

5-1-2014

Population Mean Estimation with Sub Sampling the Non-Respondents Using Two Phase Sampling


Sunil Kumar

Alliance University, Karnataka, India, sunilbhoughal06@gmail.com

M Viswanathaiah

Shantha Group of Institutions, Karnataka, India, vmatam@gmail.com

Follow this and additional works at: <http://digitalcommons.wayne.edu/jmasm>

 Part of the [Applied Statistics Commons](#), [Social and Behavioral Sciences Commons](#), and the [Statistical Theory Commons](#)

Recommended Citation

Kumar, Sunil and Viswanathaiah, M (2014) "Population Mean Estimation with Sub Sampling the Non-Respondents Using Two Phase Sampling," *Journal of Modern Applied Statistical Methods*: Vol. 13 : Iss. 1 , Article 12.

DOI: 10.22237/jmasm/1398917460

Available at: <http://digitalcommons.wayne.edu/jmasm/vol13/iss1/12>

This Regular Article is brought to you for free and open access by the Open Access Journals at DigitalCommons@WayneState. It has been accepted for inclusion in Journal of Modern Applied Statistical Methods by an authorized editor of DigitalCommons@WayneState.

Population Mean Estimation with Sub Sampling the Non-Respondents Using Two Phase Sampling

Sunil Kumar
Alliance University
Karnataka, India

M. Viswanathaiah
Shantha Group of Institutions
Karnataka, India

The problem of non-response in double (or two phase) sampling is dealt with combined ratio, product and regression estimators. Expressions of bias and MSE for these estimators are obtained. Comparisons of a proposed strategy with a usual unbiased estimator and other estimators are carried out and results obtained are illustrated numerically using an empirical sample.

Keywords: Study variable, auxiliary variable, bias, mean squared error, non-response

Introduction

In surveys regarding human populations, it is common for some information to be missing, even after some callbacks. Hansen and Hurwitz (1946) considered the problem of non-response while estimating a population mean by taking a sub sample from the non-respondent group and proposed an estimator by considering the information available from response and non-response groups. In estimating population parameters such as the mean, total or ratio, product and regression, sample survey experts sometimes use auxiliary information to improve the precision of estimates. Using Hansen and Hurwitz's (1946) technique, several authors including Cochran (1977), Rao (1986, 1987), Khare and Srivastava (1993, 1995, 1997), Okafor and Lee (2000), Lundström and Särndal (2001), Särndal and Lundström (2005), Tabasum and Khan (2004, 2006), Singh and Kumar (2008, 2009a, b, 2010), and Singh, et al. (2010) have suggested improvements to the population mean estimation procedure in the presence of non-response using an auxiliary variable.

Sunil Kumar is an Associate Professor. Email at sunilbhoulgal06@gmail.com. M. Viswanathaiah is a Professor. Email at vmatam@gmail.com.

POPULATION MEAN ESTIMATION WITH SUB SAMPLING

Following Singh and Ruiz Espejo (2007), a class of ratio-product estimators in two phase sampling in the presence of non-response is suggested in this article, and its properties studied. An estimator was studied by using one auxiliary variable for two phase sampling, which is the combined regression with Okafor and Lee's (2000) estimator and Singh and Ruiz Espejo's (2007) estimator for no information case. The conditions for attaining minimum mean squared error of the proposed classes of estimators were obtained. A comparison of the proposed estimator with other estimators was conducted and a numerical illustration is provided to support the proposed estimator.

Double Sampling Ratio, Product and Regression Estimator

Let y and x be the study and auxiliary variables with population means \bar{Y} and \bar{X} respectively. The population is divided into N_1 (responding) and N_2 (non-responding) units such that $N_1 + N_2 = N$. When the population mean \bar{X} of the auxiliary variable x is unknown, it is suggested that a first phase sample of size n' be selected from the population of size N using the simple random sampling without replacement (SRSWOR) method, and observing the information on variable x . From these selected n' units, a second phase sample size $n (< n')$ is selected for the study variable y , and it is observed that n_1 units respond and n_2 units do not respond in the sample of size n . Further, from n_2 non-responding units, select a sub sample of size $r \left(= \frac{n_2}{k} \right); k > 1$ using SRSWOR. Hence, there are $n_1 + r$ responding units on y . Consequently, to estimate \bar{Y} using the sub sampling scheme suggested by Hansen and Hurwitz (1946),

$$\bar{y}^* = w_1 \bar{y}_1 + w_2 \bar{y}_{2r},$$

where $w_1 = (n_1/n)$, $w_2 = (n_2/n)$; \bar{y}_1 and \bar{y}_{2r} denotes the sample means of the y variable based on n_1 and r units, respectively.

Similarly, to estimate the population mean \bar{X} of the auxiliary variable x , the estimator \bar{x}^* ,

$$\bar{x}^* = w_1 \bar{x}_1 + w_2 \bar{x}_{2r} \quad (1)$$

with variance

$$Var(\bar{x}^*) = \left(\frac{1}{n} - \frac{1}{N}\right) S_x^2 + \frac{W_2(k-1)}{n} S_{x(2)}^2$$

where S_x^2 and $S_{x(2)}^2$ are the population mean square of the auxiliary variable y for the entire population and for the non-responding portion of the population.

Khare and Srivastava (1993) proposed ratio and product methods for estimators respectively as:

$$t_{1R} = \bar{y}^* \left(\frac{\bar{x}'}{\bar{x}^*} \right)$$

and

$$t_{1P} = \bar{y}^* \left(\frac{\bar{x}^*}{\bar{x}'} \right).$$

The *MSE*'s of the estimators t_{1R} and t_{1P} to the first degree of approximation, are

$$MSE(t_{1R}) = \left[\left(\frac{1}{n} - \frac{1}{n'} \right) \{ S_y^2 + R(R - 2\beta_{yx}) S_x^2 \} + \left(\frac{1}{n'} - \frac{1}{N} \right) S_y^2 + \frac{W_2(k-1)}{n} \{ S_{y(2)}^2 + R(R - 2\beta_{yx(2)}) S_{x(2)}^2 \} \right] \quad (2)$$

$$MSE(t_{1P}) = \left[\left(\frac{1}{n} - \frac{1}{n'} \right) \{ S_y^2 + R(R + 2\beta_{yx}) S_x^2 \} + \left(\frac{1}{n'} - \frac{1}{N} \right) S_y^2 + \frac{W_2(k-1)}{n} \{ S_{y(2)}^2 + R(R + 2\beta_{yx(2)}) S_{x(2)}^2 \} \right] \quad (3)$$

where $R = (\bar{Y}/\bar{X})$, $S_{yx} = \rho_{yx} S_y S_x$, $S_{yx(2)} = \rho_{yx(2)} S_{y(2)} S_{x(2)}$, $\beta_{yx} = (S_{yx}/S_x^2)$,

$K_{yx(2)} = (S_{yx(2)}/S_{x(2)}^2)$, and ρ_{yx} and $\rho_{yx(2)}$ respectively denote the correlation coefficients between x and y for the whole population and for the non-response group of the population.

POPULATION MEAN ESTIMATION WITH SUB SAMPLING

Okafor and Lee (2000) proposed a double (two phase) sampling regression estimator in the presence of non-response on study as well as auxiliary variables, as

$$t_{3Re} = \bar{y}^* + b(\bar{x}' - \bar{x}^*) \quad (4)$$

where

$$\hat{b} = (s_{xy}^* / s_x^{*2}), \quad s_{xy}^* = \frac{1}{n-1} \left(\sum_{i=1}^n x_i y_i + k \sum_{i=1}^r x_i y_i - n \bar{x} \bar{y}^* \right), \quad s_x^{*2} = \frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 + k \sum_{i=1}^r x_i^2 - n \bar{x} \bar{x}^* \right).$$

The MSE of the estimator t_{3Re} is

$$MSE(t_{3Re}) = \left[\left(\frac{1}{n} - \frac{1}{n'} \right) (1 - \rho_{yx}^2) S_y^2 + \left(\frac{1}{n'} - \frac{1}{N} \right) S_y^2 + \frac{W_2(k-1)}{n} \{ S_{y(2)}^2 + \beta_{yx} (\beta_{yx} - 2\beta_{yx(2)}) S_{x(2)}^2 \} \right] \quad (5)$$

Singh and Ruiz Espejo (2007) defined an estimator in presence of non-response as

$$t_{SR}^* = \bar{y}^* \left\{ \alpha \frac{\bar{x}'}{\bar{x}^*} + (1 - \alpha) \frac{\bar{x}^*}{\bar{x}'} \right\} \quad (6)$$

where α is any suitably chosen constant.

For $\alpha = 0, 1$, the class of estimators t_{SR}^* reduces to the Khare and Srivastava (1993, 1995) and Tabasum and Khan (2004) product and ratio type estimators, that is, t_{1P} and t_{1R} .

The MSE of the estimator t_{SR}^* to the first degree of approximation is

$$MSE(t_{SR}^*) = \left[\left(\frac{1}{n'} - \frac{1}{N} \right) S_y^2 + \left(\frac{1}{n} - \frac{1}{n'} \right) \{ S_y^2 + R(2\alpha - 1)(R(2\alpha - 1) - 2\beta_{yx}) S_x^2 \} \right. \\ \left. + \frac{W_2(k-1)}{n} \{ S_{y(2)}^2 + R(2\alpha - 1)(R(2\alpha - 1) - 2\beta_{yx(2)}) S_{x(2)}^2 \} \right] \quad (7)$$

which is the minimum, when

$$\alpha = \frac{1}{2} \left(1 + \frac{D^*}{RD} \right),$$

where

$$D = \left\{ \left(\frac{1}{n} - \frac{1}{n'} \right) S_x^2 + \frac{W_2(k-1)}{n} S_{x(2)}^2 \right\}, D^* = \left\{ \left(\frac{1}{n} - \frac{1}{n'} \right) K_{yx} S_x^2 + \frac{W_2(k-1)}{n} K_{yx(2)} S_{x(2)}^2 \right\}.$$

Thus, the minimum MSE of t_{SR}^* is given by

$$\begin{aligned} MSE(t_{SR(opt)}^*) = & \left[\left(\frac{1}{n'} - \frac{1}{N} \right) S_y^2 + \left(\frac{1}{n} - \frac{1}{n'} \right) \left\{ S_y^2 + \frac{D^*}{D} \left(\frac{D^*}{D} - 2\beta_{yx} \right) S_x^2 \right\} \right. \\ & \left. + \frac{W_2(k-1)}{n} \left\{ S_{y(2)}^2 + \frac{D^*}{D} \left(\frac{D^*}{D} - 2\beta_{yx(2)} \right) S_{x(2)}^2 \right\} \right]. \end{aligned} \quad (8)$$

The Proposed Estimator

An estimator was developed using one auxiliary variable for two phase sampling for estimating the population mean \bar{Y} of a study variable y in the presence of non-response. Okafor and Lee's (2000) estimator is combined with the estimator t_{SR}^* . Thus, the proposed estimator is:

$$\begin{aligned} t_{CR}^* &= \left\{ \bar{y}^* + b(\bar{x}' - \bar{x}^*) \right\} \left\{ \varphi \frac{\bar{x}'}{\bar{x}^*} + (1-\varphi) \frac{\bar{x}^*}{\bar{x}'} \right\} \\ &= t_{3Re} \left\{ \varphi \frac{\bar{x}'}{\bar{x}^*} + (1-\varphi) \frac{\bar{x}^*}{\bar{x}'} \right\} \end{aligned} \quad (9)$$

where φ is any suitably chosen constant and t_{3Re} is defined at (4).

To obtain the bias and mean squared error of t_{CR}^* ,

$$\bar{y}^* = \bar{Y}(1 + \varepsilon_0), \bar{x}^* = \bar{X}(1 + \varepsilon_1), \bar{x}' = \bar{X}(1 + \varepsilon_2), s_{xy}^* = S_{xy}(1 + \varepsilon_3), s_x^{*2} = S_x^2(1 + \varepsilon_4),$$

such that

$$E(\varepsilon_i) = 0 \quad \forall i = 0 \text{ to } 4;$$

POPULATION MEAN ESTIMATION WITH SUB SAMPLING

$$E(\varepsilon_0^2) = \left(\frac{1}{n} - \frac{1}{N}\right) S_y^2 + \frac{W_2(k-1)}{n} S_{y(2)}^2; \quad E(\varepsilon_1^2) = \left(\frac{1}{n} - \frac{1}{N}\right) S_x^2 + \frac{W_2(k-1)}{n} S_{x(2)}^2;$$

$$E(\varepsilon_2^2) = \left(\frac{1}{n'} - \frac{1}{N}\right) S_x^2; \quad E(\varepsilon_0\varepsilon_1) = \left(\frac{1}{n} - \frac{1}{N}\right) S_{yx} + \frac{W_2(k-1)}{n} S_{yx(2)};$$

$$E(\varepsilon_0\varepsilon_2) = \left(\frac{1}{n'} - \frac{1}{N}\right) S_{yx}; \quad E(\varepsilon_1\varepsilon_2) = \left(\frac{1}{n'} - \frac{1}{N}\right) S_x^2;$$

$$E(\varepsilon_1\varepsilon_3) = \frac{N(N-n)}{(N-1)(N-2)} \frac{\mu_{21}}{n\bar{X}S_{xy}} + \frac{W_2(k-1)}{n} \frac{\mu_{21(2)}}{\bar{X}S_{xy}};$$

$$E(\varepsilon_2\varepsilon_3) = \frac{N(N-n')}{(N-1)(N-2)} \frac{\mu_{21}}{n'\bar{X}S_{xy}};$$

$$E(\varepsilon_1\varepsilon_4) = \frac{N(N-n)}{(N-1)(N-2)} \frac{\mu_{30}}{n\bar{X}S_x^2} + \frac{W_2(k-1)}{n} \frac{\mu_{30(2)}}{\bar{X}S_x^2};$$

$$E(\varepsilon_2\varepsilon_4) = \frac{N(N-n')}{(N-1)(N-2)} \frac{\mu_{30}}{n'\bar{X}S_x^2};$$

where

$$\mu_{rs} = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{X})^r (y_i - \bar{Y})^s; \quad \mu_{rs(2)} = \frac{1}{N_2} \sum_{i=1}^{N=N_1+N_2} (x_i - \bar{X}_2)^r (y_i - \bar{Y}_2)^s;$$

$$\bar{X}_2 = \frac{1}{N_2} \sum_{i=1}^{N=N_1+N_2} x_i; \quad \bar{Y}_2 = \frac{1}{N_2} \sum_{i=1}^{N=N_1+N_2} y_i; \quad (r, s)$$

being non negative integers.

Expanding t_{CR}^* in terms of ε 's results in

$$t_{CR}^* = \bar{Y} \left\{ 1 + \varepsilon_0 + A_0(\varepsilon_2 - \varepsilon_1)(1 + \varepsilon_3)(1 + \varepsilon_4)^{-1} \right\} \left\{ \varphi(1 + \varepsilon_2)(1 + \varepsilon_1)^{-1} + (1 + \varphi)(1 + \varepsilon_1)(1 + \varepsilon_2)^{-1} \right\} \quad (10)$$

where $A_0 = (\beta/R)$; $R = (\bar{Y}/\bar{X})$.

Assume that $|\varepsilon_4| < 1$, $|\varepsilon_1| < 1$ and $|\varepsilon_2| < 1$ so that $(1 + \varepsilon_4)^{-1}$, $(1 + \varepsilon_1)^{-1}$ and $(1 + \varepsilon_2)^{-1}$ are expandable in terms of ε 's. Expanding the right hand side of (10) in terms of ε 's and neglecting terms of ε 's with power greater than two results in:

$$(t_{CR}^* - \bar{Y}) = \bar{Y} \left\{ \varepsilon_0 - \varepsilon_2 + \varepsilon_1 + \varepsilon_2^2 - \varepsilon_1\varepsilon_2 - \varepsilon_0\varepsilon_2 + \varepsilon_0\varepsilon_1 + \varphi(2\varepsilon_2 - 2\varepsilon_1 + \varepsilon_1^2 - \varepsilon_2^2) + 2\varphi(\varepsilon_0\varepsilon_2 - \varepsilon_0\varepsilon_1) \right. \\ \left. + 2A_0\varphi(\varepsilon_2^2 + \varepsilon_1^2 - 2\varepsilon_1\varepsilon_2) + A_0(\varepsilon_2 - \varepsilon_1 - \varepsilon_2^2 - \varepsilon_1^2 + \varepsilon_1\varepsilon_2 - \varepsilon_2\varepsilon_4 + \varepsilon_1\varepsilon_4 + \varepsilon_2\varepsilon_3 - \varepsilon_1\varepsilon_3) \right\}. \quad (11)$$

Taking expectations of both sides of (11) results in the bias of t_{CR}^* to the first degree of approximation, as

$$B(t_{CR}^*) = \left[\begin{aligned} & \frac{1}{\bar{X}} \left(\frac{1}{n} - \frac{1}{n'} \right) \left\{ (1 - 2\varphi) S_{yx} - (R - \varphi(R + 2\beta_{yx})) S_x^2 \right\} \\ & + \frac{W_2(k-1)}{n\bar{X}} \left\{ (1 - 2\varphi) S_{yx(2)} + \varphi(R + 2\beta_{yx}) S_{x(2)}^2 \right\} \\ & - K_{yx} \left\{ \frac{N^2}{(N-1)(N-2)} \left(\frac{1}{n} - \frac{1}{n'} \right) \left(\frac{\mu_{21}}{S_{xy}} - \frac{\mu_{30}}{S_x^2} \right) + \frac{W_2(k-1)}{n} \left(\frac{\mu_{21(2)}}{S_{xy}} - \frac{\mu_{30(2)}}{S_x^2} \right) \right\} \end{aligned} \right]. \quad (12)$$

Squaring both sides of (11) and neglecting terms of ε 's with power greater than two results in

$$(t_{CR}^* - \bar{Y})^2 = \bar{Y}^2 \left\{ \varepsilon_0 + (2\varphi - 1)(\varepsilon_2 - \varepsilon_1) + A_0(\varepsilon_2 - \varepsilon_1) \right\}^2 \\ = \bar{Y}^2 \left\{ \varepsilon_0^2 + (2\varphi - 1)^2 (\varepsilon_2^2 + \varepsilon_1^2 - 2\varepsilon_1\varepsilon_2) + A_0^2 (\varepsilon_2^2 + \varepsilon_1^2 - 2\varepsilon_1\varepsilon_2) \right. \\ \left. + 2(2\varphi - 1)(\varepsilon_0\varepsilon_2 - \varepsilon_0\varepsilon_1) + 2A_0(2\varphi - 1)(\varepsilon_2^2 + \varepsilon_1^2 - 2\varepsilon_1\varepsilon_2) + 2A_0(\varepsilon_0\varepsilon_2 - \varepsilon_0\varepsilon_1) \right\}. \quad (13)$$

Taking expectations of both sides of (13) results in the *MSE* of t_{CR}^* to the first degree of approximation as:

$$MSE(t_{CR}^*) = \left[\left(\frac{1}{n'} - \frac{1}{N} \right) S_y^2 + \left(\frac{1}{n} - \frac{1}{n'} \right) \left\{ S_y^2 + (2\varphi - 1 + A_0)^2 R^2 S_x^2 - 2(2\varphi - 1 + A_0) R S_{yx} \right\} \right. \\ \left. + \frac{W_2(k-1)}{n} \left\{ S_{y(2)}^2 + (2\varphi - 1 + A_0)^2 R^2 S_{x(2)}^2 - 2(2\varphi - 1 + A_0) R S_{yx(2)} \right\} \right] \quad (14)$$

POPULATION MEAN ESTIMATION WITH SUB SAMPLING

which is the minimum

$$\varphi = \frac{1}{2} \left(1 - A_0 + \frac{D^*}{RD} \right).$$

The resulting minimum mean squared error of t_{CR}^* is therefore given by:

$$\begin{aligned} MSE(t_{CR(opt)}^*) &= \left[\left(\frac{1}{n'} - \frac{1}{N} \right) S_y^2 + \left(\frac{1}{n} - \frac{1}{n'} \right) \left\{ S_y^2 + \frac{D^*}{D} \left(\frac{D^*}{D} - 2\beta_{yx} \right) S_x^2 \right\} \right. \\ &\quad \left. + \frac{W_2(k-1)}{n} \left\{ S_{y(2)}^2 + \frac{D^*}{D} \left(\frac{D^*}{D} - 2\beta_{yx(2)} \right) S_{x(2)}^2 \right\} \right] \\ &= MSE(t_{SR(opt)}^*). \end{aligned} \quad (15)$$

From (1), (2), (3), (5), (7) and (14),

$$Var(\bar{y}^*) - MSE(t_{CR}^*) = \left\{ (2\varphi - 1 + A_0)^2 R^2 D - 2R(2\varphi - 1 + A_0) D^* \right\} \quad (16)$$

$$MSE(t_{1R}) - MSE(t_{CR}^*) = \left\{ R^2 D - 2RD^* - (2\varphi - 1 + A_0)^2 R^2 D + 2(2\varphi - 1 + A_0) RD^* \right\} \quad (17)$$

$$MSE(t_{1P}) - MSE(t_{CR}^*) = \left\{ R^2 D + 2RD^* - (2\varphi - 1 + A_0)^2 R^2 D + 2(2\varphi - 1 + A_0) RD^* \right\} \quad (18)$$

$$MSE(t_{3Re}) - MSE(t_{CR}^*) = - \left\{ 4\varphi^2 R^2 D - 4\varphi R(RD - \beta_{yx} D + D^*) + R^2 D \right\} \quad (19)$$

$$MSE(t_{SR}^*) - MSE(t_{CR}^*) = RD \left\{ 2A_0 - A_0^2 - 4\alpha A_0 \right\} + 2A_0 D^* \quad (20)$$

The differences given by (16), (17), (18), (19) and (20) are positive, respectively, if

$$\left. \begin{aligned} & \text{either } 0 < \varphi < \frac{1}{2} (1 - A_0) + \frac{D^*}{RD} \\ & \text{or } \frac{1}{2} (1 - A_0) + \frac{D^*}{RD} < \varphi < 0 \end{aligned} \right\} \quad (21)$$

$$\left. \begin{aligned} \text{either } \frac{2D^* - \beta_{yx}D}{2RD} < \varphi < \left(1 - \frac{\beta_{yx}}{2R}\right) \\ \text{or } \left(1 - \frac{\beta_{yx}}{2R}\right) < \varphi < \frac{2D^* - \beta_{yx}D}{2RD} \end{aligned} \right\} \quad (22)$$

$$\left. \begin{aligned} \text{either } \frac{-\beta_{yx}}{2R} < \varphi < \left(1 + \frac{D^*}{RD} - \frac{\beta_{yx}}{2R}\right) \\ \text{or } \left(1 + \frac{D^*}{RD} - \frac{\beta_{yx}}{2R}\right) < \varphi < \frac{-\beta_{yx}}{2R} \end{aligned} \right\} \quad (23)$$

$$\left. \begin{aligned} \text{either } -\frac{1}{2} < \varphi < \left(\frac{1}{2} + \frac{D^*}{RD}\right) \\ \text{or } \left(\frac{1}{2} + \frac{D^*}{RD}\right) < \varphi < -\frac{1}{2} \end{aligned} \right\} \quad (24)$$

$$\left. \begin{aligned} \text{either } 0 < \varphi < \frac{1}{2} \left(1 - \frac{A_0}{2}\right) + \frac{D^*}{2RD} \\ \text{or } \frac{1}{2} \left(1 - \frac{A_0}{2}\right) + \frac{D^*}{2RD} < \varphi < 0 \end{aligned} \right\} \quad (25)$$

The proposed estimator t_{CR}^* is more robust than estimators \bar{y}^* , t_{1R} , t_{1P} and t_{SR}^* respectively, if (21) – (25) hold true.

Empirical Study

To examine the robustness of the proposed estimators, consider the following data sets (Khare & Sinha, 2004, p. 53) from a survey on physical growth of an upper socioeconomic group of 95 Varanasi school children under an Indian Council of Medical Research (ICMR) study, Department of Pediatrics, Banaras Hindu University (BHU) during 1983-1984. The first 25% (24 children) were considered non-response units.

POPULATION MEAN ESTIMATION WITH SUB SAMPLING

The parameter values related to the study variable y (weight in kg) and the auxiliary variable x (chest circumference in cm) were:

$$\bar{Y} = 19.4968, \quad \bar{X} = 55.8611w, \quad S_y = 3.0435, \quad S_x = 3.2735, \quad S_{y(2)} = 2.3552, \\ S_{x(2)} = 2.5137, \quad \rho_{yx} = 0.8460, \quad \rho_{yx(2)} = 0.7290, \quad R = 0.3490, \quad \beta_{yx} = 0.7865, \\ \beta_{yx(2)} = 0.6829, \quad W_2 = 0.25, \quad N_2 = 24, \quad N_1 = 71, \quad N = 95, \quad n = 35, \quad n' = 70.$$

The percent relative efficiencies (PREs) of different suggested estimators were computed with respect to a usual unbiased estimator \bar{y}^* for different values of k .

Table 1: Percent relative efficiency of different \bar{Y} estimators with respect to \bar{y}^* .

Estimators	(1/k)			
	(1/5)	(1/4)	(1/3)	(1/2)
\bar{y}^*	100.00	100.00	100.00	100.00
t_{1R}	165.65	165.35	164.95	164.41
t_{3Re}	218.74	220.25	222.29	225.16
$t_{SR(opt)}^* = t_{CR(opt)}^*$	220.00	221.16	222.81	225.18

Table 1 shows that

- (i) the PRE's of the estimators t_{3Re} and $t_{CR(opt)}^*$ increase as the value of k increases, while the PRE's of the estimator t_{1R} decrease as the value of k increases.
- (ii) the performance of the proposed estimator $t_{CR(opt)}^*$ is the best among all other estimators \bar{y}^* , t_{1R} and t_{3Re} because it has the largest gain in efficiency.

Based on these study results, the proposed estimator t_{CR}^* is recommended for use in practice.

References

Cochran, W. G. (1977). *Sampling Techniques, 3rd Edition*. New York, NY: John Wiley and Sons.

Hansen, M. H., & Hurwitz, W. N. (1946). The problem of non-response in sample surveys. *Journal of the American Statistical Association, 41*: 517- 529.

Khare, B. B., & Sinha, R. R. (2004). Estimation of finite population ratio using two phase sampling scheme in the presence of non-response. *Aligarh Journal of Statistics, 24*: 43-56.

Khare, B. B., & Srivastava, S. (1993). Estimation of population mean using auxiliary character in presence of non-response. *National Academy Science Letters, 16*: 111-114.

Khare, B. B., & Srivastava, S. (1995). Study of conventional and alternative two-phase sampling ratio, product and regression estimators in presence of non-response. *Proceedings of the Indian National Science Academy, 65*: 195-203.

Khare, B. B., & Srivastava, S. (1997). Transformed ratio type estimators for the population mean in the presence of non response. *Communication in Statistics - Theory and Methods, 26*: 1779-179.

Lundström, S., & Särndal, C. E. (2001). *Estimation in the presence of non response and frame imperfections*. Örebro: SCB-Tryck.

Okafor, F. C., & Lee, H. (2000). Double sampling for ratio and regression estimation with sub-sampling the non-respondents. *Survey Methodology, 26*(2): 183-188.

Rao, P. S. R. S. (1986). Ratio estimation with sub sampling the non-respondents. *Survey Methodology, 12*: 217-230.

Rao, P. S. R. S. (1987). *Ratio and regression estimates with sub sampling the non-respondents*. Paper presented at a special contributed session of the International Statistical Association Meeting, Sept., 2-16, 1987, Tokyo, Japan.

Särndal, C. E. & Lundström, S. (2005). *Estimation in surveys with non-response*. New York: John Wiley and Sons.

Singh, H. P., & Ruiz Espejo, M. (2007). Double sampling ratio-product estimator of a finite population mean in sample survey. *Journal of Applied Statistics, 34*(1): 71-85.

Singh, H. P., & Kumar, S. (2008). A regression approach to the estimation of finite population mean in presence of non-response. *Australian and New Zealand Journal of Statistics, 50*(4): 395-408.

POPULATION MEAN ESTIMATION WITH SUB SAMPLING

Singh, H. P., & Kumar, S. (2009a). A general class of estimators of the population mean in survey sampling using auxiliary information with sub sampling the non-respondents. *The Korean Journal of Applied Statistics*, 22(2): 387-402.

Singh, H. P., & Kumar, S. (2009b). A general procedure of estimating the population mean in the presence of non-response under double sampling using auxiliary information. *SORT*, 33(1): 71-84.

Singh, H. P., & Kumar, S. (2010). Estimation of mean in presence of non-response using two phase sampling scheme. *Statistical Papers*, 51: 559 -582.

Singh, H. P., Kumar, S., & Kozak, M. (2010). Improved estimation of finite-population mean when sub-sampling is employed to deal with non-response. *Communication in Statistics - Theory and Methods*, 39(5): 791-802.

Tabasum, R., & Khan, I. A. (2004). Double sampling for ratio estimation with non- response. *Journal of the Indian Society of Agricultural Statistics*, 58: 300-306.

Tabasum, R., & Khan, I. A. (2006). Double sampling ratio estimator for the population mean in presence of non-response. *Assam Statistical Review*, 20: 73-83.