STATISTICS IN TRANSITION-new series, March 2012
Vol. 13, No. 1, pp. 65—78

ESTIMATION OF POPULATION MEAN IN POST-STRATIFIED SAMPLING USING KNOWN VALUE OF SOME POPULATION PARAMETER(S)

Aloy C. Onyeka[1]

ABSTRACT

Following Khoshnevisan et.al. (2007) and Koyuncu and Kadilar (2009), this paper develops a general family of combined estimators of the population mean in post-stratified sampling (PSS) scheme, using known values of some population parameters of an auxiliary variable. Properties of the proposed family of estimators, including conditions for optimal efficiency, are obtained up to first order approximations. The results are illustrated empirically.

Key words: Auxiliary information, general family of estimators, post-stratified sampling, mean squared errors
2000 AMS Classification: 62D05

1. Introduction

The use of auxiliary information in constructing estimators of population parameters of the variable of interest is highly encouraged in surveys, especially when an auxiliary variable is highly correlated with the study variable. Apart from using the known population mean $\bar{X}$ of an auxiliary variable x, many authors have ventured into the use of other known population parameters of x. Notable studies along this line, under the simple random sampling without replacement (SRSWOR) scheme, include Searls (1964), who used the coefficient of variation (CV) of x in estimating the population mean $\bar{Y}$ of the study variable y, and Sisodia and Dwivedi (1981), who used the CV of x in ratio estimation of $\bar{Y}$. Singh et.al. (1973) used known coefficient of kurtosis in estimating the population variance of y. Sen (1978) and Searls and Intarapanich (1990) also used known coefficient of kurtosis in estimating $\bar{Y}$. Singh and Tailor (2003) used known correlation coefficient in ratio estimation of $\bar{Y}$. Singh (2003) used known

1 Department of Statistics. Federal University of Technology. PMB 1526, Owerri, Nigeria. E-mail: aloyonyeka@yahoo.com.

standard deviation of the auxiliary variable x in estimating $\bar{Y}$. Khoshnevisan et.al. (2007) proposed a general family of estimators of $\bar{Y}$ under the SRSWOR scheme, which uses known parameters of the auxiliary variable x such as standard deviation, coefficient of variation, skewness, kurtosis and correlation coefficient. Motivated by Khoshnevisan et.al. (2007), the present study intends to develop a general family of estimators of $\bar{Y}$ under the post-stratified sampling scheme.

Under the stratified random sampling scheme, Cochran (1977) discussed the usual stratified sampling estimator, $\bar{y}_{st}$, and also separate and combined ratio-type estimators of $\bar{Y}$. Kadilar and Cingi (2003), motivated by the works done under the SRSWOR scheme by Sisodia and Dwivedi (1981), Singh and Kakran (1993), and Upadhyaya and Singh (1999), proposed some estimators of $\bar{Y}$ in stratified random sampling using known values of population mean $\bar{X}_h$, coefficient of variation C_{xh}, and coefficient of kurtosis $\beta_{2h}(x)$ of the auxiliary variable x in stratum h. Kadilar and Cingi (2003) restricted their work to ratio estimation of $\bar{Y}$ in stratified random sampling. Koyuncu and Kadilar (2009) proposed a more general family of combined estimators of $\bar{Y}$ in stratified random sampling along the line of Khoshnevisan et.al. (2007). Chaudhary et.al. (2009) considered, in a more recent paper, a general family of combined estimators of $\bar{Y}$ in stratified random sampling under non-response. Motivated by Khoshnevisan et.al. (2007) and Koyuncu and Kadilar (2009), we consider in the present study, a general class of combined-type estimators of $\bar{Y}$ in post-stratified sampling scheme, using information on known values of some population parameters of an auxiliary variable.

2. The Proposed Estimators

Let y_{hi} (x_{hi}) denote the ith observation in stratum h for the study (auxiliary) variate in post-stratified sampling scheme. Let a random sample of size n be drawn from a population of N units using SRSWOR method, and let the sampled units be allocated to their respective strata, where n_h (a random variable) is the number of units that fall into stratum h such that $\sum_{h=1}^{L} n_h = n$. We assume that n is large enough such that $P(n_h = 0) = 0, \forall\ h$. Following Khoshnevisan et.al. (2007) and Koyuncu and Kadilar (2009), we propose a general family of combined estimators of the population mean $\bar{Y}$ in post-stratified sampling scheme as

$$\bar{y}_{pss} = \bar{y}_{ps}\left(\frac{a\bar{X}+b}{\alpha(a\bar{x}_{ps}+b)+(1-\alpha)(a\bar{X}+b)}\right)^{g} \tag{2.1}$$

where,

$\bar{y}_{ps} = \sum_{h=1}^{L}\omega_h\bar{y}_h$ is the usual post-stratified estimator of $\bar{Y}$

$\bar{x}_{ps} = \sum_{h=1}^{L}\omega_h\bar{x}_h$ is the usual post-stratified estimator of $\bar{X}$

$\bar{X} = \sum_{h=1}^{L}\omega_h\bar{X}_h$ is the known population mean of the auxiliary variate x.

$a(\neq 0)$, b are either constants or functions of known population parameters of the auxiliary variate, such as Standard deviation (σ_x), Coefficient of variation (C_x), Skewness $(\beta_1(x))$, Kurtosis $(\beta_2(x))$, and Correlation coefficient (ρ_{yx}), and

$\omega_h = N_h/N$ is stratum weight, L is the number of strata in the population, N_h is the number of units in stratum h, N is the number of units in the population, $\bar{X}_h$ is the population mean of the auxiliary variate in stratum h, and $\bar{y}_h$ $(\bar{x}_h)$ is the sample mean of the study (auxiliary) variate in stratum h.

Under the conditional argument, that is, for the achieved sample configuration, $\underline{n} = (n_1, n_2, \cdots, n_L)$, the post-stratified estimator, $\bar{y}_{ps}$ is unbiased for $\bar{Y}$ with variance,

$$V_2(\bar{y}_{ps}) = \sum_{h=1}^{L}\omega_h^2\left(1-\frac{n_h}{N_h}\right)\frac{S_{yh}^2}{n_h} = \sum_{h=1}^{L}\frac{\omega_h^2 S_{yh}^2}{n_h} - \frac{1}{N}\sum_{h=1}^{L}\omega_h S_{yh}^2 \tag{2.2}$$

where V_2 refers to "conditional variance" and S_{yh}^2 is the population variance of y in stratum h.

For repeated samples of fixed size n, we obtain the unconditional variance of $\bar{y}_{ps}$ by taking the expectation of equation (2.2). This gives the unconditional variance of $\bar{y}_{ps}$ as:

$$V(\bar{y}_{ps}) = E(V_2(\bar{y}_{ps})) = \sum_{h=1}^{L}\omega_h^2\, E\left(\frac{1}{n_h}\right)S_{yh}^2 - \frac{1}{N}\sum_{h=1}^{L}\omega_h S_{yh}^2 \tag{2.3}$$

Following Stephan (1945), we obtain, to terms of order n^{-2},

$$E\left(\frac{1}{n_h}\right)=\frac{1}{n\omega_h}+\frac{1-\omega_h}{n^2\omega_h^2} \tag{2.4}$$

Consequently, the unconditional variance of $\bar{y}_{ps}$ obtained up to first order approximation is

$$V(\bar{y}_{ps})=\left(\frac{1-f}{n}\right)\sum_{h=1}^{L}\omega_h S_{yh}^2 \tag{2.5}$$

Similarly, the unconditional variance of $\bar{x}_{ps}$ and the unconditional covariance of $\bar{y}_{ps}$ and $\bar{x}_{ps}$ are obtained respectively as

$$V(\bar{x}_{ps})=\left(\frac{1-f}{n}\right)\sum_{h=1}^{L}\omega_h S_{xh}^2 \tag{2.6}$$

and

$$Cov(\bar{y}_{ps},\bar{x}_{ps})=\left(\frac{1-f}{n}\right)\sum_{h=1}^{L}\omega_h S_{yxh} \tag{2.7}$$

where $f=n/N$ is the population sampling fraction, S_{xh}^2 is the population variance of x in stratum h, and S_{yxh} is the population covariance of y and x in stratum h. Let

$$e_0=\frac{\bar{y}_{ps}-\bar{Y}}{\bar{Y}} \tag{2.8}$$

and

$$e_1=\frac{\bar{x}_{ps}-\bar{X}}{\bar{X}} \tag{2.9}$$

Under the unconditional argument, we obtain

$$E(e_0)=E(e_1)=0 \tag{2.10}$$

$$E(e_0^2)=\frac{V(\bar{y}_{ps})}{\bar{Y}^2}=\frac{1}{\bar{Y}^2}\left(\frac{1-f}{n}\right)\sum_{h=1}^{L}\omega_h S_{yh}^2 \tag{2.11}$$

$$E(e_1^2) = \frac{V(\bar{x}_{ps})}{\bar{X}^2} = \frac{1}{\bar{X}^2}\left(\frac{1-f}{n}\right)\sum_{h=1}^{L}\omega_h S_{xh}^2 \tag{2.12}$$

and

$$E(e_0 e_1) = \frac{Cov(\bar{y}_{ps}, \bar{x}_{ps})}{\overline{YX}} = \frac{1}{\overline{YX}}\left(\frac{1-f}{n}\right)\sum_{h=1}^{L}\omega_h S_{yxh} \tag{2.13}$$

We can rewrite equation (2.1) in terms of e_0 and e_1 as

$$\bar{y}_{pss} = \bar{Y}(1+e_0)(1+\alpha\lambda e_1)^{-g} \tag{2.14}$$

where $\lambda = \dfrac{a\bar{X}}{a\bar{X}+b}$. Assuming $|\alpha\lambda e_1| < 1$, so that the series $(1+\alpha\lambda e_1)^{-g}$ converges, and expanding equation (2.14) up to first order approximations in expected value, we obtain the expressions:

$$(\bar{y}_{pss} - \bar{Y}) = \bar{Y}(e_0 - \alpha\lambda g e_1 - \alpha\lambda g e_0 e_1 + \tfrac{1}{2}\alpha^2\lambda^2 g(g+1)e_1^2) \tag{2.15}$$

and

$$(\bar{y}_{pss} - \bar{Y})^2 = \bar{Y}^2(e_0^2 + \alpha^2\lambda^2 g^2 e_1^2 - 2\alpha\lambda g e_0 e_1) \tag{2.16}$$

To obtain the unconditional bias and mean squared error of the proposed estimators $\bar{y}_{pss}$ we take the unconditional expectations of equations (2.15) and (2.16), and use equations (2.10) – (2.13) to make the necessary substitutions. This gives the unconditional bias and mean squared error of $\bar{y}_{pss}$, respectively as

$$B(\bar{y}_{pss}) = \frac{\alpha\lambda g}{2\bar{X}}\left(\frac{1-f}{n}\right)\left(\sum_{h=1}^{L}\omega_h(\alpha\lambda(g+1)RS_{xh}^2 - 2S_{yxh})\right) \tag{2.17}$$

and

$$MSE(\bar{y}_{pss}) = \left(\frac{1-f}{n}\right)\sum_{h=1}^{L}\omega_h(S_{yh}^2 + \alpha^2\lambda^2 g^2 R^2 S_{xh}^2 - 2\alpha\lambda g R S_{yxh}) \tag{2.18}$$

where $R = \bar{Y}/\bar{X}$.

3. Special Cases

The proposed estimator, $\bar{y}_{pss}$ is a general class of estimators capable of generating an infinite number of combined estimators of $\bar{Y}$ by making appropriate choices of the values of α, g, a and b in equation (2.1). The following are some special cases of the proposed estimators, $\bar{y}_{pss}$, of $\bar{Y}$ in post-stratified sampling scheme.

Estimator	Values of α	g	a	b
1. Usual stratified estimator, $\bar{y}_{pss}(1) = \bar{y}_{ps} = \sum_{h=1}^{L} \omega_h \bar{y}_h$	–	0	–	–
2. Usual Combined ratio-type estimator, $\bar{y}_{pss}(2) = \bar{y}_{psRC} = \frac{\bar{y}_{ps}}{\bar{x}_{ps}} \bar{X}$	1	1	1	0
3. Sisodia-Dwivedi (1981) estimator, $\bar{y}_{pss}(3) = \bar{y}_{psSD} = \bar{y}_{ps} \frac{\bar{X} + C_x}{\bar{x}_{ps} + C_x}$	1	1	1	C_x
4. Singh-Kakran (1993) estimator (1), $\bar{y}_{pss}(4) = \bar{y}_{psSK1} = \bar{y}_{ps} \frac{\bar{X} + \beta_2(x)}{\bar{x}_{ps} + \beta_2(x)}$	1	1	1	$\beta_2(x)$
5. Upadhyaya-Singh (1999) estimator (1), $\bar{y}_{pss}(5) = \bar{y}_{psUS1} = \bar{y}_{ps} \frac{\bar{X}\beta_2(x) + C_x}{\bar{x}_{ps}\beta_2(x) + C_x}$	1	1	$\beta_2(x)$	C_x
6. Upadhyaya-Singh (1999) estimator (2), $\bar{y}_{pss}(6) = \bar{y}_{psUS2} = \bar{y}_{ps} \frac{\bar{X}C_x + \beta_2(x)}{\bar{x}_{ps}C_x + \beta_2(x)}$	1	1	C_x	$\beta_2(x)$

Estimator	**Values of**			
	α	g	a	b
7. Singh-Tailor (2003) estimator (1), $\bar{y}_{pss}(7)=\bar{y}_{psST1}=\bar{y}_{ps}\frac{\bar{X}+\rho_{yx}}{\bar{x}_{ps}+\rho_{yx}}$	1	1	1	ρ_{yx}
8. The usual combined product-type estimator, $\bar{y}_{pss}(8)=\bar{y}_{psPC}=\frac{\bar{y}_{ps}\bar{x}_{ps}}{\bar{X}}$	1	–1	1	0
9. Pandey-Dubey (1988) estimator, $\bar{y}_{pss}(9)=\bar{y}_{psPD}=\bar{y}_{ps}\frac{\bar{x}_{ps}+C_x}{\bar{X}+C_x}$	1	–1	1	C_x
10. Upadhyaya-Singh (1999) estimator (3), $\bar{y}_{pss}(10)=\bar{y}_{psUS3}=\bar{y}_{ps}\frac{\bar{x}_{ps}\beta_2(x)+C_x}{\bar{X}\beta_2(x)+C_x}$	1	–1	$\beta_2(x)$	C_x
11. Upadhyaya-Singh (1999) estimator (4), $\bar{y}_{pss}(11)=\bar{y}_{psUS4}=\bar{y}_{ps}\frac{\bar{x}_{ps}C_x+\beta_2(x)}{\bar{X}C_x+\beta_2(x)}$	1	–1	C_x	$\beta_2(x)$
12. G.N. Singh (2003) estimator (1), $\bar{y}_{pss}(12)=\bar{y}_{psGNS1}=\bar{y}_{ps}\frac{\bar{x}_{ps}+\sigma_x}{\bar{X}+\sigma_x}$	1	–1	1	σ_x
13. G.N. Singh (2003) estimator (2), $\bar{y}_{pss}(13)=\bar{y}_{psGNS2}=\bar{y}_{ps}\frac{\bar{x}_{ps}\beta_1(x)+\sigma_x}{\bar{X}\beta_1(x)+\sigma_x}$	1	–1	$\beta_1(x)$	σ_x
14. G.N. Singh (2003) estimator (3), $\bar{y}_{pss}(14)=\bar{y}_{psGNS3}=\bar{y}_{ps}\frac{\bar{x}_{ps}\beta_2(x)+\sigma_x}{\bar{X}\beta_2(x)+\sigma_x}$	1	–1	$\beta_2(x)$	σ_x
15. Singh-Tailor (2003) estimator (2), $\bar{y}_{pss}(15)=\bar{y}_{psST2}=\bar{y}_{ps}\frac{\bar{x}_{ps}+\rho_{yx}}{\bar{X}+\rho_{yx}}$	1	–1	1	ρ_{yx}
16. Singh-Kakran (1993) estimator (2), $\bar{y}_{pss}(16)=\bar{y}_{psSK2}=\bar{y}_{ps}\frac{\bar{x}_{ps}+\beta_2(x)}{\bar{X}+\beta_2(x)}$	1	–1	1	$\beta_2(x)$

Notice that the usual post-stratified estimator $\bar{y}_{ps}$ is a special case of the proposed family of estimators $\bar{y}_{pss}$ if and only if we choose $g = 0$, no matter the values of α, a and b. Again, we observe that the next six (6) special cases, $\bar{y}_{pss}(i)$, $i = 2, \cdots, 7$ are ratio-type estimators, while the remaining nine (9) special cases $\bar{y}_{pss}(i)$, $i = 8, \cdots, 16$ are examples of product-type estimators of $\bar{Y}$ in post-stratified sampling scheme.

4. Efficiency Comparisons

Using equation (2.18), we obtain the unconditional variance/mean squared errors of the estimators, $\bar{y}_{pss}(i)$, $i = 1, 2, \cdots, 16$ as follows:

$$V(\bar{y}_{pss}(1)) = V(\bar{y}_{ps}) = \left(\frac{1-f}{n}\right)\sum_{h=1}^{L}\omega_h S_{yh}^2 \tag{4.1}$$

$$MSE(\bar{y}_{pss}(2)) = MSE(\bar{y}_{psRC}) = \left(\frac{1-f}{n}\right)\sum_{h=1}^{L}\omega_h (S_{yh}^2 + R^2 S_{xh}^2 - 2RS_{yxh}) \tag{4.2}$$

$$MSE(\bar{y}_{pss}(3)) = MSE(\bar{y}_{psSD}) = \left(\frac{1-f}{n}\right)\sum_{h=1}^{L}\omega_h (S_{yh}^2 + \theta_1^2 S_{xh}^2 - 2\theta_1 S_{yxh}) \tag{4.3}$$

$$MSE(\bar{y}_{pss}(4)) = MSE(\bar{y}_{psSK1}) = \left(\frac{1-f}{n}\right)\sum_{h=1}^{L}\omega_h (S_{yh}^2 + \theta_2^2 S_{xh}^2 - 2\theta_2 S_{yxh}) \tag{4.4}$$

$$MSE(\bar{y}_{pss}(5)) = MSE(\bar{y}_{psUS1}) = \left(\frac{1-f}{n}\right)\sum_{h=1}^{L}\omega_h (S_{yh}^2 + \theta_3^2 S_{xh}^2 - 2\theta_3 S_{yxh}) \tag{4.5}$$

$$MSE(\bar{y}_{pss}(6)) = MSE(\bar{y}_{psUS2}) = \left(\frac{1-f}{n}\right)\sum_{h=1}^{L}\omega_h (S_{yh}^2 + \theta_4^2 S_{xh}^2 - 2\theta_4 S_{yxh}) \tag{4.6}$$

$$MSE(\bar{y}_{pss}(7)) = MSE(\bar{y}_{psST1}) = \left(\frac{1-f}{n}\right)\sum_{h=1}^{L}\omega_h (S_{yh}^2 + \theta_5^2 S_{xh}^2 - 2\theta_5 S_{yxh}) \tag{4.7}$$

$$MSE(\bar{y}_{pss}(8)) = MSE(\bar{y}_{psPC}) = \left(\frac{1-f}{n}\right)\sum_{h=1}^{L}\omega_h (S_{yh}^2 + R^2 S_{xh}^2 + 2RS_{yxh}) \tag{4.8}$$

$$MSE(\bar{y}_{pss}(9)) = MSE(\bar{y}_{psPD}) = \left(\frac{1-f}{n}\right)\sum_{h=1}^{L}\omega_h (S_{yh}^2 + \theta_1^2 S_{xh}^2 + 2\theta_1 S_{yxh}) \tag{4.9}$$

$$MSE(\bar{y}_{pss}(10)) = MSE(\bar{y}_{psUS3}) = \left(\frac{1-f}{n}\right)\sum_{h=1}^{L}\omega_h (S_{yh}^2 + \theta_3^2 S_{xh}^2 + 2\theta_3 S_{yxh}) \tag{4.10}$$

$$MSE(\bar{y}_{pss}(11)) = MSE(\bar{y}_{psUS4}) = \left(\frac{1-f}{n}\right)\sum_{h=1}^{L}\omega_h(S_{yh}^2 + \theta_4^2 S_{xh}^2 + 2\theta_4 S_{yxh}) \tag{4.11}$$

$$MSE(\bar{y}_{pss}(12)) = MSE(\bar{y}_{psGNS1}) = \left(\frac{1-f}{n}\right)\sum_{h=1}^{L}\omega_h(S_{yh}^2 + \theta_6^2 S_{xh}^2 + 2\theta_6 S_{yxh}) \tag{4.12}$$

$$MSE(\bar{y}_{pss}(13)) = MSE(\bar{y}_{psGNS2}) = \left(\frac{1-f}{n}\right)\sum_{h=1}^{L}\omega_h(S_{yh}^2 + \theta_7^2 S_{xh}^2 + 2\theta_7 S_{yxh}) \tag{4.13}$$

$$MSE(\bar{y}_{pss}(14)) = MSE(\bar{y}_{psGNS3}) = \left(\frac{1-f}{n}\right)\sum_{h=1}^{L}\omega_h(S_{yh}^2 + \theta_8^2 S_{xh}^2 + 2\theta_8 S_{yxh}) \tag{4.14}$$

$$MSE(\bar{y}_{pss}(15)) = MSE(\bar{y}_{psST2}) = \left(\frac{1-f}{n}\right)\sum_{h=1}^{L}\omega_h(S_{yh}^2 + \theta_5^2 S_{xh}^2 + 2\theta_5 S_{yxh}) \tag{4.15}$$

$$MSE(\bar{y}_{pss}(16)) = MSE(\bar{y}_{psSK2}) = \left(\frac{1-f}{n}\right)\sum_{h=1}^{L}\omega_h(S_{yh}^2 + \theta_2^2 S_{xh}^2 + 2\theta_2 S_{yxh}) \tag{4.16}$$

where

$$R = \frac{\bar{Y}}{\bar{X}}, \qquad \theta_1 = \frac{\bar{Y}}{\bar{X} + C_x}, \qquad \theta_2 = \frac{\bar{Y}}{\bar{X} + \beta_2(x)}, \qquad \theta_3 = \frac{\bar{Y}\beta_2(x)}{\bar{X}\beta_2(x) + C_x},$$

$$\theta_4 = \frac{\bar{Y}C_x}{\bar{X}C_x + \beta_2(x)},$$

and

$$\theta_5 = \frac{\bar{Y}}{\bar{X} + \rho_{yx}}, \quad \theta_6 = \frac{\bar{Y}}{\bar{X} + \sigma_x}, \quad \theta_7 = \frac{\bar{Y}\beta_1(x)}{\bar{X}\beta_1(x) + \sigma_x}, \quad \theta_8 = \frac{\bar{Y}\beta_2(x)}{\bar{X}\beta_2(x) + \sigma_x}$$

Applying the least squares method, the (optimum) choice of α that minimizes equation (2.18), is obtained as

$$\alpha_{opt} = \frac{\beta_0}{\lambda g R} \tag{4.17}$$

and the resulting optimum unconditional mean squared error of $\bar{y}_{pss}$ is obtained as

$$MSE_{opt}(\bar{y}_{pss}) = \left(\frac{1-f}{n}\right)(1-\rho_0^2)\sum_{h=1}^{L}\omega_h S_{yh}^2 \tag{4.18}$$

where

$$\beta_0 = \frac{\sum_{h=1}^{L} \omega_h S_{yxh}}{\sum_{h=1}^{L} \omega_h S_{xh}^2}, \quad \rho_0 = \frac{\sum_{h=1}^{L} \omega_h S_{yxh}}{\sqrt{\left(\sum_{h=1}^{L} \omega_h S_{yh}^2\right)\left(\sum_{h=1}^{L} \omega_h S_{xh}^2\right)}} \tag{4.19}$$

Notice that equation (4.18) is the same as the unconditional variance of the usual combined post-stratified regression estimator, $\bar{y}_{psREG} = \bar{y}_{ps} - \beta_0(\bar{x}_{ps} - \bar{X})$. This implies that the efficiency of the proposed general family of estimators may not be improved beyond the efficiency of the customary combined regression-type estimator in post-stratified sampling. However, using equations (4.1) and (4.18), we observe that:

$$V(\bar{y}_{ps}) - MSE_{opt}(\bar{y}_{pss}) = \left(\frac{1-f}{n}\right)\rho_0^2 \sum_{h=1}^{L} \omega_h S_{yh}^2 > 0 \tag{4.20}$$

This shows that under optimum conditions, the proposed family of estimators is more efficient than the usual post-stratified estimator, $\bar{y}_{ps}$ in terms of having a smaller mean squared error.

Again, let $A_0 = \sqrt{\sum_{h=1}^{L} \omega_h S_{yh}^2}$ and $A_1 = \sqrt{\sum_{h=1}^{L} \omega_h S_{xh}^2}$. Then, using equations (4.2) and (4.18), we observe that

$$MSE(\bar{y}_{pss}(2)) - MSE_{opt}(\bar{y}_{pss}) = \left(\frac{1-f}{n}\right)(\rho_0 A_0 - RA_1)^2 > 0 \tag{4.21}$$

This shows that the optimum estimator in the proposed family of estimators is more efficient than the usual combined ratio-type estimator $\bar{y}_{psRC}$. Similarly, it could be shown that none of the special cases of the proposed family of estimators is more efficient than the optimum estimator in the proposed general family of estimators.

5. Empirical Illustration

We have applied the proposed general family of estimators on the data on the academic performance of 96 students of Statistics department, Federal University of Technology, Owerri, 2008/2009 academic session. (Source: Department of Statistics, Federal University of Technology, Owerri, Nigeria). Here, we used the cumulative grade point average (CGPA) as the study variate, and performance in a general (pretest) Statistics examination as the auxiliary variate. Stratification was carried out by gender, and we assumed, hypothetically, that the number of male and female students to be included in the sample might not be determined until after sample selection. Consequently, we first took a random sample of size $n = 20$, which showed the distribution of male and female students, after sample selection, as 8 males and 12 females. The data statistics, consisting mainly of population parameters, are shown in Table 1, while Table 2 shows the percentage relative efficiencies (PRE) of the proposed estimators over the customary post-stratified estimator $\bar{y}_{ps}$ of the population mean, $\bar{Y}$ in post-stratified sampling scheme.

Table 1. Data Statistics

POPULATION	MALES = STRATUM 1	FEMALES = STRATUM 2
$N = 96$	$N_1 = 72$	$N_2 = 24$
$n = 20$	$n_1 = 8$	$n_2 = 12$
$\bar{X} = 68.13$	$\bar{X}_1 = 68.11$	$\bar{X}_2 = 68.17$
$\bar{Y} = 2.44$	$\bar{Y}_1 = 2.44$	$\bar{Y}_2 = 2.46$
$S_x = 7.03$	$S_{x1} = 7.28$	$S_{x2} = 6.36$
$S_x^2 = 49.37$	$S_{x1}^2 = 52.97$	$S_{x2}^2 = 40.41$
$S_y = 0.57$	$S_{y1} = 0.60$	$S_{y2} = 0.50$
$S_y^2 = 0.33$	$S_{y1}^2 = 0.35$	$S_{y2}^2 = 0.25$
$S_{yx} = 3.26$	$S_{yx1} = 3.43$	$S_{yx2} = 2.75$
$\rho_{yx} = 0.82$	$\rho_{yx1} = 0.80$	$\rho_{yx2} = 0.90$
$\rho_{yx}^2 = 0.67$	$\rho_{yx1}^2 = 0.64$	$\rho_{yx2}^2 = 0.80$
$C_x = 0.10$	$C_{x1} = 0.11$	$C_{x2} = 0.09$
$C_y = 0.23$	$C_{y1} = 0.24$	$C_{y2} = 0.20$
$\beta_1(x) = -1.10$	$\beta_{11}(x) = -1.23$	$\beta_{12}(x) = -0.50$
$\beta_1(y) = -0.11$	$\beta_{11}(y) = -0.14$	$\beta_{12}(y) = 0.14$
$\beta_2(x) = 3.83$	$\beta_{21}(x) = 4.33$	$\beta_{22}(x) = 1.34$
$\beta_2(y) = 1.27$	$\beta_{21}(y) = 1.40$	$\beta_{22}(y) = 0.31$
$\gamma = 0.04$	$\gamma_1 = 0.05$	$\gamma_2 = 0.16$
–	$\omega_1 = 0.75$	$\omega_2 = 0.25$
–	$\omega_1^2 = 0.56$	$\omega_2^2 = 0.06$

$R = 0.035814$, $\theta_1 = 0.035761$, $\theta_2 = 0.033908$,
$\theta_3 = 0.035800$, $\theta_4 = 0.022926$, $\theta_5 = 0.035388$,
$\theta_6 = 0.032464$, $\theta_7 = 0.039521$, $\theta_8 = 0.034874$,

Table 2. PRE of some post-stratified combined estimators of $\overline{Y}$ over $\overline{y}_{ps}$

Estimator	Variance / MSE	PRE over $\overline{y}_{st}$
$\overline{y}_{pss}(1)$	0.012864	100
$\overline{y}_{pss}(2)$	0.006151	209.14
$\overline{y}_{pss}(3)$	0.006158	208.90
$\overline{y}_{pss}(4)$	0.006381	201.60
$\overline{y}_{pss}(5)$	0.006153	209.07
$\overline{y}_{pss}(6)$	0.007984	161.12
$\overline{y}_{pss}(7)$	0.006202	207.42
$\overline{y}_{pss}(8)$	0.024637	52.21
$\overline{y}_{pss}(9)$	0.024616	52.26
$\overline{y}_{pss}(10)$	0.024632	52.22
$\overline{y}_{pss}(11)$	0.019818	64.91
$\overline{y}_{pss}(12)$	0.023322	55.16
$\overline{y}_{pss}(13)$	0.026145	49.20
$\overline{y}_{pss}(14)$	0.024264	53.02
$\overline{y}_{pss}(15)$	0.024468	52.57
$\overline{y}_{pss}(16)$	0.023883	53.86
$\overline{y}_{pss}(opt)$	0.004422	290.91

From Table 2, we observe that not every estimator in the proposed general family of estimators performed better than the usual post-stratified combined estimator $\overline{y}_{ps}$. The table also confirms that the optimum estimator $\overline{y}_{pss}(opt)$ is the most efficient estimator in the proposed family of combined estimators of $\overline{Y}$ in post-stratified sampling scheme. Again, we observe that for the given data set, the combined ratio-type estimators, $\overline{y}_{ps}(i)$, $i = 2, 3, \cdots, 7$ performed better than the customary post-stratified estimator $\overline{y}_{ps}$ in terms of having smaller mean squared errors, while the combined product-type estimators, $\overline{y}_{ps}(i)$, $i = 8, 9, \cdots, 16$ did not perform better than $\overline{y}_{ps}$. This is expected since for the given data set, there is a strong positive correlation (0.82) between the study and auxiliary variables. The product-type estimators would perform better than $\overline{y}_{ps}$ and the ratio-type estimators when there is a strong negative correlation between the study and auxiliary variables.

6. Concluding Remark

We have proposed a general family of combined estimators of $\bar{Y}$, in post-stratified sampling (PSS) scheme, which is found, under some optimum conditions, to be as efficient as the post-stratified regression estimator $\bar{y}_{psREG}$, but more efficient, in terms of having a smaller mean squared error, than the usual post-stratified combined estimator, $\bar{y}_{ps}$ Properties of the proposed general family of estimators are obtained up to first order approximations and illustrated empirically.

REFERENCES

CHAUDHARY, M.K., SINGH, R., SHUKLA, R.K., KUMAR, M. and SMARANDACHE, F. (2009): A family of estimators for estimating population mean in stratified sampling under non-response. Pak. J. Stat. Oper. Res. V(1), 47-54.

COCHRAN, W.G. (1977): Sampling Techniques, John Wiley and Sons, New York.

KADILAR, C. and CINGI, H. (2003): Ratio Estimators in Stratified Random Sampling. Biometrical Journal 45(2), 218-225.

KHOSHNEVISAN, M., SINGH, R., CHAUHAN, P., SAWAN, N., and SMARANDACHE, F. (2007). A general family of estimators for estimating population mean using known value of some population parameter(s), Far East Journal of Theoretical Statistics, 22, 181–191.

KOYUNCU, N. and KADILAR, C. (2009). Ratio and product estimators in stratified random sampling, Journal of Statistical Planning and Inference 139 (8), 2552-2558.

PANDY, B.N. and DUBEY, VYAS (1988): Modified product estimator using coefficient of variation of auxiliary variate, Assam Statistical Rev., 2(2), 64-66.

SEARLS, D.T. (1964): The utilization of known coefficient of variation in the estimation procedure. Journal of American Statistical Association, 59, 1125-1126.

SEARLS, D.T. and INTARAPANICH, P. (1990): A note on an estimator for the variance that utilizes the kurtosis. The American Statistician, 44(4), 295-296.

SEN, A.R. (1978): Estimation of the population mean when the coefficient of variation is known. Commun. Statist., Theory-Meth. A(7), 657-672.

SINGH, G.N. (2003): On the improvement of product method of estimation in sample surveys. Jour. Ind. Soc. Agric. Statistics, 56(3), 267-275.

SINGH, H.P. AND KAKRAN, M.S. (1993): A Modified Ratio Estimator Using Known Coefficients of Kurtosis of an Auxiliary Character (unpublished).

SINGH, H.P. AND TAILOR, R. (2003): Use of known correlation coefficient in estimating the finite population mean. Statistics in Transition, 6(4), 555-560.

SINGH, J., PANDEY, B.N. and HIRANO, K. (1973): On the utilization of a known coefficient of kurtosis in the estimation procedure of variance. Ann. Inst. Stat. Math., 25, 51-55.

SISODIA, B.V.S. and DWIVEDI, V.K. (1981): A Modified Ratio Estimator Using Coefficient of Variation of Auxiliary Variable. Journal of Indian Society Agricultural Statistics 33, 13-18.

STEPHAN, F. (1945): The expected value and variance of the reciprocal and other negative powers of a positive Bernoulli variate. Ann. Math. Stat., 16, 50-61.

UPADHYAYA, L.N. and SINGH, H.P. (1999): Use of Transformed Auxiliary Variable in Estimating the Finite Population Mean. Biometrical Journal 41(5), 627-636.