Ratio-cum-product estimation in presence of non-response in successive sampling

A. V. GARCÍA LUENGO

Abstract

The problem of estimating the finite population mean on the samples selected over two occasions has been considered, when there is non-response (i) on both the occasions, (ii) only on the second occasion for both matched and unmatched portions of the sample, and (iii) only on the second occasion for unmatched portion of the sample. For the case when two auxiliary variables are positively and negatively correlated with the study variable, a double sampling ratio-cum-product estimate from the matched portion of the sample is presented. Expressions for optimum matching fraction and of the combined estimate have been derived. The gain in efficiency, the sample sizes and the cost function of the combined estimators over the direct estimate using no information gathered on the first occasion is computed. The comparison between the proposed strategy with other estimators is also carried out. An empirical study is also included for illustration.

1   INTRODUCTION

Usually, in many national sample surveys, information collected regularly on the same population from one period to the next period. In such repetitive surveys, the next sampling procedure may be used:

Performing a partial replacement of units from one occasion to another, sampling on successive occasions, also called rotation sampling when the units are constructed in the number of stages in which they are to become part of the sample, as it happens with the EPA-Spanish survey of working population-which are performed quarterly, and most of the family surveys carried out by the INE-Spanish Statistics Institute. (see García Luengo, A. V. and Oña, I. (2010), pp. 2).

In surveys, the problem of non-response often happens due to the refusal of the subject, absenteeism and sometimes due to the lack of information. Hansen and Hurwitz (1946) considered the problem of non-response while estimating the population mean by taking a sub-sample from the non-respondents group with the help of some extra efforts and an estimator was proposed by combining the

DE
G
DE GRUYTER
OPEN

information available from response and non-response groups. The method has been applied by Okafor and Lee (2000) in double sampling for ratio and regression estimation. Further, Choudhary, Bathla and Sud (2004) and Singh and Kumar (2009, 2010) used the Hansen and Hurwitz technique for estimating the population mean for current occasion in the context of sampling on two occasions in presence of non-response. (see Singh, Kumar and Bhougal (2011), pp. 592).

In repetitive surveys for estimating the same characteristic at different points of time, it is possible to use the information collected on previous occasion to improve upon the usual estimators for the current period by using ratio and product methods of estimation. Generally the correlation between such observations is positive but situations do exist when the value of study variable at two successive occasions are negatively correlated e.g. amount of fertilizers contains required by the soil in two successive seasons and some biological phenomenon in medical sciences. (see Gupta (1979)). The study of environmental issues also involves negatively correlated characteristics.

It is to be mentioned that when several auxiliary variables are positively correlated with the study variable, Artés, García Luengo and Oña (2001) developed an optimum estimate of the population mean on the second occasion by combining (i) a double-sampling multivariate ratio estimate from the matched portion of the sample and (ii) a sample mean based on a random sample from the unmatched portion of the sample on the second occasion assuming that total response from the sample units are available on both the occasions.

And when several auxiliary variables are negatively correlated with the study variable, Artés and García Luengo (2005) developed an optimum estimate of the population mean on the second occasion by combining (i) a double-sampling multivariate product estimate from the matched portion of the sample and (ii) a sample mean based on a random sample from the unmatched portion of the sample on the second occasion assuming that total response from the sample units are available on both the occasions.

Theory of successive sampling had been developed by Artés, Rueda and Arcos (1999) by using a double sampling ratio-cum-product estimate from the matched portion of the sample, for the case when two auxiliary variables are positively and negatively correlated with the study variable.

The present study looks at the situation where some sample units do not supply the necessary information. The estimate due to Artés, Rueda and Arcos (1999) of

population mean on second occasion has been adjusted for the case when two auxiliary variables are positively and negatively correlated with the study variable, for unit non-response using Hansen and Hurwitz (1946) technique.

In this conditions, estimators of the population mean on second occasion, by using a double sampling ratio-cum-product estimate from the matched portion of the sample, when there is non-response, have been proposed. Three different possible cases:

(i) on both the occasions.

(ii) only on the second occasion for both matched and unmatched portions of the sample.

(iii) only on the second occasion for unmatched portion of the sample.

Expressions for optimum matching fraction and of the combined estimate have been derived. The gain in efficiency, the sample sizes and the cost function of the combined estimators over the direct estimate using no information gathered on the first occasion is computed. The comparison between the proposed strategy with other estimators is also carried out. An empirical study is also included for illustration.

2 THE TECHNIQUE

In some of practical problems, we may have information from a antecedent occasion on the value of two auxiliary variables x and z where one of the auxiliary variables may be value of the study variable y on the previous occasion. Let y be positively and negatively correlated with the auxiliary variables x and z respectively, whose population means are unknown. Suppose a simple random sample of size n be drawn on both occasions from a population of size N. For the sake of simplicity, we assume that the population size N is large enough so that finite population correction (fpc) terms can be ignored. We assume that the population can be divided into two classes, those who will respond at the first attempt and those who will not. Let the sizes of these two classes be N_1 and N_2, respectively. Let on the first occasion, schedules through mail are sent to n units selected by simple random sampling. On the second occasion, a simple random sample of $m = np$ units, for $0 < p < 1$, is retained while an independent sample of $u = nq = n - m$ units, for $q = 1 - p$, is selected (unmatched with the first occasion). We assume that in the unmatched portion of the sample on two occasions, u_1 units respond and u_2 units do not. Similarly, in the matched portion m_1 units respond and m_2 units do not.

Let $m_{h_2} = (m_2/v); (v > 1)$ denotes the size of the subsample drawn from the non-response class from the matched portion of the sample on the two occasions for

collecting information through personal interview. Similarly, denote by $u_{h_2} = (u_2/v); (v > 1)$ the size of the subsample drawn from the non-response class in the unmatched portion of the sample on the two occasions. (see Singh and Priyanka (2007)).

Let

$$\bar{x}_m^* = \frac{m_1 \bar{x}_{m_1} + m_2 \bar{x}_{m_{h_2}}}{m} :$$

be the Hansen and Hurwitz estimator of the population mean $\bar{X}$ of the auxiliary variable x for the matched portion of the sample on the first occasion,

$$\bar{x}_u^* = \frac{u_1 \bar{x}_{u_1} + u_2 \bar{x}_{u_{h_2}}}{u} :$$

be the Hansen and Hurwitz estimator of the population mean $\bar{X}$ of the auxiliary variable x for the unmatched portion of the sample on the first occasion, where

$$\bar{x}_{m_1} = \frac{1}{m_1} \sum_{t=1}^{m_1} x_t; \quad \bar{x}_{m_{h_2}} = \frac{1}{m_{h_2}} \sum_{l=1}^{m_{h_2}} x_l; \quad \bar{x}_{u_1} = \frac{1}{u_1} \sum_{\alpha=1}^{u_1} x_\alpha; \quad \bar{x}_{u_{h_2}} = \frac{1}{u_{h_2}} \sum_{\beta=1}^{u_{h_2}} x_\beta,$$

$$\bar{z}_m^* = \frac{m_1 \bar{z}_{m_1} + m_2 \bar{z}_{m_{h_2}}}{m} :$$

be the Hansen and Hurwitz estimator of the population mean $\bar{Z}$ of the auxiliary variable z for the matched portion of the sample on the first occasion,

$$\bar{z}_u^* = \frac{u_1 \bar{z}_{u_1} + u_2 \bar{z}_{u_{h_2}}}{u} :$$

be the Hansen and Hurwitz estimator of the population mean $\bar{Z}$ of the auxiliary variable z for the unmatched portion of the sample on the first occasion, where

$$\bar{z}_{m_1} = \frac{1}{m_1} \sum_{t=1}^{m_1} z_t; \quad \bar{z}_{m_{h_2}} = \frac{1}{m_{h_2}} \sum_{l=1}^{m_{h_2}} z_l; \quad \bar{z}_{u_1} = \frac{1}{u_1} \sum_{\alpha=1}^{u_1} z_\alpha; \quad \bar{z}_{u_{h_2}} = \frac{1}{u_{h_2}} \sum_{\beta=1}^{u_{h_2}} z_\beta,$$

$$\bar{y}_m^* = \frac{m_1 \bar{y}_{m_1} + m_2 \bar{y}_{m_{h_2}}}{m} :$$

be the Hansen and Hurwitz estimator of the population mean $\bar{Y}$ of the study variable y for the matched portion of the sample on the second occasion,

$$\bar{y}_u^* = \frac{u_1 \bar{y}_{u_1} + u_2 \bar{y}_{u_{h_2}}}{u} :$$

be the Hansen and Hurwitz estimator of the population mean $\bar{Y}$ of the study variable y for the unmatched portion of the sample on the second occasion, where

$$\bar{y}_{m_1} = \frac{1}{m_1} \sum_{t=1}^{m_1} y_t; \quad \bar{y}_{m_{h_2}} = \frac{1}{m_{h_2}} \sum_{l=1}^{m_{h_2}} y_l; \quad \bar{y}_{u_1} = \frac{1}{u_1} \sum_{\alpha=1}^{u_1} y_\alpha; \quad \bar{y}_{u_{h_2}} = \frac{1}{u_{h_2}} \sum_{\beta=1}^{u_{h_2}} y_\beta,$$

$\bar{Y}$ is the population mean of the study variable y, on the second occasion,

$\bar{Y}_{(2)}$ is the population mean of the study variable y, pertaining to the non-response class,

$\bar{X}$ is the population mean of the auxiliary variable x, on the first occasion,

$\bar{X}_{(2)}$ is the population mean of the study variable x, pertaining to the non-response class,

$\bar{Z}$ is the population mean of the auxiliary variable z, on the first occasion,

$\bar{Z}_{(2)}$ is the population mean of the study variable z, pertaining to the non-response class,

$\bar{x}$ is the sample mean of the auxiliary variable x on the first occasion based on the large sample of size n,

$\bar{z}$ is the sample mean of the auxiliary variable z on the first occasion based on the large sample of size n,

S_y^2 is the population variance of the study variable y,

$S_{y(2)}^2$ is the population variance of the study variable y pertaining to the non-response class,

S_x^2 is the population variance of the auxiliary variable x,

$S_{x(2)}^2$ is the population variance of the auxiliary variable x pertaining to the non-response class,

S_z^2 is the population variance of the auxiliary variable z,

$S_{z(2)}^2$ is the population variance of the auxiliary variable z pertaining to the non-response class,

ρ_{xy} is the correlation coefficient between study variable y and the auxiliary variable x for matched portion,

$\rho_{xy(2)}$ is the correlation coefficient between study variable y and the auxiliary variable x pertaining to the non-response class for matched portion,

ρ_{yz} is the correlation coefficient between study variable y and the auxiliary variable z for matched portion,

$\rho_{yz(2)}$ is the correlation coefficient between study variable y and the auxiliary variable z pertaining to the non-response class for matched portion,

ρ_{xz} is the correlation coefficient between the auxiliary variable x and the auxiliary variable z for matched portion,

$\rho_{xz(2)}$ is the correlation coefficient between the auxiliary variable x and the auxiliary variable z pertaining to the non-response class for matched portion,

$p = \frac{m}{n}$ matched proportion of the sample at the second occasion,

$q = (1 - p)$ unmatched proportion of the sample at the second occasion.

$\theta = W_2(\nu - 1)$

where $W_2 = (N_2/N)$ and $\nu = (m_2/m_{h_2}) = (u_2/u_{h_2})$

$C_x = \frac{S_x}{\bar{X}}, \quad C_y = \frac{S_y}{\bar{Y}}, \quad C_z = \frac{S_z}{\bar{Z}}$

$C_{x(2)} = \frac{S_{x(2)}}{\bar{X}_{(2)}}, \quad C_{y(2)} = \frac{S_{y(2)}}{\bar{Y}_{(2)}}, \quad C_{z(2)} = \frac{S_{z(2)}}{\bar{Z}_{(2)}}$

$\Delta_1 = \frac{C_x}{C_y}, \quad \Delta_2 = \frac{C_z}{C_y}$

$\Delta_{1(2)} = \frac{C_{x(2)}}{C_{y(2)}}, \quad \Delta_{2(2)} = \frac{C_{z(2)}}{C_{y(2)}}$

3 THE MULTIVARIATE RATIO-CUM-PRODUCT METHOD OF ESTIMATING THE MEAN ON THE SECOND OCCASION IN PRESENCE OF NON-RESPONSE

The unmatched (u units) and matched (m units) portions of the current (second) occasion sample in presence of non-response provide two independent Hansen and Hurwitz (1946) estimates: $\bar{y}_m^*$ and $\bar{y}_u^*$ for the population mean $\bar{Y}$ of the study variable y on the second occasion.

Thus, agree with Singh (1967) and proceeding as Cochran (1977), for the matched portion an improved estimate of population mean $\bar{Y}$ of the study variable y may be obtained using a double sampling ratio-cum-product estimate as:

$$\bar{t}_m^* = \bar{y}_m^* \frac{\bar{x}}{\bar{x}_m^*} \frac{\bar{z}_m^*}{\bar{z}} \tag{1}$$

$$Var(\bar{t}_m^*) = \frac{S_y^2}{m} + \frac{S_y^2}{n} \frac{u}{n-u} \left(\Delta_1^2 + \Delta_2^2 - 2\rho_{xy}\Delta_1 + 2\rho_{yz}\Delta_2 - 2\rho_{xz}\Delta_1\Delta_2 \right) +$$

$$+ \frac{\theta S_{y(2)}^2}{m} + \frac{\theta S_{y(2)}^2}{m} \left(\Delta_{1(2)}^2 + \Delta_{2(2)}^2 - 2\rho_{xy(2)}\Delta_{1(2)} + 2\rho_{yz(2)}\Delta_{2(2)} - 2\rho_{xz(2)}\Delta_{1(2)}\Delta_{2(2)} \right)$$

(see Artés, Rueda and Arcos (1999), pp. 22).

$$Var(\bar{t}_m^*) = \frac{S_y^2}{m} + \frac{S_y^2}{n} \frac{u}{n-u} \left(Z_2 - Z_1 - 2\rho_{xz}\Delta_1\Delta_2 \right) + \tag{2}$$

$$+\frac{\theta S_{y(2)}^2}{m} + \frac{\theta S_{y(2)}^2}{m}\left(Z_{2(2)} - Z_{1(2)} - 2\rho_{xz(2)}\Delta_{1(2)}\Delta_{2(2)}\right)$$

$$\mathrm{Var}(\bar{t}_m^*) = \frac{S_y^2}{m}(1+qA) + \frac{\theta}{m}(1+A_{(2)})S_{y(2)}^2 \qquad (3)$$

where

$$A = Z_2 - Z_1 - 2\rho_{xz}\Delta_1\Delta_2, \quad A_{(2)} = Z_{2(2)} - Z_{1(2)} - 2\rho_{xz(2)}\Delta_{1(2)}\Delta_{2(2)}$$

$$Z_1 = \Delta_1\left(2\rho_{xy} - \Delta_1\right), \qquad Z_2 = \Delta_2\left(2\rho_{yz} + \Delta_2\right).$$

$$Z_{1(2)} = \Delta_{1(2)}\left(2\rho_{xy(2)} - \Delta_{1(2)}\right), \qquad Z_{2(2)} = \Delta_{2(2)}\left(2\rho_{yz(2)} + \Delta_{2(2)}\right)$$

If, however, the direct estimate based on the m sampling units, its variance would be

$$\frac{S_y^2}{m} + \frac{\theta}{m}S_{y(2)}^2$$

we obtain that $\bar{t}_m^*$ is more efficient than the direct estimate based on the m sampling units if

$$Z_2 - Z_1 - 2\rho_{xz}\Delta_1\Delta_2 \leq 0 \quad \text{and} \quad Z_{2(2)} - Z_{1(2)} - 2\rho_{xz(2)}\Delta_{1(2)}\Delta_{2(2)} \leq 0.$$

3.1 Estimation of the mean population for current occasion in the presence of non-response on both occasions

Hence, we suggest an estimate of the population mean on the second occasion in presence of non-response, t, by combining the two independent estimates, $\bar{t}_m^*$ and $\bar{t}_u^*$ with weights φ and $(1-\varphi)$.

$$t = \varphi\bar{t}_m^* + (1-\varphi)\bar{t}_u^* \qquad (4)$$

where $\bar{t}_m^*$ is defined at (1) and $\bar{t}_u^* = \bar{y}_u^*$ is the unmatched sample mean on second occasion estimating $\bar{Y}$ in presence of non-response. The variance of t is given by

$$\mathrm{Var}(t) = \varphi^2\,\mathrm{Var}(\bar{t}_m^*) + (1-\varphi)^2\,\mathrm{Var}(\bar{t}_u^*) \qquad (5)$$

where $\mathrm{Var}(\bar{t}_u^*)$ is given by

$$\mathrm{Var}(\bar{t}_u^*) = \frac{S_y^2}{u} + \frac{\theta}{u}S_{y(2)}^2 \qquad (6)$$

The best estimate of the mean $\bar{Y}$ on the second occasion in presence of non-response is obtained by using the values of φ that minimizes $\text{Var}(t)$. Differentiating (5) with respect to φ and equating to zero, we get the optimum value of φ as

$$\varphi_0 = \frac{\text{Var}(\bar{t}_u^*)}{\text{Var}(\bar{t}_u^*) + \text{Var}(\bar{t}_m^*)} = \frac{p\left(S_y^2 + \theta S_{y(2)}^2\right)}{\left(S_y^2 + \theta S_{y(2)}^2\right) + q\left(qAS_y^2 + \theta A_{(2)}S_{y(2)}^2\right)} \tag{7}$$

Substituting the value of φ_0 in place of φ in (4) yields the optimum estimator for the population mean $\bar{Y}$ of the study variable y on second occasion as

$$t_0 = \varphi_0 \bar{t}_m^* + (1 - \varphi_0)\bar{t}_u^*$$

Substituting the variances of $\bar{t}_m^*$ and $\bar{t}_u^*$ and the optimum value of φ from (7), we get the variance of t_0 as

$$\text{Var}(t_0) = \frac{\text{Var}(\bar{t}_m^*)\,\text{Var}(\bar{t}_u^*)}{\text{Var}(\bar{t}_m^*) + \text{Var}(\bar{t}_u^*)} = \left(\frac{S_y^2 + \theta S_{y(2)}^2}{n}\right) \frac{S_y^2(1 + qA) + \theta S_{y(2)}^2(1 + A_{(2)})}{S_y^2(1 + q^2A) + \theta S_{y(2)}^2(1 + qA_{(2)})}, \tag{8}$$

where $\theta = W_2(v - 1)$. If $\theta = 0$, i.e., there is no non-response, the variance of t_0 reduces to

$$\text{Var}(t_0) = \frac{S_y^2}{n} \frac{(1 + qA)}{(1 + q^2A)}$$

where t_0 is the usual estimator of the mean for the current occasion in the context of sampling on two occasions when there is complete response, which is same as obtained by Artes, Rueda and Arcos (1999).

Minimizing (8) with respect to u and equating to zero, we get the optimum matching fraction as

$$p_0 = \frac{\left(S_y^2(A + 1) + \theta S_{y(2)}^2(A_{(2)} + 1)\right) - \sqrt{\left(S_y^2 + \theta S_{y(2)}^2\right)\left(S_y^2(1 + A) + \theta S_{y(2)}^2(1 + A_{(2)})\right)}}{S_y^2 A}$$

or

$$q_0 = -\frac{\left(S_y^2 + \theta S_{y(2)}^2(A_{(2)} + 1)\right) - \sqrt{\left(S_y^2 + \theta S_{y(2)}^2\right)\left(S_y^2(1 + A) + \theta S_{y(2)}^2(1 + A_{(2)})\right)}}{S_y^2 A}$$

The minimum variance of t_0 is given by

$$\text{min.Var}(t_0) = \left(\frac{S_y^2 + \theta S_{y(2)}^2}{n}\right).$$

$$\cdot \left[\frac{S_y^2 A}{2\sqrt{\left(S_y^2 + \theta S_{y(2)}^2\right)\left(S_y^2(1+A) + \theta S_{y(2)}^2(1+A_{(2)})\right)} - \left(2S_y^2 + \theta S_{y(2)}^2(A_{(2)}+2)\right)}\right]$$

We can compute the gain in precision of the combined estimate t_0, obtained by using a double sampling ratio-cum-product estimate from the matched portion of the sample on the second occasion, over direct estimator $\bar{y}^*$ of population mean $\bar{Y}$ on second occasion, using no information collected on the first occasion,

$$G_1 = \frac{\text{Var}(\bar{y}^*)}{\text{Var}(t_0)} - 1 \times 100 = \frac{-p\left(qS_y^2 A + \theta S_{y(2)}^2 A_{(2)}\right)}{S_y^2(1+qA) + \theta S_{y(2)}^2(1+A_{(2)})} \times 100.$$

$$\text{where}\quad \text{Var}(\bar{y}^*) = \frac{S_y^2}{n} + \frac{\theta}{n}S_{y(2)}^2.$$

or

$$G_{opt(1)} = \frac{\text{Var}(\bar{y}^*)}{\text{min.Var}(t_0)} - 1 \times 100 =$$

$$= \frac{2\sqrt{(S_y^2 + \theta S_{y(2)}^2)(S_y^2(1+A) + \theta S_{y(2)}^2(1+A_{(2)}))} - (2+A_{(2)})\theta S_{y(2)}^2 - (2+A)S_y^2}{AS_y^2} \times 100.$$

Consider the special case where $-\rho_{xy} = \rho_{yz} = \rho_1$, $-\rho_{xy(2)} = \rho_{yz(2)} = \rho_{1(2)}$, $\rho_{xz} = \rho_0$, $\rho_{xz(2)} = \rho_{0(2)}$, $C_x = C_y = C_z = C$, $C_{x(2)} = C_{y(2)} = C_{z(2)} = C_{(2)}$ which gives an expression more simple for the variance

$$\text{Var}(t_0) = \left(\frac{S_y^2 + \theta S_{y(2)}^2}{n}\right)\frac{S_y^2(1+qA^*) + \theta S_{y(2)}^2(1+A_{(2)}^*)}{S_y^2(1+q^2A^*) + \theta S_{y(2)}^2(1+qA_{(2)}^*)}$$

where

$$A^* = 2 - 4\rho_1 - 2\rho_0, \quad A_{(2)}^* = 2 - 4\rho_{1(2)} - 2\rho_{0(2)}.$$

3.2 Estimation of the mean population for current occasion in the presence of non-response only on the second occasion for both matched and unmatched portions of the sample

When there is non-response only on the second occasion, we define an estimator for population mean $\bar{Y}$ of the study variable y on the current occasion is defined by

$$t_1 = \varphi \bar{t}_{1m}^* + (1 - \varphi) \bar{t}_u^* \tag{9}$$

where

$$\bar{t}_{1m}^* = \bar{y}_m^* \frac{\bar{x}}{\bar{x}_m} \frac{\bar{z}_m}{\bar{z}} \quad \text{and} \quad \bar{t}_u^* = \bar{y}_u^*$$

The variance of the estimator t_1 is given by

$$\text{Var}(t_1) = \varphi^2 \text{Var}(\bar{t}_{1m}^*) + (1 - \varphi)^2 \text{Var}(\bar{t}_u^*) \tag{10}$$

where

$$\text{Var}(\bar{t}_{1m}^*) = \frac{S_y^2}{m}(1 + qA) + \frac{\theta}{m} S_{y(2)}^2 \tag{11}$$

and $\text{Var}(\bar{t}_u^*)$ is given at (6).

Differentiating (10) with respect to φ and equating to zero, we get the optimum value of φ as

$$\varphi_0^{(1)} = \frac{\text{Var}(\bar{t}_u^*)}{\text{Var}(\bar{t}_u^*) + \text{Var}(\bar{t}_{1m}^*)} = \frac{p\left(S_y^2 + \theta S_{y(2)}^2\right)}{(1 + q^2 A) S_y^2 + \theta S_{y(2)}^2} \tag{12}$$

Substituting the value of φ_0 in place of φ in (9) yields the optimum estimator for the population mean $\bar{Y}$ of the study variable y on second occasion as

$$t_1^{(0)} = \varphi_0^{(1)} \bar{t}_{1m}^* + (1 - \varphi_0^{(1)}) \bar{t}_u^*$$

Substituting the variances of $\bar{t}_{1m}^*$ and $\bar{t}_u^*$ and the optimum value of φ from (12), we get the variance of $t_1^{(0)}$ as

$$\text{Var}\left(t_1^{(0)}\right) = \frac{\text{Var}(\bar{t}_{1m}^*) \text{Var}(\bar{t}_u^*)}{\text{Var}(\bar{t}_{1m}^*) + \text{Var}(\bar{t}_u^*)} = \left(\frac{S_y^2 + \theta S_{y(2)}^2}{n}\right) \frac{(1 + qA) S_y^2 + \theta S_{y(2)}^2}{(1 + q^2 A) S_y^2 + \theta S_{y(2)}^2} \tag{13}$$

Differentiating (13) with respect to u and equating to zero, we get the optimum value

of p as

$$p_0^{(1)} = \frac{\left(S_y^2(A+1) + \theta S_{y(2)}^2\right) - \sqrt{\left(S_y^2 + \theta S_{y(2)}^2\right)\left(S_y^2(1+A) + \theta S_{y(2)}^2\right)}}{S_y^2 A}$$

or

$$q_0^{(1)} = -\frac{\left(S_y^2 + \theta S_{y(2)}^2\right) - \sqrt{\left(S_y^2 + \theta S_{y(2)}^2\right)\left(S_y^2(1+A) + \theta S_{y(2)}^2\right)}}{S_y^2 A}$$

$$\min. \mathrm{Var}\left(t_1^{(0)}\right) = \left(\frac{S_y^2 + \theta S_{y(2)}^2}{n}\right).$$

$$\cdot \frac{\left(S_y^2 + \theta S_{y(2)}^2\right) + \sqrt{\left(S_y^2 + \theta S_{y(2)}^2\right)\left(S_y^2(1+A) + \theta S_{y(2)}^2\right)}}{2(S_y^2 + \theta S_{y(2)}^2)}$$

We can compute the gain in precision of the combined estimate $t_1^{(0)}$, obtained by using a double sampling ratio-cum-product estimate from the matched portion of the sample on the second occasion, over direct estimator $\bar{y}^*$ of population mean $\bar{Y}$ on second occasion, using no information collected on the first occasion,

$$G_2 = \frac{\mathrm{Var}(\bar{y}^*)}{\mathrm{Var}(t_1^{(0)})} - 1 \times 100 = \frac{-pqAS_y^2}{(1+qA)S_y^2 + \theta S_{y(2)}^2} \times 100.$$

$$\text{where} \quad \mathrm{Var}(\bar{y}^*) = \frac{S_y^2}{n} + \frac{\theta}{n}S_{y(2)}^2.$$

or

$$G_{opt(2)} = \frac{\mathrm{Var}(\bar{y}^*)}{\min. \mathrm{Var}\left(t_1^{(0)}\right)} - 1 \times 100 =$$

$$= \frac{2\sqrt{(S_y^2 + \theta S_{y(2)}^2)(S_y^2(1+A) + \theta S_{y(2)}^2)} - 2\theta S_{y(2)}^2 - (2+A)S_y^2}{AS_y^2} \times 100.$$

Consider the special case where $-\rho_{xy} = \rho_{yz} = \rho_1$, $\rho_{xz} = \rho_0$, $C_x = C_y = C_z = C$, which

gives an expression more simple for the variance

$$\text{Var}\left(t_1^{(0)}\right) = \left(\frac{S_y^2 + \theta S_{y(2)}^2}{n}\right) \frac{(1 + qA^*) S_y^2 + \theta S_{y(2)}^2}{(1 + q^2 A^*) S_y^2 + \theta S_{y(2)}^2}$$

where

$$A^* = 2 - 4\rho_1 - 2\rho_0.$$

3.3   Estimation of the mean population for current occasion in the presence of non-response only on the second occasion for unmatched portion of the sample

When there is non-response only on the second occasion for unmatched portion of the sample, we define an estimator for population mean $\bar{Y}$ of the study variable y on the current occasion is defined by

$$t_2 = \varphi \bar{t}_m + (1 - \varphi) \bar{t}_u^* \tag{14}$$

where

$$\bar{t}_m = \bar{y}_m \frac{\bar{x}}{\bar{x}_m} \frac{\bar{z}_m}{\bar{z}} \quad \text{and} \quad \bar{t}_u^* = \bar{y}_u^*$$

The variance of the estimator t_2 is given by

$$\text{Var}(t_2) = \varphi^2 \text{Var}(\bar{t}_m) + (1 - \varphi)^2 \text{Var}(\bar{t}_u^*) \tag{15}$$

where

$$\text{Var}(\bar{t}_m) = \frac{S_y^2}{m}(1 + qA) \tag{16}$$

and $\text{Var}(\bar{t}_u^*)$ is given at (6).

Differentiating (15) with respect to φ and equating to zero, we get the optimum value of φ as

$$\varphi_0^{(2)} = \frac{\text{Var}(\bar{t}_u^*)}{\text{Var}(\bar{t}_u^*) + \text{Var}(\bar{t}_m)} = \frac{p\left(S_y^2 + \theta S_{y(2)}^2\right)}{(1 + q^2 A) S_y^2 + (1 - q)\theta S_{y(2)}^2} \tag{17}$$

Substituting the value of $\varphi_0^{(2)}$ in place of φ in (14) yields the optimum estimator for the population mean $\bar{Y}$ of the study variable y on second occasion as

$$t_2^{(0)} = \varphi_0^{(2)} \bar{t}_m + (1 - \varphi_0^{(2)}) \bar{t}_u^*$$

Substituting the variances of $\bar{t}_m$ and $\bar{t}_u^*$ and the optimum value of φ from (17), we get

the variance of $t_2^{(0)}$ as

$$\text{Var}\left(t_2^{(0)}\right) = \frac{\text{Var}(\bar{t}_m)\,\text{Var}(\bar{t}_u^*)}{\text{Var}(\bar{t}_m) + \text{Var}(\bar{t}_u^*)} = \left(\frac{S_y^2}{n}\right)\frac{(1+qA)\,(S_y^2 + \theta S_{y(2)}^2)}{(1+q^2 A)\,S_y^2 + (1-q)\theta S_{y(2)}^2} \qquad (18)$$

Differentiating (18) with respect to u and equating to zero, we get the optimum value of p as

$$p_0^{(2)} = \frac{S_y^2(A+1) - \sqrt{S_y^2(S_y^2 + \theta S_{y(2)}^2)(1+A)}}{S_y^2 A}$$

or

$$q_0^{(2)} = -\frac{S_y^2 - \sqrt{S_y^2(S_y^2 + \theta S_{y(2)}^2)(1+A)}}{S_y^2 A}$$

$$\text{min.}\,\text{Var}\left(t_2^{(0)}\right) = \left(\frac{S_y^2}{n}\right)\cdot\frac{A(S_y^2 + \theta S_{y(2)}^2)}{2\sqrt{S_y^2(S_y^2 + \theta S_{y(2)}^2)(1+A)} - 2S_y^2 - \theta S_{y(2)}^2}$$

We can compute the gain in precision of the combined estimate $t_2^{(0)}$, obtained by using a double sampling ratio-cum-product estimate from the matched portion of the sample on the second occasion, over direct estimator $\bar{y}^*$ of population mean $\bar{Y}$ on second occasion, using no information collected on the first occasion,

$$G_3 = \frac{\text{Var}(\bar{y}^*)}{\text{Var}(t_2^{(0)})} - 1 \times 100 = \frac{p\left(\theta S_{y(2)}^2 - qA S_y^2\right)}{S_y^2(1+qA)} \times 100.$$

$$\text{where}\quad \text{Var}(\bar{y}^*) = \frac{S_y^2}{n} + \frac{\theta}{n}S_{y(2)}^2.$$

or

$$G_{opt(3)} = \frac{\text{Var}(\bar{y}^*)}{\text{min.}\,\text{Var}\left(t_2^{(0)}\right)} - 1 \times 100 =$$

$$= \frac{2\sqrt{S_y^2(S_y^2 + \theta S_{y(2)}^2)(1+A)} - \theta S_{y(2)}^2 - (2+A)S_y^2}{A S_y^2} \times 100.$$

Consider the special case where $-\rho_{xy} = \rho_{yz} = \rho_1$, $\rho_{xz} = \rho_0$, $C_x = C_y = C_z = C$, which

gives an expression more simple for the variance

$$\text{Var}\left(t_2^{(0)}\right) = \left(\frac{S_y^2}{n}\right) \frac{(1+qA^*)(S_y^2+\theta S_{y(2)}^2)}{(1+q^2A^*)S_y^2+(1-q)\theta S_{y(2)}^2}$$

where

$$A^* = 2 - 4\rho_1 - 2\rho_0.$$

The gain in precision of t_0, $t_1^{(0)}$ and $t_2^{(0)}$ over the direct estimate $\bar{y}^*$ for different values of S_y^2, $S_{y(2)}^2$, ρ_{xy}, $\rho_{xy(2)}$, ρ_{yz}, $\rho_{yz(2)}$, ρ_{xz}, $\rho_{xz(2)}$ are presented in tables I-II. It is assumed that $N = 300$ and $n = 50$. From these tables, we obtain the following conclusions:

(i) For the case $S_y^2 > S_{y(2)}^2$, the gain in precision of t_0 and $t_2^{(0)}$ with respect to $\bar{y}^*$ decreases as the values of S_y^2 increases, whereas the gain in precision of $t_1^{(0)}$ with respect to $\bar{y}^*$ increases as the values of S_y^2 increase.

(ii) For the case $S_y^2 < S_{y(2)}^2$, the gain in precision of t_0 and $t_1^{(0)}$ with respect to $\bar{y}^*$ increases as the values of $S_{y(2)}^2$ increase, whereas the gain in precision of $t_2^{(0)}$ with respect to $\bar{y}^*$ decreases as values of $S_{y(2)}^2$ increase.

(iii) For the case $S_y^2 = S_{y(2)}^2$, the gain in precision of t_0, $t_1^{(0)}$ and $t_2^{(0)}$ with respect to $\bar{y}^*$ decreases as the values of S_y^2 and $S_{y(2)}^2$ increase.

(iv) For the case $\rho_{xy} > \rho_{xy(2)}$, the gain in precision of t_0 with respect to $\bar{y}^*$ increases as the values of $\rho_{xy(2)}$ increase, whereas the gain in precision of $t_1^{(0)}$ and $t_2^{(0)}$ with respect to $\bar{y}^*$ remains constant as values of $\rho_{xy(2)}$ increase.

(v) For the case $\rho_{xy} < \rho_{xy(2)}$, the gain in precision of t_0, $t_1^{(0)}$ and $t_2^{(0)}$ with respect to $\bar{y}^*$ increases as the values of ρ_{xy} increase.

(vi) For the case $\rho_{xy} = \rho_{xy(2)}$, the gain in precision of t_0, $t_1^{(0)}$ and $t_2^{(0)}$ with respect to $\bar{y}^*$ increases as the values of ρ_{xy} and $\rho_{xy(2)}$ increase.

(vii) For the case $\rho_{yz} > \rho_{yz(2)}$, the gain in precision of t_0 with respect to $\bar{y}^*$ decreases as the values of $\rho_{yz(2)}$ increase, whereas the gain in precision of $t_1^{(0)}$ and $t_2^{(0)}$ with respect to $\bar{y}^*$ remains constant as values of $\rho_{yz(2)}$ increase.

(viii) For the case $\rho_{yz} < \rho_{yz(2)}$, the gain in precision of t_0, $t_1^{(0)}$ and $t_2^{(0)}$ with respect to $\bar{y}^*$ decreases as the values of ρ_{yz} increase.

(ix) For the case $\rho_{yz} = \rho_{yz(2)}$, the gain in precision of t_0, $t_1^{(0)}$ and $t_2^{(0)}$ with respect to $\bar{y}^*$ decreases as the values of ρ_{yz} and $\rho_{yz(2)}$ increase.

(x) For the case $\rho_{xz} > \rho_{xz(2)}$, the gain in precision of t_0 with respect to $\bar{y}^*$ increases as the values of $\rho_{xz(2)}$ increase, whereas the gain in precision of $t_1^{(0)}$ and $t_2^{(0)}$ with respect to $\bar{y}^*$ remains constant as values of $\rho_{xz(2)}$ increase.

(xi) For the case $\rho_{xz} < \rho_{xz(2)}$, the gain in precision of t_0, $t_1^{(0)}$ and $t_2^{(0)}$ with respect to $\bar{y}^*$ increases as the values of ρ_{xz} increase.

(xii) For the case $\rho_{xz} = \rho_{xz(2)}$, the gain in precision of t_0, $t_1^{(0)}$ and $t_2^{(0)}$ with respect to $\bar{y}^*$ increases as the values of ρ_{xz} and $\rho_{xz(2)}$ increase.

(xiii) The gain in precision of t_0 and $t_1^{(0)}$ with respect to $\bar{y}^*$ decreases as the values of W_2 increase, whereas the gain in precision of $t_2^{(0)}$ with respect to $\bar{y}^*$ increases as the values of W_2 decrease.

(xiv) The gain in precision of t_0 and $t_1^{(0)}$ with respect to $\bar{y}^*$ decreases as the values of $(v-1)$ increase, whereas the gain in precision of $t_2^{(0)}$ with respect to $\bar{y}^*$ increases as the values of $(v-1)$ increase.

(xv) The gain in precision of t_0 and $t_1^{(0)}$ with respect to $\bar{y}^*$ first increases and after decreases as values of p increase, whereas the gain in precision of $t_2^{(0)}$ with respect to $\bar{y}^*$ decreases as the values of p increase.

Table I. Gain in precision, $G_{(1)}$, $G_{(2)}$ and $G_{(3)}$ of the proposed estimates t_0, $t_1^{(0)}$ and $t_2^{(0)}$ over the direct estimate $\bar{y}^*$ for different values of S_y^2, $S_{y(2)}^2$, ρ_{xy}, $\rho_{xy(2)}$, ρ_{yz}, $\rho_{yz(2)}$, ρ_{xz} and $\rho_{xz(2)}$ (when $\Delta_1 = \Delta_2 = \Delta_{1(2)} = \Delta_{2(2)} = 1$)

ρ_{xy}	$\rho_{xy(2)}$	ρ_{yz}	$\rho_{yz(2)}$	ρ_{xz}	$\rho_{xz(2)}$	$(v-1)$	W_2	S_y^2	$S_{y(2)}^2$	p	$G_{(1)}$	$G_{(2)}$	$G_{(3)}$
						$S_y^2 > S_{y(2)}^2$							
0.2	0.3	-0.4	-0.8	0.5	0.2	1.5	0.8	0.5	0.4	0.3	17.27	2.31	38.37
0.2	0.3	-0.4	-0.8	0.5	0.2	1.5	0.8	0.6	0.4	0.3	15.76	2.53	32.79
0.2	0.3	-0.4	-0.8	0.5	0.2	1.5	0.8	0.7	0.4	0.3	14.58	2.72	28.80
						$S_y^2 < S_{y(2)}^2$							
0.8	0.5	-0.4	-0.8	0.3	-0.7	1.5	0.4	0.5	0.6	0.2	2.99	17.39	152
0.8	0.5	-0.4	-0.8	0.3	-0.7	1.5	0.4	0.6	0.7	0.2	3.29	17.78	150
0.8	0.5	-0.4	-0.8	0.3	-0.7	1.5	0.4	0.7	0.8	0.2	3.51	18.06	148.57
						$S_y^2 = S_{y(2)}^2$							
0.4	0.4	-0.4	-0.8	0.8	0.2	1.5	0.4	0.5	0.5	0.2	180	30	780
0.3	0.4	-0.4	-0.8	0.8	0.2	1.5	0.4	0.6	0.6	0.3	84.29	23.33	130
0.2	0.4	-0.4	-0.8	0.8	0.2	1.5	0.4	0.7	0.7	0.4	60	17.14	83.08
						$\rho_{xy} > \rho_{xy(2)}$							
0.8	0.5	-0.4	-0.8	0.3	-0.7	1.5	0.5	0.5	0.5	0.2	2.58	16.84	155
0.8	0.6	-0.4	-0.8	0.3	-0.7	1.5	0.5	0.5	0.5	0.2	5	16.84	155
0.8	0.7	-0.4	-0.8	0.3	-0.7	1.5	0.5	0.5	0.5	0.2	8	16.84	155
						$\rho_{xy} < \rho_{xy(2)}$							
0.5	0.8	-0.4	-0.8	0.3	-0.7	1.5	0.5	0.5	0.5	0.2	2.15	4.48	31.47
0.6	0.8	-0.4	-0.8	0.3	-0.7	1.5	0.5	0.5	0.5	0.2	4.65	7.56	47.31
0.7	0.8	-0.4	-0.8	0.3	-0.7	1.5	0.5	0.5	0.5	0.2	7.78	11.53	77.22
						$\rho_{xy} = \rho_{xy(2)}$							
0.7	0.7	-0.4	-0.8	0.3	-0.7	1.5	0.5	0.5	0.5	0.2	4.82	11.53	77.22
0.8	0.8	-0.4	-0.8	0.3	-0.7	1.5	0.5	0.5	0.5	0.2	11.82	16.84	155
0.9	0.9	-0.4	-0.8	0.3	-0.7	1.5	0.5	0.5	0.5	0.2	24.30	24.30	855
						$\rho_{yz} > \rho_{yz(2)}$							
0.2	0.3	-0.5	-0.8	0.5	0.2	0.7	0.5	0.4	0.6	0.4	22.89	7.47	40.26
0.2	0.3	-0.5	-0.7	0.5	0.2	0.7	0.5	0.4	0.6	0.4	16.74	7.47	40.26
0.2	0.3	-0.5	-0.6	0.5	0.2	0.7	0.5	0.4	0.6	0.4	11.70	7.47	40.26
						$\rho_{yz} < \rho_{yz(2)}$							
0.2	0.3	-0.8	-0.5	0.5	0.2	0.7	0.5	0.4	0.6	0.4	25.95	25.95	112.5
0.2	0.3	-0.7	-0.5	0.5	0.2	0.7	0.5	0.4	0.6	0.4	18.37	18.37	77.31
0.2	0.3	-0.6	-0.5	0.5	0.2	0.7	0.5	0.4	0.6	0.4	12.36	12.36	55.31
						$\rho_{yz} = \rho_{yz(2)}$							
0.2	0.3	-0.8	-0.8	0.5	0.2	0.7	0.5	0.4	0.6	0.4	60	25.95	112.5
0.2	0.3	-0.7	-0.7	0.5	0.2	0.7	0.5	0.4	0.6	0.4	33.05	18.37	77.31
0.2	0.3	-0.6	-0.6	0.5	0.2	0.7	0.5	0.4	0.6	0.4	17.55	12.36	55.31
						$\rho_{xz} > \rho_{xz(2)}$							
0.2	0.3	-0.8	-0.2	0.5	0.2	0.7	0.4	0.5	0.6	0.2	16.23	29.85	113.6
0.2	0.3	-0.8	-0.2	0.5	0.3	0.7	0.4	0.5	0.6	0.2	19.86	29.85	113.6
0.2	0.3	-0.8	-0.2	0.5	0.4	0.7	0.4	0.5	0.6	0.2	24.30	29.85	113.6
						$\rho_{xz} < \rho_{xz(2)}$							
0.2	0.3	-0.8	-0.2	0.2	0.5	0.7	0.4	0.5	0.6	0.2	6.30	6.30	19.30
0.2	0.3	-0.8	-0.2	0.3	0.5	0.7	0.4	0.5	0.6	0.2	11.22	11.22	31.39
0.2	0.3	-0.8	-0.2	0.4	0.5	0.7	0.4	0.5	0.6	0.2	18.39	18.39	54.22
						$\rho_{xz} = \rho_{xz(2)}$							
0.2	0.3	-0.8	-0.2	0.3	0.3	0.7	0.4	0.5	0.6	0.2	6.98	11.21	31.38
0.2	0.3	-0.8	-0.2	0.4	0.4	0.7	0.4	0.5	0.6	0.2	15.01	18.39	54.22
0.2	0.3	-0.8	-0.2	0.5	0.5	0.7	0.4	0.5	0.6	0.2	29.85	29.85	113.6

Table II. Gain in precision, $G_{(1)}$, $G_{(2)}$ and $G_{(3)}$ of the proposed estimates t_0, $t_1^{(0)}$ and $t_2^{(0)}$ over the direct estimate $\bar{y}^*$ for different values of W_2, $(v-1)$ and p (when $\Delta_1 = \Delta_2 = \Delta_{1(2)} = \Delta_{2(2)} = 1$).

ρ_{xy}	$\rho_{xy(2)}$	ρ_{yz}	$\rho_{yz(2)}$	ρ_{xz}	$\rho_{xz(2)}$	$(v-1)$	W_2	S_y^2	$S_{y(2)}^2$	p	$G_{(1)}$	$G_{(2)}$	$G_{(3)}$
						W_2							
0.2	0.3	-0.8	-0.2	0.5	0.2	0.5	0.3	0.3	0.4	0.2	26.15	40	100
0.2	0.3	-0.8	-0.2	0.5	0.2	0.5	0.4	0.3	0.4	0.2	20.43	34.29	106.67
0.2	0.3	-0.8	-0.2	0.5	0.2	0.5	0.5	0.3	0.4	0.2	16.36	30	113.33
						$(v-1)$							
0.8	0.6	-0.4	-0.8	0.5	-0.7	1.0	0.4	0.5	0.3	0.4	51.18	84	270
0.8	0.6	-0.4	-0.8	0.5	-0.7	1.5	0.4	0.5	0.3	0.4	33.91	64.62	300
0.8	0.6	-0.4	-0.8	0.5	-0.7	2.0	0.4	0.5	0.3	0.4	23.79	52.5	330
						p							
0.2	0.3	-0.8	-0.2	0.5	0.2	0.5	0.7	0.8	0.7	0.2	17.86	31.60	110.63
0.2	0.3	-0.8	-0.2	0.5	0.2	0.5	0.7	0.8	0.7	0.3	19.60	34.64	100.62
0.2	0.3	-0.8	-0.2	0.5	0.2	0.5	0.7	0.8	0.7	0.4	18.71	33.98	90.63

4 EFFICIENCY COMPARISON OF THE ESTIMATORS T_0, $T_1^{(0)}$ AND $T_2^{(0)}$

We have compared the variances of the estimators t_0, $t_1^{(0)}$ and $t_2^{(0)}$ obtained in three different possible situations:

4.1 Comparison of the estimators t_0 and $t_1^{(0)}$

From (8) and (13), we have

$$\text{Var}\left(t_0\right) - \text{Var}\left(t_1^{(0)}\right) = \frac{\text{Var}\left(\bar{t}_m^*\right)\text{Var}\left(\bar{t}_u^*\right)}{\text{Var}\left(\bar{t}_m^*\right) + \text{Var}\left(\bar{t}_u^*\right)} - \frac{\text{Var}\left(\bar{t}_{1m}^*\right)\text{Var}\left(\bar{t}_u^*\right)}{\text{Var}\left(\bar{t}_{1m}^*\right) + \text{Var}\left(\bar{t}_u^*\right)} =$$

$$= \frac{(\text{Var}\left(\bar{t}_u^*\right))^2(\text{Var}\left(\bar{t}_m^*\right) - \text{Var}\left(\bar{t}_{1m}^*\right))}{(\text{Var}\left(\bar{t}_m^*\right) + \text{Var}\left(\bar{t}_u^*\right))(\text{Var}\left(\bar{t}_{1m}^*\right) + \text{Var}\left(\bar{t}_u^*\right))} =$$

$$= \frac{\theta A_{(2)}(\text{Var}\left(\bar{t}_u^*\right))^2 S_{y(2)}^2}{m(\text{Var}\left(\bar{t}_m^*\right) + \text{Var}\left(\bar{t}_u^*\right))(\text{Var}\left(\bar{t}_{1m}^*\right) + \text{Var}\left(\bar{t}_u^*\right))} > 0$$

which is always positive. Thus

$$\text{Var}\left(t_0\right) > \text{Var}\left(t_1^{(0)}\right). \tag{19}$$

4.2 Comparison of the estimators t_0 and $t_2^{(0)}$

From (8) and (18), we have

$$\text{Var}\left(t_0\right) - \text{Var}\left(t_2^{(0)}\right) = \frac{\text{Var}\left(\bar{t}_m^*\right)\text{Var}\left(\bar{t}_u^*\right)}{\text{Var}\left(\bar{t}_m^*\right) + \text{Var}\left(\bar{t}_u^*\right)} - \frac{\text{Var}\left(\bar{t}_m\right)\text{Var}\left(\bar{t}_u^*\right)}{\text{Var}\left(\bar{t}_m\right) + \text{Var}\left(\bar{t}_u^*\right)} =$$

$$= \frac{(\text{Var}\left(\bar{t}_u^*\right))^2(\text{Var}\left(\bar{t}_m^*\right) - \text{Var}\left(\bar{t}_1\right))}{(\text{Var}\left(\bar{t}_m^*\right) + \text{Var}\left(\bar{t}_u^*\right))(\text{Var}\left(\bar{t}_m\right) + \text{Var}\left(\bar{t}_u^*\right))} =$$

$$= \frac{\theta(\text{Var}\left(\bar{t}_u^*\right))^2(1 + A_{(2)})S_{y(2)}^2}{m(\text{Var}\left(\bar{t}_m^*\right) + \text{Var}\left(\bar{t}_u^*\right))(\text{Var}\left(\bar{t}_{1m}^*\right) + \text{Var}\left(\bar{t}_u^*\right))} > 0$$

which is always positive. Thus

$$\text{Var}\left(t_0\right) > \text{Var}\left(t_2^{(0)}\right). \tag{20}$$

4.3   Comparison of the estimators $t_1^{(0)}$ and $t_2^{(0)}$

From (13) and (18), we have

$$\mathrm{Var}\left(t_1^{(0)}\right) - \mathrm{Var}\left(t_2^{(0)}\right) = \frac{\mathrm{Var}\left(\bar{t}_{1m}^*\right)\mathrm{Var}\left(\bar{t}_u^*\right)}{\mathrm{Var}\left(\bar{t}_{1m}^*\right) + \mathrm{Var}\left(\bar{t}_u^*\right)} - \frac{\mathrm{Var}\left(\bar{t}_m\right)\mathrm{Var}\left(\bar{t}_u^*\right)}{\mathrm{Var}\left(\bar{t}_m\right) + \mathrm{Var}\left(\bar{t}_u^*\right)} =$$

$$= \frac{(\mathrm{Var}\left(\bar{t}_u^*\right))^2(\mathrm{Var}\left(\bar{t}_{1m}^*\right) - \mathrm{Var}\left(\bar{t}_m\right))}{(\mathrm{Var}\left(\bar{t}_{1m}^*\right) + \mathrm{Var}\left(\bar{t}_u^*\right))(\mathrm{Var}\left(\bar{t}_m\right) + \mathrm{Var}\left(\bar{t}_u^*\right))} =$$

$$= \frac{\theta(\mathrm{Var}\left(\bar{t}_u^*\right))^2 S_{y(2)}^2}{m(\mathrm{Var}\left(\bar{t}_{1m}^*\right) + \mathrm{Var}\left(\bar{t}_u^*\right))(\mathrm{Var}\left(\bar{t}_m\right) + \mathrm{Var}\left(\bar{t}_u^*\right))} > 0$$

which is always positive. Thus

$$\mathrm{Var}\left(t_1^{(0)}\right) > \mathrm{Var}\left(t_2^{(0)}\right). \tag{21}$$

From (19), (20) and (21), we have the following inequality,

$$\mathrm{Var}\left(t_2^{(0)}\right) < \mathrm{Var}\left(t_1^{(0)}\right) < \mathrm{Var}\left(t_{(0)}\right),$$

which shows that the estimator $t_2^{(0)}$ is more efficient than t_0 and $t_1^{(0)}$.

5   EFFICIENCY COMPARISON OF THE ESTIMATORS T_0, $T_{0(R)}$ AND $T_{0(P)}$

5.1   Estimator ratio

In presence of non-response, the optimum estimate of the population mean on second occasion in sampling on two occasions by using a double sampling ratio estimate using one auxiliary variable from the matched portion of the sample is given by

$$\mathrm{Var}\left(t_{0(r)}\right) = \varphi_0{}^2 \mathrm{Var}\left(\bar{t}_{m(r)}^*\right) + (1 - \varphi_0)^2 \mathrm{Var}\left(\bar{t}_u^*\right) \quad \text{where}$$

$$\mathrm{Var}(\bar{t}_{m(r)}^*) = \frac{S_y^2}{m}(1 + qV_1) + \frac{\theta}{m}(1 + V_{1(2)})S_{y(2)}^2; \quad V_1 = -Z_1, \quad V_{1(2)} = -Z_{1(2)}.$$

(see Singh, Kumar and Bhougal (2011), pp. 595)

The optimum estimate of the population mean on second occasion in sampling on two occasions by using a double sampling ratio-cum-product estimate in presence of non-response will be more efficient than the optimum estimate of the population mean on second occasion in sampling on two occasions by using a double sampling ratio

estimate using one auxiliary variable from the matched portion of the sample if

$$\text{Var}\left(t_{0(r)}\right) - \text{Var}\left(t_0\right) \geq 0$$

$$\text{Var}\left(t_{0(r)}\right) - \text{Var}\left(t_0\right) = \frac{\text{Var}\left(\bar{t}^*_{m(r)}\right)\text{Var}\left(\bar{t}^*_u\right)}{\text{Var}\left(\bar{t}^*_{m(r)}\right) + \text{Var}\left(\bar{t}^*_u\right)} - \frac{\text{Var}\left(\bar{t}^*_m\right)\text{Var}\left(\bar{t}^*_u\right)}{\text{Var}\left(\bar{t}^*_m\right) + \text{Var}\left(\bar{t}^*_u\right)} =$$

$$= \frac{(\text{Var}\left(\bar{t}^*_u\right))^2(\text{Var}\left(\bar{t}^*_{m(r)}\right) - \text{Var}\left(\bar{t}^*_m\right))}{(\text{Var}\left(\bar{t}^*_m\right) + \text{Var}\left(\bar{t}^*_u\right))(\text{Var}\left(\bar{t}^*_{m(r)}\right) + \text{Var}\left(\bar{t}^*_u\right))} \geq 0 \quad \text{or} \quad \text{equivalently}$$

$$Z_2 - 2\rho_{xz}\Delta_1\Delta_2 \leq 0 \qquad \text{and} \qquad Z_{2(2)} - 2\rho_{xz(2)}\Delta_{1(2)}\Delta_{2(2)} \leq 0.$$

5.2 Estimator product

In presence of non-response, the optimum estimate of the population mean on second occasion in sampling on two occasions by using a double sampling product estimate using one auxiliary variable from the matched portion of the sample is given by

$$\text{Var}\left(t_{0(p)}\right) = \varphi_0{}^2\,\text{Var}\left(\bar{t}^*_{m(p)}\right) + (1 - \varphi_0)^2\,\text{Var}\left(\bar{t}^*_u\right) \quad \text{where}$$

$$\text{Var}(\bar{t}^*_{m(p)}) = \frac{S_y^2}{m}(1 + qV_2) + \frac{\theta}{m}(1 + V_{2(2)})S_{y(2)}^2; \quad V_2 = Z_2, \quad V_{2(2)} = Z_{2(2)}.$$

(see Singh and Kumar (2009), pp. 160)

The optimum estimate of the population mean on second occasion in sampling on two occasions by using a double sampling ratio-cum-product estimate in presence of non-response will be more efficient than the optimum estimate of the population mean on second occasion in sampling on two occasions by using a double sampling product estimate using one auxiliary variable from the matched portion of the sample if

$$\text{Var}\left(t_{0(p)}\right) - \text{Var}\left(t_0\right) \geq 0$$

$$\text{Var}\left(t_{0(p)}\right) - \text{Var}\left(t_0\right) = \frac{\text{Var}\left(\bar{t}^*_{m(p)}\right)\text{Var}\left(\bar{t}^*_u\right)}{\text{Var}\left(\bar{t}^*_{m(p)}\right) + \text{Var}\left(\bar{t}^*_u\right)} - \frac{\text{Var}\left(\bar{t}^*_m\right)\text{Var}\left(\bar{t}^*_u\right)}{\text{Var}\left(\bar{t}^*_m\right) + \text{Var}\left(\bar{t}^*_u\right)} =$$

$$= \frac{(\text{Var}\,(\bar{t}_u^*))^2(\text{Var}\left(\bar{t}_{m(p)}^*\right) - \text{Var}\,(\bar{t}_m^*))}{(\text{Var}\,(\bar{t}_m^*) + \text{Var}\,(\bar{t}_u^*))(\text{Var}\left(\bar{t}_{m(p)}^*\right) + \text{Var}\,(\bar{t}_u^*))} \geq 0 \quad \text{or} \quad \text{equivalently}$$

$$Z_1 + 2\rho_{xz}\Delta_1\Delta_2 \geq 0 \qquad \text{and} \qquad Z_{1(2)} + 2\rho_{xz(2)}\Delta_{1(2)}\Delta_{2(2)} \geq 0.$$

6 COMPARING ESTIMATORS IN TERMS OF SURVEY COST

We give some ideas about how saving in cost through mail surveys in the context of successive sampling on two occasions for different assumed values of ρ_{xy}, $\rho_{xy(2)}$, ρ_{yz}, $\rho_{yz(2)}$, ρ_{xz}, $\rho_{xz(2)}$, W_2, $(\nu - 1)$ and p.

Let $N = 300$, $n = 50$, $c_0 = 1$, $c_1 = 4$, and $c_2 = 45$ (see Choudhary et al., p. 339), where c_0, c_1, and c_2 denote the cost per unit for mailing a questionnaire, processing the results from the first attempt respondents, and collecting data through personal interview, respectively. In addition, C_{00}^* is the total cost incurred for collecting the data by personal interview from the whole sample, i.e., when there is no non-response. The cost function in this case is given by (assuming the cost incurred on data collection for the matched and unmatched portion of the sample are same and cost incurred on the data collection on both occasions is same)

$$C_{00}^* = 2nc_2. \tag{22}$$

Substituting the values of n and c_2 in Eq. (22), the total cost work out to be 4500.

Let n_1 denotes the number of units which respond at the first attempt and n_2 denotes the number of units which do not respond. Thus,

(i) The cost function for the case when there is non-response on both occasions is given by

$$C_0^* = 2\left[c_0 n + c_1 n_1 + \frac{c_2 n_2}{\nu - 1}\right].$$

The expected cost is given by

(ii) The cost function for the case when there is only non-response on the second occasion is

$$C_1^* = 2c_0 n + c_1 n + \left[c_1 n_1 + \frac{c_2 n_2}{\nu - 1}\right]$$

and the expected cost is given by where

$$n_1^* = n \left[\frac{(1+qA)S_y^2 + \theta S_{y(2)}^2}{(1+q^2A)S_y^2 + \theta S_{y(2)}^2} \right]$$

(iii) The cost function for the case when there is non-response on second occasion for unmatched portion of the sample only is given by

$$C_2^* = \left[c_1 n_1 + \frac{c_2 n_2}{v-1} \right] + 2c_0 n + c_1 n,$$

which expected cost is expressed as where

$$n_2^* = \frac{n(1+qA)S_y^2}{(1+q^2A)S_y^2 + (1-q)\theta S_{y(2)}^2}$$

By equating the variances t_0, $t_1^{(0)}$ and $t_2^{(0)}$ respectively, to $\mathrm{Var}(\bar{y}^*)$ and using the assumed values of different parameters, the values of the sample size for the three cases and the corresponding expected cost of survey were determined with respect of t_0, $t_1^{(0)}$ and $t_2^{(0)}$. The sample sizes associated with the three estimators which provide equal precision to the estimator $\bar{y}^*$ are denoted by n_0^*, n_1^* and n_2^*. The results of this exercise are presented in tables III-IV. From these tables, we obtain the following conclusions:

(i) For the case $S_y^2 > S_{y(2)}^2$, the saving in cost for t_0 first increases and after decreases as the values of S_y^2 increases, and the saving in cost for $t_1^{(0)}$ and $t_2^{(0)}$ increases as the values of S_y^2 increase.
    The sample sizes for t_0 and $t_1^{(0)}$, which have the same precision than $\bar{y}^*$, as the values of S_y^2 increases and the sample sizes for $t_2^{(0)}$, which have the same precision than $\bar{y}^*$, increases as the values of S_y^2 increase.

(ii) For the case $S_y^2 < S_{y(2)}^2$, the saving in cost for t_0, $t_1^{(0)}$ and $t_2^{(0)}$ increases as the values of $S_{y(2)}^2$ increase.
    The sample sizes for t_0 and $t_1^{(0)}$ which have the same precision than $\bar{y}^*$ remains constant as the values of $S_{y(2)}^2$ increase and the sample sizes for $t_2^{(0)}$, which have the same precision than $\bar{y}^*$, first increases and after remains constant as the values of $S_{y(2)}^2$ increase.

(iii) For the case $S_y^2 = S_{y(2)}^2$, the saving in cost for t_0 and $t_1^{(0)}$ increases as the values of S_y^2 and $S_{y(2)}^2$ increase whereas the saving in cost for $t_2^{(0)}$ decreases as the values of S_y^2 and $S_{y(2)}^2$ increase.

The sample sizes for t_0 which have the same precision than $\bar{y}^*$, increases as the values of S_y^2 and $S_{y(2)}^2$ increase whereas the sample sizes for $t_1^{(0)}$, which have the same precision than $\bar{y}^*$ first increases and after remains constant as the values of S_y^2 and $S_{y(2)}^2$ increase and the sample sizes for $t_2^{(0)}$, which have the same precision than $\bar{y}^*$, decreases as the values of S_y^2 and $S_{y(2)}^2$ increase.

(iv) For the case $\rho_{xy} > \rho_{xy(2)}$, the saving in cost for t_0 increases as the values of $\rho_{xy(2)}$ increase, whereas the saving in cost for $t_1^{(0)}$ and $t_1^{(0)}$ remains constant as the values of $\rho_{xy(2)}$ increase.

The sample sizes for t_0, which have the same precision than $\bar{y}^*$, increases as the values of $\rho_{xy(2)}$ increase and the sample sizes for $t_1^{(0)}$ and $t_2^{(0)}$, which have the same precision than $\bar{y}^*$, remains constant as the values of $\rho_{xy(2)}$ increase.

(v) For the case $\rho_{xy} < \rho_{xy(2)}$, the saving in cost for t_0, $t_1^{(0)}$ and $t_2^{(0)}$ increases as the values of ρ_{xy} increase.

The sample sizes for t_0, $t_1^{(0)}$ and $t_2^{(0)}$, which have the same precision than $\bar{y}^*$, increases as the values of ρ_{xy} increase.

(vi) For the case $\rho_{xy} = \rho_{xy(2)}$, the saving in cost for t_0, $t_1^{(0)}$ and $t_2^{(0)}$ increases as the values of ρ_{xy} and $\rho_{xy(2)}$ increase.

The sample sizes for t_0, $t_1^{(0)}$ and $t_2^{(0)}$, which have the same precision than $\bar{y}^*$, increases as the values of ρ_{xy} and $\rho_{xy(2)}$ increase.

(vii) For the case $\rho_{yz} > \rho_{yz(2)}$, the saving in cost for t_0 decreases as the values of $\rho_{yz(2)}$ increase, whereas the saving in cost for $t_1^{(0)}$ and $t_1^{(0)}$ remains constant as the values of $\rho_{yz(2)}$ increase.

The sample sizes for t_0, which have the same precision than $\bar{y}^*$, decreases as the values of $\rho_{yz(2)}$ increase and the sample sizes for $t_1^{(0)}$ and $t_2^{(0)}$, which have the same precision than $\bar{y}^*$, remains constant as the values of $\rho_{yz(2)}$ increase.

(viii) For the case $\rho_{yz} < \rho_{yz(2)}$, the saving in cost for t_0, $t_1^{(0)}$ and $t_2^{(0)}$ with respect to $\bar{y}^*$ decreases as the values of ρ_{yz} increase.

The sample sizes for t_0, $t_1^{(0)}$ and $t_2^{(0)}$ which have the same precision than $\bar{y}^*$, decreases as the values of ρ_{yz} increase.

(ix) For the case $\rho_{yz} = \rho_{yz(2)}$, the saving in cost for t_0, $t_1^{(0)}$ and $t_2^{(0)}$ with respect to $\bar{y}^*$ decreases as the values of ρ_{yz} and $\rho_{yz(2)}$ increase.

The sample sizes for t_0, $t_1^{(0)}$ and $t_2^{(0)}$ which have the same precision than $\bar{y}^*$, decreases as the values of ρ_{yz} and $\rho_{yz(2)}$ increase.

(x) For the case $\rho_{xz} > \rho_{xz(2)}$, the saving in cost for t_0 increases as the values of $\rho_{xz(2)}$ increase, whereas the saving in cost for $t_1^{(0)}$ and $t_2^{(0)}$ remains constant as the values of $\rho_{xz(2)}$ increase.

The sample sizes for t_0, which have the same precision than $\bar{y}^*$, increases as the values of of $\rho_{xz(2)}$ increase and the sample sizes for $t_1^{(0)}$ and $t_2^{(0)}$, which have the same precision than $\bar{y}^*$, remains constant as the values of $\rho_{xz(2)}$ increase.

(xi) For the case $\rho_{xz} < \rho_{xz(2)}$, the saving in cost for t_0, $t_1^{(0)}$ and $t_2^{(0)}$ with respect to $\bar{y}^*$ increases as the values of ρ_{xz} increase.

The sample sizes for t_0, $t_1^{(0)}$ and $t_2^{(0)}$ which have the same precision than $\bar{y}^*$, increases as the values of ρ_{xz} increase.

(xii) For the case $\rho_{xz} = \rho_{xz(2)}$, the saving in cost for t_0, $t_1^{(0)}$ and $t_2^{(0)}$ with respect to $\bar{y}^*$ increases as the values of ρ_{xz} and $\rho_{xz(2)}$ increase.

The sample sizes for t_0, $t_1^{(0)}$ and $t_2^{(0)}$ which have the same precision than $\bar{y}^*$, increases as the values of ρ_{xz} and $\rho_{xz(2)}$ increase.

(xiii) The saving in cost for t_0, $t_1^{(0)}$ and $t_2^{(0)}$ with respect to $\bar{y}^*$ increases as the values of W_2 increase.

The sample sizes for t_0 and $t_1^{(0)}$ which have the same precision than $\bar{y}^*$ remains constant as the values of W_2 increase, whereas the sample sizes for $t_2^{(0)}$ which have the same precision than $\bar{y}^*$ decreases as the values of W_2 increase.

(xiv) The saving in cost for t_0, $t_1^{(0)}$ and $t_2^{(0)}$ with respect to $\bar{y}^*$ decreases as the values of $(v-1)$ increase.

The sample sizes for t_0, $t_1^{(0)}$ and $t_2^{(0)}$ which have the same precision than $\bar{y}^*$ decreases as the values of $(v-1)$ increase.

(xv) The saving in cost for t_0, $t_1^{(0)}$ and $t_2^{(0)}$ with respect to $\bar{y}^*$ increases as the values of p increase.

The sample sizes for t_0, $t_1^{(0)}$ and $t_2^{(0)}$ which have the same precision than $\bar{y}^*$ increases as the values of p increase.

Table III. Sample sizes and corresponding expected cost of survey, which have the same precision than the proposed estimates t_0, $t_1^{(0)}$ and $t_2^{(0)}$ over the direct estimate $\bar{y}^*$ for different values of S_y^2, $S_{y(2)}^2$, ρ_{xy}, $\rho_{xy(2)}$, ρ_{yz} and $\rho_{yz(2)}$ (when $\Delta_1 = \Delta_2 = \Delta_{1(2)} = \Delta_{2(2)} = 1$)

						$S_y^2 > S_{y(2)}^2$										
0.2	0.3	-0.4	-0.8	0.5	0.2	1.5	0.8	0.5	0.4	0.3	43	49	36	2200.1	1505.3	1112.6
0.2	0.3	-0.4	-0.8	0.5	0.2	1.5	0.8	0.6	0.4	0.3	54	51	43	2794.8	1574.1	1312.1
0.2	0.3	-0.4	-0.8	0.5	0.2	1.5	0.8	0.7	0.4	0.3	54	51	44	2786.0	1576.3	1346.6
						$S_y^2 < S_{y(2)}^2$										
0.8	0.5	-0.4	-0.8	0.3	-0.7	1.5	0.4	0.7	0.6	0.2	51	53	50	1576.3	1089.2	1029.1
0.8	0.5	-0.4	-0.8	0.3	-0.7	1.5	0.4	0.6	0.7	0.2	51	53	51	1579.1	1089.7	1031.5
0.8	0.5	-0.4	-0.8	0.3	-0.7	1.5	0.4	0.5	0.8	0.2	51	53	51	1581.0	1090.2	1033.1
						$S_y^2 = S_{y(2)}^2$										
0.4	0.4	-0.4	-0.8	0.8	0.2	1.5	0.4	0.5	0.5	0.2	55	54	52	1701.2	1102.7	1058.9
0.3	0.4	-0.4	-0.8	0.8	0.2	1.5	0.4	0.6	0.6	0.3	57	55	51	1764.7	1122.5	1038.3
0.2	0.4	-0.4	-0.8	0.8	0.2	1.5	0.4	0.7	0.7	0.4	59	55	48	1811.8	1123.7	988
						$\rho_{xy} > \rho_{xy(2)}$										
0.8	0.5	-0.4	-0.8	0.3	-0.7	1.5	0.5	0.5	0.5	0.2	51	53	50	1837.7	1227	1156.4
0.8	0.6	-0.4	-0.8	0.3	-0.7	1.5	0.5	0.5	0.5	0.2	52	53	50	1862.1	1227	1156.4
0.8	0.7	-0.4	-0.8	0.3	-0.7	1.5	0.5	0.5	0.5	0.2	53	53	50	1884	1227	1156.4
						$\rho_{xy} < \rho_{xy(2)}$										
0.5	0.8	-0.4	-0.8	0.3	-0.7	1.5	0.5	0.5	0.5	0.2	51	51	47	1832.4	1186.7	1079.7
0.6	0.8	-0.4	-0.8	0.3	0.7	1.5	0.5	0.5	0.5	0.2	52	52	48	1859	1201.7	1109.2
0.7	0.8	-0.4	-0.8	0.3	-0.7	1.5	0.5	0.5	0.5	0.2	53	53	49	1882.4	1215.1	1134.8
						$\rho_{xy} = \rho_{xy(2)}$										
0.7	0.7	-0.4	-0.8	0.3	-0.7	1.5	0.5	0.5	0.5	0.2	52	52	49	1860.5	1215.1	1134.8
0.8	0.8	-0.4	-0.8	0.3	-0.7	1.5	0.5	0.5	0.5	0.2	53	53	50	1903.1	1227	1156.4
0.9	0.9	-0.4	-0.8	0.3	-0.7	1.5	0.5	0.5	0.5	0.2	54	45	51	1937.3	1237.7	1175.2
						$\rho_{yz} > \rho_{yz(2)}$										
0.2	0.3	-0.5	-0.8	0.5	0.2	0.7	0.5	0.4	0.6	0.4	56	53	46	3934.2	2122.6	1838.2
0.2	0.3	-0.5	-0.7	0.5	0.2	0.7	0.5	0.4	0.6	0.4	55	53	46	3866.7	2122.6	1838.2
0.2	0.3	-0.5	-0.6	0.5	0.2	0.7	0.5	0.4	0.6	0.4	54	53	46	3794.3	2122.6	1838.2
						$\rho_{yz} < \rho_{yz(2)}$										
0.2	0.3	-0.8	-0.5	0.5	0.2	0.7	0.5	0.4	0.6	0.4	56	56	51	3961.7	2262.7	2045.5
0.2	0.3	-0.7	-0.5	0.5	0.2	0.7	0.5	0.4	0.6	0.4	55	55	49	3886.5	2219.7	1983
0.2	0.3	-0.6	-0.5	0.5	0.2	0.7	0.5	0.4	0.6	0.4	54	54	48	3805	2173.2	1914.2
						$\rho_{yz} = \rho_{yz(2)}$										
0.2	0.3	-0.8	-0.8	0.5	0.2	0.7	0.5	0.4	0.6	0.4	59	56	51	4134	2263	2045.5
0.2	0.3	-0.7	-0.7	0.5	0.2	0.7	0.5	0.4	0.6	0.4	57	55	49	4014.5	2219.7	1983
0.2	0.3	-0.6	-0.6	0.5	0.2	0.7	0.5	0.4	0.6	0.4	55	54	48	3876.6	2173.2	1914.2

Table IV. Sample sizes and corresponding expected cost of survey, which have the same precision than the proposed estimates t_0, $t_1^{(0)}$ and $t_2^{(0)}$ over the direct estimate $\bar{y}^*$ for different values of ρ_{xz}, $\rho_{xz(2)}$, W_2, $(v-1)$ and p (when $\Delta_1 = \Delta_2 = \Delta_{1(2)} = \Delta_{2(2)} = 1$).

						$\rho_{xz} > \rho_{xz(2)}$										
0.2	0.3	-0.8	-0.2	0.5	0.2	0.7	0.4	0.5	0.6	0.2	52	54	52	3103.4	1843.8	1798.4
0.2	0.3	-0.8	-0.2	0.5	0.3	0.7	0.4	0.5	0.6	0.2	53	54	52	3118.9	1843.8	1798.4
0.2	0.3	-0.8	-0.2	0.5	0.4	0.7	0.4	0.5	0.6	0.2	54	54	52	3133.4	1843.8	1798.4
						$\rho_{xz} < \rho_{xz(2)}$										
0.2	0.3	-0.8	-0.2	0.2	0.5	0.7	0.4	0.5	0.6	0.2	51	51	50	3028.5	1774.3	1701.6
0.2	0.3	-0.8	-0.2	0.3	0.5	0.7	0.4	0.5	0.6	0.2	52	52	51	3073.9	1800.9	1739.6
0.2	0.3	-0.8	-0.2	0.4	0.5	0.7	0.4	0.5	0.6	0.2	53	53	52	3113.1	1823.9	1771.4
						$\rho_{xz} = \rho_{xz(2)}$										
0.2	0.3	-0.8	-0.2	0.3	0.3	0.7	0.4	0.5	0.6	0.2	52	52	51	3036.2	1800.9	1739.6
0.2	0.3	-0.8	-0.2	0.4	0.4	0.7	0.4	0.5	0.6	0.2	53	53	52	3097.3	1823.9	1771.4
0.2	0.3	-0.8	-0.2	0.5	0.5	0.7	0.4	0.5	0.6	0.2	54	54	53	3147.2	1843.8	1798.4
						W_2										
0.2	0.3	-0.8	-0.2	0.5	0.2	0.5	0.3	0.3	0.4	0.2	53	54	54	3320.2	1934.7	1917.9
0.2	0.3	-0.8	-0.2	0.5	0.2	0.5	0.4	0.3	0.4	0.2	53	54	53	4223.5	2406.3	2359.8
0.2	0.3	-0.8	-0.2	0.5	0.2	0.5	0.5	0.3	0.4	0.2	53	54	52	5117.6	2864.9	2794.9
						$(v-1)$										
0.8	0.6	-0.4	-0.8	0.5	-0.7	1.0	0.4	0.5	0.3	0.4	58	60	57	2499.4	1574.3	1518
0.8	0.6	-0.4	-0.8	0.5	-0.7	1.5	0.4	0.5	0.3	0.4	57	59	56	1761.6	1203.9	1138.8
0.8	0.6	-0.4	-0.8	0.5	-0.7	2.0	0.4	0.5	0.3	0.4	56	58	54	1391.2	1017.3	943.9
						p										
0.2	0.3	-0.8	-0.2	0.5	0.2	0.5	0.7	0.8	0.7	0.2	53	54	52	6966.6	3798.6	3713.7
0.2	0.3	-0.8	-0.2	0.5	0.2	0.5	0.7	0.8	0.7	0.3	54	56	53	7125.5	3920.4	3772.1
0.2	0.3	-0.8	-0.2	0.5	0.2	0.5	0.7	0.8	0.7	0.4	55	57	54	7217.7	4015.6	3788.2

7   CONCLUSION

In this article, the problem of estimating the finite population mean on the samples selected over two occasions by using a double sampling ratio-cum-product estimate from the matched portion of the sample has been presented to take care of unit non-response.

Three different possible cases when there is non-response (i) on both the occasions, (ii) only on the second occasion for both matched and unmatched portions of the sample, and (iii) only on the second occasion for unmatched portion of the sample, have been discussed. The gain in precision by t_0 (optimum estimate in case (i)), $t_1^{(0)}$ (optimum estimate in case (ii)) and $t_2^{(0)}$ (optimum estimate in case (iii)) over direct estimator, $\bar{y}^*$, using no information gathered on the first occasion have been compared.

Also, we have derived the sample sizes and the cost function associated with the three proposed estimates t_0, $t_1^{(0)}$ and $t_2^{(0)}$ which provide equal precision to the estimator $\bar{y}^*$.

It is perceived that the gain in precision due to $t_2^{(0)}$ over $\bar{y}^*$ is larger as compared to t_0 and $t_1^{(0)}$ It has also been observed that the expected cost of the survey is minimum for the estimator, $t_2^{(0)}$ and maximum for the estimator, t_0. Thus, the proposed study is recommended when there is need to correct for non-response in sampling over two occasions.

ACKNOWLEDGMENTS

REFERENCES

Artés, E. M., Rueda, M. M. and Arcos A., 1999. Aportaciones al muestreo sucesivo. Metodología de Encuestas, 1(1), 19-28.

Artés, E. M., García Luengo, A. V., 2001. Estimation of current population ratio in successive sampling. Journal of the Indian Society of Agricultural Statistics, 54(3), 342-354.

Artés, E. M., García Luengo, A. V. and Oña, I., 2001. Una alternativa a la estimación de razón en dos ocasiones, Publicaciones del Ministerio de Defensa, 123-134.

Artés, E., García Luengo, A. V., 2005. Multivariate indirect methods of estimation in successive sampling. J. Indian Soc. Agricultural Statist., 59(2), 97-103.

Choudhary, R.K, Bathla, H.V.L, Sud, U.C., 2004. On non-response in sampling on two occasions. Journal of the Indian Society of Agricultural Statistics, 58, 331-343.

Cochran, W. G., 1977. Sampling Techniques. Third edition. John Wiley & Sons, New York.

Eckler, A. R., 1955. Rotation sampling. The Annals of Mathematical Statistics, 26, 664-685.

García Luengo, A. V., Artés, E. M., 2002. Improvement on estimating of current population ratio in successive sampling. Brazilian Journal of Probability and Statistics, 16(2), 107-122.

García Luengo, A. V., 2008. Estimation of current population product in successive sampling. Pakistan Journal of Statistics, 24(2), 87-98.

García Luengo, A. V., Oña, I., 2010. The non-response in the change of mean and the sum of mean for current occasion in sampling on two occasions. Chilean Journal of Statistics Vol. 2, No. 1, 63-84.

Garcia Luengo, A. V., 2013. Improvement on the non- response in the population ratio of mean for current occasion in sampling on two occasions. Pakistan Journal of Statistics and Operation Research, Vol. 9, No. 1, 25-51.

Garcia Luengo, A. V., 2014. A Study on Non-Response for the Estimation of Current Population Ratio in Sampling on Two Occasions. Journal of Advanced Computing. Volumen 3, No. 1, 35-55.

Gupta, P. C., 1979. Some estimation problems in sampling using auxiliary information, Ph. D. thesis IARS. New Delhi.

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

Jessen, R. J., 1942. Statistical investigation of a sample survey for obtaining farm facts. Iowa Agricultural Experiment Statistical Research Bulletin, 304, 54-59 (for SPR material).

Karna, J. P., Nath D. CH., 2015.Successive sampling strategy under non response. JAMSI, 11, No.1, 115-139.

Okafor, F. C., Arnab, R., 1987. Some strategies of two-stage sampling for estimating population ratios over two occasions. The Australian Journal of Statistics, 29(2), 128-142.

Okafor, F. C., 1992. The theory and application of sampling over two occasions for the estimation of current population ratio. Statistica, 1, 137-147.

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

Okafor, F. C., 2001. Treatment of non-response in successive sampling. Statistica, 61, 195-204.

Patterson, H. D., 1950. Sampling on successive occasions with partial replacement of units. Journal of The Royal Statistical Society Series B–Statistical Methodology, B12, 241-255.

Raj, D., 1968. Sampling Theory. McGraw Hill, New York.

Rao, J. N. K., 1957. Double ratio estimate in forest surveys. Journal of the Indian Society of Agricultural Statistics, 9, 191-204.

Rao, J. N. K., Pereira, N.P., 1968. On double ratio estimators. Sankhya, A, 30, 83-90.

Singh, M. P. 1967. Ratio–cum–Product Method of Estimation. Metrika, 12, 34–42.

Singh, V. K., Singh, G. N., Shukla, D., 1991. An efficient family of ratio-cum-difference type estimators in successive sampling over two occasions. Jour. Sci. Res., 41, C, 149-159.

Singh, G. N., Priyanka, K., 2007. Estimation of population mean at current occasion in successive sampling under a super-population model. Model Assisted Statistics and Applications, 2, 189-200.

Singh, H. P., Kumar, S., 2009. Estimation of multivariate indirect methods in presence of non-response in successive sampling. Metron, LXVII(2), 153-175.

Singh, H. P., Kumar, S., 2010. Estimation of population product in presence of non-response in successive sampling. Statistical Papers, 51, 4, Pages 975-996.

Singh, H. P., Kumar, S. and Bhougal, S., 2011: Multivariate Ratio Estimation in Presence of Non-Response in Successive Sampling, Journal of Statistical Theory and Practice, 5:4, 591-611

Tikkiwal, B. D., 1951. Theory of Successive Sampling. Unpublished thesis for diploma I.C.A.R., New Delhi, India.

Yates, F., 1949. Sampling Methods for Censuses and Surveys. Griffin, London.

Amelia Victoria García Luengo
Department of Mathematics
University of Almería
Almería, Spain
e-mail:amgarcia@ual.es