Available online at www.isas.org.in/jisas
JOURNAL OF THE INDIAN SOCIETY OF
AGRICULTURAL STATISTICS 75(3) 2021 205–213

Rescaling Bootstrap Variance Estimation of Level-0 Ranked Set Sampling under Finite Population Framework

Vinaykumar L.N., Tauqueer Ahmad, Anil Rai and Ankur Biswas

ICAR-Indian Agricultural Statistics Research Institute, New Delhi

Received 14 June 2021; Revised 29 January 2022; Accepted 07 February 2022

SUMMARY

McIntyre (1952) introduced Ranked Set Sampling (RSS) to advance upon Simple Random Sampling (SRS) for circumstances where any preliminary ranking of sampled units is possible for variable of interest using visual inspection or some other means without physically measuring the units. Further, the RSS was classified into three sampling protocols named as Level-0, Level-1 and Level-2 (Deshpande *et al.*, 2006). The Level-0 sampling protocol of RSS is considered in this article. Estimating the variance of the Level-0 RSS estimator under the finite population framework was found to be cumbersome. In this article, two distinct rescaling bootstrap with replacement methods known as Strata-based rescaling bootstrap with-replacement (SRBWR) method and Cluster-based rescaling bootstrap with-replacement (CRBWR) method have been proposed to unbiasedly estimate the variance of Level-0 RSS estimator of finite population mean. Rescaling factors are obtained for both the proposed methods to estimate the variance of the Level-0 RSS estimator unbiasedly. The results of the simulation analysis, together with real data application support, proposed methods are capable of estimating the variance of the Level-0 RSS estimator almost unbiasedly. The developed SRBWR method performs better than the CRBWR method considering Relative stability (RS) and percentage Relative Bias (%RB) for various combinations of set size (m) and several cycles (r).

Keywords: Ranks; Strata; Cluster; Level-0 Ranked Set Sampling; Rescaling bootstrap; Resampling.

1. INTRODUCTION

Ranked Set Sampling (RSS) first proposed by McIntyre (1952) to obtain sample observations from the population that are expected to be more representative of the population than the equal number of sample observations got through Simple Random Sampling (SRS) (Wolfe, 2012). The RSS is highly effective if the precise measurement of the study variable is either more expensive or problematic in terms of resources, labour, or time, but samples of units can be ranked precisely with reduced cost based on visual inspection or some other basic approach which does not require measurement. McIntyre's work was inspired by the issue of estimating average agricultural forage yields. Making precise yield quantification requires the harvesting of crops, but a specialist can perfectly rank the yields in a small set of fields based on a visual inspection and select the less number of fields for actual quantification of yield by costly measures. RSS has also been used successfully in agricultural studies (Halls and Dell 1985; Cobby *et al.* 1985; Chen *et al.* 2004; Bocci *et al.* 2010), ecological and environmental studies (Martin *et al.* 1980; Al-Saleh and Zheng, 2002; Ozturk *et al.* 2005), medical studies (Samawi and Al-Sagheer 2001; Nahhas *et al.* 2002), etc. Halls and Dell (1966) given the term Ranked Set Sampling (RSS), Takahasi and Futatsuya (1968) given the mathematical foundation for RSS and independently by Dell (1969).The mathematical foundation for RSS developed through obtaining unbiased estimation of the population mean based on the sample stratified by means of ordering. The concomitant variables are used for ranking in RSS design and Stokes (1977) referred it as "Ranked Set Sampling with concomitant variables". See Patil *et al.* (1994),Wolfe (2012) and Arnab (2017) for an historical review of the theory, methods, and applications of ranked set sampling. Most of the research work on RSS has been aimed at estimating unknown parameters especially population mean in the context of an infinite population (Wolfe, 2012; Biswas

Corresponding author: Ankur Biswas
E-mail address: ankur.biswas@icar.gov.in

et al., 2020). Though, under the finite population, the unbiased population mean estimator and its variance expression for RSS design (i.e. Level-2 RSS design) have been developed by Patil *et al.* (1995). Several works have been performed in the RSS within a finite population framework (Takahasi and Futatsuya 1988, 1998; Krishna, 2002, Rai and Krishna, 2013, Kankure and Rai, 2008).

The RSS was classified into three sampling protocols named as Level-0, Level-1 and Level-2 under finite population framework by Deshpande *et al.* (2006). The Level-0 produces the whole RSS sample by replacing all units back into the population before to selection of the next set, whereas, in case of Level-1, sampling units selected for actual measurement is not replaced, but the other units used for ranking are replaced prior to selecting the next set. However, Level-2 the RSS sample is selected by replacing none of the sampling units back into population, prior to selection. Several attempts were also made to explore the possibilities of inclusion probabilities and respective Horvitz-Thompson (HT) type estimators of finite population parameters under RSS design (Al-Saleh and Samawi, 2007; Ozdemir and Gokpinar 2007, 2008; Gokpinar and Ozdemir, 2010, 2012, 2014; Jozani and Johnson, 2011, 2012; Frey, 2011; Ozturk and Jozani, 2014; Ozturk, 2014, 2016a, 2016b).

1.1 Level-0 sampling protocol of RSS design

The RSS sample obtained through Level-0 sampling protocol by first, draw an SRSWOR sample of size *m* units (i.e. set) from the finite population Ω and rank the sampled units based on characteristics of interest or any other method not requiring actual quantification i.e. visual inspection or using auxiliary information. After ranking select the i^{th} ranked unit for actual measurement then replace all *m* selected units to the population before the second draw. In general, the smallest ranked unit is quantified from 1^{st} set and 2^{nd} smallest unit is quantified from 2^{nd} set and so forth, till the unit with the largest rank i.e. m^{th} rank from the m^{th} set is quantified. This reflects one cycle of Level-0 RSS design. This complete cycle repeated *r* times by following the same procedure of sample selection given by Deshpande *et al.* (2006). Then we obtain the RSS sample of size $n = mr$ units are quantified from originally selected $m^2 r$ units.

Let consider a finite population $\Omega = \{Y_1, Y_2, \ldots, Y_N\}$ of size *N* with mean μ is linear in nature and variance σ^2. Without loss of generality, we assume $Y_1 \le Y_2 \le \ldots \le Y_N$. Under Level-0 RSS design define the event that the i^{th} ranked unit in the subset is the s^{th} ranked unit in the population. If quantified i^{th} ranked unit from the i^{th} set of size *m* is $y_{(i:m)}$, then the first two moments of the order statistics are $E[y_{(i:m)}] = \mu_{(i:m)}$ and $V(y_{(i:m)}) = \sigma^2_{(i:m)}$ which are mean and variance of i^{th} order statistics in the population respectively and the covariance term is zero, as sets are drawn independently. Similar notations have been used in Patil *et al.* (1994, 1995). Suppose, *mr* sets, each of *m* size, are randomly selected and replaced for the finite population Ω. Let in each of the first *r* sets, the lowest-ranked unit be quantified i.e. $y_{(1:m)1}, y_{(1:m)2}, y_{(1:m)3}, \ldots, y_{(1:m)r}$. The second-ranked unit is quantified in each of the next *r* sets to yield $y_{(2:m)1}, y_{(2:m)2}, y_{(2:m)3}, \ldots, y_{(2:m)r}$. This process continues until each of the last *r* sets quantifies the highest-ranked unit i.e. $y_{(m:m)1}, y_{(m:m)2}, y_{(m:m)3}, \ldots, y_{(m:m)r}$. The RSS estimator of the population mean $(\hat{\mu}_{RSS} \; or \; \bar{y}_{RSS})$ is the average of these quantifications:

$$\hat{\mu}_{RSS} = \bar{y}_{RSS} = \frac{1}{mr} \sum_{i=1}^{m} \sum_{k=1}^{r} y_{(i:m)k} \tag{1}$$

where $y_{(i:m)j}$ is the i^{th} ranked unit in the i^{th} set of size *m* at k^{th} cycle. The expectation and variance of $\hat{\mu}_{RSS}$, from the equation (1), is given by

$$E(\hat{\mu}_{RSS}) = E(\bar{y}_{RSS}) = \frac{1}{mr} \sum_{i=1}^{m} \sum_{k=1}^{r} E(y_{(i:m)k}) = \frac{1}{m} \sum_{i=1}^{m} \mu_{(i:m)} = \mu$$

and

$$V(\hat{\mu}_{RSS}) = V(\bar{y}_{RSS}) = V\left[\frac{1}{mr} \sum_{i=1}^{m} \sum_{k=1}^{r} y_{(i:m)k}\right]$$

$$= \frac{1}{m^2 r^2}\left[r\sigma^2_{(1:m)} + r\sigma^2_{(2:m)} + \ldots + r\sigma^2_{(m:m)}\right]$$

$$= \frac{1}{m^2 r} \sum_{i=1}^{m} \sigma^2_{(i:m)} = \frac{1}{mr}\left[\sigma^2 - \frac{1}{m} \sum_{i=1}^{m} \left\{\mu_{(i:m)} - \mu\right\}^2\right]. \tag{2}$$

However, some efforts were made to obtain the variance estimation for with replacement RSS design (i.e. Level-0 sampling protocol) under an infinite population framework (Stokes, 1980; Sinha *et al.* 1996). Due to the scaling issue under the finite population setting, the Naïve bootstrap technique (Efron, 1979) may not be able to give unbiased variance estimates.

Therefore, to overcome this scaling problem rescaling bootstrap with replacement techniques (Rao and Wu, 1988; Rao *et al.* 1992) and rescaling bootstrap without replacement techniques (Ahmad, 1997) have been developed to obtain unbiased variance estimation of the estimators of finite population parameters. Chen *et al.* (2004) and Modarres *et al.* (2006) developed the Bootstrapping technique for RSS design in an infinite population setting without the use of any rescaling and/or finite correction factor. Biswas *et al.* (2013) using the Jackknife technique to develop variance estimation methods for RSS design in a finite population framework. In the Level-2 RSS design under the finite population framework, Biswas *et al.* (2020) proposed variance estimation methods using the rescaling bootstrap technique.

In the vast literature of RSS, it has been observed that unbiased variance estimation for Level-0 RSS sampling under finite population framework has been found to be cumbersome and received less attention. In this article, therefore, an effort has been made to develop rescaling bootstrap methods to obtain the unbiased estimate of variance for the Level-0 RSS estimator of finite population mean using ranks and cycles of the RSS sample. In Section 2, the proposed techniques for obtaining an unbiased estimate of variance for the RSS Level-0 design estimator of the finite population mean are discussed. Proposed techniques viz. SRBWR and CRBWR have been addressed here. The results and discussion of the simulation study were given in Section 3 and Section 4 respectively. Using a well-known real dataset, Section 5 presents the findings of the real data application on proposed techniques. Section 6 provides the concluding remarks.

2. THE PROPOSED BOOTSTRAP VARIANCE ESTIMATION METHODS FOR THE LEVEL-0 RSS DESIGN

To obtain unbiased variance estimation for the Level-0 RSS estimator of the population mean under the finite population framework, two different rescaling bootstrap with replacement methods have been proposed in the following sub sections.

2.1 Strata-based rescaling bootstrap with-replacement (SRBWR) method

Let, $n=mr$ be the size of the final Level-0 RSS sample, it contains r observations for each of m ranks. Further, ranks are considered as strata and each rank's observations are considered as units within a stratum since RSS procedure generates a stratified sample "artificially" (Stokes and Sager 1988). In general, the Level-0 RSS sample comprises m strata, and each stratum is consists of r units. The steps involved in the proposed SRBWR method as follows:

1. Draw a simple random sample $\left\{y^*_{(i:m)k}\right\}_{k=1}^{p}$ of size $p(<r)$ with replacement from the observed sample values $\left\{y_{(i:m)k}\right\}_{k=1}^{r}$ of the i^{th} stratum.
2. Independently implement Step 1 by replacing the sample into the original sample for all strata $i=1,2,...,m$ and obtain a bootstrap resample as $\left\{y^*_{(i:m)k}\right\}$, where $i=1,2,...,m$ and $k=1,2,...,p$.
3. Then, determine,

   $\tilde{y}_{(i:m)k} = \bar{y}_{(i)} + f_1^{1/2}\left(y^*_{(i:m)k} - \bar{y}_{(i)}\right)$, for all $k=1,2,...,p$,

   where, $\bar{y}_{(i)} = \frac{1}{r}\sum_{k=1}^{r} y_{(i:m)k}$ and $f_1 = \frac{p}{r-1}$ is the rescaling factor.
4. Calculate,

   $$\tilde{\tilde{y}}_{(i)} = \frac{1}{p}\sum_{k}^{r} \tilde{y}_{(i:m)k} \text{ and } \tilde{\tilde{y}}_{RSS,st} = \frac{1}{m}\sum_{i=1}^{m} \tilde{\tilde{y}}_{(i)} \quad (3)$$

5. Repeat Step 1 to Step 4 independently after replacing the drawn resample with the original sample. Repeat this process many times, say B and compute the corresponding $\tilde{\tilde{y}}^1_{RSS,st}, \tilde{\tilde{y}}^2_{RSS,st}, \ldots, \tilde{\tilde{y}}^B_{RSS,st}$,
6. The bootstrap variance estimator of $\tilde{\tilde{y}}_{RSS,st}$ is given by

   $$\hat{\hat{V}}_{b,st} = V_*\left(\tilde{\tilde{y}}_{RSS,st}\right) = E_*\left(\tilde{\tilde{y}}_{RSS,st} - E_*\tilde{\tilde{y}}_{RSS,st}\right)^2, \quad (4)$$

   where E_* and V_* indicates the expectation and variance from a given original sample corresponding to the bootstrap sampling. The Monte Carlo (MC) estimator $\hat{\hat{V}}_{b,st}(a)$ as an approximation to $\hat{\hat{V}}_{b,st}$ is given by

   $$\hat{\hat{V}}_{b,st}(a) = \frac{1}{B-1}\sum_{b=1}^{B}\left(\tilde{\tilde{y}}^b_{RSS,st} - \tilde{\tilde{y}}_{RSS,st,a}\right)^2, \quad (5)$$

   where $\tilde{\tilde{y}}_{RSS,\,st,\,a} = \frac{1}{B}\sum_{b=1}^{B}\left(\tilde{\tilde{y}}^b_{RSS,st}\right)$ is the MC mean.

It can be easily shown that by taking design based expectation at the sampling stage on the bootstrap variance estimator of the estimator ($\hat{\hat{V}}_{b,st}$) results in the variance of the Lvel-0 RSS estimator ($V(\bar{y}_{RSS})$). Therefore, the proposed SRBWR method results in approximately unbiased estimator of the variance of the Level-0 RSS estimator of population mean.

2.2 Cluster-based rescaling bootstrap with-replacement (CRBWR) method

Let, $n=mr$ be the size of the final Level-0 RSS sample, it consists of r cycles consisting of one observation from each of the m ranks. Further, these r cycles are considered as clusters and each cluster of m units. Therefore, the Level-0 RSS sample consists of r clusters of m units each. The steps involved in the proposed CRBWR method as follows

1. A random sample of size $p(<r)$ clusters of m units each drawn from the observed r clusters using SRSWR technique and all the units of the chosen clusters are enumerated to get a bootstrap sample as $\left\{y^*_{(i:m)k}\right\}$, $i=1,2,...,m$ and $k=1,2,...,p$.
2. Then, calculate

$$\tilde{y}_{(i:m)k} = \bar{y}_{RSS} + f_2^{1/2}\left(y^*_{(i:m)k} - \bar{y}_{RSS}\right)$$

$$\tilde{\bar{y}}_k = \frac{1}{m}\sum_{i=1}^{m}\tilde{y}_{(i:m)k} \text{ and } \tilde{\bar{y}}_{RSS,\,cl} = \frac{1}{p}\sum_{k=1}^{r_1}\tilde{\bar{y}}_k \tag{6}$$

where $\bar{y}_{RSS} = \frac{1}{mr}\sum_{i=1}^{m}\sum_{k=1}^{r}y_{(i:m)k} = \frac{1}{r}\sum_{k=1}^{r}\tilde{\bar{y}}_k$ and $f_2 = \frac{p}{r-1}$ is the rescaling factor.

3. Replace the selected sample of $p(<r)$ clusters to the original sample of r clusters and replicate Step 1 and Step 2 independently. Repeat this process many times, say B, and compute the corresponding $\tilde{\bar{y}}^1_{RSS,cl}$, $\tilde{\bar{y}}^2_{RSS,cl}$, …, $\tilde{\bar{y}}^B_{RSS,cl}$.
4. The bootstrap variance estimator of $\tilde{\bar{y}}_{RSS,\,cl}$ is given by

$$\hat{\hat{V}}_{b,cl} = V_*\left(\tilde{\bar{y}}_{RSS,cl}\right) = E_*\left(\tilde{\bar{y}}_{RSS,cl} - E_*\tilde{\bar{y}}_{RSS,cl}\right)^2 \tag{7}$$

where E_* and V_* indicates the expectation and variance from a given original sample corresponding to the bootstrap sampling.

5. Monte Carlo (MC) estimator $\hat{\hat{V}}_{b,cl}(a)$ as an approximation to $\hat{\hat{V}}_{b,cl}$ is given by

$$\hat{\hat{V}}_{b,cl}(a)\frac{1}{B-1}\sum_{b=1}^{B}\left(\tilde{\bar{y}}^b_{RSS,cl} - \tilde{\bar{y}}_{RSS,\,cl,\,a}\right)^2 \tag{8}$$

where $\tilde{\bar{y}}_{RSS,\,cl,\,a} = \frac{1}{B}\sum_{b=1}^{B}\left(\tilde{\bar{y}}^b_{RSS,cl}\right)$.

It can be easily shown that by taking design based expectation at the sampling stage on the bootstrap variance estimator of the estimator ($\hat{\hat{V}}_{b,cl}$) results in the variance of the Level-0 RSS estimator $V\left(\bar{y}_{RSS}\right)$. Therefore, the proposed CRBWR method results in approximately unbiased estimator of the variance of the Level-0 RSS estimator of population mean.

3. SIMULATION STUDY

A simulation study has been conducted to compare the performance of two proposed rescaling bootstrap with replacement methods i.e. SRBWR and CRBWR for unbiasedly estimating the variance of Level-0 RSS design. A bivariate normal population of size 1000 units has been generated using SAS software. The parameters of the generated population are; mean of study variable Y is $\bar{Y}$=35, mean of auxiliary variable X is $\bar{X}$=30, the standard deviation of Y is σ_y=7, the standard deviation of X is σ_x, and the population correlation coefficient between X and Y *is* ρ = 0.85. Besides, 1000 independent Level-0 RSS samples of different sample sizes, i.e. 60, 120, and 180, with corresponding cycles (r) and ranks (m) have been drawn from the simulated population. Then, variance, the estimates finite population mean, percent CV, Skewness and kurtosis were determined for each of the 1000 Level-0 RSS samples. The percentage relative bias was computed by using the relation, $\%Bias = \left(\bar{y}_{RSS} - \bar{Y}\right)/\bar{Y}$, where, $\bar{y}_{RSS}$ population mean dependent on the Level-0 RSS estimator. Further, 1000 SRSWR samples were generated to compare the Level-0 RSS estimator with the traditional SRSWR estimator. The percentage gain in efficiency of the Level-0 RSS estimator about the SRSWR estimator was calculated using the expression $\%GE = \left[V\left(\bar{y}_{SRS}\right)/V\left(\bar{y}_{RSS}\right) - 1\right]\times 100$, where, $V\left(\bar{y}_{SRS}\right)$ is the variance of the SRSWR estimator and $V\left(\bar{y}_{RSS}\right)$ is the variance obtained based on 1000 iterations.

Then, 200 independent bootstrap resamples have been selected from each of the selected Level-0 RSS samples. Then, the proposed SRBWR and CRBWR methods are used to obtain the unbiased variance estimates of the Level-0 RSS estimator of the population mean. Also, variance estimate of Level-0 RSS estimator of a population mean and Monte Carlo (MC) bootstrap estimates of the population mean were obtained. Then, Percentage Relative Bias (%RB) and Relative Stability (RS) have been computed by using the formula,

$$\%RB = \left[\left\{\frac{1}{s}\sum_s \left\{\hat{V}_s\left(\bar{y}_{RSS}\right)\right\} - V\left(\bar{y}_{RSS}\right)\right\} / V\left(\bar{y}_{RSS}\right)\right] \times 100$$

and

$$RS = \left[\frac{1}{s}\left[\sum_s \left\{\hat{V}_s\left(\bar{y}_{RSS}\right) - V\left(\bar{y}_{RSS}\right)\right\}^2\right]^{1/2} / V\left(\bar{y}_{RSS}\right)\right]$$

where, $\hat{V}_s\left(\bar{y}_{RSS}\right)$ is the estimate of the proposed estimator at s^{th} sample.

4. SIMULATION RESULTS AND DISCUSSION

In this section, we discuss the simulated results of the proposed bootstrap methods of variance estimation viz. SRBWR and CRBWR. A comparative study has been done between Level-0 RSS estimator and usual SRSWR estimator of a population mean based on statistical properties such as Variance, %Bias, Percentage Coefficient of variation (%CV), Skewness, Kurtosis, and Percentage Gain in Efficiency (% GE) and presented in Table 1. Furthermore, a comparison has also been made between proposed methods of variance estimation using bootstrap techniques by considering the presence and absence of a rescaling factor. Finally, the comparison has been done among the proposed methods on basis of statistical properties such as Monte Carlo (MC) mean, an estimate of variance, Percentage Relative Bias (%RB), and Relative Stability (RS) and presented in Table 2.

Table 1. Statistical properties of Level-0 RSS estimator of a population mean for different sample size combinations ($n = mr$) from a finite population under bivariate normal distribution

m	*r*	Mean	Variance	% Bias	% CV	Skewness	Kurtosis	% GE
2	30	35.20	0.65	-0.040	2.30	0.04	0.21	21.14
	60	35.19	0.32	-0.069	1.62	0.06	-0.07	12.49
	90	35.21	0.19	-0.003	1.25	-0.06	-0.10	18.33
3	20	35.24	0.57	0.072	2.14	0.11	-0.07	39.65
	40	35.22	0.24	0.005	1.41	0.07	-0.08	48.93
	60	35.22	0.20	0.017	1.27	0.01	0.20	14.88
4	15	35.25	0.57	0.086	2.15	-0.06	-0.19	38.60
	30	35.18	0.27	-0.092	1.49	-0.08	-0.03	31.99
	45	35.22	0.16	0.000	1.15	0.10	-0.13	40.00

Note: *m* is set size of RSS sample, *r* is number of cycles.

It can be seen from Table 1 that the Level-0 RSS design estimator is unbiased, since the average value of the estimator is almost equal to the population mean. The %Bias of Level-0 RSS estimator was found to be negligible for all sample size combinations with the increase in sample size (n) as well as with the increase in set size (m) for the fixed sample size (n). The Level-0 RSS estimator is consistent since it has the least variance. Similarly, the Level-0 RSS estimator is more stable since %CV get decreases with an increase in n and an increase in m for fixed n. The outcomes of simulation also show that the Level-0 RSS estimator is comparatively symmetric and almost Mesokurtic. The Level-0 RSS estimator is more efficient than the SRSWR estimator since the %GE of this estimator more as compared than the SRS estimator of a population mean.

From Table 2, it can be noticed that both the standard SRBWR and CRBWR bootstrap methods with no rescaling factors give a high %RB and RS. Quite the reverse, when the suggested rescaling factors are used, then both proposed methods display very less %RB and RS. Thus, the rescaling factors suggested are very successful in significantly reducing the %RB compared to the usual standard bootstrap methods without using any rescaling factor. Therefore, as established theoretically and through simulation outcomes, both the proposed variance estimation techniques are almost unbiased to estimate the variance of the Level-0 RSS estimator. It can also be observed that the RS values of the SRBWR method are consistently lower than the CRBWR method for various sample size combinations.

5. REAL DATA APPLICATION

A real data application was carried out to study the performance of the proposed methods viz. SRBWR and CRBWR to estimate the variance of Level-0 RSS estimator of finite population mean using data set of a truncated version of 399 conifer (*Pinus palustris*) trees provided in Chen *et al.* (2004). This data set information consists *of X* as the diameter in centimetres at breast height of trees and *Y* as the complete height of trees in feet. This dataset's parameters are $\bar{X} = 21.062$, $\bar{Y} = 52.677$, $\sigma_x^2 = 320.539$, $\sigma_y^2 = 3253.446$ and $\rho(X,Y) = 0.899$. Here, this dataset was considered to be the study population and 1000 Level-0 RSS samples of different sample sizes (i.e. 36, 60, and 96) have been drawn with distinct ranks (m) and cycles (r) combinations. Then, variance, the estimates finite population mean, percent CV, Skewness and kurtosis were determined for each of the 1000 Level-0 RSS samples. Then, 200 independent bootstrap resamples have been selected from each of the selected Level-0 RSS samples. Then, the proposed

Table 2. Statistical properties of proposed methods viz. SRBWR and CRBWR for different sample sizes $(n = mr)$ with corresponding bootstrap sample sizes (mp) under the simulation study

m	r	p	Variance of RSS mean	Standard Bootstrap method (without rescaling factor)				Proposed Rescaling Bootstrap method (with rescaling factor)			
				MC mean	Estimate of variance	%RB	RS	MC mean	Estimate of variance	%RB	RS
Strata-based rescaling bootstrap with replacement (SRBWR)											
2	30	9	0.65	35.20	2.07	214.54	2.24	35.20	0.64	-2.38	0.20
	60	18	0.32	35.19	1.05	224.21	2.30	35.19	0.32	-1.08	0.16
	90	27	0.19	35.22	0.70	263.99	2.68	35.22	0.21	10.4	0.18
3	20	6	0.57	35.24	1.72	201.94	2.11	35.24	0.54	-4.64	0.20
	40	12	0.24	35.22	0.87	254.85	2.60	35.22	0.26	9.18	0.19
	60	18	0.20	35.22	0.59	197.77	2.02	35.22	0.18	-9.15	0.16
4	15	5	0.57	35.25	1.33	132.10	1.40	35.25	0.47	-17.1	0.24
	30	9	0.27	35.18	0.77	179.01	1.84	35.18	0.24	-13.4	0.19
	45	14	0.16	35.22	0.50	206.43	2.10	35.22	0.16	-2.49	0.14
Cluster-based rescaling bootstrap with replacement (CRBWR)											
2	30	9	0.65	35.21	2.06	212.46	2.29	35.21	0.64	-3.03	0.27
	60	18	0.32	35.19	1.06	223.22	2.34	35.19	0.32	-1.39	0.21
	90	27	0.19	35.22	0.70	260.62	2.68	35.22	0.21	9.40	0.21
3	20	6	0.57	35.24	1.72	199.69	2.22	35.25	0.54	-5.36	0.31
	40	12	0.24	35.22	0.87	253.26	2.67	35.22	0.27	8.69	0.27
	60	18	0.20	35.23	0.59	195.28	2.05	35.23	0.18	-9.92	0.21
4	15	5	0.57	35.25	1.57	172.93	2.03	35.25	0.56	-2.52	0.38
	30	9	0.27	35.19	0.93	232.79	2.51	35.19	0.29	3.28	0.29
	45	14	0.16	35.23	0.50	205.97	2.19	35.22	0.16	-2.65	0.24

Note: *m* is the set size of RSS sample, *r* is number of cycles, *p* is bootstrap sample size for rank/cycle.

SRBWR and CRBWR methods are used to obtain the unbiased variance estimates of the Level-0 RSS estimator of a population mean, Monte Carlo (MC) bootstrap estimates of the population mean, percentage Relative Bias (%RB) and Relative Stability (RS) have been obtained from the generated 200 bootstrap samples. The statistical properties of proposed methods under real data application are shown in Table 3.

It is visible from Table 3 that both the standard SRBWR and CRBWR bootstrap methods (with no rescaling factors) give a high %RB and RS. On the contrary, when the suggested rescaling factors are used, then both proposed methods display very less %RB and RS. Thus, the rescaling factors suggested are very successful in significantly reducing the %RB compared to the usual standard bootstrap methods without using any rescaling factor. Therefore, as established theoretically and through real data application, both the proposed variance estimation techniques are almost unbiased to estimate the variance of the Level-0 RSS estimator. It can also be observed that the RS values of the SRBWR method are consistently lower than the CRBWR method for various sample size combinations.

6. CONCLUSIONS

In this article, two rescaling bootstrap variance estimation techniques are proposed in Level-0 RSS design under finite population framework named as Strata-based rescaling bootstrap with-replacement

Table 3. Statistical properties of both SRBWR and CRBWR methods for different sample sizes $(n = mr)$ with corresponding bootstrap sample sizes (mp) under real dataset given in Chen *et al.* (2004)

m	*r*	*p*	Variance of RSS mean	Standard Bootstrap method (without rescaling factor)				Proposed Rescaling Bootstrap method (with rescaling factor)			
				MC mean	Estimate of variance	%RB	RS	MC mean	Estimate of variance	%RB	RS
Strata-based rescaling bootstrap with replacement (SRBWR)											
2	18	6	65.50	52.36	195.73	198.82	3.14	52.35	69.08	5.46	1.10
	30	12	43.66	52.85	121.91	179.19	2.87	52.84	42.04	-3.72	0.98
	48	16	26.92	52.86	77.16	186.62	2.90	52.86	26.26	-2.42	0.97
3	12	4	56.01	52.54	163.27	191.46	3.05	52.55	59.37	5.98	1.10
	20	7	21.37	52.77	94.89	117.31	2.23	52.77	34.96	-19.93	0.81
	32	11	34.89	52.74	61.60	188.25	2.91	52.74	21.86	2.28	1.01
4	9	3	49.81	52.71	137.33	175.68	2.91	52.70	51.50	3.38	1.08
	15	6	29.32	52.78	70.99	142.11	2.48	52.78	30.42	3.76	1.05
	24	8	18.86	52.79	56.04	197.09	3.01	52.78	19.49	3.33	1.02
Cluster based rescaling with-replacement bootstrap (CRBWR)											
2	18	6	65.50	52.34	198.98	203.78	3.24	52.35	70.23	7.21	1.13
	30	12	43.66	52.81	122.65	180.89	2.91	52.82	42.30	-3.14	0.99
	48	16	26.92	52.84	78.67	192.22	2.98	52.85	26.78	-0.52	0.99
3	12	4	56.01	52.59	164.08	192.91	3.17	52.58	59.67	6.51	1.14
	20	7	21.37	52.79	95.97	119.78	2.31	52.78	35.35	1.31	1.05
	32	11	34.89	52.72	62.00	190.08	2.98	52.73	22.00	2.93	1.03
4	9	3	49.81	52.72	138.87	178.78	3.12	52.72	52.07	4.54	1.16
	15	6	29.32	52.74	71.02	142.22	2.58	52.75	30.43	3.81	1.09
	24	8	18.86	52.75	55.92	196.49	3.07	52.76	19.45	3.13	1.04

Note: *m is set size of RSS sample, r is number of cycles, p is bootstrap sample size for rank/cycle.*

(SRBWR) method and the Cluster-based rescaling bootstrap with-replacement (CRBWR) method. Rescaling factors are developed under these proposed methods to theoretically demonstrate that the proposed variance estimators become almost unbiased for the variance of the Level-0 RSS design estimator of a population mean. Furthermore, a comparison was made between the proposed methods. To unbiasedly estimate the variance of the Level-0 RSS estimator, both proposed methods show a very small amount of percentage Relative Bias (%RB) and Relative Stability (RS). Thus, it can be concluded from the results of simulation study and real data application that with the use of suggested rescaling factors, the proposed methods are quite effective in significantly reducing %RB and RS. Also, the proposed SRBWR method performs better than the CRBWR method considering %RB and RS for different sample size combinations i.e. with a different combination of set size (*m*) and several cycles (*r*). It can therefore be inferred that the variance estimate obtained by the SRBWR procedure is better and more stable than the CRBWR method.

ACKNOWLEDGEMENTS

Authors are very thankful to the reviewer for giving valuable suggestions which led to the improvement of the article.

REFERENCES

Ahmad, T. (1997). A resampling technique for complex survey data. *J. Ind. Soc. Agril. Statist.*, **50(3)**, 364-379.

Al-Saleh, M.F. and H.M. Samawi. (2007). A note on inclusion probability in ranked set sampling and some of its variations. *Test*, **16(1)**, 198-209.

Al-Saleh, M.F. and Zheng, G. (2002). Estimation of bivariate characteristics using Ranked Set Sampling. *Australian and New Zealand Journal of Statistics*, **44(2)**, 221-232.

Arnab, R. (2017). *Survey sampling theory and applications*. Academic Press, Oxford, UK.

Biswas, A., Ahmad, T. and Rai, A. (2013). Variance estimation using Jackknife method in ranked set sampling under finite population framework. *Journal of the Indian Society of Agricultural Statistics*, **67(3)**, 345-353.

Biswas, A., Rai, A. and Ahmad, T. (2020). Rescaling bootstrap technique for variance estimation for ranked set samples in finite population. *Communications in Statistics-Simulation and Computation*, **49(10)**, 2704-2718.

Bocci, C., Petrucci, A. and Rocco, E. (2010). Ranked set sampling allocation models for multiple skewed variables: an application to agricultural data. *Environmental and Ecological Statistics*, **17(3)**, 333-345.

Chen, Z., Bai, Z. and Sinha, B. (2004). *Ranked set sampling:Theory and applications*. New York: Springer.

Deshpande, J.V., Frey, J. and Ozturk, O. (2006). Nonparametric ranked-set sampling confidence intervals for a finite population. *Environmental and Ecological Statistics,* **13(1)**, 25-40.

Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, **7(1)**, 1-26.

Frey, J. (2011). Recursive computation of inclusion probabilities in ranked set sampling. *Journal of Statistical Planning and Inference,* **141(11)**, 3632-3639.

Gokpinar, F. and Ozdemir, Y.A. (2010). Generalization of inclusion probabilities in ranked set sampling. *Hacettepe Journal of Mathematics and Statistics*, **39(1)**, 89-95.

Gokpinar, F. and Ozdemir, Y.A. (2012). A Horvitz-Thompson estimator of the population mean using inclusion probabilities of ranked set sampling. *Communications in Statistics - Theory and Methods*, **41(6)**, 1029-1039.

Gokpinar, F. and Ozdemir, Y.A. (2014). Simple computational formulas for inclusion probabilities in ranked set sampling. *Hacettepe Journal of Mathematics and Statistics*, **43(1)**, 117-30.

Halls, L.S. and T.R. Dell. (1985). Trial of ranked set sampling for forage yields. *Forest Science*, **12(1)**, 22-26.

Jozani, M.J. and Johnson, B.C. (2011). Design based estimation for ranked set sampling infinite population. *Environmental and Ecological Statistics*, **18(4)**, 663-685.

Jozani, M.J. and Johnson, B.C. (2012). Randomized nomination sampling in finite populations. *Journal of Statistical Planning and Inference*, **142(7)**, 2103-2115.

Kankure, A.and Rai, A. (2008). Spatial ranked set sampling from spatially correlated population. *J. Ind. Soc. Agril. Statist.*, **62(3)**, 221-230.

Krishna, P. (2002). Some aspects of ranked set sampling from finite population. *Unpublished M.Sc. Thesis, Indian Agricultural Research Institute*, New Delhi.

Martin, W.L., Sharik, T.L., Oderwald, R.G. and Smith, D.W. (1980). *Evaluation of ranked set sampling for estimating shrub phytomass in Appalachian oak forests*. *School of forestry and wild life resources.* Blacksburg, VA: Virginia Polytechnic Institute and State University.

McIntyre, G. (1952). A method of unbiased selective sampling using rank sets. *Australian Journal of Agricultural Research,* **3(4)**, 385-390.

Modarres, R., Hui, T.P. and Zheng, G. (2006). Resampling methods for ranked set samples. *Computational Statistics and Data Analysis*, **51(2)**, 1039-50.

Nahhas, R.W., Wolfe, D.A. and Chen, H. (2002). Ranked set sampling: cost and optimal set size. *Biometrics*, **58(4)**, 964-971.

Ozdemir, Y.A. and Gokpinar, F. (2007). A generalized formula for inclusion probabilities in ranked set sampling. *Hacettepe Journal of Mathematics and Statistics,* **36(1)**, 89-99.

Ozdemir, Y.A. and Gokpinar, F. (2008). A new formula for inclusion probabilities in median ranked set sampling. *Communications in Statistics – Theory and Methods*, **37(13)**, 2022-33.

Ozturk, O. (2014). Estimation of population mean and total in finite population setting using multiple auxiliary variables. *Journal of Agricultural, Biological, and Environmental Statistics*, **19(2)**, 161-84.

Ozturk, O. (2016a). Estimation of a finite population mean and total using population ranks of sample units. *Journal of Agricultural, Biological, and Environmental Statistics*, **21(1)**, 181-202.

Ozturk, O. (2016b). Statistical inference based on judgment post-stratified samples in finite population. *Survey Methodology*, **42(2)**, 239-262.

Ozturk, O. Bilgin, O. and Wolfe, D.A. (2005). Estimation of population mean and variance in flock management: A ranked set sampling approach in a finite population setting. *Journal of Statistical Computation and Simulation*, **75(11)**, 905-19.

Ozturk, O. and Jozani, M.J.(2014). Inclusion probabilities in partially rank ordered set sampling. *Computational Statistics and Data Analysis*, **69**, 122-32.

Patil, G.P., A.K. Sinha, and C. Taillie. (1994). Ranked set sampling. *Handbook of statistics*, 12. ed. G.P. Patil, and C.R. Rao, pp. 167–198, Amsterdam: North-Holland.

Patil, G.P., Sinha, A.K. and Taillie, C. (1995). Finite population corrections for ranked set sampling. *Annals of Institute of Statistical Mathematics*, **47(4)**, 621-636.

Rai, A. and Krishna, P. (2013). Ranked set sampling from finite population under randomization framework. *Journal of the Indian Society of Agricultural Statistics*, **67(3)**, 363-369.

Rao, J.N.K. and Wu. C.F.J. (1988). Resampling inference with complex survey data. *Journal of the American Statistical Association*, **83(401)**, 231-241.

Rao, J.N.K., Wu. C.F.J. and Yue, K. (1992). Some recent work on resampling methods for complex surveys. *Survey Methodology*, **18(2)**, 209-217.

Samawi, H.M. and Al-Sagheer, O.A.M. (2001). On the estimation of distribution function using extreme and median ranked set sampling. *Biometrical Journal*, **43(3)**, 357-373.

Sinha, B.K. and Purkayastha, S. (1996). On some aspects of ranked set sampling for estimation of normal and exponential parameters. *Statist. Decisions*, **14(3)**, 223-240.

Stokes, S.L. (1977). Ranked set sampling with concomitant variables. *Communication in Statistics-Theory and Methods*, **6(12)**, 1207-1211.

Stokes, S.L. (1980). Estimation of variance using judgement order ranked set samples. *Biometrics*, **36**, 35-42.

Stokes, S.L. and Sager, T.W. (1988). Characterization of a ranked-set sample with application to estimating distribution functions. *Journal of the American Statistical Association*, **83(402)**, 374-381.

Takahashi, K. and Wakimoto, K. (1968). On unbiased estimates of the population mean based on the sample stratified by means of ordering. *Annals of Institute of Statistical Mathematics*, **20(1)**, 1-31.

Takahasi, K. and Futatsuya, M. (1998). Dependence between order statistics in samples from finite population and its application to ranked set sampling. *Annals of the Institute of Statistical Mathematics*, **50(1)**, 49-70.

Takahasi, K. and Futatsuya, M. (1988). Ranked set sampling from a finite population. *Proceedings of the Institute of Statistical Mathematics*, **36(1)**, 55-68.

Wolfe, D.A. (2012). Ranked set sampling: Its relevance and impact on statistical inference. *ISRN Probability and Statistics*, *2012*.