

# ON A NEW METHOD OF BIAS REDUCTION: ALTERNATIVE TO APPROXIMATELY UNBIASED RATIO ESTIMATORS

**F .B. ADEBOLA<sup>1</sup> and I.O. OSHUNGADE<sup>2</sup>**

<sup>1</sup> *Department Of Mathematical Sciences, Federal University Of Technology, Akure, Ondo State- Nigeria.*

*Emails:femi\_adebola@yahoo.Com*

*Phone:08038380448*

<sup>2</sup> *Department Of Statistics; University Of Ilorin, Ilorin, Kwara State- Nigeria.*

## **Abstract**

This paper proposes a new method of bias reduction from order  $n^{-1}$  to order  $n^{-2}$  resulting in a new approximately unbiased ratio estimator. The efficiency of this estimator for the ratio of population means of two characters is compared with the existing six other Ratio estimators under a linear regression model.

**Keywords:** *Approximately Unbiased, Mean Square error, Efficiency, Regression model, Ratio Estimator, Bias, Auxiliary variable.*

## **1.0 Introduction**

Over the years survey samplers [1 -7] among many others have been interested in methods of improving the precision of the estimates of population parameters both at the selection and estimation stages by making use of auxiliary information. Ratio estimators are often employed by these samplers to estimating the population mean of the characteristic of interest of the population ratio.

Let  $y$  and  $x$  be real variates taking  $y_i$  and  $x_i$  ( $1 \leq i \leq N$ ) for  $i^{th}$  unit of a population of Size  $N$  with means  $\mu_y$  and  $\mu_x$  respectively. Suppose that a simple random sample of size  $n$  units is drawn without replacement from the population. A commonly employed estimator in this context is traditional or classical ratio estimator  $r = \frac{\bar{y}}{\bar{x}}$  where  $\bar{y}$  and  $\bar{x}$  denote the sample means of  $y$  and  $x$  values, respectively.

## **2.0 Approximately Unbiased Ratio Type Estimators**

Generally, the classical ratio estimator( $r$ ) is biased; therefore, many authors [3-5, 8] and others have attempted to reduce this biasness in a situation where freedom from bias is important. Many authors [3-5], had also, worked on adjusting the classical ratio estimator by the term that is capable of reducing it from order  $n^{-1}$  to order  $n^{-2}$  to attain small or moderate gain in efficiency. Several ratio type estimators which satisfy this criterion are called approximately or almost unbiased ratio type estimators.

The following approximately unbiased ratio estimators in the literature [2 - 5] and others are considered for comparison.

Pascual [1] came up with the approximately unbiased ratio estimator:

$$\bar{Y}_p = \bar{r} + \frac{(N-1)(\bar{y} - \bar{r}\bar{x})}{(n-1)N} \quad (2.1)$$

This estimator has been shown to be efficient as that of combined bias ratio estimator in stratified sampling. Beale [2] proposed another approximately unbiased estimator of order  $O(n^{-2})$  as

$$\hat{R}_B = \hat{R}(1 + \theta \frac{S_{xy}}{\bar{x}\bar{y}}) / (1 + \theta \frac{S_x^2}{\bar{x}^2}) = \hat{R}(1 + \theta C_{xy}) / (1 + \theta C_x^2) \quad (2.2)$$

Tin [3] derived another approximately unbiased ratio estimator closely related to that of Beale [2] which was called

Modified ratio estimator, defined as

$$\hat{R}_T = \hat{R} \left[ 1 + \theta \left( \frac{S_{xy}}{\bar{x}\bar{y}} - \frac{S_x^2}{\bar{x}^2} \right) \right] = \hat{R}(1 + \theta(C_{xy} - C_x^2)) \quad (2.3)$$

, where

$$S_{xy} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}), \quad (2.4)$$

$$S_x^2 / \bar{x}^2 = C_x^2 \quad (2.5)$$

$$S_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (2.6)$$

$$S_{xy} / (\bar{y} \bar{x}) = C_{xy} \quad (2.7)$$

$$\text{And } \theta = \frac{1}{n} - \frac{1}{N} \quad (2.8)$$

$\hat{R}_B$  and  $\hat{R}_T$  have the same variance to order  $O(n^{-2})$ .

Sahoo [4] proposed another approximately unbiased ratio estimator termed almost unbiased ratio estimator, defined as

$$\hat{R}_S = \hat{R} / [1 + \theta (C_x^2 - C_{xy})] \quad (2.9)$$

It is said to be more efficient than  $\hat{R}_B$  and  $\hat{R}_T$ .

Sahoo [5] further derived a class of almost unbiased ratio estimators, among which we have the following as its members:

$$\hat{R}_{s1} = \hat{R} (1 + \theta C_{xy}) (1 - \theta C_x^2) \quad (2.10)$$

$$\hat{R}_{s2} = \hat{R} \left( \frac{1 - \theta C_x^2}{1 - \theta C_{xy}} \right) \quad (2.11)$$

$$\hat{R}_{s3} = \frac{\hat{R}}{[(1 - \theta C_{xy})(1 + \theta C_x^2)]} \quad (2.12)$$

### 3.0 THE PROPOSED METHOD OF BIAS REDUCTION: ALTERNATIVE TO APPROXIMATELY UNBIASED RATIO (AAUR) ESTIMATORS

The proposal of this estimator was based on the following standard regularity conditions:

Let  $v = (C_{xy}, C_x^2)$  assume values in a bounded, closed convex subset, S, of two dimensional real spaces containing the point  $V = (C_{xy}, C_x^2)$ .

Let  $f(v)$  be a function of  $v$  (which in particular may be a polynomial in  $(C_{xy}, C_x^2)$ ) satisfying the following conditions:

- (a) The function  $f(v)$  is continuous and bounded in S
- (b) The first and second order partial derivatives of  $f(v)$  exist and are continuous and bounded in S.

(c) After expansion under the given conditions, we get

$$f(v) = 1 + (C_{xy} - C_x^2) + O(v^{-2})$$

Then, we have the following theorem:

Theorem: If  $f(v)$  is differentiable in  $C_{xy}$  and  $C_x^2$  and fulfils the above regularity conditions then  $\hat{R}^* = \hat{R}f(v)$  is an asymptotically unbiased ratio estimator.

The unbiased ratio estimator proposed is  $\hat{R}_A = \hat{R} [1 - \theta (C_{xy} + C_x^2)]$  (3.1)

where  $S_{xy}$ ,  $S_x^2/\bar{x}^2$ ,  $S_y^2/\bar{y}^2$  and  $\theta$  are as defined in equations (2.4) to (2.8)

Conventionally,  $R_n = \frac{\bar{y}_n}{\bar{x}_n}$  is taken as a biased estimate of  $R_N = \frac{\bar{y}_N}{\bar{x}_N}$ , since both  $\bar{y}_n$  and  $\bar{x}_n$  are unbiased

estimates of  $\bar{y}_N$  and  $\bar{x}_N$  respectively.

Proof:

$$\text{Let } R_n = \frac{\bar{y}_n}{\bar{x}_n} \quad (3.2)$$

$$\Rightarrow \bar{y}_n = R_n \bar{x}_n \quad (3.3)$$

$$R_N = \frac{\bar{y}_N}{\bar{x}_N} = \frac{E(\bar{y}_n)}{E(\bar{x}_n)} \quad (3.4)$$

Substituting (3.3) in (3.4), we have

$$R_N = \frac{E(R_n \bar{x}_n)}{E(\bar{x}_n)} \quad (3.5)$$

Now

$$\text{Bias in } R_n = E(R_n) - R_N \quad (3.6)$$

$$B(R_n) = E(R_n) - \frac{E(R_n \bar{x}_n)}{E(\bar{x}_n)} \quad (3.7)$$

That is

$$B(R_n) = \{E(R_n)E(\bar{x}_n) - R_n \bar{x}_n\} / E(\bar{x}_n) \quad (3.8)$$

since the negative covariance between  $R_n$  and  $\bar{x}_n = E(R_n)E(\bar{x}_n) - R_n \bar{x}_n$  and  $E(\bar{x}_n) = \bar{x}_N$ , then

$$B(R_n) = - \frac{\text{Cov}(R_n, \bar{x}_n)}{\bar{x}_N} \quad (3.9)$$

Obtaining an upper bound to (3.8), we have

$$|\text{Bias in } R_n| \leq \frac{\sigma_{R_n} \sigma_{\bar{x}_n}}{\bar{x}_N} = \sigma_{R_n} \sqrt{\frac{N-n}{Nn}} C_x \quad (3.10)$$

Where  $C_x = \frac{S_x}{\bar{x}_N}$  is the coefficient of variation of  $x$  and  $\sigma_{R_n}$  and  $\sigma_{\bar{x}_n}$  are the standard errors of  $R_n$  and  $\bar{x}_n$

respectively.

From (3.10) above, we can see that if  $n$  is sufficiently large, the bias in the ratio estimate  $R_n$  is negligible as compared to its standard deviation.

$$\text{Let } y_i = \bar{y}_N + \Delta_i \quad (3.11)$$

So that

$$\bar{y}_n = \bar{y}_N + \bar{\Delta}_n \quad (3.12)$$

Where

$$E(\bar{\Delta}_n) = 0 \text{ and } E(\bar{\Delta}_n^2) = (1-f) \frac{S_y^2}{n} \quad (3.13)$$

Similarly,

$$\text{Let } x_i = \bar{x}_N + \bar{\Delta}_i^* \quad (3.14)$$

$$\text{So that } \bar{x}_n = \bar{x}_N + \bar{\Delta}_n^* \quad (3.15)$$

where

$$E(\bar{\Delta}_n^*) = 0 \text{ and } E(\bar{\Delta}_n^{*2}) = (1-f) \frac{S_x^2}{n} \quad (3.16)$$

To obtain the expected value of  $R_n$ , it is convenient to express it in terms of  $\Delta_n$  and  $\Delta_n^*$  we have

$$R_n = \frac{\overline{y_N} \left(1 + \frac{\overline{\Delta_n}}{x_N}\right)}{\overline{x_N} \left(1 + \frac{\overline{\Delta_n^*}}{x_N}\right)} \quad (3.17)$$

It is assumed that  $\overline{x_n} \neq 0$ ,  $\overline{x_N} \neq 0$ ,  $\overline{y_N} \neq 0$  and  $\left| \frac{\overline{\Delta_n}}{x_N} \right| < 1$  so that we may expand  $\left(1 + \frac{\overline{\Delta_n^*}}{x_N}\right)^{-1}$

Using Taylor's series expansion in powers of  $\overline{\Delta_n^*}$ , and expanding and multiplying out, we have

$$R_n = R_N \left\{ 1 + \frac{\overline{\Delta_n}}{y_N} - \frac{\overline{\Delta_n^*}}{x_N} + \frac{\overline{\Delta_n^{*2}}}{x_N^2} - \frac{\overline{\Delta_n}}{y_N} \frac{\overline{\Delta_n^*}}{x_N} + \frac{\overline{\Delta_n}}{y_N} \frac{\overline{\Delta_n^{*2}}}{x_N^2} - \frac{\overline{\Delta_n^{*3}}}{x_N^3} + \frac{\overline{\Delta_n^{*4}}}{x_N^4} - \frac{\overline{\Delta_n}}{y_N} \frac{\overline{\Delta_n^{*3}}}{x_N^3} + \dots \right\} \quad (3.18)$$

Now taking the expectation of (3.18) term by term we obtain,

$$E(R_n) = R_N + R_N E \left\{ \frac{\overline{\Delta_n^{*2}}}{x_N^2} - \frac{\overline{\Delta_n}}{y_N} \frac{\overline{\Delta_n^*}}{x_N} + \frac{\overline{\Delta_n}}{y_N} \frac{\overline{\Delta_n^{*2}}}{x_N^2} - \frac{\overline{\Delta_n^{*3}}}{x_N^3} + \frac{\overline{\Delta_n^{*4}}}{x_N^4} - \frac{\overline{\Delta_n}}{y_N} \frac{\overline{\Delta_n^{*3}}}{x_N^3} + \dots \right\} \dots \quad (3.19)$$

neglecting terms in  $\overline{\Delta_n}$  and  $\overline{\Delta_n^*}$  higher than the second term i.e.  $\overline{\Delta_n} \overline{\Delta_n^{*2}}$ ,  $\overline{\Delta_n^{*3}}$ ,  $\overline{\Delta_n} \overline{\Delta_n^{*3}}$

This approximation gives

$$E(R_n) = R_N + R_N \left[ \frac{E(\overline{\Delta_n^{*2}})}{x_N^2} - E\left(\frac{\overline{\Delta_n} \overline{\Delta_n^*}}{y_N x_N}\right) \right] \quad (3.20)$$

$$= R_N \left\{ 1 + \frac{N-n}{Nn} \left( \frac{S_x^2}{x_N^2} - \rho \frac{S_y}{y_N} \frac{S_x}{x_N} \right) \right\}$$

$$= R_N \left\{ 1 + \frac{N-n}{Nn} (C_x^2 - \rho C_x C_y) \right\} \quad (3.21)$$

$$\text{Where } C_x = \frac{S_x}{x_N} \text{ and } C_y = \frac{S_y}{y_N}$$

$$\therefore E[R_n] = R_N [1 + \theta(C_x^2 - C_{xy})] \quad (3.22)$$

Now, following the approach of Tin [3] in equation (2.3), by adjusting equation (3.22) for bias and by subtracting an estimate of the bias from  $E[R_n]$ ,

we obtain the estimator:

$$\hat{R}_A = E[R_n] - 2R_N \theta C_x^2 \quad (3.23)$$

$$\hat{R}_A = \hat{R} [1 + \theta(C_x^2 - C_{xy})] - 2R_N \theta C_x^2$$

$$\therefore \hat{R}_A = \hat{R} [1 - \theta(C_{XY} + C_x^2)] \quad (3.24)$$

#### 4.0 DERIVATIONS OF THE MEAN SQUARE ERROR FOR THE ALTERNATIVE UNBIASED RATIO TYPE ESTIMATOR UNDER THE REGRESSION MODEL:

$$y_i = \beta x_i + e_i, \quad i = 1, 2, \dots, N. \quad (4.1)$$

where  $\beta$  is unknown real constant and  $e_i$ 's are random variables with the following conditional expectations:

$$E(e_i/x_i) = 0 \quad (4.2)$$

$$E(e_i^2/x_i) = \delta x_i^t, \quad (4.3)$$

With

$$\delta > 0,$$

$$0 \leq t \leq 2. \quad (4.4)$$

and  $E(e_i e_j / x_i x_j) = 0, i \neq j.$

It is also assumed that  $x$  has gamma distribution with parameter  $k$  as often encountered in real life survey situations.

Now, assuming (4.1), we have

$$\mu_x = k; \quad \mu_y = \beta k; \quad R = \beta; \quad \bar{y} = \beta \bar{x} + \bar{e} \quad \text{and} \quad S_{xy} = \beta s_x^2 + s_{ex} \quad (4.5)$$

Where

$$\bar{e} = \frac{\sum e_i}{n}; \text{ and } s_{ex} = \frac{\sum e_i(x_i - \bar{x})}{n-1} \quad (4.6)$$

Let  $E(. / x)$  and  $E_x(.)$  denote respectively the expectation operators for given  $x$  and with respect to the distribution of  $x$ , so that  $E(.) = E_x(E(. / x))$  denoting the mean square error for  $R_i$  as  $MSE(R_i)$  under this model, we have

$$MSE(R_i) = E(R_i - \beta)^2 = E_x[E(R_i - \beta)^2 / X_i] \quad i=1,2, \dots, 8. \quad (4.7)$$

Let  $Z_i = x_i^t$  so that we have

$$E(Z) = Z = \frac{\Gamma(k+t)}{\Gamma(k)} \quad (4.8)$$

$$E(e_2 / x) = \frac{\delta}{n} \bar{Z} \quad (4.9)$$

$$E(\bar{e} S_{ex} / x) = \frac{\delta}{n} S_{zx} \quad (4.10)$$

and

$$E(S_{ex}^2 / x) = \frac{\delta}{n} q. \quad (4.11)$$

Where

$$\bar{Z} = \frac{\sum Z_i}{n}; S_{zx} = \frac{\sum Z_i(x_i - \bar{x})}{n-1} \text{ and } q = \frac{n \sum Z_i(x_i - \bar{x})^2}{(n-1)^2}. \quad (4.12)$$

Considering (1.13) and the regression model

$$\bar{y} = \beta \bar{x} + \bar{e} \quad (4.13)$$

We have

$$\beta = \frac{(\bar{y} - \bar{e})}{\bar{x}} \quad (4.14)$$

and

$$S_{xy} = \beta s_x^2 + s_{ex}$$



$$S_{xy} = \left[ \frac{(\bar{y} - \bar{e})}{x} \right] s_x^2 + s_{ex} \quad (4.15)$$

Now, substituting directly into the regression model and expanding under the assumptions and approximations

Tin [3], we obtain,

$$\text{MSE}(R_i) = E(R_i - \beta)^2 = E_X [E(R_i - \beta)^2 / X] \quad (4.16)$$

For  $R_i = \hat{R}_A$ , we have,

$$\begin{aligned} (\hat{R}_A - \beta)^2 &= \left\{ R(1 - \theta(C_{xy} - C_X^2)) - \beta \right\}^2 = \left\{ \frac{\bar{y}}{x} \left( 1 - \theta \frac{S_{xy}}{xy} - \theta \frac{S_X^2}{X^2} \right) - \left[ \frac{\bar{y}}{x} - \frac{\bar{e}}{x} \right] \right\}^2 \\ &= \left\{ \frac{\bar{y}}{x} - \theta \frac{S_{xy}}{x^2} - \theta \bar{y} \frac{S_X^2}{x^3} - \frac{\bar{y}}{x} + \frac{\bar{e}}{x} \right\}^2 = \left\{ \theta \frac{S_{xy}}{x^2} - \theta \bar{y} \frac{S_X^2}{x^3} + \frac{\bar{e}}{x} \right\}^2 \end{aligned} \quad (4.17)$$

Substituting for  $S_{xy}$ , we obtain,

$$\begin{aligned} (\hat{R}_A - \beta)^2 &= \left\{ \frac{\bar{e}}{x} - \frac{\theta}{x^2} \left[ \frac{\bar{y}}{x} - \frac{\bar{e}}{x} \right] s_x^2 + s_{ex} \right\} - \theta \bar{y} \frac{S_X^2}{x^3} \right\}^2 \\ &= \left\{ \frac{\bar{e}}{x} - \theta \frac{\bar{y}}{x^3} s_x^2 + \theta \frac{\bar{e}}{x^3} s_x^2 - \frac{\theta}{x^2} s_{ex} - \theta \bar{y} \frac{S_X^2}{x^3} \right\}^2 = \left\{ \frac{\bar{e}}{x} + \theta \frac{\bar{e}}{x^3} s_x^2 - \frac{\theta}{x^2} s_{ex} - 2\theta \bar{y} \frac{S_X^2}{x^3} \right\}^2 \\ &= \left\{ \frac{\bar{e}}{x} \left[ 1 + \theta \frac{s_x^2}{x^2} \right] - \frac{\theta}{x^2} [s_{ex} + 2\bar{y} \frac{S_X^2}{x}] \right\}^2 \end{aligned} \quad (4.18)$$

Recall as N tends to infinity,  $\theta$  becomes  $\frac{1}{n}$ , then

$$\begin{aligned} (\hat{R}_A - \beta)^2 &= \left\{ \frac{\bar{e}}{x} \left[ 1 + \frac{s_x^2}{nx^2} \right] - \frac{1}{nx^2} [s_{ex} + 2\bar{y} \frac{S_X^2}{x}] \right\}^2 = \left\{ \frac{\bar{e}}{x} + \frac{\bar{e}}{nx^3} s_x^2 - \frac{1}{nx^2} s_{ex} - 2\bar{y} \frac{S_X^2}{nx^3} \right\}^2 \\ &= \frac{\bar{e}^2}{x^2} + \frac{\bar{e}^2 s_x^2}{nx^4} - \frac{\bar{e} s_{ex}}{nx^3} + \frac{\bar{e}^2 s_x^2}{nx^4} + \frac{\bar{e}^2 (s_x^2)^2}{n^2 x^6} - \frac{\bar{e} s_x^2 s_{ex}}{n^2 x^5} - \frac{\bar{e} s_{ex}}{nx^3} - \frac{\bar{e} s_x^2 s_{ex}}{n^2 x^5} + \frac{s_{ex}^2}{n^2 x^4} + \text{terms involving } y\text{'s.} \\ &= \frac{\bar{e}^2}{x^2} + 2 \frac{\bar{e}^2 s_x^2}{nx^4} - 2 \frac{\bar{e} s_{ex}}{nx^3} + \frac{\bar{e}^2 (s_x^2)^2}{n^2 x^6} - 2 \frac{\bar{e} s_x^2 s_{ex}}{n^2 x^5} + \frac{s_{ex}^2}{n^2 x^4} + \text{terms involving } y\text{'s.} \end{aligned}$$

$$= \frac{\bar{e}^2}{x^2} - \frac{2}{n} \left[ e^{-\frac{s_{zx}}{x^3}} - e^{-\frac{s_x^2}{x^4}} \right] + \frac{1}{n^2} \left[ \frac{s_{zx}^2}{x^4} - 2e^{-\frac{s_{zx}}{x^3}} \frac{s_{zx}}{x^5} + e^{-\frac{s_x^2}{x^4}} \frac{(s_x^2)^2}{x^6} \right] + \text{terms involving } y\text{'s} \quad (4.19)$$

Substituting the (4.9), (4.10) and (4.11) in (4.19) we obtain,

$$E [(\hat{R}_A - \beta)^2 / X] = \frac{\delta \bar{z}}{nx^2} - \frac{2\delta}{n^2} \left[ \frac{s_{zx}}{x^3} - \bar{z} \frac{s_x^2}{x^4} \right] + \frac{\delta}{n^3} \left[ \frac{q}{x^4} - \frac{2s_{zx}}{x^5} s_x^2 + \bar{z} \frac{(s_x^2)^2}{x^6} \right] \quad (4.20)$$

Applying the well known Taylor's series expansion method, we have,

$$E_x \left( \frac{\bar{z}}{x^2} \right) = \frac{Z}{k^2} \left( 1 + \frac{3}{nk} - \frac{2t}{nk} + \frac{7}{n^2 k^2} - \frac{9t}{n^2 k^2} + \frac{3t^2}{n^2 k^2} \right) + o(n^{-3}) \quad (4.21)$$

$$E_x \left( \frac{s_{zx}}{x^3} - \bar{z} \frac{s_x^2}{x^4} \right) = \frac{Z}{k^2} \left( -\frac{1}{k} + \frac{t}{k} - \frac{2}{nk^2} + \frac{6t}{nk^2} - \frac{4t^2}{nk^2} \right) + o(n^{-2}) \quad (4.22)$$

$$E_x \left( \frac{q}{x^4} \right) = \frac{Z}{k^4} (k + t + t^2) + o(n^{-1}) \quad (4.23)$$

$$E_x \left( \bar{z} \frac{(s_x^2)^2}{x^6} \right) = \frac{Z}{k^4} + o(n^{-1}) \quad (4.24)$$

and

$$E_x \left( \frac{s_{zx}}{x^5} s_x^2 \right) = \frac{Z}{k^4} t + o(n^{-1}) \quad (4.25)$$

Now, the Mean Squared Error is given by:

$$\begin{aligned} E [(\hat{R}_A - \beta)^2 / X] &= \frac{\delta \bar{z}}{nx^2} - \frac{2\delta}{n^2} \left[ \frac{s_{zx}}{x^3} - \bar{z} \frac{s_x^2}{x^4} \right] + \frac{\delta}{n^3} \left[ \frac{q}{x^4} - \frac{2s_{zx}}{x^5} s_x^2 + \bar{z} \frac{(s_x^2)^2}{x^6} \right] \\ &= \frac{\delta Z}{nk^2} \left( 1 + \frac{3}{nk} - \frac{2t}{nk} + \frac{7}{n^2 k^2} - \frac{9t}{n^2 k^2} + \frac{3t^2}{n^2 k^2} \right) - \frac{2\delta Z}{n^2 k^2} \left( -\frac{1}{k} + \frac{t}{k} - \frac{2}{nk^2} + \frac{6t}{nk^2} - \frac{4t^2}{nk^2} \right) \\ &\quad + \frac{\delta Z}{n^3 k^4} (k + t + t^2) \frac{2\delta Z t}{n^3 k^4} + \frac{\delta Z}{n^3 k^4} = \frac{\delta Z}{n^3 k^4} [n^2 k^2 + 5nk - 4nkt + 12 - 22t + 12t^2 + k] \end{aligned} \quad (4.26)$$

Thus,

$$\text{MSE} (\hat{R}_A) = \frac{\delta Z}{n^3 k^4} \{ nk(nk + 1) + 4nk(1 - t) + 12(1 + t^2) - 22t + k \} \quad (4.27)$$

Similarly, the mean square errors of all the estimators are shown below:

$$\text{MSE} (\hat{R}_B) = \frac{\delta Z}{n^3 k^4} \{nk(nk + 1) + k + 4(1 + t^2) + 2\} \quad (4.28)$$

$$\text{MSE} (\hat{R}_T) = \frac{\delta Z}{n^3 k^4} \{nk(nk + 1) + k + 4(1 + t^2) + 2t\} \quad (4.29)$$

$$\text{MSE} (\hat{R}_S) = \frac{\delta Z}{n^3 k^4} \{nk(nk + 1) + k + 4(1 + t^2) + 2t\} \quad (4.30)$$

$$\text{MSE} (\hat{R}_{S1}) = \frac{\delta Z}{n^3 k^4} \{nk(nk + 1) + k + 4(1 + t^2)\} \quad (4.31)$$

$$\text{MSE} (\hat{R}_{S2}) = \frac{\delta Z}{n^3 k^4} \{nk(nk + 1) + k + 4(1 + t^2) + 2(2t - 1)\} \quad (4.32)$$

$$\text{MSE} (\hat{R}_{S3}) = \frac{\delta Z}{n^3 k^4} \{nk(nk + 1) + k + 4(1 + t^2) + 2(2t + 1)\} \quad (4.33)$$

## 5.0 CONDITION UNDER WHICH $\hat{R}_A$ IS BETTER THAN THE EXISTING ESTIMATORS

We intend to establish the conditions under which our proposed estimator is better than the existing estimators [2-5]. In comparison of two estimators  $e_1$  and  $e_2$ , when  $V(e_1) < V(e_2)$  or  $\text{MSE}(e_1) < \text{MSE}(e_2)$ , then,  $e_1$  is better than  $e_2$ . Hence, in order to establish the conditions when  $\hat{R}_A$  is better than other existing estimators [1-5], we compared the mean square errors of our alternative estimator with the existing ones.

Past researchers [9-11] have shown that  $\hat{R}_{S1}$  is better than others; therefore we compared the mean square errors of  $\hat{R}_A$  and  $\hat{R}_{S1}$ .

In our comparison it was discovered that  $\text{MSE}(\hat{R}_A) < \text{MSE}(\hat{R}_{S1})$  whenever,  $1 < t \leq 2.0$  and  $nk > 8$

## 6.0 EMPIRICAL INVESTIGATIONS

In this section, an empirical study is carried out using a Monte Carlo technique to compare the performance of our alternative unbiased ratio estimator along side with the classical one. We shall be considering the usual model for ratio estimator  $y_i = \beta x_i + e_i$ ,  $i = 1, 2, 3, \dots, N$ .

This work shall be viewed from the angle of real life situation which is always encountered in sampling practice by assuming that the auxiliary variable  $x$  follows a gamma (i.e. skewed population).

The approximately unbiased ratio estimators shall be compared under the following assumed model:

the regression of  $y$  on  $x$  is linear i.e,  $y = \alpha + \beta x + e$  with  $x$  having a gamma distribution with parameter  $(2, 1)$ , that is,  $x_i \sim G(2, 1)$  and  $e$  having a gamma distribution with parameter  $(0.25, 1)$ , that is,  $e_i \sim G(0.25x_i, 1.0)$ .

Under varying values of intercept  $(\alpha) = 0, 0.5, 1.2$ .

We considered the simulation of Variance, Bias, mean square error (MSE) and efficiency of the proposed alternative approximately unbiased ratio estimator  $(\hat{R}_A)$  along side with that of  $(\hat{R}_{s1})$  which has been established to be the best in previous studies [9-11]

We shall consider the values  $t = 1, 1.5, 2.0$  For  $n = 20, 40, 100, 200$  and  $k = 2$  under the regression model.

(i) For  $t = 1$ , we have,

$$y_i = 0.25x_i + e_i; e_i \sim G(0.25x_i, 1.0); x_i \sim G(2, 1).$$

(ii) For  $t = 1.5$ , we have,

$$y_i = 0.25x_i + e_i \sqrt{x_i}; e_i \sim G(0.25 \sqrt{x_i}, 1.0); x_i \sim G(2, 1).$$

(iii) For  $t = 2.0$ , we have,

$$y_i = 0.25x_i + e_i x_i; e_i \sim G(0.25, 1.0); x_i \sim G(2, 1).$$

**Table 4.1: Simulations for  $t = 1$**

EST	N	Variance	Bias	MSE
R <sub>S1</sub>	20	0.02229141	0.4957136	0.26802338225
R <sub>A</sub>	20	0.02035064	0.4512071	0.2239384870904
R <sub>S1</sub>	40	0.01163586	0.4911259	0.25284050965081
R <sub>A</sub>	40	0.1121222	0.4682516	0.33138176090256
R <sub>S1</sub>	100	0.00477786	0.5061191	0.2609344033848
R <sub>A</sub>	100	0.00471523	0.4985879	0.25330512402641
R <sub>S1</sub>	200	0.001758906	0.5378979	0.2910930568244

R <sub>A</sub>	200	0.001743158	0.534885	0.287845121225
----------------	-----	-------------	----------	----------------

**Table 4.2: Simulations for t = 1.5**

EST	N	Variance	Bias	MSE
R <sub>S1</sub>	20	0.05887116	0.5420939	0.35273695641721
R <sub>A</sub>	20	0.04969045	0.4892565	0.289070690225
R <sub>S1</sub>	40	0.01163586	0.4911259	0.25284050965081
R <sub>A</sub>	40	0.01121222	0.4682516	0.2304717809026
R <sub>S1</sub>	100	0.00477786	0.5061191	0.26093432938481
R <sub>A</sub>	100	0.00471523	0.4985879	0.2533051240264
R <sub>S1</sub>	200	0.003271526	0.80595569	0.6528361002433761
R <sub>A</sub>	200	0.003247954	0.8021401	0.646676694028

**Table 4.3: Simulations for t = 2.0**

EST	N	Variance	Bias	MSE
R <sub>S1</sub>	20	0.01470531	0.3578919	0.14279192208561
R <sub>A</sub>	20	0.0151919	0.3350565	0.1274547581923
R <sub>S1</sub>	40	0.00531964	0.3857655	0.15413466099025
R <sub>A</sub>	40	0.005348987	0.3753843	0.1462623596865
R <sub>S1</sub>	100	0.002499927	0.3738429	0.14225844088041
R <sub>A</sub>	100	0.002493612	0.3699303	0.1393420388581
R <sub>S1</sub>	200	0.0006446923	0.364413	0.133441526869
R <sub>A</sub>	200	0.0006427695	0.3626546	0.1321611284012

**CONCLUSION:** The simulations results in this study have confirmed that  $\hat{R}_A$  estimator is better than the existing ones [6-8] in terms of the bias and MSE whenever  $1 \leq t \leq 2.0$  and  $n_k > 8$ .

## References

- [1] Pascual, J.N. (1961): Unbiased ratio estimators in stratified sampling. Journal of the American Statistical Association, 56(293):70-87.
- [2] Beale, E.M.L.(1962): Some uses of computers in operational research. Industrielle organization , 31 , 51 – 52.
- [3] Tin, M (1965): comparison of some ratio estimators. Journal of the American Statistical Association, 60:294-307.
- [4] Sahoo, L. N (1983): On a method of bias reduction in Ratio estimation. Journal of Statistical Research. Vol.17,1-6.
- [5] Sahoo, L N (1987): On a class of almost unbiased estimators for population ratio. Statistics, vol. 1, P.119-121.
- [6] Adewara A, A (2006): On efficiency of some sampling ratio estimators. Journal of the Nigerian Statistical Association. Vol. 18, p.65-70.
- [7] Oshungade I. O. (1986): Non response and Ratio estimation Problems in Sample survey.Unpublished Ph D thesis, University of Essex, United Kingdom.
- [8] Cochran W. G. (1977): Sampling Techniques. New York. John Wiley.
- [9] Dalabehera ,M and Sahoo,L.N.(1994):Comparison of Six almost unbiased ratio estimators.Questio,vol18,3 Pp.369-375.
- [10] Dalabehera ,M and Sahoo,L.N.(1995):Efficiencies of Six almost unbiased ratio estimators under a particular model.Statistical Papers,36,61-67.
- [11] Sahoo, L. N, Das B.C and Sahoo, J (2006): A Note on the performance of several almost unbiased ratio estimators. Pakistan Journal of Statistics. Vol. 22 p, 209 – 215.