*International Journal of Statistical Sciences* *ISSN 1683-5603*
*Vol.24(1), March, 2024, pp 75-84*
*DOI: https://doi.org/10.3329/ijss.v24i1.72021*


# Estimating Gain in Efficiency in Complicated Randomized Response Surveys versus Simpler Alternatives

**Arijit Chaudhuri[1*] and Dipika Patra[2]**

[1]Indian Statistical Institute, Kolkata, India
Email: arijitchaudhuri1@rediffmail.com
ORCID: https://orcid.org/0000-0002-4305-7686

[2]Seth Anandram Jaipuria College, Kolkata, India
Email: dipikapatra1988@gmail.com
ORCID: https://orcid.org/0000-0003-4318-1123

[*]Correspondence should be addressed to Arijit Chaudhuri
(Email: arijitchaudhuri1@rediffmail.com)

[Received November 7, 2023; Accepted February 8, 2024]

## Abstract

Hartley and Ross's (1954) ratio-type unbiased estimator for a finite population total based on a Simple Random Sample taken Without Replacement (SRSWOR) is examined for its performance versus the expansion estimator from the sample data at hand by Chaudhuri and Samaddar (2022). They also examined how Des Raj (1956) estimator based on PPSWOR performs against SRSWOR combined with expansion estimator using PPSWOR sample values. Here we study the expansion of them to Randomized Response survey data.

**Keywords and Phrases:** PPSWOR; Qualitative & Quantitative Randomized Response Surveys; Ratio-type estimator of Hartley-Ross; SRSWOR; Symmetrized Des Raj estimator; URL; Warner.

**AMS Classification:** 62D05.

## 1. Introduction

Cochran (1953, 1963, 1977) gave a method to use stratified simple random sampling without replacement (SRSWOR) survey data at hand in examining if and how this may improve upon an alternative of unstratified SRSWOR might be contemplated for use but not actually employed. Rao (1961) gave an alternative procedure for the same purpose. Chaudhuri and Samaddar (2022) showed Rao's (1961) procedure is promising enough to examine if and how the SRSWOR-based survey data at hand could be used to examine the efficacy of Hartley and Ross's (1954) estimator for a finite population total over the traditional expansion estimator for this total. They also showed there PPSWOR-based survey data at hand could be used to check if and how Symmetrized Des Raj (SDR) estimator (1956) for a finite population total might outperform the expansion estimator for the same total if it might be based on an SRSWOR, the sample size in each case being taken the same.

In this paper we extend these two investigations in case rather than direct survey data, certain Randomized Response (RR) Techniques (RRT) might be employed instead supposing some sensitive and stigmatizing situations might be involved. We illustrate both qualitative situations employing Warner's (1965) model and Simmons's (vide Horvitz et al. (1967) and Greenberg et al's (1969)) URL model and quantitative illustrating Chaudhuri's (2011) device I.

## 2. Theoretical Details

Based on an SRSWOR of $n\ (2 \leq n \leq N)$ units taken from a finite population of N units, the expansion estimator for $Y = \sum_{i=1}^{N} y_i$, the population total of a real variable y taking values $y_i$ on units i in the finite population $U = (1,2,\ldots,i,\ldots,N)$ is $N\bar{y}$, writing $\bar{y} = \frac{1}{n}\sum_{i\in s} y_i$, the mean of a sample s of units. Then, an unbiased estimator of variance $V(N\bar{y})$ is $v(N\bar{y}) = \left(\frac{1}{n} - \frac{1}{N}\right)\frac{N^2}{n-1}\sum_{i\in s}(y_i - \bar{y})^2$.

For Y, Hartley and Ross's (1954) unbiased estimator based on SRSWOR is

$$\hat{Y}_{HR} = N\left[\bar{r} + \left(\frac{N-1}{N}\right)\left(\frac{n}{n-1}\right)\frac{(\bar{y} - \bar{r}\bar{x})}{\bar{X}}\right]\bar{X}$$
$$= N[\bar{r} + c(\bar{y} - \bar{r}\bar{x})]\bar{X},$$

say with x as a positively correlated variable with values $x_i$ for $i = 1,\ldots,N$ and $\bar{X} = \frac{X}{N}$, with $X = \sum_{i=1}^{N} x_i$ and $c = \left(\frac{N-1}{N}\right)\left(\frac{n}{n-1}\right)\frac{1}{\bar{X}}$. Also, $\bar{r} = \frac{1}{n}\sum_{i\in s} r_i = \frac{1}{n}\sum_{i\in s}\frac{y_i}{x_i}$.

Supposing y is a sensitive variable as say, habitual income tax dodging, drug addictiveness and such qualitative features or loss or gain in gambling, cost of treatment of venereal diseases, fine paid for speed violation in motor driving and such other quantitative variables bearing social stigmas, alternatives are studied here.

To gather responses from sampled persons i in such cases Warner's (1965) RRT enjoins from a box of cards marked A and its complement $A^c$ in proportions $p:(1-p)$, with $(0 < p < 1$ but $p \neq \frac{1}{2})$ asking for a response

$$I_i = \begin{cases} 1 & \text{if i bears A or } A^c \text{ matching his or her true feature A or } A^c \\ 0 & \text{if 'no match' in it.} \end{cases}$$

The true value is

$$y_i = \begin{cases} 1 & \text{if i bears A} \\ 0 & \text{if i bears } A^c. \end{cases}$$

Writing $E_R, V_R$ as operators for expectation, variance for an RRT generically, one gets

$E_R(I_i) = py_i + (1-p)(1-y_i) = (1-p) + (2p-1)y_i$ and
$V_R(I_i) = p(1-p)$, since $I_i = 1,0$ and $y_i = 1,0$ , for Warner's device.

Then, $z_i = \frac{I_i-(1-p)}{2p-1}$ has $E_R(z_i) = y_i$ and $V_R(z_i) = \frac{p(1-p)}{(2p-1)^2} = V_i$, say.
So, $t_1 = N\bar{y}|_{y_i=z_i}$ having $E(t_1) = E_P E_R(t_1) = Y$ is an unbiased estimator for Y and $V(t_1) = V_P(N\bar{y}) + \sum_{i=1}^{N} V_i$, writing $E_P, V_P$ as expectation, variance operators generically in respect of sampling design P and $E = E_P E_R = E_R E_P$ and $V = E_P V_R + V_P E_R = E_R V_P + V_R E_P$ as overall expectation, variance operators (generically) taking $E_P, V_P, E_R, V_R$ as commutative as they generally are in practice.

**Warner's RRT**

For Warner's (1965) RRT an unbiased estimator for $V(t_1)$ is

$$v(t_1) = \left(\frac{1}{n} - \frac{1}{N}\right)\frac{N^2}{n-1}\sum_{i\in s}(z_i - \bar{z})^2 + \frac{N^2}{n}\frac{p(1-p)}{(2p-1)^2}$$

because $\sum_{i=1}^{N} V_i = \frac{N^2}{n}\frac{p(1-p)}{(2p-1)^2}$.

Again, $t_2 = \hat{Y}_{HR}|_{y_i=z_i}$ is the version of Hartley-Ross estimator (generically for an RRT) based on Warner's RRT is

$$v(t_2) = [\hat{Y}_{HR}^2 - \{\frac{N}{n}\sum_{i\in s} y_i^2 + \frac{N(N-1)}{n(n-1)}\sum_{i\neq j} y_i y_j\}]_{y_i=z_i} + \frac{Np(1-p)}{(2p-1)^2}$$

vide Chaudhuri (2010) and Chaudhuri and Samaddar (2022) noting the formula for unbiased estimator of $V_P(\hat{Y}_{HR})$ in the above two references.

**Simmons's URL**

Simmons URL is a version of RRT for which a sampled person i is requested to randomly choose a card from a box of cards marked A and B in proportions $p_1 : (1-p_1)$, $(0 < p_1 < 1)$ and independently and similarly one from a second box of cards marked A and B in proportions $p_2 : (1-p_2)$, $(0 < p_2 < 1, p_1 \neq p_2)$. Here A is a stigmatizing characteristic similar to what is treated by Warner and B is an innocuous feature like preferring music to painting which is quite unrelated to A. As usual, $A^c$ is complementary to A and $B^c$ is complementary to B. Also y is valued $y_i$ which is 1 if i bears A and 0 if i bears $A^c$. Also x is a variable valued for i as $x_i$ which is 1 if i bears B and 0 if i bears $B^c$.

An RR from i is then

$$I_i = \begin{cases} 1 & \text{if for i, A or B matches for the card type drawn and the person feature is A or B} \\ 0 & \text{if there is no match when drawn from the 1st box} \end{cases}$$

and independently and similarly

$$J_i = \begin{cases} 1 & \text{if there is 'match'} \\ 0 & \text{if 'no match', when the 2nd box is used.} \end{cases}$$

Then, $E_R(I_i) = p_1 y_i + (1-p_1)x_i$
$E_R(J_i) = p_2 y_i + (1-p_2)x_i$

Thus, $z_i = \frac{(1-p_2)I_i-(1-p_1)J_i}{(p_1-p_2)}$ yielding $E_R(z_i) = y_i$ and $V_R(z_i) = \frac{(1-p_1)(1-p_2)(p_1+p_2-2p_1p_2)}{(p_1-p_2)^2}(y_i - x_i)^2 = V_i$, say.

But since $V_R(z_i) = E_R(z_i^2) - (E_R(z_i))^2$
$= E_R(z_i^2) - y_i^2$
$= E_R(z_i^2) - y_i$
$= E_R(z_i^2) - E_R(z_i)$
$= E_R(z_i(z_i - 1))$

So, $v_i = z_i(z_i - 1)$ is an unbiased estimator for $V_i = V_R(z_i)$.

Now if SRSWOR in n $(2 \le n < N)$ draws is taken from U of size N and if for such a sample s URL-based RR data are gathered as $z_i$ as above for i in s, then for the expansion estimator $t_1 = \frac{N}{n}\sum_{i\in s} z_i$, $E_R(t_1) = N\bar{y}$, $\bar{y} = \frac{1}{n}\sum_{i\in s} y_i$

and $t_1 = N\bar{z}$ with $\bar{z} = \frac{1}{n}\sum_{i\in s} z_i$ ;

then $E(t_1) = E_P(N\bar{y}) = Y$ and

$$V(t_1) = V_P(N\bar{y}) + E_P\left(\left(\frac{N}{n}\right)^2 \sum_{i\in s} V_i\right)$$
$$= N^2\left(\frac{1}{n} - \frac{1}{N}\right)\frac{1}{N-1}\sum_{i=1}^N (y_i - \bar{Y})^2 + \frac{N}{n}\sum_{i=1}^N V_i \,.$$

To find an unbiased estimator for $V(t_1)$, let us try

$$a = N^2\left(\frac{1}{n} - \frac{1}{N}\right)\frac{1}{n-1}\sum_{i\in s}(z_i - \bar{z})^2 + \frac{N}{n}\sum_{i\in s} v_i$$

and $E(a) = E_P E_R(a)$.

So, this is an unbiased estimator for $V(t_1)$ because

$$E_R(a) = N^2\left(\frac{1}{n} - \frac{1}{N}\right)\frac{1}{n-1}\sum_{i\in s} E_R\left[(z_i - y_i) - (\bar{z} - \bar{y}) + (y_i - \bar{y})\right]^2 + \frac{N}{n}\sum_{i\in s} V_i$$
$$= N^2\frac{\left(\frac{1}{n} - \frac{1}{N}\right)}{(n-1)}\sum_{i\in s}\left[V_i + \frac{\sum_s V_i}{n^2} - 2\frac{V_i}{n} + (y_i - \bar{y})^2\right] + \frac{N}{n}\sum_{i\in s} V_i$$
$$= N^2\frac{\left(\frac{1}{n} - \frac{1}{N}\right)}{(n-1)}\left[\sum_{i\in s} V_i + \frac{\sum_{i\in s} V_i}{n} - 2\frac{\sum_{i\in s} V_i}{n} + \sum_{i\in s}(y_i - \bar{y})^2\right] + \frac{N}{n}\sum_{i\in s} V_i$$
$$= N^2\left(\frac{1}{n} - \frac{1}{N}\right)\left[\frac{1}{n-1}\sum_{i\in s}(y_i - \bar{y})^2\right] + N^2\frac{\left(\frac{1}{n} - \frac{1}{N}\right)}{(n-1)}\left[\left(1 - \frac{1}{n}\right)\sum_{i\in s} V_i + \frac{N}{n}\sum_{i\in s} V_i\right.$$

and

$$E(a) = N^2\left(\frac{1}{n} - \frac{1}{N}\right)\frac{1}{N-1}\sum_{i=1}^N (y_i - \bar{Y})^2 + N^2\left(\frac{1}{n} - \frac{1}{N}\right)\frac{1}{N}\sum_{i=1}^N V_i + \sum_{i=1}^N V_i$$
$$= N^2\left(\frac{1}{n} - \frac{1}{N}\right)\frac{1}{N-1}\sum_{i=1}^N (y_i - \bar{Y})^2 + \sum_{i=1}^N V_i\left(\frac{N}{n}\right) = V(t_1)\,.$$

**Chaudhuri's Device I**

Next we consider Chaudhuri's Device I (ref. Chaudhuri (2011)) to illustrate a case of a quantitative characteristic. Here a sampled person i is requested to randomly choose a card from a box containing numerous cards marked $a_1, a_2, \dots, a_M$ such that their mean is $\frac{1}{M}\sum_1^M a_j = \mu \neq 0$ and the variance is $\sigma^2 = \frac{1}{M-1}\sum_{i=1}^M (a_j - \mu)^2$, both $\mu$ and $\sigma\ (> 0)$ are thus known. The person i is further requested to independently and randomly draw one card from a second box with numerous cards marked $b_1, \dots, b_T$ with mean $\gamma = \frac{1}{T}\sum_{j=1}^T b_j$ and known variance $\varphi^2 = \frac{1}{T-1}\sum_{i=1}^T (b_j - \gamma)^2$.

Then, a requested RR from i is $I_i = a_j y_i + b_k$ , if $a_j$ marked card is drawn from 1st box and $b_k$ marked card from the 2nd box and $y_i$ is the true value for i th person on the stigmatizing quantitative variable $y$.

Then, $E_R(I_i) = \mu y_i + \gamma$ and
$V_R(I_i) = y_i^2\sigma^2 + \varphi^2$.
Then, $z_i = \frac{I_i - \gamma}{\mu}$ has $E_R(z_i) = y_i$
and $V_R(z_i) = y_i^2 \frac{\sigma^2}{\mu^2} + \frac{\varphi^2}{\mu^2} = V_i$, say.

Then $v_i = \frac{z_i^2 \frac{\sigma^2}{\mu^2} + \frac{\varphi^2}{\mu^2}}{1 + \frac{\sigma^2}{\mu^2}} = \frac{z_i^2\sigma^2 + \varphi^2}{\mu^2 + \sigma^2}$ has $E_R(v_i) = V_i$.

Then, $t_1 = \frac{N}{n}\sum_{i\in s} z_i$ has for an SRSWOR of n units from U of size N,
$E(t_1) = E_P\left(\frac{N}{n}\sum_{i\in s} y_i\right) = Y$
$V(t_1) = V_P(N\bar{y}) + E_P(\frac{N^2}{n^2}\sum_{i\in s} V_i)$
$v(t_1) = \frac{N^2(\frac{1}{n} - \frac{1}{N})}{n-1}\sum_{i\in s}(z_i - \bar{z})^2 + \frac{N}{n}\sum_{i\in s} v_i$ .

Again, $t_2 = N[\frac{1}{n}\sum_{i\in s}\frac{z_i}{x_i} + \frac{N-1}{N}\left(\frac{n}{n-1}\right)\frac{1}{\bar{X}}\left(\bar{z} - \bar{x}(\frac{1}{n}\sum_{i\in s}\frac{z_i}{x_i})\right)]\bar{X}$
$\hat{V}(t_2) = t_2^2 - \left[\frac{N}{n}\sum_{i\in s} z_i^2 + \frac{N(N-1)}{n(n-1)}\sum_{i\neq j}\sum_{\in s} z_i z_j\right] + \frac{N}{n}\sum_{i\in s} v_i = v(t_2)$ .

**(PPSWOR, Symmetrized Des Raj estimator) versus (SRSWOR, Expansion estimator)**

Let $s = (i_1, i_2, \dots, i_n)$ be an ordered PPSWOR sample of size n from a finite population chosen with the probability
$P(s) = p_{i_1}\frac{p_{i_2}}{1 - p_{i_1}} \dots \frac{p_{i_n}}{1 - p_{i_1} - p_{i_2} - \dots - p_{i_{n-1}}}$ .
Then, $t_D(z) = \frac{1}{n}(t_1 + t_2 + \dots + t_n)$ is an unbiased estimator of Y where
$t_1 = \frac{z_{i_1}}{p_{i_1}}$,
$t_2 = z_{i_1} + \frac{z_{i_2}}{p_{i_2}}(1 - p_{i_1})$,
...

$t_j = z_{i_i} + z_{i_2} + \cdots + z_{i_{j-1}} + \frac{z_{i_j}}{p_{i_j}}(1 - p_{i_1} - \cdots - p_{i_{j-1}})$ .

Then, $v(t_D) = \frac{1}{2n^2(n-1)} \sum_j \sum_{\neq k} (t_j - t_k)^2 + (\widehat{\sum_{i=1}^N V_i})$

$= \frac{1}{2n^2(n-1)} \sum_j \sum_{\neq k} (t_j - t_k)^2 + t_D(z)|_{z_i = v_i}$ where $v_i = r_{i_j}(r_{i_j} - 1)$.

Now, the expansion estimator $N\bar{y} = \frac{N}{n} \sum_{i \in s} z_i$ has the variance

$$V(N\bar{y}) = N^2 \left(\frac{1}{n} - \frac{1}{N}\right) \frac{1}{N-1} \sum_{i=1}^{N} (y_i - \bar{Y})^2 + \sum_{i=1}^{N} V_i.$$

We know, $V(t_D) = E(t_D^2) - Y^2$ . Thus, $\widehat{Y^2} = t_D^2 - v(t_D)$ .

Also, $\widehat{\sum_{i=1}^N y_i^2} = t_D(z)|_{z_i = y_i^2} = t_D(y^2)$ .

Therefore, $\hat{v}(N\bar{y}) = N^2 \left(\frac{1}{n} - \frac{1}{N}\right) \frac{1}{N-1} \left[t_D(y^2) - \frac{\widehat{Y^2}}{N}\right] = \hat{v}_D$ (say) .

Now, let $s^* = \{i_1, i_2, \dots, i_n\}$ be the unordered sample correspond to the above ordered sample s and $p(s^*) = \sum_{s \to s^*} p(s)$, writing $\sum_{s \to s^*}$ to denote the sum over all possible samples in the set $s^*$.

So, $t_{SD}^* = t_{SD}^*(s^*) = \frac{\sum_{s \to s^*} p(s) t_D(s)}{\sum_{s \to s^*} p(s)}$

$V(t_{SD}^*) = V(t_D(s)) - E(t_D - t_{SD}^*)^2$

$v(t_{SD}^*) = v(t_D) - (t_D - t_{SD}^*)^2$

and $\hat{v}(N\bar{y}) = \left(\frac{1}{n} - \frac{1}{N}\right) \frac{N^2}{N-1} \left[t_{SD}^*(y^2) - \frac{\widehat{Y^2}}{N}\right] = \hat{v}_{SD}$ (say) where $\widehat{Y^2} = (t_{SD}^*)^2 - v(t_{SD}^*)$ .

## 3. Numerical Computation

We use numerical data here borrowed from the comprehensive work by Chaudhuri, Christofides and Saha (2009), Chaudhuri and Christofides (2013). Let $y$ be a sensitive variable having values $y_i$ for $i^{th}$ $(i = 1,2, \dots, N)$ individual. Here $N = 116$ and $Y = \sum_{i=1}^N y_i = \begin{cases} 96 & \text{when y is qualitative variable} \\ 105336 & \text{when y is quantitative variable} \end{cases}$ .For URL model, an innocuous character is also considered.

Table 1 shows the performance of RRTs with Hartley-Ross (1954) estimator based on an SRSWOR sample at hand and Table 2 examines the performances for Symmetrized Des Raj (1956) estimator versus Des Raj estimator. The tables consider different sample sizes with different parameter values to judge the efficacy.

**Table 1: Numerical findings for Hartley-Ross Estimator**

| RR | Sample Size | Parameter values | $t_1$ | $v(t_1)$ | $t_2$ | $v(t_2)$ | $G_{RR,HR}$ |
|---|---|---|---|---|---|---|---|
| Warner | **30** | $\mathbf{p = 0.3}$ | 87 | 1104.7 | 86.64 | 606.60 | 45.09 |
| | **35** | $\mathbf{p = 0.7}$ | 78.71 | 927.59 | 78.16 | 488.50 | 47.34 |
| | **40** | $\mathbf{p = 0.4}$ | 101.5 | 3399.43 | 103.61 | 2509.51 | 26.18 |
| | **45** | $\mathbf{p = 0.8}$ | 73.04 | 259.75 | 72.2 | 56.90 | 78.09 |
| URL | **30** | $(\mathbf{p_1, p_2}) = (\mathbf{0.25, 0.55})$ | 99.18 | 1189.58 | 97.66 | 339.92 | 71.43 |
| | **35** | $(\mathbf{p_1, p_2}) = (\mathbf{0.25, 0.65})$ | 95.87 | 508.89 | 95.07 | 171.55 | 66.29 |
| | **40** | $(\mathbf{p_1, p_2}) = (\mathbf{0.4, 06})$ | 95.7 | 1249.38 | 96.62 | 869.04 | 30.44 |
| | **45** | $(\mathbf{p_1, p_2}) = (\mathbf{0.3, 0.6})$ | 96.58 | 643.37 | 96.09 | 277.11 | 56.93 |
| Chaudhuri' s Device I | **30** | $(\boldsymbol{\mu, \gamma}) = (\mathbf{4.75, 22.5})$ | 106812 | 577235456 | 106319 | 266805775 | 53.78 |
| | **35** | $(\boldsymbol{\mu, \gamma}) = (\mathbf{4.75, 25})$ | 103755 | 459099897 | 103056 | 138061094 | 69.92 |
| | **40** | $(\boldsymbol{\mu, \gamma}) = (\mathbf{5.75, 21.75})$ | 106775 | 291770389 | 106407 | 119696330 | 58.97 |
| | **45** | $(\boldsymbol{\mu, \gamma}) = (\mathbf{5.75, 25})$ | 103435 | 220581427 | 103374 | 133519932 | 39.47 |

**Table 2: Numerical findings for Symmetrized Des Raj estimator**

| RR | Sample Size | Parameter values | $t_D$ | $v(t_D)$ | $\hat{v}_D$ | $t_{SD}$ | $v(t_{SD})$ | $\hat{v}_{SD}$ | $G_{RR,SD}$ |
|---|---|---|---|---|---|---|---|---|---|
| Warner | 5 | $p = 0.3$ | 89.66 | 5685.46 | 5085.64 | 91.83 | 5680.75 | 1525.09 | 73.15 |
| | 6 | $p = 0.7$ | 102.02 | 3946.35 | 3800.04 | 105.41 | 3934.84 | 805.27 | 79.53 |
| | 7 | $p = 0.4$ | 91.85 | 13713.49 | 13690.88 | 89.16 | 13706.23 | 2179.91 | 84.10 |
| | 8 | $p = 0.8$ | 88.39 | 1275.36 | 1123.71 | 86.85 | 1272.97 | 446.62 | 64.92 |
| URL | 5 | $(p_1, p_2) = (0.25, 0.55)$ | 73.95 | 3582.99 | 3252.47 | 72.58 | 3581.12 | 1299.65 | 63.71 |
| | 6 | $(p_1, p_2) = (0.25, 0.65)$ | 92.94 | 2743.22 | 2035.581 | 93.65 | 2742.71 | 770.92 | 71.89 |
| | 7 | $(p_1, p_2) = (0.4, 0.6)$ | 100.43 | 3849.89 | 2374.53 | 98.79 | 3847.23 | 751.03 | 80.48 |
| | 8 | $(p_1, p_2) = (0.3, 0.6)$ | 79.67 | 2153.23 | 1908.82 | 80.03 | 2153.11 | 590.71 | 72.56 |
| Chaudhuri’ s Device I | 5 | $(\mu, \gamma) = (4.75, 22.5)$ | 117754 | 2675231283 | 2485195963 | 115648 | 2670796028 | 1233601645 | 53.81 |
| | 6 | $(\mu, \gamma) = (4.75, 25)$ | 109939 | 2582962403 | 2746710178 | 112238 | 2577677604 | 1570746904 | 39.06 |
| | 7 | $(\mu, \gamma) = (5.75, 21.75)$ | 110619 | 695860959 | 714461335 | 109957 | 695422280 | 246060324 | 64.62 |
| | 8 | $(\mu, \gamma) = (5.75, 25)$ | 111598 | 1360005437 | 1382713124 | 112908 | 1358288910 | 834817064 | 38.54 |

In Table 1, $t_1 = N\bar{z}$ is the unbiased estimator of population total $Y$ when an SRSWOR of $n$ units taken from the finite population and $v(t_1)$ is the unbiased variance estimator. To evaluate Hartley-Ross estimator $t_2$ and its unbiased variance estimator $v(t_2)$, we have taken the RR survey data that already are obtained. Different parameter values are shown in third column of the table. For Warner RRT, we have considered $p$, the portion of cards marked A in the given box, as $0.3, 0.7, 0.4$ and $0.8$ with SRSWOR sample of size $30, 35, 40,$ and $45$ respectively. Similarly, for URL model, $p_i$ is the portion of cards marked A in the $i^{th}$ $(i = 1,2)$ box and we have taken $(p_1, p_2)$ as $(0.25, 0.55), (0.25, 0.65), (0.4, 0.6)$ and $(0.3, 0.6)$ with sample sizes $30, 35, 40,$ and $45$, respectively.

It has been shown there that the unbiased variance estimate $v(t_2)$ is much lesser than $v(t_1)$. The relative gain in efficiency is also computed using $G_{RR,HR} = 100(\frac{v(t_1)-v(t_2)}{v(t_1)})$ and shown in the last column of the mentioned table. Therefore, compared to the strategy $(\text{SRSWOR}, N\bar{y})$, $(\text{SRSWOR}, \text{Hartley} - \text{Ross Ratio estimator})$ is more gainful for randomized response as well as direct response, shown in Chaudhuri and Samaddar (2022).

In Table 2, we present the values of Symmetrized Des Raj estimator $(t_{SD})$ with its unbiased variance estimate $v(t_{SD})$ and the unbiased variance estimator $\hat{v}_{SD}$, obtained by the RR survey data at hand. The values of Des Raj estimator $(t_D)$ and its unbiased variance estimator $v(t_D)$ are also reported in this table. Similar to Table 1, the relative gain in efficiency is computed. Here, the formula is $G_{RR,SD} = 100(\frac{v(t_{SD})-\hat{v}_{SD}}{v(t_{SD})})$.

It can be seen that $\hat{v}_{SD}$ is much lesser than $v(t_{SD})$ and $G_{RR,SD}$ is positive. So, we may conclude that Symmetrized Des Raj estimator combined with PPSWOR outperforms the expansion estimator based on SRSWOR.

## 4. Conclusion

Our presumption is corroborated by real data that a more complex data analytical procedure may in practice lead to be more efficacious than a simpler alternative for direct as well as randomized response based indirect procedure of data gathering.

## References

**[1]** Chaudhuri, A. (2010): Essentials of Survey Sampling. Delhi, India: Prentice Hall of India.

[2] Chaudhuri, A. (2011): Randomized response and indirect questioning techniques in surveys. Boca Raton: CRC Press.

[3] Chaudhuri, A., and Christofides, T. C. (2013): Indirect questioning in sample surveys. Berlin, Germany: Springer Verlag.

[4] Chaudhuri, A., and Samaddar, S. (2022): Estimating the population mean using a complex sampling design dependent on an auxiliary variable. Statistics in Transition New Series , 23 (1), 39-54.

[5] Chaudhuri, A., Christofides, T. C., and Saha, A. (2009): Protection of privacy in efficient application of randomized response techniques. Statistical Methods & Applications , 18, 389-418.

[6] Cochran, W. G. (1953): Sampling Techniques (1st ed.). John Wiley & Sons. N.Y.

[7] Cochran, W. G. (1963): Sampling Techniques (2nd ed.). John Wiley & Sons. N.Y.

[8] Cochran, W. G. (1977): Sampling Techniques. (3rd ed.). John Wiley & Sons. N.Y.

[9] Des Raj. (1956): Some estimators in sampling with varying probabilities without replacement. Journal of American Statistical Association , 51, 269-284.

[10] Greenberg, B. G., Abul-Ela, A. L., Simmons, W. R., and Horvitz, D. G. (1969): The unrelated question randomized response model: theoretical framework. Journal of American Statistical Association , 64, 520-539.

[11] Hartley, H.,and Ross, A. (1954): Unbiased ratio estimators. Nature , 174, 270-271.

[12] Horvitz, D. G., Shah, B. V., and Simmons, W. R. (1967): The unrelated question randomized response model. Proceedings of Social Statistics Section, American Statistical Association, (pp. 65-72).

[13] Rao, J. N. (1961): On the estimate of variance in unequal probability sampling. Annals of the Institute of Statistical Mathematics , 13, 57-60.

[14] Warner, S. L. (1965): Randomized response: a survey technique for eliminating evasive answer bias. Journal of American Statistical Association , 60, 63-69.