

A class of ratio type estimators in double sampling using an auxiliary variable with some known population parameters

Toluwalase J. Akingbade and Fabian C. Okafor¹.

Department of Statistics, University of Nigeria, Nsukka, Nigeria

Many authors, to improve on the ratio (regression) estimator of the population mean of the study variable, have made use of some known population parameters of the auxiliary variable including the population mean of the auxiliary variable. In their work these authors have used the ratio (regression) – type estimator. In practical situations, the population mean of the auxiliary variable may not be known. In this paper therefore, we have suggested a class of regression - type ratio estimator when only the population mean of the auxiliary variable is unknown using double sampling procedure. The expression for the bias and the mean square error of the proposed estimators were derived and sub-members of the class of estimators were also identified. The conditions for which the proposed estimators perform better than the sample mean, classical ratio and existing ratio- type estimator in double sampling are derived. From the analysis, it was observed that the proposed estimators perform better than the existing estimators considered in this study.

Keywords: Bias, Mean Square error, ratio estimator, double sampling, Percent relative efficiency.

1. Introduction

Proper use of auxiliary variable is always known to improve the performance of estimators. Ratio, product and regression sampling estimators are the most common and widely discussed in sampling theory literature. Cochran (1940, 1942) proposed the classical ratio and regression estimator and Murthy (1964) proposed the classical product estimator under simple random sampling with assumed known population mean. When the correlation between the study variable (Y) and the auxiliary variable (X) is highly positive and the regression line of Y on X is linear and passes through the origin, ratio estimator is appropriate but when the correlation is highly negative, product method of estimation is quite effective in this case. In most practical situations, the regression line does not pass through the origin and this makes the ratio estimator having limitation of not performing equally well as regression estimator.

In order to improve on the performance of ratio estimator of the population mean, a lot of authors have suggested the use of what is usually called ratio-type estimator or regression type ratio estimator. Among the authors are: Sisodia, and Dwivedi (1981), Upadhyaya and Singh (1999), Yan and Tian (2010), Singh and Tailor (2003) and Subramani and Kumarapandiyam (2013). The general form of the estimators given by the above authors is:

¹ Corresponding author. Email: fabian.okafor@gmail.com

$$\hat{T} = \bar{y} \left[\frac{A\bar{X} + G}{A\bar{x} + G} \right] \quad (1)$$

A and G are known population parameters like population correlation coefficient, ρ , population skewness, β_1 , kurtosis, β_2 and coefficient of variation of the auxiliary variable C_x and deciles D_k etc. Other authors like Kadilar and Cingi (2004, 2006), Yan and Tian (2010) and Raja *et al.* (2017) gave the regression type ratio estimator having the form:

$$\hat{R} = \frac{\bar{y} - b_{xy}(\bar{x} - \bar{X})}{(A\bar{x} + G)} (A\bar{X} + G), \quad (2)$$

where b_{xy} is the sample regression coefficient of y on x .

All these authors assumed that the population mean of X is known; but this is not always the case. When the population means of the auxiliary variable is unknown, we adopt the procedure of double sampling scheme. In double sampling scheme, a large initial sample of size n' ($n' < N$) is drawn from the population of size N using simple random sample without replacement (SRSWOR) from where we measure x . In the second phase, we draw a sample subsample of size n from first phase sample of size n' , i.e. ($n < n'$) by using (SRSWOR), and observe the study variable y . With the information from the first phase sample an estimator \bar{x}' for \bar{X} is obtained.

2. Double Sampling Ratio-Type Estimator of Population Mean

Malik and Tailor (2013), by setting $A=1, G=\rho$, and $\bar{x}' = \bar{X}$ in (1), obtained the double sampling estimator of \bar{Y} as

$$\hat{T}_{MK} = \bar{y} \left[\frac{\bar{x}' + \rho}{\bar{x} + \rho} \right] \quad (3)$$

Following Kadilar and Cingi (2004, 2006), Raja *et al.* (2017), we proposed a class of regression type ratio estimator when the population mean of the auxiliary variable (X) is unknown given by

$$\hat{R}_d = \frac{\bar{y} - t(\bar{x}' - \bar{x}'')}{(A\bar{x} + G)^\alpha} (A\bar{x}' + G)^\alpha, \quad (4)$$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i, \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \text{ and } \bar{x}' = \frac{1}{n'} \sum_{i=1}^{n'} x_i, \text{ are the sample means.}$$

A and G are the either real values or function of any known population parameters of an auxiliary variable such as skewness (β_1), kurtosis (β_2), coefficient of variation (C_x), deciles (D_1, D_2, \dots, D_{10}) of an auxiliary variable, correlation coefficient (ρ), γ is real value. These can be obtained from past surveys

and therefore considered to be constants. Let t be a known constant. The scalar α takes value zero or -1 (for product-type estimator) or + 1 (for ratio-type estimator).

In this work, we are only interested in the ratio estimator when the variable of interest and auxiliary variable are positively correlated, i.e. $\alpha = 1$

2.1 Bias and mean square error of \hat{R}_d

To obtain the Bias and MSE of \hat{R}_d , up to the first order of approximation, let

$$\bar{y} = \bar{Y}(1 + \Delta_y), \bar{x} = \bar{X}(1 + \Delta_x), \bar{x}' = \bar{X}(1 + \Delta'_x), \bar{x}'^\gamma = [\bar{X}(1 + \Delta'_x)]^\gamma,$$

Such that $E(\Delta_x) = E(\Delta_y) = E(\Delta'_x) = 0$ and

$$E(\Delta_y^2) = \omega_1 C_y^2, E(\Delta_x^2) = \omega_1 C_x^2, E(\Delta_x'^2) = E(\Delta_x \Delta_x') = \omega_2 C_x^2, E(\Delta_x \Delta_y) = \omega_1 \rho_{xy} C_x C_y$$

$$E(\Delta_x' \Delta_y) = \omega_2 \rho_{xy} C_x C_y;$$

where $C_y^2 = \frac{S_y^2}{\bar{Y}^2}$ and $C_x^2 = \frac{S_x^2}{\bar{X}^2}$ are the square of the population coefficient of variation of their

respective subscripts. $\omega_1 = \frac{1}{n} - \frac{1}{N}$ and $\omega_2 = \frac{1}{n'} - \frac{1}{N}$

The proposed estimator \hat{R}_d can be written in terms of Δ as

$$\hat{R}_d = \frac{\left[\bar{Y}(1 + \Delta_y) - t\gamma \bar{X}^\gamma \Delta_x - \frac{t\gamma(\gamma - 1)\bar{X}^\gamma \Delta_x^2}{2} + t\gamma \bar{X}^\gamma \Delta'_x + \frac{t\gamma(\gamma - 1)\bar{X}^\gamma \Delta_x'^2}{2} \right] [1 + \lambda \Delta_x']^\alpha}{[1 + \lambda \Delta_x]^\alpha}, \tag{5}$$

where $\lambda = \frac{A\bar{X}}{A\bar{X} + G}$

Expanding (5) and ignoring order higher than two, we have

$$\begin{aligned} \hat{R}_d = & \bar{Y} + \bar{Y}\Delta_y - t\gamma \bar{X}^\gamma - \frac{t\gamma(\gamma - 1)\bar{X}^\gamma \Delta_x^2}{2} + t\gamma \bar{X}^\gamma \Delta'_x + \frac{t\gamma(\gamma - 1)\bar{X}^\gamma \Delta_x'^2}{2} + \alpha \lambda \bar{Y} \Delta_x' \\ & + \alpha \lambda \bar{Y} \Delta_x' \Delta_y - t\alpha \lambda \gamma \bar{X}^\gamma \Delta_x' \Delta_x - t\alpha \lambda \gamma \bar{X}^\gamma \Delta_x'^2 + \frac{\alpha(\alpha - 1)\lambda^2 \bar{Y} \Delta_x'^2}{2} - \alpha \lambda \bar{Y} \Delta_x \\ & - \alpha \lambda \Delta_x \Delta_y + t\alpha \lambda \gamma \bar{X}^\gamma \Delta_x^2 - t\alpha \lambda \gamma \bar{X}^\gamma \Delta_x \Delta_x' - \alpha^2 \lambda^2 \Delta_x \Delta_x' + \frac{\alpha(\alpha + 1)\lambda^2 \bar{Y} \Delta_x'^2}{2}. \end{aligned} \tag{6}$$

From (6) the bias of \hat{R}_d is

$$\begin{aligned}
 B(\hat{R}_d) &= E(\hat{R}_d - \bar{Y}) \\
 &= -\frac{\omega_1 t \gamma (\gamma - 1) \bar{X}^\gamma C_x^2}{2} + \frac{\omega_2 t \gamma (\gamma - 1) \bar{X}^\gamma C_x^2}{2} + (\omega_2 - \omega_1) \alpha \lambda \bar{Y} \rho_{xy} C_x C_y + \frac{\omega_2 \alpha (\alpha - 1) \lambda^2 \bar{Y} C_x^2}{2} \\
 &\quad + (\omega_1 - \omega_2) t \alpha \lambda \gamma \bar{X}^\gamma C_x^2 - \omega_2 \alpha^2 \lambda^2 \bar{Y} C_x^2 + \frac{\omega_1 \alpha (\alpha + 1) \lambda^2 \bar{Y} C_x^2}{2} \tag{7}
 \end{aligned}$$

$$\begin{aligned}
 &= \omega_3 \left\{ t \alpha \lambda \gamma \bar{X}^\gamma C_x^2 - \frac{t \gamma (\gamma - 1) \bar{X}^\gamma C_x^2}{2} - \alpha \lambda \bar{Y} \rho_{xy} C_x C_y \right\} - \omega_2 \alpha^2 \lambda^2 \bar{Y} C_x^2 \\
 &\quad + \frac{\alpha \lambda^2 \bar{Y} C_x^2 \{ \omega_2 (\alpha - 1) + \omega_1 (\alpha + 1) \}}{2} \\
 &= \omega_3 \left\{ t \alpha \lambda \gamma \bar{X}^\gamma C_x^2 - \frac{t \gamma (\gamma - 1) \bar{X}^\gamma C_x^2}{2} - \alpha \lambda \bar{Y} \rho_{xy} C_x C_y + \frac{\lambda^2 \bar{Y} C_x^2 (\alpha^2 + \alpha)}{2} \right\} \tag{8}
 \end{aligned}$$

where $\omega_3 = \omega_1 - \omega_2 = \left(\frac{1}{n} - \frac{1}{n'} \right)$.

The mean square error is given, from (6) by

$$\text{MSE}(\hat{R}_d) = E(\hat{R}_d - \bar{Y})^2 =$$

$$\begin{aligned}
 &E[\bar{Y}^2 \Delta_y^2 + t^2 \gamma^2 \bar{X}^{2\gamma} \Delta_x^2 + t^2 \gamma^2 \bar{X}^{2\gamma} \Delta_x'^2 + \alpha^2 \lambda^2 \bar{Y}^2 \Delta_x'^2 + \alpha^2 \lambda^2 \gamma^2 \bar{Y}^2 \Delta_x^2 - 2t \gamma \bar{X}^\gamma \bar{Y} \Delta_x \Delta_y + 2t \gamma \bar{X}^\gamma \bar{Y} \Delta_x' \Delta_y + 2\alpha \lambda \bar{Y}^2 \Delta_x' \Delta_y - \\
 &2\alpha \lambda \bar{Y}^2 \Delta_x \Delta_y - 2t^2 \gamma^2 \bar{X}^{2\gamma} \Delta_x \Delta_x' - 2t \alpha \lambda \gamma \bar{X}^\gamma \bar{Y} \Delta_x \Delta_x' + 2t \alpha \lambda \gamma \bar{X}^\gamma \Delta_x^2 + 2t \alpha \lambda \gamma \bar{X}^\gamma \bar{Y} \Delta_x'^2 - 2t \alpha \lambda \gamma \bar{X}^\gamma \Delta_x \Delta_x' - 2\alpha^2 \lambda^2 \bar{Y}^2 \Delta_x \Delta_x']. \tag{9}
 \end{aligned}$$

Simplifying (9), we obtain the mean square error as

$$\begin{aligned}
 \text{MSE}(\hat{R}_d) &= \omega_1 \bar{Y}^2 C_y^2 + \omega_3 [t^2 \gamma^2 \bar{X}^{2\gamma} C_x^2 + \alpha^2 \lambda^2 \bar{Y}^2 C_x^2 - 2(\alpha \lambda \bar{Y}^2 \rho_{xy} C_x C_y + t \gamma \bar{X}^\gamma \bar{Y} \rho_{xy} C_x C_y \\
 &\quad - t \alpha \lambda \gamma \bar{X}^\gamma \bar{Y} C_x^2)]. \tag{10}
 \end{aligned}$$

Let $v_1 = \gamma \bar{X}^\gamma$, and substituting this in (10), we have

$$\begin{aligned}
 \text{MSE}(\hat{R}_d) &= \omega_1 \bar{Y}^2 C_y^2 + \omega_3 [t^2 v_1^2 C_x^2 + \alpha^2 \lambda^2 \bar{Y}^2 C_x^2 - 2(\alpha \lambda \bar{Y}^2 \rho_{xy} C_x C_y + t v_1 \bar{Y} \rho_{xy} C_x C_y \\
 &\quad - t \alpha \lambda v_1 \bar{Y} C_x^2)] \tag{11}
 \end{aligned}$$

To obtain the optimum t we differentiate (11) with respect to t and equate it to zero, we have

$$t_o = (\bar{Y}\rho_{xy}C_y - \alpha\lambda\bar{Y}C_x)/v_1C_x. \quad (12)$$

Substituting t_o in (11) we obtain the minimum MSE of \hat{R}_d as

$$MSE(\hat{R}_d)_{\min} = \bar{Y}^2 C_y^2 (\omega_1 - \omega_3 \rho_{xy}^2) \quad (13)$$

This is equivalent to the approximate MSE of the classical linear regression estimator in double sampling. This implies that, the proposed class of estimator at t optimum is equally efficient as classical linear regression estimator in double sampling. If we substitute the optimum value of t given in (12) in (8) the bias optimum becomes

$$B_{t_o}(\hat{R}_d) = \frac{\omega_3 \bar{Y}}{2} \{ \alpha(1-\alpha)\lambda^2 C_x^2 + (\gamma-1)(\alpha\lambda C_x^2 - \rho_{xy} C_x C_y) \} \quad (14)$$

Members of the Proposed Estimator

By setting A , G and t to certain parameters given in Table 1, and $\alpha = 1$ and γ ($0 < \gamma \leq 1$) in (4) we have the Class 1 ratio estimators in double sampling in Table 1 and Class 2 regression type ratio estimators in Table 2 below. When $t=0$, $\alpha = 1$ in (4), we have the generalized form of (3), which is under class 1 as

$$\hat{T}_d = \bar{y} \frac{A\bar{x}' + G}{A\bar{x} + G} \quad (15)$$

Its MSE is obtained by similar substitution in (11), i.e.

$$MSE(\hat{T}_d) = \omega_1 \bar{Y}^2 C_y^2 + \omega_3 [\lambda^2 \bar{Y}^2 C_x^2 - 2\lambda \bar{Y}^2 \rho_{xy} C_x C_y] \quad (16)$$

The bias from (8) is given by

$$Bias(\hat{T}_d) = \omega_3 \left\{ -\lambda \bar{Y} \rho_{xy} C_x C_y + \lambda^2 \bar{Y} C_x^2 \right\} \quad (17)$$

Appropriate substitution of A and G in (15) gives some members of class 1 ratio estimator in doubling sampling presented in Table 1. While for the class 2 estimators, we set $t = b_{xy}$ in (4). This gives us the members of regression type ratio estimators in double sampling presented in Table 2

Table 1: Members of class 1 ratio type estimator in double sampling

S/No	Estimators	t	α	A	G
1	$\hat{T}_{d1} = \bar{y}(\bar{x}' + C_x)/(\bar{x} + C_x)$	0	1	1	C_x
2	$\hat{T}_{d2} = \bar{y}(\bar{x}' + \beta_2)/(\bar{x} + \beta_2)$	0	1	1	β_2
3	$\hat{T}_{d3} = \bar{y}(\beta_2\bar{x}' + C_x)/(\beta_2\bar{x} + C_x)$	0	1	β_2	C_x
4	$\hat{T}_{d4} = \bar{y}(C_x\bar{x}' + \beta_2)/(C_x\bar{x} + \beta_2)$	0	1	C_x	β_2
5	$\hat{T}_{d5} = \bar{y}(\bar{x}' + \beta_1)/(\bar{x} + \beta_1)$	0	1	1	β_1
6	$\hat{T}_{d6} = \bar{y}(\beta_1\bar{x}' + \beta_2)/(\beta_1\bar{x} + \beta_2)$	0	1	β_1	β_2
7	$\hat{T}_{d7} = \bar{y}(\beta_2\bar{x}' + \beta_1)/(\beta_2\bar{x} + \beta_1)$	0	1	β_2	β_1
8	$\hat{T}_{d8} = \bar{y}(C_x\bar{x}' + \beta_1)/(C_x\bar{x} + \beta_1)$	0	1	C_x	β_1
9	$\hat{T}_{d9} = \bar{y}(\bar{x}' + \rho_{xy})/(\bar{x} + \rho_{xy})$	0	1	1	ρ_{xy}
10	$\hat{T}_{d10} = \bar{y}(\bar{x}' + D_1)/(\bar{x} + D_1)$	0	1	1	D_1
11	$\hat{T}_{d11} = \bar{y}(\bar{x}' + D_2)/(\bar{x} + D_2)$	0	1	1	D_2
12	$\hat{T}_{d12} = \bar{y}(\bar{x}' + D_3)/(\bar{x} + D_3)$	0	1	1	D_3
13	$\hat{T}_{d13} = \bar{y}(\bar{x}' + D_4)/(\bar{x} + D_4)$	0	1	1	D_4
14	$\hat{T}_{d14} = \bar{y}(\bar{x}' + D_5)/(\bar{x} + D_5)$	0	1	1	D_5
15	$\hat{T}_{d15} = \bar{y}(\bar{x}' + D_6)/(\bar{x} + D_6)$	0	1	1	D_6
16	$\hat{T}_{d16} = \bar{y}(\bar{x}' + D_7)/(\bar{x} + D_7)$	0	1	1	D_7
17	$\hat{T}_{d17} = \bar{y}(\bar{x}' + D_8)/(\bar{x} + D_8)$	0	1	1	D_8
18	$\hat{T}_{d18} = \bar{y}(\bar{x}' + D_9)/(\bar{x} + D_9)$	0	1	1	D_9
19	$\hat{T}_{d19} = \bar{y}(\bar{x}' + D_{10})/(\bar{x} + D_{10})$	0	1	1	D_{10}

Table 2: Members of class 2 regression type ratio estimators in double sampling

S/No	Estimators with $0 < \gamma \leq 1$	A	G
1	$\hat{R}_{d1} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)]\bar{x}'/\bar{x}$	1	0
2	$\hat{R}_{d2} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](\bar{x}' + C_x)/(\bar{x} + C_x)$	1	C_x
3	$\hat{R}_{d3} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](\bar{x}' + \beta_2)/(\bar{x} + \beta_2)$	1	β_2
4	$\hat{R}_{d4} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](\beta_2\bar{x}' + C_x)/(\beta_2\bar{x} + C_x)$	β_2	C_x
5	$\hat{R}_{d5} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](C_x\bar{x}' + \beta_2)/(C_x\bar{x} + \beta_2)$	C_x	β_2
6	$\hat{R}_{d6} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](\bar{x}' + \rho_{xy})/(\bar{x} + \rho_{xy})$	1	ρ_{xy}
7	$\hat{R}_{d7} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](C_x\bar{x}' + \rho_{xy})/(C_x\bar{x} + \rho_{xy})$	C_x	ρ_{xy}
8	$\hat{R}_{d8} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](\rho_{xy}\bar{x}' + C_x)/(\rho_{xy}\bar{x} + C_x)$	ρ_{xy}	C_x
9	$\hat{R}_{d9} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](\beta_2\bar{x}' + \rho_{xy})/(\beta_2\bar{x} + \rho_{xy})$	β_2	ρ_{xy}
10	$\hat{R}_{d10} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](\rho_{xy}\bar{x}' + \beta_2)/(\rho_{xy}\bar{x} + \beta_2)$	ρ_{xy}	β_2
11	$\hat{R}_{d11} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](\bar{x}' + \beta_1)/(\bar{x} + \beta_1)$	1	β_1
12	$\hat{R}_{d12} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](\beta_1\bar{x}' + \beta_2)/(\beta_1\bar{x} + \beta_2)$	β_1	β_2
13	$\hat{R}_{d13} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](D_1\bar{x}' + C_x)/(D_1\bar{x} + C_x)$	D_1	C_x
14	$\hat{R}_{d14} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](D_2\bar{x}' + C_x)/(D_2\bar{x} + C_x)$	D_2	C_x
15	$\hat{R}_{d15} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](D_3\bar{x}' + C_x)/(D_3\bar{x} + C_x)$	D_3	C_x
16	$\hat{R}_{d16} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](D_4\bar{x}' + C_x)/(D_4\bar{x} + C_x)$	D_4	C_x
17	$\hat{R}_{d17} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](D_5\bar{x}' + C_x)/(D_5\bar{x} + C_x)$	D_5	C_x
18	$\hat{R}_{d18} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](D_6\bar{x}' + C_x)/(D_6\bar{x} + C_x)$	D_6	C_x
19	$\hat{R}_{d19} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](D_7\bar{x}' + C_x)/(D_7\bar{x} + C_x)$	D_7	C_x
20	$\hat{R}_{d20} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](D_8\bar{x}' + C_x)/(D_8\bar{x} + C_x)$	D_8	C_x
21	$\hat{R}_{d21} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](D_9\bar{x}' + C_x)/(D_9\bar{x} + C_x)$	D_9	C_x
22	$\hat{R}_{d22} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](D_{10}\bar{x}' + C_x)/(D_{10}\bar{x} + C_x)$	D_{10}	C_x
23	$\hat{R}_{d23} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](C_x\bar{x}' + D_1)/(C_x\bar{x} + D_1)$	C_x	D_1
24	$\hat{R}_{d24} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](C_x\bar{x}' + D_2)/(C_x\bar{x} + D_2)$	C_x	D_2
25	$\hat{R}_{d25} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](C_x\bar{x}' + D_3)/(C_x\bar{x} + D_3)$	C_x	D_3
26	$\hat{R}_{d26} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](C_x\bar{x}' + D_4)/(C_x\bar{x} + D_4)$	C_x	D_4
27	$\hat{R}_{d27} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](C_x\bar{x}' + D_5)/(C_x\bar{x} + D_5)$	C_x	D_5
28	$\hat{R}_{d28} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](C_x\bar{x}' + D_6)/(C_x\bar{x} + D_6)$	C_x	D_6
29	$\hat{R}_{d29} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](C_x\bar{x}' + D_7)/(C_x\bar{x} + D_7)$	C_x	D_7
30	$\hat{R}_{d30} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](C_x\bar{x}' + D_8)/(C_x\bar{x} + D_8)$	C_x	D_8
31	$\hat{R}_{d31} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](C_x\bar{x}' + D_9)/(C_x\bar{x} + D_9)$	C_x	D_9
32	$\hat{R}_{d32} = [\bar{y} - b_{xy}(\bar{x}^\gamma - \bar{x}'^\gamma)](C_x\bar{x}' + D_{10})/(C_x\bar{x} + D_{10})$	C_x	D_{10}

3. Theoretical Comparison

Our interest is on the estimation of the population mean using auxiliary variable, therefore, we shall compare our proposed estimator with the sample mean, double sampling linear regression and double sampling classical ratio estimator. Comparison will also be done among the subclass members presented in Tables 1 and 2.

Sample mean

$$\hat{T}_0 = \frac{1}{n} \sum_{i=1}^n y_i \quad (18)$$

$$\text{Its variance is } V(\hat{T}_0) = \omega_1 \bar{Y}^2 C_y^2 \quad (19)$$

Double sampling linear regression

This is given by Cochran (1977, p339) as

$$\hat{T}_{dlr} = \bar{y} - b_{xy} (\bar{x} - \bar{x}') \quad (20)$$

This can also be obtained by setting $t = b_{xy}$ and $\alpha = 0$ in (4).

Its MSE is

$$MSE(\hat{T}_{dlr}) = \bar{Y}^2 C_y^2 [\omega_1 - \omega_3 \rho_{xy}^2] \quad (21)$$

Double sampling ratio:

$$\hat{T}_{dr} = \frac{\bar{y}}{\bar{x}} \bar{x}' \quad (22)$$

MSE as in Cochran (1977, p343) is

$$MSE(\hat{T}_{dr}) = \omega_1 \bar{Y}^2 C_y^2 + \omega_3 [\bar{Y}^2 C_x^2 - 2\bar{Y}^2 \rho_{xy} C_x C_y] \quad (23)$$

From (16) and (19) the double sampling ratio type estimator \hat{T}_d will perform better than \hat{T}_0 , in the sense of having smaller MSE if and only if $MSE(\hat{T}_d) \leq MSE(\hat{T}_0)$. That will happen if

$$\rho_{xy} \geq \lambda C_x / 2C_y \quad (24)$$

Considering (15) and (20), $MSE(\hat{T}_d) \leq MSE(\hat{T}_{dlr})$ if and only if

$$(\lambda C_x - \rho_{xy} C_y)^2 \leq 0 \quad (25)$$

This is a contradiction; therefore the proposed double sampling ratio estimator can never perform better than the classical double sampling regression estimator. They will perform equally well if

$$\rho_{xy} = \lambda C_x / C_y$$

Finally, using (16) and (23), it can be concluded that \hat{T}_d will be more precise than T_{dr} if

$$\rho_{xy} \geq (\lambda + 1)C_x/2C_y \quad (26)$$

We shall now compare the double sampling regression type ratio estimator \hat{R}_d , in (4), for $\alpha = 1$, with the other usual estimators as in above.

From (11) and (18) \hat{R}_d will perform better than \hat{T}_0 if and only if

$$\rho_{xy} \geq C_x(tv_1 + \lambda\bar{Y})/2\bar{Y}C_y \quad (27)$$

Using (11) and (20), \hat{R}_d will be more precise than \hat{T}_{dr} if and only if

$$(tv_1 + \lambda\bar{Y})^2 \left\{ 1 - C_y^2/C_x^2 \right\} \leq -\left\{ \rho\bar{Y} - (tv_1 + \lambda\bar{Y})C_y/C_x \right\}^2 \quad (28)$$

A simple and quick check is to see whether $C_x > C_y$; if so then \hat{R}_d cannot be more precise than \hat{T}_{dr} .

From (11) and (23) we deduce that the proposed double sampling regression type ratio estimator, \hat{R}_d will be more efficient than the double sampling ratio estimator \hat{T}_{dr} if

$$\rho \geq C_x(tv_1 + \lambda\bar{Y} + \bar{Y})/2\bar{Y}C_y \quad (29)$$

Furthermore, \hat{R}_d will be more precise than \hat{T}_d if

$$\rho_{xy} \geq C_x tv_1 (tv_1 + 2\lambda\bar{Y})/2\bar{Y}\{tv_1 + \lambda(\lambda - 1)\bar{Y}\} \quad (30)$$

4. Empirical Comparison

In this section, we considered a real life data set by Singh and Chaudary (1986) and details of the data set are shown below. Using the data we calculated the MSEs, biases and percentage relative efficiencies with respect to sample mean, \hat{T}_0 of the proposed class of estimators in Tables 1 and 2. The results of the computation are presented in Tables 3, 4, 5 and 6 below.

Y - Number of animals in the i^{th} unit; X - size of the i^{th} unit.

$$N = 34 \quad n' = 15 \quad n = 10 \quad \bar{Y} = 856.4117 \quad \bar{X} = 199.4412 \quad \rho = 0.4453 \quad S_y = 733.1407 \quad S_x = 150.2150 \quad C_y = 0.8561 \quad C_x = 0.7531 \quad \beta_2 = 1.0445 \quad \beta_1 = 1.1823$$

$$D_1 = 60.600 \quad D_2 = 83.0000 \quad D_3 = 102.7000 \quad D_4 = 111.2000 \quad D_5 = 142.5000$$

$$D_6 = 210.2000 \quad D_7 = 264.5000 \quad D_8 = 304.4000 \quad D_9 = 373.2000 \quad D_{10} = 634.0000$$

Table 3: The MSE and bias values of the estimators (Class 1) in Table 1 and their PRE with respect to the sample mean estimator.

Estimators	MSE	PRE	Bias
\hat{T}_0 (Sample mean)	37944.29	100	0
\hat{T}_{dlr} (double sampling linear regression)	34391.27	110.33	-
\hat{T}_{dr} (double sampling classical ratio)	37772.28	100.46	7.995
\hat{T}_{d1}	37720.96	100.59	7.904
\hat{T}_{d2}	37701.31	100.64	7.869
\hat{T}_{d3}	37723.13	100.59	7.908
\hat{T}_{d4}	37678.37	100.71	7.829
\hat{T}_{d5}	37692.06	100.67	7.853
\hat{T}_{d6}	37712.16	100.62	7.889
\hat{T}_{d7}	37695.44	100.66	7.859
\hat{T}_{d8}	37666.18	100.74	7.807
\hat{T}_{d9} (Malik and Tailor, 2013)	37741.84	100.54	7.941
\hat{T}_{d10}	35334.08	107.39	3.238
\hat{T}_{d11}	34945.52	108.58	2.286
\hat{T}_{d12}	34719.65	109.29	1.645
\hat{T}_{d13}	34647.09	109.52	1.412
\hat{T}_{d14}	34473.61	110.07	0.728
\hat{T}_{d15}	34396.45	110.31	0.152
\hat{T}_{d16}	34472.03	110.07	0.531
\hat{T}_{d17}	34560.15	109.79	0.707
\hat{T}_{d18}	34737.07	109.23	0.891
\hat{T}_{d19}	35379.05	107.25	1.034

Table 4: The MSE and PRE values of the sub- members of the proposed class of ratio estimator in Table 2 (Class 2)

\hat{R}_d	$\gamma = 0.1$		$\gamma = 0.3$		$\gamma = 0.5$		$\gamma = 0.8$		$\gamma = 1$	
	MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE	MSE	PRE
\hat{R}_{d1}	37778.2	100.4	37823.5	100.3	38022.1	99.8	39968.3	95.09	48255.1	78.63
\hat{R}_{d2}	37726.8	100.6	37771.8	100.5	37968.9	99.94	39902.4	95.15	48151	78.8
\hat{R}_{d3}	37707.2	100.6	37752.0	100.5	37948.6	99.99	39877.1	95.09	48111	78.87
\hat{R}_{d4}	37729.0	100.6	37774.0	100.5	37971.2	99.93	39905.2	95.22	48155.4	78.8
\hat{R}_{d5}	37684.2	100.7	37728.9	100.6	37924.8	100.1	39847.5	95.03	48064.2	78.94
\hat{R}_{d6}	37747.7	100.5	37792.9	100.4	37990.6	99.88	39929.2	95.06	48193.4	78.73
\hat{R}_{d7}	37737.8	100.5	37782.9	100.4	37980.3	99.91	39916.5	95.29	48173.2	78.77
\hat{R}_{d8}	37663.9	100.7	37708.5	100.6	37903.8	100.1	39821.4	95.02	48022.9	79.01
\hat{R}_{d9}	37749.0	100.5	37794.2	100.4	37991.9	99.87	39930.9	95.42	48196	78.73
\hat{R}_{d10}	37620.7	100.9	37665.0	100.7	37859.0	100.2	39765.8	95.18	47934.6	79.16
\hat{R}_{d11}	37697.9	100.7	37742.7	100.5	37939.0	100.0	39865.2	95.12	48092.2	78.9
\hat{R}_{d12}	37718.0	100.6	37763.0	100.5	37959.8	99.96	39891	94.94	48133.1	78.83
\hat{R}_{d13}	37777.3	100.4	37822.7	100.3	38021.2	99.8	39967.2	94.94	48253.4	78.64
\hat{R}_{d14}	37777.6	100.4	37822.9	100.3	38021.5	99.8	39967.5	94.94	48253.8	78.63
\hat{R}_{d15}	37777.7	100.4	37823	100.3	38021.6	99.8	39967.7	94.94	48254.1	78.63
\hat{R}_{d16}	37777.7	100.4	37823.1	100.3	38021.6	99.8	39967.7	94.94	48254.1	78.63
\hat{R}_{d17}	37777.8	100.4	37823.2	100.3	38021.7	99.8	39967.9	94.94	48254.4	78.63
\hat{R}_{d18}	37777.9	100.4	37823.3	100.3	38021.9	99.8	39968.0	94.94	48254.6	78.63
\hat{R}_{d19}	37778.0	100.4	37823.3	100.3	38021.9	99.8	39968.1	94.94	48254.7	78.63
\hat{R}_{d20}	37778.0	100.4	37823.4	100.3	38021.9	99.8	39968.1	94.94	48254.7	78.63
\hat{R}_{d21}	37778.0	100.4	37823.4	100.3	38022.0	99.8	39968.2	94.94	48254.8	78.63
\hat{R}_{d22}	37778.1	100.4	37823.4	100.3	38022.0	99.8	39968.2	105.2	48254.9	78.63
\hat{R}_{d23}	34984.0	108.5	35003.0	108.4	35088.5	108.1	36058.2	107	41429.3	91.59
\hat{R}_{d24}	34656.5	109.5	34669.3	109.4	34727.8	109.3	35465.1	108.1	40142.1	94.53
\hat{R}_{d25}	34499.0	110	34507.2	110	34546.0	109.8	35112.8	108.4	39280.9	96.6
\hat{R}_{d26}	34457.0	110.1	34463.4	110.1	34494.6	110	34995.8	109.4	38968.1	97.37
\hat{R}_{d27}	34392.0	110.3	34392.9	110.3	34399.8	110.3	34692.4	110.2	38041.4	99.74
\hat{R}_{d28}	34501.1	110	34493.1	110	34462.2	110.1	34427.3	110.3	36798.7	103.1
\hat{R}_{d29}	34677.1	109.4	34664.2	109.5	34611.7	109.6	*34391.5	110.3	36209.4	104.8
\hat{R}_{d30}	34817.7	109	34801.8	109	34736.9	109.2	34408.6	110.1	35904.2	105.7
\hat{R}_{d31}	35055.1	108.2	35035.2	108.3	34953.2	108.6	34477.4	109	35532.5	106.8
\hat{R}_{d32}	35760.5	106.1	35731.9	106.2	35612.3	106.5	34812.4	95.09	34899.5	108.7
$(\hat{T}_{dp})_{opt}$	MSE =34391.27		PRE = 110.33							

Table 5: The bias values of the sub- members of the proposed class of ratio estimator in Table 2 (class 2) when $t = b_{xy}$

\hat{R}_d	$\gamma = 0.1$	$\gamma = 0.3$	$\gamma = 0.5$	$\gamma = 0.8$	$\gamma = 1$
\hat{R}_{d1}	8.01	8.079	8.36	10.5	16.19
\hat{R}_{d2}	7.92	7.988	8.268	10.4	16.07
\hat{R}_{d3}	7.88	7.953	8.233	10.36	16.02
\hat{R}_{d4}	7.92	7.991	8.272	10.4	16.08
\hat{R}_{d5}	7.84	7.912	8.191	10.32	15.97
\hat{R}_{d6}	7.95	8.025	8.305	10.44	16.12
\hat{R}_{d7}	7.94	8.007	8.288	10.42	16.1
\hat{R}_{d8}	7.8	7.876	8.155	10.28	15.92
\hat{R}_{d9}	7.96	8.027	8.308	10.44	16.12
\hat{R}_{d10}	7.73	7.799	8.077	10.19	15.82
\hat{R}_{d11}	7.87	7.936	8.216	10.34	16
\hat{R}_{d12}	7.9	7.972	8.252	10.38	16.05
\hat{R}_{d13}	8.01	8.077	8.358	10.5	16.19
\hat{R}_{d14}	8.01	8.077	8.359	10.5	16.19
\hat{R}_{d15}	8.01	8.078	8.359	10.5	16.19
\hat{R}_{d16}	8.01	8.078	8.359	10.5	16.19
\hat{R}_{d17}	8.01	8.078	8.359	10.5	16.19
\hat{R}_{d18}	8.01	8.078	8.359	10.5	16.19
\hat{R}_{d19}	8.01	8.078	8.359	10.5	16.19
\hat{R}_{d20}	8.01	8.078	8.359	10.5	16.19
\hat{R}_{d21}	8.01	8.078	8.36	10.5	16.19
\hat{R}_{d22}	8.01	8.078	8.36	10.5	16.19
\hat{R}_{d23}	2.39	2.445	2.661	4.229	8.221
\hat{R}_{d24}	1.45	1.499	1.698	3.13	6.717
\hat{R}_{d25}	0.85	0.901	1.089	2.422	5.712
\hat{R}_{d26}	0.64	0.693	0.876	2.171	5.346
\hat{R}_{d27}	0.07	0.11	0.28	1.452	4.264
\hat{R}_{d28}	0.6	0.56	0.41	0.572	2.812
\hat{R}_{d29}	0.8	0.8	0.67	0.207	2.124
\hat{R}_{d30}	0.9	0.9	0.77	0.038	1.768
\hat{R}_{d31}	1	0.8	0.86	0.14	1.333
\hat{R}_{d32}	1	0.94	0.85	0.31	0.594

Table 6: The bias values of the sub- members of the proposed class of double sampling ratio estimator in Table 2 (class 2) at t optimum value

\hat{R}_d	$\gamma = 0.1$	$\gamma = 0.3$	$\gamma = 0.5$	$\gamma = 0.8$	$\gamma = 1$
\hat{R}_{d1}	3.6	2.8	2	0.8	0
\hat{R}_{d2}	3.57	2.78	1.98	0.79	0
\hat{R}_{d3}	3.56	2.77	1.98	0.79	0
\hat{R}_{d4}	3.57	2.78	1.98	0.79	0
\hat{R}_{d5}	3.55	2.76	1.97	0.79	0
\hat{R}_{d6}	3.58	2.79	1.99	0.8	0
\hat{R}_{d7}	3.58	2.78	1.99	0.79	0
\hat{R}_{d8}	3.54	2.75	1.96	0.79	0
\hat{R}_{d9}	3.58	2.79	1.99	0.8	0
\hat{R}_{d10}	3.51	2.73	1.95	0.78	0
\hat{R}_{d11}	3.55	2.76	1.97	0.79	0
\hat{R}_{d12}	3.57	2.77	1.98	0.79	0
\hat{R}_{d13}	3.6	2.8	2	0.8	0
\hat{R}_{d14}	3.6	2.8	2	0.8	0
\hat{R}_{d15}	3.6	2.8	2	0.8	0
\hat{R}_{d16}	3.6	2.8	2	0.8	0
\hat{R}_{d17}	3.6	2.8	2	0.8	0
\hat{R}_{d18}	3.6	2.8	2	0.8	0
\hat{R}_{d19}	3.6	2.8	2	0.8	0
\hat{R}_{d20}	3.6	2.8	2	0.8	0
\hat{R}_{d21}	3.6	2.8	2	0.8	0
\hat{R}_{d22}	3.6	2.8	2	0.8	0
\hat{R}_{d23}	1.5	1.17	0.84	0.33	0
\hat{R}_{d24}	1	0.78	0.56	0.22	0
\hat{R}_{d25}	0.64	0.5	0.36	0.14	0
\hat{R}_{d26}	0.5	0.39	0.28	0.11	0
\hat{R}_{d27}	0.05	0.04	0.03	0.01	0
\hat{R}_{d28}	0.652	0.507	0.362	0.145	0
\hat{R}_{d29}	1.049	0.816	0.583	0.233	0
\hat{R}_{d30}	1.281	0.996	0.712	0.285	0
\hat{R}_{d31}	1.597	1.242	0.887	0.355	0
\hat{R}_{d32}	2.293	1.783	1.274	0.509	0

5. Results and Discussions

From Table 3, all the estimators under class 1, \hat{T}_{dj} , $j = 1, \dots, 19$ including \hat{T}_{d9} by Malik and Tailor (2013) are more efficient and have less biased than the sample mean, \hat{T}_0 and classical double sampling ratio estimator, \hat{T}_{dr} . We observe that the estimators, \hat{T}_{dj} , $j = 10, \dots, 19$ that utilize the deciles of an auxiliary variable prove most efficient having smaller biases and MSEs. The estimator \hat{T}_{d15} that uses the 6th decile proves to be most efficient estimator because it has the highest gain in percent relative efficiency (110.31) and less bias (0.152). The MSE of this estimator also tends close to that of classical double sampling regression estimator, \hat{T}_{d1r} for the population data used. However, other estimators in this class do not prove better than classical double sampling regression estimator.

Note: The bias of \hat{T}_{d1r} is not given because it requires re-sampling technique to produce the estimate of its bias, which is outside the scope of this present research.

From Table 4, we observe that the class 2 estimators have minimal gain in efficiency for $\gamma \leq 0.3$; when $\gamma > 0.3$ the sample mean is preferable. For $\gamma \leq 0.3$, the double sampling regression type ratio estimator \hat{R}_d has almost the same gain in efficiency as \hat{T}_d . When $\gamma > 0.3$, \hat{R}_d is less efficient than \hat{T}_d . It also has higher bias than \hat{T}_d . We notice that estimators \hat{R}_{d1} to \hat{R}_{d12} are not better than \hat{T}_{dr} for all values of γ . The class of estimators \hat{R}_d is less efficient than \hat{T}_0 when $\gamma = 1$. However, estimators \hat{R}_{d23} to \hat{R}_{d32} have a gain of efficiency over \hat{T}_{dr} and \hat{T}_0 , and are also less biased than \hat{T}_{dr} when $\gamma \leq 0.8$. Estimators \hat{R}_{d28} to \hat{R}_{d32} are all preferable to double sampling ratio estimator \hat{T}_{dr} for all values of $\gamma \leq 1.0$; and have same gain in efficiency as \hat{T}_{d1r} $\gamma = 0.8$. The proposed estimator \hat{R}_{d27} at $\gamma = [0.1, 0.5]$ has the least MSE and least bias and equally efficient as the proposed class of regression type ratio estimator when t is optimum. It also has a MSE close to that of \hat{T}_{d1r} , the double sampling regression estimator.

From Table 5, the regression type ratio estimators \hat{R}_{d1} to \hat{R}_{d22} have higher biases than \hat{T}_{dr} for all values of γ , the bias is doubled when $\gamma = 1$. In contrast \hat{R}_{d24} to \hat{R}_{d32} have very low biases in comparison with \hat{T}_{dr} . \hat{R}_{d30} has the least bias (0.038) at $\gamma = 0.8$, followed by \hat{R}_{d27} (0.07) at $\gamma = 0.1$. In other words the estimators that use the deciles have the least bias. The bias increases as γ increases.

The results in Table 6 give the biases of the proposed regression type ratio estimators obtained using (14). We clearly see that these biases are lower than when $t = b_{xy}$. The biases gave zero when $\gamma = 1$ giving it advantage over classical ratio and regression estimators.

6. Conclusion

This study has proposed a class of ratio type estimators in double sampling using an auxiliary variable with known population parameters. The proposed class of regression type ratio estimator in double sampling at t optimum value is the most efficient estimator because it has the least MSE and highest PRE and is equally as efficient as the classical linear regression in double sampling. At this minimum MSE, the bias is zero when $\gamma = 1$ which is an additional advantage. This class of estimators is more preferred if the deciles used stand alone without being the coefficient of the sample mean of the auxiliary variable.

The other ratio type estimator \hat{T}_d is equally as good as the classical double sample ratio estimator. It is preferred to the classical ratio estimator if the deciles are used in the estimator with higher deciles having the least biases. It is therefore highly recommended if ratio estimation is to be used in the estimation of the population mean because of its smaller MSE and its substantial reduction in bias.

References

- Cochran, W. G. (1940). The estimation of the yields of the cereal experiments by sampling for the ratio of grain to total produce. *Jour. Agri. Sci.*, **59**, 1225 - 1226.
- Cochran, W. G. (1942). Sampling theory when the sampling units are unequal sizes. *J. Amer. Stat. Assoc.*, **37**, 199-212.
- Cochran W.G. (1977). *Sampling Techniques* (3rd edition), John Wiley Eastern Limited, New York.
- Kadilar, C. and Cingi, H. (2004). Ratio estimators in simple random sampling. *Applied Mathematics and Computation*, **151**(3), 893 - 902.
- Kadilar, C. and Cingi, H. (2006). An improvement in estimating the population mean by using the correlation coefficient. *Hacettepe Journal of Mathematics and Statistics*, **35**(1), 103 - 109.
- Malik, K. A. and Tailor, R. (2013). Ratio type estimator of population mean in double sampling. *International Journal of Advanced Mathematics and Statistics*, **1**(1), pp. 34 – 39.
- Murthy, M.N. (1964). Product method of estimation. *Sankhya A*, **26**, 66 - 74.
- Raja, T. A., Subair, M., Maqbool, S. and Hakak, A. (2017). Enhancing the Mean Ratio Estimator for Estimating Population Mean Using Conventional Parameters. *International Journal of Mathematics and Statistics Invention (IJSM)*, **5**(1), 58 - 61.
- Singh, D. and Chaudhary, F. S. (1986). *Theory and analysis of sample survey designs*. (1st ed.), New Age International Publisher, India.

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

Singh, H. P. and Tailor, R. (2003). Use of known correlation coefficient in estimating the finite population means. *Statistics in Transition*, 6(4), 555 - 560.

Subramani, J. & Kumarapandiyan, G. (2013). Estimation of Finite Population Mean Using Deciles of an Auxiliary Variable. *Statistics in Transition-new series*, 14(1), pp. 75–88.

Upadhyaya, L. N. and Singh, H. (1999). Use of transformed auxiliary variable in estimating the finite population mean. *Biometrical Journal*, 41(5), 627 - 636.

Yan, Z. and Tian, B. (2010). Ratio Method to the Mean Estimation Using Coefficient of Skewness of Auxiliary Variable. ICICA 2010, Part II, CCIS 106, pp. 103-110.