Improvement on the Non-response in the Population Ratio of Mean for Current Occasion in Sampling on Two Occasions

Amelia Victoria Garca Luengo
Departamento de Estadísticay Matemática Aplicada
Universidad de Almería. Almería. Spain
amgarcia@ual.es

Abstract

In this article, we attempt the problem of estimation of the population ratio 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 gain in efficiency of all the estimators over the direct estimate using no information gathered on the first occasion. We derive the sample sizes and the saving in cost for all the estimators, which have the same precision than the direct estimate using no information gathered on the first occasion. An empirical study that allows us to investigate the performance of the proposed strategy is carried out.

Keywords: Successive sampling, Non-response, Estimator of the ratio, Gain in efficiency.

AMS Subject Classification: 62D05.

1. Introduction

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. In many practical situations the estimate of the population ratio and product of two characters for the most recent occasion may be of considerable interest. The theory of estimation of the population ratio of two characters over two occasions has been considered by Rao (1957), Rao and Pereira (1968), Okafor (1992), Artés and García (2001), García and Artés (2002) among others. Further, García (2008) presented some sampling strategies for estimating, by a linear estimate, the population product of two characters over two occasions.

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 and García and Oña (2011) used 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 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 population ratio 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. An empirical study that allows us to investigate the performance of the proposed strategy is carried out.

2. The technique

Consider a finite population of N identifiable units. Let (x_i, y_i) be, for $i = 1, 2, \mathrm{K}, N$, 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 $\sigma_{x_j}^2$, $\sigma_{y_j}^2$; $j = 1, 2$ and $\sigma_{x_j}^2(2)$, $\sigma_{y_j}^2(2)$; $j = 1, 2$ denote the population variance and population variance pertaining to the non-response class, respectively. In addition, let $\bar{x}_{1m}^*$, $\bar{y}_{1m}^*$, $\bar{x}_{1u}^*$ and $\bar{y}_{1u}^*$ 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 $\bar{x}_{2m}^*$, $\bar{y}_{2m}^*$, $\bar{x}_{2u}^*$ and $\bar{y}_{2u}^*$. Thus, have the following setup:

x_i (y_i), the variable x (y) on ith occasion, $i = 1, 2$,

$R_1 = \dfrac{\bar{Y}_1}{\bar{X}_1}$ ($R_2 = \dfrac{\bar{Y}_2}{\bar{X}_2}$), the population ratio on the first (second) occasion,

$\hat{R}_1 = \dfrac{\bar{y}_1}{\bar{x}_1}$ ($\hat{R}_2 = \dfrac{\bar{y}_2}{\bar{x}_2}$), the estimator of the population ratio on the first (second) occasion,

$\hat{R}_{1m}^* = \dfrac{\bar{y}_{1m}^*}{\bar{x}_{1m}^*}$ ($\hat{R}_{2m}^* = \dfrac{\bar{y}_{2m}^*}{\bar{x}_{2m}^*}$), the estimator of the population ratio on the first (second) occasion based on the matched sample of m units,

$\hat{R}_{1u}^* = \dfrac{\bar{y}_{1u}^*}{\bar{x}_{1u}^*}$ ($\hat{R}_{2u}^* = \dfrac{\bar{y}_{2u}^*}{\bar{x}_{2u}^*}$), the estimator of the population ratio on the first (second) occasion based on the unmatched sample of u units.

ρ_1 (ρ_2), the correlation coefficients between the variables y_1 and x_1 (y_2 and x_2),

ρ_3 (ρ_4), the correlation coefficients between the variables y_2 and x_1 (y_1 and x_2),

ρ_5 (ρ_6), the correlation coefficients between the variables x_1 and x_2 (y_1 and y_2),

$\rho_{1(2)}$ ($\rho_{2(2)}$), the correlation coefficients between the variables $y_{1(2)}$ and $x_{1(2)}$ ($y_{2(2)}$ and $x_{2(2)}$),

$\rho_{3(2)}$ ($\rho_{4(2)}$), the correlation coefficients between the variables $y_{2(2)}$ and $x_{1(2)}$ ($y_{1(2)}$ and $x_{2(2)}$),

$\rho_{5(2)}$ ($\rho_{6(2)}$), the correlation coefficients between the variables $x_{1(2)}$ and $x_{2(2)}$ ($y_{1(2)}$ and $y_{2(2)}$).

1^{st} occasion $\rightarrow$	$\bar{R}^*_{1u}$	$\bar{R}^*_{1m}$	
2^{nd} occasion $\rightarrow$		$\bar{R}^*_{2m}$	$\bar{R}^*_{2u}$

where

$$\bar{y}^*_{1m} = \frac{m_1\bar{y}_{1m_1} + m_2\bar{y}_{1m_{h_2}}}{m}; \quad \bar{y}^*_{2m} = \frac{m_1\bar{y}_{2m_1} + m_2\bar{y}_{2m_{h_2}}}{m}$$

$$\bar{y}^*_{1u} = \frac{u_1\bar{y}_{1u_1} + u_2\bar{y}_{1u_{h_2}}}{u}; \quad \bar{y}^*_{2u} = \frac{u_1\bar{y}_{2u_1} + u_2\bar{y}_{2u_{h_2}}}{u}$$

$$\bar{x}^*_{1m} = \frac{m_1\bar{x}_{1m_1} + m_2\bar{x}_{1m_{h_2}}}{m}; \quad \bar{x}^*_{2m} = \frac{m_1\bar{x}_{2m_1} + m_2\bar{x}_{2m_{h_2}}}{m}$$

$$\bar{x}^*_{1u} = \frac{u_1\bar{x}_{1u_1} + u_2\bar{x}_{1u_{h_2}}}{u}; \quad \bar{x}^*_{2u} = \frac{u_1\bar{x}_{2u_1} + u_2\bar{x}_{2u_{h_2}}}{u}$$

$$i, j = 1,2 \begin{cases} \bar{y}_{jm_1} = \frac{1}{m_1}\sum_{i=1}^{m_1} y_{ji} \qquad \bar{x}_{jm_1} = \frac{1}{m_1}\sum_{i=1}^{m_1} x_{ji} \qquad \bar{y}_{jm_{h_2}} = \frac{1}{m_{h_2}}\sum_{l=1}^{m_{h_2}} y_{jl} \\ \bar{x}_{jm_{h_2}} = \frac{1}{m_{h_2}}\sum_{l=1}^{m_{h_2}} x_{jl} \qquad \bar{y}_{ju_1} = \frac{1}{u_1}\sum_{\alpha=1}^{u_1} y_{j\alpha} \qquad \bar{x}_{ju_1} = \frac{1}{u_1}\sum_{\alpha=1}^{u_1} x_{j\alpha} \\ \bar{y}_{ju_{h_2}} = \frac{1}{u_{h_2}}\sum_{\beta=1}^{u_{h_2}} y_{j\beta} \quad \bar{x}_{ju_{h_2}} = \frac{1}{u_{h_2}}\sum_{\beta=1}^{u_{h_2}} x_{j\beta} \end{cases}$$

It can be easily seen that (see Singh and Kumar 2010, p. 979)

$$Cov(\bar{R}^*_{1u}, \bar{R}^*_{1m}) = Cov(\bar{R}^*_{1u}, \bar{R}^*_{2m}) = Cov(\bar{R}^*_{1u}, \bar{R}^*_{2u}) = Cov(\bar{R}^*_{1m}, \bar{R}^*_{2u}) = Cov(\bar{R}^*_{2m}, \bar{R}^*_{2u}) = 0.$$

3. Estimation of the population ratio of mean

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

We wish to estimate, R_2, the population ratio for the second period by a linear estimate (HH technique) of the form

$$\hat{R}_2^* = a\hat{R}_{1u}^* + b\hat{R}_{1m}^* + c\hat{R}_{2m}^* + d\hat{R}_{2u}^*$$

We have

$$E(\hat{R}_{1u}^*) = E(\hat{R}_{1m}^*) = R_1 \quad and \quad E(\hat{R}_{2u}^*) = E(\hat{R}_{2m}^*) = R_2$$

we find that

$$E(\hat{R}_2^*) = (a+b)R_1 + (c+d)R_2$$

If we now require that $\hat{R}_2^*$ be an unbiased estimate of R_2, we must have

$$a+b=0 \qquad and \qquad c+d=1$$

so that

$$\hat{R}_2^* = a\left(\hat{R}_{1u}^* - \hat{R}_{1m}^*\right) + c\hat{R}_{2m}^* + (1-c)\hat{R}_{2u}^*$$

The variance of $\hat{R}_2^*$ is given by

$$V(\hat{R}_2^*) = a^2\left(\frac{1}{q}+\frac{1}{p}\right)\frac{1}{n\bar{X}_1^2}D^* + c^2\frac{1}{pn\bar{X}_2^2}E^* + (1-c)^2\frac{1}{qn\bar{X}_2^2}E^* - 2ac\frac{1}{pn\bar{X}_1\bar{X}_2}C^*$$

where

$$D^* = \{A + W_2(k-1)A_{(2)}\}; \quad E^* = \{B + W_2(k-1)B_{(2)}\}$$

$$W_2 = N_2/N; \quad k = m_2/m_{h_2} = u_2/u_{h_2}$$

$$A = S_{y_1}^2 + R_1^2 S_{x_1}^2 - 2R_1 Cov(y_1, x_1); \quad A_{(2)} = S_{y_1}^2(2) + R_1^2 S_{x_1}^2(2) - 2R_1 Cov(y_1, x_1)_{(2)}$$

$$B = S_{y_2}^2 + R_2^2 S_{x_2}^2 - 2R_2 Cov(y_2, x_2); \quad B_{(2)} = S_{y_2}^2(2) + R_2^2 S_{x_2}^2(2) - 2R_2 Cov(y_2, x_2)_{(2)}$$

$$C^* = [Cov(y_1, y_2) - R_1 Cov(y_2, x_1) - R_2 Cov(y_1, x_2) + R_1 R_2 Cov(x_1, x_2)]$$
$$+W_2(k-1)[Cov(y_1, y_2)_{(2)} - R_1 Cov(y_2, x_1)_{(2)} - R_2 Cov(y_1, x_2)_{(2)} + R_1 R_2 Cov(x_1, x_2)_{(2)}]$$

We wish to choose values of a and c that minimize $V(\hat{R}_2^*)$. Equating to zero the derivatives of $V(\hat{R}_2^*)$ with respect to a and c, it follows that the optimum values are

$$a_{opt} = \frac{pq\bar{X}_1 E^* C^*}{\bar{X}_2(D^* E^* - q^2 C^{*2})} \quad and \quad c_{opt} = \frac{pD^* E^*}{D^* E^* - q^2 C^{*2}}$$

Thus, the estimate with optimum values for a and c may be written

$$\hat{R}_2^{**} = \frac{pq\bar{X}_1 E^* C^*}{\bar{X}_2(D^*E^* - q^2C^{*2})}(\hat{R}_{1u}^* - \hat{R}_{1m}^*) + \frac{pD^*E^*}{D^*E^* - q^2C^{*2}}\hat{R}_{2m}^* + \left(1 - \frac{pD^*E^*}{D^*E^* - q^2C^{*2}}\right)\hat{R}_{2u}^* \quad (1)$$

and its variance is

$$V(\hat{R}_2^{**}) = \frac{E^*}{\bar{X}_2^2 n}\frac{D^*E^* - qC^{*2}}{D^*E^* - q^2C^{*2}} \quad (2)$$

Note that if $q = 0$, $p = 1$, complete matching or $p = 0$, $q = 1$, no matching this variance Eq. (2) has the same value,

$$V(\hat{R}_2^{**}) = \frac{E^*}{\bar{X}_2^2 n}$$

Thus, for current estimates, equal precision is obtained either by keeping the same sample or by changing it on every occasion. If $\bar{X}_1 = \bar{X}_2$, the estimate give by Eq. (1) is somewhat simplified

$$\hat{R}_2^{**} = \frac{pqE^* C^*}{(D^*E^* - q^2C^{*2})}(\hat{R}_{1u}^* - \hat{R}_{1m}^*) + \frac{pD^*E^*}{D^*E^* - q^2C^{*2}}\hat{R}_{2m}^* + \left(1 - \frac{pD^*E^*}{D^*E^* - q^2C^{*2}}\right)\hat{R}_{2u}^*$$

but its variance is unchanged, that is,

$$V(\hat{R}_2^{**}) = \frac{E^*}{\bar{X}_2^2 n}\frac{D^*E^* - qC^{*2}}{D^*E^* - q^2C^{*2}}$$

Note, also, that an estimate for the first occasion is given by Eq. (1) simply by interchanging R_1's and R_2's if the estimate for the first occasion can await a time until data for both occasions are available.

$$\hat{R}_1^{**} = \frac{pq\bar{X}_2 E^* C^*}{\bar{X}_1(D^*E^* - q^2C^{*2})}(\hat{R}_{2u}^* - \hat{R}_{2m}^*) + \frac{pD^*E^*}{D^*E^* - q^2C^{*2}}\hat{R}_{1m}^* + \left(1 - \frac{pD^*E^*}{D^*E^* - q^2C^{*2}}\right)\hat{R}_{1u}^*$$

Its variance is

$$V(\hat{R}_1^{**}) = \frac{D^*}{\bar{X}_1^2 n}\frac{D^*E^* - qC^{*2}}{D^*E^* - q^2C^{*2}}$$

Equating to zero the derivative of $V(\hat{R}_2^{**})$ with respect to q, we find that the variance of $V(\hat{R}_2^{**})$ will have its minimum value if we choose

$$q_{opt}^{(0)} = \frac{D^*E^* - \sqrt{D^{*2}E^{*2} - C^{*2}D^*E^*}}{C^{*2}} \quad (3)$$

and

$$V_{min}(\hat{R}_2^{**}) = \frac{E^*}{\bar{X}_2^2 n}\frac{D^*E^* + \sqrt{D^{*2}E^{*2} - C^{*2}D^*E^*}}{2D^*E^*}$$

However, if only the estimate using information gathered on the second occasion is considered, the estimator of the population ratio is

$$\hat{R}^{*} = p\hat{R}^{*}_{2m} + q\hat{R}^{*}_{2u}$$

and its variance is

$$V(\hat{R}^{*}) = \frac{E^{*}}{\bar{X}_2^2 n}$$

and we find

$$\frac{E^{*}}{\bar{X}_2^2 n}\frac{D^{*}E^{*} + \sqrt{D^{*2}E^{*2} - C^{*2}D^{*}E^{*}}}{2D^{*}E^{*}} \leq \frac{E^{*}}{\bar{X}_2^2 n}$$

3.2 Estimation of the population ratio 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 for the population ratio on current occasion can be obtained as follows:

$$\hat{R}^{*}_{21} = a\left(\hat{R}^{*}_{1u} - \hat{R}^{*}_{1m}\right) + c\hat{R}_{2m} + (1-c)\hat{R}_{2u} \quad where \quad \hat{R}_{2m} = \frac{\bar{y}_{2m}}{\bar{x}_{2m}} \quad and \quad \hat{R}_{2u} = \frac{\bar{y}_{2u}}{\bar{x}_{2u}}$$

The variance of $\hat{R}^{*}_{21}$ is given by

$$V(\hat{R}^{*}_{21}) = a^2\left(\frac{1}{q} + \frac{1}{p}\right)\frac{1}{n\bar{X}_1^2}D^{*} + c^2\frac{1}{pn\bar{X}_2^2}B + (1-c)^2\frac{1}{qn\bar{X}_2^2}B - 2ac\frac{1}{pn\bar{X}_1\bar{X}_2}k_1$$

which is minimum when

$$a_{opt} = \frac{pq\bar{X}_1 Bk_1}{\bar{X}_2(D^{*}B - q^2k_1^2)} \quad and \quad c_{opt} = \frac{pD^{*}B}{D^{*}B - q^2k_1^2}$$

where

$$k_1 = [Cov(y_1, y_2) - R_1Cov(y_2, x_1) - R_2Cov(y_1, x_2) + R_1R_2Cov(x_1, x_2)]$$

$$V(\hat{R}_{2m}) = \frac{1}{pn\bar{X}_2^2}B; \quad V(\hat{R}_{2u}) = \frac{1}{qn\bar{X}_2^2}B$$

Thus the estimator $\hat{R}^{*}_{21}$ turns out to be

$$\hat{R}^{**}_{21} = \frac{pq\bar{X}_1 Bk_1}{\bar{X}_2(D^{*}B - q^2k_1^{*2})}\left(\hat{R}^{*}_{1u} - \hat{R}^{*}_{1m}\right) + \frac{pD^{*}B}{D^{*}B - q^2k_1^{*2}}\hat{R}_{2m} + \left(1 - \frac{pD^{*}B}{D^{*}B - q^2k_1^{*2}}\right)\hat{R}_{2u}$$

with the variance

$$V(\hat{R}^{**}_{21}) = \frac{B}{\bar{X}_2^2 n}\frac{D^{*}B - qk_1^2}{D^{*}B - q^2k_1^2}$$

The optimum fraction to be unmatched is given by

$$q_{opt}^{(1)} = \frac{D^*B - \sqrt{D^{*2}B^2 - k_1^2 D^*B}}{k_1^2} \tag{4}$$

and thus the minimum variance of $\hat{R}_{21}^{**}$ is

$$V_{min}(\hat{R}_{21}^{**}) = \frac{B}{\bar{X}_2^2 n}\frac{D^*B + \sqrt{D^{*2}B^2 - k_1^2 D^*B}}{2D^*B}$$

3.3 Estimation of the population ratio 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 for the population ratio on current occasion can be obtained as follows:

$$\hat{R}_{22}^{*} = a\left(\hat{R}_{1u} - \hat{R}_{1m}\right) + c\hat{R}_{2m}^{*} + (1-c)\hat{R}_{2u}^{*} \quad where \quad \hat{R}_{1m} = \frac{\bar{y}_{1m}}{\bar{x}_{1m}} \quad and \quad \hat{R}_{1u} = \frac{\bar{y}_{1u}}{\bar{x}_{1u}}$$

The variance of $\hat{R}_{22}^{*}$ is given by

$$V(\hat{R}_{22}^{*}) = a^2\left(\frac{1}{q} + \frac{1}{p}\right)\frac{1}{n\bar{X}_1^2}A + c^2\frac{1}{pn\bar{X}_2^2}E^* + (1-c)^2\frac{1}{qn\bar{X}_2^2}E^* - 2ac\frac{1}{pn\bar{X}_1\bar{X}_2}k_1$$

which is minimum when

$$a_{opt} = \frac{pq\bar{X}_1E^*k_1}{\bar{X}_2(AE^* - q^2k_1^2)} \quad and \quad c_{opt} = \frac{pAE^*}{AE^* - q^2k_1^2}$$

where

$$k_1 = [Cov(y_1, y_2) - R_1Cov(y_2, x_1) - R_2Cov(y_1, x_2) + R_1R_2Cov(x_1, x_2)]$$

$$V(\hat{R}_{1m}) = \frac{1}{pn\bar{X}_1^2}A; \quad V(\hat{R}_{1u}) = \frac{1}{qn\bar{X}_1^2}A$$

Thus the estimator $\hat{R}_{22}^{*}$ turns out to be

$$\hat{R}_{22}^{**} = \frac{pq\bar{X}_1E^*k_1}{\bar{X}_2(AE^* - q^2k_1^2)}\left(\hat{R}_{1u} - \hat{R}_{1m}\right) + \frac{pAE^*}{AE^* - q^2k_1^2}\hat{R}_{2m}^{*} + \left(1 - \frac{pAE^*}{AE^* - q^2k_1^2}\right)\hat{R}_{2u}^{*}$$

with the variance

$$V(\hat{R}_{22}^{**}) = \frac{E^*}{\bar{X}_2^2 n}\frac{E^*A - qk_1^2}{E^*A - q^2k_1^2}$$

The optimum fraction to be unmatched is given by

$$q_{opt}^{(2)} = \frac{AE^* - \sqrt{E^{*2}A^2 - k_1^2E^*A}}{k_1^2} \tag{5}$$

and thus the minimum variance of $\hat{R}_{22}^{**}$ is

$$V_{min}(\hat{R}_{22}^{**}) = \frac{E^*}{\bar{X}_2^2 n}\frac{E^*A+\sqrt{E^{*2}A^2-k_1^2E^*A}}{2E^*A}$$

3.4 Comparison between variances of the estimators, $\hat{R}^*$, $\hat{R}_2^{**}$, $\hat{R}_{21}^{**}$ and $\hat{R}_{22}^{**}$

In this subsection, we carry out an analysis based on the gain in precision of $\hat{R}_2^{**}$, $\hat{R}_{21}^{**}$ and $\hat{R}_{22}^{**}$ with respect to $\hat{R}^*$.

$$G_{(1)} = \frac{V(\hat{R}^*)}{V(\hat{R}_2^{**})} = \frac{D^*E^*-q^2C^{*2}}{D^*E^*-qC^{*2}} \tag{6}$$

or

$$G_{opt(1)} = \frac{V(\hat{R}^*)}{V_{min}(\hat{R}_2^{**})} = \frac{2D^*E^*}{D^*E^*+\sqrt{D^{*2}E^{*2}-C^{*2}D^*E^*}} \tag{7}$$

$$G_{(2)} = \frac{V(\hat{R}^*)}{V(\hat{R}_{21}^{**})} = \frac{E^*}{B}\frac{D^*B-q^2k_1^2}{D^*B-qk_1^2} \tag{8}$$

or

$$G_{opt(2)} = \frac{V(\hat{R}^*)}{V_{min}(\hat{R}_{21}^{**})} = \frac{2D^*E^*}{D^*B+\sqrt{D^{*2}B^2-k_1^2D^*B}} \tag{9}$$

$$G_{(3)} = \frac{V(\hat{R}^*)}{V(\hat{R}_{22}^{**})} = \frac{E^*A-q^2k_1^2}{E^*A-qk_1^2} \tag{10}$$

or

$$G_{opt(3)} = \frac{V(\hat{R}^*)}{V_{min}(\hat{R}_{22}^{**})} = \frac{2E^*A}{E^*A+\sqrt{E^{*2}A^2-k_1^2E^*A}} \tag{11}$$

Now, we assume that

$$C_{y_1} = C_{y_2} = C_{x_1} = C_{x_2} = C_0, \quad C_{y_1(2)} = C_{y_2(2)} = C_{x_1(2)} = C_{x_2(2)} = C_{0(2)}$$
$$\rho_1 = \rho_2 = \rho_3 = \rho_4 = \rho, \quad \rho_{1(2)} = \rho_{2(2)} = \rho_{3(2)} = \rho_{4(2)} = \rho_{(2)}$$
$$\rho_5 = \rho_6 = \rho_0, \quad \rho_{5(2)} = \rho_{6(2)} = \rho_{0(2)}$$

The expressions of D^*, E^* and C^* becomes

$$D^* = 2\bar{Y}_1^2 d, \quad E^* = 2\bar{Y}_2^2 d, \quad \textit{and} \quad C^* = 2\bar{Y}_1\bar{Y}_2 t$$

where

$d = (1-\rho)C_0^2 + W_2(k-1)(1-\rho_{(2)})C_{0(2)}^2$, $t = (\rho_0-\rho)C_0^2 + W_2(k-1)(\rho_{0(2)}-\rho_{(2)})C_{0(2)}^2$ and the expressions (6), (7), (8), (9), (10) and (11) are given by

$$G_{(1)} = \frac{d^2 - q^2 t^2}{d^2 - qt^2} \qquad and \qquad G_{opt(1)} = \frac{2d}{d + \sqrt{d^2 - t^2}}$$

$$G_{(2)} = \frac{d\left((1-\rho)d - (\rho_0 - \rho)^2 q^2 C_0^2\right)}{(1-\rho)\left((1-\rho)d - (\rho_0 - \rho)^2 qC_0^2\right)C_0^2}$$

$$G_{opt(2)} = \frac{2d^2}{(1-\rho)d + \sqrt{(1-\rho)^2 d^2 - (1-\rho)(\rho_0 - \rho)^2 dC_0^2}}$$

$$G_{(3)} = \frac{(1-\rho)d - (\rho_0 - \rho)^2 q^2 C_0^2}{(1-\rho)d - (\rho_0 - \rho)^2 qC_0^2}$$

$$G_{opt(3)} = \frac{2(1-\rho)d}{(1-\rho)d + \sqrt{(1-\rho)^2 d^2 - (1-\rho)(\rho_0 - \rho)^2 dC_0^2}}$$

Also, the expressions (3), (4) and (5) becomes

$$q_{opt}^{(0)} = \frac{d^2 - \sqrt{d^4 - 2d^2 t^2}}{t^2}, \qquad q_{opt}^{(1)} = q_{opt}^{(2)} = \frac{(1-\rho)d - \sqrt{(1-\rho)^2 d^2 - (1-\rho)(\rho_0 - \rho)^2 dC_0^2}}{(\rho_0 - \rho)^2 C_0^2}$$

The gain in precision of $\hat{R}_2^{**}$, $\hat{R}_{21}^{**}$ and $\hat{R}_{22}^{**}$ with respect to $\hat{R}^{*}$ for different values of C_0, $C_{0(2)}$, ρ, ρ_0, $\rho_{(2)}$ and $\rho_{0(2)}$, 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) For the case $C_0 < C_{0(2)}$, the gain in precision of $\hat{R}_2^{**}$ and $\hat{R}_{22}^{**}$ with respect to $\hat{R}^{*}$ decreases as the values of $C_{0(2)}$ increases, whereas the gain in precision of $\hat{R}_{21}^{**}$ with respect to $\hat{R}^{*}$ increases as the values of $C_{0(2)}$ increase; see Figure 1 (a).

(ii) For the case $C_0 > C_{0(2)}$, the gain in precision of $\hat{R}_2^{**}$ and $\hat{R}_{22}^{**}$ with respect to $\hat{R}^{*}$ increases as the values of C_0 increases, whereas the gain in precision of $\hat{R}_{21}^{**}$ with respect to $\hat{R}^{*}$ decreases as the values of C_0 increase; see Figure 1 (b).

(iii) For the case $C_0 = C_{0(2)}$, the gain in precision of all the estimators with respect to $\hat{R}^{*}$ remain constant as the values of C_0 and $C_{0(2)}$ increase; see Figure 1 (c).

(iv) For the case $\rho > \rho_0$, the gain in precision of all the estimators with respect to $\hat{R}^{*}$ decreases as the values of ρ_0 increase; see Figure 1 (d).

(v) For the case $\rho < \rho_0$, the gain in precision of all the estimators with respect to $\hat{R}^{*}$ decreases as the values of ρ increase; see Figure 1 (e).

(vi) For the case $\rho = \rho_0$, the gain in precision of $\hat{R}_{22}^{**}$ with respect to $\hat{R}^{*}$ remains constant as the values of ρ and ρ_0 increase, whereas the gain in precision of $\hat{R}_{2}^{**}$ and $\hat{R}_{21}^{**}$ with respect to $\hat{R}^{*}$ increases as the values of ρ and ρ_0 increase; see Figure 1(f).

(vii) For the case $\rho_{(2)} > \rho_{0(2)}$, the gain in precision of $\hat{R}_{2}^{**}$ with respect to $\hat{R}^{*}$ decreases as the values of $\rho_{0(2)}$ increase, whereas the gain in precision of $\hat{R}_{21}^{**}$ and $\hat{R}_{22}^{**}$ with respect to $\hat{R}^{*}$ remains constant as the values of $\rho_{0(2)}$ increase; see Figure 1 (g).

Table 1: Gain in precision, $G_{(1)}$, $G_{(2)}$ and $G_{(3)}$ of the estimate proposed over the direct estimate for different values of C_0, $C_{0(2)}$, ρ, ρ_0, $\rho_{(2)}$ and $\rho_{0(2)}$.1

ρ	ρ_0	$\rho_{(2)}$	$\rho_{0(2)}$	$(k-1)$	W_2	C_0	$C_{0(2)}$	q	$G_{(1)}$	$G_{(2)}$	$G_{(3)}$
				$C_0 < C_{0(2)}$							
0.7	0.2	0.5	0.3	1.5	0.8	0.4	1	0.7	1.06	14.18	1.05
0.7	0.2	0.5	0.3	1.5	0.8	0.4	1.5	0.7	1.05	29.75	1.02
0.7	0.2	0.5	0.3	1.5	0.8	0.4	2	0.7	1.04	51.61	1.01
				$C_0 > C_{0(2)}$							
0.6	0.4	0.4	0.2	0.5	0.6	0.3	0.2	0.5	1.062	1.27	1.058
0.6	0.4	0.4	0.2	0.5	0.6	0.5	0.2	0.5	1.068	1.14	1.066
0.6	0.4	0.4	0.2	0.5	0.6	0.7	0.2	0.5	1.069	1.11	1.068
				$C_0 = C_{0(2)}$							
0.6	0.3	0.5	0.3	1	0.7	0.1	0.1	0.3	1.08	2.00	1.07
0.6	0.3	0.5	0.3	1	0.7	0.3	0.3	0.3	1.08	2.00	1.07
0.6	0.3	0.5	0.3	1	0.7	0.8	0.8	0.3	1.08	2.00	1.07
				$\rho > \rho_0$							
0.7	0.3	0.6	0.7	0.5	0.8	0.5	0.6	0.4	1.06	2.63	1.34
0.7	0.4	0.6	0.7	0.5	0.8	0.5	0.6	0.4	1.02	2.26	1.15
0.7	0.5	0.6	0.7	0.5	0.8	0.5	0.6	0.4	1.00	2.08	1.06
				$\rho < \rho_0$							
0.2	0.7	0.6	0.7	0.5	0.4	0.7	0.3	0.3	1.09	1.11	1.09
0.4	0.7	0.6	0.7	0.5	0.4	0.7	0.3	0.3	1.05	1.08	1.06
0.6	0.7	0.6	0.7	0.5	0.4	0.7	0.3	0.3	1.01	1.05	1.01
				$\rho = \rho_0$							
0.2	0.2	0.5	0.1	1.5	0.5	0.4	0.6	0.6	1.04	2.05	1.00
0.5	0.5	0.5	0.1	1.5	0.5	0.4	0.6	0.6	1.07	2.69	1.00
0.8	0.8	0.5	0.1	1.5	0.5	0.4	0.6	0.6	1.13	5.22	1.00
				$\rho_{(2)} > \rho_{0(2)}$							
0.7	0.6	0.5	0.2	1.5	0.8	0.5	0.6	0.4	1.08	3.91	1.01
0.7	0.6	0.5	0.3	1.5	0.8	0.5	0.6	0.4	1.04	3.91	1.01
0.7	0.6	0.5	0.4	1.5	0.8	0.5	0.6	0.4	1.01	3.91	1.01
				$\rho_{(2)} < \rho_{0(2)}$							
0.5	0.6	0.1	0.4	2.5	0.6	0.5	0.6	0.8	1.02	4.89	1.00
0.5	0.6	0.2	0.4	2.5	0.6	0.5	0.6	0.8	1.01	4.46	1.00
0.5	0.6	0.3	0.4	2.5	0.6	0.5	0.6	0.8	1.00	4.03	1.00
				$\rho_{(2)} = \rho_{0(2)}$							
0.7	0.2	0.5	0.5	1.5	0.6	0.2	0.6	0.5	1.00	15.27	1.05
0.7	0.2	0.7	0.7	1.5	0.6	0.2	0.6	0.5	1.01	9.92	1.09
0.7	0.2	0.8	0.8	1.5	0.6	0.2	0.6	0.5	1.02	7.29	1.14

Table 2: Gain in precision, $G_{(1)}$, $G_{(2)}$ and $G_{(3)}$ of the estimate proposed over the direct estimate for different values of W_2, $(k-1)$ and q. 2

ρ	ρ_0	$\rho_{(2)}$	$\rho_{0(2)}$	$(k-1)$	W_2	C_0	$C_{0(2)}$	q	$G_{(1)}$	$G_{(2)}$	$G_{(3)}$
				W_2							
0.7	0.2	0.5	0.5	1.5	0.2	0.2	0.6	0.8	1.015	6.25	1.14
0.7	0.2	0.5	0.5	1.5	0.4	0.2	0.6	0.8	1.004	10.57	1.06
0.7	0.2	0.5	0.5	1.5	0.8	0.2	0.6	0.8	1.001	19.50	1.03
				$(k-1)$							
0.8	0.3	0.4	0.6	0.5	0.8	0.2	0.6	0.4	1.00	13.70	1.16
0.8	0.3	0.4	0.6	1.0	0.8	0.2	0.6	0.4	1.01	24.29	1.07
0.8	0.3	0.4	0.6	1.5	0.8	0.2	0.6	0.4	1.02	35.02	1.05
				q							
0.8	0.3	0.4	0.6	1.5	0.8	0.2	0.6	0.2	1.010	34.44	1.03
0.8	0.3	0.4	0.6	1.5	0.8	0.2	0.6	0.4	1.015	35.02	1.05
0.8	0.3	0.4	0.6	1.5	0.8	0.2	0.6	0.6	1.015	35.09	1.05

(viii) For the case $\rho_{(2)} < \rho_{0(2)}$, the gain in precision of $\bar{R}_2^{**}$ and $\bar{R}_{21}^{**}$ with respect to $\bar{R}^{*}$ decreases as the values of $\rho_{(2)}$ increase, whereas the gain in precision of $\bar{R}_{22}^{**}$ with respect to $\bar{R}^{*}$ remains constant as the values of $\rho_{(2)}$ increase; see Figure 1 (h).

(ix) For the case $\rho_{(2)} = \rho_{0(2)}$, the gain in precision of $\bar{R}_2^{**}$ and $\bar{R}_{22}^{**}$ with respect to $\bar{R}^{*}$ increases as the values of $\rho_{(2)}$ and $\rho_{0(2)}$ increase, whereas the gain in precision of $\bar{R}_{21}^{**}$ with respect to $\bar{R}^{*}$ decreases as the values of $\rho_{(2)}$ and $\rho_{0(2)}$ increase; see Figure 1 (i).

(x) The gain in precision of $\bar{R}_2^{**}$ and $\bar{R}_{22}^{**}$ with respect to $\bar{R}^{*}$ decreases as the values of W_2 increases, whereas the gain in precision of $\bar{R}_{21}^{**}$ with respect to $\bar{R}^{*}$ increases as the values of W_2 increase; see Figure 1(j).

(xi) The gain in precision of $\bar{R}_{22}^{**}$ with respect to $\bar{R}^{*}$ decreases as the values of $k-1$ increase, whereas the gain in precision of $\bar{R}_2^{**}$ and $\bar{R}_{21}^{**}$ with respect to $\bar{R}^{*}$ increases as the values of $k-1$ increase; see Figure 1 (k).

(xii) The gain in precision of $\bar{R}_2^{**}$, $\bar{R}_{21}^{**}$, $\bar{R}_{22}^{**}$ with respect to $\bar{R}^{*}$ increases as the values of q increase; see Figure 1 (l).

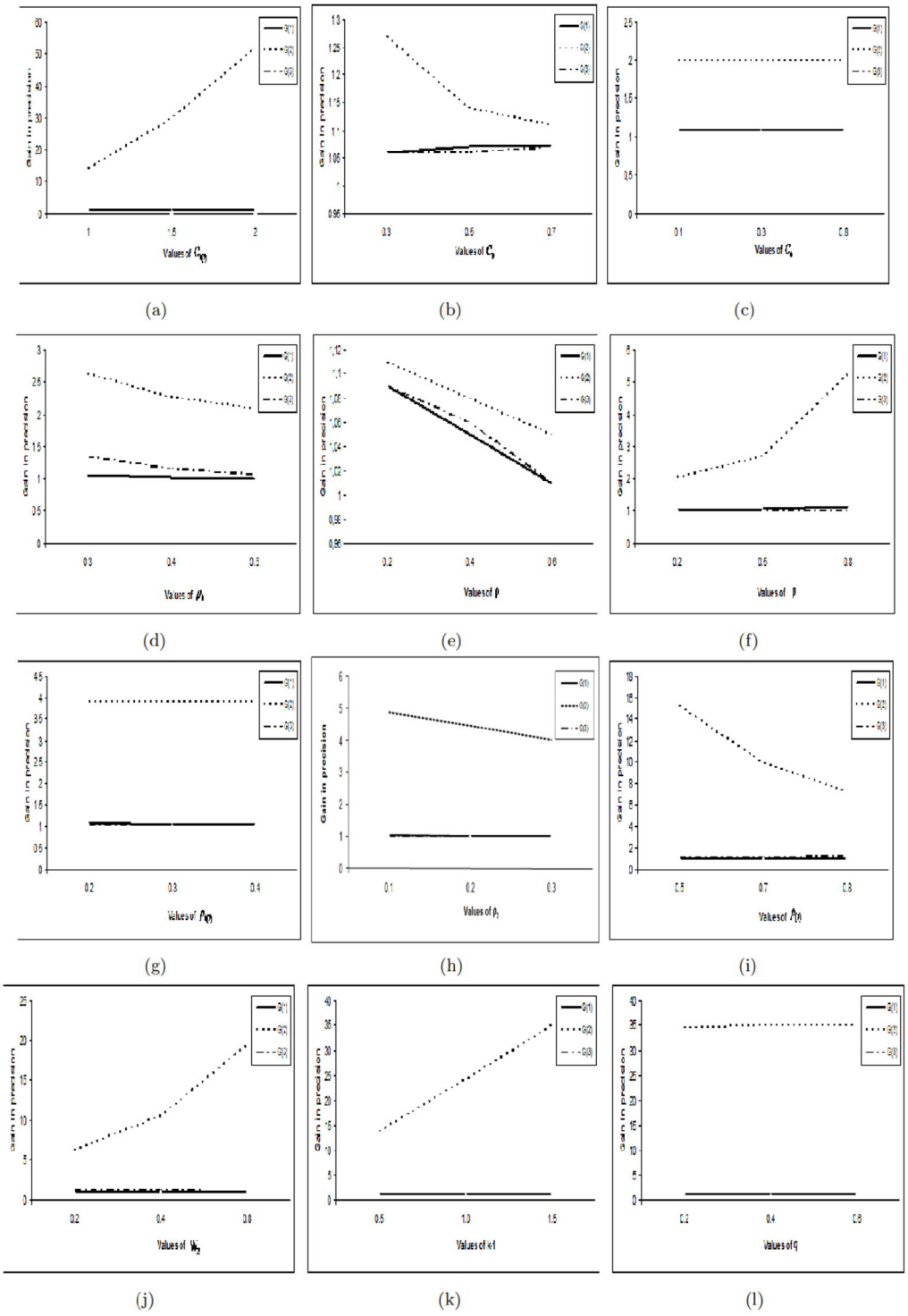

Figure 1: Gain in precision, $G_{(1)}$, $G_{(2)}$ and $G_{(3)}$ of the estimate proposed over the direct estimate for (a)-(b) different values of $C_{0(2)}$ and C_0, (c) the case $C_0 = C_{0(2)}$, (d)-(e) different values of ρ_0 and ρ, (f) the case $\rho = \rho_0$, (g)-(h) different values of $\rho_{0(2)}$ and ρ_2, (i) the case $\rho_2 = \rho_{0(2)}$, (j)-(k) different values of W_2 and $k-1$, and (l) different values of q.

The optimum gain of $\hat{R}_2^{**}$, $\hat{R}_{21}^{**}$ and $\hat{R}_{22}^{**}$ with respect to $\hat{R}^{*}$ and the optimum fraction to be unmatched for different values of C_0, $C_{0(2)}$, ρ, ρ_0, $\rho_{(2)}$ and $\rho_{0(2)}$ are presented in tables 3-4 and in Figure 2. From these tables, we obtain the following conclusions:

(i) For the case $C_0 < C_{0(2)}$, the optimum gain of $\hat{R}_2^{**}$ and $\hat{R}_{22}^{**}$ with respect to $\hat{R}^{*}$ decreases as the values of $C_{0(2)}$ increases, whereas the optimum gain of $\hat{R}_{21}^{**}$ with respect to $\hat{R}^{*}$ increases as the values of $C_{0(2)}$ increase.

(ii) For the case $C_0 > C_{0(2)}$, the optimum gain of $\hat{R}_2^{**}$, $\hat{R}_{21}^{**}$ and $\hat{R}_{22}^{**}$ with respect to $\hat{R}^{*}$ increases as the values of C_0 increases.

(iii) For the case $C_0 = C_{0(2)}$, the optimum gain of $\hat{R}_2^{**}$ and $\hat{R}_{22}^{**}$ with respect to $\hat{R}^{*}$ remain constant as the values of C_0 and $C_{0(2)}$ increase, whereas the optimum gain of $\hat{R}_{21}^{**}$ with respect to $\hat{R}^{*}$ increases as the values of C_0 and $C_{0(2)}$ increase.

(iv) For the case $\rho > \rho_0$, the optimum gain of all the estimators with respect to $\hat{R}^{*}$ decreases as the values of ρ_0 increase.

(v) For the case $\rho < \rho_0$, the optimum gain of all the estimators with respect to $\hat{R}^{*}$ decreases as the values of ρ increase.

(vi) For the case $\rho = \rho_0$, the optimum gain of $\hat{R}_2^{**}$, $\hat{R}_{21}^{**}$ and $\hat{R}_{22}^{**}$ with respect to $\hat{R}^{*}$ increases as the values of ρ and ρ_0 increase.

(vii) For the case $\rho_{(2)} > \rho_{0(2)}$, the optimum gain of $\hat{R}_2^{**}$ with respect to $\hat{R}^{*}$ decreases as the values of $\rho_{0(2)}$ increase, whereas the optimum gain of $\hat{R}_{21}^{**}$ and $\hat{R}_{22}^{**}$ with respect to $\hat{R}^{*}$ remains constant as the values of $\rho_{0(2)}$ increase.

(viii) For the case $\rho_{(2)} < \rho_{0(2)}$, the optimum gain of $\hat{R}_2^{**}$ and $\hat{R}_{21}^{**}$ with respect to $\hat{R}^{*}$ decreases as the values of $\rho_{(2)}$ increase, whereas the optimum gain of $\hat{R}_{22}^{**}$ with respect to $\hat{R}^{*}$ remains constant as the values of $\rho_{(2)}$ increase.

(ix) For the case $\rho_{(2)} = \rho_{0(2)}$, the optimum gain of $\hat{R}_2^{**}$ and $\hat{R}_{22}^{**}$ with respect to $\hat{R}^{*}$ increases as the values of $\rho_{(2)}$ and $\rho_{0(2)}$ increase, whereas the optimum gain of $\hat{R}_{21}^{**}$ with respect to $\hat{R}^{*}$ decreases as the values of $\rho_{(2)}$ and $\rho_{0(2)}$ increase.

(x) The optimum gain of $\hat{R}_2^{**}$ and $\hat{R}_{22}^{**}$ with respect to $\hat{R}^{*}$ decreases as the values of W_2 increases, whereas the optimum gain of $\hat{R}_{21}^{**}$ with respect to $\hat{R}^{*}$ increases as the values of W_2 increase.

(xi) The optimum gain of $\hat{R}_2^{**}$ and $\hat{R}_{21}^{**}$ with respect to $\hat{R}^{*}$ increases as the values of $k-1$ increase, whereas the optimum gain of $\hat{R}_{22}^{**}$ with respect to $\hat{R}^{*}$ decreases as the values of $k-1$ increase.

(xii) The optimum gain of $\hat{R}_2^{**}$, $\hat{R}_{21}^{**}$ and $\hat{R}_{22}^{**}$ with respect to $\hat{R}^{*}$ remains constant as the values of q increase.

Table 3: Optimum gain, $G_{opt(1)}$, $G_{opt(2)}$, $G_{opt(3)}$ and the optimum fraction to be unmatched, $q_{opt}^{(0)}$, $q_{opt}^{(1)}$ and $q_{opt}^{(2)}$ of the estimate proposed over the direct estimate for different values of C_0, $C_{0(2)}$, ρ, ρ_0, $\rho_{(2)}$ and $\rho_{0(2)}$.

ρ	ρ_0	$\rho_{(2)}$	$\rho_{0(2)}$	$(k-1)$	W_2	C_0	$C_{0(2)}$	q	$q_{opt}^{(0)}$	$q_{opt}^{(1,2)}$	$G_{opt(1)}$	$G_{opt(2)}$	$G_{opt(3)}$
				$C_0 < C_{0(2)}$									
0.7	0.2	0.5	0.3	1.5	0.8	0.4	1.0	0.7	1.2	0.5	1.07	2.28	1.06
0.7	0.2	0.5	0.3	1.5	0.8	0.4	1.5	0.7	1.1	0.5	1.05	4.78	1.03
0.7	0.2	0.5	0.3	1.5	0.8	0.4	2.0	0.7	1.1	0.5	1.04	8.27	1.01
				$C_0 > C_{0(2)}$									
0.6	0.4	0.4	0.2	0.5	0.6	0.3	0.2	0.5	1.1	0.5	1.06	0.11	1.06
0.6	0.4	0.4	0.2	0.5	0.6	0.5	0.2	0.5	1.2	0.5	1.07	0.29	1.07
0.6	0.4	0.4	0.2	0.5	0.6	0.7	0.2	0.5	1.2	0.5	1.07	0.54	1.07
				$C_0 = C_{0(2)}$									
0.6	0.3	0.5	0.3	1.0	0.7	0.1	0.1	0.3	1.3	0.5	1.11	0.02	1.09
0.6	0.3	0.5	0.3	1.0	0.7	0.3	0.3	0.3	1.3	0.5	1.11	0.18	1.09
0.6	0.3	0.5	0.3	1.0	0.7	0.8	0.8	0.3	1.3	0.5	1.11	1.31	1.09
				$\rho > \rho_0$									
0.7	0.3	0.6	0.7	0.5	0.8	0.5	0.6	0.4	1.2	0.8	1.07	0.75	1.53
0.7	0.4	0.6	0.7	0.5	0.8	0.5	0.6	0.4	1.1	0.6	1.02	0.58	1.18
0.7	0.5	0.6	0.7	0.5	0.8	0.5	0.6	0.4	1.0	0.5	1.00	0.52	1.06
				$\rho < \rho_0$									
0.2	0.7	0.6	0.7	0.5	0.4	0.7	0.3	0.3	1.3	0.6	1.12	0.56	1.12
0.4	0.7	0.6	0.7	0.5	0.4	0.7	0.3	0.3	1.2	0.5	1.07	0.54	1.07
0.6	0.7	0.6	0.7	0.5	0.4	0.7	0.3	0.3	1.0	0.5	1.02	0.51	1.02
				$\rho = \rho_0$									
0.2	0.2	0.5	0.1	1.5	0.5	0.4	0.6	0.6	1.1	-	1.05	0.33	1.00
0.5	0.5	0.5	0.1	1.5	0.5	0.4	0.6	0.6	1.2	-	1.07	0.43	1.00
0.8	0.8	0.5	0.1	1.5	0.5	0.4	0.6	0.6	1.4	-	1.13	0.83	1.00
				$\rho_{(2)} > \rho_{0(2)}$									
0.7	0.6	0.5	0.2	1.5	0.8	0.5	0.6	0.4	1.2	0.5	1.08	0.98	1.01
0.7	0.6	0.5	0.3	1.5	0.8	0.5	0.6	0.4	1.1	0.5	1.04	0.98	1.01
0.7	0.6	0.5	0.4	1.5	0.8	0.5	0.6	0.4	1.0	0.5	1.01	0.98	1.01
				$\rho_{(2)} < \rho_{0(2)}$									
0.5	0.6	0.1	0.4	2.5	0.6	0.5	0.6	0.8	1.1	0.5	1.02	1.22	1.00
0.5	0.6	0.2	0.4	2.5	0.6	0.5	0.6	0.8	1.0	0.5	1.01	1.12	1.00
0.5	0.6	0.3	0.4	2.5	0.6	0.5	0.6	0.8	1.2	0.5	1.01	1.01	1.00
				$\rho_{(2)} = \rho_{0(2)}$									
0.7	0.2	0.5	0.5	1.5	0.6	0.2	0.6	0.5	1.0	0.5	1.00	0.61	1.05
0.7	0.2	0.7	0.7	1.5	0.6	0.2	0.6	0.5	1.0	0.5	1.01	0.40	1.09
0.7	0.2	0.8	0.8	1.5	0.6	0.2	0.6	0.5	1.0	0.6	1.02	0.29	1.14

Table 4: Optimum gain in precision, $G_{opt(1)}$, $G_{opt(2)}$, $G_{opt(3)}$ and the optimum fraction to be unmatched, $q_{opt}^{(0)}$, $q_{opt}^{(1)}$ and $q_{opt}^{(2)}$ of the estimate proposed over the direct estimate for different values of W_2, $(k-1)$ and q.

ρ	ρ_0	$\rho_{(2)}$	$\rho_{0(2)}$	$(k-1)$	W_2	C_0	$C_{0(2)}$	q	$q_{opt}^{(0)}$	$q_{opt}^{(1,2)}$	$G_{opt(1)}$	$G_{opt(2)}$	$G_{opt(3)}$
				W_2									
0.7	0.2	0.5	0.5	1.5	0.2	0.2	0.6	0.8	1.1	0.6	1.02	0.26	1.17
0.7	0.2	0.5	0.5	1.5	0.4	0.2	0.6	0.8	1.0	0.5	1.01	0.43	1.08
0.7	0.2	0.5	0.5	1.5	0.8	0.2	0.6	0.8	1.0	0.5	1.00	0.79	1.04
				$(k-1)$									
0.8	0.3	0.4	0.6	0.5	0.8	0.2	0.6	0.4	1.0	0.6	1.00	0.56	1.19
0.8	0.3	0.4	0.6	1.0	0.8	0.2	0.6	0.4	1.0	0.5	1.01	0.98	1.08
0.8	0.3	0.4	0.6	1.5	0.8	0.2	0.6	0.4	1.0	0.5	1.02	1.41	1.05
				q									
0.8	0.3	0.4	0.6	1.5	0.8	0.2	0.6	0.2	1.0	0.5	1.02	1.41	1.05
0.8	0.3	0.4	0.6	1.5	0.8	0.2	0.6	0.4	1.0	0.5	1.02	1.41	1.05
0.8	0.3	0.4	0.6	1.5	0.8	0.2	0.6	0.6	1.0	0.5	1.02	1.41	1.05

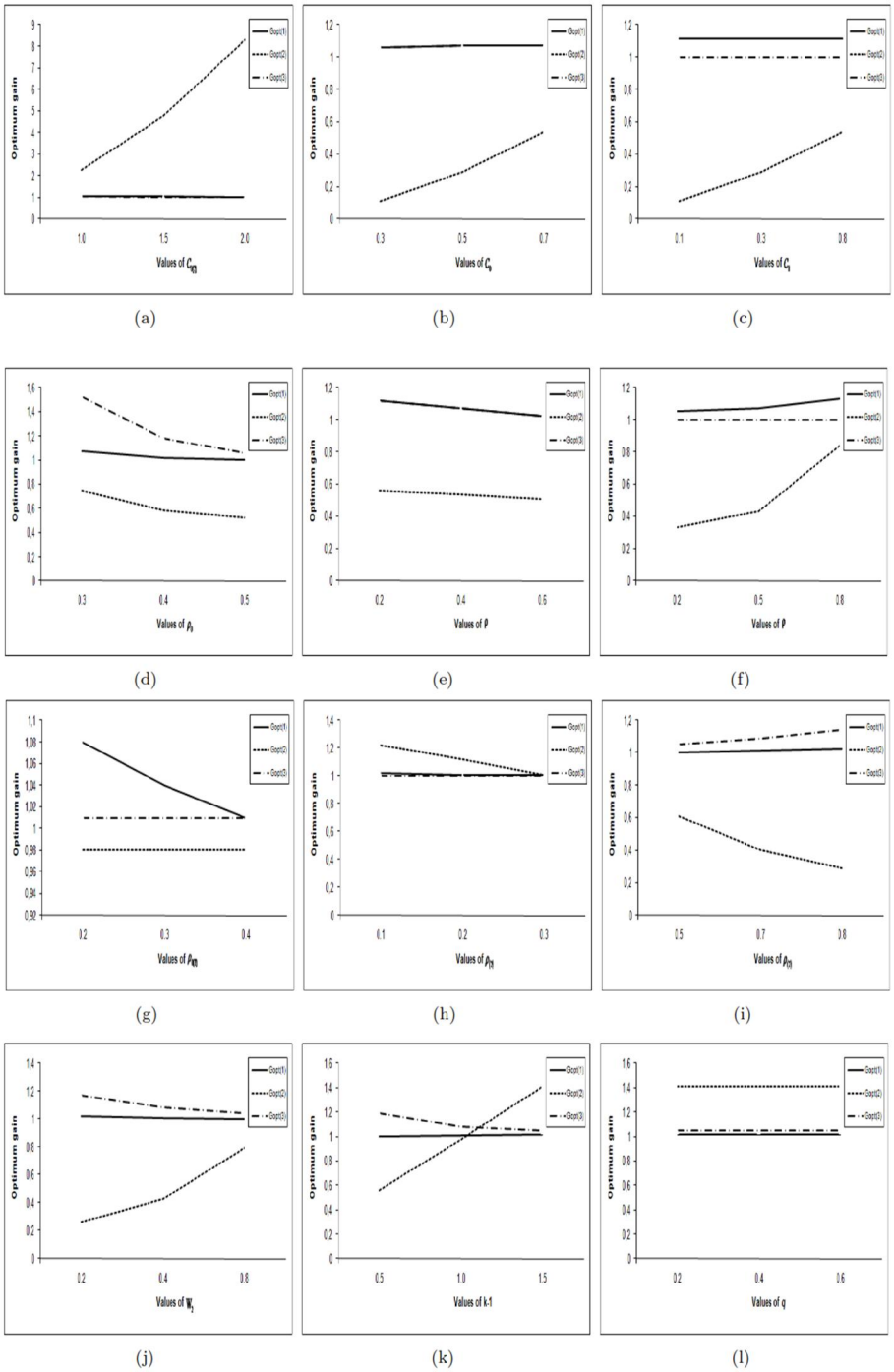

Figure 2: Optimum gain, $G_{opt(1)}$, $G_{opt(2)}$, $G_{opt(3)}$ of the estimate proposed over the direct estimate for (a)-(b) different values of $C_{0(2)}$ and C_0, (c) the case $C_0 = C_{0(2)}$, (d)-(e) different values of ρ_0 and ρ, (f) the case $\rho = \rho_0$, (g)-(h) different values of $\rho_{0(2)}$ and ρ_2, (i) the case $\rho_2 = \rho_{0(2)}$, (j)-(k) different values of W_2 and $k-1$, and (l) different values of q.

4. 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 C_0, $C_{0(2)}$, ρ, ρ_0, $\rho_{(2)}$, $\rho_{0(2)}$, W_2, $(k-1)$ 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)

$$C_{00}^{*} = 2nc_2. \tag{12}$$

Substituting the values of n and c_2 in Eq. (12), 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

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

The expected cost is given by

$$\mathrm{E}(C_0^{*}) = 2n_0^{*}\left[c_0 + c_1 W_1 + \frac{c_2 W_2}{k-1}\right],$$

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

$$n_0^{*} = \frac{n(d^2 - qt^2)}{(d^2 - q^2 t^2)}$$

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

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

and the expected cost is given by

$$\mathrm{E}(C_1^{*}) = n_1^{*}\left[2c_0 + c_1(W_1 + 1) + \frac{c_2 W_2}{k-1}\right].$$

where

$$n_1^{*} = \frac{n(1-\rho)\left((1-\rho)d - (\rho_0 - \rho)^2 qC_0^2\right)C_0^2}{d\left((1-\rho)d - (\rho_0 - \rho)^2 q^2 C_0^2\right)}$$

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

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

which expected cost is expressed as

$$\mathrm{E}(C_2^*) = n_2^* \left[2c_0 + c_1(W_1 + 1) + \frac{c_2 W_2}{k-1} \right].$$

where

$$n_2^* = \frac{n((1-\rho)d - (\rho_0 - \rho)^2 q C_0^2)}{(1-\rho)d - (\rho_0 - \rho)^2 q^2 C_0^2}$$

By equating the variances $\hat{R}_2^{**}$, $\hat{R}_{21}^{**}$, and $\hat{R}_{22}^{**}$, respectively, to $V(\hat{R}^*)$ 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 $\hat{R}_2^{**}$, $\hat{R}_{21}^{**}$, and $\hat{R}_{22}^{**}$. The sample sizes associated with the three estimators which provide equal precision to the estimator $\hat{R}^*$ are denoted by n_0^*, n_1^* and n_2^*. The results of this exercise are presented in tables 5-6 and in Figures 3-4. From these tables, we obtain the following conclusions:

(i) For the case $C_0 < C_{0(2)}$, the saving in cost for $\hat{R}_2^{**}$ and $\hat{R}_{22}^{**}$ increases as the values of $C_{0(2)}$ increase, whereas for $\hat{R}_{21}^{**}$ the saving in cost decreases as the values of $C_{0(2)}$ increase; see Figure 3(a).

The sample sizes for $\hat{R}_2^{**}$ and $\hat{R}_{22}^{**}$, which have the same precision than $\hat{R}^*$, increases as the values of $C_{0(2)}$ increase, whereas the sample size for $\hat{R}_{21}^{**}$, which have the same precision than $\hat{R}^*$ decreases as the values of $C_{0(2)}$ increase; see Figure 3(b).

(ii) For the case $C_0 > C_{0(2)}$, the saving in cost for $\hat{R}_2^{**}$ and $\hat{R}_{22}^{**}$ decreases as the values of C_0 increase, whereas for $\hat{R}_{21}^{**}$ the saving in cost increases as the values of C_0 increase; see Figure 3(c).

The sample sizes for $\hat{R}_2^{**}$ and $\hat{R}_{22}^{**}$ remains constant as the values of C_0 increase, whereas for $\hat{R}_{21}^{**}$ the saving in cost increases as the values of C_0 increase; see Figure 3 (d).

(iii) For the case $C_0 = C_{0(2)}$ the saving in cost for all the estimators remains constant as the values of C_0 and $C_{0(2)}$ increase; see Figure 3(e).

The sample sizes for all the estimators, which have the same precision than $\hat{R}^*$, remain constant as the values of C_0 and $C_{0(2)}$ increase; see Figure 3(f).

(iv) For the case $\rho > \rho_0$, the saving in cost for all the estimators increases as the values of ρ_0 increase; see Figure 3(g).

The sample sizes for the three estimators, which have the same precision than $\hat{R}^*$, increases as the values of ρ_0 increase; see Figure 3(h).

(v) For the case $\rho < \rho_0$, the saving in cost for all the estimators increases as the values of ρ increase; see Figure 3(i).

The sample sizes for the three estimators, which have the same precision than $\hat{R}^*$, increases as the values of ρ increase; see Figure 3(j).

(vi) For the case $\rho = \rho_0$, the saving in cost for $\hat{R}^{**}_{22}$ remains constant as the values of ρ and ρ_0 increase, whereas for $\hat{R}^{**}_{2}$ and $\hat{R}^{**}_{21}$ the saving in cost decreases as the values of ρ and ρ_0 increase; see Figure 3 (k).

The sample sizes for $\hat{R}^{**}_{2}$ and $\hat{R}^{**}_{21}$, which give equal precision to $\hat{R}^*$ decrease as the values of ρ and ρ_0 increase, whereas the sample size for $\hat{R}^{**}_{22}$, which has the same precision than $\hat{R}^*$, remains constant as the values of ρ and ρ_0 increase; see Figure 3 (l).

(vii) For the case $\rho_{(2)} > \rho_{0(2)}$, the saving in cost for $\hat{R}^{**}_{2}$ increases as the values of $\rho_{0(2)}$ increase, whereas for $\hat{R}^{**}_{21}$ and $\hat{R}^{**}_{22}$ the saving in cost remains constant as the values of $\rho_{0(2)}$ increase; see Figure 4(a).

The sample size for $\hat{R}^{**}_{2}$, which have the same precision than $\hat{R}^*$, increases as the values of $\rho_{0(2)}$ increase, whereas for $\hat{R}^{**}_{21}$ and $\hat{R}^{**}_{22}$ which give equal precision to $\hat{R}^*$ remains constant as the values of $\rho_{0(2)}$ increase; see Figure 4(b).

(viii) For the case $\rho_{(2)} < \rho_{0(2)}$, the saving in cost for $\hat{R}^{**}_{2}$ and $\hat{R}^{**}_{21}$ increases as the values of $\rho_{(2)}$ increase whereas for $\hat{R}^{**}_{22}$ the saving in cost decreases as the values of $\rho_{(2)}$ increase; see Figure 4(c).

The sample size for $\hat{R}_{22}^{**}$, which have the same precision than $\hat{R}^{*}$ remains constant as the values of $\rho_{(2)}$ increase, whereas the sample size for $\hat{R}_{2}^{**}$ and $\hat{R}_{21}^{**}$, which have the same precision than $\hat{R}^{*}$, increases as the values of $\rho_{(2)}$ increase; see Figure 4(d).

(ix) For the case $\rho_{(2)} = \rho_{0(2)}$, the saving in cost for $\hat{R}_{2}^{**}$ and $\hat{R}_{22}^{**}$ decreases as the values of $\rho_{(2)}$ and $\rho_{0(2)}$ increase, whereas for $\hat{R}_{21}^{**}$ the saving in cost increases as the values of $\rho_{(2)}$ and $\rho_{0(2)}$ increase; see Figure 4(e).

The sample sizes for $\hat{R}_{2}^{**}$ and $\hat{R}_{22}^{**}$, which have the same precision than $\hat{R}^{*}$, decreases as the values of $\rho_{(2)}$ and $\rho_{0(2)}$ increase, whereas the sample size for $\hat{R}_{21}^{**}$, which have the same precision than $\hat{R}^{*}$ increases as the values of $\rho_{(2)}$ and $\rho_{0(2)}$ increase; see Figure 4(f).

(x) The saving in cost for $\hat{R}_{2}^{**}$ and $\hat{R}_{22}^{**}$ increases as the values of W_2 increase, whereas for $\hat{R}_{21}^{**}$ the saving in cost decreases as the values of W_2 increase; see Figure 4(g).

The sample sizes for $\hat{R}_{2}^{**}$ and $\hat{R}_{22}^{**}$, which have the same precision than $\hat{R}^{*}$, increases as the values of W_2 increase, whereas the sample size for $\hat{R}_{21}^{**}$, which have the same precision than $\hat{R}^{*}$ decreases as the values of W_2 increase; see Figure 4(h).

(xi) The saving in cost for all the estimators decreases as the values of $k-1$ increase; see Figure 4(i).

The sample sizes for $\hat{R}_{2}^{**}$ and $\hat{R}_{21}^{**}$, which give equal precision to $\hat{R}^{*}$ decrease as the values of $k-1$ increase, whereas the sample size for $\hat{R}_{22}^{**}$, which has the same precision than $\hat{R}^{*}$, increases as the values of $k-1$ increase; see Figure 4(j).

(xii) The saving in cost for all the estimators decreases as the values of q increase; see Figure 4(k).

The samples sizes for $\hat{R}_{2}^{**}$ and $\hat{R}_{21}^{**}$, which give equal precision to $\hat{R}^{*}$ remain constant as the values of q increase, whereas the sample size for $\hat{R}_{22}^{**}$, which has the same precision than $\hat{R}^{*}$, decreases as the values of q increase; see Figure 4 (l).

Table 5: Sample sizes and corresponding expected cost of survey, which have the same precision than $\hat{R}_2^{**}$, $\hat{R}_{21}^{**}$ and $\hat{R}_{22}^{**}$, with respect to $\hat{R}^{*}$ for different values of C_0, $C_{0(2)}$, ρ, ρ_0, $\rho_{(2)}$ and $\rho_{0(2)}$

ρ	ρ_0	$\rho_{(2)}$	$\rho_{0(2)}$	$(k-1)$	W_2	C_0	$C_{0(2)}$	q	n_0^*	n_1^*	n_2^*	E(C_0^*)	E(C_1^*)	E(C_2^*)
				$C_0 < C_{0(2)}$										
0.7	0.2	0.5	0.3	1.5	0.8	0.4	1.0	0.7	47	3	47	2429.94	108.59	1465.99
0.7	0.2	0.5	0.3	1.5	0.8	0.4	1.5	0.7	48	2	49	2462.07	51.76	1507.64
0.7	0.2	0.5	0.3	1.5	0.8	0.4	2.0	0.7	48	1	49	2472.72	29.84	1521.90
				$C_0 > C_{0(2)}$										
0.6	0.4	0.4	0.2	0.5	0.6	0.3	0.2	0.5	47	39	47	5325.83	2425.64	2910.77
0.6	0.4	0.4	0.2	0.5	0.6	0.5	0.2	0.5	47	44	47	5300.43	2695.25	2889.31
0.6	0.4	0.4	0.2	0.5	0.6	0.7	0.2	0.5	47	45	47	5292.10	2780.28	2882.41
				$C_0 = C_{0(2)}$										
0.6	0.3	0.5	0.3	1	0.7	0.1	0.1	0.3	46	25	47	3118.63	965.18	1809.71
0.6	0.3	0.5	0.3	1	0.7	0.3	0.3	0.3	46	25	47	3118.63	965.18	1809.71
0.6	0.3	0.5	0.3	1	0.7	0.8	0.8	0.3	46	25	47	3118.63	965.18	1809.71
				$\rho > \rho_0$										
0.7	0.3	0.6	0.7	0.5	0.8	0.5	0.6	0.4	47	19	37	6948.27	1498.32	2936.71
0.7	0.4	0.6	0.7	0.5	0.8	0.5	0.6	0.4	49	22	43	7202.24	1742.18	3414.67
0.7	0.5	0.6	0.7	0.5	0.8	0.5	0.6	0.4	50	24	47	7343.04	1896.69	3717.50
				$\rho < \rho_0$										
0.2	0.7	0.6	0.7	0.5	0.4	0.7	0.3	0.3	46	45	46	3612.49	1998.08	2034.78
0.4	0.7	0.6	0.7	0.5	0.4	0.7	0.3	0.3	47	46	47	3733.53	2053.39	2103.68
0.6	0.7	0.6	0.7	0.5	0.4	0.7	0.3	0.3	49	48	49	3887.99	2114.08	2191.74
				$\rho = \rho_0$										
0.2	0.2	0.5	0.1	1.5	0.5	0.4	0.6	0.6	48	24	50	1722.44	559.69	1150
0.5	0.5	0.5	0.1	1.5	0.5	0.4	0.6	0.6	47	19	50	1680.10	427.91	1150
0.8	0.8	0.5	0.1	1.5	0.5	0.4	0.6	0.6	44	9	50	1587.30	220.36	1150
				$\rho_{(2)} > \rho_{0(2)}$										
0.7	0.6	0.5	0.2	1.5	0.8	0.5	0.6	0.4	46	13	50	2396.96	394.17	1529.37
0.7	0.6	0.5	0.3	1.5	0.8	0.5	0.6	0.4	48	13	50	2487.08	394.17	1529.37
0.7	0.6	0.5	0.4	1.5	0.8	0.5	0.6	0.4	49	13	50	2545.69	394.17	1529.37
				$\rho_{(2)} < \rho_{0(2)}$										
0.5	0.6	0.1	0.4	2.5	0.6	0.5	0.6	0.8	49	10	50	1318.64	187.97	918.79
0.5	0.6	0.2	0.4	2.5	0.6	0.5	0.6	0.8	50	11	50	1327.31	206.16	918.67
0.5	0.6	0.3	0.4	2.5	0.6	0.5	0.6	0.8	50	12	50	1334.63	228.26	918.53
				$\rho_{(2)} = \rho_{0(2)}$										
0.7	0.2	0.5	0.5	1.5	0.6	0.2	0.6	0.5	50	3	47	2053.17	83.83	1215.61
0.7	0.2	0.7	0.7	1.5	0.6	0.2	0.6	0.5	49	5	46	2042.58	129.04	1174.25
0.7	0.2	0.8	0.8	1.5	0.6	0.2	0.6	0.5	49	7	44	2024.47	175.66	1124.21

Table 6: Sample sizes and corresponding expected cost of survey, which have the same precision than $\hat{R}_2^{**}$, $\hat{R}_{21}^{**}$ and $\hat{R}_{22}^{**}$, with respect to $\hat{R}^{*}$ for different values of W_2, $(k-1)$ and q

ρ	ρ_0	$\rho_{(2)}$	$\rho_{0(2)}$	$(k-1)$	W_2	C_0	$C_{0(2)}$	q	n_0^*	n_1^*	n_2^*	$E(C_0^*)$	$E(C_1^*)$	$E(C_2^*)$
				W_2										
0.7	0.2	0.5	0.5	1.5	0.2	0.2	0.6	0.8	49	8	44	1004.08	121.68	669.25
0.7	0.2	0.5	0.5	1.5	0.4	0.2	0.6	0.8	50	5	47	1533.03	96.49	964.86
0.7	0.2	0.5	0.5	1.5	0.8	0.2	0.6	0.8	50	2	49	2576.81	78.96	1500.26
				$(k-1)$										
0.8	0.3	0.4	0.6	0.5	0.8	0.2	0.6	0.4	50	4	43	7364.59	287.52	3392.78
0.8	0.3	0.4	0.6	1.0	0.8	0.2	0.6	0.4	49	2	46	3740.49	88.11	1991.39
0.8	0.3	0.4	0.6	1.5	0.8	0.2	0.6	0.4	49	1	48	2541.38	43.97	1468.70
				q										
0.8	0.3	0.4	0.6	1.5	0.8	0.2	0.6	0.2	49	1	48	2554.44	44.72	1493.54
0.8	0.3	0.4	0.6	1.5	0.8	0.2	0.6	0.4	49	1	48	2541.38	43.97	1468.70
0.8	0.3	0.4	0.6	1.5	0.8	0.2	0.6	0.6	49	1	47	2540.89	43.89	1465.84

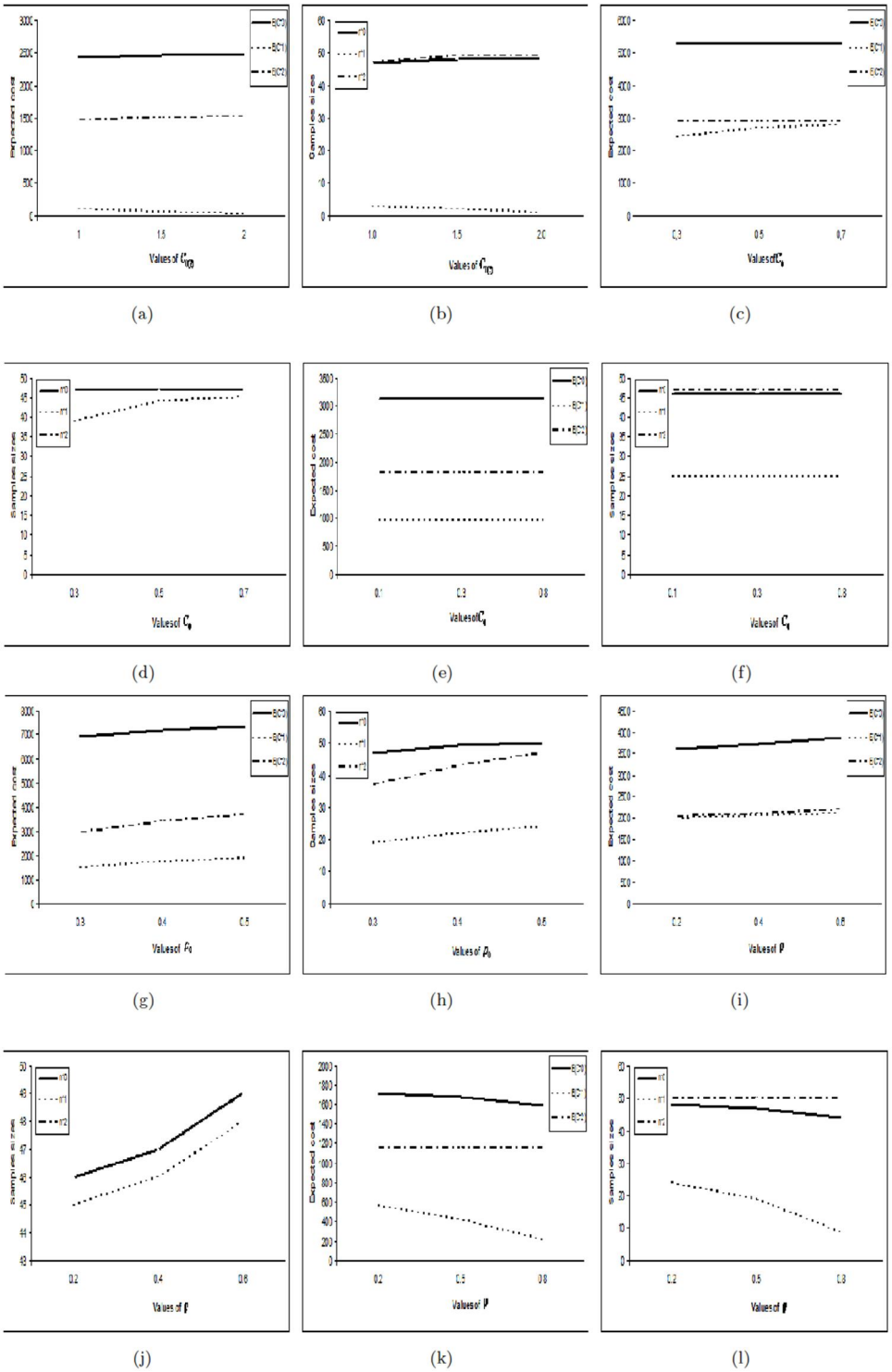

Figure 3: Corresponding expected cost of survey and sample sizes, which have the same precision than $\mathbb{R}_2^{**}$, $\mathbb{R}_{21}^{**}$ and $\mathbb{R}_{22}^{**}$, with respect to $\mathbb{R}^{*}$ for (a)-(b) different values of $C_{0(2)}$, (c)-(d) different values of C_0, (e)-(f) the case $C_{0(2)}$ and C_0, (g)-(h) different values of ρ_0, (i)-(j) different values of ρ, (k)-(l) the case $\rho = \rho_0$.

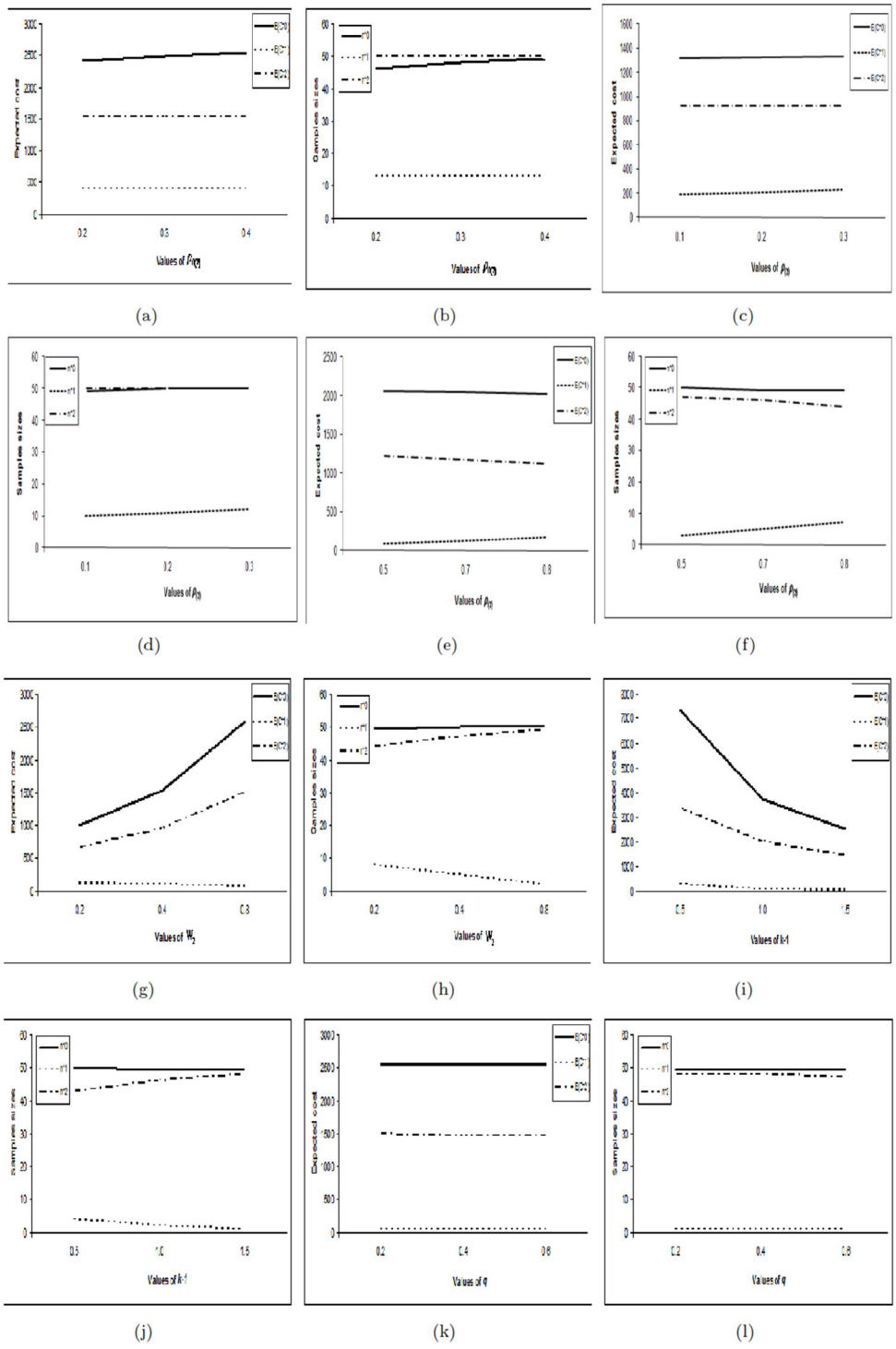

Figure 4: Corresponding expected cost of survey and sample sizes, which have the same precision than $\hat{R}_2^{**}$, $\hat{R}_{21}^{**}$ and $\hat{R}_{22}^{**}$, with respect to $\hat{R}^{*}$ for (a)-(b) different values of $\rho_{0(2)}$, (c)-(d) different values of ρ_2, (e)-(f) the case $\rho_{0(2)} = \rho_2$, (g)-(h) different values of W_2, (i)-(j) different values of $k-1$, and (k)-(l) different values of q.

5. Conclusions

In this paper, we have used the HH technique for estimating the population ratio 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. The results obtained reveals that the gain in precision is maximum for the estimation of the ratio of mean when there is non-response only on the first occasion, whereas it is least for the estimation of the ratio of mean when there is non-response only on the second occasion and when there is non-response on both occasions. Also, we have derived the sample sizes and the saving in cost for all the estimators that have the same precision than the direct estimate using no information gathered on the first occasion. In the majority of the cases the sample sizes and the saving in cost is maximum for the estimation of the ratio of mean when there is non-response on both occasions, whereas it is least for the estimation of the ratio of mean when there is non-response only on the first occasion and when there is non-response only on the second occasion.

References

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

2. 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.

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

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

5. García, 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.

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

7. Garcia, 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.

8. 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.

9. 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).

10. 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.

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

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

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

14. 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.

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

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

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

18. 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.

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

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