Generalized Ratio-Product Estimator under Two-Phase Stratified Sampling in the Presence of Non-Response Using Two Auxiliary Variables

A. E. Anieting

Department of Mathematics and Statistics, University of Uyo, Nigeria

Abstract: **In this article, a new generalized ratio-product estimator under two-phase stratified Sampling in the presence of nonresponse using two auxiliary variables has been developed. The mean squared error of the estimator formulated using large sample approximation. The estimator is efficient based on the empirical study done and better than other estimators in its class.**

Keywords: **Stratified Sampling, Nonresponse, Auxiliary Variables**

I. INTRODUCTION

Classical two-phase stratified sampling is a method that uses two random samples where the second sample is a stratified subsample of the first sample. If the problem of nonresponse is there then this subsample may be divided into classes of respondents and non-respondents Cochran (1977). Information on characters correlated with the main character under study is referred to as auxiliary information which can be used at either planning stage or design stage to obtain an improved estimator. It is also known to enhance efficiency in population parameter estimates. The presence of nonresponse in a survey distorts parameter estimation by increasing bias in estimates resulting in large variance or mean square error. Hansen and Hurwitz (1946) proposed a technique of subsampling from nonresponse in dealing with the problem of nonresponse. Upadhyaya and Singh (1999),Kadilar andCingi(2003),Khare and Sinha(2004), shibbir and Gupta(2005), Khoshnevisan et al. (2007), Chaudhary et al. (2009), Grover and kaur(2011), Chaudhary et al. (2012), Sanuallah et al.(2014,2015,2018), Chaudhary and Kumar (2015) and Shabbir et al. (2018) have all suggested estimators in stratified double sampling under nonresponse using either single or two auxiliary variables. In this article, a generalized ratio-product estimator for population mean in two-phase stratified sampling in the presence of nonresponse using two-auxiliary variables has been proposed. A real data set is used to access the performance of the estimator.

II. SAMPLING DESIGN

Consider a finite population of N units which is stratified into K homogenous strata. Let the size of each stratum be N_i(i=1, 2,,,,K) such that $\sum_i^k N_i = N$. Let y_{ij} and (x_{ij} ,z_{ij}) be values of the study variable and the auxiliary variables respectively on the j^{th} unit in the i^{th} stratum. Let $\bar{y}_i = \frac{\sum_j^{n_i} y_{ij}}{n_i}$, $\bar{x}_i = \frac{\sum_j^{n_i} x_{ij}}{n_i}$ and $\bar{z}_i = \frac{\sum_j^{n_i} z_{ij}}{n_i}$ be the sample means corresponding to population means $\bar{Y}_i = \frac{\sum_j^{N_i} y_{ij}}{N_i}$, $\bar{X}_i = \frac{\sum_j^{N_i} x_{ij}}{N_i}$ and $\bar{Z}_i = \frac{\sum_j^{N_i} z_{ij}}{N_i}$ respectively in thei^{th} stratum. When both $\bar{X}_i$ and $\bar{Z}_i$ are unknown, it is desirable to take a preliminary large sample of size $n_i^{'}$ (<N_i) from the i^{th} stratum to estimate $\bar{X}_i$ and $\bar{Z}_i$ in the first phase sample such that $\sum_i^k n_i^{'} = n^{'}$. In the second phase, a subsample of size $n^{'}(< n_i^{'})$ is drawn using (SRSWOR) scheme such that $\sum_i^k n_i$= n and information is collected on y_i, x_i and z_i. Let $\bar{y}_{st} = \sum_i^k P_i \bar{y}_i$, $\bar{x}_{st} = \sum_i^k P_i \bar{x}_i$, and$\bar{z}_{st} = \sum_i^k P_i \bar{z}_i$ be the sample means based on second phase sampling and $\bar{x}_{st}^{'} = \sum_i^k P_i \bar{x}_i^{'}$ and $\bar{z}_{st}^{'} = \sum_i^k P_i \bar{z}_i^{'}$ be the sample means based on the first phase sampling where $\bar{y}_i$, $\bar{x}_i$ and $\bar{z}_i$ are sample means of y, x and z respectively in the i^{th} stratum and $P_i = \frac{N_i}{N}$ are the known stratum weights. It is assumed that in the first phase a sample of size$n_i^{'}$ units are selected from the i^{th} stratum by using SRSWOR and the auxiliary variables are observed. At second phase, a subsample of size n_i units are selected and observations are made on both the study and the auxiliary variables. For the second phase sample of size n_i, it is assumed that n_{i1}units lead to responses and n_{i2} units lead to non-responses. Using Hansen and Hurwitz (1946) technique, a subsample of size $r_{i2} = \frac{n_{i2}}{k_{i2}}$ $(k_{i2} > 1)$ units is selected randomly and a response is obtained by interview.

III. SOME EXISTING ESTIMATORS

3.1. Rao (1991) Estimator

Rao (1991) suggested a difference type estimator as

$$\bar{y}_{P_{14(st)}} = d_{14}\bar{y}_{st}^{*} + d_{15}(\bar{x}_{st}^{'} - \bar{x}_{st}^{*}) + d_{16}(\bar{z}_{st}^{'} - \bar{z}_{st}^{*}) \quad (1)$$

Where d_{14}, d_{15} and $d_{16}$are constants

$$MSE\ (\bar{y}_{P_{14(st)}}) = (d_{14} - 1)^2\bar{Y}^2 + d_{14}^2\bar{Y}^2 A_0 + d_{15}^2\bar{X}^2 B_0 + d_{16}^2\bar{Z}^2 C_0 - 2d_{14}d_{15}\bar{X}\bar{Y}D_0 - 2d_{14}\bar{Y}\bar{Z}E_0 + 2d_{15}d_{16}\bar{X}\bar{Z}F_0 \quad (2)$$

Where $A_0 = \frac{1}{\bar{Y}^2}\sum_i^k P_i^2[\left(\frac{1}{n_i} - \frac{1}{N_i}\right)S_{y_i}^2 + \frac{(K_{i2}-1)W_{i2}S_{y_{i2}}^2}{n_i}]$

$B_0 = \frac{1}{\bar{X}^2}\sum_i^k P_i^2 [\left(\frac{1}{n_i}-\frac{1}{N_i}\right) S_{x_i}^2 + \frac{(K_{i2}-1)W_{i2}S_{x_{i2}}^2}{n_i}]$

$C_0 = \frac{1}{\bar{Z}^2}\sum_i^k P_i^2 [\left(\frac{1}{n_i}-\frac{1}{N_i}\right) S_{z_i}^2 + \frac{(K_{i2}-1)W_{i2}S_{z_{i2}}^2}{n_i}]$

$D_0 = \frac{1}{\bar{Y}\bar{X}}\sum_i^k P_i^2 [\left(\frac{1}{n_i}-\frac{1}{N_i}\right) S_{yx_i} + \frac{(K_{i2}-1)W_{i2}S_{yx_{i2}}}{n_i}]$

$E_0 = \frac{1}{\bar{Y}\bar{Z}}\sum_i^k P_i^2 [\left(\frac{1}{n_i}-\frac{1}{N_i}\right) S_{yz_i} + \frac{(K_{i2}-1)W_{i2}S_{yz_{i2}}}{n_i}]$

$F_0 = \frac{1}{\bar{Z}\bar{X}}\sum_i^k P_i^2 [\left(\frac{1}{n_i}-\frac{1}{N_i}\right) S_{xz_i} + \frac{(K_{i2}-1)W_{i2}S_{xz_{i2}}}{n_i}]$

3.2. Sanaullah et al (2015) Estimators

Sanaullah et al (2015) proposed a generalized exponential-type ratio-cum-ratio estimator of population mean as

$$\hat{\bar{Y}}_{RR}^{G} = \sum_h^L P_h \bar{y}_h^* \exp \frac{\sum_h^L P_h(\bar{X}_h-\bar{x}_h')}{\sum_h^L P_h(\bar{X}_h+(a-1)\bar{x}_h')} \exp \frac{\sum_h^L P_h(\bar{Z}_h-\bar{z}_h^*)}{\sum_h^L P_h(\bar{Z}_h+(b-1)\bar{z}_h^*)} \quad (3)$$

Where a and b are some suitably chosen scalars

The MSE ($\hat{\bar{Y}}_{RR}^{G}$) =

$$\bar{Y}^2\left[V_{020}^* + \frac{1}{a^2}V_{200}' + \frac{1}{b^2}V_{002}^* - 2\left(\frac{1}{a}V_{110}' + \frac{1}{b}V_{011}^* - \frac{1}{ab}V_{101}'\right)\right] \quad (4)$$

Where

$V_{020}^* = \frac{1}{\bar{Y}^2}\sum_i^k P_i^2 [\left(\frac{1}{n_i}-\frac{1}{n_i'}\right) S_{y_i}^2 + \frac{(K_{i2}-1)W_{i2}S_{y_{i2}}^2}{n_i}]$

$V_{200}' = \frac{1}{\bar{X}^2}\sum_i^k P_i^2 [\left(\frac{1}{n_i}-\frac{1}{N_i}\right) S_{x_i}^2$

$V_{002}^* = \frac{1}{\bar{Z}^2}\sum_i^k P_i^2 [\left(\frac{1}{n_i}-\frac{1}{n_i'}\right) S_{z_i}^2 + \frac{(K_{i2}-1)W_{i2}S_{z_{i2}}^2}{n_i}]$

$V_{110}' = \frac{1}{\bar{Y}\bar{X}}\sum_i^k P_i^2 [\left(\frac{1}{n_i}-\frac{1}{N_i}\right) S_{yx_i}$

$V_{011}^* = \frac{1}{\bar{Y}\bar{Z}}\sum_i^k P_i^2 [\left(\frac{1}{n_i}-\frac{1}{n_i'}\right) S_{yz_i} + \frac{(K_{i2}-1)W_{i2}S_{yz_{i2}}}{n_i}]$

$V_{101}' = \frac{1}{\bar{Z}\bar{X}}\sum_i^k P_i^2 [\left(\frac{1}{n_i}-\frac{1}{N_i}\right) S_{xz_i}$

Sanaullah et al (2015) also proposed a generalized exponential-type product-cum-product estimator of population mean as

$$\hat{\bar{Y}}_{PP}^{G} = \sum_h^L P_h \bar{y}_h^* \exp \frac{\sum_h^L P_h(\bar{x}_h'-\bar{X}_h)}{\sum_h^L P_h(\bar{x}_h'+(c-1)\bar{X}_h)} \exp \frac{\sum_h^L P_h(\bar{z}_h^*-\bar{Z}_h)}{\sum_h^L P_h(\bar{z}_h^*+(d-1)\bar{Z}_h)} \quad (5)$$

Where c and d are some suitably chosen scalars

The MSE ($\hat{\bar{Y}}_{PP}^{G}$) =

$$\bar{Y}^2\left[V_{020}^* + \frac{1}{c^2}V_{200}' + \frac{1}{d^2}V_{002}^* + 2\left(\frac{1}{c}V_{110}' + \frac{1}{d}V_{011}^* + \frac{1}{cd}V_{101}'\right)\right] \quad (6)$$

3.3. Shabbir et al (2018) Estimator

Shabbir et al (2018) proposed a generalized class of estimators when non response exists on the study variable as well as on the two auxiliary variables under two-phase stratified sampling in estimating the finite population mean as

$$\bar{y}_{p2(st)} = [k_1\bar{y}_{st}^* + k_2(\bar{x}_{st}' - \bar{x}_{st}^*) + k_3(\bar{z}_{st}' - \bar{z}_{st}^*)] \left[\left\{\exp\left(\frac{\bar{x}_{st}'-\bar{x}_{st}^*}{\bar{x}_{st}'+\bar{x}_{st}^*}\right)\right\}^{\lambda_1}\left(\frac{\bar{x}_{st}'}{\bar{x}_{st}^*}\right)^{\lambda_2}\right] \quad (7)$$

The MSE ($\bar{y}_{p2(st)}$) = $(k_1-1)^2\bar{Y}^2 + k_1^2\bar{Y}^2\text{A} + k_2^2\bar{X}^2B + k_3^2\bar{Z}^2C - 2k_1\bar{Y}^2D - 2k_2\bar{Y}\bar{X}E - 2k_3\bar{Y}\bar{Z}F - 2k_1k_2\bar{Y}\bar{X}G - 2k_1k_2\bar{Y}\bar{Z}H + 2k_2k_3\bar{X}\bar{Z}I$ (8)

Where $A = \frac{1}{\bar{Y}^2}\sum_i^k P_i^2 [\left(\frac{1}{n_i}-\frac{1}{N_i}\right) S_{y_i}^2 + \frac{(K_{i2}-1)W_{i2}S_{y_{i2}}^2}{n_i}] + \left\{\frac{\lambda_1}{2}(1+\lambda_1) + \lambda_2(1+\lambda_1+2\lambda_2)\right\}\frac{1}{\bar{X}^2}\sum_i^k P_i^2 [\left(\frac{1}{n_i}-\frac{1}{N_i}\right) S_{x_i}^2 + \frac{(K_{i2}-1)W_{i2}S_{x_{i2}}^2}{n_i}] - 4\left(\frac{\lambda_1}{2}+\lambda_2\right)\frac{1}{\bar{Y}\bar{X}}\sum_i^k P_i^2 [\left(\frac{1}{n_i}-\frac{1}{n_i'}\right) S_{yx_i} + \frac{(K_{i2}-1)W_{i2}S_{yx_{i2}}}{n_i}]$

$B = \frac{1}{\bar{X}^2}\sum_i^k P_i^2 [\left(\frac{1}{n_i}-\frac{1}{N_i}\right) S_{x_i}^2 + \frac{(K_{i2}-1)W_{i2}S_{x_{i2}}^2}{n_i}]$

$C = \frac{1}{\bar{Z}^2}\sum_i^k P_i^2 [\left(\frac{1}{n_i}-\frac{1}{N_i}\right) S_{z_i}^2 + \frac{(K_{i2}-1)W_{i2}S_{z_{i2}}^2}{n_i}]$

$D = \{\frac{\lambda_1}{4}\left(1+\frac{\lambda_1}{2}\right) + \frac{\lambda_2}{4}(3+\lambda_1)\}\frac{1}{\bar{X}^2}\sum_i^k P_i^2 [\left(\frac{1}{n_i}-\frac{1}{N_i}\right) S_{x_i}^2 + \frac{(K_{i2}-1)W_{i2}S_{x_{i2}}^2}{n_i}] - \left(\frac{\lambda_1}{2}+\lambda_2\right)\frac{1}{\bar{Y}\bar{X}}\sum_i^k P_i^2 [\left(\frac{1}{n_i}-\frac{1}{n_i'}\right) S_{yx_i} + \frac{(K_{i2}-1)W_{i2}S_{yx_{i2}}}{n_i}]$

$E = \left(\frac{\lambda_1}{2}+\lambda_2\right)\frac{1}{\bar{X}^2}\sum_i^k P_i^2 [\left(\frac{1}{n_i}-\frac{1}{N_i}\right) S_{x_i}^2 + \frac{(K_{i2}-1)W_{i2}S_{x_{i2}}^2}{n_i}]$

$F = \left(\frac{\lambda_1}{2}+\lambda_2\right)\frac{1}{\bar{Z}\bar{X}}\sum_i^k P_i^2 [\left(\frac{1}{n_i}-\frac{1}{n_i'}\right) S_{xz_i} + \frac{(K_{i2}-1)W_{i2}S_{xz_{i2}}}{n_i}]$

IV. THE PROPOSED ESTIMATOR

The proposed estimator for population mean in stratified double sampling in the presence of non response is given as

$$T_a = \bar{y}_{st}^*\left(\frac{\bar{x}_{st}'+\varphi}{\bar{x}_{st}-\varphi}\right)^{\alpha}\left(\frac{\bar{z}_{st}'+\varphi}{\bar{z}_{st}-\varphi}\right)^{\beta}$$

Where $\varphi = \sum_i^k \rho_i$ and α *and* β are constants. $\bar{y}_{st}^* = \sum_i^k p_i\bar{y}_i^*$; $\bar{y}_i^* = \frac{n_{i1}\bar{y}_{n_{i1}} + n_{i2}\bar{y}_{h_{i2}}}{n_i}$.

$\bar{y}_{n_{i1}}$ and $\bar{y}_{h_{i2}}$ are the means based on n_{i1} respondent units and h_{i2} subsampled non-respondent units respectively for the study variable and $\bar{x}_{st}' = \sum_i^k p_i\bar{x}_i'$, $\bar{x}_{st} = \sum_i^k p_i\bar{x}_i$, $\bar{z}_{st}' = \sum_i^k p_i\bar{z}_i'$, $\bar{z}_{st} = \sum_i^k p_i\bar{z}_i$

Using the following symbols

$\bar{y}^*_{st} = \bar{Y}(1+e_0)$, $\bar{x}_{st} = \bar{X}(1+e_1)$, $\bar{x}'_{st} = \bar{X}(1+e_2)$,$\bar{z}_{st} = \bar{Z}(1+e_1)$, $\bar{z}'_{st} = \bar{Z}(1+e_2)$ where the e_is are the relative error terms defined as

$e_0 = \frac{\bar{y}^*_{st}-\bar{Y}}{\bar{Y}}$, $e_1 = \frac{\bar{x}_{st}-\bar{X}}{\bar{X}}$, $e_2 = \frac{\bar{x}'_{st}-\bar{X}}{\bar{X}}$, $e_3 = \frac{\bar{z}_{st}-\bar{Z}}{\bar{Z}}$, $e_4 = \frac{\bar{z}'_{st}-\bar{X}}{\bar{X}}$

Such that the following expectations are applied

E (e_0) = (e_1) = (e_2) = 0

E (e_0^2) = $\frac{1}{\bar{Y}^2}\sum_i^k \left[\left(\frac{1}{n_i}-\frac{1}{N_i}\right)p_i^2 S_{Y_i}^2 + \frac{(k_{i2}-1)}{n_i} p_i^2 W_{i2} S_{Y_{i2}}^2\right]$

E (e_1^2) = $\frac{1}{\bar{X}^2}\sum_i^k \left[\left(\frac{1}{n_i}-\frac{1}{N_i}\right)p_i^2 S_{X_i}^2\right]$

E (e_2^2) = $\frac{1}{\bar{X}^2}\sum_i^k \left[\left(\frac{1}{n_i'}-\frac{1}{N_i}\right)p_i^2 S_{X_i}^2\right]$ = E ($e_1 e_2$)

E (e_3^2) = $\frac{1}{\bar{Z}^2}\sum_i^k \left[\left(\frac{1}{n_i}-\frac{1}{N_i}\right)p_i^2 S_{Z_i}^2\right]$

E (e_4^2) = $\frac{1}{\bar{Z}^2}\sum_i^k \left[\left(\frac{1}{n_i'}-\frac{1}{N_i}\right)p_i^2 S_{Z_i}^2\right]$ = E ($e_3 e_4$)

E ($e_0 e_1$) = $\frac{1}{\bar{X}\bar{Y}}\sum_i^k \left(\frac{1}{n_i}-\frac{1}{N_i}\right)p_i^2 S_{yxi}$

E ($e_0 e_2$) = $\frac{1}{\bar{X}\bar{Y}}\sum_i^k \left(\frac{1}{n_i'}-\frac{1}{N_i}\right)p_i^2\, S_{yxi}$

E ($e_0 e_3$) = $\frac{1}{\bar{Z}\bar{Y}}\sum_i^k \left(\frac{1}{n_i}-\frac{1}{N_i}\right)p_i^2\, S_{yzi}$

E ($e_0 e_4$) = $\frac{1}{\bar{Z}\bar{Y}}\sum_i^k \left(\frac{1}{n_i'}-\frac{1}{N_i}\right)p_i^2\, S_{yzi}$

E ($e_1 e_3$) = $\frac{1}{\bar{Z}\bar{X}}\sum_i^k \left(\frac{1}{n_i}-\frac{1}{N_i}\right)p_i^2\, S_{xzi}$

E ($e_1 e_4$) = $\frac{1}{\bar{Z}\bar{X}}\sum_i^k \left(\frac{1}{n_i'}-\frac{1}{N_i}\right)p_i^2\, S_{xzi}$ = E ($e_2 e_4$) = E ($e_2 e_3$)

Using large sample approximation to obtain the mean squared error(MSE)up to the terms of order (1/n) of the proposed estimator, hence, the MSE is given as

MSE (T_a)

$$=\bar{Y}^2\left\{\beta^2 d^2 \frac{1}{\bar{Z}^2}\sum_i^k\left[\left(\frac{1}{n_i}-\frac{1}{N_i}\right)p_i^2 S_{Z_i}^2\right] - 2\beta^2 cd\frac{1}{\bar{Z}^2}\sum_i^k\left[\left(\frac{1}{n_i'}-\frac{1}{N_i}\right)p_i^2 S_{Z_i}^2\right] + 2\alpha\beta bd\frac{1}{\bar{Z}\bar{X}}\sum_i^k\left(\frac{1}{n_i}-\frac{1}{N_i}\right)p_i^2\, S_{xzi} - 2\alpha\beta ad\frac{1}{\bar{Z}\bar{X}}\sum_i^k\left(\frac{1}{n_i'}-\frac{1}{N_i}\right)p_i^2\, S_{xzi} - 2\beta d\frac{1}{\bar{Z}\bar{Y}}\sum_i^k\left(\frac{1}{n_i}-\frac{1}{N_i}\right)p_i^2\, S_{yzi} + \beta^2 c^2\frac{1}{\bar{Z}^2}\sum_i^k\left[\left(\frac{1}{n_i'}-\frac{1}{N_i}\right)p_i^2 S_{Z_i}^2\right] - 2\alpha\beta cb\frac{1}{\bar{Z}\bar{X}}\sum_i^k\left(\frac{1}{n_i'}-\frac{1}{N_i}\right)p_i^2\, S_{xzi} + 2\alpha\beta ca\frac{1}{\bar{Z}\bar{X}}\sum_i^k\left(\frac{1}{n_i'}-\frac{1}{N_i}\right)p_i^2\, S_{xzi} + 2\beta c\frac{1}{\bar{Z}\bar{Y}}\sum_i^k\left(\frac{1}{n_i'}-\frac{1}{N_i}\right)p_i^2\, S_{yzi} + \alpha^2 b^2\frac{1}{\bar{X}^2}\sum_i^k\left[\left(\frac{1}{n_i}-\frac{1}{N_i}\right)p_i^2 S_{X_i}^2\right] - 2\alpha^2 ab\frac{1}{\bar{X}^2}\sum_i^k\left[\left(\frac{1}{n_i'}-\frac{1}{N_i}\right)p_i^2 S_{X_i}^2\right] - 2\alpha b\frac{1}{\bar{X}\bar{Y}}\sum_i^k\left(\frac{1}{n_i}-\frac{1}{N_i}\right)p_i^2 S_{yxi} + \alpha^2 a^2\frac{1}{\bar{X}^2}\sum_i^k\left[\left(\frac{1}{n_i'}-\frac{1}{N_i}\right)p_i^2 S_{X_i}^2\right] + 2\alpha a\frac{1}{\bar{X}\bar{Y}}\sum_i^k\left(\frac{1}{n_i'}-\frac{1}{N_i}\right)p_i^2\, S_{yxi} + \frac{1}{\bar{Y}^2}\sum_i^k\left[\left(\frac{1}{n_i}-\frac{1}{N_i}\right)p_i^2 S_{Y_i}^2 + \frac{(k_{i2}-1)}{n_i}p_i^2 W_{i2} S_{Y_{i2}}^2\right]\right\}$$

Where a = $\frac{\bar{X}}{\bar{X}+\varphi}$, b = $\frac{\bar{X}}{\bar{X}-\varphi}$, c = $\frac{\bar{Z}}{\bar{Z}+\varphi}$, d = $\frac{\bar{Z}}{\bar{Z}-\varphi}$ and a,b,c,d are approximately 1 using population 1, hence

MSE (T_a) =

$$\bar{Y}^2\left\{\beta^2\frac{1}{\bar{Z}^2}\sum_i^k\left[\left(\frac{1}{n_i}-\frac{1}{N_i}\right)p_i^2 S_{Z_i}^2\right] - \beta^2\frac{1}{\bar{Z}^2}\sum_i^k\left[\left(\frac{1}{n_i'}-\frac{1}{N_i}\right)p_i^2 S_{Z_i}^2\right] + 2\alpha\beta\frac{1}{\bar{Z}\bar{X}}\sum_i^k\left(\frac{1}{n_i}-\frac{1}{N_i}\right)p_i^2\, S_{xzi} - 2\alpha\beta\frac{1}{\bar{Z}\bar{X}}\sum_i^k\left(\frac{1}{n_i'}-\frac{1}{N_i}\right)p_i^2\, S_{xzi} - 2\beta\frac{1}{\bar{Z}\bar{Y}}\sum_i^k\left(\frac{1}{n_i}-\frac{1}{N_i}\right)p_i^2\, S_{yzi} + 2\beta\frac{1}{\bar{Z}\bar{Y}}\sum_i^k\left(\frac{1}{n_i'}-\frac{1}{N_i}\right)p_i^2\, S_{yzi} + \alpha^2\frac{1}{\bar{X}^2}\sum_i^k\left[\left(\frac{1}{n_i}-\frac{1}{N_i}\right)p_i^2 S_{X_i}^2\right] - \alpha^2\frac{1}{\bar{X}^2}\sum_i^k\left[\left(\frac{1}{n_i'}-\frac{1}{N_i}\right)p_i^2 S_{X_i}^2\right] - 2\alpha b\frac{1}{\bar{X}\bar{Y}}\sum_i^k\left(\frac{1}{n_i}-\frac{1}{N_i}\right)p_i^2 S_{yxi} + 2\alpha a\frac{1}{\bar{X}\bar{Y}}\sum_i^k\left(\frac{1}{n_i'}-\frac{1}{N_i}\right)p_i^2\, S_{yxi} + \frac{1}{\bar{Y}^2}\sum_i^k\left[\left(\frac{1}{n_i}-\frac{1}{N_i}\right)p_i^2 S_{Y_i}^2 + \frac{(k_{i2}-1)}{n_i}p_i^2 W_{i2} S_{Y_{i2}}^2\right]\right\}$$

To obtain the optimum value of MSE (T_a), MSE (T_a) is differentiated with respect to α and β and equated to zero i. e.

$$\frac{\partial \text{MSE } (T_a)}{\partial \alpha} = 0 \qquad (9)$$

$$, \frac{\partial \text{MSE } (T_a)}{\partial \beta} = 0 \qquad (10)$$

Solving (9) and (10) simultaneously gives the optimum values of α and β to be

$$\alpha_{opt} = \frac{\frac{1}{\bar{Y}\bar{X}}\sum_i^k\left(\frac{1}{n_i}-\frac{1}{n_i'}\right)p_i^2\,S_{yxi}\,\frac{1}{\bar{Z}^2}\sum_i^k\left[\left(\frac{1}{n_i}-\frac{1}{n_i'}\right)p_i^2 S_{Z_i}^2\right]-\frac{1}{\bar{Y}\bar{Z}}\sum_i^k\left(\frac{1}{n_i}-\frac{1}{n_i'}\right)p_i^2\,S_{yzi}\,\frac{1}{\bar{Z}\bar{X}}\sum_i^k\left(\frac{1}{n_i}-\frac{1}{n_i'}\right)p_i^2\,S_{xzi}}{\frac{1}{\bar{X}^2}\sum_i^k\left[\left(\frac{1}{n_i}-\frac{1}{n_i'}\right)p_i^2 S_{X_i}^2\right]\frac{1}{\bar{Z}^2}\sum_i^k\left[\left(\frac{1}{n_i}-\frac{1}{n_i'}\right)p_i^2 S_{Z_i}^2\right]-\left[\frac{1}{\bar{Z}\bar{X}}\sum_i^k\left(\frac{1}{n_i}-\frac{1}{n_i'}\right)p_i^2\,S_{xzi}\right]^2}$$

$$\beta_{opt} = \frac{\frac{1}{\bar{Y}\bar{Z}}\sum_i^k\left(\frac{1}{n_i}-\frac{1}{n_i'}\right)p_i^2\,S_{yzi}\,\frac{1}{\bar{X}^2}\sum_i^k\left[\left(\frac{1}{n_i}-\frac{1}{n_i'}\right)p_i^2 S_{X_i}^2\right]-\frac{1}{\bar{Y}\bar{X}}\sum_i^k\left(\frac{1}{n_i}-\frac{1}{n_i'}\right)p_i^2\,S_{yxi}\,\frac{1}{\bar{Z}\bar{X}}\sum_i^k\left(\frac{1}{n_i}-\frac{1}{n_i'}\right)p_i^2\,S_{xzi}}{\frac{1}{\bar{X}^2}\sum_i^k\left[\left(\frac{1}{n_i}-\frac{1}{n_i'}\right)p_i^2 S_{X_i}^2\right]\frac{1}{\bar{Z}^2}\sum_i^k\left[\left(\frac{1}{n_i}-\frac{1}{n_i'}\right)p_i^2 S_{Z_i}^2\right]-\left[\frac{1}{\bar{Z}\bar{X}}\sum_i^k\left(\frac{1}{n_i}-\frac{1}{n_i'}\right)p_i^2\,S_{xzi}\right]^2}$$

Hence the optimum value of MSE (T_a) is given as

$$\begin{aligned}\text{MSE}\,(T_a)_{opt} = {} & \bar{Y}^2\Big\{\beta_{opt}^2\frac{1}{\bar{Z}^2}\sum_i^k\left[\left(\frac{1}{n_i}-\frac{1}{N_i}\right)p_i^2S_{Z_i}^2\right]-\beta_{opt}^2\frac{1}{\bar{Z}^2}\sum_i^k\left[\left(\frac{1}{n_i'}-\frac{1}{N_i}\right)p_i^2S_{Z_i}^2\right]+2\alpha_{opt}\beta_{opt}\frac{1}{\bar{Z}\bar{X}}\sum_i^k\left(\frac{1}{n_i}-\frac{1}{N_i}\right)p_i^2\,S_{xzi} \\ & -2\alpha_{opt}\beta_{opt}\frac{1}{\bar{Z}\bar{X}}\sum_i^k\left(\frac{1}{n_i'}-\frac{1}{N_i}\right)p_i^2\,S_{xzi}-2\beta_{opt}\frac{1}{\bar{Z}\bar{Y}}\sum_i^k\left(\frac{1}{n_i}-\frac{1}{N_i}\right)p_i^2\,S_{yzi}+2\beta_{opt}\frac{1}{\bar{Z}\bar{Y}}\sum_i^k\left(\frac{1}{n_i'}-\frac{1}{N_i}\right)p_i^2\,S_{yzi} \\ & +\alpha_{opt}^2\frac{1}{\bar{X}^2}\sum_i^k\left[\left(\frac{1}{n_i}-\frac{1}{N_i}\right)p_i^2S_{X_i}^2\right]-\alpha_{opt}^2\frac{1}{\bar{X}^2}\sum_i^k\left[\left(\frac{1}{n_i'}-\frac{1}{N_i}\right)p_i^2S_{X_i}^2\right]-2\alpha_{opt}\frac{1}{\bar{X}\bar{Y}}\sum_i^k\left(\frac{1}{n_i}-\frac{1}{N_i}\right)p_i^2S_{yxi} \\ & +2\alpha_{opt}\frac{1}{\bar{X}\bar{Y}}\sum_i^k\left(\frac{1}{n_i'}-\frac{1}{N_i}\right)p_i^2\,S_{yxi}+\frac{1}{\bar{Y}^2}\sum_i^k\left[\left(\frac{1}{n_i}-\frac{1}{N_i}\right)p_i^2S_{Y_i}^2+\frac{(k_{i2}-1)}{n_i}p_i^2W_{i2}S_{Y_{i2}}^2\right]\Big\}\end{aligned}$$

V. EMPIRICAL STUDY

To study the performance of the proposed estimator, secondary data set (called population 1) used by Sanuallah et al. (2015) is employed. The situation II of selection here is that only the study variable undergoes nonresponse at the second phase.

Table 1

Percent Relative Efficiencies (PREs) of Estimators with respect to $\bar{y}_{st}^*$ for Different Values of k_{i2} each at a particular Rate of Non-Response under Situation II using Population 1

W_{i2}	k_{i2}	$\bar{y}_{st}^*$	T_a	$\bar{\bar{Y}}_{RR}^G$	$\bar{\bar{Y}}_{cr}^G$
10%	2	100	1794	493.14	205.02
	2.5	100	763.8	377.77	186.29
	3	100	518.1	314.75	173.24
	3.5	100	407.9	275.03	163.61

VI. CONCLUSION

Based on table 1 values, the proposed estimator is more efficient than other estimators in the same situation, hence T_a is recommended for use in practice in the situation explained therein.

REFERENCES

[1]. M. H. Hansen, W. N. Hurwitz (1946).The problem of non-response in sample surveys. Journal of The American Statistical Association, , no. 41, pp. 517–529

[2]. Cochran , W. G. (1977). Sampling Techniques. New York: Wiley and Sons

[3]. L. N. Upadhyaya, H. P. Singh (1999). Use of transformed auxiliary variable in estimating the finite population mean. Biometrical Journal, 41, no. 5, pp. 627–636..

[4]. C. Kadilar, H. Cingi (2003). Ratio estimators in stratified random sampling.Biometrical journal, 2, no. 45, pp. 218–225.

[5]. B. B. Khare, R. R. Sinha (2004). Estimation of finite population ratio using two-phase sampling scheme in the presence of non-response. Aligarh J. Stat., 24, 43-56., 24, pp. 43–56.

[6]. J. Shabbir, S. Gupta (2005). Improved ratio estimators in stratified sampling. American Journal of Mathematical and Management Sciences, 25, no. 3-4, pp. 293–311.

[7]. M. Khoshnevisan, R. Singh, P. Chauhan, N. Sawanand, F. Smarandache (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, pp. 181–191.

[8]. M. K. Chaudhary, R. Singh, R. K. Shukla, M. Kumar (2009). A family of estimators for estimating population mean in stratified sampling under non- response. Pakistan Journal of Statistics and Operation Research, 5, no. 1, pp. 47–54.

[9]. Grover, L. K., and P. Kaur.(2011). An improved estimator of the finite population mean in simple random sampling. Model Assisted Statistics and Application 6 (1):47–55.

[10]. M. K. Chaudhary, V. K. Singh, R. K. Shukla (2012). Combined-type fam- ily of estimators of population mean in stratified random sampling under non-response. Journal of Reliability and Statistical Studies (JRSS), 5, no. 2, pp. 133–142.

[11]. M. K. Chaudhary and A. Kumar(2015) Use of double sample scheme in estimating the mean of

[12]. Stratified population Under nonresponse. STATISTICA, anno, Lxxv,n.4

[13]. Sanaullah, A., H. M. Ali, M. Noor-ul-Amin, and M. Hanif.(2014). Generalized exponential chain ratio estimators under stratified two-phase random sampling. Applied Mathematics and Computation 226:541–47.

[14]. Sanaullah, A., M. Noor-ul-Amin, M. Hanif, and R. Singh.(2015). Generalized exponential chain estimators using two auxiliary variables for stratified sampling with non-response. Science International (Lahore) 27 (2):901–15.

[15]. Sanaullah, A., M. Noor-ul-Amin, M. Hanif, and N. Koyuncu.(2018). Generalized exponential-type estimators for population mean taking two auxiliary variables for unknown means in stratified sampling with sub-sampling the non-respondents. International Journal of Applied and Computational Mathematics 4 (2):1–28.

[16]. JavidShabbir, Sat Gupta &Shakeel Ahmed (2018): A generalized class of estimators under two-phase stratified sampling for nonresponse, Communications in Statistics - Theory and Methods, DOI: 10.1080/03610926.2018.1481969