Scientific
Research
Publishing

# Estimation of Population Ratio in Post-Stratified Sampling Using Variable Transformation

**Aloy Chijioke Onyeka, Chinyeaka Hostensia Izunobi, Iheanyi Sylvester Iwueze**

Department of Statistics, Federal University of Technology, Owerri, Nigeria
Email: aloyonyeka@futo.edu.ng, chiyeaka2007@yahoo.com, isiwueze@yahoo.com

## Abstract

Extending the work carried out by [1], this paper proposes six combined-type estimators of population ratio of two variables in post-stratified sampling scheme, using variable transformation. Properties of the proposed estimators were obtained up to first order approximations, $o\left(n^{-1}\right)$, both for achieved sample configurations (conditional argument) and over repeated samples of fixed size $n$ (unconditional argument). Efficiency conditions were obtained. Under these conditions the proposed combined-type estimators would perform better than the associated customary combined-type estimator. Furthermore, optimum estimators among the proposed combined-type estimators were obtained both under the conditional and unconditional arguments. An empirical work confirmed the theoretical results.

## Keywords

## 1. Introduction

The use of information on auxiliary character to improve estimates of population parameters of the study variable is a common practice in sample survey, and sometimes, information on several variables is used to estimate or predict a characteristic of interest. The investigators often collect observations from more than one variable, including the variable of interest $y$ and some auxiliary variables $x$. The use of these variables (known as auxiliary information in sample survey design) often results in efficient estimate of population parameters (e.g.

mean, ratio, proportion, etc.) under some realistic conditions, especially when there is a strong correlation between the study variables and the auxiliary variables. Many authors have made contributions in this regard, including [2] and [3]. In this context, ratio, product and regression methods of estimation are good examples. Ratio and product-type estimators take advantage of the correlation between the auxiliary variable and the study variable, to improve the estimate of the characteristic of interest. For example, when information is available on the auxiliary variable that is highly positively correlated with the study variable, the ratio method of estimation proposed by [4] is a suitable estimator to estimate the population mean, and when the correlation is negative, the product method of estimation, as envisaged by [5] and [6], is appropriate. However, in some studies, the ratio of the population means (or totals) of the study and auxiliary variables might be of great significance, hence the need to estimate such ratios.

The customary estimator of the population ratio $\left(R = \bar{Y}/\bar{X}\right)$ of the population means of two variables, $y$ and $x$, under the simple random sampling scheme, is given as $\hat{R} = \bar{y}/\bar{x}$, which is the ratio of the sample means of the two variables ([2] and [7]). The estimator, $\hat{R} = \bar{y}/\bar{x}$, uses information on only two variables, namely the study variable $(y)$ and one auxiliary variable $(x)$. However, several authors, like [7] and [8], have contributed to the problem of estimating the population ratio of two means, often utilizing additional information on one or more auxiliary variables, say $z_i \left(i = 1, 2, \cdots\right)$. While it is possible to record increased efficiency by introducing such additional auxiliary variables, it is obvious that extra cost is involved in order to obtain information on such additional auxiliary variables. References [1] and [9] have argued that such extra cost could be avoided by using variable transformation of the already observed auxiliary variable, instead of introducing additional (new) auxiliary variables. However, the works carried out by [1] [9] were restricted to estimation of population ratio in simple random sampling scheme. The present study is necessitated by the need to extend to post-stratified sampling scheme, the works on ratio estimation carried out by [1] [9] under the simple random sampling scheme. This is in order to extend to other sampling schemes, the obvious advantage of reduced cost in the use of variable transformation instead of introducing additional (new) auxiliary variables when estimating population ratio of two population parameters.

## 2. The Proposed Combined-Type Estimators

Let $n$ units be drawn from a population of $N$ units using simple random sampling method and let the sampled units be allocated to their respective strata, where $n_h$ is the number of units that fall into stratum $h$ such that $\sum_{h=1}^{L} n_h = n$. Let $y_{hi}$ and $x_{hi}$ be the $i^{\text{th}}$ observation on the study and auxiliary variables, respectively. Consider the following variable transformation of the auxiliary variable, $x$, under post-stratified sampling scheme.

$$x_{hi}^* = \frac{N\bar{X} - nx_{hi}}{N - n}, \quad h = 1, 2, \cdots, L \text{ and } i = 1, 2, \cdots, N \tag{2.1}$$

An equivalent of the transformation (2.1), in simple random sampling scheme, has been used by authors like [1] [8]-[13]. The associated sample mean estimator of the transformed variable (2.1), in post-stratified sampling scheme, can be written as

$$\bar{x}_{ps}^* = \left(1 - \pi\right)\bar{X} - \pi\bar{x}_{ps}, \quad \text{where} \quad \pi = \frac{n}{N - n} \tag{2.2}$$

and $\bar{x}_{ps} = \sum_{h=1}^{L} \omega_h \bar{x}_h$ and $\bar{y}_{ps} = \sum_{h=1}^{L} \omega_h \bar{y}_h$ are sample mean estimators based on $x_{hi}$ and $y_{hi}$ respectively. Using the sample means $\bar{y}_{ps}$, $\bar{x}_{ps}$ and $\bar{x}_{ps}^*$, and assuming that the population mean, $\bar{X}$ of the auxiliary variable $x_{hi}$, is known, we proposed six combined-type estimators of the population ratio $R = \bar{Y}/\bar{X}$ in post stratified sampling scheme as

$$\hat{R}_{1c} = \frac{\bar{y}_{ps}}{\bar{x}_{ps} - b\left(\bar{x}_{ps}^* - \bar{X}\right)} \tag{2.3}$$

$$\hat{R}_{2c} = \frac{\overline{y}_{ps}}{\left(\dfrac{\overline{x}_{ps}}{\overline{x}_{ps}^*}\,\overline{X}\right)} = \frac{\overline{y}_{ps}\overline{x}_{ps}^*}{\overline{x}_{ps}\overline{X}} \tag{2.4}$$

$$\hat{R}_{3c} = \frac{\overline{y}_{ps}}{\left(\dfrac{\overline{x}_{ps}\overline{x}_{ps}^*}{\overline{X}}\right)} = \frac{\overline{y}_{ps}\overline{X}}{\overline{x}_{ps}\overline{x}_{ps}^*} \tag{2.5}$$

$$\hat{R}_{4c} = \frac{\overline{y}_{ps}}{\overline{x}_{ps}^*} \tag{2.6}$$

$$\hat{R}_{5c} = \frac{\overline{y}_{ps}}{\overline{x}_{ps}^* - b\left(\overline{x}_{ps} - \overline{X}\right)} \tag{2.7}$$

$$\hat{R}_{6c} = \frac{\overline{y}_{ps}}{\left(\dfrac{\overline{x}_{ps}^*}{\overline{x}_{ps}}\,\overline{X}\right)} = \frac{\overline{y}_{ps}\overline{x}_{ps}}{\overline{x}_{ps}^*\overline{X}}. \tag{2.8}$$

## 2.1. Conditional Properties of the Proposed Estimators

Reference [14] defined that under the conditional argument, that is, for the achieved sample configuration, $\underline{n} = \left(n_1, n_2, n_3, \cdots, n_L\right)$ the post stratified estimator, $\overline{y}_{ps}$ is unbiased for the population mean, $\overline{Y}$, with variance

$$V_2\left(\overline{y}_{ps}\right) = \sum_{h=1}^{L} \omega_h^2 \left(1 - f_h\right)\frac{S_{yh}^2}{n_h} = \sum_{h=1}^{L} \frac{\omega_h^2 S_{yh}^2}{n_h} - \frac{1}{N}\sum_{h=1}^{L} \omega_h S_{yh}^2 \tag{2.9}$$

where $V_2$ refers to conditional variance and $S_{yh}^2$ is the population variance of $y$ in stratum $h$. Similarly, Onyeka (2012) obtained the conditional variance of $\overline{x}_{ps}$ and the conditional covariance of $\overline{y}_{ps}$ and $\overline{x}_{ps}$ respectively as:

$$V_2\left(\overline{x}_{ps}\right) = \sum_{h=1}^{L} \omega_h^2 \left(1 - f_h\right)\frac{S_{xh}^2}{n_h} = \sum_{h=1}^{L} \frac{\omega_h^2 S_{xh}^2}{n_h} - \frac{1}{N}\sum_{h=1}^{L} \omega_h S_{xh}^2 \tag{2.10}$$

and

$$C_2\left(\overline{y}_{ps}, \overline{x}_{ps}\right) = \sum_{h=1}^{L} \omega_h^2 \left(1 - f_h\right)\frac{S_{yxh}}{n_h} = \sum_{h=1}^{L} \frac{\omega_h^2 S_{yxh}}{n_h} - \frac{1}{N}\sum_{h=1}^{L} \omega_h S_{yxh} \tag{2.11}$$

where $S_{xh}^2$ is the population variance of $x$ in stratum $h$, $S_{yxh}$ is the covariance of $y$ and $x$ in stratum $h$, and $C_2$ refers to conditional covariance.

Let

$$e_0 = \frac{\overline{y}_{ps} - \overline{Y}}{\overline{Y}} \quad \text{and} \quad e_1 = \frac{\overline{x}_{ps} - \overline{X}}{\overline{X}}. \tag{2.12}$$

Then, under the conditional argument,

$$E_2\left(e_0\right) = E_2\left(e_1\right) = 0 \tag{2.13}$$

$$E_2\left(e_0^2\right) = \frac{V_2\left(\overline{y}_{ps}\right)}{\overline{Y}^2} = \frac{1}{\overline{Y}^2}\sum_{h=1}^{L} \omega_h^2 \left(1 - f_h\right)\frac{S_{yh}^2}{n_h} \tag{2.14}$$

$$E_2\left(e_1^2\right) = \frac{V_2\left(\overline{x}_{ps}\right)}{\overline{X}^2} = \frac{1}{\overline{X}^2}\sum_{h=1}^{L} \omega_h^2 \left(1 - f_h\right)\frac{S_{xh}^2}{n_h} \tag{2.15}$$

$$E_2\left(e_0 e_1\right) = \frac{C_2\left(\overline{y}_{ps}, \overline{x}_{ps}\right)}{\overline{Y}\overline{X}} = \frac{1}{\overline{Y}\overline{X}} \sum_{h=1}^{L} \omega_h^2 \left(1 - f_h\right) \frac{S_{yxh}}{n_h}. \tag{2.16}$$

Using (2.12), the first proposed estimator, $\hat{R}_{1C}$, given in (2.3), can be re-written up to first order approximation, $o\left(n^{-1}\right)$, in expected value, as

$$\left(\hat{R}_{1c} - R\right) = R\left[e_0 - \left(1 + b\pi\right)e_1 - \left(1 + b\pi\right)e_0 e_1 + \left(1 + b\pi\right)^2 e_1^2\right] \tag{2.17}$$

and

$$\left[\hat{R}_{1c} - R\right]^2 = R^2\left[e_0^2 + \left(1 + b\pi\right)^2 e_1^2 - 2\left(1 + b\pi\right)e_0 e_1\right]. \tag{2.18}$$

We take conditional expectation of (2.17) and (2.18), and use (2.13) to (2.16) to make the necessary substitutions. This gives the conditional bias and mean square error of $\hat{R}_{1C}$ respectively as

$$B_2\left(\hat{R}_{1C}\right) = \frac{1}{\overline{X}^2}\left[\left(1 + b\pi\right)^2 RA_{22} - \left(1 + b\pi\right)A_{12}\right] \tag{2.19}$$

and

$$\mathrm{MSE}_2\left(\hat{R}_{1C}\right) = \frac{1}{\overline{X}^2}\left[A_{11} + \left(1 + b\pi\right)^2 R^2 A_{22} - 2\left(1 + b\pi\right)RA_{12}\right] \tag{2.20}$$

where

$$A_{11} = \sum_{h=1}^{L} \frac{\omega_h^2\left(1 - f_h\right)S_{yh}^2}{n_h}, \quad A_{22} = \sum_{h=1}^{L} \frac{\omega_h^2\left(1 - f_h\right)S_{xh}^2}{n_h}, \quad A_{12} = \sum_{h=1}^{L} \frac{\omega_h^2\left(1 - f_h\right)S_{yxh}}{n_h}. \tag{2.21}$$

Following similar procedure, we obtain the conditional biases and mean square errors of the six proposed estimators, together with those of the customary combined-type estimator, $\hat{R}_C = \overline{y}_{ps}/\overline{x}_{ps}$, of population ratio $(R)$, in post-stratified sampling, up to first order approximation, $o\left(n^{-1}\right)$, as:

$$B_2\left(\hat{R}_C\right) = \frac{1}{\overline{X}^2}\left[RA_{22} - A_{12}\right] \tag{2.22}$$

$$B_2\left(\hat{R}_{1C}\right) = \frac{1}{\overline{X}^2}\left(1 + b\pi\right)\left[\left(1 + b\pi\right)RA_{22} - A_{12}\right] \tag{2.23}$$

$$B_2\left(\hat{R}_{2C}\right) = \frac{1}{\overline{X}^2}\left(1 + \pi\right)\left[\left(1 + \pi\right)RA_{22} - A_{12}\right] \tag{2.24}$$

$$B_2\left(\hat{R}_{3C}\right) = \frac{1}{\overline{X}^2}\left[\left(1 - \pi + \pi^2\right)RA_{22} - \left(1 - \pi\right)A_{12}\right] \tag{2.25}$$

$$B_2\left(\hat{R}_{4C}\right) = \frac{1}{\overline{X}^2}\left[\pi^2 RA_{22} + \pi A_{12}\right] \tag{2.26}$$

$$B_2\left(\hat{R}_{5C}\right) = \frac{1}{\overline{X}^2}\left(\pi + b\right)\left[\left(\pi + b\right)RA_{22} + A_{12}\right] \tag{2.27}$$

$$B_2\left(\hat{R}_{6C}\right) = \frac{1}{\overline{X}^2}\left(1 + \pi\right)\left[\pi RA_{22} + A_{12}\right] \tag{2.28}$$

and

$$\mathrm{MSE}_2\left(\hat{R}_C\right) = \frac{1}{\overline{X}^2}\left[A_{11} + R^2 A_{22} - 2RA_{12}\right] \tag{2.29}$$

$$\mathrm{MSE}_2\left(\hat{R}_{1C}\right) = \frac{1}{\overline{X}^2}\left[A_{11} + \left(1 + b\pi\right)^2 R^2 A_{22} - 2\left(1 + b\pi\right)RA_{12}\right] \tag{2.30}$$

$$\text{MSE}_2\left(\hat{R}_{2C}\right) = \frac{1}{\overline{X}^2}\left[A_{11} + \left(1+\pi\right)^2 R^2 A_{22} - 2\left(1+\pi\right)RA_{12}\right] \tag{2.31}$$

$$\text{MSE}_2\left(\hat{R}_{3c}\right) = \frac{1}{\overline{X}^2}\left[A_{11} + \left(1-\pi\right)^2 R^2 A_{22} - 2\left(1-\pi\right)RA_{12}\right] \tag{2.32}$$

$$\text{MSE}_2\left(\hat{R}_{4C}\right) = \frac{1}{\overline{X}^2}\left[A_{11} + \pi^2 R^2 A_{22} + 2\pi RA_{12}\right] \tag{2.33}$$

$$\text{MSE}_2\left(\hat{R}_{5C}\right) = \frac{1}{\overline{X}^2}\left[A_{11} + \left(\pi+b\right)^2 R^2 A_{22} + 2\left(\pi+b\right)RA_{12}\right] \tag{2.34}$$

$$\text{MSE}_2\left(\hat{R}_{6C}\right) = \frac{1}{\overline{X}^2}\left[A_{11} + \left(1+\pi\right)^2 R^2 A_{22} + 2\left(1+\pi\right)RA_{12}\right]. \tag{2.35}$$

Generally, we have for the proposed six combined-type estimators,

$$\text{MSE}_2\left(\hat{R}_{qc}\right) = \frac{1}{\overline{X}^2}\left[A_{11} + \theta_q^2 R^2 A_{22} - 2\theta_q RA_{12}\right] \tag{2.36}$$

where $q = 1, \cdots, 6$ and

$$\theta_1 = \left(1+b\pi\right), \quad \theta_2 = \left(1+\pi\right), \quad \theta_3 = \left(1-\pi\right), \quad \theta_4 = -\pi, \quad \theta_5 = -\left(\pi+b\right), \quad \theta_6 = -\left(1+\pi\right). \tag{2.37}$$

## 2.2. Unconditional Properties of the Proposed Estimators

Following [14] we obtain the following (unconditional) variances and covariance, for repeated samples of fixed size *n*.

$$V\left(\overline{y}_{ps}\right) = \left(\frac{1-f}{n}\right)\sum_{h=1}^{L}\omega_h S_{yh}^2 \tag{2.38}$$

$$V\left(\overline{x}_{ps}\right) = \left(\frac{1-f}{n}\right)\sum_{h=1}^{L}\omega_h S_{xh}^2 \tag{2.39}$$

and

$$\text{Cov}\left(\overline{y}_{ps}, \overline{x}_{ps}\right) = \left(\frac{1-f}{n}\right)\sum_{h=1}^{L}\omega_h S_{yxh} \tag{2.40}$$

where $f = n/N$ is the population sampling fraction. By taking unconditional expectations of (2.17) and (2.18), and using (2.38)-(2.40) to make the necessary substitutions, we obtain the unconditional bias and mean square errors of the first proposed estimator, $\hat{R}_{1c}$, up to first order approximation, $o\left(n^{-1}\right)$, as:

$$B\left(\hat{R}_{1C}\right) = \frac{1}{\overline{X}^2}\left(1+b\pi\right)\left(\frac{1-f}{n}\right)\left[\left(1+b\pi\right)RA_{22}' - A_{12}'\right] \tag{2.41}$$

and

$$\text{MSE}\left(\hat{R}_{2C}\right) = \frac{1}{\overline{X}^2}\left(\frac{1-f}{n}\right)\left[A_{11}' + \left(1+\pi\right)^2 R^2 A_{22}' - 2\left(1+\pi\right)RA_{12}'\right] \tag{2.42}$$

where

$$A_{11}' = \sum_{h=1}^{L}\omega_h S_{yh}^2, \quad A_{22}' = \sum_{h=1}^{L}\omega_h S_{xh}^2, \quad A_{12}' = \sum_{h=1}^{L}\omega_h S_{yxh}. \tag{2.43}$$

Following similar procedure, we obtain the unconditional biases and mean square errors of the six proposed estimators, together with those of the customary combined-type estimator, $\hat{R}_C = \overline{y}_{ps}/\overline{x}_{ps}$, of population ratio $(R)$, in post-stratified sampling, up to first order approximation, $o\left(n^{-1}\right)$, as:

$$B\left(\hat{R}_C\right) = \frac{1}{\overline{X}^2}\left(\frac{1-f}{n}\right)\left[RA_{22}' - A_{12}'\right] \tag{2.44}$$

$$B\left(\hat{R}_{1C}\right) = \frac{1}{\bar{X}^2}\left(\frac{1-f}{n}\right)\left(1+b\pi\right)\left[\left(1+b\pi\right)RA'_{22} - A'_{12}\right] \tag{2.45}$$

$$B\left(\hat{R}_{2C}\right) = \frac{1}{\bar{X}^2}\left(\frac{1-f}{n}\right)\left(1+\pi\right)\left[\left(1+\pi\right)RA'_{22} - A'_{12}\right] \tag{2.46}$$

$$B\left(\hat{R}_{3C}\right) = \frac{1}{\bar{X}^2}\left(\frac{1-f}{n}\right)\left[\left(1-\pi+\pi^2\right)RA'_{22} - \left(1-\pi\right)A'_{12}\right] \tag{2.47}$$

$$B\left(\hat{R}_{4C}\right) = \frac{1}{\bar{X}^2}\left(\frac{1-f}{n}\right)\left[R\pi^2 A'_{12} + \pi A'_{12}\right] \tag{2.48}$$

$$B\left(\hat{R}_{5C}\right) = \frac{1}{\bar{X}^2}\left(\frac{1-f}{n}\right)\left(b+\pi\right)\left[\left(b+\pi\right)RA'_{22} + A'_{12}\right] \tag{2.49}$$

$$B\left(\hat{R}_{6C}\right) = \frac{1}{\bar{X}^2}\left(\frac{1-f}{n}\right)\left(1+\pi\right)\left[\pi RA'_{22} + A'_{12}\right] \tag{2.50}$$

and,

$$\text{MSE}\left(\hat{R}_{C}\right) = \frac{1}{\bar{X}^2}\left(\frac{1-f}{n}\right)\left[A'_{11} + R^2 A'_{22} - 2RA'_{12}\right] \tag{2.51}$$

$$\text{MSE}\left(\hat{R}_{1C}\right) = \frac{1}{\bar{X}^2}\left(\frac{1-f}{n}\right)\left[A'_{11} + R^2\left(1+\pi b\right)A'_{22} - 2RA'_{12}\right] \tag{2.52}$$

$$\text{MSE}\left(\hat{R}_{2C}\right) = \frac{1}{\bar{X}^2}\left(\frac{1-f}{n}\right)\left[A'_{11} + \left(1+\pi\right)^2 R^2 A'_{22} - 2\left(1+\pi\right)RA'_{12}\right] \tag{2.53}$$

$$\text{MSE}\left(\hat{R}_{3C}\right) = \frac{1}{\bar{X}^2}\left(\frac{1-f}{n}\right)\left[A'_{11} + \left(1-\pi\right)^2 R^2 A'_{22} - 2\left(1-\pi\right)RA'_{12}\right] \tag{2.54}$$

$$\text{MSE}\left(\hat{R}_{4C}\right) = \frac{1}{\bar{X}^2}\left(\frac{1-f}{n}\right)\left[A'_{11} + \pi^2 R^2 A'_{22} + 2R\pi A'_{12}\right] \tag{2.55}$$

$$\text{MSE}\left(\hat{R}_{5C}\right) = \frac{1}{\bar{X}^2}\left(\frac{1-f}{n}\right)\left[A'_{11} + \left(b+\pi\right)^2 R^2 A'_{22} + 2\left(b+\pi\right)RA'_{12}\right] \tag{2.56}$$

$$\text{MSE}\left(\hat{R}_{6C}\right) = \frac{1}{\bar{X}^2}\left(\frac{1-f}{n}\right)\left[A'_{11} + \left(1+\pi\right)^2 R^2 A'_{22} + 2\left(1+\pi\right)RA'_{12}\right]. \tag{2.57}$$

Generally, the unconditional mean square errors of the proposed combined-type estimators is obtained as

$$\text{MSE}\left(\hat{R}_{qC}\right) = \frac{1}{\bar{X}^2}\left[\frac{1-f}{n}\right]\left[A'_{11} + \theta_q^2 R^2 A'_{22} - 2\theta_q RA'_{12}\right] \tag{2.58}$$

where $\theta_q$, $q = 1,\cdots,6$ is as given in (2.37).

## 3. Efficiency Comparison

The efficiencies of the six proposed combined-type estimators are first compared with that of the customary combined ratio estimator $\hat{R}_C$ in estimating the population ratio $R$ of two population means under the conditional and unconditional arguments in post-stratified random sampling scheme. Secondly, the performances of the proposed estimators among themselves are investigated. Furthermore, the optimum estimators among the proposed estimators are also obtained. The efficiency comparison is carried out using the mean square errors of the estimators and the results are shown in **Table 1**.

## 4. Numerical Illustration

Here, we use the final year GPA $(y)$ and the level of absenteeism $(x)$ of 2012/2013 graduating students of

6

Statistics Department, Federal University of Technology Owerri to illustrate the properties of the estimators proposed in the present study. Absenteeism is measured as the average number of days absent from lectures in a month. The class consists of 50 students, with 32 and 18 students respectively falling into low-absenteeism (0 - 3 days per month) and high-absenteeism (4 - 6 days per month) groups or strata. Our interest is to estimate the ratio of final year GPA to absenteeism from lectures, based on a post-stratified sample of 20 out of the 50 graduating students in the class. The data statistics, consisting mainly of population parameters are shown in **Table 2**.

**Table 3** shows the percentage relative efficiencies (PRE-1) of the proposed combined-type estimators, $\hat{R}_{qc}$,

**Table 1.** Efficiency conditions under conditional and unconditional arguments.

| Estimator | Conditional argument | Unconditional argument |
|---|---|---|
| $R_{qc}$ is better than $R_c$ if: | 1) $\|\theta_q\| < 1$ and $\beta < R$ <br> or <br> 2) $\|\theta_q\| > 1$ and $\beta > R$ | 1) $\|\theta_q\| < 1$ and $\beta' < R$ <br> or <br> 2) $\|\theta_q\| > 1$ and $\beta' > R$ |
| $R_{kc}$ is better than $R_{jc}$ if: | 1) $\|\theta_j\| < \|\theta_k\|$ and $\|\theta_j\| < \beta/R$ <br> or <br> 2) $\|\theta_j\| > \|\theta_k\|$ and $\|\theta_j\| > \beta/R$ | 1) $\|\theta_j\| < \|\theta_k\|$ and $\|\theta_j\| < \beta'/R$ <br> or <br> 2) $\|\theta_j\| > \|\theta_k\|$ and $\|\theta_j\| > \beta'/R$ |
| $R_{qc}$ is optimum if: | $\|\theta_q^0\| = \beta/R$ | $\|\theta_q^0\| = \beta'/R$ |

Where $\beta = A_{12}/A_{22}$, $\beta' = A'_{12}/A'_{22}$ and $\theta_q$, $q = 1, \cdots, 6$ is as given in (2.37).

**Table 2.** Data statistics for final year GPA $(y)$ and absenteeism from lectures $(x)$.

| Population/sample parameters | Stratum 1 (low-absenteeism) | Stratum 2 (high-absenteeism) |
|---|---|---|
| $N = 50$ | $N_1 = 32$ | $N_2 = 18$ |
| $n = 20$ | $n_1 = 12$ | $n_2 = 8$ |
| $(1 - f) = 0.60$ | $(1 - f_1) = 0.625$ | $(1 - f_2) = 0.556$ |
| $\overline{Y} = 2.98$ | $\overline{Y}_1 = 3.16$ | $\overline{Y}_2 = 2.65$ |
| $\overline{X} = 3.16$ | $\overline{X}_1 = 2.03$ | $\overline{X}_2 = 5.17$ |
| $R = 0.94$ | $R_1 = 1.56$ | $R_2 = 0.51$ |
| $\pi = 0.67$ | $S_{y1}^2 = 0.2422$ | $S_{y2}^2 = 0.0389$ |
| $\omega_1 = 0.64$ | $S_{x1}^2 = 0.9990$ | $S_{x2}^2 = 0.6176$ |
|  | $S_{yx1} = -0.2124$ | $S_{yx2} = -0.0161$ |
|  | $\omega_1 = 0.64$ | $\omega_2 = 0.36$ |

**Table 3.** Percentage relative efficiencies under conditional and unconditional arguments.

| Estimator | $\theta$ | Conditional argument | | | Unconditional argument | | |
|---|---|---|---|---|---|---|---|
|  |  | MSE | PRE-1 (%) | PRE-2 (%) | MSE | PRE-1 (%) | PRE-2 (%) |
| $\hat{R}_{1c}$ | 0.464 | 0.00148 | 259 | 100 | 0.00091 | 262 | 100 |
| $\hat{R}_{2c}$ | 1.670 | 0.00872 | 44 | 588 | 0.00548 | 44 | 600 |
| $\hat{R}_{3c}$ | 0.330 | 0.00111 | 346 | 75 | 0.00068 | 352 | 74 |
| $\hat{R}_{4c}$ | −0.670 | 0.00104 | 370 | 70 | 0.00067 | 360 | 73 |
| $\hat{R}_{5c}$ | 0.130 | 0.00071 | 539 | 48 | 0.00043 | 553 | 47 |
| $\hat{R}_{6c}$ | −1.670 | 0.00576 | 67 | 388 | 0.00370 | 65 | 405 |
| $\hat{R}_{c}$ | 1.000 | 0.00384 | 100 | 259 | 0.00240 | 100 | 262 |
| $\hat{R}_{qc}^0$ |  | 0.00046 | 836 | 31 | 0.00028 | 854 | 31 |

over the customary combined-type estimator, $\hat{R}_c$, under the conditional and under the unconditional arguments. The table also shows the percentage relative efficiency (PRE-2) of the proposed combined-type estimators, $\hat{R}_{1c}$, over the other combined-type estimators, under the conditional and under the unconditional arguments.

**Table 3** shows that apart from the estimators, $\hat{R}_{2c}$ and $\hat{R}_{6c}$, the remaining four proposed combined-type estimators, under the conditional and under the unconditional arguments, are more efficient than the customary combined-type estimator, $\hat{R}_c$, for the data under consideration, and their gains in efficiency (PRE-1) are relatively large. Also, using PRE-2, we observe that the proposed combined-type estimator, $\hat{R}_{1c}$, is more efficient than the estimators, $\hat{R}_{2c}$, $\hat{R}_{6c}$, and $\hat{R}_c$, under the conditional and unconditional arguments. The optimum estimator, as expected, has the highest gain in efficiency, both under the conditional and unconditional arguments. However, the customary combined-type estimator, on the other hand, is found to be more efficient than some of the proposed combined-type estimators for the given set of data. This confirms the theoretical results, which showed that the proposed estimators are not always more efficient than the customary combined-type estimator. Notice that $\beta' = -0.16$ and $R = 0.94$ showing that $\beta' < R$ and from the theoretical results in **Table 1**, the proposed estimators would be more efficient than the customary combined-type estimator, under the unconditional argument, if $|\theta_q| < 1$. The empirical results in **Table 3** show that $|\theta_2| > 1$ and $|\theta_6| > 1$, and the proposed estimators $\hat{R}_2$ (PRE-1 = 44%) and $\hat{R}_6$ (PRE-1 = 65%) under the unconditional argument, are less efficient than the customary combined-type estimator, $\hat{R}_c$. Hence the empirical results confirm the theoretical results.

## 5. Concluding Remarks

The study extends use of variable transformation in estimating population ratio in simple random sampling scheme to post-stratified sampling scheme. Efficiency conditions for preferring the proposed estimators to the customary combined-type estimator are obtained. The study shows that in any given survey, these efficiency conditions should be employed in order to determine the appropriate proposed combined-type estimators to use for the purpose of estimating the population ratio of two variables in post-stratified sampling scheme, using variable transformation.

## References

[1] Onyeka, A.C., Nlebedim, V.U. and Izunobi, C.H. (2013) Estimation of Population Ratio in Simple Random Sampling Using Variable Transformation. *Global Journal of Science Frontier Research*, **13**, 57-65.

[2] Sukhatme, P.V. and Sukhatme, B.V. (1970) Sampling Theory of Surveys with Applications. Iowa State University Press, Ames.

[3] Cochran, W.G. (1977) Sampling Techniques. 3rd Edition, John Wiley & Sons, New York.

[4] Cochran, W.G. (1940) The Estimation of the Yields of the Cereal Experiments by Sampling for the Ratio of Grain to Total Produce. *The Journal of Agricultural Science*, **30**, 262-275.
http://dx.doi.org/10.1017/S0021859600048012

[5] Robson, D.S. (1957) Application of Multivariate Polykays to the Theory of Unbiased Ratio-Type Estimation. *Journal of the American Statistical Association*, **52**, 511-522.
http://dx.doi.org/10.1080/01621459.1957.10501407

[6] Murthy, M.N. (1964) Product Method of Estimation. *Sankhya Series A*, **26**, 294-307.

[7] Singh, M.P. (1965) On the Estimation of Ratio and Product of the Population Parameters. *Sankhya Series B*, **27**, 321-328.

[8] Upadhyaya, L.N., Singh, G.N. and Singh, H.P. (2000) Use of Transformed Auxiliary Variable in the Estimation of Population Ratio in Sample Survey. *Statistics in Transition*, **4**, 1019-1027.

[9] Onyeka, A.C., Nlebedim, V.U. and Izunobi, C.H. (2014) A Class of Estimators for Population Ratio in Simple Random Sampling Using Variable Transformation. *Open Journal of Statistics*, **4**, 284-291.
http://dx.doi.org/10.4236/ojs.2014.44029

[10] Srivenkataramana, T. (1980) A Dual of Ratio Estimator in Sample Surveys. *Biometrika*, **67**, 199-204.
http://dx.doi.org/10.1093/biomet/67.1.199

[11] Singh, H.P. and Tailor, R. (2005) Estimation of Finite Population Mean Using Known Correlation Coefficient between Auxiliary Characters. *Statistica*, **4**, 407-418.

[12] Tailor, R. and Sharma, B.K. (2009) A Modified Ratio-Cum-Product Estimator of Finite Population Mean Using Known Coefficient of Variation and Coefficient of Kurtosis. *Statistics in Transition—New Series*, **10**, 15-24.

[13] Sharma, B. and Tailor, R. (2010) A New Ratio-Cum-Dual to Ratio Estimator of Finite Population Mean in Simple Random Sampling. *Global Journal of Science Frontier Research*, **10**, 27-31.

[14] Onyeka, A.C. (2012) Estimation of Population Mean in Post-Stratified Sampling Using Known Value of Some Population Parameter(s). *Statistics in Transition—New Series*, **13**, 65-78.

Scientific Research Publishing (SCIRP) is one of the largest Open Access journal publishers. It is currently publishing more than 200 open access, online, peer-reviewed journals covering a wide range of academic disciplines. SCIRP serves the worldwide academic communities and contributes to the progress and application of science with its publication.

Other selected journals from SCIRP are listed as below. Submit your manuscript to us via either submit@scirp.org or Online Submission Portal.