Pak. J. Statist.
2015 Vol. 31(3), 295-306

GENERALIZED ESTIMATORS FOR POPULATION MEAN IN THE PRESENCE OF NON-RESPONSE FOR TWO-PHASE SAMPLING

Muhammad Ismail[1], **Muhammad Hanif**[2] and **Muhammad Qaiser Shahbaz**[3]

[1] Department of Statistics, COMSATS Institute of Information Technology Lahore. Pakistan. Email: drismail39@gmail.com

[2] National College of Business Administration and Economic Lahore. Pakistan. Email: drmianhanif@gmail.com

[3] Department of Statistics, King Abdul Aziz University, KSA. Email: qshahbaz@gmail.com

ABSTRACT

In this paper some generalized estimators for population mean in the presence of non–response for two-phase sampling using information of one auxiliary variable x for different situations have been proposed. Special cases of these estimators have also been given. Empirical and theoretical study is carried out to compare the efficiency of the suggested estimators.

KEY WORDS

Non–Response; Mean Square Error; Auxiliary Variable.

1. INTRODUCTION

It has been observed that non-response creates problem mostly in all types of surveys and this cannot be eliminated simply by increasing the size of sample. Non-response may be non-ignorable or ignorable it depends whether it is correlated with the target variable or not (Glynn, Laird, & Rubin, 1993; Little, 1982). It has been observed that non-response increases bias in estimates which is the main reason of reducing efficiency of the surveys. Many methods have been considered by different survey statisticians for estimating the population characteristics in the presence of non–response for two-phase sampling. In this aspect the two-phase sampling method has been a popular one.

Suppose a sample of size " n " is drawn from a population of size " N " by using simple random sample without replacement (SRSWOR). From the " n " sample units, " r_1 " sample units respond to survey variable " Y " and " r_2 " are non-respondents. The population divided in respondents and non-respondents containing " N_1 " and " N_2 " units corresponds to sample respondents and non–respondents. From r_2 non–respondents unit, a sub-sample of " k " $\left(k = r_2 / h, h > 1\right)$ sample units is drawn and obtained information about study variable from these k units. Hansen and Hurwitz (1946), proposed an unbiased estimator of population mean by using sub-sampling technique to overcome the problem of non-response:

The Hansen and Hurwitz (1946) estimator in presence of non-response is

© 2015 Pakistan Journal of Statistics

$$\bar{y}^{*} = (r_1/n)\,\bar{y}_{r_1 1} + (r_2/n)\,\bar{y}_{k2}, \qquad (1.1)$$

where $\bar{y}_1 = r_1^{-1}\sum_{i=1}^{r_1} y_i$ and $\bar{y}_{k2} = k^{-1}\sum_{i=1}^{k} y_i$ are means of study variable " y ". The estimator (1.1) is unbiased with variance:

$$Var\left(\bar{y}^{*}\right) = \lambda S_y^2 + \theta S_{y_2}^2, \qquad (1.2)$$

where $S_y^2 = \sum_{i=1}^{N}\left(y_i - \bar{Y}\right)^2 \Big/ (N-1)$, $S_{y_2}^2 = \sum_{i=1}^{N_2}\left(y_i - \bar{Y}_2\right)^2 \Big/ (N_2 - 1)$, $\lambda = \left(\frac{1}{n} - \frac{1}{N}\right)$, $\theta = \frac{N_2(h-1)}{Nn}$, $\bar{Y} = N^{-1}\sum_{i=1}^{N} y_i$ and $\bar{Y}_2 = N_2^{-1}\sum_{i=1}^{N_2} y_i$. Some other notable estimators in this area are those proposed by Cochran (2007), Rao (1986), Naik and Gupta (1991), Khare and Srivastava (1993, 1995), Tripathi and Khare (1997), Tabasum and Khan (2004, 2006), Singh and Kumar (2008a, 2008b, 2008c, 2009, 2011), Khare and Srivastava (2010), Ismail Shahbaz, and Hanif (2011) and Ismail, Hanif and Shahbaz (2013).

2. PROCEDURE OF TWO-PHASE SAMPLING IN THE PRESENCE OF NON-RESPONSE

The use of Two-phase sampling procedure in the presence of non–response has been recently emerged as more efficient method of constructing the estimators. The procedure is described as under:

i) Draw a sample of size " n_1 " from population of size " N " by using simple random sampling without replacement (SRSWOR) and record information about the auxiliary variable say " X ". It is the first phase sample.

ii) Draw another sample of size " n_2 " from the first phase sample of size " n_1 " by using simple random sampling without replacement (SRSWOR).Suppose " r_1 " elements are respondents and " r_2 " elements are non-respondents.

iii) Draw a sub-sample of size " k " from non-respondents where $\left(k = r_2 / h, h > 1\right)$ and collect information on study variable " Y ".

Survey statisticians have proposed estimators of population mean using two phase sampling in presence of non-response by considering following two situations when population mean of auxiliary variables x is unknown:

Situation 1:

The population mean of auxiliary variable x is unknown, incomplete information on study variable y and incomplete information of auxiliary variable x is available [Khare and Srivastava (1993)].

Situation 2:

The population mean of auxiliary variable x is unknown, incomplete information on study variable y and complete information of auxiliary variable x is available [Khare and Srivastava (1993)].

Using the notations $\lambda_1 = n_1^{-1} - N^{-1}$, $\lambda_3 = n_2^{-1} - n_1^{-1}$, $R = \bar{Y}/\bar{X}$, $C = \beta/R$, $\beta = S_{xy}/S_x^2$, $C_{(2)} = \beta_{(2)}/R$, $\beta_{(2)} = S_{xy(2)}/S_{x_2}^2$, $\lambda_2 = n_2^{-1} - N^{-1}$ and " $\bar{x}_1$ " is the mean of first phase sample. Several estimators of population mean have been proposed from time to time. These are discussed below.

Khare and Srivastava (1993) proposed the following classical ratio estimator by using situation 1:

$$t_2 = \bar{y}^*\left(\bar{x}_1/\bar{x}^*\right). \tag{2.1}$$

The mean square error of (2.1) is

$$MSE(t_2) \approx \lambda_3\left\{S_y^2 + R^2S_x^2(1-2C)\right\} + \lambda_1 S_y^2 + \theta\left\{S_{y(2)}^2 + R^2S_{x(2)}^2\left(1-2C_{(2)}\right)\right\}. \tag{2.2}$$

Khare and Srivastava (1993) proposed product type estimator for situation 1 which is given as:

$$t_3 = \bar{y}^*\left(\frac{\bar{x}^*}{\bar{x}_1}\right). \tag{2.3}$$

The mean square error of (2.3) is

$$MSE(t_3) \approx \lambda_3\left\{S_y^2 + R^2S_x^2(1+2C)\right\} + \lambda_1 S_y^2 + \theta\left\{S_{y(2)}^2 + R^2S_{x(2)}^2\left(1+2C_{(2)}\right)\right\}. \tag{2.4}$$

Khare and Srivastava (1995) also proposed regression estimator for situation 1 in presence of non-response and by using two-phase sampling procedure as

$$t_4 = \bar{y}^* + b^*\left(\bar{x}_1 - \bar{x}^*\right), \tag{2.5}$$

where $b^* = \frac{s_{xy}^*}{s_x^{*2}}$ and the estimates s_x^{*2} and s_{xy}^* are based on the available information under the given sampling design. The mean squares errors of (2.5) is

$$MSE(t_4) \approx \lambda_3 S_y^2\left(1-\rho^2\right) + \lambda_1 S_y^2 + \theta\left[S_{y(2)}^2 + \beta^2 S_{x(2)}^2 - 2\beta S_{xy(2)}^2\right]. \tag{2.6}$$

Singh and Kumar (2008a) considered the situation in which information on the auxiliary variable x is obtained from all the sample units and population mean of the auxiliary variable is unknown, but some units fail to supply information on the study variable y. The proposed estimator is:

$$t_5 = \bar{y}^*\left(\frac{\bar{x}_1}{\bar{x}^*}\right)\left(\frac{\bar{x}_1}{\bar{x}}\right), \tag{2.7}$$

where $\bar{x}^*$ and $\bar{x}$, both are unbiased estimators of the population mean $\bar{X}$ of the auxiliary variable X.

The mean square errors of (2.7) is

$$MSE(t_5) \approx \lambda_3\left[S_y^2 + 4RS_x^2(R-\beta)\right] + \theta\left[S_{y(2)}^2 + RS_{x(2)}^2\left(R - 2\beta_{(2)}\right)\right] + \lambda_1 S_y^2. \quad (2.8)$$

Another estimator of for mentioned situation is:

$$t_6 = \bar{y}^*\left(\frac{\bar{x}^*}{\bar{x}_1}\right)\left(\frac{\bar{x}}{\bar{x}_1}\right). \quad (2.9)$$

The mean square errors of (2.9) is

$$MSE(t_6) \approx \lambda_3\left[S_y^2 + 4RS_x^2(R+\beta)\right] + \theta\left[S_{y(2)}^2 + RS_{x(2)}^2\left(R + 2\beta_{(2)}\right)\right] + \lambda_1 S_y^2. \quad (2.10)$$

Khare and Srivastava (2010) have proposed following generalized two-phase sampling estimator for the population mean in the presence of non-response under the situation 1:

$$t_7 = \bar{y}^*\left(\frac{\bar{x}^*}{\bar{x}_1}\right)^{\alpha}, \quad (2.11)$$

where α is a constant.

The mean square error of (2.11) is

$$MSE(t_7) \approx \lambda_3\left(S_y^2 + \alpha^2R^2S_x^2 + 2\alpha RS_{yx}\right) + \lambda_1 S_y^2 + \theta\left(S_{y(2)}^2 + \alpha^2R^2S_{x(2)}^2 + 2\alpha RS_{yx(2)}\right). \quad (2.12)$$

Khare and Srivastava (1993) proposed ratio estimator in presence of non response for two-phase sampling under situation 2 is:

$$t_8 = \bar{y}^*\left(\frac{\bar{x}_1}{\bar{x}}\right). \quad (2.13)$$

The mean square error of (2.13) is

$$MSE(t_8) \approx \lambda_3\left\{S_y^2 + R^2S_x^2(1-2C)\right\} + \lambda_1 S_y^2 + \theta S_{y(2)}^2. \quad (2.14)$$

Khare and Srivastava (1993) proposed product estimator under situation 2 is:

$$t_9 = \bar{y}^*\left(\frac{\bar{x}}{\bar{x}_1}\right), \quad (2.15)$$

The mean square error of (2.15) is

$$MSE(t_9) \approx \lambda_3\left\{S_y^2 + R^2S_x^2(1+2C)\right\} + \lambda_1 S_y^2 + \theta S_{y(2)}^2. \quad (2.16)$$

Khare and Srivastava (1995) proposed following regression estimator for situation 2:

$$t_{10} = \bar{y}^{*} + b^{**}\left(\bar{x}_1 - \bar{x}\right), \tag{2.17}$$

where $b^{**} = \frac{\hat{S}_{xy}}{s_x^2}$ and the estimates and $\hat{S}_{xy}$ are based on the available information under the given sampling design. The mean squares error of (2.17) is

$$MSE\left(t_{10}\right) \approx \lambda_3 S_y^2\left(1-\rho^2\right) + \lambda_1 S_y^2 + \theta S_{y(2)}^2. \tag{2.18}$$

Khare and Srivastava (2010) proposed generalized estimator for the population mean in the presence of non-response for situation 2 given by is

$$t_{11} = \bar{y}^{*}\left(\frac{\bar{x}}{\bar{x}_1}\right)^{\alpha}, \tag{2.19}$$

where α is a constant. The mean square error of (2.19)

$$MSE\left(t_{11}\right) \approx \lambda_3\left(S_y^2 + \alpha^2 R^2 S_x^2 + 2\alpha R S_{yx}\right) + \lambda_1 S_y^2 + \theta S_{y(2)}^2. \tag{2.20}$$

We now propose some new estimators using two phase sampling procedure in presence of non-response in the following section.

3. PROPOSED ESTIMATORS

In this section we have suggested four generalized estimators of population mean for two-phase sampling in the presence of non–response along with their special cases. All the proposed generalized estimators and different choices of constants have been taken from the motivation of Tripathi et al. (1994).

3.1 Generalized Estimators for Situation 1

We propose following generalized estimator using single auxiliary variable in presence of non-response:

$$t_{ng1(1)} = \frac{\bar{y}^{*} - p\left(\bar{x}^{*b} - \bar{x}_1^{b}\right)}{\left[\bar{x}^{*} - q\left(\bar{x}^{*b} - \bar{x}_1^{b}\right)\right]^{a}}\bar{x}^{*a}; \tag{3.1}$$

where a, b, p and q are suitably chosen values.

Using $\bar{y}^{*} = \bar{Y} + \bar{e}_y^{*}$, $\bar{x}^{*} = \bar{X} + \bar{e}_x^{*}$ and $\bar{x}_1 = \bar{X} + \bar{e}_{x_1}$ the above estimator can be written as:

$$t_{ng1(1)} - \bar{Y} = \bar{e}_y^{*} + b\bar{X}^{b-1}\left(aq\frac{\bar{Y}}{\bar{X}} - p\right)\left(\bar{e}_x^{*} - \bar{e}_{x_1}\right).$$

Squaring above equation, applying expectation and using following results:

$$\left.\begin{array}{l} E\left(\bar{e}_y^{*2}\right)=\lambda_2 S_y^2+\theta S_{y_2}^2\ ;\ E\left(\bar{e}_x^{*2}\right)=\lambda_2 S_x^2+\theta S_{x_2}^2\ ; E\left(\bar{e}_{x_1}^2\right)=\lambda_1 S_x^2 \\ E\left(\bar{e}_y^{*}\bar{e}_x^{*}\right)=\lambda_2 S_{xy}+\theta S_{xy(2)}\ ; E\left(\bar{e}_y^{*}\bar{e}_{x_1}\right)=\lambda_1 S_{xy}\ ; E\left(\bar{e}_x^{*}\bar{e}_{x_1}\right)=\lambda_1 S_x^2 \end{array}\right\}, \tag{3.2}$$

The mean square error of (3.1) is

$$MSE\left(t_{ng1(1)}\right)\approx\lambda_3\left[S_y^2+b^2\bar{X}^{2(b-1)}\left(aqR-p\right)^2 S_x^2+2b\bar{X}^{b-1}\left(aqR-p\right)S_{xy}\right]+\lambda_1 S_y^2$$
$$+\theta\left[S_{y(2)}^2+b^2\bar{X}^{2(b-1)}\left(aqR-p\right)^2 S_{x(2)}^2+2b\bar{X}^{b-1}\left(aqR-p\right)S_{xy(2)}\right], \tag{3.3}$$

The mean square error (3.3) can also be written as:

$$MSE\left(t_{ng1(1)}\right)\approx\lambda_3\left[S_y^2+K^2 S_x^2+2KS_{xy}\right]+\lambda_1 S_y^2+\theta\left[S_{y(2)}^2+K^2 S_{x(2)}^2+2KS_{xy(2)}\right], \tag{3.4}$$

where $K=b\bar{X}^{b-1}\left(aqR-p\right)$.

3.2 Generalized Estimators for Situation 2

The proposed generalized estimator using single auxiliary variable for situation 2 is:

$$t_{ng1(2)}=\frac{\bar{y}^*-p\left(\bar{x}^b-\bar{x}_1^b\right)}{\left[\bar{x}-q\left(\bar{x}^b-\bar{x}_1^b\right)\right]^a}\bar{x}^a; \tag{3.5}$$

where a, b, p and q are suitably chosen values.

Using the relations $\bar{y}^*=\bar{Y}+\bar{e}_y^*$, $\bar{x}=\bar{X}+\bar{e}_x$ and $\bar{x}_1=\bar{X}+\bar{e}_{x_1}$ we can write (3.5) as:

$$t_{ng1(2)}-\bar{Y}\approx\bar{e}_y^*+b\bar{X}^{b-1}\left(aq\frac{\bar{Y}}{\bar{X}}-p\right)\left(\bar{e}_x-\bar{e}_{x_1}\right).$$

Squaring above equation, applying expectation and using (3.2), the mean square error of (3.5) is

$$MSE\left(t_{ng1(2)}\right)\approx\lambda_3\left[S_y^2+b^2\bar{X}^{2(b-1)}\left(aqR-p\right)^2 S_x^2+2b\bar{X}^{b-1}\left(aqR-p\right)\mathrm{S}_{\mathrm{xy}}\right]$$
$$+\lambda_1 S_y^2+\theta S_{y(2)}^2, \tag{3.6}$$

The mean square error (3.6) can also be written as

$$MSE\left(t_{ng1(2)}\right)\approx\lambda_3\left[S_y^2+K^2 S_x^2+2KS_{xy}\right]+\lambda_1 S_y^2+\theta S_{y(2)}^2, \tag{3.7}$$

where $\mathrm{K}=b\bar{X}^{b-1}\left(aqR-p\right)$.

We have given some special cases of above estimators in the following table.

Table 3.1
Special Cases of Generalized Estimator $t_{ng1(1)}$ for Situation 1
for Different Values of a, b, p and q, (3.1) and (3.3)

S#	Value of Constant	Estimator (3.1)	MSE (3.3)
1	$p=0, q=0,$ $a=a, b=b$	$\bar{y}^*$ (1.1) Hansen and Hurwitz(1946)	$Var\left(\bar{y}^*\right) = \lambda_2 S_y^2 + \theta S_{y(2)}^2$ (1.2) Hansen and Hurwitz(1946)
2	$p=b, q=q,$ $a=0, b=1$	$t_4 = \bar{y}^* + b\left(\bar{x}_1 - \bar{x}^*\right)$ (2.7) Khare and Srivastava (1995)	$\lambda_3\left(S_y^2+\beta^2 S_x^2 - 2\beta S_{xy}\right)+\lambda_1 S_y^2$ $+\theta\left(S_{y(2)}^2 + \beta^2 S_{x(2)}^2 - 2\beta S_{xy(2)}\right)$ (2.8) Khare and Srivastava (1995)
3	$p=p, q=q,$ $a=1, b=1$	$t_{n1(1)} = \dfrac{\bar{y}^* - p\left(\bar{x}^* - \bar{x}_1\right)}{\bar{x}^* - q\left(\bar{x}^* - \bar{x}_1\right)}\bar{x}^*$	$\lambda_3\left[S_y^2 + \left(qR-p\right)^2 S_x^2 + 2\left(qR-p\right)S_{xy}\right]+\lambda_1 S_y^2$ $+\theta\left[S_{y(2)}^2 + \left(qR-p\right)^2 S_{x(2)}^2 + 2\left(qR-p\right)S_{xy(2)}\right]$
4	$p=0,$ $q=1-w,$ $a=1, b=1$	$t_{n2(1)} = \dfrac{\bar{y}^*}{\bar{x}^* - (1-w)\left(\bar{x}^* - \bar{x}_1\right)}\bar{x}^*$	$\lambda_3\left[S_y^2 + R^2\left(1-w\right)^2 S_x^2 + 2R\left(1-w\right)S_{xy}\right]+\lambda_1 S_y^2$ $+\theta\left[S_{y(2)}^2 + R^2\left(1-w\right)^2 S_{x(2)}^2 + 2R\left(1-w\right)S_{xy(2)}\right]$
5	$p=w\lambda,$ $q=q,$ $a=0, b=1$	$t_{n3(1)} = \bar{y}^* - w\lambda\left(\bar{x}^* - \bar{x}_1\right)$	$\lambda_3\left(S_y^2 + w^2\lambda^2 S_x^2 - 2w\lambda S_{xy}\right) + \lambda_1 S_y^2$ $+\theta\left(S_{y(2)}^2 + w^2\lambda^2 S_{x(2)}^2 - 2w\lambda S_{xy(2)}\right)$
6	$p=-1,$ $q=0,$ $a=1,$ $b=\dfrac{1}{2}$	$t_{n4(1)} = \bar{y}^* + \left(\sqrt{\bar{x}^*} - \sqrt{\bar{x}_1}\right)$	$\lambda_3\left(S_y^2 + \dfrac{S_x^2}{4\bar{X}} + \dfrac{S_{xy}}{\sqrt{\bar{X}}}\right) + \lambda_1 S_y^2$ $+\theta\left(S_{y(2)}^2 + \dfrac{S_{x(2)}^2}{4\bar{X}} + \dfrac{S_{xy(2)}}{\sqrt{\bar{X}}}\right)$
7	$p=-1,$ $q=0,$ $a=1,$ $b=-\dfrac{1}{2}$	$t_{n5(1)} = \bar{y}^* + \left(\dfrac{1}{\sqrt{\bar{x}^*}} - \dfrac{1}{\sqrt{\bar{x}_1}}\right)$	$\lambda_3\left(S_y^2 + \dfrac{S_x^2}{4\bar{X}^3} - \dfrac{S_{xy}}{\bar{X}\sqrt{\bar{X}}}\right)+\lambda_1 S_y^2$ $+\theta\left(S_{y(2)}^2 + \dfrac{S_{x(2)}^2}{4\bar{X}^3} - \dfrac{S_{xy(2)}}{\bar{X}\sqrt{\bar{X}}}\right)$

Table 3.2
Special Cases of Generalized Estimator $t_{ng1(2)}$ for Situation 2
for Different Values of a, b, p and q, (3.5) and (3.6)

S#	Value of Constant	Estimator (3.5)	MSE (3.6)
1	$p = 0, q = 0, a = a, b = b$	$\bar{y}^*$ (1.1) Hansen and Hurwitz(1946)	$Var\left(\bar{y}^*\right) = \lambda_2 S_y^2 + \theta S_{y(2)}^2$ (1.2) Hansen and Hurwitz(1946)
2	$p = b^{**}$, $q = q$, $a = 0, b = 1$	$t_{10} = \bar{y}^* + b^{**}\left(\bar{x}_1 - \bar{x}\right)$ (2.21) Khare and Srivastava (1995)	$\lambda_3\left(S_y^2 + \beta^2 S_x^2 - 2\beta S_{xy}\right) + \lambda_1 S_y^2 + \theta S_{y(2)}^2$ (2.22) Khare and Srivastava (1995)
3	$p = p, q = q, a = 1, b = 1$	$t_{n1(2)} = \dfrac{\bar{y}^* - p\left(\bar{x} - \bar{x}_1\right)}{\bar{x} - q\left(\bar{x} - \bar{x}_1\right)}\bar{x}$	$\lambda_3\left[S_y^2 + \left(qR - p\right)^2 S_x^2 + 2\left(qR - p\right)S_{xy}\right] + \lambda_1 S_y^2 + \theta S_{y(2)}^2$
4	$p = 0$, $q = 1 - w$, $a = 1, b = 1$	$t_{n2(2)} = \dfrac{\bar{y}^*}{\bar{x} - (1-w)\left(\bar{x} - \bar{x}_1\right)}\bar{x}$	$\lambda_3\left[S_y^2 + R^2\left(1-w\right)^2 S_x^2 + 2R\left(1-w\right)S_{xy}\right] + \lambda_1 S_y^2 + \theta S_{y(2)}^2$
5	$p = w\lambda$, $q = q, a = 0$, $b = 1$	$t_{n3(2)} = \bar{y}^* - w\lambda\left(\bar{x} - \bar{x}_1\right)$	$\lambda_3\left(S_y^2 + w^2\lambda^2 S_x^2 - 2w\lambda S_{xy}\right) + \lambda_1 S_y^2 + \theta S_{y(2)}^2$
6	$p = -1$, $q = 0, a = 1$, $b = \dfrac{1}{2}$	$t_{n4(2)} = \bar{y}^* + \left(\sqrt{\bar{x}} - \sqrt{\bar{x}_1}\right)$	$\lambda_3\left(S_y^2 + \dfrac{S_x^2}{4\bar{X}} + \dfrac{S_{xy}}{\sqrt{\bar{X}}}\right) + \lambda_1 S_y^2 + \theta S_{y(2)}^2.$
7	$p = -1, q = 0$, $a = 1$, $b = -\dfrac{1}{2}$	$t_{n5(2)} = \bar{y}^* + \left(\dfrac{1}{\sqrt{\bar{x}}} - \dfrac{1}{\sqrt{\bar{x}_1}}\right)$	$\lambda_3\left(S_y^2 + \dfrac{S_x^2}{4\bar{X}^3} - \dfrac{S_{xy}}{\bar{X}\sqrt{\bar{X}}}\right) + \lambda_1 S_y^2 + \theta S_{y(2)}^2$

4. COMPARISON OF PROPOSED ESTIMATORS WITH EXISTING ESTIMATORS

Comparison of some new estimators with some existing estimators for situation 1 is given below.

Khare and Srivastava (1993) vs. Proposed Estimator $t_{n2(1)}$ for Situation 1

For comparison of t_2 and $t_{n2(1)}$, $t_{n2(1)}$ is better than t_2 if $MSE\left(t_2\right) - MSE\left(t_{n2(1)}\right) > 0$.

It is cleared that in situation 1 proposed estimator $t_{n2(1)}$ is better than Khare and Srivastava (1993) estimator in the presence of non-response for two-phase sampling if $2C < w < 2$ and $2C_{(2)} < w < 2$.

Khare and Srivastava (1993) Product Estimator vs. Proposed Estimator $t_{n10(1)}$ for Situation 1

For comparison of t_3 and $t_{n10(1)}$, $t_{n10(1)}$ is better than t_3 if $MSE(t_3) - MSE(t_{n10(1)}) > 0.$

By taking the difference, $MSE(t_3) - MSE(t_{n10(1)}) = 0$.

It is cleared that proposed estimator $t_{n10(1)}$ is equal to the Khare and Srivastava (1993) product estimator (3.4.5) in the presence of non-response for two-phase sampling.

Comparison of Estimators for Situation 2

Comparison of some new estimators with some existing estimators for situation 2 is given below.

- **Khare and Srivastava (1993) with Proposed Estimator $t_{n2(2)}$.**

  For comparison of t_8 and $t_{n2(2)}$, $t_{n2(2)}$ is better than t_{18} if $MSE(t_8) - MSE(t_{n2(2)}) > 0.$

  It is cleared that our proposed estimator $t_{n2(2)}$ is better than Khare and Srivastava (1993) estimator, if $2C < w < 2$.

- **Khare and Srivastava (1993) Product Estimator t_9 with Proposed Estimator $t_{n10(2)}$ for Situation 2**

  For comparison of t_9 and $t_{n10(2)}$, $t_{n10(2)}$ is better than t_9 if $MSE(t_9) - MSE(t_{n10(2)}) > 0.$

  Taking difference

  $$MSE(t_9) - MSE(t_{n10(2)}) = 0$$

  It is cleared that the proposed estimator $t_{n10(2)}$ equal to Khare and Srivastava (1993) product estimator t_9 .

5. EMPIRICAL COMPARISON BETWEEN SOME PROPOSED ESTIMATORS WITH SOME OF THE EXISTING ESTIMATORS

In this section empirical comparison has been made by taking some of the proposed estimators and some of the existing estimators. The data of Khare and Sinha (2007) and Singh and Kumar (2011) have been considered. Description of the population data is given below:

The data on physical growth of upper socioeconomic group of 95 school children of Varanasi under an ICMR study, Department of pediatrics, B.H.U., during 1983-1984 has been taken under study. The first 25% units have been considered as non responding units. Let us consider the study and auxiliary variable as follows:

y : weight in kg of the children,

x : skull circumference in cm of the children,

z : chest circumference in cm of the children.

$\bar{Y} = 19.4968$, $C_y = 0.15613$, $C_{y_2} = 0.12075$,

$\bar{X} = 51.1726$, $C_x = 0.03006$, $C_{x_2} = 0.02478$,

$\bar{Z} = 55.8611$, $C_z = 0.05860$, $C_{z_2} = 0.05402$,

$\rho_{yx} = 0.328$, $\rho_{yx_2} = 0.477$, $\rho_{yz_2} = 0.729$,

$\rho_{yz} = 0.846$, $\rho_{xz_2} = 0.570$ $\rho_{xz} = 0.297$,

$n_2 = 24$, $n_1 = 35$, $W_2 = 0.25$,

$N = 95$,

Table 5.1
Mean square errors of the different estimators
for situation 1 at different values of h

Estimator	$(1/h)$					**Average**	**Rank**
	(1/2)	**(1/3)**	**(1/4)**	**(1/5)**	**(1/10)**		
$\bar{y}^*$	0.346	0.4036	0.4613	0.5189	0.8073	0.5074	5th
t_2	0.6335	0.6821	0.7309	0.7797	0.8024	0.7257	6th
t_3 , $t_{n5(1)}$	0.9621	1.0335	1.1048	1.1763	1.5333	1.1619	7th
t_4	0.3310	0.3868	0.4426	0.4985	0.7776	0.4873	2nd
t_5	1.8282	1.8771	1.9252	1.9764	2.2187	1.9651	8th
t_6	2.4632	2.5346	2.6059	2.6773	3.0343	2.6631	9h
t_7	0.3371	0.3941	0.4511	0.508	0.7927	0.4966	3th
$t_{n2(1)}$, $t_{n4(1)}$,	0.3309	0.3866	0.4421	0.4975	0.7729	0.4860	1st
$t_{n3(1)}$	0.3454	0.4031	0.46086	0.5186	0.8071	0.5070	4th

Discussion for Situation 1

After empirical studies it has been observed that estimators $t_{n2(1)}$ and $t_{n4(1)}$ are equal at first position at different values of h whereas estimators t_4 is at 2nd position, t_7 , $t_{n3(1)}$ and $\bar{y}^*$ are at 3rd, 4th and 5th positions respectively. It has been noted that most of our proposed estimators give minimum mean square errors as compared to existing estimators.

Table 5.2
Mean square errors of the different estimators for situation 2 at different values of h

Estimator	$(1/h)$					Average	Rank
	(1/2)	(1/3)	(1/4)	(1/5)	(1/10)		
$\bar{y}^*$	0.3459	0.4036	0.4613	0.5189	0.8073	0.5074	5th
t_8	0.6421	0.6998	0.7575	0.8152	1.1035	0.80362	8th
t_9	0.9484	1.006	1.0637	1.1214	1.4097	1.10984	9th
t_{10}	0.3329	0.3905	0.4483	0.5059	0.5043	0.44438	1st
t_{11}	0.3379	0.3955	0.4532	0.5109	0.7992	0.49934	4th
$t_{n2(2)}$	0.385	0.4427	0.5004	0.558	0.8464	0.5465	6th
$t_{n1(2)}$	0.3309	0.3905	0.4482	0.5059	0.7943	0.49396	2nd
$t_{n4(2)}$	0.3891	0.4468	0.5045	0.5621	0.8505	0.5506	7th
$t_{n5(2)}$	0.3328	0.3905	0.4482	0.5059	0.7942	0.49432	3rd

Discussion for Situation 2

After numerical calculation it has been observed that estimator t_{10} are at 1st position and after that $t_{n1(2)}$ with the different values of h, is at 2nd place. After these estimators $t_{n3(2)}$ is at 3rd position. It has been noted that most of our proposed estimators give minimum mean square errors as compared to existing estimators.

ACKNOWLEDGEMENT

We are thankful to the referees for their useful comments for improving the text of this paper.

REFERENCES

1. Cochran, W.G. (1977). *Sampling Techniques*. 3rd ed., John Wiley and Sons, New York.
2. Glynn, R.J., Laird, N.M. and Rubin, D.B. (1993). Multiple imputation in mixture models for non-ignorable non response with follow-ups. *J. Amer. Statist. Assoc.*, 88(423), 984-993.
3. Hansen, M.H. and Hurwitz, W.N. (1946). The problem of non response in sample surveys. *J. Amer. Statist. Assoc.*, 41, 517-529.
4. Ismail, M., Shahbaz, M.Q. and Hanif, M. (2011). A General class of estimator of population mean in the presence of non-response. *Pak. J. Statist.*, 27(4), 467-476.
5. Ismail, M., Shahbaz, M.Q. and Hanif, M. (2013). A New Cost Effective Estimator in the Presence of Non-Response for Two-Phase Sampling. *Jurnal Teknologi.,* 63(2), 1-4.

6. Khare, B.B. and Srivastava, S. (1993). Estimation of population mean using auxiliary character in presence of non-response. *Nat. Acad. Sci. Lett. India*, 16, 111-114.
7. Khare, B.B. and Srivastava, S. (1995). Study of conventional and alternative two-phase sampling ratio, product and regression estimators in presence of non-response. *Proc. Indian. Nat. Sci. Acad.*, 65, 195-203.
8. Khare, B.B. and Srivastava, S. (1997). Transformed ratio type estimators for the population mean in presence of non-response. *Comm. Statist. Theory Methods*, 26, 1779-1791.
9. Khare, B.B. and Srivastava, S. (2010). Generalized two phase estimators for the population mean in the presence of non response. *Aligarh J. Statist.*, 30, 39-54.
10. Little, R.J.A. (1982). Models for non response in sample surveys. *J. Amer. Statist. Assoc*., 77, 237-250.
11. Naik, V.D. and Gupta, P.C. (1991). A general class of estimators for estimating population mean using auxiliary information, *Metrika*, 38, 11-17.
12. Rao, P.S.R.S. (1986). Ratio estimation with sub sampling the non respondents. *Surv. Methodology*, 12(2), 217-230.
13. Rao, P.S.R.S. (1990). Regression estimators with sub sampling of non-respondents, In-*Data Quality Control*. Theory and Pragmatics, (Gunar E. Liepins and V.R.R. Uppuluri, eds.) Marcel Dekker, New York, 191-208.
14. Singh, H.P. and Kumar, S. (2008a). Estimation of mean in presence of non response of non response using two-phase sampling scheme. *Statist. Pap*. DOI 10.1007/s00362-008-040-5.
15. Singh, H.P. and Kumar, S. (2008b). A regression approach to the estimation of finite population mean in presence of non-response. *Aust. N.Z.J. Statist.*, 50(4), 395-408.
16. Singh, H.P. and Kumar, S. (2008c). A general family of estimators of finite population ratio, product and mean using two phase sampling scheme in presence of non-response. *J. Statist. Theo. and Prac.*, 2(4), 677-692.
17. Singh, H.P. and Kumar, S. (2009). A general procedure for estimating the population mean in presence of non-response under double sampling using auxiliary information. *SORT*, 33(1), 71-83.
18. Singh, H.P. and Kumar, S. (2011). Combination of regression and ratio estimate in presence of non response. *Braz. J. Statist. Assoc.*, 25(2), 205-217.
19. Tabasum, R. and Khan, I.A. (2004). Double sampling for ratio estimation with non-response. *J. Ind. Soc. Agricult. Statist*., 58(3), 300-306.
20. Tabasum, R. and Khan, I.A. (2006). Double sampling ratio estimator for the population mean in presence of non response. *Assam Statist. Rev*., 20(1), 73-83.
21. Tripathi, T.P., Das A.K. and Khare, B.B. (1994).Use of auxiliary information in sample surveys. *A review. Aligarh J. Stat.*, 14, 79-134.
22. Tripathi, T.P. and Khare, B.B. (1997). Estimation of mean vector in the presence of non-response. *Commun. Statist. Theor. Meth*., 26(9), 2255-2269.