Chilean Journal of Statistics
Vol. 2, No. 1, April 2011, 63–84

Sampling Theory
Research Paper

The non-response in the change of mean and the sum of mean for current occasion in sampling on two occasions

Amelia V. García-Luengo* and Inmaculada Oña Casado

Departamento de Estadística y Matemática Aplicada, Universidad de Almería, Almería, Spain

(Received: 15 August 2009 · Accepted in final form: 13 January 2010)

Abstract

In this article, we attempt the problem of estimation of the change of mean and the sum of mean in mail surveys. This problem is conducted for current occasion in the context of sampling on two occasions when there is non-response (i) on both occasions, (ii) only on the first occasion and (iii) only on the second occasion. We obtain the loss in precision of all the estimators with respect to the estimator of the change of mean and the sum of mean when there is no non-response. We derive the sample sizes and the saving in cost for all the estimators, which have the same precision than the estimator of the change of mean and the sum of mean when there is no non-response. An empirical study that allows us to investigate the performance of the proposed strategy is carried out.

Keywords: Estimator of the change · Estimator of the sum · Non-response · Successive sampling.

Mathematics Subject Classification: 62D05.

1. Introduction

A fact that cannot be underestimated when samples are analyzed is the moment or spell in which the sample results refer. There exist two major reasons to explain why the time factor must be taken into account in this issue, which are (i) the population characteristics, since these may be modified through time or (ii) the population composition, since this may be modified due to the fact that individuals can increase it (births) or decrease it (death). If the composition and characteristics of the sample units remain unchanged, a single occasion would be enough to perform a sampling, as the results would always be valid. In practice, the mentioned changes prevent us from that simplification and, at the same time, give rise to a set of targets –such as cross estimation of population parameters and net changes, estimations of average values of parameters through time, etc.– that can be analyzed by means of continuous surveys.

The survey circumstances and the study characteristics are the key to choose the appropriate sampling design. These are some of the choices:

*Corresponding author. Amelia Victoria García Luengo. Departamento de Estadística y Matemática Aplicada, Universidad de Almería, Carretera de Sacramento s/n, La Cañada de San Urbano, 04120 Almería, España. Email: amgarcia@ual.es

ISSN: 0718-7912 (print)/ISSN: 0718-7920 (online)

http://www.soche.cl/chjs

(i) To extract a new sample on every occasion (repeated sampling). To estimate the sum, the better thing is using a new sample in every occasion.
(ii) To use the same sample in every occasion (panel sampling). To estimate the change, the better thing is using the same sample in every occasion.
(iii) To perform a partial replacement of units from one occasion to another (sampling on successive occasions, which is also called rotation sampling when the units are constructed by the number of stages in which they 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–).

If a population unit value in a occasion can be related to the same unit in the next occasion, then we are enabled to use the information obtained in the preceding occasion in order to improve current estimation of the population parameter. To this effect, the sample must be obtained in such a way that the sample units in the two successive occasions have some common units so that the preceding sampling information is used.

Some of the reasons that explain the use of the partial replacement of sample units are the following:

(i) Cost reduction (using totally new samples at each time can be unduly expensive).
(ii) Increase of the estimators' accuracy.
(iii) The evasion of indefinite presence of the same units in the sample, since this can result in failures and efficiency reduction of the estimators.

For instance, using panel sampling for family surveys are biased due to the lack of cooperation of some families that belong to the home panel. For this reason, INE frequently uses surveys consisting of rotating sampling because it takes advantage of the two other surveys (repeated and panel surveys).

Jessen (1942), Tikkiwal (1951), Yates (1949), Patterson (1950), Eckler (1955) and Raj (1968) contributed towards the development of the theory of unbiased estimation of mean of characteristics in successive sampling. Hansen and Hurwitz (1946) suggested a technique for handling the non-response in mail surveys. These surveys have the advantage that the data can be collected in a relatively inexpensive way. Okafor (2001) extended these surveys to the estimation of the population total in element sampling on two successive occasions. Later, Choudhary et al. (2004) used the Hansen and Hurwitz (HH) technique to estimate the population mean for current occasion in the context of sampling on two occasions when there is non-response on both occasions. More recently, Singh and Kumar (2010) used the HH technique to estimate the population product for current occasion in the context of sampling on two occasions when there is non-response on both occasions. However, non-response is a common problem with mail surveys. Cochran (1977) and Okafor and Lee (2000) extended the HH technique to the case when the information on the characteristic under study is also available on auxiliary characteristic.

In this article, we develop the HH technique to estimate the change of mean and the sum of mean for current occasion in the context of sampling on two occasions when there is non-response (i) on both occasions, (ii) only on the first occasion and (iii) only on the second occasion.

The rest of this paper is organized as follows. Section 2 describes the HH technique. Section 3 discusses about the estimation of the change of mean. Section 4 is focussed on the estimation of the sum of mean. In this section, an empirical study that allows us to investigate the performance of the proposed strategy is carried out. Section 5 compares proposed estimators in terms of the survey cost. Finally, Section 6 sketches some conclusions.

2. The Technique

Consider a finite population of N identifiable units. Let (x_i, y_i) be, for $i = 1, \ldots, N$, the values of the characteristic on the first and second occasions, respectively. We assume that the population can be divided into two classes, those who respond at the first attempt and those who 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} 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} the size of the subsample drawn from the non-response class in the unmatched portion of the sample on the two occasions. Also, let σ^2 and σ_2^2 denote the population variance and population variance pertaining to the non-response class, respectively. Similarly, ρ and ρ_2 denote correlation between units belonging to the matched portion and the correlation between non-respondents belonging to the matched portion. In addition, let $\bar{x}_m^*$ and $\bar{x}_u^*$ denote the estimator for matched and unmatched portions of the sample on the first occasion, respectively. Let the corresponding estimator for the second occasion be denoted by $\bar{y}_m^*$ and $\bar{y}_u^*$. Thus, have the following setup:

$$\begin{array}{lcc} 1^{\text{st}} \text{ occasion} \longrightarrow & \bar{x}_u^* & \bar{x}_m^*, \\ 2^{\text{nd}} \text{ occasion} \longrightarrow & & \bar{y}_m^* \quad \bar{y}_u^*, \end{array}$$

where

$$\begin{aligned}
\bar{x}_m^* &= \frac{m_1 \bar{x}_{m_1} + m_2 \bar{x}_{m_{h_2}}}{m}, \\
\bar{x}_u^* &= \frac{u_1 \bar{x}_{u_1} + u_2 \bar{x}_{u_{h_2}}}{u}, \\
\bar{y}_m^* &= \frac{m_1 \bar{y}_{m_1} + m_2 \bar{y}_{m_{h_2}}}{m}, \quad \text{and} \\
\bar{y}_u^* &= \frac{u_1 \bar{y}_{u_1} + u_2 \bar{y}_{u_{h_2}}}{u}.
\end{aligned}$$

It can be easily seen that

$$\begin{aligned}
\mathrm{Cov}(\bar{x}_m^*, \bar{x}_u^*) = \mathrm{Cov}(\bar{x}_m^*, \bar{y}_u^*) &= \mathrm{Cov}(\bar{y}_m^*, \bar{x}_u^*) = \mathrm{Cov}(\bar{y}_m^*, \bar{y}_u^*) = \mathrm{Cov}(\bar{y}_u^*, \bar{x}_u^*) = 0, \\
\mathrm{Cov}(\bar{x}_m^*, \bar{x}_m^*) = \mathrm{Var}(\bar{x}_m^*) &= \frac{\sigma^2}{m} + \frac{f N_2 \sigma_2^2}{N m}, \\
\mathrm{Cov}(\bar{x}_u^*, \bar{x}_u^*) = \mathrm{Var}(\bar{x}_u^*) &= \frac{\sigma^2}{u} + \frac{f N_2 \sigma_2^2}{N u}, \\
\mathrm{Cov}(\bar{y}_m^*, \bar{y}_m^*) = \mathrm{Var}(\bar{y}_m^*) &= \frac{\sigma^2}{m} + \frac{f N_2 \sigma_2^2}{N m}, \\
\mathrm{Cov}(\bar{y}_u^*, \bar{y}_u^*) = \mathrm{Var}(\bar{y}_u^*) &= \frac{\sigma^2}{u} + \frac{f N_2 \sigma_2^2}{N u}, \quad \text{and} \\
\mathrm{Cov}(\bar{y}_m^*, \bar{x}_m^*) &= \frac{\rho \sigma^2}{m} + \frac{\rho_2 f N_2 \sigma_2^2}{N m},
\end{aligned}$$

where $W_2 = N_2/N$, $A = f\, W_2\, \sigma^2$, and $f = m_2/m_{h_2} = u_2/u_{h_2}$.

3. Estimation of the Change of Mean

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

Consider the following minimum variance linear unbiased estimator of the change:

$$\Delta_{12} = a\,\bar{x}_u^* + b\,\bar{x}_m^* + c\,\bar{y}_m^* + d\,\bar{y}_u^*, \tag{1}$$

which expected value is given by

$$\begin{aligned}\mathrm{E}(\Delta_{12}) &= a\,\mathrm{E}(\bar{x}_u^*) + b\mathrm{E}(\bar{x}_m^*) + c\,\mathrm{E}(\bar{y}_m^*) + d\mathrm{E}(\bar{y}_u^*)\\ &= a\bar{X}^* + b\,\bar{X}^* + c\,\bar{Y}^* + d1, \bar{Y}^* = (a+b)\,\bar{X}^* + (c+d)\,\bar{Y}^* = \bar{Y}^* - \bar{X}^*.\end{aligned}$$

Unbiasedness of Δ_{12} implies $a+b=-1$ and $c+d=1$, so that $b=-(a+1)$ an $d=1-c$. Substituting the value of b and d in Equation (1), we obtain

$$\Delta_{12} = a\,\bar{x}_u^* - (a+1)\,\bar{x}_m^* + c\,\bar{y}_m^* + (1-c)\,\bar{y}_u^*. \tag{2}$$

The variance of Δ_{12} is given by

$$\mathrm{V}(\Delta_{12}) = a^2\,\mathrm{V}(\bar{x}_u^*) + (a+1)^2\,\mathrm{V}(\bar{x}_m^*) + c^2\,\mathrm{V}(\bar{y}_m^*) + (1-c)^2\,\mathrm{V}(\bar{y}_u^*) - 2(a+1)\,c\,\mathrm{Cov}(\bar{x}_m^*, \bar{y}_m^*).$$

We wish to choose whose values of a and c that minimize $\mathrm{V}(\Delta_{12})$. Equating the derivatives of $\mathrm{V}(\Delta_{12})$ with respect to a and c to zero, it follows that the optimum values are

$$\begin{aligned} a_{\mathrm{opt}} &= \frac{p\,q\,(\sigma^2+A)(\rho\,\sigma^2+\rho_2 A)}{(\sigma^2+A)^2 - q^2(\rho\,\sigma^2+\rho_2 A)^2} - \frac{q\,((\sigma^2+A)^2 - q\,(\rho\,\sigma^2+\rho_2 A)^2)}{(\sigma^2+A)^2 - q^2(\rho\,\sigma^2+\rho_2 A)^2} \quad \text{and}\\ c_{\mathrm{opt}} &= \frac{p\,(\sigma^2+A)^2}{(\sigma^2+A)^2 - q^2(\rho\,\sigma^2+\rho_2 A)^2} + \frac{p\,q\,(\sigma^2+A)(\rho\,\sigma^2+\rho_2 A)}{(\sigma^2+A)^2 - q^2(\rho\,\sigma^2+\rho_2 A)^2}.\end{aligned}$$

Substituting the optimum values of a and c in Equation (2), we obtain

$$\begin{aligned}\Delta_{12} &= \frac{q\,((\sigma^2+A)^2 - q\,(\rho\,\sigma^2+\rho_2 A)^2)}{(\sigma^2+A)^2 - q^2(\rho\,\sigma^2+\rho_2 A)^2}(\bar{y}_u^* - \bar{x}_u^*) + \frac{p\,(\sigma^2+A)^2}{(\sigma^2+A)^2 - q^2(\rho\,\sigma^2+\rho_2 A)^2}(\bar{y}_m^* - \bar{x}_m^*)\\ &\quad + \frac{p\,q\,(\sigma^2+A)(\rho\,\sigma^2+\rho_2 A)}{(\sigma^2+A)^2 - q^2(\rho\,\sigma^2+\rho_2 A)^2}\,[(\bar{x}_u^* - \bar{x}_m^*) + (\bar{y}_u^* - \bar{y}_m^*)]\\ &= \frac{p\,(\sigma^2+A)}{(\sigma^2+A) - q\,(\rho\,\sigma^2+\rho_2 A)}(\bar{y}_m^* - \bar{x}_m^*) + \frac{q\,((\sigma^2+A) - (\rho\,\sigma^2+\rho_2 A))}{(\sigma^2+A) - q\,(\rho\,\sigma^2+\rho_2 A)}(\bar{y}_u^* - \bar{x}_u^*).\end{aligned}$$

Thus, the optimum variance of Δ_{12} is given by

$$\mathrm{V}(\Delta_{12}) = \frac{2}{n}(\sigma^2+A)\frac{(\sigma^2+A) - (\rho\,\sigma^2+\rho_2 A)}{(\sigma^2+A) - q\,(\rho\,\sigma^2+\rho_2 A)}. \tag{3}$$

We note that, for $(\rho\,\sigma^2+\rho_2 A)/(\sigma^2+A) > 0$, Equation (3) is minimum for $q=0$, i.e., the variance de Δ_{12} is minimized if the units on both occasions are identical. In this case,

$$\mathrm{V}(\Delta_{12}) = \frac{2}{n}(\sigma^2+A) - (\rho\,\sigma^2+\rho_2 A).$$

For $\rho = \rho_2$, $\mathrm{V}(\Delta_{12})$ reduces to

$$\mathrm{V}(\Delta_{12}) = \frac{2}{n}(\sigma^2 + A)\frac{(1-\rho)}{(1-q\,\rho)},$$

while if $A = 0$, i.e., there is non-response, the $\mathrm{V}(\Delta_{12})$ reduces to

$$\mathrm{V}(\Delta_0) = \frac{2\sigma^2}{n}\frac{(1-\rho)}{(1-q\,\rho)},$$

where Δ_0 is the usual estimator of the change for the current occasion in the context of sampling on two occasions when there is complete response, that is,

$$\Delta_0 = a\,\bar{x}_u + b\,\bar{x}_m + c\,\bar{y}_m + d\,\bar{y}_u.$$

3.2 Estimation of the change of mean for the current occasion in the presence of non-response on the first occasion

When there is non-response only on the first occasion, the minimum variance linear unbiased estimator of the change can be obtained as

$$\Delta_1 = a\,\bar{x}^*_u + b\,\bar{x}^*_m + c, \bar{y}_m + d\,\bar{y}_u, \quad \text{where} \quad \bar{y}_m = \frac{1}{m}\sum_{i=1}^{m} y_i \quad \text{and} \quad \bar{y}_u = \frac{1}{u}\sum_{i=1}^{u} y_i.$$

Imposing the unbiasedness and minimum variance unbiased conditions, the optimum values of constants a and c are given by

$$\begin{aligned} a_{\text{opt}} &= \frac{p\,q\,\sigma^2\rho}{(\sigma^2 + A) - q^2\,\rho^2\sigma^2} - \frac{q\,((\sigma^2 + A) - q\,\rho^2\sigma^2)}{(\sigma^2 + A) - q^2\,\rho^2\sigma^2} \quad \text{and} \\ c_{\text{opt}} &= \frac{p\,(\sigma^2 + A)}{(\sigma^2 + A) - q^2\,\rho^2\sigma^2} + \frac{p\,q\,(\sigma^2 + A)\rho}{(\sigma^2 + A) - q^2\,\rho^2\sigma^2}. \end{aligned}$$

Thus,

$$\Delta_1 = a\,\bar{x}^*_u - (a+1)\,\bar{x}^*_m + c\,\bar{y}_m + (1-c)\,\bar{y}_u$$

and its corresponding minimum variance is given by

$$\begin{aligned} \mathrm{V}(\Delta_1) &= a^2\mathrm{V}(\bar{x}^*_u) + (a+1)^2\,\mathrm{V}(\bar{x}^*_m) + c^2\,\mathrm{V}(\bar{y}_m) + (1-c)^2\,\mathrm{V}(\bar{y}_u) - 2(a+1)\,c\,\mathrm{Cov}(\bar{x}^*_m, \bar{y}_m) \\ &= a^2\left(\frac{\sigma^2}{q\,n} + \frac{A}{q\,n}\right) + (a+1)^2\left(\frac{\sigma^2}{p\,n} + \frac{A}{p\,n}\right) + c^2\left(\frac{\sigma^2}{p\,n}\right) + (1-c)^2\left(\frac{\sigma^2}{q\,n}\right) \\ &\quad -2(a+1)\,c\left(\frac{\sigma^2\rho}{p\,n}\right). \end{aligned}$$

3.3 Estimation of the change of mean for the current occasion in the presence of non-response on the second occasion

When there is non-response only on the second occasion, the minimum variance linear unbiased estimator of the change can be obtained as

$$\Delta_2 = a\,\bar{x}_u + b\,\bar{x}_m + c\,\bar{y}_m^* + d\,\bar{y}_u^*, \quad \text{where} \quad \bar{x}_m = \frac{1}{m}\sum_{i=1}^{m} x_i \quad \text{and} \quad \bar{x}_u = \frac{1}{u}\sum_{i=1}^{u} x_i.$$

Imposing the unbiasedness and minimum variance unbiased conditions, the optimum values of constants a and c are given by

$$a_{\text{opt}} = \frac{p\,q\,(\sigma^2 + A)\rho}{(\sigma^2 + A) - q^2\,\rho^2\sigma^2} - \frac{q\,((\sigma^2 + A) - q\,\rho^2\sigma^2)}{(\sigma^2 + A) - q^2\,\rho^2\sigma^2} \quad \text{and}$$

$$c_{\text{opt}} = \frac{p\,(\sigma^2 + A)}{(\sigma^2 + A) - q^2\,\rho^2\sigma^2} + \frac{p\,q\,\sigma^2\rho}{(\sigma^2 + A) - q^2\,\rho^2\sigma^2}.$$

Thus,

$$\Delta_2 = a\,\bar{x}_u - (a+1)\,\bar{x}_m + c\,\bar{y}_m^* + (1-c)\,\bar{y}_u^*$$

and its corresponding minimum variance is given by

$$\begin{aligned}
\mathrm{V}(\Delta_2) &= a^2\mathrm{V}(\bar{x}_u) + (a+1)^2\,\mathrm{V}(\bar{x}_m) + c^2\,\mathrm{V}(\bar{y}_m^*) + (1-c)^2\,\mathrm{V}(\bar{y}_u^*) - 2(a+1)\,c\,\mathrm{Cov}(\bar{x}_m, \bar{y}_m^*) \\
&= a^2\left(\frac{\sigma^2}{q\,n}\right) + (a+1)^2\left(\frac{\sigma^2}{p\,n}\right) + c^2\left(\frac{\sigma^2}{p\,n} + \frac{A}{p\,n}\right) + (1-c)^2\left(\frac{\sigma^2}{q\,n} + \frac{A}{q\,n}\right) \\
&\quad -2(a+1)\,c\left(\frac{\sigma^2\rho}{p\,n}\right).
\end{aligned}$$

3.4 Comparison between variances of the estimators of the change, Δ_0, Δ_{12}, Δ_1 and Δ_2

In this subsection, we carry out an analysis based on the loss in precision of Δ_{12}, Δ_1 and Δ_2 with respect to Δ_0. This loss is expressed in percentage and given by

$$L_{12} = \left[\frac{\mathrm{V}(\Delta_{12})}{\mathrm{V}(\Delta_0)} - 1\right] \times 100, \quad L_1 = \left[\frac{\mathrm{V}(\Delta_1)}{\mathrm{V}(\Delta_0)} - 1\right] \times 100, \quad \text{and} \quad L_2 = \left[\frac{\mathrm{V}(\Delta_2)}{\mathrm{V}(\Delta_0)} - 1\right] \times 100,$$

respectively. The losses in precision of Δ_{12}, Δ_1, Δ_2 with respect to Δ_0 for different values of ρ, ρ_2, σ_2^2, σ^2, W_2, f, and q are presented in Tables 1-2 and in Figure 1. It is assumed that $N = 300$ and $n = 50$. From these tables, we obtain the following conclusions:

(i) In the majority of the cases, the loss in precision is maximum at Δ_2 and minimum at Δ_1. Also, it can be seen that, in the majority of the cases, the loss in precision of Δ_{12} is less than that of Δ_2.

(ii) For the case $\sigma^2 < \sigma_2^2$, the loss in precision of all the estimators with respect to Δ_0 increases as the the values of σ_2^2 increase; see Figure 1(a).

(iii) For the case $\sigma^2 > \sigma_2^2$, the loss in precision of all the estimators with respect to Δ_0 decreases as the values of σ^2 increase; see Figure 1(b).

(iv) For the case $\sigma^2 = \sigma_2^2$, the loss in precision of all the estimators with respect to Δ_0 remain constant as the values of σ^2 and σ_2^2 increase; see Figure 1(c).

(v) For the case $\rho < \rho_2$, the loss in precision of all the estimators with respect to Δ_0 increases as the values of ρ increase; see Figure 1(d).

(vi) For the case $\rho > \rho_2$, the loss in precision of Δ_1 and Δ_2 with respect to Δ_0 remains constant as the values of ρ_2 increase, whereas the loss in precision of Δ_{12} with respect to Δ_0 decreases as the values of ρ_2 increase; see Figure 1(e).

(vii) For the case $\rho = \rho_2$, the loss in precision of Δ_{12} with respect to Δ_0 remains constant as the values of ρ and ρ_2 increase, whereas the loss in precision of Δ_1 and Δ_2 with respect to Δ_0 increases as the values of ρ and ρ_2 increase; see Figure 1(f).

(viii) The loss in precision of Δ_{12}, Δ_1, Δ_2 with respect to Δ_0 increases as the values of W_2 increase; see Figure 1(g).

(ix) The loss in precision of Δ_{12}, Δ_1, Δ_2 with respect to Δ_0 increases as the values of f increase; see Figure 1(h).

(x) The loss in precision of Δ_{12}, Δ_1, Δ_2 with respect to Δ_0 decreases as the values of q increase; see Figure 1(i).

Table 1. Loss in precision, expressed in percentage, of Δ_{12}, Δ_1, Δ_2 with respect to Δ_0 for several values of ρ, ρ_2, σ_2^2, σ^2.

ρ	ρ_2	q	f	W_2	σ_2^2	σ^2	L_{12}	L_1	L_2
				$\sigma^2 < \sigma_2^2$					
0.7	0.2	0.7	2.5	0.8	0.4	0.3	441.1	246.5	419.4
0.7	0.2	0.7	2.5	0.8	0.6	0.3	653.2	361.5	611.3
0.7	0.2	0.7	2.5	0.8	0.9	0.3	970.4	532.7	896.9
				$\sigma^2 > \sigma_2^2$					
0.6	0.2	0.3	1.5	0.6	0.2	0.3	108.5	65.8	128.2
0.6	0.2	0.3	1.5	0.6	0.2	0.7	47.0	28.7	55.8
0.6	0.2	0.3	1.5	0.6	0.2	0.9	36.7	22.4	43.5
				$\sigma^2 = \sigma_2^2$					
0.8	0.3	0.7	2.0	0.7	0.1	0.1	303.0	189.2	337.9
0.8	0.3	0.7	2.0	0.7	0.3	0.3	303.0	189.2	337.9
0.8	0.3	0.7	2.0	0.7	0.8	0.8	303.0	189.2	337.9
				$\rho < \rho_2$					
0.1	0.7	0.6	2.5	0.5	0.5	0.4	98.4	81.7	118.4
0.3	0.7	0.6	2.5	0.5	0.5	0.4	103.2	92.6	144.9
0.6	0.7	0.6	2.5	0.5	0.5	0.4	130.3	134.4	237.0
				$\rho > \rho_2$					
0.8	0.1	0.3	2.0	0.5	0.5	0.4	474.3	261.2	530.5
0.8	0.4	0.3	2.0	0.5	0.5	0.4	336.7	261.2	530.5
0.8	0.9	0.3	2.0	0.5	0.5	0.4	66.1	261.2	530.5
				$\rho = \rho_2$					
0.2	0.2	0.8	1.5	0.5	0.5	0.5	75	39.6	50.5
0.5	0.5	0.8	1.5	0.5	0.5	0.5	75	47.6	69.0
0.9	0.9	0.8	1.5	0.5	0.5	0.5	75	172.9	298.3

Table 2. Loss in precision, expressed in percentage, of Δ_{12}, Δ_1, Δ_2 with respect to Δ_0 for different values of W_2, f and q.

ρ	ρ_2	q	f	W_2	σ_2^2	σ^2	L_{12}	L_1	L_2
				W_2					
0.7	0.2	0.6	2.5	0.1	0.4	0.6	34.5	20.8	38.8
0.7	0.2	0.6	2.5	0.3	0.4	0.6	99.0	58.8	108.9
0.7	0.2	0.6	2.5	0.6	0.4	0.6	191.3	111.9	205.7
				f					
0.8	0.3	0.4	1.0	0.5	0.4	0.6	101.8	70.4	139.9
0.8	0.3	0.4	3.0	0.5	0.4	0.6	292.3	196.7	394.1
0.8	0.3	0.4	3.5	0.5	0.4	0.6	338.8	226.9	455.4
				q					
0.8	0.2	0.1	1.5	0.4	0.7	0.5	323.4	200.3	404.5
0.8	0.2	0.5	1.5	0.4	0.7	0.5	254.9	154.0	302.1
0.8	0.2	0.7	1.5	0.4	0.7	0.5	203.7	121.2	219.2
0.8	0.2	0.9	1.5	0.4	0.7	0.5	131.9	74.5	108.2

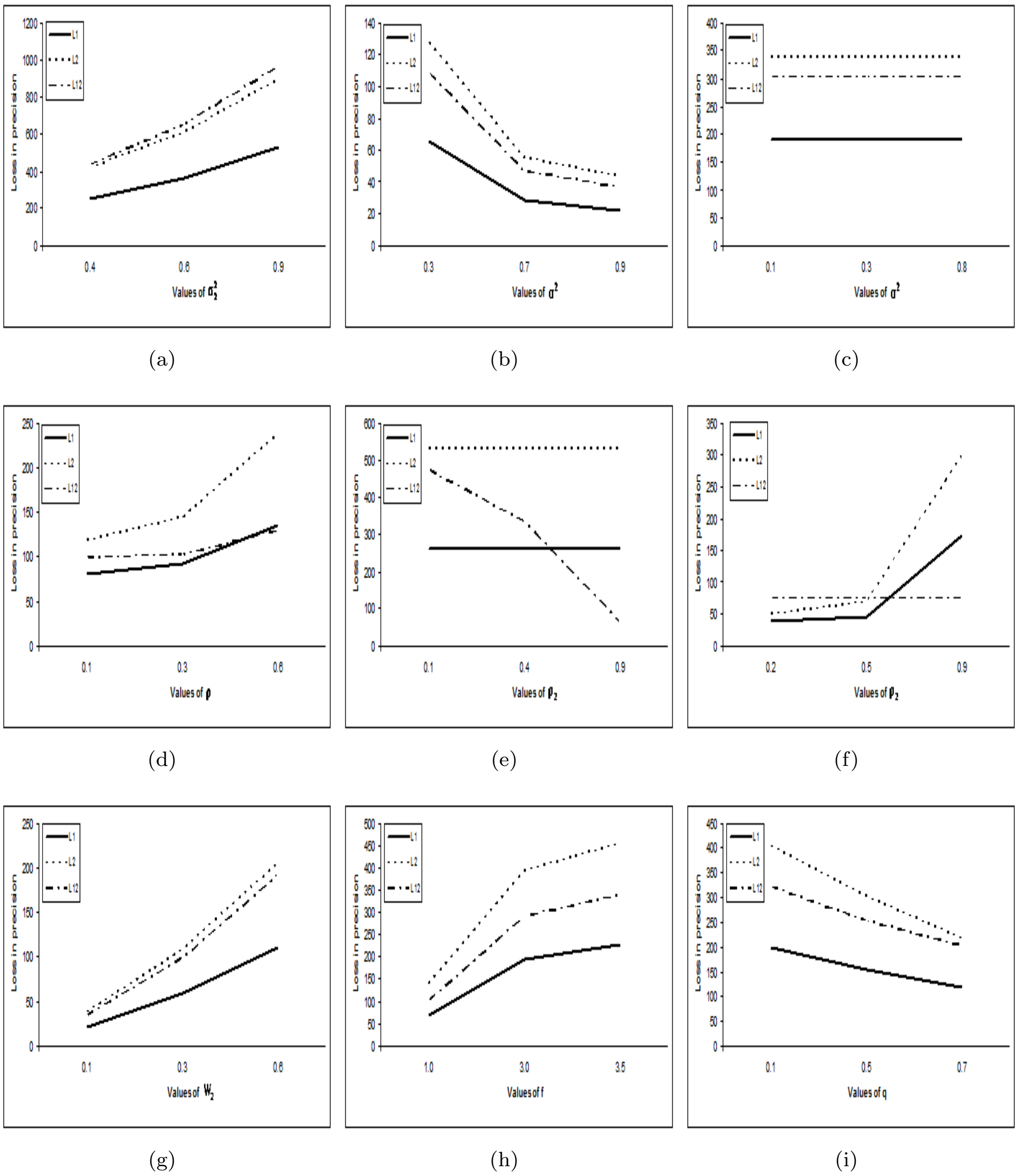


Figure 1. Loss in precision, expressed in percentage, of Δ_{12}, Δ_1, Δ_2 with respect to Δ_0 for (a)-(b) different values of σ_2^2 and σ^2, (c) the case σ^2=σ_2^2, (d)-(e) different values of ρ and ρ_2, (f) the case ρ=ρ_2, (g)-(h) different values of W_2 and f, and (i) different values of q.

4. Estimation of the Sum of Mean

4.1 Estimation of the sum of mean for current occasion in the presence of non-response on both occasions

Consider the following minimum variance linear unbiased estimator of the sum

$$z_{12} = a\,\bar{x}_u^* + b\,\bar{x}_m^* + c\,\bar{y}_m^* + d\,\bar{y}_u^*, \tag{4}$$

which expected value is given by

$$\begin{aligned}\mathrm{E}(z_{12}) &= a\mathrm{E}(\bar{x}_u^*) + b\mathrm{E}(\bar{x}_m^*) + c\mathrm{E}(\bar{y}_m^*) + d\mathrm{E}(\bar{y}_u^*)\\ &= a\bar{X}^* + b\bar{X}^* + c\bar{Y}^* + d\bar{Y}^* = (a+b)\bar{X}^* + (c+d)\bar{Y}^* = \bar{X}^* + \bar{Y}^*.\end{aligned}$$

Unbiasedness of z_{21} implies $a+b=1$ and $c+d=1$, so that $b = 1-a$ and $d = 1-c$. Substituting the value of b and d in Equation (4), we obtain

$$z_{12} = a\,\bar{x}_u^* + (1-a)\bar{x}_m^* + c\,\bar{y}_m^* + (1-c)\,\bar{y}_u^*. \tag{5}$$

The variance of z_{21} is

$$\mathrm{V}(z_{12}) = a^2\mathrm{V}(\bar{x}_u^*) + (1-a)^2\,\mathrm{V}(\bar{x}_m^*) + c^2\,\mathrm{V}(\bar{y}_m^*) + (1-c)^2\,\mathrm{V}(\bar{y}_u^*) + 2(1-a)\,c\,\mathrm{Cov}(\bar{x}_m^*, \bar{y}_m^*).$$

We wish to choose whose values of a and c that minimize $\mathrm{V}(z_{21})$. Equating the derivatives of $\mathrm{V}(z_{21})$ with respect to a and c to zero, it follows that the optimum values are

$$\begin{aligned}a_{\mathrm{opt}} &= \frac{p\,q\,(\sigma^2+A)(\rho\,\sigma^2+\rho_2 A)}{(\sigma^2+A)^2 - q^2(\rho\,\sigma^2+\rho_2 A)^2} + \frac{q\,((\sigma^2+A)^2 - q\,(\rho\,\sigma^2+\rho_2 A)^2)}{(\sigma^2+A)^2 - q^2(\rho\,\sigma^2+\rho_2 A)^2} \quad \text{and}\\ c_{\mathrm{opt}} &= \frac{p\,(\sigma^2+A)^2}{(\sigma^2+A)^2 - q^2(\rho\,\sigma^2+\rho_2 A)^2} - \frac{p\,q\,(\sigma^2+A)(\rho\,\sigma^2+\rho_2 A)}{(\sigma^2+A)^2 - q^2(\rho\,\sigma^2+\rho_2 A)^2}.\end{aligned}$$

Substituting the optimum values of a and c in Equation (5), we obtain

$$\begin{aligned}z_{12} &= \frac{q\,((\sigma^2+A)^2 - q\,(\rho\,\sigma^2+\rho_2 A)^2)}{(\sigma^2+A)^2 - q^2(\rho\,\sigma^2+\rho_2 A)^2}(\bar{y}_u^* + \bar{x}_u^*) + \frac{p\,(\sigma^2+A)^2}{(\sigma^2+A)^2 - q^2(\rho\,\sigma^2+\rho_2 A)^2}(\bar{y}_m^* + \bar{x}_m^*)\\ &\quad + \frac{p\,q\,(\sigma^2+A)(\rho\,\sigma^2+\rho_2 A)}{(\sigma^2+A)^2 - q^2(\rho\,\sigma^2+\rho_2 A)^2}\,[(\bar{x}_u^* - \bar{x}_m^*) + (\bar{y}_u^* - \bar{y}_m^*)]\\ &= \frac{p\,(\sigma^2+A)}{(\sigma^2+A) + q\,(\rho\,\sigma^2+\rho_2 A)}(\bar{y}_m^* + \bar{x}_m^*) + \frac{q\,((\sigma^2+A) + (\rho\,\sigma^2+\rho_2 A))}{(\sigma^2+A) + q\,(\rho\,\sigma^2+\rho_2 A)}(\bar{y}_u^* + \bar{x}_u^*).\end{aligned}$$

Thus, the optimum variance of z_{21} is given by

$$\mathrm{V}(z_{12}) = \frac{2}{n}(\sigma^2+A)\frac{(\sigma^2+A) + (\rho\,\sigma^2+\rho_2 A)}{(\sigma^2+A) + q\,(\rho\,\sigma^2+\rho_2 A)}. \tag{6}$$

We note that, for $(\rho\,\sigma^2+\rho_2 A)/(\sigma^2+A) > 0$, Equation (6) is minimum for $q=0$, i.e., the variance de z_{12} is minimized if the units on both occasions are independent. In this case

$$\mathrm{V}(z_{12}) = \frac{2}{n}(\sigma^2+A).$$

In case $\rho = \rho_2$, $\mathrm{V}(z_{21})$ reduces to

$$\mathrm{V}(z_{12}) = \frac{2}{n}(\sigma^2 + A)\frac{(1+\rho)}{(1+q\,\rho)},$$

while if $A = 0$, i.e., there is non-response, the $\mathrm{V}(z_{21})$ reduces to

$$\mathrm{V}(z_0) = \frac{2\sigma^2}{n}\frac{(1+\rho)}{(1+q\,\rho)},$$

where z_0 is the usual estimator of the sum for current occasion in the context of sampling on two occasions when there is complete response, that is,

$$z_0 = a\,\bar{x}_u + b\,\bar{x}_m + c\,\bar{y}_m + d\,\bar{y}_u.$$

4.2 Estimation of the sum of mean for current occasion in the presence of non-response on the first occasion

When there is non-response only on the first occasion, the minimum variance linear unbiased estimator of the change can be obtained as

$$z_1 = a\,\bar{x}_u^* + b\,\bar{x}_m^* + c\,\bar{y}_m + d\,\bar{y}_u, \quad \text{where} \quad \bar{y}_m = \frac{1}{m}\sum_{i=1}^{m} y_i \quad \text{and} \quad \bar{y}_u = \frac{1}{u}\sum_{i=1}^{u} y_i.$$

Imposing the unbiasedness and minimum variance unbiased conditions, the optimum values of constants a and c are given by

$$\begin{aligned} a_{\text{opt}} &= \frac{p\,q\,\sigma^2\rho}{(\sigma^2 + A) - q^2\,\rho^2\sigma^2} + \frac{q\,((\sigma^2 + A) - q\,\rho^2\sigma^2)}{(\sigma^2 + A) - q^2\,\rho^2\sigma^2} \quad \text{and} \\ c_{\text{opt}} &= \frac{p\,(\sigma^2 + A)}{(\sigma^2 + A) - q^2\,\rho^2\sigma^2} - \frac{p\,q\,(\sigma^2 + A)\rho}{(\sigma^2 + A) - q^2\,\rho^2\sigma^2}. \end{aligned}$$

Thus,

$$z_1 = a\,\bar{x}_u^* + (1-a)\bar{x}_m^* + c\,\bar{y}_m + (1-c)\,\bar{y}_u$$

and its corresponding minimum variance is given by

$$\begin{aligned} \mathrm{V}(z_1) &= a^2\mathrm{V}(\bar{x}_u^*) + (1-a)^2\,\mathrm{V}(\bar{x}_m^*) + c^2\,\mathrm{V}(\bar{y}_m) + (1-c)^2\,\mathrm{V}(\bar{y}_u) + 2(1-a)\,c\,\mathrm{Cov}(\bar{x}_m^*, \bar{y}_m) \\ &= a^2\left(\frac{\sigma^2}{q\,n} + \frac{A}{q\,n}\right) + (1-a)^2\left(\frac{\sigma^2}{p\,n} + \frac{A}{p\,n}\right) + c^2\left(\frac{\sigma^2}{p\,n}\right) + (1-c)^2\left(\frac{\sigma^2}{q\,n}\right) \\ &\quad + 2(1-a)\,c\left(\frac{\sigma^2\rho}{p\,n}\right). \end{aligned}$$

4.3 Estimation of the sum of mean for current occasion in the presence of non-response on the second occasion

When there is non-response only on the first occasion, the minimum variance linear unbiased estimator of the change can be obtained as

$$z_2 = a\,\bar{x}_u + b\,\bar{x}_m + c\,\bar{y}^*_m + d\,\bar{y}^*_u, \quad \text{where} \quad \bar{x}_m = \frac{1}{m}\sum_{i=1}^{m} x_i \quad \text{and} \quad \bar{x}_u = \frac{1}{u}\sum_{i=1}^{u} x_i.$$

Imposing the unbiasedness and minimum variance unbiased conditions, the optimum values of constants a and c are given by

$$a_{\rm opt} = \frac{p\,q\,(\sigma^2+A)\rho}{(\sigma^2+A) - q^2\,\rho^2\sigma^2} + \frac{q\,((\sigma^2+A) - q\,\rho^2\sigma^2)}{(\sigma^2+A) - q^2\,\rho^2\sigma^2} \quad \text{and}$$

$$c_{\rm opt} = \frac{p\,(\sigma^2+A)}{(\sigma^2+A) - q^2\,\rho^2\sigma^2} - \frac{p\,q\,\sigma^2\rho}{(\sigma^2+A) - q^2\,\rho^2\sigma^2}.$$

Thus,

$$z_2 = a\,\bar{x}_u + (1-a)\bar{x}_m + c\,\bar{y}^*_m + (1-c)\,\bar{y}^*_u$$

and its corresponding minimum variance is given by

$$\begin{aligned} \mathrm{V}(z_2) &= a^2\mathrm{V}(\bar{x}_u) + (1-a)^2\,\mathrm{V}(\bar{x}_m) + c^2\,\mathrm{V}(\bar{y}^*_m) + (1-c)^2\,\mathrm{V}(\bar{y}^*_u) + 2(1-a)\,c\,\mathrm{Cov}(\bar{x}_m, \bar{y}^*_m) \\ &= a^2\left(\frac{\sigma^2}{q\,n}\right) + (1-a)^2\left(\frac{\sigma^2}{p\,n} + \frac{A}{p\,n}\right) + c^2\left(\frac{\sigma^2}{p\,n} + \frac{A}{p\,n}\right) + (1-c)^2\left(\frac{\sigma^2}{q\,n} + \frac{A}{q\,n}\right) \\ &\quad + 2(1-a)\,c\left(\frac{\sigma^2\rho}{p\,n}\right). \end{aligned}$$

4.4 Comparison between variances of the estimators of the sum, z_0, z_{12}, z_1 and z_2

Once again, in this subsection, we carry out an analysis based on the loss in precision of of z_{12}, z_1 and z_2 with respect to z_0. This loss is expressed in percentage and given by

$$L_{12} = \left[\frac{\mathrm{V}(z_{12})}{\mathrm{V}(z_0)} - 1\right] \times 100, \quad L_1 = \left[\frac{\mathrm{V}(z_1)}{\mathrm{V}(z_0)} - 1\right] \times 100, \quad \text{and} \quad L_2 = \left[\frac{\mathrm{V}(z_2)}{\mathrm{V}(z_0)} - 1\right] \times 100,$$

respectively. The losses in precision of z_{12}, z_1, z_2 with respect to z_0 for different values of ρ, ρ_2, σ_2^2, σ^2, W_2, f and q are presented in Tables 3-4 and in Figure 2. It is assumed that $N = 300$ and $n = 50$. From these tables, we obtain the following conclusions:

(i) The loss in precision is maximum at z_{12} and minimum at z_1.

(ii) For the case $\sigma^2 < \sigma_2^2$, the loss in precision of all the estimators with respect to z_0 increases as the values of σ_2^2 increase; see Figure 2(a).

(iii) For the case $\sigma^2 > \sigma_2^2$ the loss in precision of all the estimators with respect to z_0 decreases as the values of σ^2 increase; see Figure 2(b).

(iv) For the case $\sigma^2 = \sigma_2^2$ the loss in precision of all the estimators with respect to z_0 remains constant as the values of σ^2 and σ_2^2 increase; see Figure 2(c).

(v) For the case $\rho < \rho_2$, the loss in precision of all the estimators with respect to z_0 decreases as the values of ρ increase; see Figure 2(d).

(vi) For the case $\rho > \rho_2$ the loss in precision of z_1 and z_2 with respect to z_0 remains constant as the values of ρ_2 increase, whereas the loss in precision of z_{12} with respect to z_0 increases as the values of ρ_2 increase; see Figure 2(e).

(vii) For the case $\rho = \rho_2$ the loss in precision of z_{12} with respect to z_0 remains constant as the values of ρ and ρ_2 increase, whereas the loss in precision of z_1 and z_2 with respect to z_0 decreases as the values of ρ and ρ_2 increase; see Figure 2(f).

(viii) The loss in precision of z_{12}, z_1, z_2 with respect to z_0 increases as the values of W_2 increase; see ; see Figure 2(g).

(ix) The loss in precision of z_{12}, z_1, z_2 with respect to z_0 increases as the values of f increase; see Figure 2(h).

(x) The loss in precision of z_{12} and z_1 with respect to z_0 increases as the values of q increase and the loss in precision of z_2 with respect to z_0 first decreases and after increase as values of q increase; see Figure 2(i).

Table 3. Loss in precision, expressed in percentage of z_{12}, z_1, z_2 with respect to z_0 for different values of ρ, ρ_2, σ_2^2 and σ^2.

ρ	ρ_2	q	f	W_2	σ_2^2	σ^2	L_{12}	L_1	L_2
				$\sigma^2 < \sigma_2^2$					
0.7	0.2	0.7	2.5	0.8	0.4	0.3	247.6	118.1	128.5
0.7	0.2	0.7	2.5	0.8	0.6	0.3	370.8	176.6	191.7
0.7	0.2	0.7	2.5	0.8	0.9	0.3	555.5	264.3	286.3
				$\sigma^2 > \sigma_2^2$					
0.6	0.2	0.3	1.5	0.6	0.2	0.3	50.7	22.9	33.7
0.6	0.2	0.3	1.5	0.6	0.2	0.7	21.8	9.9	14.6
0.6	0.2	0.3	1.5	0.6	0.2	0.9	16.9	7.7	11.4
				$\sigma^2 = \sigma_2^2$					
0.8	0.3	0.7	2.0	0.7	0.1	0.1	131.4	61.8	66.4
0.8	0.3	0.7	2.0	0.7	0.3	0.3	131.4	61.8	66.4
0.8	0.3	0.7	2.0	0.7	0.8	0.8	131.4	61.8	66.4
				$\rho < \rho_2$					
0.1	0.7	0.6	2.5	0.5	0.5	0.4	182.9	75.3	102.0
0.3	0.7	0.6	2.5	0.5	0.5	0.4	170.7	71.3	90.9
0.6	0.7	0.6	2.5	0.5	0.5	0.4	159.0	67.5	79.6
				$\rho > \rho_2$					
0.8	0.1	0.3	2.0	0.5	0.5	0.4	94.7	44.7	63.0
0.8	0.4	0.3	2.0	0.5	0.5	0.4	108.4	44.7	63.0
0.8	0.9	0.3	2.0	0.5	0.5	0.4	128.9	44.7	63.0
				$\rho = \rho_2$					
0.2	0.2	0.8	1.5	0.5	0.5	0.5	75	36.3	41.6
0.5	0.5	0.8	1.5	0.5	0.5	0.5	75	35.4	38.4
0.9	0.9	0.8	1.5	0.5	0.5	0.5	75	34.5	35.6

Table 4. Loss in precision, expressed in percentage of z_{12}, z_1, z_2 with respect to z_0 for different values of W_2, f and q.

ρ	ρ_2	q	f	W_2	σ_2^2	σ^2	L_{12}	L_1	L_2
				W_2					
0.7	0.2	0.6	2.5	0.1	0.4	0.6	15.2	7.3	8.6
0.7	0.2	0.6	2.5	0.3	0.4	0.6	45.5	21.6	25.2
0.7	0.2	0.6	2.5	0.6	0.4	0.6	90.7	42.9	49.6
				f					
0.8	0.3	0.4	1.0	0.5	0.4	0.6	28.9	13.0	16.9
0.8	0.3	0.4	3.0	0.5	0.4	0.6	86.3	38.2	49.5
0.8	0.3	0.4	3.5	0.5	0.4	0.6	100.6	44.4	57.5
				q					
0.8	0.2	0.1	1.5	0.4	0.7	0.5	60.0	25.9	45.2
0.8	0.2	0.5	1.5	0.4	0.7	0.5	72.9	34.0	41.1
0.8	0.2	0.7	1.5	0.4	0.7	0.5	77.9	37.3	40.4
0.8	0.2	0.9	1.5	0.4	0.7	0.5	82.1	40.4	41.0

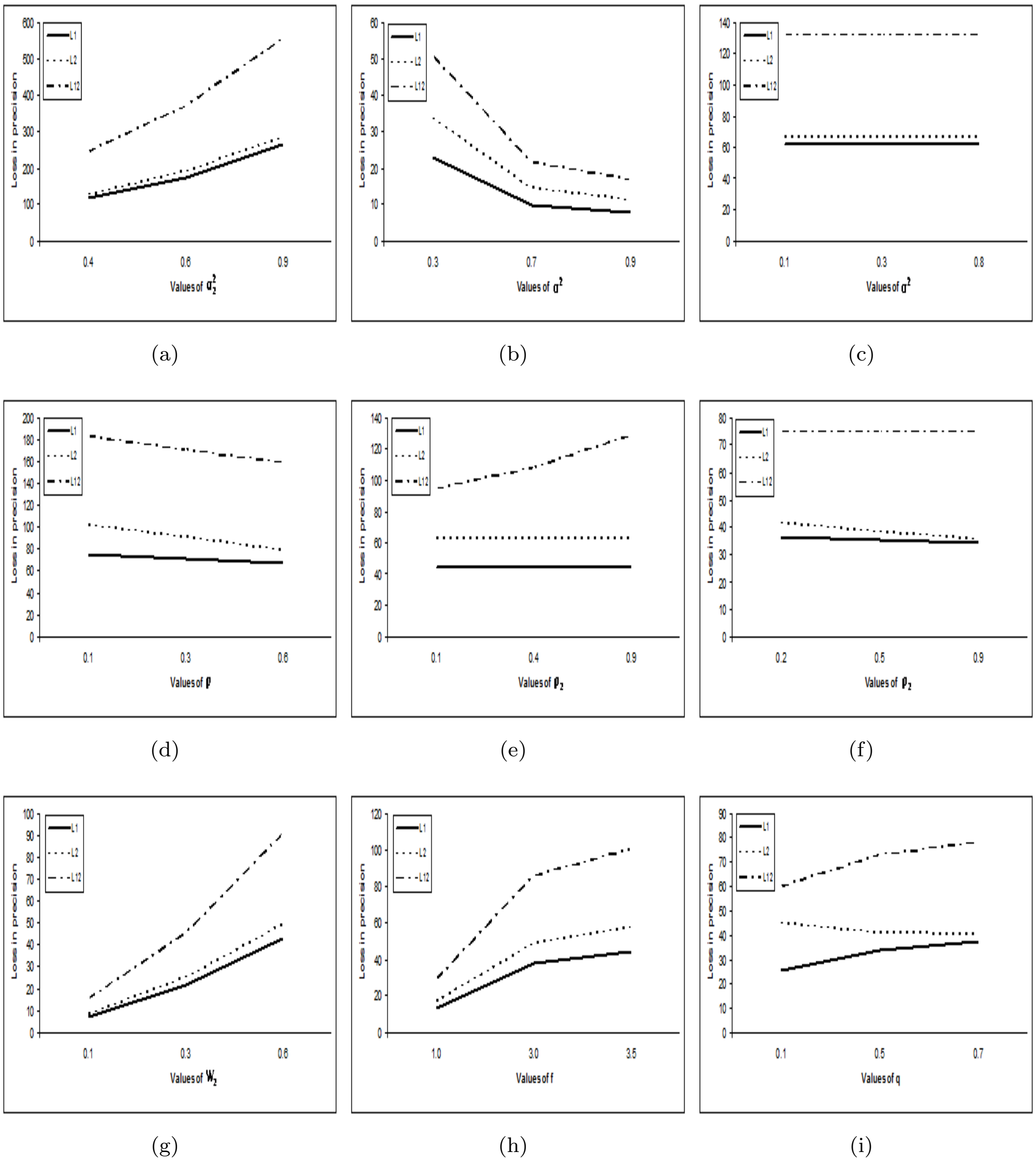


Figure 2. Loss in precision, expressed in percentage, of z_{12}, z_1, z_2 with respect to z_0 for (a)-(b) different values of σ_2^2 and σ^2, (c) the case σ^2=σ_2^2, (d)-(e) different values of ρ and ρ_2, (f) the case ρ=ρ_2, (g)-(h) different values of W_2 and f, and (i) different values of q.

5. 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 σ^2, σ_2^2, ρ, ρ_2, W_2, f and q. Let $N = 300$, $n = 50$, $c_0 = 1$, $c_1 = 4$, and $c_2 = 45$, 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)

$$C_{00} = 2nc_2. \tag{7}$$

Substituting the values of n and c_2 in Equation (7), 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

$$C_{12} = 2\left[c_o n + c_1 n_1 + \tfrac{c_2 n_2}{f}\right].$$

The expected cost is given by

$$\mathrm{E}(C_{12}) = 2n\left[c_0 + c_1 W_1 + \tfrac{c_2 W_2}{f}\right],$$

where $W_1 = N_1/N$ and $W_2 = N_2/N$, such that $W_1 + W_2 = 1$.

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

$$C_2 = 2c_0 n + c_1 n + \left[c_1 n_1 + \tfrac{c_2 n_2}{f}\right]$$

and the expected cost is given by

$$\mathrm{E}(C_2) = n\left[2c_0 + c_1(W_1 + 1) + \tfrac{c_2 W_2}{f}\right].$$

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

$$C_1 = \left[c_1 n_1 + \tfrac{c_2 n_2}{f}\right] + 2c_0 n + c_1 n,$$

which expected cost is expressed as

$$\mathrm{E}(C_1) = n\left[2c_0 + c_1(W_1 + 1) + \tfrac{c_2 W_2}{f}\right].$$

By equating the variances Δ_{12}, Δ_1, and Δ_2, respectively, to Δ_0 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 Δ_{12}, Δ_1 and Δ_2. The sample sizes associated with the three estimators which provide equal precision to the estimator $\mathrm{V}(\Delta_0)$ are denoted by n', n_1' and n_2'. The results of this exercise are presented in Tables 5-6 and in Figures 3-4. The sample sizes associated with the three estimators, which have the same precision than Δ_0, is maximum at Δ_{12} and minimum at Δ_1. It can be seen that in the majority of the cases the sample sizes for Δ_{12} is less than that of Δ_2. From these tables, we obtain the following conclusions:

(i) For the case $\sigma^2 < \sigma_2^2$, the saving in cost for all the estimators decreases as the values of σ_2^2 increase; see Figure 3(a).

(ii) The sample sizes for the three estimators, which have the same precision than Δ_0, increase as the values of σ_2^2 increase; see Figure 3(b).

(iii) For the case $\sigma^2 > \sigma_2^2$, the saving in cost for all the estimators increases as the values of σ^2 increase; see Figure 3(c).

(iv) The sample sizes for the three estimators, which have the same precision than Δ_0, decrease as the values of σ^2 increase; see Figure 3(d).

(v) For the case $\sigma^2 = \sigma_2^2$ the saving in cost for all the estimators remains constant as the values of σ^2 and σ_2^2 increase; see Figure 3(e).

(vi) The sample sizes for all the estimators, which have the same precision than Δ_0, remain constant; see Figure 3(f).

(vii) For the case $\rho < \rho_2$, the saving in cost for all the estimators decreases as the values of ρ increase; see Figure 3(g).

(viii) The sample sizes for the three estimators, which have the same precision than Δ_0, increases as the values of ρ increase; see Figure 3(h).

(ix) For the case $\rho > \rho_2$, the saving in cost for Δ_1 and Δ_2 remains constant as the values of ρ_2 increase, whereas for Δ_{12} the saving in cost increases as the values of ρ_2 increase; see Figure 3(i).

(x) The sample sizes for Δ_1 and Δ_2, which have the same precision than Δ_0, remain constant, whereas the sample size for Δ_{12}, which have the same precision than Δ_0, decreases; see Figure 3(j).

(xi) For the case $\rho = \rho_2$, the saving in cost for Δ_{12} remains constant as the values of ρ and ρ_2 increase, whereas for Δ_1 and Δ_2 the saving in cost decreases as the values of ρ and ρ_2 increase; see Figure 3(k).

(xii) The sample sizes for Δ_1 and Δ_2, which give equal precision to Δ_0 increase, whereas the sample size for Δ_{12}, which has the same precision than Δ_0, remains constant; see Figure 3(l).

(xiii) The saving in cost for all the estimators decreases as the values of W_2 increase; see Figure 4(a).

(xiv) The sample sizes associated with the three estimators, which have the same precision than Δ_0, increase as the values of W_2; see Figure 4(b).

(xv) The saving in cost increases as the values of f increase; see Figure 4(c).

(xvi) The sample sizes associated with the three estimators, which have the same precision than Δ_0, increase as the values of f increase; see Figure 4(d).

(xvii) The saving in cost increases as the values of q increase; see Figure 4(e).

(xviii) The sample sizes associated with the three estimators, which give equal precision to Δ_0, decreases as the values of q increase; see Figure 4(f).

Table 5. Sample sizes and corresponding expected cost of survey, which have the same precision than Δ_{12}, Δ_1 and Δ_2, with respect to Δ_0 for different values of ρ, ρ_2, σ_2^2 and σ^2.

ρ	ρ_2	q	f	W_2	σ_2^2	σ^2	n'	n'_1	n'_2	E(C_{12})	E(C_1)	E(C_2)
				$\sigma^2 < \sigma_2^2$								
0.7	0.2	0.5	2.5	0.4	0.4	0.3	187	130	204	3958.8	2036.3	3179.9
0.7	0.2	0.5	2.5	0.4	0.7	0.3	285	186	310	6031.6	2907.9	4844.3
0.7	0.2	0.5	2.5	0.4	0.8	0.3	317	205	346	6718.5	3195.1	5392.8
				$\sigma^2 > \sigma_2^2$								
0.6	0.2	0.3	1.5	0.3	0.2	0.3	77	67	82	1981.1	1187.1	1467.7
0.6	0.2	0.3	1.5	0.3	0.2	0.6	64	58	66	1632.8	1040.1	1181.3
0.6	0.2	0.3	1.5	0.3	0.2	0.9	59	56	61	1515.8	990.5	1084.9
				$\sigma^2 = \sigma_2^2$								
0.8	0.3	0.7	2.0	0.5	0.2	0.2	161	120	177	4587.8	2320.4	3407.6
0.8	0.3	0.7	2.0	0.5	0.6	0.6	161	120	177	4587.8	2320.4	3407.6
0.8	0.3	0.7	2.0	0.5	0.9	0.9	161	120	177	4587.8	2320.4	3407.6
				$\rho < \rho_2$								
0.1	0.7	0.6	2.5	0.5	0.4	0.6	77	72	82	1856.5	1220.4	1387.4
0.5	0.7	0.6	2.5	0.5	0.4	0.6	81	81	103	1952.1	1377.5	1748.2
0.8	0.7	0.6	2.5	0.5	0.4	0.6	107	119	181	2565.4	2021.9	3077.4
				$\rho > \rho_2$								
0.8	0.2	0.3	2.0	0.4	0.5	0.3	279	189	332	6917.1	3286.4	5780.1
0.8	0.6	0.3	2.0	0.4	0.5	0.3	175	189	332	4350.4	3286.4	5780.2
0.8	0.9	0.3	2.0	0.4	0.5	0.3	85	189	332	2114.3	3286.4	5780.2
				$\rho = \rho_2$								
0.3	0.3	0.8	1.5	0.3	0.6	0.4	84	68	75	2144	1220.8	1328.4
0.5	0.5	0.8	1.5	0.3	0.6	0.4	84	71	81	2144	1272.6	1445.4
0.8	0.8	0.8	1.5	0.3	0.6	0.4	84	91	120	2144	1629.9	2139.6

Table 6. Sample sizes and corresponding expected cost of survey, which have the same precision than Δ_{12}, Δ_1 and Δ_2, with respect to Δ_0 for different values of W_2, f and q.

ρ	ρ_2	q	f	W_2	σ_2^2	σ^2	n'	n'_1	n'_2	E(C_{12})	E(C_1)	E(C_2)
				W_2								
0.7	0.2	0.6	2.5	0.2	0.4	0.6	84	70	87	1304.6	897.5	1118.5
0.7	0.2	0.6	2.5	0.6	0.4	0.6	146	106	153	3903.7	1949.7	2812.5
0.7	0.2	0.6	2.5	0.8	0.4	0.6	176	123	184	5696.3	2608.6	3898.1
				f								
0.8	0.3	0.4	1.0	0.5	0.4	0.6	101	85	120	5146.6	2598.5	3658.5
0.8	0.3	0.4	1.5	0.5	0.4	0.6	125	101	153	4508.0	2336.0	3513.1
0.8	0.3	0.4	3.0	0.5	0.4	0.6	196	148	247	4119.2	2299.5	3829.5
				q								
0.8	0.2	0.2	1.5	0.4	0.7	0.5	205	145	243	6303.3	2959.1	4951.4
0.8	0.2	0.7	1.5	0.4	0.7	0.5	152	111	160	4676.5	2256.70	3256.6
0.8	0.2	0.9	1.5	0.4	0.7	0.5	116	87	104	3570.6	1780.3	2124.3

By equating the variances of z_{12}, z_1, and z_2 to V(z_0) 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 z_{12}, z_1 and z_2. The sample sizes associated with the three estimators, which provide the same precision of the estimator of the V(z_0), are denoted by n', n'_1 and n'_2. The results of this exercise are presented in Tables 7-8 and in Figures 5-6. The sample sizes associated with the three estimators, which give the same precision of z_0, is maximum at z_{12} and minimum at z_1. From these tables, we obtain the following conclusions:

(i) For the case $\sigma^2 < \sigma_2^2$, the saving in cost for all the estimators decreases as the values of σ_2^2 increase; see Figure 5(a).

(ii) The sample sizes for the three estimators, which have the same precision than z_0, increase as the values of σ_2^2 increase; see Figure 5(b).

(iii) For the case $\sigma^2 > \sigma_2^2$, the saving in cost for all the estimators increases as the values of σ^2 increase; see Figure 5(c).

(iv) The sample sizes for the three estimators, which have the same precision than z_0, decrease as the values of σ^2 increase; see Figure 5(d).

(v) For the case $\sigma^2 = \sigma_2^2$ the saving in cost for all the estimators remains constant as the values of σ^2 and σ_2^2 increase; see Figure 5(e).

(vi) The sample sizes for all the estimators, which have the same precision than z_0, remain constant; see Figure 5(f).

(vii) For the case $\rho < \rho_2$, the saving in cost for all the estimators increases as the values of ρ increase; see Figure 5(g).

(viii) The sample sizes for the three estimators, which have the same precision than z_0, decreases as the values of ρ increase; see Figure 5(h).

(ix) For the case $\rho > \rho_2$, the saving in cost for z_1 and z_2 remains constant as the values of ρ_2 increase, whereas for z_{12} the saving in cost decreases as the values of ρ_2 increase; see Figure 5(i).

(x) The sample sizes for z_1 and z_2, which have the same precision than z_0, remain constant, whereas the sample size for z_{12}, which has the same precision than z_0, increases; see Figure 5(j).

(xi) For the case $\rho = \rho_2$, the saving in cost for z_{12} remains constant as the values of ρ and ρ_2 increase, whereas for z_1 and z_2 the saving in cost increases as the values of ρ and ρ_2 increase; see Figure 5(k).

(xii) The sample sizes for z_1 and z_2, which have the same precision than z_0, decrease, whereas the sample size for z_{12}, which has the same precision than z_0, remains constant; see Figure 5(l).

(xiii) The saving in cost for all the estimators decreases as the values of W_2 increase; see Figure 6(a).

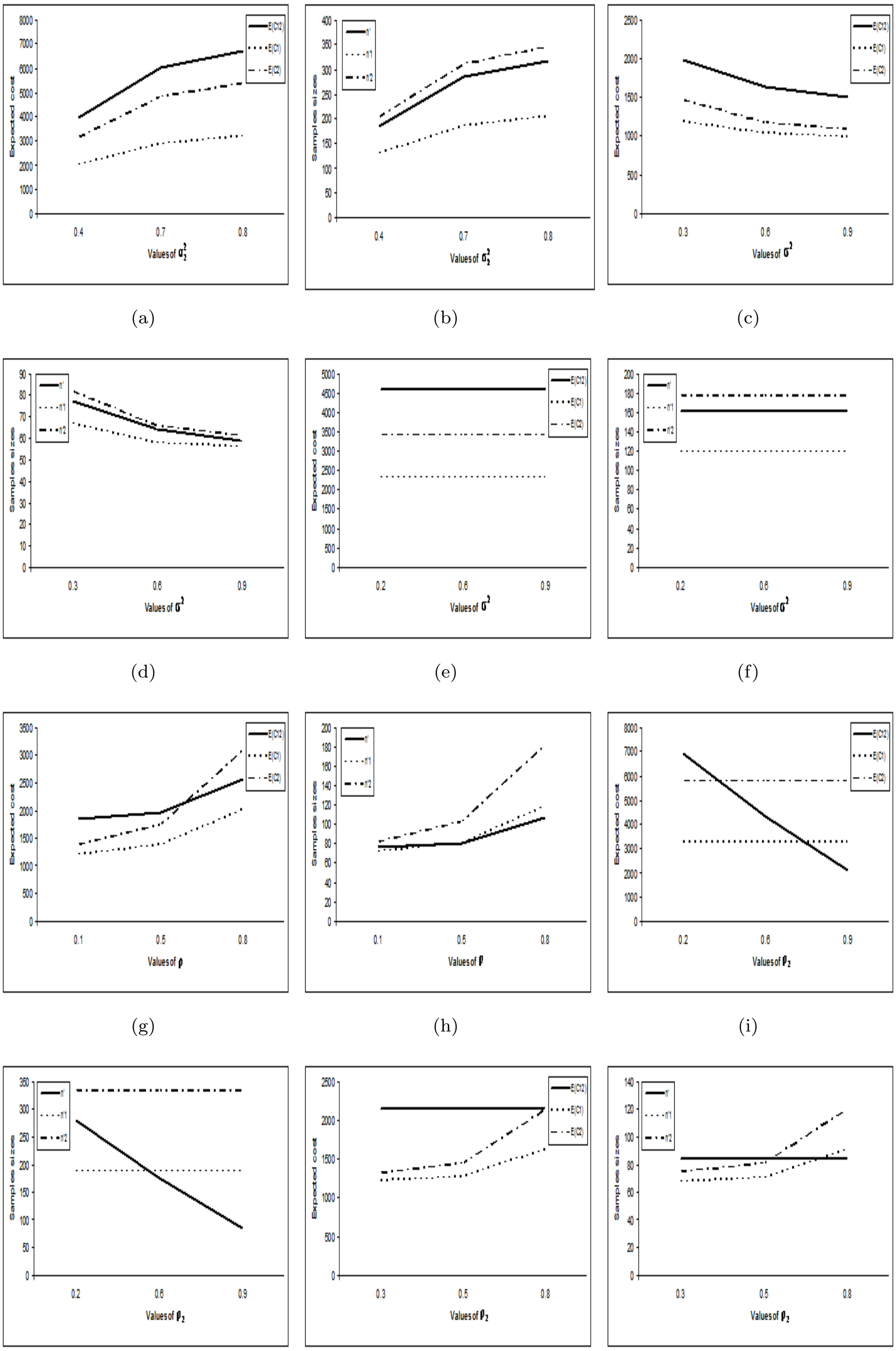
(xiv) The sample sizes associated with the three estimators, which have the same precision than z_0, increase as the values of W_2 increase; see Figure 6(b).

(xv) The sample sizes associated with the three estimators, which have the same precision than z_0, increase as the values of f increase; see Figure 6(c).

(xvi) The saving in cost for all the estimators increases as the values of f increase; see Figure 6(d).

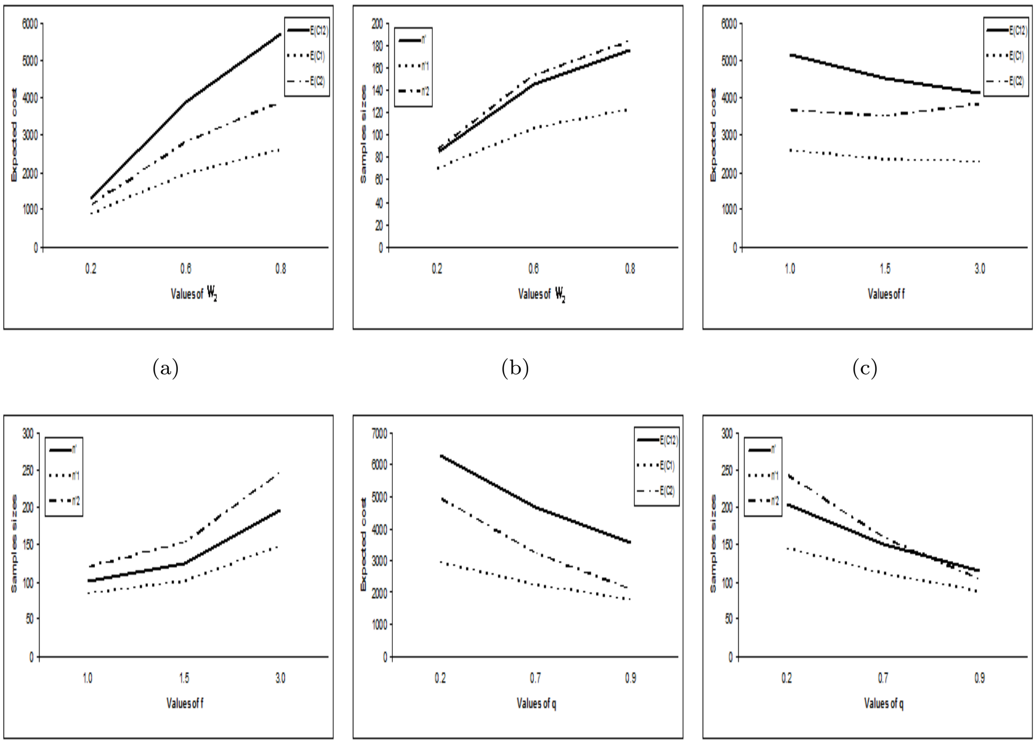
(xvii) The saving in cost for all the estimators decreases as the values of q increase. The saving in cost for z_2 increases and after decreases as q increases; see Figure 6(e).

(xviii) The sample sizes associated with the three estimators, which have the same precision than z_0, increase as the values of q increase, except the sample sizes for z_2 that give equal precision to z_0 first decreases and after increases as the values of q increase; see Figure 6(f).



(a) (b) (c) (d) (e) (f) (g) (h) (i) (j) (k) (l)

Figure 3. Sample sizes and corresponding expected cost of survey, which have the same precision than Δ_{12}, Δ_1 and Δ_2 with respect to Δ_0 for (a)-(b) different values of σ_2^2, (c)-(d) different values of σ^2, (e)-(f) the case $\sigma^2=\sigma_2^2$, (g)-(h) different values of ρ, (i)-(j) different values of ρ_2, and (k)-(l) the case $\rho=\rho_2$.



(a) (b) (c)

(d) (e) (f)

Figure 4. Sample sizes and corresponding expected cost of survey, which have the same precision than Δ_{12}, Δ_1 and Δ_2, with respect to Δ_0 for (a)-(b) different values of W_2, (c)-(d) different values of f, and (e)-(f) different values of q.

Table 7. Sample sizes and corresponding expected cost of survey, which have the same precision than z_{12}, z_1, z_2 with respect to z_0 for different values of W_2, f and q.

ρ	ρ_2	q	f	W_2	σ_2^2	σ^2	n'	n'_1	n'_2	E(C_{12})	E(C_1)	E(C_2)
				W_2								
0.7	0.2	0.6	2.5	0.2	0.4	0.6	65	57	58	1017.2	732.8	748.7
0.7	0.2	0.6	2.5	0.6	0.4	0.6	95	71	75	2555.9	1314.4	1376.5
0.7	0.2	0.6	2.5	0.8	0.4	0.6	110	78	83	3576.5	1663.4	1756.4
				f								
0.8	0.3	0.4	1.0	0.5	0.4	0.6	64	56	58	3288.4	1723.4	1784.0
0.8	0.3	0.4	1.5	0.5	0.4	0.6	72	60	63	2580.3	1372.8	1440.1
0.8	0.3	0.4	3.0	0.5	0.4	0.6	93	69	75	1956.5	1071.0	1158.6
				q								
0.8	0.2	0.2	1.5	0.4	0.7	0.5	82	62	72	2521.5	1260.2	1467.6
0.8	0.2	0.7	1.5	0.4	0.7	0.5	89	69	70	2739.0	1400.5	1431.8
0.8	0.2	0.9	1.5	0.4	0.7	0.5	91	70	71	2804.3	1432.3	1438.5

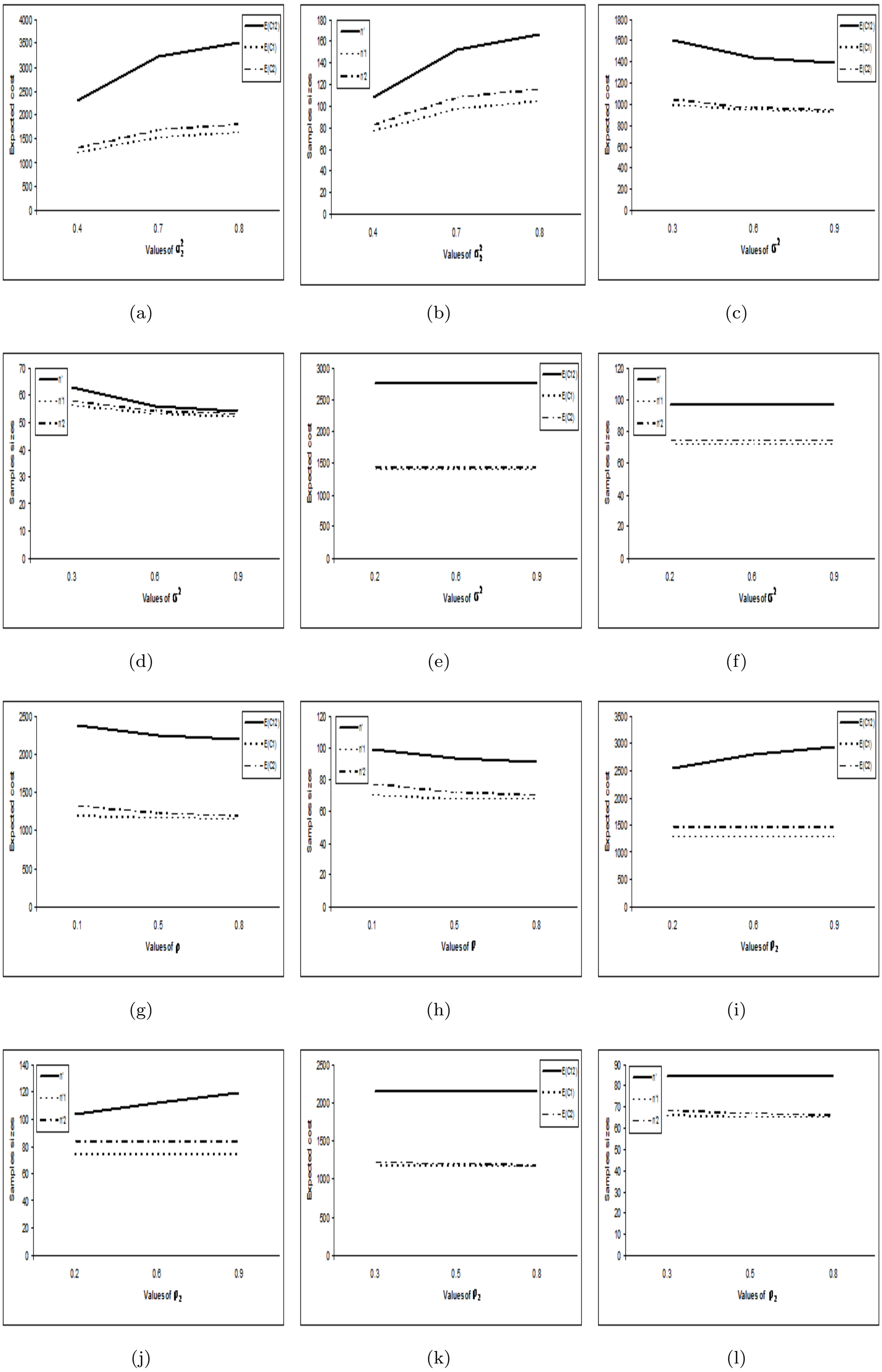
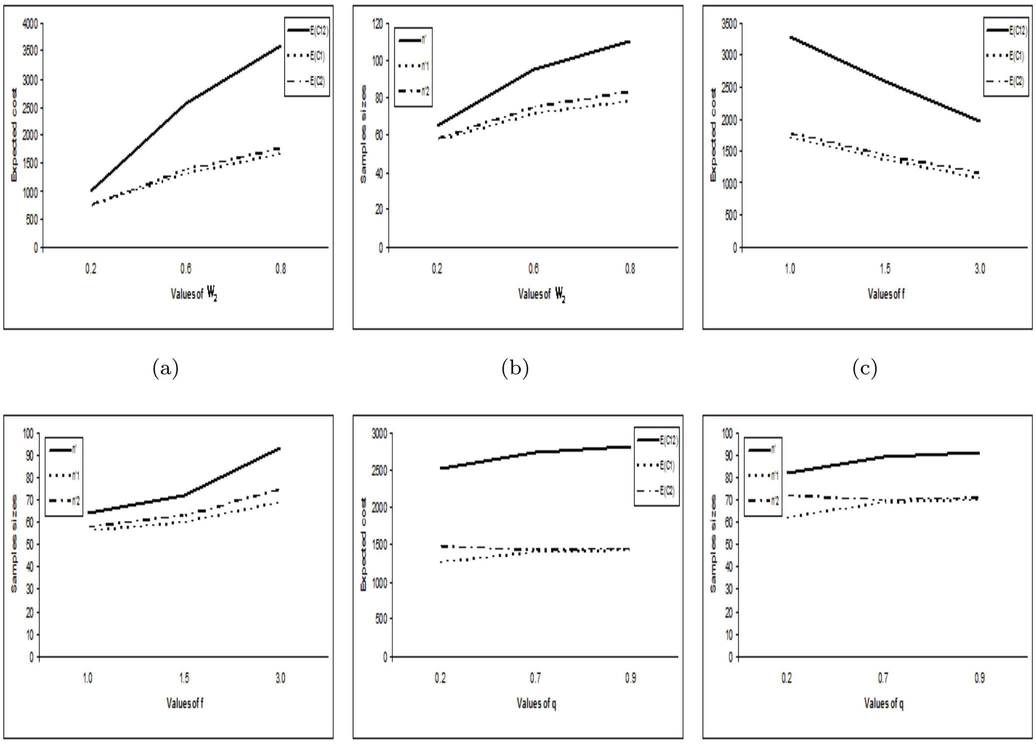


Figure 5. Sample sizes and corresponding expected cost of survey, which have the same precision than z_{12}, z_1 and z_2 with respect to z_0 for (a)-(b) different values of σ_2^2, (c)-(d) different values of σ^2, (e)-(f) the case σ^2=σ_2^2, (g)-(h) different values of ρ, (i)-(j) different values of ρ_2, and (k)-(l) the case ρ=ρ_2.



(a) (b) (c)

(d) (e) (f)

Figure 6. Sample sizes and corresponding expected cost of survey, which have the same precision than z_{12}, z_1 and z_2 with respect to z_0 for (a)-(b) different values of W_2, (c)-(d) different values of f, and (e)-(f) different values of q.

Table 8. Sample sizes and corresponding expected cost of survey, which have the same precision than z_{12}, z_1, z_2, with respect to z_0 for different values of ρ, ρ_2, σ_2^2 and σ^2.

ρ	ρ_2	q	f	W_2	σ_2^2	σ^2	n'	n'_1	n'_2	$\mathrm{E}(C_{12})$	$\mathrm{E}(C_1)$	$\mathrm{E}(C_2)$
				$\sigma^2 < \sigma_2^2$								
0.7	0.2	0.5	2.5	0.4	0.4	0.3	109	77	83	2301.5	1203.9	1301.1
0.7	0.2	0.5	2.5	0.4	0.7	0.3	152	97	108	3223.7	1515.9	1680.4
0.7	0.2	0.5	2.5	0.4	0.8	0.3	167	104	116	3530.8	1619.6	1806.4
				$\sigma^2 > \sigma_2^2$								
0.6	0.2	0.3	1.5	0.3	0.2	0.3	63	56	58	1605.7	992.5	1041.3
0.6	0.2	0.3	1.5	0.3	0.2	0.6	56	53	54	1443.1	941.5	966.1
0.6	0.2	0.3	1.5	0.3	0.2	0.9	54	52	53	1388.8	924.4	940.8
				$\sigma^2 = \sigma_2^2$								
0.8	0.3	0.7	2.0	0.5	0.2	0.2	97	72	74	2764.2	1388.9	1423.0
0.8	0.3	0.7	2.0	0.5	0.6	0.6	97	72	74	2764.2	1388.9	1423.0
0.8	0.3	0.7	2.0	0.5	0.9	0.9	97	72	74	2764.2	1388.9	1423.0
				$\rho < \rho_2$								
0.1	0.7	0.6	2.5	0.5	0.4	0.6	99	70	77	2378.3	1191.7	1312.8
0.5	0.7	0.6	2.5	0.5	0.4	0.6	93	68	72	2239.4	1162.6	1229.2
0.8	0.7	0.6	2.5	0.5	0.4	0.6	91	68	70	2184.7	1151.0	1192.2
				$\rho > \rho_2$								
0.8	0.2	0.3	2.0	0.4	0.5	0.3	103	74	83	2554.1	1284.0	1453.8
0.8	0.6	0.3	2.0	0.4	0.5	0.3	112	74	83	2786.7	1284.0	1453.8
0.8	0.9	0.3	2.0	0.4	0.5	0.3	119	74	83	2944.5	1284.0	1453.8
				$\rho = \rho_2$								
0.3	0.3	0.8	1.5	0.3	0.6	0.4	84	66	68	2144	1178.0	1213.5
0.5	0.5	0.8	1.5	0.3	0.6	0.4	84	65	67	2144	1173.5	1198.2
0.8	0.8	0.8	1.5	0.3	0.6	0.4	84	65	66	2144	1168.5	1180.9

6. Conclusions

In this article, we have used the HH technique for estimating the change of mean and the sum of mean in mail surveys. This problem has been conducted for current occasion in the context of sampling on two occasions when there is non-response (i) on both occasions, (ii) only on the first occasion and (iii) only on the second occasion. The obtained results have revealed that the loss in precision is maximum for the estimation of the sum of mean when there is non-response on both occasions. However, it is minimum for the estimation of the sum of mean when there is non-response only on the first occasion. In the majority of the cases, the loss in precision (expressed in percentage of the estimation of the change of mean), when there is non-response on both occasions, is less than that from the estimation of the change of mean when there is non-response only on the second occasion. Also, we have derived the sample sizes and the saving in cost for all the estimators that have the same precision than the estimator of the change of mean and sum of mean when there is no non-response.

Acknowledgements

The authors wish to thank the executive editor, Víctor Leiva and referees for their helpful comments that aided in improving this article.

References

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.

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).

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.

Singh, H.P., Kumar, S., 2010. Estimation of population product in presence of non-response in successive sampling. Statistical Papers (in press). `http://dx.doi.org/10.1007/s00362-008-0193-5`

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.