Thailand Statistician
July 2023; 21(3): 616-630
http://statassoc.or.th
Contributed paper

Optimal Estimation of Population Mean in the Presence of Random Non-Response

Shashi Bhushan [a] and Abhay Pratap Pandey*[b]
[a] Department of Mathematics and Statistics, Dr. Shakuntala Misra National Rehabilitation University, Lucknow, India.
[b] Department of Statistics, Ramanujan College, University of Delhi, India.
*Corresponding author; e-mail: abhaypratap.pandey@gmail.com

Received: 6 February 2021
Revised: 21 September 2021
Accepted: 28 September 2021

Abstract

In the present paper, we have proposed some improved ratio and regression (or difference) type estimators of the finite population mean utilizing the information on auxiliary variables in the presence of random non-response. Using the Searls (1964) methodology, some improved ratio and regression (or difference) type estimators have been suggested in two different situations of random non-response and studied their properties. Proposed classes of estimators are empirically compared with some contemporary estimators of population mean under similar realistic conditions. Their performances have been demonstrated through numerical illustration followed by suitable recommendations.

Keywords: Finite population, auxiliary information, random non-response, mean square error, percentage relative efficiency.

1. Introduction

The negative impact of non-response of the population means estimators are well known. In survey sampling, non-response or missingness is a significant problem encountered by the practitioners. Statisticians have proved the incompleteness in the form of missingness of data can spoil inference. The problem of non-response was first studied by Hansen and Hurvitz (1946) in relation to a mail survey. Rubin (1976) suggested three concepts of missingness: missing at random (MAR), observed at random (OAR), and parameter distribution (PD). Heitjan and Basu (1996) distinguished between missing at random (MAR) and missing completely at random (MCAR). The non-response adversely affected the estimate of the population characteristics in survey sampling. Many authors dealt with the problem of non-response and suggested different methods to estimate the population characteristics under non-response. The imputation technique is frequently used to replace the missing data to overcome missing observations or non-response in sample surveys.

Numerous authors have proposed various estimators for estimating the population mean and variance under the random non-responses. Singh and Joarder (1998) studied the properties of ratio type estimator of population variance suggested by Isaki (1983) under two different situations of random non-response (MAR) advocated by Tracy and Osahan (1994) when (i) random non-response on both the study and auxiliary variables and (ii) only on the study variable. Singh et al. (2012) revisited the family of estimators of population variance suggested by Srivastava (1981) under the above situations of random non-responses.

The practical use of the auxiliary information benefits the estimation procedure. The estimation of population means using supplementary information has drawn the attention of various influential authors, and some of them are given below. Srivastava and Jhajj (1981) suggested a generalized class of estimators of the population mean in survey sampling using auxiliary information, which incorporated many important estimators as its particular case. Singh and Tracy (2001) have suggested an estimator of the population mean in the presence of random non-response. Similarly, under random non-response, some authors suggested improving the estimator's efficiency by using auxiliary information. Singh and Joarder (1998), Singh et al. (2000), and Singh and Singh (1979, 1985) presented estimators of finite population variance and mean using random non-response in survey sampling. Singh et al. (2012) and Singh et al. (2007) suggested a family of estimators for mean, ratio, and product of finite population and finite population variance under random non-response, and Singh and Tracy (2001) and Maji et al. (2019) suggested an estimator of the population mean in the presence of random non-response. Hafeez and Shabbir (2015), Ahmed et al. (2016) and Hafeez et al. (2020) also proposed some estimation methods under different sampling strategy.

2. Distribution of Random Non-Response, Notations, and Expectations

Let $U=\left(U_1,U_2,...,U_N\right)$ denote a population of N units from which a simple random sample of size n is drawn without replacement. If $r\left(r=0,1,2,...,\left(n-2\right)\right)$ represents the number of sampling units on which information could not be obtained due to a random non-response, then the remaining $\left(n-r\right)$ units can be treated as a simple random sample from U. It is assumed that r is less than $\left(n-1\right)$, i.e., $0\le r\le\left(n-2\right)$. Singh and Joarder (1998) took that r has the following discrete distribution as

$$P\left(r\right)=\frac{n-r}{\left(nq+2p\right)}\binom{n-2}{r}p^r q^{n-2-r}, \tag{1}$$

where p is the probability of non-response, $q=1-p$ and $\binom{n-2}{r}$ represents the total number of ways to obtain r non-response out of a possible $\left(n-2\right)$. We define for the variate y_i and x_i.

$\bar{X}=N^{-1}\sum_{i=1}^{N}x_i$: Population mean of the i^{th} variate x_i.

$\bar{Y}=N^{-1}\sum_{i=1}^{N}y_i$: Population mean of the i^{th} variate y_i.

C_x, C_y : The population coefficient of variation (CV) of the variate y and x.

$\rho_{il}=S_{yx}/S_yS_x$: The population correlation coefficient between the variates y and x.

$$\left(n-1\right)S_{yx}=\sum_{i=1}^{N}\left(y_i-\bar{Y}_i\right)\left(x_i-\bar{X}\right);\;\; K_{xy}=\rho_{xy}\frac{C_y}{C_x},\;\; K=\left(K_{12}\right),\;\; d=\left(\lambda_{12}C_y+\lambda_{12}C_x\right),$$

$$\lambda_{w_1,w_2,} = \mu_{w_1,w_2} \Big/ (\mu_{20})^{\frac{w_1}{2}} (\mu_{02})^{\frac{w_2}{2}}, \ (N-1)\mu_{w_1,w_2} = \sum_{i=1}^{N} (y_i - \bar{Y})^{w_1} (x_i - \bar{X})^{w_2},$$

where w_1 and w_2 are non-negative integers. Define the following terms

$$B = C_x^2 K^2 + \frac{(C_x K \lambda_{03} - d)^2}{(\lambda_{04} - \lambda_{03}^2 - 1)} \ \text{ or } \ B = (C_x^2 K^2 + \Delta B_1^2),$$

where $B_1 = \frac{\Delta_2}{\Delta}, \Delta_2 = (C_x K \lambda_{03} - d)$ and $\Delta = (\lambda_{04} - \lambda_{03}^2 - 1)$. For the variates y_i and x_i in the sample:

$$\bar{y}_i = n^{-1} \sum_{i=1}^{n} y_i, \ \bar{y}_i^* = (n-r)^{-1} \sum_{i=1}^{n-r} y_i, \ s_x^2 = (n-1)^{-1} \sum_{i=1}^{n} (x_i - \bar{x})^2,$$

$s_x^{*2} = (n-1)^{-1} \sum_{i=1}^{n-r} (x_i - \bar{x}^*)^2$ are conditionally unbiased estimators of S_x^2, respectively, where $\theta^* = \left(\frac{1}{nq + 2p} - \frac{1}{N} \right)$ and $\theta = \left(\frac{1}{n} - \frac{1}{N} \right)$.

Singh and Joarder (1998) obtained the following maximum likelihood estimator of p (probability of non-response), as

$$\hat{p} = \frac{(n-1+r) - \sqrt{(n-1+r)^2 - \frac{4rn(n-3)}{(n-2)}}}{2(n-3)}. \tag{2}$$

If $r = 0$ then $\hat{p} = 0$, and $r = n-2$ then, $\hat{p} = 1$, thus $\hat{p}$ is an admissible estimator of response probability p.

Let us define $\bar{y}^* = \bar{Y}(1+\epsilon_0)$, $\bar{x}^* = \bar{X}(1+\epsilon_1)$, $\bar{x} = \bar{X}(1+\epsilon_2)$, $s_x^{*2} = S_x^2(1+\epsilon_3)$ and $s_x^2 = S_x^2(1+\epsilon_4)$. Then under the model

$E(\varepsilon_i) = 0, (i = 0,1,...,4)$,

$E(\varepsilon_0^2) = \theta^* C_y^2, \ E(\varepsilon_1^2) = \theta^* C_x^2, \ E(\varepsilon_2^2) = \theta C_x^2, \ E(\varepsilon_3^2) = \theta^* (\lambda_{04} - 1), \ E(\varepsilon_4^2) = \theta (\lambda_{04} - 1)$,

$E(\varepsilon_0 \varepsilon_1) = \theta^* \rho_{yx} C_y C_x, \ E(\varepsilon_0 \varepsilon_2) = \theta \rho_{yx} C_y C_x, \ E(\varepsilon_0 \varepsilon_3) = \theta^* \lambda_{12} C_y, \ E(\varepsilon_0 \varepsilon_4) = \theta \lambda_{12} C_y$,

$E(\varepsilon_1 \varepsilon_2) = \theta C_x^2, E(\varepsilon_1 \varepsilon_3) = \theta^* \lambda_{03} C_x, \ E(\varepsilon_1 \varepsilon_4) = \theta \lambda_{03} C_x, \ E(\varepsilon_2 \varepsilon_3) = \theta \lambda_{03} C_x, \ E(\varepsilon_2 \varepsilon_4) = \theta \lambda_{03} C_x$, and $E(\varepsilon_3 \varepsilon_4) = \theta (\lambda_{04} - 1)$.

3. Exiting estimators

Srivastava and Jhajj (1981) suggested a class of estimators of the population mean is defined by

$$t = \bar{y} g(u, v), \tag{3}$$

where $g(u,v)$ is a function of (u,v), where $u = \frac{\bar{x}}{\bar{X}}$ and $v = \frac{s_x^2}{S_x^2}$, such that $g(1,1) = 1$ and satisfying certain regularity conditions.

To the first degree of approximation, the bias and minimum mean square error of t is respectively, given below

$$Bias(t) = \frac{\theta\bar{Y}}{2}\begin{Bmatrix} 2K_{yx}C_x^2 g_1(1,1) + 2\lambda_{12}g_2(1,1) + C_x^2 g_{11}(1,1) + 2\lambda_{03}C_x g_{12}(1,1) \\ +(\lambda_{04}-1)g_{22}(1,1) \end{Bmatrix}, \tag{4}$$

$$MSE(t) = \theta S_y^2\left[1-\rho_{yx}^2 - (\lambda_{03}\rho_{yx} - \lambda_{12})^2/\Delta\right]. \tag{5}$$

Similarly, Singh et al. (2007) suggest a family of the estimator under three different strategies. Which provided almost all the estimators proposed till that date are their particular cases.

Strategy I: When random non-response for r units on study variable y and auxiliary variable x is present in the sample, and the population mean $\bar{X}$ and the variance S_x^2 of x are known.

We define a family of estimators $\bar{Y}$ as

$$t_3 = \bar{y}^* f\left(u^\bullet, v^\bullet\right), \tag{6}$$

where $f\left(u^\bullet, v^\bullet\right)$ is a function of $\left(u^\bullet, v^\bullet\right)$, where $u^\bullet = \frac{\bar{x}^*}{\bar{X}}$ and $v^\bullet = \frac{s_x^{*2}}{S_x^2}$, such that $f(1,1)=1$ and satisfying the following regularity conditions:

1. Whatever the sample $\left(u^\bullet, v^\bullet\right)$ assumes values in a bounded, closed convex subset, S, of the real two-dimensional space containing the point $(1,1)$.

2. In S, the function $f\left(u^\bullet, v^\bullet\right)$ is continuous and bounded.

3. The first and second-order partial derivatives of $f\left(u^\bullet, v^\bullet\right)$ existing and are continuous and bounded in S. To the first degree of approximation minimum, MSE of t_3 is given by

$$\min MSE(t_3) = \theta^* S_y^2\left[1-\rho_{yx}^2 - (\lambda_{03}\rho_{yx} - \lambda_{12})^2/\Delta\right], \tag{7}$$

where $\bar{y}^*$, $t_1 = \bar{y}^* f_1\left(u^\bullet\right)$ and $t_2 = \bar{y}^* t_2\left(v^\bullet\right)$ are the members of the above family and the mean square error of $\bar{y}^*$, t_1 and t_2 are provided by

$$Var\left(\bar{y}^*\right) = \theta^* S_y^2, \tag{8}$$

$$\min MSE(t_1) = \theta^* S_y^2\left[1-\rho_{yx}^2\right], \tag{9}$$

$$\min MSE(t_2) = \theta^* S_y^2\left[1-\lambda_{12}^2/\left(\lambda_{04}-1\right)\right]. \tag{10}$$

Strategy II: We consider the situation when information on a variable y can not be obtained for r units, while the population mean $\bar{X}$ and the variance S_x^2 of the auxiliary variable x is known.

We consider the following family of the estimator $\bar{Y}$ given by

$$t_6 = \bar{y}^* \phi(u,v), \tag{11}$$

where $\phi(u,v)$ is a function of (u,v), $u = \frac{\bar{x}}{\bar{X}}$ and $v = \frac{s_x^2}{S_x^2}$, such that $\phi(1,1)=1$ and satisfying certain regularity conditions, as defined similarly as in Strategy I. To the first degree of approximation, MSE of t_6 is given by

$$\min MSE\left(t_6\right)=\theta S_y^2\left[1-\rho_{yx}^2-(\lambda_{03}\rho_{yx}-\lambda_{12})^2/\Delta\right]+\left(\theta^*-\theta\right)S_y^2, \tag{12}$$

where $t_4=\bar{y}^*\phi_1\left(u\right)$ and $t_5=\bar{y}^*\phi_2\left(v\right)$ are the members of the above family and the minimum MSE of t_4 and t_5 are provided by

$$\min MSE\left(t_4\right)=\theta S_y^2\left(1-\rho_{yx}^2\right)+\left(\theta^*-\theta\right)S_y^2, \tag{13}$$

$$\min MSE\left(t_5\right)=\theta S_y^2\left[1-\lambda_{12}^2/\left(\lambda_{04}-1\right)\right]+\left(\theta^*-\theta\right)S_y^2. \tag{14}$$

Strategy III: We consider the situation when information on the study variable y can not be obtained for r units. At the same time, information on the auxiliary variable x is obtained for all the sample units. But the population mean $\bar{X}$ and the variance S_x^2 of the additional character x is not known.

Under these circumstances, Singh et al. (2007) defines the following class of estimator $\bar{Y}$ as

$$t_9=\bar{y}^*\psi\left(u^*,v^*\right), \tag{15}$$

where $\psi\left(u^*,v^*\right)$ is a function of $\left(u^*,v^*\right)$, $u=\dfrac{\bar{x}^*}{\bar{x}}$ and $v=\dfrac{s_x^{*2}}{s_x^2}$, such that $\psi\left(1,1\right)=1$ and satisfying the regularity conditions, as defined similarly as in Strategy I. To the first degree of approximation, MSE of t_9 is given by

$$\min MSE\left(t_9\right)=S_y^2\left[\theta^*-\left(\theta^*-\theta\right)\left\{\rho_{yx}^2+(\lambda_{03}\rho_{yx}-\lambda_{12})^2/\Delta\right\}\right], \tag{16}$$

where $t_7=\bar{y}^*\psi_1\left(u^*\right)$ and $t_8=\bar{y}^*\psi_2\left(v^*\right)$ are the members of the above family and the minimum MSE of t_7 and t_8 are defined as below

$$\min MSE\left(t_7\right)=S_y^2\left[\theta^*-\left(\theta^*-\theta\right)\rho_{yx}^2\right], \tag{17}$$

$$\min MSE\left(t_8\right)=S_y^2\left[\theta^*-\left(\theta^*-\theta\right)\lambda_{12}^2/\left(\lambda_{04}-1\right)\right], \tag{18}$$

The minimum MSE of the conventional estimators t_i $\left(i=1,2,...,9\right)$ can write as

$$\min MSE\left(T_i\right)=\bar{Y}^2\left\{1-\left(1-\frac{\min .MSE\left(t_i\right)}{\bar{Y}^2}\right)\right\}. \tag{19}$$

Singh et al. (2000) have proposed some regression type estimators t_{d_i} and Ahmed et al. (2009) have also proposed ratio type estimators t_{r_i} under random non-response, which are the special case of Singh et al. (2007).

In this paper, we have proposed improved regression (or difference) and ratio type estimators using transformation method under rndom non-response. Proposed estimators provide the better results in term of percentage relative efficiency (PRE) in comparison to the existing estimators.

4. Proposed Estimators

We propose improved difference and ratio type estimators using the Searls methodology. Searls (1964) suggested a technique for improving the conventional estimators by multiplying a constant tuning term α whose optimum value depends on the coefficient of variation, a reasonably stable quantity. We refer to this multiplication technique by a tuning constant α as Searls type transformation (STT) under three different strategies in single-phase sampling stated hereunder.

$$T_s = \alpha \bar{y}^*, \tag{20}$$

the optimum value of α is given by

$$\alpha = \frac{1}{1+\theta^* C_y^2},$$

The proposed difference and ratio type estimators under three different strategies using Searls methodology are defined below

Strategy I:

$$T_{d_1} = \alpha_1 \bar{y}^* + d_1\left(\bar{x}^* - \bar{X}\right), \tag{21}$$

$$T_{d_2} = \alpha_2 \bar{y}^* + d_2\left(s_x^{*2} - S_x^2\right), \tag{22}$$

$$T_{d_3} = \alpha_3 \bar{y}^* + d_3\left(\bar{x}^* - \bar{X}\right) + k_1\left(s_x^{*2} - S_x^2\right), \tag{23}$$

$$T_{r_1} = \gamma_1 \bar{y}^* \left(\frac{\bar{X}}{\bar{x}^*}\right)^{\beta_1}, \tag{24}$$

$$T_{r_2} = \gamma_2 \bar{y}^* \left(\frac{S_x^2}{s_x^{*2}}\right)^{\beta_2}, \tag{25}$$

$$T_{r_3} = \gamma_3 \bar{y}^* \left(\frac{\bar{X}}{\bar{x}^*}\right)^{\beta_3} \left(\frac{S_x^2}{s_x^{*2}}\right)^{\eta_1}. \tag{26}$$

Strategy II:

$$T_{d_4} = \alpha_4 \bar{y}^* + d_4\left(\bar{x} - \bar{X}\right), \tag{27}$$

$$T_{d_5} = \alpha_5 \bar{y}^* + d_5\left(s_x^2 - S_x^2\right), \tag{28}$$

$$T_{d_6} = \alpha_6 \bar{y}^* + d_6\left(\bar{x} - \bar{X}\right) + k_2\left(s_x^2 - S_x^2\right), \tag{29}$$

$$T_{r_4} = \gamma_4 \bar{y}^* \left(\frac{\bar{X}}{\bar{x}}\right)^{\beta_4}, \tag{30}$$

$$T_{r_5} = \gamma_5 \bar{y}^* \left(\frac{S_x^2}{s_x^2}\right)^{\beta_5}, \tag{31}$$

$$T_{r_6} = \gamma_6 \bar{y}^* \left(\frac{\bar{X}}{\bar{x}}\right)^{\beta_6} \left(\frac{S_x^2}{s_x^2}\right)^{\eta_2}. \tag{32}$$

Strategy III:

$$T_{d_7} = \alpha_7 \bar{y}^* + d_7\left(\bar{x}^* - \bar{x}\right), \tag{33}$$

$$T_{d_8} = \alpha_8 \bar{y}^* + d_8\left(s_x^{*2} - s_x^2\right), \tag{34}$$

$$T_{d_9} = \alpha_9 \bar{y}^* + d_9\left(\bar{x}^* - \bar{x}\right) + k_3\left(s_x^{*2} - s_x^2\right), \tag{35}$$

$$T_{r_7} = \gamma_7 \bar{y}^* \left(\frac{\bar{x}}{\bar{x}^*} \right)^{\beta_7}, \tag{36}$$

$$T_{r_8} = \gamma_8 \bar{y}^* \left(\frac{s_x^2}{s_x^{*2}} \right)^{\beta_8}, \tag{37}$$

$$T_{r_9} = \gamma_9 \bar{y}^* \left(\frac{\bar{x}}{\bar{x}^*} \right)^{\beta_9} \left(\frac{s_x^2}{s_x^{*2}} \right)^{\eta_3}. \tag{38}$$

The biases and MSE's of the proposed estimators are given in subsequent theorems.

Theorem 1 *The bias and minimum MSE of the proposed ratio estimator* T_{r_j} $(j = 1, 2, ..., 9)$ *are given by*

$$Bias_{\alpha_j}\left(T_{r_j}\right) = \bar{Y}\left(\gamma_j - 1\right) + \alpha_j Bias_1\left(t_{r_j}\right), \tag{39}$$

and

$$\min MSE\left(T_{r_j}\right) = \bar{Y}^2 \left(1 - \frac{B_j^2}{A_j}\right), \tag{40}$$

where $Bias_{\gamma_j}\left(T_{r_j}\right)$ is the first-order bias with parameter γ_j and $Bias_1\left(t_{r_j}\right)$ is the first order bias of the conventional estimator's counterpart with $\gamma_j = 1$.

Proof: See Appendix.

It is noteworthy that the simultaneous optimization w.r.t, γ_j, β_j and η_l of the expression of MSE is not possible. We use the optimum value of $\beta_j = \beta_{jopt}$ when $\gamma_j = 1$ and use this within $\gamma_j = \gamma_{jopt}$ to obtain.

Theorem 2 *The bias and minimum MSE of the proposed difference estimator* T_{d_i} *is given by*

$$Bias\left(T_{d_i}\right) = \bar{Y}\left(\alpha_i - 1\right), \tag{41}$$

and

$$\min MSE_{\alpha_i}\left(T_{d_i}\right) = \frac{\bar{Y}^2 MSE_1\left(t_i\right)}{\bar{Y}^2 + MSE_1\left(t_i\right)}, \tag{42}$$

where $MSE_{\alpha_i}\left(T_{d_i}\right)$ is the first order MSE with parameter α_i and $MSE_1\left(t_i\right)$ is the first order MSE of the conventional estimator's counterpart with $\alpha_i = 1$.

Proof: See Appendix.

Theorem 3 *The proposed difference estimators for population mean* T_{d_i} *have always lesser MSE* *than the conventional estimators for population mean* t_i $(i = 1, 2, ...9)$.

Proof: It may be easily observed from a comparison of (40) with (42). It is interesting to note that simultaneous optimization w.r.t. the characterizing scalars α_i, d_i and k_j of the expression of MSE is possible for proposed difference estimator but not proposed ratio estimators.

Theorem 4 *The proposed difference estimation method is better than proposed ratio estimation methods* $(i = 1, 2, ..., 9)$ *iff*

$$\alpha_{iopt} > \frac{B_j^2}{A_j}, \tag{43}$$

and vice versa. Otherwise, both are equally efficient in the case of equality in (43).

Proof: It may be easily observed from (42) that the MSE of proposed difference estimators T_{d_i} is given by

$$\min MSE\left(T_{d_i}\right) = \bar{Y}^2\left(1 - \alpha_{iopt}\right). \tag{44}$$

Comparing (44) with (40), we have the theorem.

Theorem 5 *The proposed ratio estimation method is better than the conventional estimator's* $t_i, \left(i = j = 1, 2, ..., 9\right)$ *iff*

$$\frac{B_j^2}{A_j} > \left\{1 - \frac{\min MSE\left(t_i\right)}{\bar{Y}^2}\right\}, \tag{45}$$

and vice versa. Otherwise, both are equally efficient in the case of equality in (45)

Proof: It may be easily observed from (40) that the MSE of proposed ratio estimators T_{r_j} is given by

$$\min MSE\left(T_{r_j}\right) = \bar{Y}^2\left(1 - \frac{B_j^2}{A_j}\right). \tag{46}$$

Comparing (46) with (19), we have the theorem.

5. Empirical Study

Suppose that a bank selected a simple random sample of twenty states without replacement from the USA during 1997 and collected information (in thousands) on real (y) and nonreal estate farm loans (x). The selected states are CA, CT, FL, IL, ME, MS, MO, NE, NJ, NM, ND, OK, SC, TN, TX, UT, VA, WA, WV, and WI. For details of the data set, please see population-1 on page 1111 in Singh (2003). However, assume the information on the real estate farm loans was unavailable for four states ME, ND, TX, and VA. Parameters of the population are given below:

$Y = 27771.73$, $X = 43908.12$, $\bar{Y} = 555.43$, $\bar{X} = 878.16$, $S_x^2 = 1176526$, $S_y^2 = 342021.5$,

$C_x^2 = 1.5256$, $C_y^2 = 1.1086$, $\lambda_{03} = 1.5936$, $S_{xy} = 509910.41$, $\rho_{xy} = 0.8038$, $N = 50$,

β = 0.4334, λ_{12} = 1.0982, λ_{21} = 0.9387, λ_{04} = 4.5247, λ_{22} = 2.8411. PRE of the proposed estimator is defined by

$$PRE = \frac{Var(\bar{y}^*)}{\min MSE(T_{d_i} \text{ or } T_{r_i})} \times 100.$$

Table 1 MSE and PRE of the proposed estimators

Estimators	MSE(PRE)	Proposed estimators	MSE(PRE)	Proposed estimators	MSE(PRE)
$\bar{y}^*$	14811.97(100)	$\bar{y}^*$	14811.97(100)	$\bar{y}^*$	14811.97(100)
		Strategy I			
$t_1 = t_{d_1} = t_{r_1}$	5,242.04(282.56)	T_{d_1}	5,154.45(287.36)	T_{r_1}	5,067.80**(292.27)**
$t_2 = t_{d_2} = t_{r_2}$	9,743.77(152.01)	T_{d_2}	9,445.44(156.81)	T_{r_2}	9,064.70**(163.40)**
$t_3 = t_{d_3} = t_{r_3}$	4,739.97(312.49)	T_{d_3}	4,668.24**(317.29)**	T_{r_3}	4,670.49(317.14)
		Strategy II			
$t_4 = t_{d_4} = t_{r_4}$	8,182.62(181.02)	T_{d_4}	7,971.19(185.81)	T_{r_4}	7,887.66**(187.78)**
$t_5 = t_{d_5} = t_{r_5}$	11,301.09(131.07)	T_{d_5}	10,901.74(135.86)	T_{r_5}	10,621.99**(139.44)**
$t_6 = t_{d_6} = t_{r_6}$	3,283.50(451.10)	T_{d_6}	3,248.92**(455.90)**	T_{r_6}	7,643.21(193.79)
		Strategy III			
$t_7 = t_{d_7} = t_{r_7}$	11,871.38(124.77)	T_{d_7}	11,431.49(129.57)	T_{r_7}	11,381.57**(130.14)**
$t_8 = t_{d_8} = t_{r_8}$	13,254.64(111.75)	T_{d_8}	12,708.62(116.55)	T_{r_8}	12,575.49**(117.78)**
$t_9 = t_{d_9} = t_{r_9}$	11,717.11(126.41)	T_{d_9}	11,288.37**(131.21)**	T_{r_9}	11,289.67(131.19)

Table 1 shows the numerical results of the MSE and PRE of the estimators. It can be easily observed that the proposed estimators T_{d_i} or $T_{r_i}, i = 1, 2, ..., 9$ always perform better than the conventional estimators $t_i, i = 1, 2, ..., 9$ under the respective strategy. And the main important finding of this study is that it is a well-known fact that the optimal ratio type estimator attains the mean square error (MSE) of the regression estimator (or optimal difference estimator). Still, while using the proposed methodology, this may not always happen. In Table 1, we have observed that proposed ratio type estimators perform better than difference type estimators under the optimality condition (43).

6. Simulation Study

Simulation has been used to calculate the MSE of the estimators of $\bar{Y}$ using R software. Consider an actual population of size N=50 for the simulation study. The following steps involved in simulations are

Step 1. Select a sample (SRSWOR) of size 20 from the population of size 50.

Step 2. Select value of r.

Step 3. Drop r units randomly from each sample in Step 1.

Step 4. Find the value of the estimator based on the $(n-r)$ observation.

Step 5. Repeat the Steps 1, 2, 3, and (4) 50,000 times. Thus, we obtain 50,000 values for an estimate of the population mean.

Step 6. The MSE of proposed estimator is obtained by $MSE(T_i) = \frac{1}{50,000}\sum_{i=1}^{50,000}(T_i - \bar{Y})^2$.

Percentage relative efficiency of the conventional and the proposed estimators is given by

$$PRE = \frac{Var(\bar{y}^*)}{\min MSE(T_{d_i} \text{ or } T_{r_i})} \times 100.$$

The following tables are obtained.

Table 2 PRE of the proposed estimators with conventional estimators for $r = 4$

Estimator	PRE	Estimator	PRE	Estimator	PRE	Estimator	PRE
$\bar{y}^*$	100	T_s	106.5175	$\bar{y}^*$	100	T_s	106.5175
Strategy I							
t_{d_1}	281.4288	T_{d_1}	288.1601	t_{r_1}	255.8096	T_{r_1}	263.3846
t_{d_2}	149.0583	T_{d_2}	155.9210	t_{r_2}	140.8327	T_{r_2}	150.5283
t_{d_3}	283.2874	T_{d_3}	290.0656	t_{r_3}	255.9930	T_{r_3}	264.6997
Strategy II							
t_{d_4}	182.1180	T_{d_4}	188.7476	t_{r_4}	173.7333	T_{r_4}	180.8168
t_{d_5}	129.8997	T_{d_5}	136.5529	t_{r_5}	125.1248	T_{r_5}	135.6213
t_{d_6}	182.5743	T_{d_6}	189.2160	t_{r_6}	173.7333	T_{r_6}	181.3345
Strategy III							
t_{d_7}	123.8624	T_{d_7}	130.5151	t_{r_7}	122.7737	T_{r_7}	129.5189
t_{d_8}	110.6685	T_{d_8}	117.3967	t_{r_8}	108.8135	T_{r_8}	116.7958
t_{d_9}	123.9794	T_{d_9}	130.6440	t_{r_9}	122.7652	T_{r_9}	129.6646

Table 3 PRE of the proposed estimators with conventional estimators for $r = 6$

Estimator	PRE	Estimator	PRE	Estimator	PRE	Estimator	PRE
$\bar{y}^*$	100	T_s	108.1070	$\bar{y}^*$	100	T_s	108.1070
Strategy I							
t_{d_1}	284.0613	T_{d_1}	291.7685	t_{r_1}	255.7177	T_{r_1}	264.4089
t_{d_2}	148.8680	T_{d_2}	156.7626	t_{r_2}	140.2486	T_{r_2}	151.1669
t_{d_3}	285.6681	T_{d_3}	293.3750	t_{r_3}	255.3914	T_{r_3}	265.4026
Strategy II							
t_{d_4}	161.0028	T_{d_4}	168.9215	t_{r_4}	154.3354	T_{r_4}	162.5339
t_{d_5}	124.0667	T_{d_5}	131.7942	t_{r_5}	120.1163	T_{r_5}	131.4047
t_{d_6}	161.3059	T_{d_6}	169.2018	t_{r_6}	154.3266	T_{r_6}	162.9504
Strategy III							
t_{d_7}	136.7236	T_{d_7}	144.5800	t_{r_7}	134.5616	T_{r_7}	142.6565
t_{d_8}	115.5755	T_{d_8}	123.6368	t_{r_8}	112.4431	T_{r_8}	122.8329
t_{d_9}	136.9203	T_{d_9}	144.7957	t_{r_9}	134.5131	T_{r_9}	142.9025

Table 4 PRE of the proposed estimators with conventional estimators for $r=8$

Estimator	PRE	Estimator	PRE	Estimator	PRE	Estimator	PRE
$\bar{y}^*$	100	T_s	109.5077	$\bar{y}^*$	100	T_s	109.5077
				Strategy I			
t_{d_1}	281.4104	T_{d_1}	292.0196	t_{r_1}	250.0481	T_{r_1}	262.0772
t_{d_2}	148.0615	T_{d_2}	158.2077	t_{r_2}	138.4214	T_{r_2}	152.6441
t_{d_3}	282.8988	T_{d_3}	293.4877	t_{r_3}	249.3356	T_{r_3}	262.9924
				Strategy II			
t_{d_4}	144.7307	T_{d_4}	154.6329	t_{r_4}	140.5065	T_{r_4}	150.7298
t_{d_5}	118.8114	T_{d_5}	128.8349	t_{r_5}	115.4661	T_{r_5}	128.9488
t_{d_6}	144.9241	T_{d_6}	154.8318	t_{r_6}	140.4502	T_{r_6}	151.0878
				Strategy III			
t_{d_7}	151.3335	T_{d_7}	161.9544	t_{r_7}	146.7905	T_{r_7}	157.8752
t_{d_8}	120.1094	T_{d_8}	130.3043	t_{r_8}	115.2867	T_{r_8}	129.2955
t_{d_9}	151.6529	T_{d_9}	162.2574	t_{r_9}	146.6897	T_{r_9}	158.2421

7. Interpretation of empirical results

Tables 1-4 shows that the proposed difference type estimators T_{d_1}, T_{d_2} and T_{d_3} for the Strategy I have the higher relative efficiency concerning the conventional estimators $\bar{y}^*$, t_1, t_2 and t_3. And proposed ratio type estimators, T_{r_1}, T_{r_2} and T_{r_3} for Strategy I, also have the higher relative efficiency concerning the traditional estimators $\bar{y}^*$, t_1, t_2 and t_3. And in Table 1, it is observed that the proposed ratio type estimators T_{r_1} and T_{r_2} have higher percent relative efficiency concerning proposed difference type estimators T_{d_1}, T_{d_2}. But in case of T_{d_3} result is reversed, means T_{d_3} have the higher percent relative efficiency concerning proposed ratio type estimators T_{r_3}. This is happening only because of the optimality condition 43. And the same result has been observed for Strategies II and III.

In simulation study section, we have performed simulation on the real data source for the different value of r. We have observed in Tables 2-4, the proposed estimators T_{d_i} or T_{r_i} have the higher PRE as comparison to the existing conventional estimators t_{d_i} or t_{r_i}. Under consider strategies, Strategy I provides the better result as compare to Strategies II and III.

8. Conclusion

This article concludes that the proposed estimators always perform better than the estimators proposed by Singh et al. (2007) under random non-response. This paper's most important conclusion is that the traditional thought that the ratio type estimator can at best match up to its regression counterpart is vitiated. We have proved that the proposed regression type estimators improve the traditional regression estimator. We have also established that the proposed ratio estimator can improve both the conventional regression estimator and proposed regression estimator provided their respective efficiency conditions (45) and (43) are satisfied. The computational result is given in Table 1 also supports this fact. And simulation results also support our findings.

Acknowledgments

The authors are deeply obliged to the honorable reviewers and the editors for the rigorous review which significantly improved the revised manuscript.

References

Ahmed MS, Al-Titi O, Al-Rawi Z, Abu-dayyeh W. Estimation of finite population mean using random non-response in survey sampling. J Appl Stat Sci. 2009; 17(1): 1-12.

Ahmed S, Shabbir J, Hafeez W. Hansen and Hurwitz estimator with scrambled response on second call in stratified random sampling. J Stat Appl Prob. 2016; 5(2): 357-369.

Hafeez W, Shabbir J. Estimation of the finite population mean, using median based estimators in stratified random sampling. J Stat Appl Prob. 2015; 4(3): 367-374.

Hafeez W, Shabbir J, Shah MT, Ahmed S. Some median type estimators to estimate the finite population mean. Asian J Prob Stat. 2020; 7(4): 48-58.

Hansen MH, Hurwitz WN. The problem of non-response in sample surveys. J Am Stat Assoc. 1946; 41(236): 517-529.

Heitjan DF, Basu S. Distinguishing "Missing at Random" and "Missing Completely at Random". Am Stat. 1996; 50(3): 207-213.

Isaki CT. Variance estimation using auxiliary information. J Am Stat Assoc. 1983; 7(381): 117–123.

Maji R, Singh GN, Bandyopadhyay A. Estimation of population mean in presence of random non-response in two-stage cluster sampling. Commun Stat A-Theor. 2019; 48(14): 3586-3608.

Rubin DB. Inference and missing data. Biometrika. 1976; 63(3): 581-592.

Searls DT. The Utilization of a known coefficient of variation in the estimation procedure. J Am Stat Assoc. 1964; 59(308): 1225-1226.

Singh HP, Tailor R, Kim JM, Singh S. Families of estimators of finite population variance using random non-response in survey sampling. Korean J Appl Stat. 2012; 25(4): 681-695.

Singh HP, Tracy DS. Estimation of population mean in the presence of random non-response in sample surveys. Statistica. 2001; 61(2):231-248.

Singh S, Joarder AH. Estimation of finite population variance using random non-response in survey sampling. Metrika. 1998; 47: 241-249.

Singh S, Joarder AH, Tracy DS. Regression type estimators for random non-response in survey sampling. Statistica. 2000; 60(1):39-44.

Singh S, Singh R. On random non-response in unequal sampling. Sankhya. 1979; 41:27-37.

Singh S, Singh R. On random non-response in double sampling with difference estimator. Commun Stat Simulat. 1985; 14(3): 747-757.

Srivastava SK. A class of estimators using auxiliary information in sample surveys. Can J Stat. 1980; 8: 253-254.

Srivastava SK, Jhajj HS. A class of estimators of the population mean in survey sampling using auxiliary information. Biometrika. 1981; 68(1): 341-343.

Tracy DS, Osahan SS. Random non-response on study variable versus on study as well as auxiliary variables. Statistica. 1994; 54(2): 163-168.

Appendix

Outline of derivation of Theorem 1. The MSE of T_{r_j} ($j=1,2,...,9$) is given by

$$MSE(T_j)=\bar{Y}^2\left[1+\gamma_j^2 A_j-2\gamma_j B_j\right].$$

The optimum values of scalars involved are tabulated below for ready reference:

$$\gamma_{jopt}=\frac{B_j}{A_j};\left(j=1,2,....9\right),$$

substituting the optimum value of γ_j in $MSE\left(T_{r_j}\right)$ we get minimum MSE

$$MSE\left(T_{r_j}\right)=\bar{Y}^2\left(1-\frac{B_j^2}{A_j}\right),$$

where

$$A_1=1+\theta^* C_y^2+2\beta_1^2\theta^* C_x^2+\beta_1\theta^*\left(C_x^2-4\rho_{yx}C_yC_x\right),$$

$$B_1=1+\frac{\beta_1^2}{2}\theta^* C_x^2+\frac{\beta_1}{2}\theta^*\left(C_x^2-2\rho_{yx}C_yC_x\right),$$

$$A_2=1+\theta^* C_y^2+2\beta_2^2\theta^*\left(\lambda_{04}-1\right)+\beta_2\theta^*\left\{\left(\lambda_{04}-1\right)-4\lambda_{12}C_y\right\},$$

$$B_2=1+\frac{\beta_2^2}{2}\theta^*\left(\lambda_{04}-1\right)+\frac{\beta_2}{2}\theta^*\left\{\left(\lambda_{04}-1\right)-2\lambda_{12}C_y\right\},$$

$$A_3=1+\theta^* C_y^2+2\beta_3^2\theta^* C_x^2+2\beta_4^2\theta^*\left(\lambda_{04}-1\right)+\beta_3\theta^*\left(C_x^2-4\rho_{yx}C_yC_x\right)$$
$$+\beta_4\theta^*\left\{\left(\lambda_{04}-1\right)-4\lambda_{12}C_y\right\}+4\beta_3\beta_4\theta^*\lambda_{03}C_x,$$

$$B_3=1+\frac{\beta_3^2}{2}\theta^* C_x^2+\frac{\beta_4^2}{2}\theta^*\left(\lambda_{04}-1\right)+\frac{\beta_3}{2}\theta^*\left(C_x^2-2\rho_{yx}C_yC_x\right)$$
$$+\frac{\beta_4}{2}\theta^*\left\{\left(\lambda_{04}-1\right)-2\lambda_{12}C_y\right\}+\beta_3\beta_4\theta^*\lambda_{03}C_x,$$

$$A_4=1+\theta^* C_y^2+2\beta_5^2\theta C_x^2+\beta_5\theta\left(C_x^2-4\rho_{yx}C_yC_x\right),$$

$$B_4=1+\frac{\beta_5^2}{2}\theta C_x^2+\frac{\beta_5}{2}\theta\left(C_x^2-2\rho_{yx}C_yC_x\right),$$

$$A_5=1+\theta^* C_y^2+2\beta_6^2\theta\left(\lambda_{04}-1\right)+\beta_6\theta\left\{\left(\lambda_{04}-1\right)-4\lambda_{12}C_y\right\},$$

$$B_5=1+\frac{\beta_6^2}{2}\theta^*\left(\lambda_{04}-1\right)+\frac{\beta_6}{2}\theta^*\left\{\left(\lambda_{04}-1\right)-2\lambda_{12}C_y\right\},$$

$$A_6=1+\theta^* C_y^2+2\beta_7^2\theta C_x^2+2\beta_8^2\theta\left(\lambda_{04}-1\right)+\beta_7\theta\left(C_x^2-4\rho_{yx}C_yC_x\right)$$
$$+\beta_8\theta\left\{\left(\lambda_{04}-1\right)-4\lambda_{12}C_y\right\}+4\beta_7\beta_8\theta\lambda_{03}C_x,$$

$$B_6 = 1+\frac{\beta_7^2}{2}\theta C_x^2 + \frac{\beta_8^2}{2}\theta\left(\lambda_{04}-1\right)+\frac{\beta_7}{2}\theta\left(C_x^2 - 2\rho_{yx}C_yC_x\right)$$
$$+\frac{\beta_8}{2}\theta\left\{\left(\lambda_{04}-1\right)-2\lambda_{12}C_y\right\}+\beta_7\beta_8\theta\lambda_{03}C_x,$$
$$A_7 = 1+\theta^* C_y^2 + \left(\theta^*-\theta\right)\left\{2\beta_9^2 C_x^2 + \beta_9\left(C_x^2 - 4\rho_{yx}C_yC_x\right)\right\},$$
$$B_7 = 1+\left(\theta^*-\theta\right)\left\{\frac{\beta_9^2}{2}C_x^2 + \frac{\beta_9}{2}\left(C_x^2 - 2\rho_{yx}C_yC_x\right)\right\},$$
$$A_8 = 1+\theta^* C_y^2 + \left(\theta^*-\theta\right)\left[2\beta_{10}^2\left(\lambda_{04}-1\right)+\beta_{10}\left\{\left(\lambda_{04}-1\right)-4\lambda_{12}C_y\right\}\right],$$
$$B_8 = 1+\left(\theta^*-\theta\right)\left[\frac{\beta_{10}^2}{2}\left(\lambda_{04}-1\right)+\frac{\beta_{10}}{2}\left\{\left(\lambda_{04}-1\right)-2\lambda_{12}C_y\right\}\right],$$
$$A_9 = 1+\theta^* C_y^2 + \left(\theta^*-\theta\right)\left[2\beta_{11}^2 C_x^2 + 2\beta_{12}^2\left(\lambda_{04}-1\right)+\beta_{11}\left(C_x^2 - 4\rho_{yx}C_yC_x\right)\right.$$
$$\left.+\beta_{12}\left\{\left(\lambda_{04}-1\right)-4\lambda_{12}C_y\right\}+4\beta_{11}\beta_{12}\lambda_{03}C_x\right],$$
$$B_9 = 1+\left(\theta^*-\theta\right)\left[\frac{\beta_{11}^2}{2}C_x^2 + \frac{\beta_{12}^2}{2}\left(\lambda_{04}-1\right)+\frac{\beta_{11}}{2}\left(C_x^2 - 2\rho_{yx}C_yC_x\right)\right.$$
$$\left.+\frac{\beta_{12}}{2}\left\{\left(\lambda_{04}-1\right)-2\lambda_{12}C_y\right\}+\beta_{11}\beta_{12}\lambda_{03}C_x\right].$$

We use the optimum value of $\beta_j = \beta_{jopt}$ and $\eta_l = \eta_{lopt}$ when $\gamma_j = 1$ and use this within $\gamma_j = \gamma_{jopt}$ to obtain. Optimum values of $\beta_j;\, j = 1,2,...,9$ and $\eta_l;\, l = 1,2,3$ are given as

$$\beta_1 = \beta_4 = \beta_7 = \rho_{yx}\frac{C_y}{C_x},\ \beta_2 = \beta_5 = \beta_8 = \frac{\lambda_{12}C_y}{\left(\lambda_{04}-1\right)} \text{ and } \beta_3 = \beta_6 = \beta_9 = \frac{C_y}{C_x}\left[\frac{\rho_{yx}\left(\lambda_{04}-1\right)-\lambda_{12}\lambda_{03}}{\left(\lambda_{04}-1-\lambda_{03}{}^2\right)}\right]$$

$$\eta_1 = \eta_2 = \eta_3 = \frac{C_y\left(\lambda_{12}-\lambda_{03}\rho_{yx}\right)}{\left(\lambda_{04}-1-\lambda_{03}{}^2\right)}.$$

Outline of Derivation of Theorem 2. The MSE of T_{d_1} is given by

$$MSE\left(T_{d_1}\right) = \frac{\bar{Y}^2\theta^* S_y^2\left(1-\rho_{yx}^2\right)}{\left[\bar{Y}^2 + \theta^* S_y^2\left(1-\rho_{yx}^2\right)\right]},\quad MSE\left(T_{d_1}\right) = \frac{\bar{Y}^2\left[MSE\left(t_1^{(1)}\right)\right]}{\left[\bar{Y}^2 + MSE\left(t_1^{(1)}\right)\right]}.$$

The optimum values of scalars involved are given below:

$$\alpha_1 = \frac{\bar{Y}^2}{\left[\bar{Y}^2 + \theta^* S_y^2\left(1-\rho_{yx}^2\right)\right]},\ d_1 = -\alpha\rho_{yx}\frac{S_y}{S_x},$$

$$\alpha_2 = \frac{\bar{Y}^2}{\left[\bar{Y}^2 + \theta^* S_y^2\left(1-\lambda_{12}^2/\left(\lambda_{04}-1\right)\right)\right]},\ d_2 = -\alpha_2\frac{\lambda_{12}S_y}{S_x^2\left(\lambda_{04}-1\right)},$$

$$\alpha_3=\frac{\bar{Y}^2}{\left[\bar{Y}^2+\theta^* S_y^2\left\{\left(1-\rho_{yx}^2\right)-\left(\lambda_{03}\rho_{yx}-\lambda_{12}\right)^2/\left((\lambda_{04}-1-\lambda_{03}^2\right)\right\}\right]},$$

$$d_3=-\alpha_3\frac{S_y}{S_x}\left\{\rho_{yx}+\frac{\lambda_{03}\left(\lambda_{03}\rho_{yx}-\lambda_{12}\right)}{\left(\lambda_{04}-1-\lambda_{03}^2\right)}\right\},$$

$$k_1=-\alpha_3\frac{S_y}{S_x^2}\left\{\frac{\left(\lambda_{03}\rho_{yx}-\lambda_{12}\right)}{\left(\lambda_{04}-1-\lambda_{03}^2\right)}\right\},$$

$$\alpha_4=\frac{\bar{Y}^2}{\left[\bar{Y}^2+S_y^2\left(\theta^*-\theta\rho_{yx}^2\right)\right]},\ d_4=-\alpha_4\rho_{yx}\frac{S_y}{S_x},$$

$$\alpha_5=\frac{\bar{Y}^2}{\left[\bar{Y}^2+S_y^2\left(\theta^*-\theta\frac{\lambda_{12}^2}{\left(\lambda_{04}-1\right)}\right)\right]},\ d_5=-\alpha_5\frac{\lambda_{12}S_y}{S_x^2\left(\lambda_{04}-1\right)},$$

$$\alpha_6=\frac{\bar{Y}^2}{\left[\bar{Y}^2+S_y^2\left\{\theta^*-\theta\left(\rho_{yx}^2+\frac{\left(\lambda_{03}\rho_{yx}-\lambda_{12}\right)^2}{\left((\lambda_{04}-1-\lambda_{03}^2\right)}\right)\right\}\right]},$$

$$d_6=-\alpha_6\frac{S_y}{S_x}\left\{\rho_{yx}+\frac{\lambda_{03}\left(\lambda_{03}\rho_{yx}-\lambda_{12}\right)}{\left(\lambda_{04}-1-\lambda_{03}^2\right)}\right\},$$

$$k_2=-\alpha_6\frac{S_y}{S_x^2}\left\{\frac{\left(\lambda_{03}\rho_{yx}-\lambda_{12}\right)}{\left(\lambda_{04}-1-\lambda_{03}^2\right)}\right\},$$

$$\alpha_7=\frac{\bar{Y}^2}{\left[\bar{Y}^2+\theta^* S_y^2-\left(\theta^*-\theta\right)S_y^2\rho_{yx}^2\right]},\ d_7=-\alpha_7\rho_{yx}\frac{S_y}{S_x},$$

$$\alpha_8=\frac{\bar{Y}^2}{\left[\bar{Y}^2+\theta^* S_y^2-\left(\theta^*-\theta\right)\frac{\lambda_{12}^2 S_y^2}{\left(\lambda_{04}-1\right)}\right]},\ d_8=-\alpha_8\frac{\lambda_{12}S_y}{S_x^2\left(\lambda_{04}-1\right)},$$

$$\alpha_9=\frac{\bar{Y}^2}{\left[\bar{Y}^2+S_y^2\left\{\theta^*-\left(\theta^*-\theta\right)\left(\rho_{yx}^2+\frac{\left(\lambda_{03}\rho_{yx}-\lambda_{12}\right)^2}{\left((\lambda_{04}-1-\lambda_{03}^2\right)}\right)\right\}\right]},$$

$$d_9=-\alpha_9\frac{S_y}{S_x}\left\{\rho_{yx}+\frac{\lambda_{03}\left(\lambda_{03}\rho_{yx}-\lambda_{12}\right)}{\left(\lambda_{04}-1-\lambda_{03}^2\right)}\right\},$$

$$k_3=-\alpha_9\frac{S_y}{S_x^2}\left\{\frac{\left(\lambda_{03}\rho_{yx}-\lambda_{12}\right)}{\left(\lambda_{04}-1-\lambda_{03}^2\right)}\right\}.$$