Punjab University
Journal of Mathematics (ISSN 1016-2526)
Vol. 49(2)(2017) pp. 59-73

A Modified Factor-type Estimator under Two-phase Sampling

A. Audu
Department of Mathematics,
Usmanu Danfodiyo University, P. M. B. 2346, Sokoto, Nigeria,
Email: ahmed.audu@udusok.edu.ng

A. A. Adewara
Department of Statistics,
University of Ilorin, P. M. B 1515, Ilorin, Kwara State, Nigeria,
Email: aaadewara@gmail.com

Received: 12 August, 2016 / Accepted: 19 December, 2016 / Published online: 25 May, 2017

Abstract. In this paper, a modified factor-type estimator under two-phase sampling has been suggested. The suggested estimator is obtained by incorporating information like coefficient of variation, kurtosis, skewness and correlation coefficient in Shukla [9] estimator. The bias and MSE of the estimator have been obtained under the conditions of optimality. The efficiency of the proposed estimator with respect to the conventional estimators considered in this study was compared empirically and the results showed that the proposed estimator is more efficient. It is also observed that the suggested estimator is more efficient than the corresponding estimators in case of fixed cost and incurred minimum cost for specified precision.

AMS (MOS) Subject Classification Codes: 35S29; 40S70; 25U09
Key Words: Estimator, Auxiliary variable, Mean square error (MSE), Two-phase sampling.

1. Introduction

The efficiency of ratio and product estimators solely depends on strong-positive and strong-negative correlations respectively between the study and auxiliary variables. In 1993, Singh and Shukla suggested conventional factor-type estimator which is applicable when correlations between the study and auxiliary variables are either positive or negative. Ratio, product, factor-type and dual-to-ratio estimators completely depend on the knowledge of population mean $\bar{X}$ of auxiliary variable. However, knowledge about population mean $\bar{X}$ may be unknown before the commencement of a survey due to limited resources

(time, money and human) and lack of accessibility to all units of the population. In such circumstances, the usual things to do is to estimate it by sample mean $\bar{x}_1$ based on the preliminary sample of size n_1 such that $n_1 > n_2$. Also, judicious utilization of other information of auxiliary variable like coefficient of variation, kurtosis and correlation coefficient either at planning stage or at the design stage or at estimation stage helps to arrive at an improved estimator compared to those, not utilizing auxiliary information (Choudhury and Singh [1]). In this paper, a modified factor-type estimator under two-phase sampling using functions of auxiliary variable mentioned above has been suggested and its properties have been established.
Let $\Omega = \{1, 2, ..., N\}$ be a population of size N and $Y,\ X > 0$ be two real valued vectors in $\Re^N$ on the i^{th} unit of $\Omega\,(1 \le i \le N)$. Consider a preliminary large sample S_1 of size n_1 drawn from population Ω by SRSWOR and secondary sample S_2 of size $n_2\,(n_2 < n_1)$ drawn either of the following manners:
Case I: as a subset from the preliminary sample i.e. $S_2 \subset S_1$.
Case II: as an independent sample from population i.e. $S_2 \subset \Omega$.
Shukla [9] suggested a factor-type estimator for population mean under two-phase sampling as;

$$\bar{y}_{FTd} = \bar{y}\frac{(A+C)\,\bar{x}_1 + fB\bar{x}}{(A+fB)\,\bar{x}_1 + C\bar{x}} \tag{1. 1}$$

where $\bar{x} = \frac{1}{n_2}\sum_{i=1}^{n_2} x_i,\ \ \bar{x}_1 = \frac{1}{n_1}\sum_{i=1}^{n_1} x_i\ \ \bar{y} = \frac{1}{n_2}\sum_{i=1}^{n_2} y_i$
d is an unknown positive real number to be estimated i.e $d \in \Re^+$
The bias and MSE of $\bar{y}_{FTd}$ under case I and II are respectively

$$Bias\,(\bar{y}_{FTd})_I = \bar{Y}P\theta_3\left[\rho_{xy}C_xC_y - \psi_4 C_x^2\right] \tag{1. 2}$$

$$Bias\,(\bar{y}_{FTd})_{II} = \bar{Y}P\left[(\theta_1\psi_3 - \theta_2\psi_4)\,C_x^2 + \theta_2\rho_{xy}C_xC_y\right] \tag{1. 3}$$

$$MSE\,(\bar{y}_{FTd})_I = \bar{Y}^2\left[\theta_2 C_y^2 + \theta_3 P^2 C_x^2 + 2\theta_3 P\rho_{xy}C_xC_y\right] \tag{1. 4}$$

$$MSE\,(\bar{y}_{FTd})_{II} = \bar{Y}^2\left[\theta_2 C_y^2 + \theta_4 P^2 C_x^2 + 2\theta_2 P\rho_{xy}C_xC_y\right] \tag{1. 5}$$

where $\theta_1 = \frac{1}{n_1} - \frac{1}{N}\ ,\ \theta_2 = \frac{1}{n_2} - \frac{1}{N}\ ,\ \theta_3 = \frac{1}{n_2} - \frac{1}{n_1}\ ,\ \theta_4 = \theta_1 + \theta_2$

$$\psi_1 = \frac{A+C}{A+fB+C},\ \psi_2 = \frac{fB}{A+fB+C},\ \psi_3 = \frac{A+fB}{A+fB+C},\ \psi_4 = \frac{C}{A+fB+C}$$

$P = \psi_3 - \psi_1 = \psi_2 - \psi_4$
We observed that the factor-type estimator $\bar{y}_{FTd}$ was more efficient than classical ratio estimator $\bar{y}_r^d$ if $-2C_{yx} < P < 0$ in case I and if $-2C_{yx}\,(1+\delta)^{-1} < P < 0$ in case II where $\delta = \theta_1\theta_2^{-1}$.

2. SUGGESTED CLASS OF ESTIMATOR

After studying the related estimators mentioned in section 1 and motivated by the work of Singh and Agnihotri [5] defined as

$$\bar{y}_{SA} = \bar{y}\left[\delta\left(\frac{a_x\bar{X}+b_x}{a_x\bar{x}+b_x}\right)+(1-\delta)\left(\frac{a_x\bar{x}+b_x}{a_x\bar{X}+b_x}\right)\right] \tag{2. 6}$$

where $a_x > 0$ and $b_x > 0$ are some known parameters of auxiliary variable X such as standard deviation σ_x, coefficient of variation C_x, skewness $\beta_{1(x)}$, kurtosis $\beta_{2(x)}$ and correlation coefficient ρ_{xy}, for instance, see Upadhyaya and Singh [11] and Singh and Tailor [8]. We therefore suggest the estimator for population mean given by

$$\bar{y}^{(d)}_{FTAA} = \bar{y}\frac{\left[(A+C)\,\bar{x}_1+fB\bar{x}\right]a_x+\left[A+C+fB\right]b_x}{\left[(A+fB)\,\bar{x}_1+C\bar{x}\right]a_x+\left[A+C+fB\right]b_x} \tag{2. 7}$$

TABLE 1. Some Member of the Family of Estimator $\bar{y}^{(d)}_{FTAA}$

SN	a_x	b_x	d	Deduced Estimator
1	$a_x \neq 0$	0	1	$\bar{y}_2\frac{\bar{x}_1}{\bar{x}_2}$ Sukhatme [9]
2	1	ρ_{xy}	1	$\bar{y}_2\frac{\bar{x}_1+\rho_{xy}}{\bar{x}_2+\rho_{xy}}$ Malik and Tailor [4]
3	$a_0 \neq 0$	0	2	$\bar{y}_2\frac{\bar{x}_2}{\bar{x}_1}$ product estimator
4	$a_0 \neq 0$	0	3	$\bar{y}_2\frac{N\bar{x}_1-n\bar{x}_2}{(N-n)\bar{x}_1}$ Two-phase Dual-to-ratio estimator
5	1	0	d	$\bar{y}_2\frac{(A+C)\bar{x}_1+fB\bar{x}_2}{(A+fB)\bar{x}_1+C\bar{x}_2}$ Shukla [8]
6	$a_0 \neq 0$	0	4	$\bar{y}$

3. BIAS AND MEAN SQUARE ERROR OF THE SUGGESTED ESTIMATOR

In order to study the properties of the suggested estimator, we define $\Delta_{\bar{y}_2} = \left(\bar{y}_2-\bar{Y}\right)/\bar{Y},\ \Delta_{\bar{x}_1} = \left(\bar{x}_1-\bar{X}\right)/\bar{X},\ \Delta_{\bar{x}_2} = \left(\bar{x}_2-\bar{X}\right)/\bar{X}$ such that $|\Delta_{\bar{y}_2}|<1,\ |\Delta_{\bar{x}_1}|<1,\ |\Delta_{\bar{x}_2}|<1.$

Equation (2.7) can be expressed in terms of $\Delta_{\bar{y}_2}$, $\Delta_{\bar{x}_1}$ and $\Delta_{\bar{x}_2}$ as

$$\bar{y}^{(d)}_{FTAA} = \bar{Y}\left[1+\Delta_{\bar{y}_2}\right]\left[1+\delta_x\psi_1\Delta_{\bar{x}_1}+\delta_x\psi_2\Delta_{\bar{x}_2}\right]\left[1+\delta_x\psi_3\Delta_{\bar{x}_1}+\delta_x\psi_4\Delta_{\bar{x}_2}\right]^{-1} \tag{3. 8}$$

where $\delta_x = \frac{a_x\bar{X}}{a_x\bar{X}+b_x}$

We now assume that $|\delta_x\psi_3\Delta_{\bar{x}_1}+\delta_x\psi_4\Delta_{\bar{x}_2}|<1$ and $(1+\delta_x\psi_3\Delta_{\bar{x}_1}+\delta_x\psi_4\Delta_{\bar{x}_2})^{-1}$ is expandable.

Using power series expansion, the simplification of equation (3.8) up to first order approximation $O\left(n^{-1}\right)$ is given by

$$\begin{aligned}\bar{y}^{(d)}_{FTAA}-\bar{Y} = \bar{Y}\big[&\Delta_{\bar{y}_2}-\delta_xP\Delta_{\bar{x}_1}+\delta_xP\Delta_{\bar{x}_2}+\delta_x^2\psi_3P\Delta^2_{\bar{x}_1}-\delta_x^2\psi_4P\Delta^2_{\bar{x}_1}\\ &+\delta_x^2\left(\psi_4-\psi_3\right)P\Delta_{\bar{x}_1}\Delta_{\bar{x}_2}-\delta_xP\Delta_{\bar{y}_2}\Delta_{\bar{x}_1}+\delta_xP\Delta_{\bar{y}_2}\Delta_{\bar{x}_2}\big]\end{aligned} \tag{3. 9}$$

Under Case I
In case I, we have

$$\left.\begin{array}{l} E\left(\Delta_{\bar{y}_2}\right)=E\left(\Delta_{\bar{x}_2}\right)=E\left(\Delta_{\bar{x}_1}\right)=0,\ E\left(\Delta^2_{\bar{x}_1}\right)=\theta_1 C_x^2 \\ E\left(\Delta^2_{\bar{x}_2}\right)=\theta_2 C_x^2,\ E\left(\Delta^2_{\bar{y}_2}\right)=\theta_2 C_y^2,\ E\left(\Delta_{\bar{y}_2}\Delta_{\bar{x}_1}\right)=\theta_1\rho_{xy}C_xC_y \\ E\left(\Delta_{\bar{y}_2}\Delta_{\bar{x}_2}\right)=\theta_2\rho_{xy}C_xC_y,\ \ E\left(\Delta_{\bar{x}_2}\Delta_{\bar{x}_1}\right)=\theta_1 C_x^2 \end{array}\right\} \quad (3.\ 10)$$

Taking expectation of (3.9) and using results of (3.10), the bias of the suggested estimator to terms of order n^{-1} is obtained as

$$Bias\left(\bar{y}^{(d)}_{FTAA}\right)_I=\bar{Y}P\theta_3\left[\delta_x\rho_{xy}C_xC_y-\delta_x^2\psi_4C_x^2\right] \quad (3.\ 11)$$

Squaring both sides of (3.9), then taking expectation and using results in (3.10), we obtain the MSE of the suggested estimator to terms of order n^{-1} as

$$MSE\left(\bar{y}^{(d)}_{FTAA}\right)_I=\bar{Y}^2\left[\theta_2C_y^2+\theta_3\delta_x^2P^2C_x^2+2\theta_3\delta_xP\rho_{xy}C_xC_y\right] \quad (3.\ 12)$$

Under Case II
In case II, we have

$$\left.\begin{array}{l} E\left(\Delta_{\bar{y}_2}\right)=E\left(\Delta_{\bar{x}_2}\right)=E\left(\Delta_{\bar{x}_1}\right)=0,\ E\left(\Delta^2_{\bar{x}_1}\right)=\theta_1 C_x^2 \\ E\left(\Delta^2_{\bar{x}_2}\right)=\theta_2 C_x^2,\ E\left(\Delta^2_{\bar{y}_2}\right)=\theta_2 C_y^2,\ E\left(\Delta_{\bar{y}_2}\Delta_{\bar{x}_1}\right)=\theta_1\rho_{xy}C_xC_y \\ E\left(\Delta_{\bar{y}_2}\Delta_{\bar{x}_2}\right)=0,\ \ E\left(\Delta_{\bar{x}_2}\Delta_{\bar{x}_1}\right)=0 \end{array}\right\} \quad (3.\ 13)$$

Taking expectation of (3.9) and using results of (3.13), the bias of the suggested estimator to terms of order n^{-1} is obtained as

$$Bias\left(\bar{y}^{(d)}_{FTAA}\right)_{II}=\bar{Y}P\left[\theta_2\delta_x\rho_{xy}C_xC_y+\left(\theta_1\psi_3-\theta_2\psi_4\right)\delta_x^2C_x^2\right] \quad (3.\ 14)$$

Squaring both sides of (3.9), then taking expectation and using results in (3.13), we obtain the MSE of the suggested estimator to terms of order n^{-1} as

$$MSE\left(\bar{y}^{(d)}_{FTAA}\right)_{II}=\bar{Y}^2\left[\theta_2C_y^2+\theta_4\delta_x^2P^2C_x^2+2\theta_2\delta_xP\rho_{xy}C_xC_y\right] \quad (3.\ 15)$$

4. Minimum MSE of $\bar{y}^{(d)}_{FTAA}$

Differentiate (3.12) partially with respect to and equate the result to zero as

$$\frac{\partial}{\partial P}MSE\left(\bar{y}^{(d)}_{FTAA}\right)_I=\bar{Y}^2\theta_3\left[2P\delta_xC_x^2+2\delta_x\rho_{xy}C_xC_y\right]=0 \quad (4.\ 16)$$

$$P=-C_{yx}/\delta_x \quad (4.\ 17)$$

Substitute (4.17) in (3.12), the minimum $MSE\left(\bar{y}^{(d)}_{FTAA}\right)_I$ written as $MSE\left(\bar{y}^{(d)}_{FTAA}\right)_{I\min}$ is obtained as

$$MSE\left(\bar{y}^{(d)}_{FTAA}\right)_{I\min}=\bar{Y}^2C_y^2\left[\theta_2-\theta_3\rho_{xy}^2\right] \quad (4.\ 18)$$

In order to estimate unknown value $d \in \Re^+$, $P = -C_{yx}/\delta_x$ obtained in section 4 and $P = \psi_3 - \psi_1$ defined in section 1 are equated to obtain (4.19)

$$\left(-\frac{C_{yx}}{\delta_x} - 1\right) d^3 + \left(-\frac{C_{yx}}{\delta_x} f + f + 8\frac{C_{yx}}{\delta_x} + 9\right) d^2 - \left(-5f\frac{C_{yx}}{\delta_x} + 5f + 23\frac{C_{yx}}{\delta_x} + 26\right) d + \left(-4f\frac{C_{yx}}{\delta_x} + 4f + 22\frac{C_{yx}}{\delta_x} + 24\right) = 0 \tag{4. 19}$$

Also, differentiate (3.15) partially with respect to P and equate the result to zero as

$$\frac{\partial}{\partial P} MSE\left(\bar{y}_{FTAA}^{(d)}\right)_{II} = \bar{Y}^2 \left[2\theta_4 P \delta_x^2 C_x^2 + 2\theta_2 \delta_x \rho_{xy} C_x C_y\right] = 0 \tag{4. 20}$$

$$P = -\theta_2 C_{yx}/\delta_x \theta_4 \tag{4. 21}$$

Substitute (4.21) in (3.15), the minimum $MSE\left(\bar{y}_{FTAA}^{(d)}\right)_{II}$ written as $MSE\left(\bar{y}_{FTAA}^{(d)}\right)_{II\,\min}$ is obtained as

$$MSE\left(\bar{y}_{FTAA}^{(d)}\right)_{II\,\min} = \bar{Y}^2 C_y^2 \left[\theta_2 - \frac{\theta_2^2}{\theta_4}\rho_{xy}^2\right] \tag{4. 22}$$

In order to estimate unknown value $d \in \Re^+$, $P = -\theta_2 C_{yx}/\delta_x \theta_4$ obtained in section 4 and $P = \psi_3 - \psi_1$ defined in section 1 are equated to obtain (4.23)

$$\left(\frac{-\theta_2 C_{yx}}{\delta_x \theta_4} - 1\right) d^3 + \left(-f\frac{\theta_2 C_{yx}}{\delta_x \theta_4} + f + 8\frac{\theta_2 C_{yx}}{\delta_x \theta_4} + 9\right) d^2 - \left(-5f\frac{\theta_2 C_{yx}}{\delta_x \theta_4} + 5f + 23\frac{\theta_2 C_{yx}}{\delta_x \theta_4} + 26\right) d + \left(-4f\frac{\theta_2 C_{yx}}{\delta_x \theta_4} + 4f + 22\frac{\theta_2 C_{yx}}{\delta_x \theta_4} + 24\right) = 0 \tag{4. 23}$$

By solving (4.19) and (4.23), at most 3 zeros d_1, d_2 and d_3 of the polynomials for which (2.7) is optimal under case I and II respectively will be obtained.

5. Efficiency comparisons

In this section efficiency of the suggested estimator is compared with efficiency of traditional factor-type estimator under case I and case II.
The efficiency of suggested estimator $\bar{y}_{FTAA}^{(d)}$ and estimator $\bar{y}_{FTd}$ suggested by Shukla [9] are compared as,

$$MSE\left(\bar{y}_{FTd}\right)_I - MSE\left(\bar{y}_{FTAA}^{(d)}\right)_I > 0 \tag{5. 24}$$

$$\bar{Y}^2\left[\theta_2 C_y^2 + \theta_3 P^2 C_x^2 + 2\theta_3 P \rho_{xy} C_x C_y\right] - \bar{Y}^2\left[\theta_2 C_y^2 + \theta_3 \delta_x^2 P^2 C_x^2 + 2\theta_3 \delta_x P \rho_{xy} C_x C_y\right] > 0$$

$$P < -\frac{2\rho_{xy} C_y}{[\delta_x + 1]\, C_x} \tag{5. 25}$$

$$MSE\left(\bar{y}_{FTd}\right)_{II} - MSE\left(\bar{y}_{FTAA}^{(d)}\right)_{II} > 0 \tag{5. 26}$$

$$\bar{Y}^2\left[\theta_2 C_y^2+\theta_4 P^2 C_x^2+2\theta_2 P\rho_{xy}C_xC_y\right]- \\ \bar{Y}^2\left[\theta_2 C_y^2+\theta_4\delta_x^2 P^2 C_x^2+2\theta_2\delta_x P\rho_{xy}C_xC_y\right]>0$$

$$P<-\frac{2\frac{\theta_2}{\theta_4}\rho_{xy}C_y}{[\delta_x+1]\,C_x} \tag{5. 27}$$

The suggested estimator $\bar{y}_{FTAA}^{(d)}$ is more efficient than estimator $\bar{y}_{FTd}$ proposed by Shukla [9] if $P<-\frac{2\rho_{xy}C_y}{[\delta_x+1]C_x}$ and $P<-\frac{2\frac{\theta_2}{\theta_4}\rho_{xy}C_y}{[\delta_x+1]C_x}$ when $S_2\subset S_1$ and $S_2\subset\Omega_N$ respectively

6. Determination of n_1 and n_2 for the Fixed Cost C

Let c_0 be the overhead cost, c_1 be the cost of selecting and processing a single unit in the first phase sampling n_1 and c_2 be the cost of selecting and processing a single unit in the second phase sampling n_2. The expected total cost of the survey is given by

$$C=c_0+c_1n_1+c_2n_2 \tag{6. 28}$$

Let us define a function L_1 for minimizing the $MSE\left(\bar{y}_{FTAA}^{(d)}\right)$ and to obtain the optimum values of n_1 and n_2 for the Fixed Cost C which is given as

$$L_1=MSE\left(\bar{y}_{FTAA}^{(d)}\right)-\lambda_1(C-c_0-c_1n_1-c_2n_2) \tag{6. 29}$$

where λ_1 is the Lagrange multiplier.
Now, differentiating L_1 with respect to n_1 and n_2 and equate to zero, we have

$$n_1=\bar{Y}\sqrt{\frac{-2E_2-E_1}{\lambda_1c_1}} \tag{6. 30}$$

$$n_2=\bar{Y}\sqrt{\frac{C_y^2+E_1+2E_2}{\lambda_1c_2}} \tag{6. 31}$$

where $E_1=\delta_x^2P^2C_x^2\,,\ E_2=\delta_xP\rho_{xy}C_xC_y$
By substituting the values of n_1 and n_2 from (6.30) and (6.31) in (6.28), we have

$$\sqrt{\lambda_1}=\bar{Y}E_3/\left(C-c_0\right) \tag{6. 32}$$

where $E_3=\sqrt{c_1\left[-2E_2-E_1\right]}+\sqrt{c_2\left[C_y^2+E_1+2E_2\right]}$
Substituting (6.32) in (6.30) and (6.31), we have

$$n_1=\frac{(C-c_0)}{E_3}\sqrt{\frac{-2E_2-E_1}{c_1}} \tag{6. 33}$$

$$n_2=\frac{(C-c_0)}{E_3}\sqrt{\frac{C_y^2+E_1+2E_2}{c_2}} \tag{6. 34}$$

The optimum n_1 and n_2 are obtained by putting the optimum value of P from (4.17) in (6.33) and (6.34) and we have

$$n_1^{opt} = \frac{(C - c_0)\,\rho_{xy}}{\left(c_1\rho_{xy} + \sqrt{c_1c_2\left[1 - \rho_{xy}^2\right]}\right)} \tag{6. 35}$$

$$n_2^{opt} = \frac{(C - c_0)\sqrt{1 - \rho_{xy}^2}}{\left(\sqrt{c_1c_2}\rho_{xy} + c_2\sqrt{1 - \rho_{xy}^2}\right)} \tag{6. 36}$$

Hence, for the optimum values of n_1 and n_2, the minimum MSE of $\bar{y}_{FTAA}^{(d)}$ having neglecting the term of order $1/N$ is

$$MSE\left(\bar{y}_{FTAA}^{(d)}\right)_{\min} = \bar{Y}^2C_y^2\left[\frac{1}{(C - c_0)}\left\{\sqrt{\left(1 - \rho_{xy}^2\right)c_2} + \sqrt{c_1}\rho_{xy}\right\}^2\right] \tag{6. 37}$$

7. Determination of n_1 and n_2 for specified precision V

Let V be the specified variance of the estimator $\bar{y}_{FTAA}^{(d)}$ fixed in advance. Then we have

$$V = MSE\left(\bar{y}_{FTAA}^{(d)}\right) \tag{7. 38}$$

To minimize the cost C for the specified variance $MSE\left(\bar{y}_{FTAA}^{(d)}\right) = V$ and to obtain the optimum values of n_1 and n_1, we defined a function φ_1 given as

$$\varphi_1 = (c_0 + c_1n_1 + c_2n_2) - \mu_1\left[V - MSE\left(\bar{y}_{FTAA}^{(d)}\right)\right] \tag{7. 39}$$

where μ_1 is the Lagrange multiplier.
Now, differentiating φ_1 with respect to n_1 and n_2 and equate to zero, we have

$$n_1 = \bar{Y}\sqrt{\frac{-\mu_1\left(E_1 + 2E_2\right)}{c_1}} \tag{7. 40}$$

$$n_2 = \bar{Y}\sqrt{\frac{\mu_1\left(C_y^2 + E_1 + 2E_2\right)}{c_2}} \tag{7. 41}$$

By substituting the values of n_1 and n_2 from (7.40) and (7.41) in (7.38), we have

$$\sqrt{\mu_1} = \bar{Y}E_3/\left(V + \bar{Y}^2C_y^2/N\right) \tag{7. 42}$$

Substituting (7.42) in (7.40) and (7.41), we have

$$n_1 = \bar{Y}^2\frac{E_3}{V + \bar{Y}^2C_y^2/N}\sqrt{\frac{-E_1 - 2E_2}{c_1}} \tag{7. 43}$$

$$n_2 = \bar{Y}^2\frac{E_3}{V + \bar{Y}^2C_y^2/N}\sqrt{\frac{C_y^2 + E_1 + 2E_2}{c_2}} \tag{7. 44}$$

The optimum n_1 and n_2 are obtained by putting the optimum value of P from (4.17) in (7.43) and (7.44) and we have

$$n_1^{opt} = \bar{Y}^2 C_y^2 \frac{\sqrt{c_2\left[1-\rho_{xy}^2\right]}+\sqrt{c_1\rho_{xy}^2}}{V+\bar{Y}^2C_y^2/N}\frac{\rho_{xy}}{\sqrt{c_1}} \tag{7. 45}$$

$$n_2^{opt} = \bar{Y}^2 C_y^2 \frac{\sqrt{c_2\left[1-\rho_{xy}^2\right]}+\sqrt{c_1\rho_{xy}^2}}{V+\bar{Y}^2C_y^2/N}\sqrt{\frac{\left[1-\rho_{xy}^2\right]}{c_2}} \tag{7. 46}$$

Hence, for the optimum values of n_1 and n_2 , the minimum total cost to be incurred on the use of $\bar{y}_{FTAA}^{(d)}$ for the specified precision having neglecting the term of order ${}^1\!/_N$ under case I is

$$C\left(\bar{y}_{FTAA}^{(d)}\right)_{\min} = c_0 + \bar{Y}^2 C_y^2\left\{\sqrt{c_2\left[1-\rho_{xy}^2\right]}+\sqrt{c_1\rho_{xy}^2}\right\}^2/V \tag{7. 47}$$

8. Empirical Study

In order to investigate the efficiency of the suggested estimator, these two real populations were considered

Data1: Source (Cochran [2])

$\bar{X} = 58.80,\ \bar{Y} = 101.1,\ C_x = 0.1281,\ C_y = 0.1445,\ \rho = 0.65,\ N = 20,$
$n = 8,\ n_1 = 12$

Data 2: Source (Das [3])

$\bar{X} = 25.11,\ \bar{Y} = 39.07,\ C_x = 1.6198,\ C_y = 1.4451,\ \rho = 0.72,\ N = 278,$
$n = 60,\ n_1 = 180$

Data 3: Source (Singh [7])

$$\bar{Y} = 555.43,\ \bar{X} = 878.16,\ C_x^2 = 1.5256,\ C_y^2 = 1.1086,\ \rho_{xy} = 0.8038$$

$$N = 50,\ C = \not{S}5000,\ c_0 = \not{S}2,000,\ c_1 = \not{S}50,\ c_2 = \not{S}500,\ V = 27765.22$$

TABLE 2. Mean Square Error of some Estimators under Case I using Data 1

Pre. sample	Sec. sample	Estimators				
n_1	n	Ratio estimator	Dual-to-ratio estimator	Malik and Tailor [4] estimator	Suggested estimator	Sample mean
3	2	83.00082	91.82996	82.84109	81.01121	96.03964
6	4	36.16488	38.43371	36.08501	35.17007	42.68428
9	6	20.55289	20.75449	20.49965	19.88969	24.89917
12	8	12.7469	12.28045	12.70697	12.2495	16.00661

TABLE 3. Mean Square Error of some Estimators under Case II using Data 1

Pre. sample	Sec. sample	Estimators				
n_1	n	Ratio estimator	Dual-to-ratio estimator	Malik and Tailor [4] estimator	Suggested estimator	Sample mean
3	2	108.3572	85.2602	106.8924	71.14027	96.03964
6	4	46.60575	33.70591	45.98851	31.29432	42.68428
9	6	26.02192	18.07799	25.68719	17.99548	24.89917
12	8	15.73001	11.78436	15.53654	11.32467	16.00661

TABLE 4. MSE of Shukla [9] and Suggested Factor-type estimators using Data 1

Values of d	Estimators			
	Shukla [9]		Suggested	
	Case I	Case II	Case I	Case II
1	3.8543	10.09715	3.8144	9.943618
5	5.5232	6.01545	5.5382	6.023696
10	3.3704	6.776089	3.3754	6.722319
15	3.4045	7.866532	3.3948	7.777846
20	3.4869	8.440236	3.4694	8.334295
30	3.5954	9.01004	3.5702	8.887422
40	3.6563	9.289849	3.6273	9.159174
50	3.6944	9.455469	3.6631	9.32006

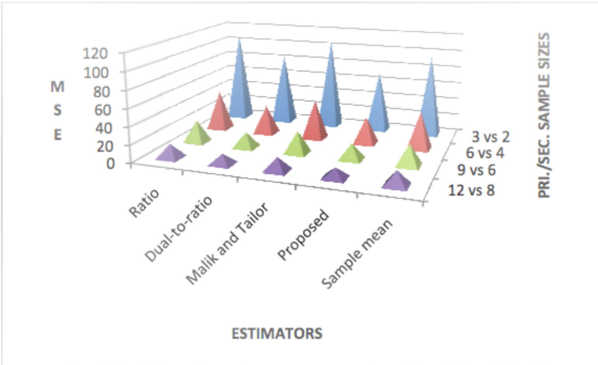

FIGURE 1. 3-Dimensional Pyramid for MSE in Table 2

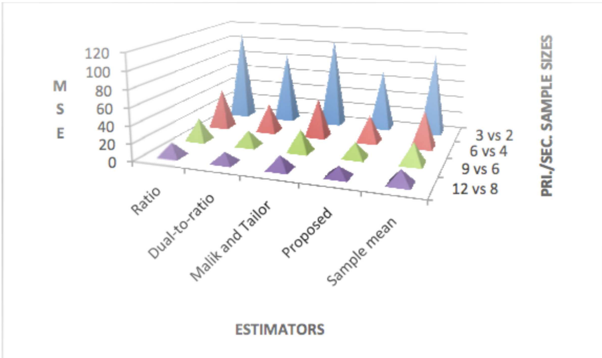

FIGURE 2. 3-Dimensional Pyramid for MSE in Table 3

TABLE 5. Mean Square Error of some Estimators under Case I using Data 2

Pre. sample	Sec. sample	Estimators				
n_1	n	Ratio estimator	Dual-to-ratio estimator	Malik and Tailor [4] estimator	Suggested estimator	Sample mean
45	15	62.11222	77.89603	61.18462	52.70719	83.0441
90	30	28.68793	33.97873	28.22413	23.98542	39.1539
135	45	17.5465	19.35707	17.2373	14.41149	24.5238
180	60	11.97579	12.10185	11.74389	9.62453	17.2088

From Tables 2, 3, 5 and 6 and figures 1, 2, 5, and 6 the results showed that the suggested estimator has the least MSE among all the conventional estimators considered in this study.

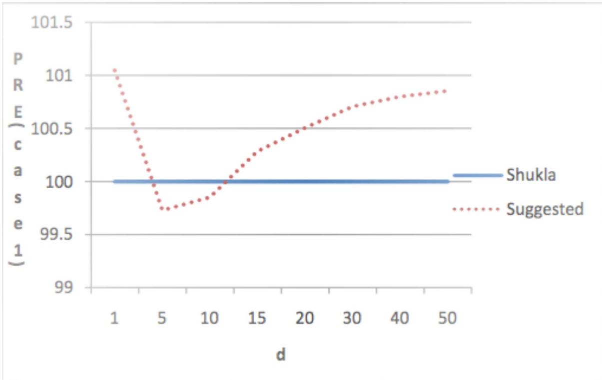

FIGURE 3. PRE of Suggested estimator over Shukla [9] estimator in table 4 for case I

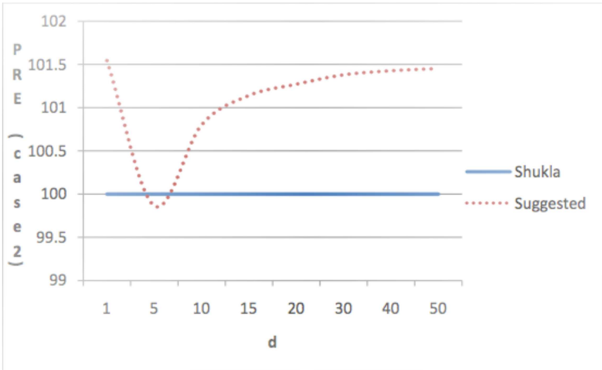

FIGURE 4. PRE of Suggested estimator over Shukla [9] estimator in table 4 for case II

This implies that the suggested estimator has higher accuracy than other traditional estimators in the study in estimating finite population mean. The results of Tables 4 and 7 and figures 3, 4, 7 and 8 revealed that the suggested estimator is more efficient than Shukla [9] estimator for all values of $d \in (1,\ 50)$ except at $d = 5$. From table 8, we observed that for fixed cost, the suggested estimator has smaller MSE and more efficient than other estimators. From table 9, we also observed that the expected costs incurred in using suggested estimator is less in comparison to the expected costs incurred for other estimators in the case of specified precision with exception of Shukla [9] estimator. Conclusively, information like coefficient of variation, kurtosis, skewness and correlation coefficient of auxiliary variable, if judiciously utilized, play important role in description of central values of the study variable.

TABLE 6. Mean Square Error of some Estimators under Case II using Data 2

Pre. sample	Sec. sample	Estimators				
n_1	n	Ratio estimator	Dual-to-ratio estimator	Malik and Tailor [4] estimator	Suggested estimator	Sample mean
45	15	84.15203	75.83887	81.73073	49.80879	83.04413
90	30	37.57953	32.41075	36.51312	22.95083	39.15388
135	45	22.05536	18.26333	21.44058	13.96978	24.5238
180	60	14.29328	11.44713	13.90432	9.450321	17.20876

TABLE 7. MSE of Shukla [9] and Proposed Factor-type estimators using Data 2

Values of d	Estimators							
	Shukla [9]				Suggested			
	Case I		Case II		Case I		Case II	
	MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE
1	11.976	100	14.293	100	11.744	101.976	13.904	102.798
5	13.802	100	13.290	100	13.855	99.617	13.349	99.558
10	9.650	100	9.814	100	9.636	100.145	9.744	100.7184
15	10.135	100	11.010	100	10.050	100.846	10.833	101.634
20	10.533	100	11.778	100	10.409	101.191	11.546	102.009
30	10.988	100	12.602	100	10.827	101.487	12.315	102.331
40	11.228	100	13.023	100	11.049	101.620	12.711	102.455
50	11.376	100	13.278	100	11.186	101.699	12.949	102.541

TABLE 8. Relative Efficiency with respect to sample mean for Fixed Cost C

Estimators	n_1^{opt}	n_2^{opt}	R.E (.)
Sample mean	NA	6	100.00 (50163.15)*
Ratio estimator	15	5	130.92 (38315.86)
Malik and Tailor estimator [4]	15	5	131.03(38281.52)
Shukla estimator [9]	16	4	145.76 (34413.96)
Suggested estimator	17	4	149.02 (33663.02)

9. AUTHORS' CONTRIBUTIONS

The suggested estimator was developed by A. Audu in collaboration with A. A. Adewara and likewise the derivation of the properties as well as numerical results.

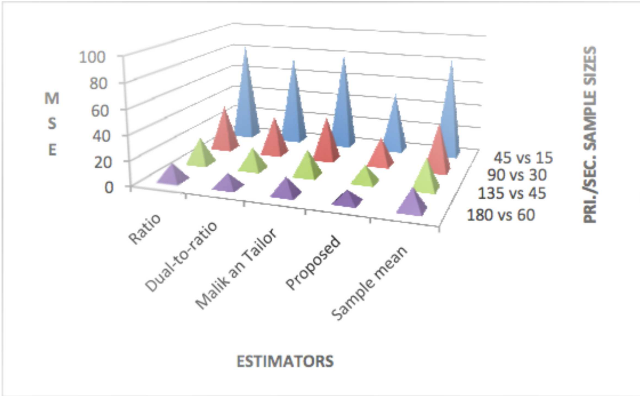

FIGURE 5. 3-Dimensional Pyramid for MSEs in Table 5

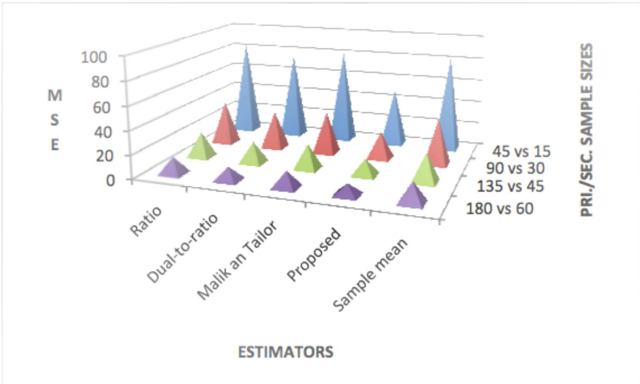

FIGURE 6. 3-Dimensional Pyramid for MSE in Table 6

TABLE 9. Expected cost of the Estimators for the specified Precision V

Estimators	n_1^{opt}	n_2^{opt}	Expected Cost $\not{S}$
Sample mean	NA	9	4,500
Ratio estimator	21	6	4,050
Malik and Tailor estimator [4]	21	6	4,050
Shukla estimator [9]	21	5	3,550
Suggested estimator	21	5	3,550

10. ACKNOWLEDGMENTS

I would like to thank my supervisor, Dr. A. A. Adewara, for the patience, guidance, encouragement and contribution he has provided while writing this paper.

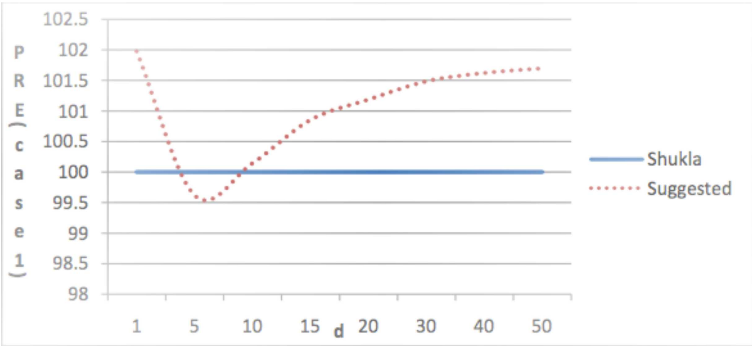

FIGURE 7. PRE of Suggested estimator over Shukla [9] estimator in table 7 for case I

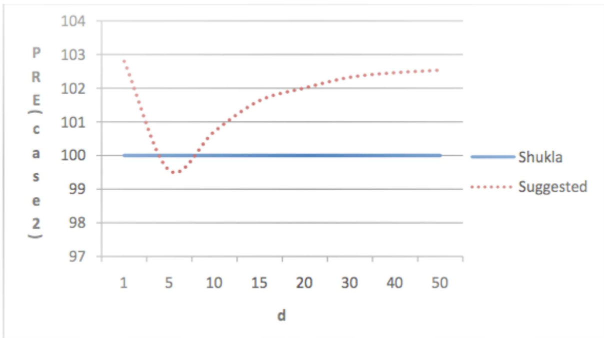

FIGURE 8. PRE of Suggested estimator over Shukla [9] estimator in table 7 for case II

REFERENCES

[1] S. Choudhury and B. K. Singh, *A class of chain ratio-product type estimators with two auxiliary variables under double sampling scheme*, Jour. Of the Korean Stat. Soc. **41**, (2012) 247-256.
[2] W. G. Cochran, *Sampling Techniques*, John Wiley and Son, New York, 1977.
[3] A. K. Das, *Contribution to the Theory of Sampling Strategies Based on Auxiliary Information*, Ph. D. thesis, BCKV, West Bengal, India, 1988.
[4] K. Malik and R. Tailor, *Ratio Type Estimator of Population Mean in Double Sampling*, International Journal of Advanced Mathematics and Statistics **1**, No. 1 (2013) 34-39.
[5] H. P. Singh and N. Agnihotri, *A general procedure of estimating population mean using auxiliary information in sample survey*, Statistics in Transition **9**, No. 1 (2008) 71-87.
[6] V. K. Singh and D. Shukla, *An efficient one-parameter family of factor-type estimator in sample surveys*, Metron International Journal of Statistics **55**,(1993) 139-159.
[7] S. Singh, *Advanced Sampling Theory with Application*, Springer Science+Business Media Dordrecht, 2003.
[8] H. P. Singh and R. Tailor, *Use of Known correlation coefficient in estimating the finite population mean*, Statistics in Transition **6**, No. 4 (2003) 555-560.
[9] D. Shukla, *F-T estimator under two phase sampling*, Metron. **59**, No. (1-2)(2002) 253-263.

[10] B. Y. Sukhatme, *Some ratio type estimators in two-phase sampling*, J. Amer. Statist. Assoc. **57**, (1962) 628-632.

[11] L. N. Upadhyaya and H. P. Singh, *Use of transformed auxiliary variable in estimating the finite population mean*, Biometrical journal **41**, No. 5 (1999) 627-636.