables one at a time. It is also established that, if the regression line of the variable under study and the auxiliary variable passes through the origin and are positively correlated the best estimator to be used is ratio estimator (Singh and Espejo, 2003). In the same vein, if the regression line of the variable under study and the auxiliary variable pass through the origin but are negatively correlated the best estimator to be used is product estimator. On the other hand when the regression line does not passes through the origin

but makes an intercept along the y-axis and there is correlation either positive or negative between the auxiliary variable (x) and the variable of interest (Y) the best estimator to use is the linear regression estimator. It is observed though, that all these methods yield biased estimators and of course bias decreases with increase in sample size (Okafor, 2002).

BACKGROUND OF THE STUDY
Consider a finite population $P = (P_1, P_2, ..., P_N)$ of N units, let a sample be drawn using simple random sampling without replacement (SRSWOR), let y_i and x_i represent the values of a response variable y and auxiliary variable x respectively. The units of the finite population are identifiable in the sense that they are uniquely labeled from 1 to N and the label on each unit is known. Further, suppose in a sample survey problem, we are interested in estimating the population mean $\overline{Y}$ of y, when the population parameters of the auxiliary variable X such as population mean $\overline{X}$, knowledge of coefficient variation C_x, the coefficient of skewness $\beta_1(x)$, coefficient of Kurtosis $\beta_2(x)$, median M_d of auxiliary variable x are known. Before discussing further about the modified ratio estimators and the proposed modified ratio estimator, the notation used in this research work are described below.

- N – population size
- n – sample size
- $f = {n}/{N}$ –sampling fraction
- Y – study variable
- X – auxiliary variable
- $\overline{Y}, \overline{X}$ –population means
- $\overline{y}, \overline{x}$ –sample means
- S_y, S_x – sample standard deviations
- C_y, C_x –co-efficient of variations
- ρ – co-efficient of correlation
- $\beta_1 = \dfrac{N\sum_{i=1}^{N}\left(x_i - \overline{X}\right)^3}{(N-1)(N-2)S^3}$, coefficient of skewness of the auxiliary variable
- $\beta_2 = \dfrac{N(N+1)\sum_{i=1}^{N}\left(x_i - \overline{X}\right)^4}{(N-1)(N-2)(N-3)S^4} - \dfrac{3(N-1)^2}{(N-2)(N-3)}$, coefficient of Kurtosis of the auxiliary variable
- M_d – median of the auxiliary variable
- $B(.)$ – bias of the estimator
- $MSE(.)$ – mean square error of the estimator

The traditional ratio and product estimators for $\overline{Y}$ are respectively given as,

$$T_{R1} = \overline{y}\left(\frac{\overline{X}}{\overline{x}}\right) \tag{2.1}$$

$$T_{P1} = \overline{y}\left(\frac{\overline{x}}{\overline{X}}\right) \tag{2.2}$$

To the first degree of approximation the bias and mean squared errors of T_{R1}, T_{P1} are respectively as,

$$bias\left(T_{R1}\right) = \frac{(1-f)}{n}\overline{Y}\left[C_x^2\left\{1-K\right\}\right] \tag{2.3}$$

$$bias\left(T_{P1}\right) = \frac{(1-f)}{n}\overline{Y}KC_x^2 \tag{2.4}$$

$$MSE\left(T_{R1}\right) = \frac{(1-f)}{n}\overline{Y}^2\left[C_y^2 + C_x^2\left\{1-2K\right\}\right] \tag{2.5}$$

$$MSE\left(T_{P1}\right) = \frac{(1-f)}{n}\overline{Y}^2\left[C_y^2 + C_x^2\left\{1+2K\right\}\right] \tag{2.6}$$

The modified ratio and product estimator by Subramani and Kumarapandiyan (2012) of $\overline{Y}$ are respectively given as,

$$T_{R2} = \overline{y}\left(\frac{\overline{X}C_x + M_d}{\overline{x}C_x + M_d}\right) \tag{2.7}$$

$$T_{P2} = \overline{y}\left(\frac{\overline{x}C_x + M_d}{\overline{X}C_x + M_d}\right) \tag{2.8}$$

To the first degree of approximation the bias and mean squared errors of T_{R2}, T_{P2} are respectively as,

$$bias\left(T_{R2}\right) = \frac{(1-f)}{n}\overline{Y}\left[\theta C_x^2\left\{\theta - K\right\}\right] \tag{2.9}$$

$$bias\left(T_{R2}\right) = \frac{(1-f)}{n}\overline{Y}\left[\theta C_x^2\left\{\theta - K\right\}\right] \tag{2.10}$$

$$MSE\left(T_{R2}\right) = \frac{(1-f)}{n}\overline{Y}^2\left[C_y^2 + \theta C_x^2\left\{\theta - 2K\right\}\right] \tag{2.11}$$

$$MSE\left(T_{P2}\right) = \frac{(1-f)}{n}\overline{Y}^2\left[C_y^2 + \theta C_x^2\left\{\theta + 2K\right\}\right] \tag{2.12}$$

where $\theta = \dfrac{\overline{X}C_x}{\overline{X}C_x + M_d}$

Proposed estimator

Motivated by Housila and Neha-Agnihotri (2008) and Subramani and Kumarapandiyan (2012), we defined a ratio-product estimator of population mean using coefficient of variation and median of the auxiliary variable in simple random sampling as

$$T_{RP} = \left[\delta\overline{y}\left(\frac{\overline{X}C_x + M_d}{\overline{x}C_x + M_d}\right) + (1-\delta)\overline{y}\left(\frac{\overline{x}C_x + M_d}{\overline{X}C_x + M_d}\right)\right] \tag{3.1}$$

Where $\overline{y} = \dfrac{1}{n}\sum_{i=1}^{n} y_i$ and $\overline{x} = \dfrac{1}{n}\sum_{i=1}^{n} x_i$ are unbiased estimators of the population means $\left(\overline{Y}, \overline{X}\right)$ respectively,

where 'C_x' coefficient of variation and 'M_d' median are the known characteristics positive scalars of the auxiliary variable X respectively and δ is a real constant to be determined such that the mean square error of T_{RP} is minimum. The family of estimators T_{RP} reduces to the following set of known estimators,

(i) For $\left(C_x, M_d, \delta\right) = \left(0,1,\delta\right)$, $T_{RP} \rightarrow \overline{y}$ (usual unbiased estimator)

(ii) For $\left(C_x, M_d, \delta\right) = \left(0,1,\delta\right)$, $T_{RP} \rightarrow T_1 = \left[\delta\overline{y}\left(\dfrac{\overline{X}}{\overline{x}}\right) + (1-\delta)\overline{y}\left(\dfrac{\overline{x}}{\overline{X}}\right)\right]$

which is due to Singh and Ruiz Espejo (2003)

(iii) For $\left(C_x, M_d, \delta\right) = \left(1, C_x, \delta\right)$, $T_{RP} \rightarrow T_2 = \left[\delta\overline{y}\left(\dfrac{\overline{X} + C_x}{\overline{x} + C_x}\right) + (1-\delta)\overline{y}\left(\dfrac{\overline{x} + C_x}{\overline{X} + C_x}\right)\right]$

envisaged by Singh and Tailor (2005), where C_x is the known population coefficient of variation and M_d is the median of the auxiliary X respectively, many other ratio-product estimators can be generated from T_{RP} by putting any suitable parameters rather than values of $\left(C_x, M_d, \delta\right)$.

bias and Mean Square Error

To obtain the bias and mean square error (MSE) of the proposed estimator T_{RP} in (3.1), we write

Let $e_0 = \dfrac{(\overline{y} - \overline{Y})}{\overline{Y}}$ and $e_1 = \dfrac{(\overline{x} - \overline{X})}{\overline{X}}$

Then $\overline{y} = \left(1 + e_0\right)$ and $\overline{x} = \left(1 + e_1\right)$

Such that $E\left(e_0\right) = E\left(e_1\right) = 0$

$E\left(e_0^2\right) = \dfrac{1-f}{n}C_y^2$, $E\left(e_1^2\right) = \dfrac{1-f}{n}C_x^2$, $E\left(e_0 e_1\right) = \dfrac{1-f}{n}KC_x^2$

Where $f = \dfrac{n}{N}$, $K = \rho\dfrac{C_y}{C_x}$, $\rho = \dfrac{S_{xy}}{S_x S_Y}$, $C_y = \dfrac{S_y}{\overline{Y}}$, $C_x = \dfrac{S_x}{\overline{X}}$.

$$S_{xy} = \frac{\sum_{i=1}^{N}(x_i - \overline{X})(y_i - \overline{Y})}{(N-1)}, \ S_x^2 = \frac{\sum_{i=1}^{N}(x_i - \overline{X})^2}{(N-1)}, \ S_y^2 = \frac{\sum_{i=1}^{N}(y_i - \overline{Y})^2}{(N-1)}.$$

And expanding (3.1) in terms of e's, we have

$$T_{RP} = \overline{Y}(1 + e_0)\left[\delta(1 + \theta e_1)^{-1} + (1 - \delta)(1 + \theta e_1)\right] \tag{3.2}$$

We assume that $|\theta e_1| < 1$, so that the expression $(1 + \theta e_1)^{-1}$ can be expanded to a convergent infinite series using binomial theorem. Hence from (3.2) we have.

$$T_{RP} = \overline{Y}(1 + e_0)\left[\delta(1 - \theta e_1 + \theta^2 e_1^2 - \theta^3 e_1^3 + \theta^4 e_1^4 \ldots) + (1 - \delta)(1 + \theta e_1)\right]$$

$$= \overline{Y}\left[\delta(1 + e_0)(1 - \theta e_1 + \theta^2 e_1^2 - \theta^3 e_1^3 + \theta^4 e_1^4 \ldots) + (1 - \delta)(1 + e_0)(1 + \theta e_1)\right]$$

$$= \overline{Y}\left[\delta(1 + e_0 - \theta e_1 + \theta^2 e_1^2 - \theta e_1 e_0 + \theta^2 e_1 e_0 + \theta^2 e_1^2 e_0 - \theta^3 e_1^3 + \theta^4 e_1^4 - \theta^3 e_1^3 e_0 + \ldots) + (1 - \delta)(1 + e_0 + \theta e_1 + \theta e_1 e_0)\right]$$

$$= \overline{Y}\left[1 + e_0 + \theta e_1 + \theta e_1 e_0 + \delta(1 + e_0 - \theta e_1 - \theta e_1 e_0 + \theta^2 e_1^2 + \theta^2 e_1^2 e_0 - \theta^3 e_1^3 + \theta^4 e_1^4 - \theta^3 e_1^3 e_0 + \ldots - 1 - e_0 - \theta e_1 - \theta e_0 e_1)\right]$$

$$= \overline{Y}\left[1 + e_0 + \theta e_1 + \theta e_1 e_0 + \delta(-2\theta e_1 e_0 - 2\theta e_0 e_1 + \theta^2 e_1^2 + \theta^2 e_1^2 e_0 - \theta^3 e_1^3 + \theta^4 e_1^4 - \theta^3 e_1^3 e_0 + \ldots)\right]$$

$$= \overline{Y}\left[1 + e_0 - (1 - 2\delta)\theta e_1 + (1 - 2\delta)\theta e_1 e_0 + \delta\theta^2 e_1^2 + \delta(\theta^2 e_1^2 e_0 - \theta^3 e_1^3 + \theta^4 e_1^4 - \theta^3 e_1^3 e_0 + \ldots)\right],$$

We assume that the contribution of terms involving powers in e_0 and e_1 higher than the second is negligible, being of order $\frac{1}{n^v}$, where $v > 1$. Thus, from the above expression we write to a first order of approximation,

$$T_{RP} \cong \overline{Y}\left[1 + e_0 - (1 - 2\delta)\theta e_1 + (1 - 2\delta)\theta e_1 e_0 + \delta\theta^2 e_1^2 + \delta\theta^2 e_1^2 e_0\right], \ \text{or}$$

$$(T_{RP} - \overline{Y}) = \overline{Y}\left[e_0 - (1 - 2\delta)\theta e_1 + (1 - 2\delta)\theta e_1 e_0 + \delta\theta^2 e_1^2 + \delta\theta^2 e_1^2 e_0\right], \tag{3.3}$$

Taking the expectation of both side of (3.3), we obtained the bias of (T_{RP}) to the first degree of approximation as

$$bias(T_{RP}) = \frac{(1 - f)}{n}\theta\overline{Y}\left[K + \delta(\theta - 2K)\right]C_x^2 \tag{3.4}$$

will vanishes if

$$\delta = \frac{K}{(2K - \theta)} \tag{3.5}$$

Thus for $\delta = \frac{K}{(2K - \theta)}$, T_{RP} is almost unbiased.

Squaring both side of the equation (3.3), and neglecting the terms of $e's$ having power greater than two we have

$$(T_{RP} - \overline{Y})^2 = \overline{Y}^2\left[e_0^2 + (1 - 2\delta)\theta\left\{(1 - 2\delta)\theta e_1^2 + 2e_1 e_0\right\}\right] \tag{3.6}$$

Taking the expectation of both sides of (3.4), we get the mean square error MSE of T_{RP} to the first order of approximation as

$$MSE(T_{RP}) = \frac{(1 - f)}{n}\overline{Y}^2\left[C_y^2 + \theta(1 - 2\delta)C_x^2\left\{(1 - 2\delta)\theta + 2K\right\}\right] \tag{3.7}$$

To obtain the value of δ that minimizes the MSE of (T_{RP}), we take the partial derivative of the MSE of (T_{RP}) with respect to δ and equate it to zero.

$$\delta = \frac{1}{2}\left(1 + \frac{K}{\theta}\right) = \delta_0 \ \text{(Optimum Value)} \tag{3.8}$$

Putting (3.8) in (3.1), we get the Asymptotically Optimum Estimator (AOE) as

$$T_{RPO} = \frac{\overline{y}}{2}\left[\left(1 + \frac{K}{\theta}\right)\left(\frac{\overline{X}C_x + M_d}{\overline{x}C_x + M_d}\right) + \left(1 - \frac{K}{\theta}\right)\left(\frac{\overline{x}C_x + M_d}{\overline{X}C_x + M_d}\right)\right] \tag{3.9}$$

Substitution of (3.8) in (3.7) yield the minimum MSE of (T_{RP}) or the MSE of asymptotically optimum estimator (AOE) T_{RP} as

$$MSE_{min}\left(T_{RP}\right) = \frac{(1-f)}{n} S_y^2 \left(1 - \rho^2\right) = MSE\left(T_{RPO}\right) \tag{3.10}$$

which is equal to the approximate MSE of regression estimator

$$\overline{y}_{lr} = \overline{y} + \hat{\beta}\left(\overline{X} - \overline{x}\right) \tag{3.11}$$

Where $\hat{\beta} = s_{xy}/s_x^2$, is the sample estimate of the population regression coefficient β of y on x.

$$s_{xy} = \frac{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)\left(y_i - \overline{y}\right)}{(n-1)}, s_x^2 = \frac{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2}{(n-1)}.$$

It is to be noted the AOE of $\left(T_{RPO}\right)$ in (3.9) depends on K and θ, so the AOE of $\left(T_{RPO}\right)$ can be used in practice only when K and θ are known. Here it should be mentioned that θ is a function of known quantities $\left(C_x, M_d, \delta\right)$. So only the value of K should be known for making the use of AOE of $\left(T_{RPO}\right)$ in practice. The value of K can be made known quite accurately either from pilot study or past data or experience gathered in due course of time. This problem has been discussed among others by Murthy (1967), Reddy (1978), Srivankataramana and Tracy (1980). Thus, the value of K can be guessed quite accurately and such an estimator can be used in practice.

Allowable Departure

Let k_0 be an estimate or guessed value of K with

$$k_0 = K\left(1 + \eta\right), \text{ then}$$

$$\delta = \frac{1}{2}\left(1 + \frac{K}{\theta}\right) = \frac{1}{2}\left(1 + \frac{K}{\theta} + \frac{1}{\theta}\left(k_0 - K\right)\right) = \delta_0 + \frac{\eta K}{2\theta} \tag{4.1}$$

Putting (4.1) in (3.7) we obtain the MSE of $\left(T_{RP}\right)$ as

$$MSE(T_{RP}) = \left(T_{RPO}\right) + \left(\frac{1-f}{n}\right)\rho^2\eta^2 S_y^2$$

$$\Rightarrow MSE(T_{RP}) - MSE(T_{RPO}) = \left(\frac{1-f}{n}\right)\rho^2\eta^2 S_y^2$$

$$\Rightarrow \frac{MSE(T_{RP}) - MSE(T_{RPO})}{MSE(T_{RPO})} = \left(\frac{\rho^2\eta^2}{\left(1-\rho^2\right)}\right) \tag{4.2}$$

It follows from (3.25) that the proportional increase in MSE of $\left(T_{RP}\right)$ over that of AOE of $\left(T_{RP}\right)$ is less than γ if,

$$\frac{\rho^2\eta^2}{\left(1-\rho^2\right)} < \gamma$$

i.e. $|\eta| < \sqrt{\frac{\left(1-\rho^2\right)}{\rho^2}\gamma}$, $\tag{4.3}$

Which clearly shows that to ensure only a small relative increase in MSE of $\left(T_{RP}\right)$, $|\eta|$ must be in the neighborhood of "zero" if ρ is high but can depart substantially from "zero" if ρ is moderate.

Efficiency Comparisons

In order to compare the efficiency of the proposed a new ratio-product estimator with the various existing estimators we require the expressions of the MSE of the estimators, up to the first order of approximation. It is well known that under SRSWOR that

$$Var\left(\overline{y}\right) = \left(\frac{1-f}{n}\right)S_y^2 \tag{5.1}$$

From (3.7) and (5.1) we have

$$Var\left(\overline{y}\right) - MSE\left(T_{RP}\right) = \left(\frac{1-f}{n}\right)\theta\overline{Y}^2 C_x^2\left(2\delta - 1\right)\left[\left(1 - 2\delta\right)\theta + 2K\right]$$

which is non-negative if

$$\min\left\{\frac{1}{2}, \frac{1}{2}\left(1 + \frac{2K}{\theta}\right)\right\} < \delta < \max\left\{\frac{1}{2}, \frac{1}{2}\left(1 + \frac{2K}{\theta}\right)\right\} \tag{5.2}$$

It is to be noted that for $\delta = 1$, the estimator T_{RP} reduces to the ratio-type estimator

$$T_{R2} = \bar{y}\left(\frac{\bar{X}C_x + M_d}{\bar{x}C_x + M_d}\right) \tag{5.3}$$

where for $\delta = 0$, the estimator T_{RP} turns out to be the product-type estimator

$$T_{R2} = \bar{y}\left(\frac{\bar{x}C_x + M_d}{\bar{X}C_x + M_d}\right) \tag{5.4}$$

To the first degree of approximation the mean squared errors of T_{R2}, T_{p2} are respectively given by

$$MSE(T_{R2}) = \frac{(1-f)}{n}\bar{Y}^2\left[C_y^2 + \theta C_x^2\{\theta - 2K\}\right] \tag{5.5}$$

$$MSE(T_{P2}) = \frac{(1-f)}{n}\bar{Y}^2\left[C_y^2 + \theta C_x^2\{\theta + 2K\}\right] \tag{5.6}$$

From (3.7), (5.5) and (5.6), we have

$$MSE(T_{R2}) - MSE(T_{RP}) = \frac{4(1-f)}{n}\theta\bar{Y}^2 C_x^2\left[(1-\delta)(\delta\theta - K)\right]$$

(5.7)

$$MSE(T_{P2}) - MSE(T_{RP}) = \frac{4(1-f)}{n}\theta\bar{Y}^2\delta C_x^2\left[\theta(1-\delta) + K\right] \tag{5.8}$$

It follows from (5.7) and (5.8) that the ratio-product estimator T_{RP} is more efficient than

(i) The ratio type estimator T_{R2} if

$$\min\left(\frac{K}{\theta}, 1\right) < \delta < \max\left(\frac{K}{\theta}, 1\right) \tag{5.9}$$

(ii) The product- type estimator T_{P2} if

$$\min\left(\left(1 + \frac{K}{\theta}\right), 0\right) < \delta < \max\left(\left(1 + \frac{K}{\theta}\right), 0\right) \tag{5.10}$$

Further, if we set $(C_x, M_d) = (1, 0)$ in (5.3) and (5.4) the ratio-type estimator T_{R2} and product-type T_{P2} estimators respectively reduces to

$$T_{R2} \rightarrow \bar{y}\frac{\bar{X}}{\bar{x}} \text{ (Usual ratio estimator)} \tag{5.11}$$

$$T_{P2} \rightarrow \bar{y}\frac{\bar{x}}{\bar{X}} \text{ (Usual product estimator)} \tag{5.12}$$

Putting $(C_x, M_d) = (1, 0)$ in (5.3) and (5.6) we get the mean squared errors of usual ratio and product estimators respectively as

$$MSE(T_{R1}) = \frac{(1-f)}{n}\bar{Y}^2\left[C_y^2 + C_x^2\{1 - 2K\}\right] \tag{5.13}$$

$$MSE(T_{P1}) = \frac{(1-f)}{n}\bar{Y}^2\left[C_y^2 + C_x^2\{1 + 2K\}\right] \tag{5.14}$$

From (3.7), (5.13) and (5.14) we have

$$MSE(T_{R1}) - MSE(T_{RP}) = \frac{(1-f)}{n}\bar{Y}^2 C_x^2(1 + \theta - 2\delta\theta)(1 - \theta - 2K + 2\delta\theta) \tag{5.15}$$

$$MSE(T_{P1}) - MSE(T_{RP}) = \frac{(1-f)}{n}\bar{Y}^2 C_x^2(1 - \theta + 2\delta\theta)(1 + \theta + 2K - 2\delta\theta) \tag{5.16}$$

From (5.15) and (5.16) we note that the ratio–product estimator T_{RP} is better than

(i) The ratio type estimator T_{R1} if

$$\min\left\{\left(\frac{1+\theta}{2\theta}\right), \left(\frac{2K+\theta-1}{2\theta}\right)\right\} < \delta < \max\left\{\left(\frac{1+\theta}{2\theta}\right), \left(\frac{2K+\theta-1}{2\theta}\right)\right\} \tag{5.17}$$

(ii) The product- type estimator T_{p1} if

$$\min\left\{\left(\frac{\theta-1}{2\theta}\right), \left(\frac{2K+\theta+1}{2\theta}\right)\right\} < \delta < \max\left\{\left(\frac{\theta-1}{2\theta}\right), \left(\frac{2K+\theta+1}{2\theta}\right)\right\} \tag{5.18}$$

Estimator based on the Optimum value

The optimum value of δ in (3.8) is given by

$$\delta_0 = \frac{1}{2}\left(1 + \frac{K}{\theta}\right) \tag{6.1}$$

Where θ is a known quantity and

$$K = \rho\frac{C_y}{C_x} = \rho\frac{S_y\overline{X}}{S_x\overline{Y}} = \frac{S_{xy}}{RS_x^2} = \frac{\beta}{R}$$

Replacing β and R by their consistent estimators

$$\hat{\beta} = \frac{s_{xy}}{s_x^2} \text{ and } \widehat{R} = \frac{\overline{y}}{\overline{x}} \text{ respectively.}$$

From (6.1) we get a consistent estimator of δ_0 as

$$\hat{\delta}_0 = \frac{1}{2}\left(1 + \frac{\hat{K}}{\theta}\right) \tag{6.2}$$

Where $\hat{K} = \frac{\hat{\beta}}{\widehat{R}}$

If the experimenter is unable to guess the value of K, then it is worth advisable to replace K by $\widehat{K}$ in (3.21). Thus, the estimator based on the estimated 'optimum' value.

$$\hat{T}_{RPO} = \frac{\overline{y}}{2}\left[\left(1 + \frac{\hat{K}}{\theta}\right)\left(\frac{\overline{X}C_x + M_d}{\overline{x}C_x + M_d}\right) + \left(1 + \frac{\hat{K}}{\theta}\right)\left(\frac{\overline{x}C_x + M_d}{\overline{X}C_x + M_d}\right)\right] \tag{6.3}$$

To obtain the MSE of $\hat{T}_{RPO}$ we write

Let $e_2 = \frac{(\hat{k} - K)}{K}$, then $\hat{k} = K(1 + e_2)$

with $E(e_2) = K + o(n^{-1})$, expanding (6.3) in terms of $e's$ we have

$$\hat{T}_{RPO} = \frac{\overline{Y}}{2}(1 + e_0)\left[\left\{1 + \frac{K}{\theta}(1 + e_2)\right\}\left(1 + \theta e_1 + \theta^2 e_1^2 + \theta^3 e_1^3 + \theta^4 e_1^4 + ...\right) + \left\{1 - \frac{K}{\theta}(1 + e_2)\right\}(1 + \theta e_1)\right] \tag{6.4}$$

From (6.4) we have

$$= \frac{\overline{Y}}{2}(1 + e_0)\left[1 + \frac{K}{\theta}(1 + e_2) - \theta e_1 - K(e_1 + e_1 e_2) + K\theta(e_1^2 + e_2 e_1^2) + ... + 1 - \frac{K}{\theta}(1 + e_2) + \theta e_1 - K(e_1 + e_1 e_2)\right]$$

$$= \frac{\overline{Y}}{2}(1 + e_0)\left[2 - 2K(e_1 + e_1 e_2) + K\theta(e_1^2 + e_2 e_1^2) + ...\right]$$

$$= \overline{Y}(1 + e_0)\left[1 - K(e_1 + e_1 e_2) + \frac{K\theta}{2}(e_1^2 + e_2 e_1^2) + ...\right]$$

$$= \overline{Y}(1 + e_0)\left[1 + e_0 - K(e_1 + e_0 e_1 + e_1 e_2 + e_0 e_1 e_2) + \frac{K\theta}{2}(e_1^2 + e_2 e_1^2 + e_0 e_1^2 + e_0 e_2 e_1^2) + ...\right]$$

Neglecting the terms of $e's$ having power greater than two we have

$$\hat{T}_{RPO} = \overline{Y}\left[1 + e_0 - Ke_1 - K(e_0 e_1 + e_1 e_2) + \frac{K\theta}{2}e_1^2\right] \text{ or}$$

$$(\hat{T}_{RPO} - \overline{Y}) = \overline{Y}\left[e_0 - Ke_1 - K(e_0 e_1 + e_1 e_2) + \frac{K\theta}{2}e_1^2\right] \tag{6.5}$$

Now squaring both sides of (6.5) and neglecting the terms of $e's$ having power greater than the second we have

$$(\hat{T}_{RPO} - \overline{Y})^2 = \overline{Y}^2\left[e_0^2 + K^2 e_1^2 - 2Ke_0 e_1\right] \tag{6.6}$$

Taking the expectation of both sides of (6.6) we get the mean square error (MSE) of $\hat{T}_{RPO}$ to the first order of approximation as

$$MSE(\hat{T}_{RPO}) = E(\hat{T}_{RPO} - \overline{Y})^2 = \overline{Y}^2 E\left[e_0^2 + K^2 e_1^2 - 2K(e_0 e_1)\right]$$

$$= \frac{(1-f)}{n}\overline{Y}^2\left[C_y^2 + K^2 C_x^2 - 2K\rho C_y C_x\right]$$

$$= \frac{(1-f)}{n}\overline{Y}^2\left[C_y^2 + \left(\rho\frac{C_y}{C_x}\right)^2 C_x^2 - 2\left(\rho\frac{C_y}{C_x}\right)\rho C_y C_x\right]$$

$$= \frac{(1-f)}{n}\overline{Y}^2\left(C_y^2 - \rho^2 C_y^2\right)$$

$$= \frac{(1-f)}{n}\overline{Y}^2 C_y^2\left(1 - \rho^2\right)$$

$$= \frac{(1-f)}{n}S_y^2\left(1 - \rho^2\right) \tag{6.7}$$

Which is equal to the minimum MSE of T_{RP} or the MSE of T_{RPO} given by (3.10). Thus, we established the result that the MSE of the estimator $\hat{T}_{RPO}$ in (6.3) based on 'estimated optimum value' to the first degree of approximation is same as that of $\hat{T}_{RPO}$ in (3.9). 134 teresting to note the estimator $\hat{T}_{RPO}$ in (6.3) can be used as an alternative to the $\hat{T}_{RPO}$ in (3.9) if the value of the parameter is not known.

Numerical Comparison

To see the merit of the suggested estimator T_{RP} over $\overline{y}$, T_{R1}, T_{P1}, T_{R2} and T_{P2}.

Dataset: The populations considered in this study are real-life dataset comprises of population 1 and 2. The population 1 is taken from Singh and Chaudary (1986) given in page 177 and population 2 is taken from Murthy (1964) given in page 228. Table 1 below summarized the dataset for the study.

Table 1: Summary of the Dataset

Parameters	Population 1: Y = output X = fixed capital	Population 2: Y = output X = number of workers
N	80	80
n	20	20
$\overline{Y}$	51.8264	51.8264
$\overline{X}$	11.2646	2.8513
ρ	0.941	0.9150
C_y	0.3542	0.3542
C_x	0.7507	0.9484
M_d	7.575	1.4800

Therefore, we see the merit of the proposed estimators T_{RP} over some existing Ratio and Product estimators $\overline{y}$, T_{R1}, T_{P1}, T_{R2} and T_{P2} in **Table 2** and **Table 3**.

Table 2: Ranges of δ under which the proposed estimator T_{RP} is better than $\overline{y}$, T_{R1}, T_{P1}, T_{R2} and T_{P2}.

Populations ⟶	1	2
$\overline{y}$	$0.5 < \delta < 1.340$	$0.5 < \delta < 1.029$
T_{R2}	$0.842 < \delta < 1.000$	$0.529 < \delta < 1.000$
T_{P2}	$0.00 < \delta < 1.842$	$0.000 < \delta < 1.529$
T_{R1}	$0.394 < \delta < 1.448$	$0.255 < \delta < 1.274$
T_{P1}	$-0.448 < \delta < 2.2896$	$-0.274 < \delta < 1.802$
δ	0.921	0.764

Table 3: The variance and MSE of $\overline{y}$, T_{R1}, T_{P1}, T_{R2} and T_{P2} and the proposed estimator T_{RP}

Populations $\longrightarrow$	1	2
$\overline{y}$	12.637	12.637
T_{R2}	1.843	10.461
T_{P2}	55.018	90.498
T_{R1}	18.995	63.216
T_{P1}	119.804	143.252
T_{RP}	1.447	2.057

The Percentage relative efficiency (PRE) of different estimators T respect to $\overline{y}$ is defined as,

$$PRE\left(T., \overline{y}\right) = \frac{V\left(\overline{y}\right)}{V\left(T.\right)} X\,100$$

Table 4: Percentage relative efficiency of different estimators with respect to $\overline{y}$

Populations $\downarrow$	$\overline{y}$	T_{R1}	T_{P1}	T_{R2}	T_{P2}	T_{RP}
1	100	< 100	< 100	685.69	< 100	873.23
2	100	< 100	< 100	120.80	< 100	614.35

DISCUSSION

We have proposed an estimator of the ratio-product estimator and obtained the asymptotically optimum estimator (AOE) with its approximate MSE formula for the proposed estimator using the coefficient of variation and median of the auxiliary variable X in simple random sampling.

Theoretically, we have demonstrated that the proposed estimator is always more efficient than other estimators $\overline{y}$, T_{R1}, T_{P1}, T_{R2} and T_{P2} under the effective ranges of δ and its optimum values.

In addition, we support these theoretical results numerically using the data sets as shown in **Table 1.**

Table 2 provides the wide ranges of δ along with its optimum values for which the proposed estimator T_{RP} is more efficient than other estimators $\overline{y}$, T_{R1},

T_{P1}, T_{R2} and T_{P2} as far as mean squared error criterion is considered. It is also observed from **Table 2** that there is a scope for chosen δ to obtain better estimators than $\overline{y}$, T_{R1}, T_{P1}, T_{R2} and T_{P2}.

Table 4 provides that there is a considerable gain in efficiency by using proposed estimator T_{RP} over the estimators $\overline{y}$, T_{R1}, T_{P1}, T_{R2} and T_{P2}. This shows that even if the scalar δ deviates from its optimum values $\left(\delta_{opt}\right)$. The suggested estimator T_{RP} will yield better estimate than $\overline{y}$, T_{R1}, T_{P1}, T_{R2} and T_{P2}.

$$MSE\left(\overline{y}\right) - MSE\left(T_{RP}\right) = \frac{\left(1-f\right)}{n}\overline{Y}^2 C_y^2 > 0\,,$$

$$MSE\left(T_{R1}\right) - MSE\left(T_{RP}\right) = \frac{\left(1-f\right)}{n}\overline{Y}^2 C_x^2 \left(1+K\right)^2 > 0$$

$$MSE\left(T_{P1}\right) - MSE\left(T_{RP}\right) = \frac{\left(1-f\right)}{n}\overline{Y}^2 C_x^2 \left(1-K\right)^2 > 0$$

$$MSE\left(T_{R2}\right) - MSE\left(T_{RP}\right) = \frac{\left(1-f\right)}{n}\overline{Y}^2 C_x^2 \left(\theta+K\right)^2 > 0$$

$$MSE\left(T_{P2}\right) - MSE\left(T_{RP}\right) = \frac{\left(1-f\right)}{n}\overline{Y}^2 C_x^2 \left(\theta-K\right)^2 > 0$$

.

CONCLUSION

Evidence from the study revealed that the proposed estimator is more efficient than the already existing ratio, product and ratio- product type estimators based on some certain conditions and efficiency conditions. Therefore, there is always need to ensure that the auxiliary variable is highly correlated with the study variable and the population under consideration is homogeneously distributed, and when there is no correlation between the auxiliary variable and the study variable, the application of single- phase sample will not yield more efficient or the mean per unit will be more efficient.

Hence we conclude that the proposed class of estimator T_{RP} is more efficient than the other estimators in case of its optimality in simple random sampling. Thus, it is preferred to use the proposed estimator T_{RP} in practice over $\bar{y}, T_{R1}, T_{P1}, T_{R2}$ and T_{P2}.

REFERENCES

Cochran, W.G. (1940). The estimation of the yield of the cereal experiments by sampling for ratio of grain to total produce. *The journal of Agricultural Science*,30: 262-275.

Das, A.K. and Tripathi, T.P. (1981). A class of sampling strategies for population mean using information on mean and variance of an auxiliary character. *Proceedings of Italian Statistical Institute, Golden Jubilee International Conference of Statistics. Applications and new directions*,174-181.

Housila, P.S. and Neha-Agnihotri (2008). A general procedure of estimating population mean using auxiliary information in sample surveys. *Statistics in Transition-new series*, 9(1):71-87.

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

Murthy, M. N. (1967). *Sampling Theory and Methods.* Calcutta, India: Statistical Publishing Society.

Okafor, F. C. (2002). *Sample Survey Theory and Applications (1ˢᵗ edition).* Nsukka, Nigeria.

Perri, P.F. (2005). Combination of auxiliary variables in ratio-cum-product type estimators. In: *Proceedings of Italian Statistical Society, Intermediate meeting on Statistics and Environment, Messina, Italy, September 2005,*pp. 193-196.

Reddy, V.N. and Singh, H.P. (1978). A study on the use of prior knowledge on certain population parameters in estimation, *Sankhya C,* 40: 29-37.

Singh, H.P. and Tailor, R. (2003). Estimation of finite population mean using known correlation coefficient of auxiliary characters. *Statistica, Anno LXV*, 4:407-418.

Singh, H.P. and Tailor, R. (2005). Estimation of finite population mean with known coefficient of variation of auxiliary characteristic. *Statistica, anno LXV*, (3):301-313.

Singh, D. and Chaudary, F.S. (1986). *Theory and analysis of sample survey Designs.* New Age International Publisher.pp. 177-178.

Singh, H.P. and Ruiz Espejo, M. (2003). On linear regression and ratio-product estimation of a finite population mean. *The Statistician,* 52(1): 59-67.

Sisodia B.V.S. and Dwivedi, V.K. (1981). A modified ratio estimator using coefficient of variation of auxiliary variable*. Journal of Indian Society of Agricultural Statistics.* 33(1) :13-18.

Srivenkataramana, T. and Tracy, D. S. (1980). On ratio and product methods of estimation in sampling. *Statistica Neerlandica, 33*: 37-49.

Subramani, J. and Kumarapandiyan, G. (2012). Estimation of population mean using co-efficient of variation and median of the auxiliary variable. *International Journal of Probability Statistics,* 1(4): 111-118, DOI 10.5923/j. ijps.20120104.04.