SAHA´S RANDOMIZED RESPONSE TECHNIQUE UNDER RANKED SET SAMPLING

Carlos Bouza*, Rene Bracho**, Agustin Santiago[1]***, Jose M. Sautto***
*Universidad de La Habana, Cuba
** Universidad Nacional de Panamá, Panamá
*** Universidad Autónoma de Guerrero, México

ABSTRACT
We will study the use of Ranked Set Sampling (RSS) as alternative to SRSWR design when Saha´s RR is implemented. Estimators of the mean are derived under 3 alternative ranking criteria. The expected variances are obtained and used for comparing their accuracy. The RSS alternatives are superior in terms of having a smaller sampling error than the SRSWR design..

KEYWORDS: Randomized responses, ranked set sampling, scrambling

MSC: 62D05

RESUMEN
Estudiaremos el uso de RSS como una alternativa al diseño SRSWR cuando es usado el método de respuestas aleatorizadas de Saha. Estimadores de la media son derivados bajo 3 criterios 3 alternativos de rankeo. Las varianzas esperadas son obtenidas y se utilizan para copara su exactitud. Las alternativas RSS son superiores en términos de poseer menores error de muestreo que el diseño SRSWR.

PALABRAS CLAVE: Respuestas aleatorizadas, ranked set sampling, enmascaramiento

1. INTRODUCTION

When interviewers try to obtain honest responses, in studies where some sensitive issues are present they usually face difficulties. The work of Warner (1965) on randomized response (RR) models was seminal. The method dealt with the estimation of a proportion of positive responses to a sensitive question in a population, granting that the respondent is not declaring his/her real status. Nowadays RR models may be recommended for both decreasing evasive answer bias and providing privacy protection to the respondents. Is expected that RR will increase response rate and will reduce the response error . The investigations of sensitive question are often encountered in various areas of sample surveys. The sensitive problem refer to the problem that the privacy or interest of individuals and units may leads to having a stigma. Many persons think is inconvenient to reveal publicly the statement. Since the introduction of RR theoretical and practical studies have been developed. Some of the results on RR for quantitive variables are Arnab-Dorffner (2006), Bar-Lev et al. [2004], Gjestvang-Singh, S. (2006), Odumade-Singh (2010), Ryu et al. [2005], Singh-Chen (2009), Bouza, C. N. (2009).

Saha (2007) proposed a RR (RRS) procedure that provides the estimation of the mean of a sensitive quantitative variable. The statistical model was developed considering inferences based on homogeneous linear unbiased estimators and the use of simple random sampling with replacement (SRSWR).

McIntyre (1952) proposed a sampling method that is currently known as ranked set sampling (RSS). In this method the sampling units are partitioned into small subsets of the same size. The units of each subset are ranked with respect to the characteristic of interest Y using a concomitant variable X. Ranking is supposed to be easily made and at a low cost. RSS is to be considered as an alternative sample design, which provide gains in accuracy with respect to SRSWR. This theme is of growing importance. Recent papers as Al-Saleh, M.F. and Al-Omari, A.I. (2002, Al-Omari-Jaber (2008), Al-Nasser (2007), Kadilar –Unyazici- Cingi, (2009) may be mentioned. See Patil (2002), Chen et al (2004) for a detailed discussion on RSS.

[1] Corresponding Author: asantiago@uagro.mx

We will study the use of RSS as alternative to SRSWR design when Saha´s RR is implemented. Estimators of the mean are derived under 3 alternative ranking criteria. The expected variances are obtained and used for comparing their accuracy. The RSS alternatives are superior in terms of having a smaller sampling error than the SRSWR design..
Section 2 presents the model when Saha´s RR is used under SRSWR design. Section 3 presents the basic ideas of RSS. 3 ranking procedures are proposed and the corresponding estimators of the population mean and the expected variances are obtained. The behavior of the estimators are compared with the SRSWR model and among them.

2. SAHA´S PROPOSAL UNDER SRSWR.

Let $U = (1,\dots, i,\dots, N)$ denote a finite population of N individuals. The belonging to a certain group, say A, is stigmatizing. Take Y_i as the value of a stigmatizing variable Y. An estimate of its population mean $\mu_Y = \frac{\sum_{i=1}^{N} Y_i}{N}$ is required. The surveyor fixes two sets

$$\boldsymbol{Z} = \{\boldsymbol{Z_t} > 0; t = 1, \dots, T\}, \boldsymbol{V} = \{\boldsymbol{V_h} > 0; h = 1, \dots, H\}$$

They are independent sets of random numbers and both are independent of the identification to A, measured by the Bernoulli variable $I_A(i)$, and Y_i . As the surveyor fixes $\boldsymbol{Z}$ and $\boldsymbol{V}$ them are known their means and variances.
The RR proposed by Saha (2007) may be described as follows:

RR procedure of Saha (RRS)

Each sampled person is presented with ***Z*** and ***V***

Step 1. The interviewed selects randomly a number v* from ***V***.
Step 2. The selected number is added to the true value of Y_i.
Step 3 The interviewed selects randomly a number z* from ***Z.***
Step 4. Compute the scrambled report $B_i = z_i^*(Y_i + v_i^*)$□

Due to the independence of the involved variable the expectation under the RR procedure is

$$E_S(B_i) = Y_i\mu_Z + \mu_Z\mu_V$$

The variance is given by

$$\begin{aligned} V_S(B_i) &= Y_i^2\sigma_Z^2 + E_S(z_i^2 v_i^2) - \mu_V^2\mu_Z^2 + 2\left(Y_i \ \mu_V\, E_S(z_i^2) - Y_i \ \mu_V\, \mu_Z^2\right) \\ &= Y_i^2\sigma_Z^2 + \sigma_V^2\sigma_Z^2 + \mu_V^2\sigma_Z^2 + \mu_Z^2\sigma_V^2 + 2Y_i\mu_V^2\sigma_Z^2 \end{aligned}$$

These results sustain the validity of the following theorem:

Theorem (Saha (2007)): Select randomly a respondent and use RRS for obtaining the scrambled response B_i. Then

1. $R_i = \frac{B_i}{\mu_Z} - \mu_V$ is model unbiased for the sensitive variable Y_i
2. $V_S(R_i) = \alpha Y_i^2 + \beta Y_i + \gamma;\ \alpha = \frac{\sigma_Z^2}{\mu_Z^2},\ \beta = \frac{2\sigma_Z^2\mu_V^2}{\mu_Z^2},\ \ \gamma = \frac{\sigma_Z^2(\mu_V^2+\sigma_V^2)}{\mu_Z^2} + \sigma_V^2$.□

Therefore if simple random sampling with replacement (SRSWR) is used, an unbiased estimator for the mean of Y is

$$\hat{\mu}_Y = \frac{\sum_{i=1}^{n} R_i}{n}$$

The variance is

$$E_d V_S(\hat{\mu}_Y) = \frac{V_0}{n} = \frac{\alpha\left(\sigma_y^2 + \mu_Y^2\right) + \beta\mu_Y + \gamma}{n} = V_{srswr}$$

3. GT UNDER RSS

RSS involves randomly selecting units from the population sing SRSWR. These units are randomly allocated into m sets, each of size m. The n units of each sample may be are ranked with respect to the variable of interest. From the first set of m units, the smallest unit is to be interviewed. From the second set of units, the

second smallest unit is interviewed. The process is continued until from the *m-th* set of *m* units the largest unit should respond. The procedure is repeated $r\geq 1$ times for obtaining $n=rm$ responses.
We analyze the use of RSS when the RR proposed by GT is used for obtaining a report. The selection procedure can be sketched as follows

Procedure RSS

While *t<m, h<r* do

Select a sample s(t,h) of size m independently from *U* using srswr.

Each unit in s(t,h) is ranked and the order statistics (os) O*(1:t)h ,...., O(m:t)h* are determined.

END

If the variable of interest is O, the ranked set sample in a cycle h is composed by the elements in the diagonal of the matrix.

O(1:1)h	···O(i:1)h···	O(m:1)h
O(1:2)h	···**O(i:2)**h···	O(m:2)h
:	:::::::::::::	:
O(1:m)h	···O(i:m)h···	**O(m:m)h**

As we are using order statistics $E[O_{(ii)t}]=\mu_{(i)}$, $\sum_{i=1}^{m}$ and $(\mu_{(i)}-\mu)= \sum_{i=1}^{m} \Delta_{(i)} =0$. Therefore is unbiased the RSS-estimator of the population mean $\bar{O}_{rss} = \frac{1}{n}\sum_{i=1}^{m} O_{(i)i}$. A basic relationship in RSS was obtained by Stokes (1980)

$$V(\bar{O}_{rss}) = \frac{1}{m^2}\sum_{i=1}^{m} \sigma_{(i)}^2 = \frac{\sigma_O^2}{m} - \frac{1}{m^2}\sum_{i=1}^{m} \Delta_{(i)}^2$$

Then using RSS produces more accurate estimation than applying SRSWR and using the arithmetic mean as estimator. See Chen-Bai-Sinha (2004) and Bouza, C. N. (2005) for a detailed discussion.
Let us consider the behavior of the use of GT.The usual report is $B_i = z_i^*(Y_i + v_i^*)$.

Case 1. Ranking based on Y

The surveyor has some information that allows ranking the selected units in the set s(t)h. For example a physician may be able to rank, after a gynecological analysis, the number of abortions of women. Then the report is modeled by

$$B_{(i:i)h} = z_i^*(Y_{(i:i)h} + v_i^*)$$

Now we have

$$E_S\left(B_{(i:i)h}\right) = Y_{(i:i)h}\mu_Z + \mu_Z\mu_V$$

$$V_S\left(B_{(i:i)h}\right) = Y_{(i:i)h}^2 \quad \sigma_Z^2 + \sigma_V^2\sigma_Z^2 + \mu_V^2\sigma_Z^2 + \mu_Z^2\sigma_V^2 + 2Y_{(i:i)h}\mu_V^2\sigma_Z^2$$

The unscrambled variable is

$$R_{(i:i)h} = \frac{B_{(i:i)h}}{\mu_Z} - \mu_V$$

Note that, as the sampling design d is RSS,

$$E_d\left(R_{(i:i)h}\right) = \frac{E_d(Y_{(i:i)h})\mu_Z + \mu_Z\mu_V}{\mu_Z} - \mu_V = \mu_{Y_{(i)}}.$$

Take the estimator

$$\bar{R}_{(rss)} = \frac{\sum_{h=1}^{r} \quad \sum_{t=1}^{m} R_{(i:i)h}}{rm}$$

Its unbiasedness follows from the fact that

$$E_d\left(\bar{R}_{(rss)}\right) = \frac{\sum_{h=1}^{r} \quad \sum_{t=1}^{m} E_d\left(R_{(i:i)h}\right)}{rm} = \mu_Y$$

The design expectation of the model variance is

$$E_d V_S(B_{(i:i)h}) = \left(\sigma^2_{Y\,(i)} + \mu^2_{Y_{(i)}}\right)\sigma_Z^2 + \sigma_V^2\sigma_Z^2 + \mu_V^2\sigma_Z^2 + \mu_Z^2\sigma_V^2 + 2\mu_{Y_{(i)}}\mu_V^2\sigma_Z^2$$

$$V_S(R_i) = \alpha Y_i^2 + \beta Y_i + \gamma;\ \alpha = \frac{\sigma_Z^2}{\mu_Z^2}, \qquad \beta = \frac{2\sigma_Z^2\mu_V^2}{\mu_Z^2}, \quad \gamma = \frac{\sigma_Z^2(\mu_V^2 + \sigma_V^2)}{\mu_Z^2} + \sigma_V^2$$

Therefore

$$E_d V_S(\bar{R}_{(rss)}) = \alpha\left(\frac{\sigma_Y^2}{n} - \frac{\sum_{i=1}^m \Delta^2_{Y_{(i)}}}{nm} + \frac{\sum_{t=1}^m \mu^2_{Y_{(i)}}}{nm}\right) + \beta\frac{\mu_Y}{n} + \frac{\gamma}{n} = V_{RSS}$$

As the ranking is made on Y, $\alpha = \frac{\sigma_Z^2}{\mu_Z^2}, \quad \beta = \frac{2\sigma_Z^2\mu_V^2}{\mu_Z^2}, \quad \gamma = \frac{\sigma_Z^2(\mu_V^2 + \sigma_V^2)}{\mu_Z^2} + \sigma_V^2$

Note that

$$\mu_Y^2 = \left(\frac{\sum_{t=1}^m \mu_{Y_{(i)}}}{m}\right)^2 = \frac{\sum_{i=1}^m \mu^2_{Y_{(i)}} + \sum_{i\neq j}\mu_{Y_{(i)}}\mu_{Y_{(j)}}}{m^2}$$

Let us compare $E_d V_S(\bar{R}_{(rss)})$ with $E_d V_S(\hat{\mu}_Y)$. We have that

$$\frac{E_d V_S(\bar{R}_{(rss)}) - E_d V_S(\hat{\mu}_Y)}{\alpha} = \left(\frac{m-1}{nm^2}\right)\sum_{t=1}^m \mu^2_{Y_{(i)}} - \frac{\sum_{i=1}^m \Delta^2_{Y_{(i)}}}{nm} - \frac{\sum_{i\neq j}\mu_{Y_{(i)}}\mu_{Y_{(j)}}}{nm^2}$$

Commonly in RSS, see Chen-Bai-Sinha (2004), m is small and hence having a large n is valid considering that $\frac{m-1}{nm^2} \cong 0$. If Y>0 the last term of this expression represents a gain in accuracy due to the use of RSS.

Case 2. Ranking based on Z

The respondents communicate among them the value obtained of Z and the ranking is reported. This information is un-sensitive. Then the surveyor has the ranking and the selects the unitsto respond. Now the report is modeled by

$$B_{Z(i:i)h} = z^*_{(i:i)h}(Y_i + v_i^*)$$

We have

$$E_S(B_{Z(i:i)h}) = Y_i\mu_{Z_{(i)}} + \mu_{Z_{(i)}}\mu_V$$

$$V_S(B_{Z(i:i)h}) = Y_i^2 \quad \sigma^2_{Z_{(i)}} + \sigma_V^2\sigma^2_{Z_{(i)}} + \mu_V^2\sigma^2_{Z_{(i)}} + \mu^2_{Z_{(i)}}\sigma_V^2 + 2Y_i\mu_V^2\sigma^2_{Z_{(i)}}$$

We take as unscrambled variable

$$R_{Z(i:i)h} = \frac{B_{Z(i:i)h}}{\mu_{Z_{(i)}}} - \mu_V$$

As $E_S(R_{Z(i:i)h}) = Y_i$,

$$E_S(\bar{R}_{Z(rss)}) = E_S\left(\frac{\sum_{h=1}^r \quad \sum_{t=1}^m R_{Z(i:i)h}}{rm}\right) = \mu_Y$$

and

$$E_d V_S(R_{Z(i:i)h}) = \frac{\sigma^2_{Z_{(i)}} \quad (\mu_Y^2 \quad + \sigma_Y^2)}{\mu^2_{Z_{(i)}}} + \frac{\sigma_V^2\sigma^2_{Z_{(i)}} + \mu_V^2\sigma^2_{Z_{(i)}} + \mu^2_{Z_{(i)}}\sigma_V^2 + 2\mu_Y \quad \mu_V^2\sigma^2_{Z_{(i)}}}{\mu^2_{Z_{(i)}}} =$$
$$= \alpha_{(i)}(\sigma_Y^2 + \mu_Y^2 \quad) + \beta_{(i)}\mu_Y \quad + \gamma_{(i)}$$

Where

$$\alpha_{(i)}=\frac{\sigma^2_{Z_{(i)}}}{\mu^2_{Z_{(i)}}},\ \ \beta_{(i)}=2\frac{\sigma^2_{Z_{(i)}}\ \ \mu^2_V}{\mu^2_{Z_{(i)}}},\quad \gamma_{(i)}=\frac{\sigma^2_V\sigma^2_{Z_{(i)}}+\mu^2_V\sigma^2_{Z_{(i)}}}{\mu^2_{Z_{(i)}}}+\sigma^2_V.$$

As a result

$$E_dV_S\left(\bar{R}_{Z(rss)}\right)=\frac{\left(\sigma^2_Y+\mu^2_Y\ \right)}{nm}\sum_{i=1}^{m}\frac{\sigma^2_{Z_{(i)}}}{\mu^2_{Z_{(i)}}}+\frac{2\mu_Y\mu^2_V}{mn}\sum_{i=1}^{m}\frac{\sigma^2_{Z_{(i)}}}{\mu^2_{Z_{(i)}}}+\frac{\left(\sigma^2_V+\mu^2_V\right)}{mn}\sum_{i=1}^{m}\frac{\sigma^2_{Z_{(i)}}}{\mu^2_{Z_{(i)}}}+\frac{\sigma^2_V}{n}=V_{ZRSS}$$

Let us compare this procedure with the use of SRSWR

$$\begin{aligned}&E_dV_S(\hat{\mu}_Y)-E_dV_S\left(\bar{R}_{Z(rss)}\right)\\&\qquad=\frac{\left(\frac{\sigma^2_Z}{\mu^2_Z}-\sum_{i=1}^{m}\frac{\sigma^2_{Z_{(i)}}}{m\mu^2_{Z_{(i)}}}\right)\left(\sigma^2_y+\mu^2_Y\right)}{n}+\frac{2\mu_Y\mu^2_V}{n}\left(\frac{\sigma^2_Z}{\mu^2_Z}-\sum_{i=1}^{m}\frac{\sigma^2_{Z_{(i)}}}{m\mu^2_{Z_{(i)}}}\right)\\&\qquad+\left(\frac{\mu^2_V+\sigma^2_V}{n\mu^2_Z}+\sigma^2_V\right)\left(\frac{\sigma^2_Z}{\mu^2_Z}-\sum_{i=1}^{m}\frac{\sigma^2_{Z_{(i)}}}{m\mu^2_{Z_{(i)}}}\right)\end{aligned}$$

It is acceptable that $\frac{\sigma^2_Z}{\mu^2_Z}\le\frac{\sigma^2_{Z_{(i)}}}{\mu^2_{Z_{(i)}}}$ therefore we accept that $\frac{\sigma^2_Z}{\mu^2_Z}-\sum_{i=1}^{m}\frac{\sigma^2_{Z_{(i)}}}{m\mu^2_{Z_{(i)}}}\ge 0$. Then this difference is positive, determining that ranking on Z determines a RSS procedure that has a smaller error than the SRSWR one. The quantitative gain in accuracy due to this RSS model is measured by with respect of a ranking using information on Y is

$$\begin{aligned}&E_dV_S\left(\bar{R}_{(rss)}\right)-E_dV_S\left(\bar{R}_{Z(rss)}\right)\\&\qquad=\left(\frac{\sigma^2_Z}{\mu^2_Z}-\sum_{i=1}^{m}\frac{\sigma^2_{Z_{(i)}}}{m\mu^2_{Z_{(i)}}}\right)\frac{\sigma^2_Y}{n}+\left(\frac{\sigma^2_Z}{\mu^2_Z}\frac{\sum_{t=1}^{m}\mu^2_{Y_{(i)}}}{nm}-\frac{\mu^2_Y}{nm}\sum_{i=1}^{m}\frac{\sigma^2_{Z_{(i)}}}{\mu^2_{Z_{(i)}}}\right)+\frac{2\mu_Y\mu^2_V}{n}\left(\frac{\sigma^2_Z}{\mu^2_Z}-\sum_{i=1}^{m}\frac{\sigma^2_{Z_{(i)}}}{m\mu^2_{Z_{(i)}}}\right)\\&\qquad+\frac{\left(\sigma^2_V+\mu^2_V\right)}{mn}\left(\frac{\sigma^2_Z}{\mu^2_Z}-\sum_{i=1}^{m}\frac{\sigma^2_{Z_{(i)}}}{m\mu^2_{Z_{(i)}}}\right)=A+B+C+D\end{aligned}$$

Clearly A, C and D are non positive. The value of B is to be determined in each particular case. The surveyor is able to determine the set Z is such a way that it be negative too ensuring the preference for dealing with $\bar{R}_{(rss)}$.

Case 3. Ranking based on V

The respondents may communicate among them the value obtained of V and then ranking themselves. The surveyor selects the persons to give their report. The report obtained for a person is

$$B_{V(i:i)h}=\ z^*_i(\ Y_i+v^*_{(i:i)h})$$

Now the RRS expectation and variance of the report are

$$E_S\left(B_{V(i:i)h}\right)=Y_i\mu_Z\quad+\mu_{V_{(i)}}\mu_Z$$

$$V_S\left(B_{V(i:i)h}\right)=Y^2_i\quad\sigma^2_Z\quad+\sigma^2_Z\sigma^2_{V_{(i)}}+\mu_{V_{(i)}}\sigma^2_Z\quad+\mu^2_Z\ \ \sigma^2_{V_{(i)}}+2Y_i\mu^2_{V_{(i)}}\sigma^2_Z$$

Unscrambling we have

$$R_{V(i:i)h}=\frac{B_{V(i:i)h}}{\mu_Z}-\mu_{V_{(i)}}$$

Its model expected value also is $E_S\left(R_{V(i:i)h}\right) = Y_i$. Hence it is also unbiased for the sensitive variable and

$$E_S(\bar{R}_{V(rss)}) = E_S\left(\frac{\sum_{h=1}^{r} \quad \sum_{t=1}^{m} R_{V(i:i)h}}{rm}\right) = \mu_Y$$

The expectation of the model variance is given by

$$E_d V_S\left(R_{V(i:i)h}\right) = \frac{\sigma_Z^2 \quad \left(\mu_Y^2 \quad + \sigma_Y^2\right)}{\mu_Z^2} + \frac{\sigma_Z^2 \sigma_{V_{(i)}}^2 + \mu_{V_{(i)}} \sigma_Z^2 \quad + \mu_Z^2 \quad \sigma_{V_{(i)}}^2 + 2Y_i \mu_{V_{(i)}}^2 \sigma_Z^2}{\mu_Z^2} =$$
$$= A_{(i)}\left(\sigma_Y^2 + \mu_Y^2 \quad\right) + B_{(i)} \mu_Y \quad + C_{(i)}$$

Where

$$A_{(i)} = \frac{\sigma_Z^2}{\mu_Z^2}, \; B_{(i)} = 2\frac{\sigma_Z^2 \quad \mu_{V_{(i)}}^2}{\mu_Z^2}, \quad C_{(i)} = \frac{\sigma_{V_{(i)}}^2 \sigma_Z^2 \quad + \mu_{V_{(i)}}^2 \sigma_Z^2}{\mu_Z^2} + \sigma_{V_{(i)}}^2.$$

The expected error is derived as

$$E_d V_S\left(\bar{R}_{V(rss)}\right) = \frac{\left(\sigma_Y^2 + \mu_Y^2 \quad\right)}{n} \frac{\sigma_Z^2}{\mu_Z^2} + 2\frac{\sigma_Z^2 \quad \sum_{t=1}^{m} \mu_{V_{(i)}}^2}{mn\mu_Z^2} \mu_Y \quad + \frac{\sigma_Z^2 \quad \sum_{t=1}^{m} \quad \sigma_{V_{(i)}}^2 + \mu_{V_{(i)}}^2}{mn\mu_Z^2} + \frac{\sum_{t=1}^{m} \quad \sigma_{V_{(i)}}^2}{mn}$$
$$= V_{VRSS}$$

Let us compare it with the error when using SRSWR. We have that

$$E_d V_S(\hat{\mu}_Y) = \frac{V_0}{n} = \frac{\alpha\left(\sigma_y^2 + \mu_Y^2\right) + \beta \mu_Y + \gamma}{n}$$

$$\alpha = \frac{\sigma_Z^2}{\mu_Z^2}, \qquad \beta = \frac{2\sigma_Z^2 \mu_V^2}{\mu_Z^2}, \quad \gamma = \frac{\sigma_Z^2(\mu_V^2 + \sigma_V^2)}{\mu_Z^2} + \sigma_V^2$$

Hence

$$E_d V_S(\hat{\mu}_Y) - E_d V_S\left(\bar{R}_{V(rss)}\right)$$
$$= \frac{2\mu_Y \sigma_Z^2}{n\mu_Z^2}\left(\mu_V^2 - \sum_{i=1}^{m} \frac{\mu_{V_{(i)}}^2}{m}\right) + \left(\left(\mu_{V-}^2 \sum_{i=1}^{m} \frac{\mu_{V_{(i)}}^2}{m}\right) + \left(\sigma_V^2 - \sum_{i=1}^{m} \frac{\sigma_{V_{(i)}}^2}{m}\right)\right)\left(\frac{\sigma_Z^2}{n\mu_Z^2}\right)$$
$$+ \left(\frac{\sigma_V^2}{n} - \frac{\sum_{t=1}^{m} \quad \sigma_{V_{(i)}}^2}{mn}\right) = a + (b + c) + d$$

Note that

$$\mu_V^2 - \sum_{i=1}^{m} \frac{\mu_{V_{(i)}}^2}{m} = \sum_{i=1}^{m} \frac{\left(\mu_V - \mu_{V_{(i)}}\right)\left(\mu_V + \mu_{V_{(i)}}\right)}{m} > \left(\mu_V \; + \mu_{V_{(1)}}\right) \sum_{i=1}^{m} \frac{\left(\mu_V - \mu_{V_{(i)}}\right)}{m} = 0 \; (1)$$

As

$$\sum_{i=1}^{m} \frac{\sigma_V^2 - \sigma_{V_{(i)}}^2}{m} = \sum_{i=1}^{m} \frac{\Delta_{V_{(i)}}^2}{m} \; (2)$$

We have that, as $a > 0, b > 0, c > 0$ and $d > 0$, this RSS model is more accurate than SRSWR.

The comparison with the raking on Y leads to

$$E_dV_S(\bar{R}_{(rss)})-E_dV_S(\bar{R}_{V(rss)})=\frac{\sigma_Z^2}{n\mu_Z^2}\left(\frac{\sum_{t=1}^{m}\mu_{Y(i)}^2}{m}-\mu_Y^2\right)+\frac{2\mu_Y\ \ \sigma_Z^2}{n\mu_Z^2}\left(\mu_V^2-\frac{\sum_{t=1}^{m}\mu_{V(i)}^2}{m}\right)+\frac{\sigma_Z^2}{n\mu_Z^2}\left(\left(\sigma_V^2-\frac{\sum_{t=1}^{m}\ \ \sigma_{V(i)}^2}{m}\right)+\left(\mu_V^2-\frac{\sum_{t=1}^{m}\mu_{V(i)}^2}{m}\right)\right)+\frac{1}{n}\left(\sigma_V^2-\frac{\sum_{t=1}^{m}\ \ \sigma_{V(i)}^2}{m}\right)=a+b+(c+d)+e$$

(1) implies that b and d are non negative while (2) implies that c and e are also non negative. Then, under an adequate selection of the parameters of the distribution of V is granted that the RSS model obtained ranking V is more accurate than ranking on Y.
When this model overcome the model based on a ranking on Y it clearly is better than the ranking in Z.

3. A NUMERICAL EXPERIMENT

We developed an experiment for evaluating the behavior of the methods and the valuation of the interviews. We took a population of 310 persons. Some of them had a number of sanitary infractions in their restaurants and cafeterias. We defined Y=number of punishments in the last semester. That value was known as well as X= value of the fines in $ during the previous year. For each person we generated independently two variable with distribution uniform in (0,100). Then each person j was attached with (X_j, Y_j, V_j, Z_j). We fixed m=2, 3, 5 and n=30. The we computed the scrambled variable using the ranking determined by X, V and Z. The relative efficiencies

$$E_{pq}=\frac{V_p}{V_q}, p, q=SRSWR, RSS, Z(rss)\ and\ V(rss)$$

Were computed. The results appear in the following table

Table 1. $E_{pq}=\frac{V_p}{V_q},\ W\sim U(0{,}100), W=Z, V$

p	SRSWR	RSS	Z(rss)	V(rss)
SRSWR	1	1,852	1,943	2,114
RSS		1	0,893	1,139
Z(rss)			1	0,717
V(rss)				1

Note how good is the use of V.
We inquired the possible interviewed which was the more trusted method of maintaining the confidence of their status. Each of them was evaluated in a Likert scale 1(worst)-5(best). See the results in the table 2. They support that ranking in V is preferred for ranking by the interwed.

p	Mean rank	Variance of the ranks	$\frac{\sqrt{\text{Variance of the ranks}}}{\text{Mean rank}}$
SRSWR	1,12	7,13	2,381
RSS	3,25	10,18	0,970
Z(rss)	4,22	14,42	0,924
V(rss)	4,81	5,39	0,478

More studies on the confidence of the respondents is needed for establishing whether the preference for ranking, once information on V is shared by them for providing ranks, suggests it current use.

RECEIVED: AUGUST, 2017
REVISED: NOVEMBER, 2017

REFERENCES

.

[1]. AL-OMARI, A.I.. and JABER, K. (2008): Percentile double ranked set sampling**. Journal of Mathematics and Statistics,** 4, 60-64.

[2]. AL-SALEH, M.F. and AL-OMARI, A.I. (2002): Multistage ranked set sampling, **Journal of Statistical Planning and Inference,** 102, 273-286.

[3]. ARNAB, R. and DORFFNER, G. (2006): Randomized response technique for complex survey designs. **Statistical Papers** 48, 131–141, 2006.

[4]. BAR-LEV, S. K., E. BOBOVITCH and B. (2004): A note on randomized response models for quantitative data. **Metrika** , 60, 255--260

[5]. BOUZA, C. N. (2005): Sampling using ranked sets: concepts, results and perspectives**. Revista Investigación Operacional** , 26, 275-293.

[6]. BOUZA, C. N. (2009): Ranked set sampling and randomized response procedures for estimating the mean of a sensitive quantitative character. **Metrik**a, 70, 267–277.

[7]. BOUZA, C.N, and L. SINGH (2009): Optimal aspects of a model based randomized responses procedure under unequal selection of insensitive variables. **Revista Investigación Operacional** , 30, 234-243.

[8]. BOUZA-HERRERA, C. N. and D. COVARRUBIAS-MELGAR (2012): Ranked Set Sampling and Rao-Hartley-Cochran Strategies Using the Randomized Procedure of Eichhorn-Hayre. **International Journal of Mathematics and Computation,** 15, 31-37.

[9]. CHEN, Z., Z. BAI and K. B. SINHA (2004): **Ranked Set Sampling**. Springer, Berlin.

[10]. GJESTVANg, C. R. and Singh, S. (2006): A new randomized response model**. Journal of the Royal Statistical Society , B**, 68 :523-530

[11]. KADILAR C: , Y. UNYAZICI, and H. CING, (2009): Ratio estimator for the population mean using ranked set sampling. **Statist. Papers** 50, 301–309.

[12]. MCINTYRE, G.A. (1952): A method of unbiased selective sampling using ranked sets. **Australian J. Agricultural Research**. 3, 385-390.

[13]. ODUMADE, O. and SINGH, S. (2010): An alternative to the Bar-Lev, Bobovitch and Boukai randomized response model. **Sociological Methods & Research**, 39, 206-221.

[14]. PATIL, G.P (2002): Ranked set sampling. In **Encyclopedia of Enviromentrics**.(A.H. El- Shaarawi and W.W. Pieegoshed, (Editors): Vol. 3 2684-2690. Wiley , Chichester.

[15]. RYU, J.-B., KIM, J.-M., HEO, T.Y. and PARK, C. G. (2005): On stratified Randomized response sampling, **Model Assisted Statistics and Application**, 1, 31-36.

[16]. SINGH, S. and CHEN, C.C. (2009): Utilization of higher order moments of scrambling variables in randomized response sampling. **Journal of Statistical Planning and Inference**,139, 3377-3380.

[17]. STOKES S. L. (1980): Estimation of the variance using judgment ordered samples. **Biometrics**, 36, 35-42.

[18]. WARNER, S. L (1965): Randomized Response : a Survey Technique for Elimination Evasive Answer Bias. **J of the American Statistical Association,** 60, 63-69