Journal of Applied and Natural Science 9 (3): 1666 -1675 (2017)

Estimation of population mean in two– stage sampling under a deterministic response mechanism in the presence of non-response

Monika Devi[*] and B. V. S. Sisodia

Department of Agriculture Statistics, N.D. University of Agriculture & Technology, Faizabad, (UP), INDIA
[*]Corresponding author. E-mail: mscagstats@gmail.com

Received: June 24, 2016; Revised received: April 17, 2017; Accepted: August 9, 2017

Abstract: In the present paper, we have considered the problem of estimation of population mean in the presence of non-response under two-stage sampling. Two different models of non-response with deterministic response mechanism have been discussed in the paper. The estimators under two non-response models have been developed by using Hansen and Hurwitz (1946) technique. The expressions for the variances and estimates of variance of these estimators have been derived. The optimum values of sample sizes have been obtained by considering a suitable cost function for a fixed variance. A limited simulation study has been carried out to examine the magnitude of percent relative loss (% RL) in standard error due to non-response. An empirical study with the real populations has also been carried out to assess the % RL in standard error due to non-response.

Keywords. Population mean, Two-stage sampling, Non-response, Percent relative loss

INTRODUCTION

It is assumed, for each sampling design, that the true values of the variables of interest could be made available for the elements of the population under consideration. However, this may not be true particularly for large scale surveys. Errors can occur at almost every stage of planning and execution of a large scale survey. These errors may be attributed to various causes right from the beginning stage, when the survey is planned and designed, to the final stage when the data are collected, processed and analyzed. For large or medium scale surveys we are often faced with the scenario that the sampling frame of ultimate stage units is not available and the cost of construction of the frame is very high. Sometimes the population elements are scattered over a wide area resulting in a widely scattered sample. Therefore, not only the cost of enumeration of units in such a sample may be very high, the supervision of field work may also be very difficult. For such situations, two-stage or multi-stage sampling designs are very effective. It is also the case that, in many human surveys, information is not obtained from all the units in surveys.

Colombo (1992), Anido and Valdes (2000) and Biemer and Link (2006) proposed the call-back approach to reduce the nonresponse bias. The problem of non-response persists even after call-backs. The estimates obtained from incomplete sample data become biased. Hansen and Hurwitz (1946) proposed a technique for adjusting for non-response to address the problem of bias. The technique consists of selecting a sub-sample of non-respondents and the data are collected through specialized efforts from the non-respondents so as to obtain an estimate of non-responding units in the population**.** Tripathi and Khare (1997) extended the sub-sampling of non-respondents approach to multivariate case. Okafor (2001, 2005) further extended the approach in the context of element sampling and two-stage sampling respectively on two successive occasions. Chhikara and Sud (2009) used the sub-sampling of non-respondents approach for estimation of population and domain totals in the context of item non-response. Various authors have used auxiliary information to improve the estimate by developing ratio and regression type estimators in the presence of non-response. Notably among them are Rao (1986), Khare and Srivastava (1993), Khare and Sinha (2009), Sodipo (2010), Singh and Kumar (2011), Monika Devi et al (2014) etc. Okafor and Lee (2000), Kumar and Viswanathaiah (2014) and many others extended the approach to double sampling for ratio and regression estimation. Al Baghal and Lynn (2015), Anderson *et. al*. (2015), Burton *et. al.* (2015), Fahimi *et. al.* (2015) and Matei and Ranalli (2015) proposed different approaches to deal with the problem of non-response.

Most of the work is however, dedicated to uni-stage sampling in the presence of non-response. The present work is therefore initiated to develop the methodology for estimation of population mean in two-stage sampling under non-response with the following objectives:

To develop an efficient estimator of population mean in two-stage sampling under Deterministic Response

ISSN : 0974-9411 (Print), 2231-5209 (Online)

Mechanism.

To carry out empirical study with real data to examine the performance of the estimators.

MATERIALS AND METHODS

The estimators for estimation of population mean in two-stage sampling in the presence of non-response have been developed under two non-response models. It is assumed that the non-response is deterministic.

Consider that a finite population U consists of N primary stage units (psus) labeled 1 through N, and each psu comprises of M second stage units (ssus). Let y_{ij} be the value of study character y pertaining to j-*th* ssu in the i - *th* psu, $i = 1,2,\ldots\ldots,N, j = 1,2\ldots\ldots, M$,.

The objective is to estimate the population mean

$$\bar{Y} = \frac{1}{MN}\sum_{i=1}^{N}\sum_{j=1}^{M} y_{ij}$$

We state the first non-response model referred to as Situation-1 as follows:

Situation 1: It is assumed that the psu(s) are divided into two strata, i.e. (i) first stratum consisting of N_1 psu(s) from where we do not get response at all, and (ii) second stratum consisting of N_2 psu(s) from where we do get partial responses from ssu(s), such that $N = N_1 + N_2$. A random sample of n psus is drawn from N by simple random sampling without replacement (SRSWOR). From each selected psu a sample of m ssus from M ssu(s) is drawn by SRSWOR. Let there be complete non-response from n_1 psus. In the n_2 psu ($n_1 + n_2 = n$) m_{i1} ssus respond while m_{i2} ssus do not respond, $m_{i1} + m_{i1} = m$. A sub-sample of h_{i2} units is selected by srswor from m_{i2} and data are collected through specialized efforts. Further, a sub-sample of h_1 psus is drawn out of n_1 psus and data are collected through specialized efforts on each of m ssus in the selected h_1 psus. Let $n_1 = f_1 h_1$ and $m_{i2} = f_{i2}h_{i2}$, i = 1,2, ….n_2. It is further assumed that in n_2 psu(s) there are M_{i1} responding and M_{i2} non-responding ssu units such that $M_{i1} + M_{i2} = M$.

First, we define the following:

$$\bar{y}_{im} = \frac{1}{m}\sum_{j=1}^{m} y_{ij},$$

sample mean in i - *th* psu ($i \in n_1$)

$$\bar{y}_{m_{i1}} = \frac{1}{m_{i1}}\sum_{j=1}^{m_{i1}} y_{ij},$$

sample mean in i - *th* psu ($i \in n_2$) psu corresponding to the responsing ssu(s).

$$\bar{y}_{h_{i2}} = \frac{1}{h_{i2}}\sum_{j=1}^{h_{i2}} y_{ij},$$

sample mean in i - *th* psu ($i \in n_2$) corresponding to sub -sample of non-responsing ssu(s).

$$S_b^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(\bar{Y}_{iM} - \bar{Y}\right)^2, \qquad \bar{Y}_{iM} = \frac{1}{M}\sum_{j=1}^{M} Y_{ij},$$

$$S_{bN_1}^2 = \left(\frac{1}{N-1}\right)\sum_{i=1}^{N_1}\left(\bar{Y}_{iM} - \bar{Y}_{N_1}\right)^2, \qquad f_1 = \frac{n_1}{h_1},$$

$$\bar{Y}_{N_1} = \frac{1}{N_1}\sum_{i=1}^{N_1} Y_{iM}$$

$$S_{iM}^2 = \frac{1}{M-1}\sum_{j=1}^{M}\left(Y_{ij} - \bar{Y}_{iM}\right)^2,$$

$$S_{M_{i2}}^2 = \frac{1}{(M_{i2}-1)}\sum_{j=1}^{M_{i2}}\left(Y_{ij} - \bar{Y}_{M_{i2}}\right)^2, \qquad \bar{Y}_{M_{i2}} = \frac{1}{M_{i2}}\sum_{j=1}^{M_{i2}} Y_{ij},$$

$$s_b^2 = \frac{1}{n-1}\left[\frac{n_1}{h_1}\sum_{i=1}^{h_1}\bar{y}_{im}^2 + \sum_{i=1}^{n_2}\frac{1}{m^2}\left(m_{i1}\bar{y}_{m_{i1}} + m_{i2}\bar{y}_{h_{i2}}\right)^2 - n\bar{y}_1^2\right]$$

$$s_{b_{h_1}}^2 = \left(\frac{1}{h_1-1}\right)\sum_{i=1}^{h_1}\left(\bar{y}_{im} - \bar{y}_{h_1}\right)^2,$$

$$\bar{y}_{h_1} = \frac{1}{h_1}\sum_{i=1}^{h_1}\bar{y}_{im}, \qquad s_{1im}^2 = \frac{1}{(m-1)}\left(\sum_{j=1}^{m} y_{ij}^2 - m\bar{y}_{im}^2\right)$$ where,

$$\bar{y}_{im} = \frac{1}{m}\sum_{j=1}^{m} y_{ij}$$

$$s_{im}^2 = \frac{1}{(m-1)}\left(\sum_{j=1}^{m_{i1}} y_{ij}^2 + \frac{m_{i2}}{h_{i2}}\sum_{j=1}^{h_{i2}} y_{ij}^2 - m\bar{y}_{im}^2\right),$$

$$s_{h_{i2}}^2 = \frac{1}{(h_{i2}-1)}\sum_{j-1}^{h_{i2}}\left(y_{ij} - \bar{y}_{h_{i2}}\right)^2, \qquad \bar{y}_{h_{i2}} = \frac{1}{h_{i2}}\sum_{j=1}^{h_{i2}} y_{ij}$$

We, now, state and prove the following theorem.

Theorem 2.1: An unbiased estimator of $\bar{Y}$ is given by

$$\bar{y}_1 = \frac{1}{n}\left[\frac{n_1}{h_1}\sum_{i=1}^{h_1}\bar{y}_{im} + \sum_{i=1}^{n_2}\frac{m_{i1}\bar{y}_{m_{i1}} + m_{i2}\bar{y}_{h_{i2}}}{m}\right] \quad \ldots\ (1)$$

with variance of $\bar{y}_1$ as

$$V(\bar{y}_1) = \left(\frac{1}{n} - \frac{1}{N}\right)S_b^2 + \frac{1}{Nn}\left(\frac{1}{m} - \frac{1}{M}\right)\left\{f_1\sum_{i=1}^{N_1}S_{iM}^2 + \sum_{i=1}^{N_2}S_{iM}^2\right\} + \frac{1}{Nn}\sum_{i=1}^{N_2}\frac{M_{i2}}{Mm}(f_{i2}-1)S_{M_{i2}}^2 + \frac{N_1}{Nn}(f_1-1)S_{bN_1}^2 \quad \ldots\ (2)$$

and unbiased variance estimator as

$$\hat{V}(\bar{y}_1) = \left(\frac{1}{n} - \frac{1}{N}\right)\hat{S}_b^2 + \frac{1}{n^2}\left(\frac{1}{m} - \frac{1}{M}\right)\left\{f_1\frac{n_1}{h_1}\sum_{i=1}^{h_1}s_{1im}^2 + \sum_{i=1}^{n_2}s_{im}^2\right\}$$
$$+ \frac{1}{n^2}\frac{(M-1)}{(m-1)}\frac{m}{M}\sum_{i=1}^{n_2}\frac{m_{i2}}{m^2}(f_{i2}-1)s_{h_{i2}}^2 + \frac{n_1}{n^2}(f_1-1)\left\{s_{bh_1}^2 - \frac{1}{h_1}\sum_{i=1}^{h_1}\left(\frac{1}{m} - \frac{1}{M}\right)s_{1im}^2\right\} \quad (3)$$

Proof: By definition,

$$E(\bar{y}_1) = E_1E_2E_3E_4\frac{1}{n}\left[\frac{n_1}{h_1}\sum_{i=1}^{h_1}\bar{y}_{im} + \sum_{i=1}^{n_2}\frac{m_{i1}\bar{y}_{m_{i1}} + m_{i2}\bar{y}_{h_{i2}}}{m}\right]$$

$$= E_1 E_2 E_3 \frac{1}{n}\left[\frac{n_1}{h_1}\sum_{i=1}^{h_1}\bar{y}_{im} + \sum_{i=1}^{n_2}\bar{y}_{im}\right]$$

$$= E_1 \frac{1}{n}\left[\sum_{i=1}^{n_1}\bar{Y}_{iM} + \sum_{i=1}^{n_2}\bar{Y}_{iM}\right]$$

$$= E_1 \frac{1}{n}\left[\sum_{i=1}^{n}\bar{Y}_{iM}\right] = \frac{1}{N}\sum_{i=1}^{N}\bar{Y}_{iM} = \bar{Y}$$

Where E_4 represents conditional expectation of all possible samples of size h_{i2} drawn from a sample of size m_{i2} , E_3 represents conditional expectation of all possible samples of size m drawn from M, E_2 refers to conditional expectation arising out of selection of all possible samples of size h_1 drawn from n_1 while E_1 refers to expectation arising out of all possible samples of size n drawn from a population of size N.

To obtain the variance, we proceed as follows:

By definition,

$$V(\bar{y}_1) = V_1E_2E_3E_4(\bar{y}_1) + E_1V_2E_3E_4(\bar{y}_1) + E_1E_2V_3E_4(\bar{y}_1) + E_1E_2E_3V_4(\bar{y}_1)$$

where V_1 , V_2 , V_3 , V_4 are defined similarly as E_1 , E_2 , E_3 , E_4 .

where,

$$V_1E_2E_3E_4(\bar{y}_1) = \left(\frac{1}{n} - \frac{1}{N}\right)S_b^2$$

$$E_1V_2E_3E_4(\bar{y}_1) = \frac{N_1}{Nn}(f_1 - 1)S_{bN_1}^2$$

$$E_1E_2V_3E_4(\bar{y}_1) = \frac{1}{nN}\left(\frac{1}{m} - \frac{1}{M}\right)\left\{f_1\sum_{i=1}^{N_1}S_{iM}^2 + \sum_{i=1}^{N_2}S_{iM}^2\right\}$$

$$E_1E_2E_3V_4(\bar{y}_1) = \frac{1}{nN}\sum_{i=1}^{N_2}\frac{M_{i2}}{Mm}(f_{i2} - 1)S_{M_{i2}}^2$$

Thus, by adding all the terms we obtain the required result.

$$V(\bar{y}_1) = \left(\frac{1}{n} - \frac{1}{N}\right)S_b^2 + \frac{1}{Nn}\left(\frac{1}{m} - \frac{1}{M}\right)\left\{f_1\sum_{i=1}^{N_1}S_{iM}^2 + \sum_{i=1}^{N_2}S_{iM}^2\right\} + \frac{1}{Nn}\sum_{i=1}^{N_2}\frac{M_{i2}}{Mm}(f_{i2} - 1)S_{M_{i2}}^2 + \frac{N_1}{Nn}(f_1 - 1)S_{bN_1}^2$$

Taking the expectation and simplifying we get,

$$E_1E_2E_3E_4E_5E_6E_7\left(s_b^2\right) = S_b^2 + \frac{(n - f_1)}{N(n-1)}\sum_{i=1}^{N_1}\left(\frac{1}{m} - \frac{1}{M}\right)S_{iM}^2 + \frac{1}{N}\sum_{i=1}^{N_2}\left(\frac{1}{m} - \frac{1}{M}\right)S_{iM}^2$$

$$+ \frac{1}{N}\sum_{i=1}^{N_2}\frac{M_{i2}}{Mm}(f_{i2} - 1)S_{M_{i2}}^2 - \frac{N_1}{N(n-1)}(f_1 - 1)S_{b_{N_1}}^2$$

$$E\left(s_{im}^2\right) = S_{iM}^2 - \frac{M_{i2}}{M(m-1)}(f_{i2} - 1)S_{M_{i2}}^2$$

$$E\left(s_{1im}^2\right) = S_{iM}^2$$ and

$$E\left(s_{b_{h_1}}^2\right) = \frac{1}{N_1}\sum_{i=1}^{N_1}\left(\frac{1}{m} - \frac{1}{M}\right)S_{iM}^2 + S_{b_{N_1}}^2$$

Where E_7 represents conditional expectation of all possible samples of size h_{i2} drawn from a sample of size m_{i2}, E_6 represents conditional expectation of all possible samples of size m_{i1}, m_{i2} respectively drawn from M_{i1}, M_{i2} , respectively by keeping m_{i1}, m_{i2}, fixed. Here M_{i1}, M_{i2} denote the number of responding and non -responding units in the population, E_5 refers to conditional expectation arising out of randomness $m_{i1}, m_{i2}, M_{i1}, M_{i1}, M_{i2}$,. refers to conditional expectation of all possible samples of size m drawn from M, E_3 refers to conditional expectation of all possible samples of size h_1 drawn from n_1, E_2, refers to expectation arising out of all possible samples of size n_1, n_2 ,drawn from N_1,N_2 keeping n_1, n_2,fixed while E_1 refers to expectation arising out of randomness of n_1, n_2.

Let,

$$\hat{S}_b^2 = s_b^2 - \frac{1}{n}\frac{(n - f_1)}{(n-1)}\sum_{i=1}^{n_1}\left(\frac{1}{m} - \frac{1}{M}\right)\hat{S}_{iM}^2 - \frac{1}{n}\sum_{i=1}^{n_2}\left(\frac{1}{m} - \frac{1}{M}\right)\hat{S}_{im}^2 - \frac{1}{n}\sum_{i=1}^{n_2}\frac{m_{i2}}{m^2}(f_{i2} - 1)\hat{S}_{M_{i2}}^2 + \frac{n_1}{n(n-1)}(f_1 - 1)\hat{S}_{bN_1}^2$$

$$\hat{S}_{iM}^2 = s_{im}^2 + \frac{m_{i2}}{m(m-1)}(f_{i2} - 1)s_{h_{i2}}^2$$,

for n_2 $\hat{S}_{iM}^2 = s_{1im}^2$,

for n_1

$$\hat{S}_{bN_1}^2 = s_{bh_1}^2 - \frac{1}{h_1}\sum_{i=1}^{h_1}\left(\frac{1}{m} - \frac{1}{M}\right)s_{1im}^2$$ and $\hat{S}_{M_{i2}}^2 = s_{h_{i2}}^2$

Substituting the estimated values in the variance expression (2.2) we get the required estimate of $V(\bar{y}_1)$.

To determine the optimum values of *n, m*, and f_{i2} by minimizing the expected cost for a fixed variance, we use the relation $m_{i2} = h_{i2} f_{i2}$, i=1,2,…., n_2. To achieve this, consider the following cost function

$$C = C_1(h_1 + n_2) + C_2\sum_{i=1}^{n_2}m_{i1} + C_3\left(\sum_{i=1}^{n_2}h_{i2} + h_1 m\right)$$

where, *C*: Total cost; C_1: Per psu travel cost; C_2: Cost per ssu for collecting the information on the study character in the first attempt; C_3: Cost per ssu for collecting the information by expensive method after the first attempt has failed for obtaining information

It is envisaged that C_3 will be higher than C_1 and substantially higher than C_2.

The expected cost in this case is,

$$C^* = E(C) = \frac{n}{N}\left[C_1\left(\frac{N_1}{f_1} + N_2\right) + C_2\sum_{i=1}^{N_2}\frac{mM_{i1}}{M} + C_3\left(\sum_{i=1}^{N_2}\frac{nM_{i2}}{Mf_{i2}} + \frac{N_1}{f_1}m\right)\right]$$

To minimize the expected cost subject to fixed variance consider the function.

$$\phi = C^* + \lambda\{V(\bar{y}_1) - V_0\}$$

During optimization we have substituted f_2 in place of fi_2 for simplicity in calculations. To overcome the problem arising due to simultaneous minimization of n, m, f_1 and f_2, we assume that $n_2 = f_1 h_2$ for making the calculations simple. Thus minimization gives the optimum values as ,

$$n_{opt} = \frac{k}{\left(V_0 + \frac{S_b^2}{N}\right)}$$

$$m_{opt} = \frac{-b_1 + \sqrt{b_1^2 - 4a_1c_1}}{2a_1},$$

$$f_{2opt} = \frac{-b_2 + \sqrt{b_2^2 - 4a_2c_2}}{2a_2}$$

where,

$$k = S_b^2 + \frac{1}{N}\left(\frac{1}{m} - \frac{1}{M}\right)\left\{f_1\sum_{i=1}^{N_1} S_{iM}^2 + \sum_{i=1}^{N_2} S_{iM}^2\right\} + \frac{1}{N}\sum_{i=1}^{N_2}\frac{M_{i2}}{Mm}(f_2 - 1)S_{M_{i2}}^2 + \frac{N_1}{N}(f_1 - 1)S_{bN_1}^2$$

$$a_1 = C_3\sum_{i=1}^{N_2} M_{i2}\left[N_1 S_{bN_1}^2 - \sum_{i=1}^{N_1}\frac{1}{M}S_{iM}^2\right]$$

$$b_1 = C_3\left[\sum_{i=1}^{N_2} M_{i2}\sum_{i=1}^{N_1} S_{iM}^2 - N_1\sum_{i=1}^{N_2} M_{i2}S_{M_{i2}}^2\right],$$

$$c_1 = -C_1 N_1\sum_{i=1}^{N_2} M_{i2}S_{M_{i2}}^2$$

$$a_2 = C_2\sum_{i=1}^{N_2} M_{i2}S_{iM}^2\sum_{i=1}^{N_2}\frac{M_{i1}}{M},$$

$$c_2 = C_3\sum_{i=1}^{N_2} M_{i2}\left(\sum_{i=1}^{N_2}\frac{M_{i2}}{M}S_{M_{i2}}^2 - \sum_{i=1}^{N_2} S_{iM}^2\right)$$

$$b_2 = C_3\left[\sum_{i=1}^{N_2} M_{i2}\sum_{i=1}^{N_2}\frac{M_{i2}}{M}S_{iM_2}^2 - \left(N_1 + \sum_{i=1}^{N_2}\frac{M_{i2}}{M}\right)\sum_{i=1}^{N_2} M_{i2}S_{M_{i2}}^2 + \sum_{i=1}^{N_2} M_{i2}\sum_{i=1}^{N_1} S_{iM}^2\right]$$

Special case of Situation 1: Here we consider the case that a sample of n psus is drawn from N, within each selected psu a sample of m ssus is drawn by srswor design. This sample is divided into two parts n_1 and n_2. Let there be complete non-response in the n_1 psus, $n_1 + n_2 = n$. Let there be no non-response in n_2 psus, further a sub-sample of h_1 psus is drawn out of n_1 psus and data are collected through specialized efforts on each of m ssus in the selected h_1 psus. Let $n_1 = f_1 h_1$. Assume $N = N_1 + N_2$ where N_1 and N_2 are the number of psus in the population representing the two non-response categories considered here.

In this context, we state and prove the following theorem.

Theorem 2.2: An unbiased estimator of $\bar{Y}$is given by

$$\bar{y}_2 = \frac{1}{n}\left[\frac{n_1}{h_1}\sum_{i=1}^{h_1}\bar{y}_{im} + \sum_{i=1}^{n_2}\bar{y}_{im}\right] \quad \dots (4)$$

with variance

$$V(\bar{y}_2) = \left(\frac{1}{n} - \frac{1}{N}\right)S_b^2 + \frac{1}{Nn}\left(\frac{1}{m} - \frac{1}{M}\right)\left\{f_1\sum_{i=1}^{N_1} S_{iM}^2 + \sum_{i=1}^{N_2} S_{iM}^2\right\} + \frac{N_1}{Nn}(f_1 - 1)S_{bN_1}^2 \quad \dots (5)$$

An unbiased estimator of variance is,

$$\hat{V}(\bar{y}_2) = \left(\frac{1}{n} - \frac{1}{N}\right)\hat{S}_b^2 + \frac{1}{n^2}\left(\frac{1}{m} - \frac{1}{M}\right)\left\{f_1\frac{n_1}{h_1}\sum_{i=1}^{h_1} s_{1im}^2 + \sum_{i=1}^{n_2} s_{1im}^2\right\} + \frac{n_1}{n^2}(f_1 - 1)\left\{s_{bh_1}^2 - \frac{1}{h_1}\sum_{i=1}^{h_1}\left(\frac{1}{m} - \frac{1}{M}\right)s_{1im}^2\right\} \quad \dots (6)$$

where,

$$s_b^2 = \frac{1}{n-1}\left[\frac{n_1}{h_1}\sum_{i=1}^{h_1}\bar{y}_{im}^2 + \sum_{i=1}^{n_2}\bar{y}_{im}^2 - n\bar{y}_2^2\right],$$

while rest of the terms are defined earlier.

Proof:

$$E(\bar{y}_2) = E_1E_2E_3\frac{1}{n}\left[\frac{n_1}{h_1}\sum_{i=1}^{h_1}\bar{y}_{im} + \sum_{i=1}^{n_2}\bar{y}_{im}\right]$$

$$= E_1\frac{1}{n}\left[\sum_{i=1}^{n_1}\bar{Y}_{iM} + \sum_{i=1}^{n_2}\bar{Y}_{iM}\right]$$

$$= E_1\left[\frac{1}{n}\sum_{i=1}^{n}\bar{Y}_{iM}\right]$$

$$= \frac{1}{n}\sum_{i=1}^{n}E(\bar{Y}_{iM}) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{N}\sum_{i=1}^{N}\bar{Y}_{iM} = \bar{Y}$$

where, E_3 represents conditional expectation of all possible samples of size m drawn from M, E_2 refers to conditional expectation arising out of selection of all possible samples of size h_1 drawn from n_1 while E_1 refers to expectation arising out of all possible samples of size n drawn from a population of size N.

To obtain the variance we proceed as follows:

By definition,

$$V(\bar{y}_2) = V_1E_2E_3(\bar{y}_2) + E_1V_2E_3(\bar{y}_2) + E_1E_2V_3(\bar{y}_2)$$

where,

$$V_1E_2E_3(\bar{y}_2) = \left(\frac{1}{n} - \frac{1}{N}\right)S_b^2$$

$$E_1V_2E_3(\bar{y}_2) = \frac{N_1}{nN}(f_1 - 1)S_{bN_1}^2$$

$$E_1E_2V_3(\bar{y}_2) = \frac{1}{Nn}\left(\frac{1}{m} - \frac{1}{M}\right)\left[f_1\sum_{i=1}^{N_1} S_{iM}^2 + \sum_{i=1}^{N_2} S_{iM}^2\right]$$

Thus, by adding all the terms we obtain the required result.

$$V(\bar{y}_2) = \left(\frac{1}{n} - \frac{1}{N}\right)S_b^2 + \frac{1}{Nn}\left(\frac{1}{m} - \frac{1}{M}\right)\left\{f_1\sum_{i=1}^{N_1} S_{iM}^2 + \sum_{i=1}^{N_2} S_{iM}^2\right\} + \frac{N_1}{Nn}(f_1 - 1)S_{bN_1}^2$$

Taking the expectation and simplifying we get,

$$E_1E_2E_3E_4(s_b^2) = S_b^2 + \frac{(n - f_1)n}{N(n-1)}\sum_{i=1}^{N_1}\left(\frac{1}{m} - \frac{1}{M}\right)S_{iM}^2 + \frac{1}{N}\sum_{i=1}^{N_2}\left(\frac{1}{m} - \frac{1}{M}\right)S_{iM}^2 - \frac{N_1}{N(n-1)}(f_1 - 1)S_{bN_1}^2$$

$$E(s_{bh_1}^2) = \frac{1}{N_1}\sum_{i=1}^{N_1}\left(\frac{1}{m} - \frac{1}{M}\right)S_{iM}^2 + S_{b_{N_1}}^2$$,

$$E(s_{1im}^2) = S_{iM}^2$$, as defined earlier.

Where E_4 refers to conditional expectation of all possible samples of size m drawn from M, E_3 refers to conditional expectation of all possible samples of size h_1 drawn from n_1, E_2 refers to expectation arising out of all possible samples of size n_1, n_2 ,drawn from N_1, N_2, keeping n_1, n_2 fixed while E_1 refers to expectation arising out of randomness of n_1, n_2. Let,

$$\hat{S}_b^2 = s_b^2 - \frac{(n - f_1)n}{n(n-1)}\sum_{i=1}^{n_1}\left(\frac{1}{m} - \frac{1}{M}\right)\hat{S}_{iM}^2 - \frac{1}{n}\sum_{i=1}^{n_2}\left(\frac{1}{m} - \frac{1}{M}\right)\hat{S}_{iM}^2 + \frac{n_1}{n(n-1)}(f_1 - 1)\hat{S}_{bN_1}^2$$

$$\hat{S}_{iM}^2 = s_{1im}^2$$ and

$$\hat{S}_{bN_1}^2 = s_{bh_1}^2 - \frac{1}{h_1}\sum_{i=1}^{h_1}\left(\frac{1}{m} - \frac{1}{M}\right)s_{1im}^2$$

Substituting the estimated values in the variance expression (2.5) we get the required estimate of $V(\bar{y}_2)$.
To determine the optimum values of n, m, and f_1 we proceed as earlier, i.e. minimization of expected cost subject to fixed variance. The optimum values are determined in the same way as the previous estimator by minimizing the expected cost with respect to fixed variance.
The relevant cost function in this case is,

$$C = C_1h_1 + C_2n_2m + C_3h_1m$$

The various costs are defined here are same as defined earlier. The expected cost is,

$$C^* = E(C) = \frac{n}{N}\left[C_1\frac{N_1}{f_1} + C_2N_2m + C_3m\frac{N_1}{f_1}\right]$$

Consider the function:

$$\phi = C^* + \lambda\{V(\bar{y}_2) - V_0\}$$

The minimization gives the optimum values as

$$n_{opt} = \frac{k}{\left(V_0 + \frac{S_b^2}{N}\right)} \qquad m_{opt} = \frac{-b_3 + \sqrt{b_3^2 - 4a_3c_3}}{2a_3}$$

$$f_{1opt} = \frac{-b_4 + \sqrt{b_4^2 - 4a_4c_4}}{2a_4}$$

where,

$$k = S_b^2 + \frac{1}{N}\left(\frac{1}{m} - \frac{1}{M}\right)\left\{f_1\sum_{i=1}^{N_1} S_{iM}^2 + \sum_{i=1}^{N_2} S_{iM}^2\right\} + \frac{N_1}{N}(f_1 - 1)S_{bN_1}^2$$

$$a_3 = \left(C_2N_2 + C_3\frac{N_1}{f_1}\right)\left(N_1S_{bN_1}^2 - \frac{1}{M}\sum_{i=1}^{N_1} S_{iM}^2\right)$$

$$c_4 = (C_1N_1 + C_3mN_1)\frac{1}{m^2}\sum_{i=1}^{N_2} S_{iM}^2$$

$$b_3 = \left(C_2N_2 + C_3\frac{N_1}{f_1}\right)\sum_{i=1}^{N_1} S_{iM}^2 - C_3\frac{N_1}{f_1^2}\left(\sum_{i=1}^{N_1} f_1S_{iM}^2 + \sum_{i=1}^{N_2} S_{iM}^2\right)$$

$$c_3 = -C_1\frac{N_1}{f_1^2}\left(\sum_{i=1}^{N_1} f_1S_{iM}^2 + \sum_{i=1}^{N_2} S_{iM}^2\right)$$

$$a_4 = C_2N_2\left\{\left(\frac{1}{m} - \frac{1}{M}\right)\sum_{i=1}^{N_1} S_{iM}^2 + N_1S_{b_{N_1}^2}^2\right\}$$

$$b_4 = C_3N_1\left\{\left(\frac{1}{m} - \frac{1}{M}\right)\sum_{i=1}^{N_1} S_{iM}^2 + N_1S_{b_{N_1}^2}^2\right\} - (C_1N_1 + C_3mN_1)\frac{1}{m^2}\sum_{i=1}^{N_1} S_{iM}^2$$

Control situation: The following estimator was also considered for efficiency comparison purpose. Here we assume that srswor sample of n psus is selected from N and within each selected psu a sample of m ssus is selected from M ssus. Data are collected through specialized efforts to obtain complete response, i.e. there is no non-response. Then we give the following Theorem 2.3
Theorem 2.3 The estimator

$$\bar{y}_{nm} = \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m} y_{ij} = \frac{1}{n}\sum_{i=1}^{n}\bar{y}_{im} \quad \ldots (7)$$

is unbiased of $\bar{Y}$, ith variance,

$$V(\bar{y}_{nm}) = \left(\frac{1}{n}-\frac{1}{N}\right)S_b^2 + \frac{1}{Nn}\sum_{i=1}^{N}\left(\frac{1}{m}-\frac{1}{M}\right)S_{iM}^2 \quad \ldots (8)$$

where S^2_b and S^2_{iM} are already defined, and unbiased estimator of variance,

$$\hat{V}(\bar{y}_{nm}) = \left(\frac{1}{n}-\frac{1}{N}\right)S_b^2 + \frac{1}{Nn}\sum_{i=1}^{n}\left(\frac{1}{m}-\frac{1}{M}\right)S_{1iM}^2 \quad \ldots (9)$$

where

$$s_b^2 = \frac{1}{(n-1)}\sum_{i=1}^{n}(\bar{y}_{im}-\bar{y})^2,$$

$$s_{1im}^2 = \frac{1}{(m-1)}\left(\sum_{j=1}^{m} y_{ij}^2 - m\bar{y}_{im}^2\right)$$

Proof: The proof of unbiasedness of the given estimator and its variance and unbiased variance estimator can be found in Cochran (1997), pp. 277-278.
The cost function in this case is, $C = C_{1n} + C_{3nm}$
where, C, C_1, C_3, have been defined earlier.
To obtain optimum values of n and $m$we minimize the cost by fixing the variance. The optimum values are as follows,

$$n_{opt} = \frac{S_b^2 + \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{m}-\frac{1}{M}\right)S_{iM}^2}{\left(V_0 + \frac{S_b^2}{N}\right)}$$ and

$$m_{opt} = \sqrt{\frac{C_1\sum_{i=1}^{N}\frac{S_{iM}^2}{N}}{C_3\left(S_b^2 - \frac{1}{MN}\sum_{i=1}^{N}S_{iM}^2\right)}}$$

Simulation study: A limited simulation study has been conducted with real data to examine the relative merits of the proposed estimators $\bar{y}_1$ and $\bar{y}_2$ and in comparison to the usual estimators $\bar{y}_{nm}$ (without non-response) in two-stage sampling. A design based simulation based on real data is carried out. The following criterion was used for assessing the relative performance of these estimators:
The percent relative root mean square error (RRMSE), which is defined as,

$$\%RRMSE(\hat{\theta}) = \sqrt{\frac{1}{L}\sum_{i=1}^{L}\left(\frac{\hat{\theta}_1-\theta}{\theta}\right)^2}\times 100$$

Here, $\hat{\theta}_i$ is the value of estimator $\hat{\theta}$ ($\bar{y}_1$, $\bar{y}_2$ & $\bar{y}_{nm}$) of θ ($\bar{Y}$)in the i-th (i=1,……,L=1000) simulation run.
Further, the percent relative loss in standard error in $\bar{y}_1$ & $\bar{y}_2$ due to non-response as compared to standard error of $\bar{y}_{nm}$ has been computed as follows:

$$\%RL(\bar{y}_1) = \frac{\sqrt{MSE(\bar{y}_{nm})}-\sqrt{MSE(\bar{y}_1)}}{\sqrt{MSE(\bar{y}_1)}}\times 100$$

and

$$\%RL(\bar{y}_2) = \frac{\sqrt{MSE(\bar{y}_{nm})}-\sqrt{MSE(\bar{y}_2)}}{\sqrt{MSE(\bar{y}_2)}}\times 100,$$

where $\sqrt{MSE(\bar{y}_{nm})}$, $\sqrt{MSE(\bar{y}_1)}$ and $\sqrt{MSE(\bar{y}_2)}$ are the empirical root MSE of the estimator $\bar{y}_{nm}$(usual two-stage estimator without non-response), $\bar{y}_1$ & $\bar{y}_2$ (our proposed estimators), respectively.

Here,

$$RootMSE(\bar{y}_{nm}) = \left(\sqrt{\frac{1}{L}\sum_{i=1}^{L}(\bar{y}_{nm(i)}-\theta)^2}\right)$$

$$RootMSE(\bar{y}_1) = \left(\sqrt{\frac{1}{L}\sum_{i=1}^{L}(\bar{y}_{1(i)}-\theta)^2}\right),$$ and

$$RootMSE(\bar{y}_2) = \left(\sqrt{\frac{1}{L}\sum_{i=1}^{L}(\bar{y}_{2(i)}-\theta)^2}\right)$$

with $\bar{y}_{1(i)}$ and $\bar{y}_{2(i)}$ are the values of our proposed estimators $\bar{y}_1$ and $\bar{y}_2$ in the simulation run i(i=1,….,L).
In design based simulation study with real data, we used the data given in Appendix-B: The MU284 population (Sarndal et al (1992)). From the Appendix-B, the 1985 population (in thousands) with respect to municipalities has been considered as study variable. There are in all 284 municipalities. To form the psu(s), the first 15 municipalities constitute the first psu, and then next 15 municipalities form the second psu and so on. Therefore, we get in all 18 psu(s) each consisting of 15 ssu(s). In our study we used 270 municipalities out of 284 and remaining last 14 municipalities were left. 1000 independent random samples of size 7 psu(s) out of 18 are drawn by using simple random sampling without replacement. For each selected psu 6 ssu(s) m = 6 out of total 15 ssu(s) are drawn by using simple

random sampling without replacement. We also consider that 18 psu(s) are divided into two classes, i.e. N_1 = 6 and N_2 = 6, where N_1 constitutes the class of complete non-responding psu(s) and N_2 constitutes the partially responding class/complete response class of psu(s), i.e. ($N_1 + N_2$ = N). Again we assume that that the sample of size n=7 is also divided into two parts: i.e. n_1 = 3 and n_2 = 3 , which comes from complete non -response and partially response classes, respectively. From n_1 = 3, we further draw a subsample of size 2 (h_1 = 2) and we make use of each of the values of m ssu (s) in the selected h_1 psu(s). In the n_2 = 4 psu(s), m_{i1} = 3 ssu(s) respond while m_{i2} = 3 ssu(s) do not respond. A subsample of h_{12} = 2 units is selected by SRSWOR from m_{i2}. Here $n_1 = f_1h_1$ = 3 and $m_{i2}=h_{i2}f_{i2}=3$. We computed the values of $\bar{y}_{nm}$, $\bar{y}_1$, and $\bar{y}_2$ from one thousand samples. The true population mean $\bar{Y}$ has been computed to be 29.90. The percent relative root mean square error (%RRMSE), the percent of relative loss in standard error (%RL) have been computed for our proposed estimators $\bar{y}_1$ and $\bar{y}_2$.

These computed values are presented in the Table 1.

It is obvious that making bias adjustment in case of non-response in sample surveys, we loose efficiency of the estimators to some extent (Hansen & Hurwitz 1946). It is evident from the results of the Table 1 that the %RRMSE of the $\bar{y}_1$ & $\bar{y}_2$ have increased to about 24 percent in comparison to about 21 percent of $\bar{y}_{nm}$ (without non-response). The percent relative loss in standard error has been found more (10.36%) in case of $\bar{y}_1$ as compared to that of $\bar{y}_2$ (7.80) which is on the expected line because more sampling error is expected in situation-I than situation-II.

Empirical study: An empirical study using some real populations has also been carried out to examine the loss in standard error of the estimate due to non-response. Four populations viz. (i) P75 (1975 population (in thousands)), (ii) P85 (1985 population (in thousands)), (iii) RMT85 (Revenues from the 1985 municipal taxation (in millions of kronor)) and (iv) REV84 (Real estate values according to 1984 assessment (in millions of kronor)) have been considered from the Appendix-B of Sarndal *et al*. (1992). There are in all 284 municipalities. To form the psu(s), the first 15 municipalities constitute the first psu, and then next 15 municipalities form the second psu and so on. Therefore, we get in all 18 psu(s) each consisting of 15 ssu(s). In our study, we used 270 municipalities out of 284 and remaining last 14 municipalities were left.

For each population, we have considered N =18, N_1 =6, N_2 =12, M = 15, M_{i1} = 9, M_{i2} = 6, n = 6, n_1 = 3, n_1 = 3, m = 8, $m_{i1=}$ 4,5, $m_{i2=}$ 4,3, h_{i2} = 2, $i = 1,2.$ $f_{i2} = \dfrac{m_{i2}}{h_{i2}} = 2\,\&\,1.5$, $f_1 = \dfrac{n_1}{h_1} = \dfrac{3}{2} = 1.5$

The different population parameters involved in the variances of the estimators have been computed and are presented in the Table 2.

The variance of control estimator, i.e. $V(\bar{y}_{nm})$ has been computed for each population. The variance of the proposed estimator $\bar{y}_1$ for each population has been computed for f_{i2} = 1.5 and 2.00. Similarly, the variance of the proposed estimator $\bar{y}_2$ for each population has been computed for f_1 = 1.5 .

The percent relative loss in standard error due to non-response over complete response with respect to the proposed estimators has been computed as follows

$$\%RL = \frac{\sqrt{V(\bar{y}_{nm})} - \sqrt{V(\bar{y}_i)}}{\sqrt{V(\bar{y}_i)}} \times 100,$$

These results are summarized in the Table 3 to 5.

It is obvious that making bias adjustment in case of non-response in sample surveys, we loose efficiency of the estimators to some extent (Hansen and Hurwitz 1946). It is evident from the results of the Table 3 and 4.3 that the percent relative loss in standard error has been found more in case of f_{i2} = 2 as compared to that of f_{i2} = 1.5 for each population for the proposed estimator $\bar{y}_1$ over $\bar{y}_{nm}$. Since f_{i2} is the reciprocal of fraction of sampled non-response ssu′s so in case of f_{i2} = 2 we have more loss in percent relative standard error. So as the f_{i2} increases percent relative loss in standard error also increases and it decrease with the decreasing value of f_{i2} for each population.

The Table 5 shows the percent relative loss in standard error in case of proposed estimator $\bar{y}_2$ over $\bar{y}_{nm}$ for

$$f_1 = \frac{n_1}{h_1} = 1.5$$

.

It is obvious that %RL will be less for proposed estimator $\bar{y}_2$ for each population as compared to that of % RL of estimator $\bar{y}_1$ since estimator $\bar{y}_1$ consists two parts, one having complete non-response and, second of partial response whereas in case of estimator $\bar{y}_1$ one part has complete non-response and other has complete response.

It may also be noted that sampling $\bar{X}$ rates and sub-sampling rates of non-respondent ssu(s) from non-respondents ssu(s) in selected psu(s) in the aforesaid simulation and empirical studies have already been at high side $\bar{X}$ because of limitation of data. The decrease in $\bar{X}$ these rates would certainly increase the loss in efficiency.

RESULTS AND DISCUSSION

In order to make effective use of available sources various sampling technique have been developed from time to time which provide estimators of population characteristics of interest with high precision, reduced cost and above all will have the operational feasibility

Table 1. Percent Relative Root Mean Square Error (% RRMSE) and percent Relative Loss in Standard Error (% RL) over $\bar{y}_{nm}$.

Estimators	%RRMSE	%RL
$\bar{y}_{nm}$	21.40	-
$\bar{y}_1$	24.39	10.40
$\bar{y}_2$	23.71	7.80

on the closely related variable (auxiliary information) is utilized judiciously in the estimation procedure. Rao (1986) suggested a ratio estimator for the population mean $\bar{Y}$, when the population mean of the auxiliary variable is known. Khare and Srivastava (1993) suggested a ratio-product type exponential estimator for estimating the finite population mean in the presence of non-response in different situations viz. (i) population mean is known, and (ii) population mean is unknown. The expressions of biases and mean squared

Table 2. Values of the different population parameters

Population	S_b^2	$\sum_{i=1}^{N} S_{iM}^2$	$\sum_{i=1}^{N_1} S_{iM}^2$	$\sum_{i=1}^{N_2} S_{iM}^2$	$\sum_{i=1}^{N_2} S_{M_{i2}}^2$	$S_{bN_1}^2$
P75	246.5995	51645.0286	7406.514	44238.51	8519.5	90.2834
P85	249.5263	48784.1048	7049.476	41734.63	71087.9	91.3898
RMT85	2728.361	552067.9	257809.6	294258.3	137812	1038.076
REV84	534705.4	65061004	25919883	39141121	22849808	519388.7

Table 3. Variance & standard error of the $\bar{y}_{nm}$ and $\bar{y}_1$ and % RL when f_{i2}=1.5

S.No.	Description	Population			
		P75	P85	RMT85	REV84
1	$V(\bar{y}_{nm})$	55.295	54.075	601.336	94552.685
2	$\sqrt{V(\bar{y}_{nm})}$	7.436	7.354	24.522	307.494
3	$V(\bar{y}_1)$	61.775	74.973	1410.043	122150.977
4	$\sqrt{V(\bar{y}_1)}$	7.859	8.659	37.550	349.501
5	$\%RL$	5.390	15.073	34.697	12.019

Table 4. Variance and standard error of the $\bar{y}_{nm}$ and $\bar{y}_1$ and % RL when f_{i2}=2.

S.No.	Description	Population			
		P75	P85	RMT85	REV84
1	$V(\bar{y}_{nm})$	55.295	54.075	601.336	94552.685
2	$\sqrt{V(\bar{y}_{nm})}$	7.436	7.354	24.522	307.494
3	$V(\bar{y}_1)$	63.747	91.428	1411.903	126558.733
4	$\sqrt{V(\bar{y}_1)}$	7.984	9.562	37.575	355.751
5	$\%RL$	6.865	23.095	34.739	13.565

and practical applicability. Various authors have used auxiliary information to improve the estimate by developing ratio and regression type estimators in the presence of non-response. It may lead to much improvement in precision of estimation if the information

errors of the proposed estimators, up to the first order of approximation, have also been obtained. The results obtained were depicted with the help of numerical illustration. Singh and Kumar (2011) provided a Combination of regression and ratio estimators in presence of

Table 5. Variance and standard error of the $\bar{y}_{nm}$ and $\bar{y}_2$ and % RL when f_1=1.5

S.No.	Description	Population			
		P75	P85	RMT85	REV84
1	$V(\bar{y}_{nm})$	55.295	54.075	601.336	94552.685
2	$\sqrt{V(\bar{y}_{nm})}$	7.436	7.354	24.522	307.494
3	$V(\bar{y}_2)$	59.803	58.517	1408.182	115980.118
4	$\sqrt{V(\bar{y}_2)}$	7.733	7.650	37.526	340.559
5	$\%RL$	3.843	3.870	34.653	9.709

nonresponse. They addressed the problem of estimating the population mean of the study variable *y* using information on two auxiliary variables *x* and *z* in presence of nonresponse. Two classes of combined regression and ratio estimators were defined in two different situations along with their properties. Many others extended the approach to double sampling for ratio and regression estimation.

Most of the work is however, dedicated to uni-stage sampling in the presence of non-response. Recently, Sud et al (2012) have made an attempt to develop the estimators of population mean in two-stage sampling in the presence of non-response. They have considered three types of non-response models. There are two more possible response models which they have not considered.

In what follows, we have considered these two non-response models for the development of the estimation of population mean in two-stage sampling design in the present paper. Here, we have considered the deterministic response mechanism. In situation-1, it is assumed that the psu(s) are divided into two strata, i.e. (i) first stratum consisting of N_1 psu(s) from where we do not get response at all, and (ii) second stratum consisting of N_2 psu(s) from where we do get partial responses from ssu(s), such that $N = N_1 + N_2$. And in an special case of situation-1, we assumed that the psu(s) are divided into two strata, i.e. (i) first stratum consisting of N_1 psu(s) from where we do not get response at all, and (ii) second stratum consisting of N_2 psu(s) from where we do get complete responses from ssu(s), such that $N = N_1 + N_2$. The expressions for the variances and estimates of variance of these estimators have been derived. The optimum values of sample sizes have been obtained by considering a suitable cost function for a fixed variance.

Conclusion

The study of two-stage estimators of population mean under non-response has been presented. The optimum values of sample sizes have been obtained by considering a suitable cost function for a fixed variance. It is evident from the results of the Table 1 that the % RRMSE of the proposed estimators have increased to about 24 percent in comparison to about 21 percent of usual two-stage estimator (without non-response). The percent relative loss in standard error has been found more in situation-I as compared to that of situation-II which is on the expected line because more sampling error is expected in situation-I than situation-II. An empirical study using some real populations has also been carried out to examine the loss in standard error of the estimate due to non-response. It is also observed that the percentage relative efficiency decreases with increase in non-response. Since size of sub-sample is the reciprocal of fraction of sampled non-response ssu′s so the percent relative loss in standard error will be more in case of small size sub-sample size as compared to that of a larger sub-sample and this has been supported by our empirical study results.

ACKNOWLEDGEMENTS

The present paper is a part of Ph.D. thesis of the first author, who is grateful to the DST, Govt. of India, for providing DST Fellowship (Inspire Fellowship) vide grant No. IF-120491 for carrying out present investigation.

REFERENCES

Al Baghal, T. and Lynn, P. (2015). Using motivational statements in web-instrument design to reduce item-missing rates in a mixed-mode context. *Public Opinion Quarterly*, 79: 568 -579.

Anderson, Me., Henrikson, N., King, D. and Ulrich, K. (2015). Measuring the effects of operational designs on response rates and nonresponse bias. *The American Association for Public Opinion Research*, 70th Annual Conference. Retrieved from www.websm.org/db/12/17894/Web_Survey_Bibliography/Measurng_the_Effects_of_Operational_Designs_on_Response_Rates_and_Nonresponse_Bias/menu=1&lst=&q=search_1_111111_-1 & qdb=12 & qsort =1

Anido, C. and Valdes, T. (2000). An Iterative Estimation Procedure for Probit-type Nonresponse Models in Surveys with Call Backs. *Sociedad de Estadistica e Invetigacion Operativa*, 9(1): 233-253.

Biemer, P., and M. Link (2006). "A Latent Call-Back Model for Nonresponse," in 17th International Workshop on *Household Survey Nonresponse*, Omaha, Nebraska, USA.

Burton, J., Jaeckle, A. and Lynn, P. (2015). Going online with a face-to-face household panel: Effects of a mixed mode design on item and unit non-response. *Survey Research Methods*, 9(1): 57-70.

Chhikara, R. S. and Sud, U. C. (2009). Estimation of population and domain totals under two phase sampling in the presence of non-response. *Journal of the Indian Society of Agricultural Statistics*, 63(3): 297-304.

Cochran, W.G. (1997). Sampling Techniques, 3rd Edition. John Wiley & Sons, Inc., New York.

Colombo, R. (1992). Using Call-Backs to Adjust for Non-response Bias. Pages: 269-277. *Elsevier*, North Holland.

Fahimi, M., F. M. Barlas, R. K. Thomas and N. Buttermore (2015). Scientific Surveys Based on Incomplete Sampling Frames and High Rates of Nonresponse. *Survey Practice*, 8(5).

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

Khare, B. B. and Srivastava, S. (1993). Estimation of population mean using auxiliary character in presence of non-response. *Nat. Acad. Sc. Letters, India*, 16(3): 111-114.

Khare, B.B. and Sinha, R.R. (2009). On class of estimators for population mean using multi-auxiliary characters in the presence of non-response. *Statistics in Transition-new series.* 10(1): 3-14.

Matei, A. and Ranalli, M.G. (2015). Dealing with non-ignorable nonresponse in survey sampling: A latent modeling approach. *Survey Methodology*, 41(1), 145-164.

Monika, D., Sisodia, B.V.S. and Dube, L.K. (2014). Estimation of finite population mean under non-response using auxiliary information. *International Journal of Agricultural and Statistical Sciences.* 10(1): 127-131.

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

Okafor, F. C. (2001). Treatment of non-response in successive sampling, *Statistica*, 61(2):195-204.

Okafor, F. C. (2005). Sub-sampling the non-respondents in two-stage sampling over two successive occasions. *Journal of the Indian Statistical Association*, 43(1): 33-49.

Rao, P. S. R. S. (1986). Ratio estimation with sub sampling the nonrespondents. *Survey Methodology*, 12(2): 217–230.

S. Kumar and M. Viswanathaiah (2014). Population mean estimation with sub sampling the non-respondents using two phase sampling. *Journal of Modern Applied Statistical Methods*, 13(1): 187-198.

Sarndal, C. E., Swensson, B. and Wretman, J. (1992). Model Assisted Survey Sampling, Springer Verlag, New York.

Singh, H. P. and Kumar, S. (2011). Combination of regression and ratio estimate in presence of nonresponse. *Braz. J. Probab. Stat.*, 25(2): 205-217.

Sodipo, A.A. (2010). Difference- type and regression- type estimators for the population mean based on post stratification and subsampling of the non-respondents. *European Journal of Scientific Research.* 43(4): 445-451.

Sud, U. C., Aditya, Kaustav, Chandra H. and Parsad, Rajender (2012). Two stage sampling for estimation of population mean with sub-sampling of non-respondents, *Journal of the Indian Society of Agricultural Statistics*, 66(3): 447-457.

Tripathi, T. P. and Khare, B. B. (1997). Estimation of mean vector in presence of nonresponse. *Communications in Statistics-Theory and Methods*, 26(9): 2255 - 2269.