

# Judgment Post-stratified Sampling with Multiple Ranking: A Comparison with Ranked Set Sampling



Omer Ozturk, Jennifer Brown, and Olena Kravchuk

**Abstract** Ranked set sampling and judgment post-stratified sampling designs form groups among sample units using their relative positions (ranks) in small comparison sets. This rank information governs the decision on whether to include units in a final ranked set sample (RSS), but only supplements the primary selection of units in a judgment post-stratified sample (JPS). If the position information in the comparison sets is accurate, for both designs, the samples represent the population better than a simple random sample (SRS) of the same size. The RSS design uses the ranking information in a more direct way. However, the RSS design induces a strong structure in a sample, and the data so collected may not be suitable for studies where a multipurpose analysis is desired. The JPS design is slightly less efficient, but more flexible and enables multipurpose analyses. This paper explores the benefits of the JPS over the RSS design of the same sample size. We show that the efficiency loss in the JPS design can be reduced by using ranks from multiple comparison sets. The paper presents results from an extensive simulation study to demonstrate the benefit of the JPS design over the SRS and RSS designs when the JPS is constructed using multiple ranking methods.

---

O. Ozturk (✉)

The Ohio State University, Department of Statistics, Columbus, OH, USA

e-mail: [ozturk.4@osu.edu](mailto:ozturk.4@osu.edu)

J. Brown

University of Canterbury, School of Mathematics and Statistics, Christchurch, New Zealand

e-mail: [jennifer.brown@canterbury.ac.nz](mailto:jennifer.brown@canterbury.ac.nz)

O. Kravchuk

University of Adelaide, School of Agriculture, Food and Wine, Adelaide, SA, Australia

e-mail: [olena.kravchuk@adelaide.edu.au](mailto:olena.kravchuk@adelaide.edu.au)

## 1 Introduction

In field sampling and social science research, creating samples that are representative of the population is important. This can be achieved by using stratified sampling, cluster sampling, or post-stratified sampling designs. In certain cases, the stratification variable may be subjective, rough, and imprecise, but can still provide valuable information about the relative position of a sample unit in a small set. Such stratification variable can be used to reduce the sampling variation, and cost in ranked set and judgment post-stratified sampling designs. These designs stratify the sample into groups of homogeneous observations using sample units' relative positions (ranks) in small comparison sets.

For a ranked set sample (RSS) of size  $n$ , one first determines a set size  $H$  and then selects  $nH$  units at random from the population. These units are divided into  $n$  comparison sets, each of size  $H$ . Units in the comparison sets are then ranked from the smallest to the largest, without measurement. Ranking can be performed on either the variable of interest assessed on a less elaborate scale or an auxiliary variable. The unit judged to be the  $h$ -th smallest ( $Y_{[h]j}$ ) is measured in  $n_h$  comparison sets for  $j = 1, \dots, n_h$ ,  $\sum_{h=1}^H n_h = n$ . The measured observations  $Y_{[h]j}$ ,  $j = 1, \dots, n_h$ ;  $h = 1, \dots, H$  are called a ranked set sample. If  $n_h = d$  for all  $h = 1, \dots, H$  so that  $n = dH$ , the RSS is called balanced, and  $d$  is called the cycle size. If there is no ranking error, the square brackets are replaced with round parentheses, and the  $Y_{(h)j}$  becomes the  $h$ -th order statistic in a sample of size  $H$ .

Ranked set sampling design was introduced by McIntyre (1952, 2005). The main motivation in McIntyre's work was to enable field researchers to conduct pasture yield (and similar) field assessments in an objective and efficient way. Takahasi and Wakimoto (1968) developed the theoretical foundation of the ranked set sampling design and showed that the RSS mean is always better than a sample mean of a simple random sample (SRS). Dell and Clutter (1972) showed that even with some ranking errors, the RSS mean is as good as, or better than, the SRS mean depending on the quality of ranking information. Research activities in RSS designs then expanded in different directions, including parametric and nonparametric settings. In the parametric setting, a few representative publications are Stokes (1995), Chen and Bai (2000), Arslan and Ozturk (2013), Hatefi et al. (2014), and Hatefi et al. (2015). In the nonparametric setting, readers are referred to Bohn and Wolfe (1992, 1994), Hettmansperger (1995), Koti and Babu (1996), Ozturk (1999), and Fligner and MacEachern (2006). Two books have been published on ranked set sampling design, Chen et al. (2003), and Bouza and Al-Omari (2019). A comprehensive list of references can be found in these publications.

The RSS research activities also considered the finite population setting. Patil et al. (1995) constructed an RSS using sampling without replacement selection procedure. Deshpande et al. (2006) expanded the RSS design to three different schemes of sampling without replacement. Frey (2011), Ozturk and Jafari Jozani (2014), and Jafari Jozani and Johnson (2011) used probability sampling and constructed Horvitz-Thompson-type estimators. Ozturk and Bayramoglu Kavlak (2018) constructed inference using a superpopulation model in ranked set sampling.

MacEachern et al. (2004) introduced the judgment post-stratification design to provide the flexibility for a multipurpose analysis of sample data. For a judgment post-stratified sample (JPS), one first selects and measures an SRS of size  $n$ ,  $Y_i$ ,  $i = 1, \dots, n$ . For each measured unit  $Y_i$ , one then selects additional  $H - 1$  units from the population, without directly measuring them, to form a comparison set of size  $H$ . The units in the comparison set are ranked from the smallest to the largest, and the rank of  $Y_i$ ,  $R_i$  is recorded. The pairs  $(Y_i, R_i)$ ,  $i = 1, \dots, n$ , constitute a JPS.

In recent years, the JPS design in an infinite population setting has generated extensive research interest. Ozturk (2014) considered the estimation of the population quantile and variance from a JPS. Wang et al. (2006) used the concomitant order statistics to estimate the population mean. Frey and Feeman (2012, 2013) constructed estimators for the population mean and variance by conditioning on the judgment group sample sizes. These new estimators improve the unconditional JPS estimators. Chen et al. (2014), Frey and Ozturk (2011), Wang et al. (2012), Wang et al. (2008), and Stokes et al. (2007) constructed constrained estimators using stochastic ordering among judgment ranking groups. The main idea in the constraint estimators is to minimize the impact of ranking error by forcing judgment class means to follow the stochastic order among ranking groups. In a different direction, Ozturk (2017) constructed conditional ranks in smaller comparison sets of size  $K < H$  given the original ranks in a larger comparison set of size  $H$ . The impact of any ranking error on the estimator in this case was relatively small, and less than for the estimator based on the large comparison set of size  $H$ . Ozturk (2013) and Ozturk and Demirel (2016) used a multi-ranking approach to reduce the impact of ranking error in judgment post-stratified and ranked set samples.

In the finite population settings, Ozturk (2016a, 2016b, 2019) constructed estimators for the population mean and total for the JPS design. A JPS can be constructed by sampling with or without replacement. It is shown that the variance estimator of the sample mean requires a finite population correction factor when sampling without replacement. Ozturk and Bayramoglu Kavlak (2018, 2019, 2020) developed inference to predict the population mean and total using a superpopulation model.

In the JPS design, the ranks are constructed post-experimentally after an SRS is chosen. Hence, it is possible to have more than one rank for each measured unit in the SRS by permuting the  $n(H - 1)$  unmeasured units used in the construction of comparison sets in the first created JPS. Each permutation creates  $n$  comparison sets, each of size  $H$ , containing the measured unit. The units in the sets are ranked again, without measurement, and the ranks of the measured units in the comparison sets are determined. This permutation procedure can be done many times and each permutation creates a new set of ranks for the same measured values. Ranks from different permutations are conditionally independent given the original SRS. One may then combine all these ranks using the Rao-Blackwell theorem by conditioning on the original SRS.

A similar idea can be used in the RSS design, but the extension to multiple ranks is not as trivial as in the JPS design. In the RSS design, the measured observations,  $Y_{[h]j}$ , are not identically distributed. Hence, the units in the comparison set constructed after the permutation of  $n(H - 1)$  units are not iid since each comparison

set contains one of the  $y_{[h]j}$  from the original ranked set sample and this will have a different distribution from the other units in the set. Even though the comparison sets will be different after each permutation, the rank of  $y_{[h]j}$  will depend on the original rank  $h$ . Hence, the idea of multiple ranks in the judgment post-stratified sampling may not be easily extended to ranked set sampling.

There are a few other differences between the RSS and JPS designs. One of the major differences is whether the ranking is done before or after the units are measured for the variable of interest. In RSS, the ranking is performed before one measures the units, and the ranks guide the measurement decision. The rank and the measurement of a unit cannot be separated. Hence, an RSS cannot be reduced to an SRS, unless it is unusual situation where the ranking variable is not correlated with the measurement variable. In a JPS, ranking is performed after one measures the units in the SRS. The ranks are not the essential part of the measured units; they are the ranks of the variable of interest measured on a quicker scale (e.g., visual inspection) after the construction of an SRS. Since the auxiliary (ranking) variable is only post-associated with the response measurements, it can be ignored and a JPS can be reduced to an SRS if desired.

Another major difference is the distributional properties of the ranks. The ranks in RSS are pre-determined nonrandom constants. Hence, the ranking group sizes  $n_h$ ,  $h = 1, \dots, H$ , are nonrandom integers. In a JPS, the rank  $R_i$  is a discrete uniform random variable with the support on integers  $1, \dots, H$ . Hence, the judgment group sample size vector  $(n_1, \dots, n_H)$  has a multinomial distribution with the sample size  $n$  and the success probability vector  $(1/H, \dots, 1/H)$ .

One may look at the RSS and JPS designs in terms of the trade-off between the efficiency gain of the RSS and the adaptability of JPS for multipurpose studies. To our knowledge, this trade-off has not yet been posed and investigated. In this paper, we provide a comprehensive study to compare the RSS and JPS designs for their efficiencies and multiple ranking properties. In Sect. 2, we provide a detailed description of multi-ranking in RSS and JPS designs. In Sect. 3, we review the distributional properties of the RSS and JPS means. In Sect. 4, we present empirical results to compare the RSS and JPS designs. In Sect. 5, we illustrate the use of RSS and JPS designs with an agricultural application example. Section 6 provides concluding remarks.

## 2 Sampling Designs with Multiple Ranking Methods

We consider a finite population of size  $N$ . The population values of the variable  $Y$  are denoted as  $y_1, \dots, y_n$ . The mean and variance of the population are given by

$$\bar{y}_N = \frac{1}{N} \sum_{i=1}^N y_i, \quad S_N^2 = \frac{1}{N} \sum_{i=1}^N (y_i - \bar{y}_N)^2.$$

From this finite population, we construct RSS and JPS with multiple ranks. The samples are constructed using the sampling *with* and *without* replacement selection procedures. Unless stated otherwise, we always consider a finite population setting in this paper.

*RSS with Multiple Ranks* We first consider an RSS selected using the sampling with replacement (SWR) selection procedure. For cycle  $j$  and rank  $h$ , we construct a comparison set of size  $H$  using a sampling without replacement (SWOR) scheme. The units in the comparison set are ranked by the best ranking method available. The unit judged to be the  $h$ -th smallest,  $Y_{[h]j}$ , is measured. For the observation  $Y_{[h]j}$ , additional  $K - 1$  ranks can be constructed in two ways. If there are  $K - 1$  ( $K > 1$ ) rankers or ranking variables available, the rank of  $Y_{[h]j}$ ,  $R_{k|j,h}$ , among the units in the comparison set is determined for each method  $k$ ,  $k = 2, \dots, K$ . After these ranks are determined, all units in the comparison set are returned to the population before constructing the next comparison set. Hence, the same unit may appear in the final sample more than once, and all the observations are independent. We note that units within a comparison set are selected using the SWOR procedure to minimize the ranking error. The ranks using the first ranking method ( $k = 1$ ) are predetermined (nonrandom constants) to have a balanced ranked set sample,  $n_h = d$ , for  $h = 1, \dots, H$ . The remaining  $K - 1$  ranks are random and may not necessarily be balanced.

Even if there is only one ranker or one auxiliary variable to rank the units, we can still construct an RSS with multiple ranks. For given values of  $h$  and  $j$ ,  $Y_{[h]j}$  is measured in a comparison set. Next, we form  $K - 1$  different comparison sets by selecting  $H - 1$  additional units at random from the population without measurement,  $V_{k|h,j} = \{Y_{[h]j}, Y_{k,1}, \dots, Y_{j,H-1}\}$ ,  $k = 2, \dots, K$ , and determine the rank of  $Y_{[h]j}$ ,  $R_{k|j,h}$ , in each set for  $k = 2, \dots, K$ . The RSS with multiple ranks can be written as

$$\{(Y_{[h]j}, R_{k|j,h}), h = 1, \dots, H, j = 1, \dots, d; k = 1, \dots, K\},$$

where  $R_{k|j,h}$  is the conditional rank assigned by ranking method  $k$  given that the observation  $Y_{[h]j}$  is assigned rank  $h$ . We note that  $P(R_{1|j,h} = h) = 1$ . The ranks assigned by another ranking method are random variables, but their distributions depend on the ranks assigned by the first (best) ranking method.

An RSS with multiple ranks using a SWOR selection scheme can be constructed in a similar fashion. The only difference is that after determining the rank of  $Y_{[h]j}$ , all  $H$  units in the comparison set are removed from the population before constructing the next comparison set. Hence, for each ranking method, all comparison sets are disjoint.

The final sample cannot have repeated observations and the observations are not independent. If the population size  $N$  is large with respect to sample size  $n$ , ranked set samples constructed using SWR or SWOR selection procedures become approximately equivalent.

**Table 1** Illustration of RSS multi-ranker sampling with set size  $H = 3$ , cycle size  $d = 2$ , and the number of ranking methods  $K = 3$

Cycle ( $j$ )	h	Balanced RSS	Ranks from $K$ methods	Ranked set sample
1	1	$\{Y_{[1]1}, Y_{[2]1}, Y_{[3]1}\}$	$\{1, R_{2 1,1}, R_{3 1,1}\}$	$\{Y_{[1]1}, 1, R_{2 1,1}, R_{3 1,1}\}$
1	2	$\{Y_{[1]1}, Y_{[2]1}, Y_{[3]1}\}$	$\{2, R_{2 1,2}, R_{3 1,2}\}$	$\{Y_{[2]1}, 2, R_{2 1,2}, R_{3 1,2}\}$
1	3	$\{Y_{[1]1}, Y_{[2]1}, Y_{[3]1}\}$	$\{3, R_{2 1,3}, R_{3 1,3}\}$	$\{Y_{[3]1}, 3, R_{2 1,3}, R_{3 1,3}\}$
2	1	$\{Y_{[1]2}, Y_{[2]2}, Y_{[3]2}\}$	$\{1, R_{2 2,1}, R_{3 2,1}\}$	$\{Y_{[1]2}, 1, R_{2 2,1}, R_{3 2,1}\}$
2	2	$\{Y_{[1]2}, Y_{[2]2}, Y_{[3]2}\}$	$\{2, R_{2 2,2}, R_{3 2,2}\}$	$\{Y_{[2]2}, 2, R_{2 2,2}, R_{3 2,2}\}$
2	3	$\{Y_{[1]2}, Y_{[2]2}, Y_{[3]2}\}$	$\{3, R_{2 2,3}, R_{3 2,3}\}$	$\{Y_{[3]2}, 3, R_{2 2,3}, R_{3 2,3}\}$

The construction of ranked set samples using multiple ranking methods is illustrated in Table 1. In this table, the third column presents the comparison sets in which a balanced ranked set sample is constructed with the first ranking method. It highlights that the units are ranked using ranking method 1; the bold-faced values are measured. The fourth column lists the ranks obtained from all  $K$  ( $K = 3$ ) different ranking methods for the bold-faced values in column 3. The last column gives the ranked set sample of size 6. In this example, each entry has three ranks generated by three ranking methods.

*JPS with Multiple Ranks* We first construct a simple random sample of size  $n$  using the SWR selection procedure and measure all  $n$  units,  $Y_1, \dots, Y_n$ . For each  $Y_i$ , we then select additional  $H - 1$  units under SWOR selection from the population to form a comparison set  $V_i = \{Y_i, Y_1, \dots, Y_{H-1}\}$ . We rank these units from smallest to largest without measuring  $Y$ , using  $K$  different ranking methods, and identify the rank of  $Y_i$ ,  $R_{k|i}$ , for each ranking method  $k$ ,  $k = 1, \dots, K$ , where  $R_{k|i}$  is the rank of  $Y_i$  assigned by ranking method  $k$ . All units in the comparison set, including the one we measured, are returned to the population before the construction of the next comparison set. Hence, a JPS may have repeated observations and all  $Y_i$ ,  $i = 1, \dots, n$ , are independent. This process creates the sample

$$\{Y_i, R_{k|i}\}; i = 1, \dots, n, k = 1, \dots, K.$$

If only one ranking method is available, for each  $Y_i$ , one can create  $K$  different comparison sets,  $V_{k|i} = \{Y_i, Y_{1,k}, \dots, Y_{H-1,k}\}$  for  $k = 1, \dots, K$ , where  $Y_{h,k} \neq Y_i$  is the additional unit selected from the population to construct the  $k$ -th comparison set. These sets are ranked using the ranking method and the ranks of  $Y_i$ ,  $R_{k|i}$ , are determined in  $V_{k|i}$ , for  $k = 1, \dots, K$ .

A JPS under the SWOR selection procedure is constructed in a similar fashion. The only difference here is that all comparison sets for each ranking method are disjoint, and hence, the JPS cannot have repeated observations. For small population sizes  $N$ , observations  $Y_i$ ,  $i = 1, \dots, n$ , in the sample are negatively correlated since sample units are selected as an SRS without replacement.

The construction of a JPS with multiple ranking methods and under the SWOR selection scheme is illustrated in Table 2. In this example, the sample and set

**Table 2** Illustration of multi-ranker JPS under sampling without replacement selection with set size  $H = 3$ , sample size  $n = 6$ , and the number of ranking methods  $K = 3$

$(j)$	SRS	Comparison sets	JPS
1	$Y_1$	$\{Y_1, Y_{7,1}, Y_{8,1}\}, \{Y_1, Y_{9,2}, Y_{19,2}\}, \{Y_1, Y_{26,3}, Y_{12,3}\}$	$\{Y_1, R_{1 1}, R_{2 1}, R_{3 1}\}$
2	$Y_2$	$\{Y_2, Y_{9,1}, Y_{10,1}\}, \{Y_2, Y_{20,2}, Y_{8,2}\}, \{Y_1, Y_{17,3}, Y_{27,3}\}$	$\{Y_2, R_{1 2}, R_{2 2}, R_{3 2}\}$
3	$Y_3$	$\{Y_3, Y_{11,1}, Y_{12,1}\}, \{Y_3, Y_{21,2}, Y_{22,2}\}, \{Y_3, Y_{9,3}, Y_{10,3}\}$	$\{Y_3, R_{1 3}, R_{2 3}, R_{3 3}\}$
4	$Y_4$	$\{Y_4, Y_{13,1}, Y_{14,1}\}, \{Y_4, Y_{15,2}, Y_{23,2}\}, \{Y_4, Y_{14,3}, Y_{28,3}\}$	$\{Y_4, R_{1 4}, R_{2 4}, R_{3 4}\}$
5	$Y_5$	$\{Y_5, Y_{15,1}, Y_{16,1}\}, \{Y_5, Y_{18,2}, Y_{24,2}\}, \{Y_5, Y_{29,2}, Y_{13,2}\}$	$\{Y_5, R_{1 5}, R_{2 5}, R_{3 5}\}$
6	$Y_6$	$\{Y_6, Y_{17,1}, Y_{18,1}\}, \{Y_6, Y_{12,2}, Y_{25,2}\}, \{Y_6, Y_{8,3}, Y_{30,3}\}$	$\{Y_6, R_{1 6}, R_{2 6}, R_{3 6}\}$

sizes are 6 and 3, respectively. For each measured unit, three ranks are constructed ( $K = 3$ ). The second column presents a simple random sample of size  $n = 6$ . The third column presents three comparison sets,  $V_{k|i}$ , for each  $Y_i$ , one for each ranking method. The fourth column presents the JPS with three ranks. The comparison sets of each ranking method in Table 2, sets in block 1, 2, or 3 in column 3, are disjoint and cannot have repeated observations. Comparison sets for the different ranking methods (sets in different blocks) are not necessarily disjoint because the same ranking unit can appear in more than one set in different ranking methods. Sampling is without replacement and thus the comparison sets in different rows for the same ranking method are disjoint. We note that the sample units will not be independent if the population size  $N$  is small in relation to the sample size  $n$ .

### 3 Statistical Inference Using RSS and JPS

In this section, we provide a brief overview of statistical inference using the RSS and JPS designs. We first assume  $K = 1$ . The estimators for the population mean are given as the sample mean of the RSS and JPS:

$$\bar{Y}_{RSS} = \frac{1}{dH} \sum_{h=1}^H \sum_{j=1}^d Y_{[h]j}, \quad \bar{Y}_{JPS} = \frac{1}{d_n} \sum_{h=1}^H J_h I_h \sum_{j=1}^n Y_j I(R_j = h),$$

where  $I(a)$  is 1 if  $a$  is true,  $I_h = I(n_h > 0)$ ,  $d_n = \sum_{h=1}^H I_h$ , and  $J_h = 1/n_h$  if  $n_h > 0$  and zero otherwise. Both of these estimators are unbiased for the population mean  $\bar{y}_N$  regardless of the ranking quality as long as a consistent ranking method is used. If all units in the comparison sets are ranked with the same ranking methods, the ranking procedure is called consistent. The following theorem provides variances of the sample means under SWR and SWOR selection schemes using a consistent ranking method.

**Theorem 1** Let  $Y_{[h]j}$ ,  $h = 1, \dots, H$ ,  $j = 1, \dots, d$  and  $(Y_j, R_j)$ ,  $j = 1, \dots, n$  be RSS and JPS constructed using a consistent ranking methods, respectively.

(i) If the samples are constructed with replacement, the variances of  $\bar{Y}_{RSS}$  and  $\bar{Y}_{JPS}$  are given by

$$\sigma_{RSS}^2 = \frac{1}{dH^2} \sum_{h=1}^H S_{[h]}^2 \quad \sigma_{JPS}^2 = \frac{H}{H-1} \text{Var} \left( \frac{I_1}{d_n} \right) \sum_{h=1}^H (\bar{y}_{[h]} - \bar{y}_N)^2 + E \left( \frac{I_1^2}{d_n^2 n_1} \right) \sum_{h=1}^H S_{[h]}^2,$$

where  $S_{[h]}^2 = \text{Var}(Y_{[h]1})$ ,  $\bar{y}_{[h]} = E(Y_{[h]1})$ ,  $\text{Var}(I_1/d_n) = \frac{1}{H^2} \sum_{k=1}^{H-1} \left(\frac{k}{H}\right)^{n-1}$  and

$$E \left( \frac{I_1^2}{n_1 d_n^2} \right) = \frac{1}{H^n} \left( \frac{1}{n} + \sum_{k=2}^H \sum_{j=1}^{k-1} \sum_{t=1}^{n-k+1} \frac{(-1)^{j-1} (H-1)(k-1) \binom{n}{t} (k-j)^{n-t}}{k^2 t} \right).$$

(ii) If the samples are constructed without replacement, the variances of  $\bar{Y}_{RSS}$  and  $\bar{Y}_{JPS}$  are given by

$$\sigma_{RSS}^2 = \frac{N-1-n}{n(N-1)} S_N^2 - \frac{1}{nH} \sum_{h=1}^h (\bar{y}_{[h]} - \bar{y}_N)^2 - \frac{1}{nH} \sum_{h=1}^H S_{[h,h]}$$

$$\sigma_{JPS}^2 = C_1(n, H) \left\{ \sum_{h=1}^H S_{[h]}^2 - \sum_{h=1}^H S_{[h,h]} \right\} + C_2(n, H, N) \frac{H^2 S_N^2}{H-1},$$

where  $S_{[h,h]} = \text{Cov}(Y_{[h]1}, Y_{[h]2})$ ,

$$C_1(n, H) = \left\{ \frac{1}{H(H-1)} + E \left( \frac{I_1^2}{d_n^2 n_1} \right) - \frac{H}{H-1} E \left( \frac{I_1^2}{d_n^2} \right) \right\}$$

$$C_2(n, H, N) = \left\{ \text{Var} \left( \frac{I_1}{d_n} \right) - \frac{1}{N-1} \left\{ \frac{1}{H} - E \left( \frac{I_1^2}{d_n^2} \right) \right\} \right\}.$$

The proofs of  $\sigma_{JPS}^2$  in (i) and (ii) are given in Ozturk (2016a). The proof of  $\sigma_{RSS}^2$  in (ii) is given in Patil et al. (1995). It is clear that the variance of the JPS mean involves expected values and variances of the functions of judgment group indicator function ( $I_1$ ), sample sizes ( $n_1$ ), and the number of non-empty judgment groups ( $d_n$ ). These quantities account for the variation due to the random sample sizes in judgment post-stratified samples. Ozturk (2016b) shows that as the sample size  $n$  becomes large,  $\sigma_{JPS}^2$  approaches from above  $\sigma_{RSS}^2$ .



We now introduce unbiased estimators for  $\sigma_{JPS}^2$  and  $\sigma_{RSS}^2$ . We first define the following quantities:

$$U_1 = \frac{1}{E\left(\frac{I_1 I_2}{d_n^2}\right)} \sum_{h=1}^H \sum_{h \neq h'}^H \frac{I_h I_{h'} J_h J_{h'}}{d_n^2} \sum_{i=1}^n \sum_{j=1}^n (Y_i - Y_j)^2 I(R_i = h) I(R_j = h'),$$

$$U_2 = \sum_{h=1}^H \frac{H I_h^* J_h J_h^*}{d_n^*} \sum_{i=1}^n \sum_{j \neq i}^n (Y_i - Y_j)^2 I(R_i = h) I(R_j = h),$$

$$U_1^* = \frac{1}{2d^2 H^2} \sum_{h=1}^H \sum_{h' \neq h}^H \sum_{i=1}^d \sum_{j=1}^d (Y_{[h]i} - Y_{[h']j})^2$$

$$U_2^* = \frac{1}{2d(d-1)H^2} \sum_{h=1}^H \sum_{i=1}^d \sum_{j \neq i}^d (Y_{[h]i} - Y_{[h]j})^2,$$

where  $d_n^* = \sum_{h=1}^H I(n_h > 1)$ , and  $J_h^* = 1/(n_h - 1)$  if  $n_h > 1$  and zero otherwise.

**Theorem 2** Let  $Y_{[h]j}$ ,  $h = 1, \dots, H$ ,  $j = 1, \dots, d$  and  $(Y_j, R_j)$ ,  $j = 1, \dots, n$  be RSS and JPS constructed using a consistent ranking method, respectively.

(i) If the samples are constructed with replacement,  $d > 1$  and at least one judgment group in a JPS has at least two observations, the unbiased variance estimators for  $\bar{Y}_{RSS}$  and  $\bar{Y}_{JPS}$  are given by

$$\hat{\sigma}_{JPS}^2 = \frac{Var(I_1/d_n)}{2(H-1)} U_1 + \left\{ E\left(\frac{I_1^2}{d_n^2 n_1}\right) - Var\left(\frac{I_1}{d_n}\right) \right\} \frac{U_2}{2}$$

$$\hat{\sigma}_{RSS}^2 = \frac{U_2^*}{d}.$$

(ii) If the samples are constructed without replacement,  $d > 1$  and at least one judgment group in a JPS has at least two observations, the unbiased variance estimators for  $\hat{Y}_{RSS}$  and  $\hat{Y}_{JPS}$  are given by

$$\hat{\sigma}_{JPS}^2 = C_1(n, H) U_2 / 2 + C_2(n, H, N) \frac{(N-1)(U_1 + U_{2,2})}{2N(H-1)},$$

$$\hat{\sigma}_{RSS}^2 = \frac{U_2^*}{d} - \frac{U_1^* + U_2^*}{N}.$$

Theorem 2 provides unbiased estimators for the variance of the RSS and JPS means for an arbitrary but consistent ranking scheme when  $K = 1$ . An approximate  $(1 - \alpha)100\%$  confidence interval for the population mean can be constructed using the

normal approximation:

$$\begin{aligned}\bar{Y}_{RSS} &\pm t_{1-\alpha/2, n-H} \hat{\sigma}_{RSS} \\ \bar{Y}_{JPS} &\pm t_{1-\alpha/2, n-H} \hat{\sigma}_{JPS},\end{aligned}$$

where  $t_{a,df}$  is the  $a$ -th upper quantile of the  $t$ -distribution with degrees of freedom  $df$ . The degrees of freedom  $df = n - H$  is suggested to account for the heterogeneity among ranking groups.

There are different ways to combine the ranking information in multi-ranking RSS and JPS designs. Ozturk and Kravchuk (2021a, 2021b) provided detailed developments of these procedures. In this paper, we only consider one of the approaches, in which each observation is weighted based on the agreement scores of the  $K$  ranking methods. Let  $w_{h',i}$  be the proportion of  $K$  ranking methods which assign rank  $h'$  to the  $i$ -th observation in the sample:

$$w_{h'|i,h} = \frac{1}{K} \sum_{k=1}^K I(R_{k|i,h} = h')/K, \quad h' = 1, \dots, H, \quad \text{for the RSS}$$

and

$$w_{h'|i} = \frac{1}{K} \sum_{k=1}^K I(R_{k|i} = h')/K, \quad h' = 1, \dots, H, \quad \text{for the JPS.}$$

We estimate the population mean by allocating each observation into ranking group  $h'$  based on how strong the agreement is among the  $K$  ranking methods to assign the observation to judgment group  $h'$ :

$$\begin{aligned}\bar{Y}_{RSS,w} &= \sum_{h'=1}^H \frac{J_{w,h'}}{d_w} \sum_{h=1}^H \sum_{i=1}^d Y_{[h|i]} w_{h'|i,h}, \quad J_{w,h'} = \begin{cases} \frac{1}{n_{w,h'}} & \text{if } n_{w,h'} > 0 \\ 0 & \text{otherwise} \end{cases}, \quad n_{w,h'} = \sum_{i=1}^d \sum_{h=1}^H w_{h'|i,h}. \\ \bar{Y}_{JPS,w} &= \sum_{h'=1}^H \frac{J_{w,h'}}{d_w} \sum_{h=1}^H \sum_{i=1}^d Y_i w_{h'|i}, \quad J_{w,h'} = \begin{cases} \frac{1}{n_{w,h'}} & \text{if } n_{w,h'} > 0 \\ 0 & \text{otherwise} \end{cases}, \quad n_{w,h'} = \sum_{i=1}^d \sum_{h=1}^H w_{h'|i},\end{aligned}$$

where  $d_w = \sum_{h'=1}^H I(n_{w,h'} > 0)$ . In the expressions above,  $n_{w,h'}$  can be considered as the effective sample size for judgment group  $h'$ . The asymptotic distribution of  $\bar{Y}_{JPS,w}$  is considered in MacEachern et al. (2004) and Ozturk and Kravchuk (2021a). The asymptotic distribution of  $\bar{Y}_{RSS,w}$  is given in Ozturk and Kravchuk (2021b).

In this paper, we only consider the jackknife variance estimates of these estimators. Let  $\bar{Y}_{RSS,w}^{(-[h]i)}$  ( $\bar{Y}_{JPS,w}^{(-i)}$ ) be the RSS (JPS) estimator after the observation  $Y_{[h|i]}$  ( $Y_i$ ) and all ranks associated with it are removed from the sample. The jackknife variance estimates are given by

$$\hat{\sigma}_{RSS,J}^2 = fpc \frac{(n-1)^2}{n^2} \sum_{h=1} \sum_{i=1} \left( \bar{Y}_{RSS,w}^{(-[h]i)} - \bar{Y}_{RSS,w}^{-([.].)} \right)^2$$

$$\hat{\sigma}_{JPS,J}^2 = fpc \frac{(n-1)^2}{n^2} \sum_{i=1}^n \left( \bar{Y}_{JPS,w}^{(-i)} - \bar{Y}_{JPS,w}^{(.)} \right)^2$$

$$fpc = \begin{cases} 1 - \frac{n}{N-1} & \text{SWOR selection} \\ 1 & \text{SWR selection.} \end{cases}$$

where  $fpc$  is the finite population correction factor,  $\bar{Y}_{RSS,w}^{-([.].)} = \frac{1}{dH} \sum_{h=1}^H \sum_{i=1}^d \bar{Y}_{RSS,w}^{(-[h]i)}$  and  $\bar{Y}_{JPS,w}^{(.)} = \frac{1}{n} \sum_{i=1}^n \bar{Y}_{JPS,w}^{(-i)}$ . In the jackknife variance estimates, we used the coefficient  $(n-1)^2/n^2$  since this coefficient provides smaller bias than the usual coefficient  $(n-1)/n$ , (Ozturk and Kravchuk, 2021a,b).

An approximate  $(1-\alpha)100\%$  confidence interval for multi-ranking RSS and JPS designs can be constructed using the jackknife variance estimates:

$$\bar{Y}_{RSS,w} \pm t_{1-\alpha/2, n-H} \hat{\sigma}_{RSS,J}$$

$$\bar{Y}_{JPS,w} \pm t_{1-\alpha/2, n-H} \hat{\sigma}_{JPS,J}$$

In the next section, we compare the RSS and JPS estimators in terms of their efficiencies and coverage probabilities for a varying degree of ranking quality and different set sizes.

## 4 Comparison of RSS and JPS Designs

We performed a simulation study to investigate the contrasting features of RSS and JPS estimators. In the simulation study, samples were generated from two finite populations with large population size  $N = n + 1000$  and small population sizes  $N = nH + 50$ . We considered a normal,  $N(\mu = 50, \sigma = 5)$ , and a lognormal,  $LN(\mu = 0, \sigma = 1)$ , distribution. The population values of the response variable  $Y$  were generated using the quantile functions:

$$y_i = F_N^{-1}(i/(N+1), \mu = 50, \sigma = 5), \text{ and } y_i = F_{LN}^{-1}(i/(N+1), \mu = 0, \sigma = 1), i = 1, \dots, N,$$

where  $F_N^{-1}(y, \mu, \sigma)$  and  $F_{LN}^{-1}(y, \mu, \sigma)$  are the inverse cumulative distribution functions of a normal distribution with location parameter  $\mu$  and scale parameter  $\sigma$  and lognormal distribution with scale parameter  $exp(\mu)$  and shape parameter  $\sigma$ , respectively. The samples were generated using SWR and SWOR selection procedures for both population sizes  $N = n + 1000$  and  $N = nH + 50$ . The

quality of ranking was modeled using a ranking variable  $X$ , such that  $X = Y + \tau\epsilon$ , where  $\epsilon$  has a normal distribution with mean zero and variance 1 and independent of  $Y$ . The correlation coefficient between  $X$  and  $Y$  is given by  $\rho = \frac{1}{\sqrt{1+\tau^2/\sigma^2}}$ . The values of  $\rho$  were selected to be 0.01, 0.25, 0.5, 0.75, 0.9, 1 where values less than 1 will result in imperfect ranking. For the normal distribution, we fixed the sample size at  $n = 36$  and varied the set sizes as  $H = 2, 3, 4, 6, 12$  to explore the impact of different set sizes on the RSS and JPS designs. We purposely selected a smaller sample size  $n = 36$  to evaluate the approximation of the coverage probabilities of the confidence intervals to the nominal coverage probability 0.95. The simulation size is taken to be 5000. An R-package *RankedSetSampling* (Ozturk et al., 2021) is used to compute the estimators and construct confidence intervals. The package is available to download at <https://biometryhub.github.io/RankedSetSampling>.

We first investigate the efficiencies of the RSS and JPS estimators. The relative efficiencies are defined as the ratio of the mean square errors of the RSS and JPS estimators:

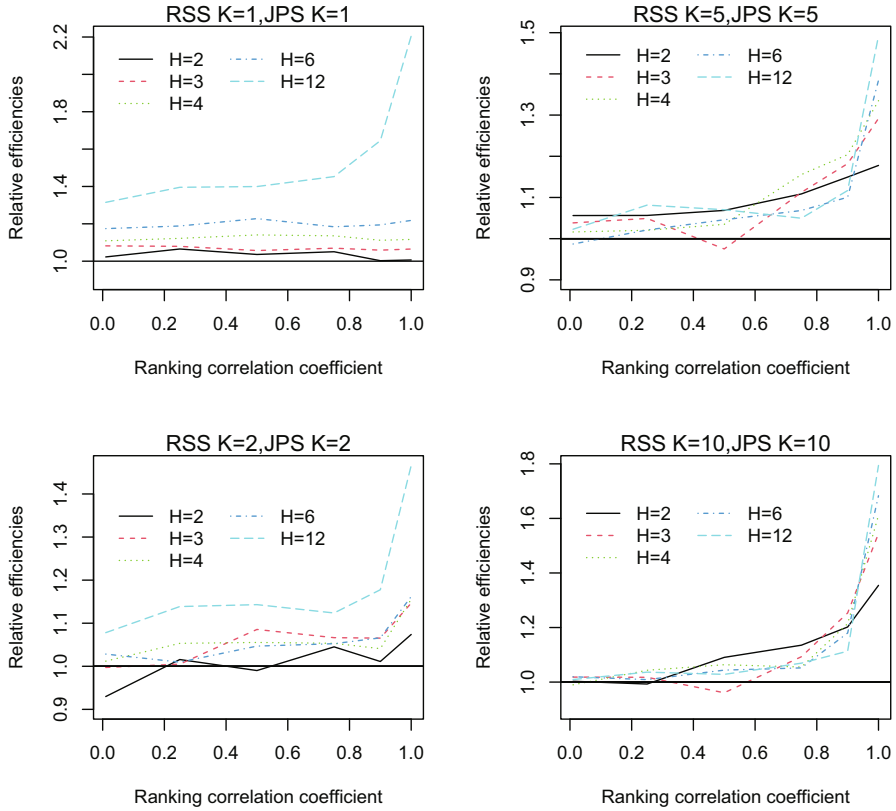
$$RE = \frac{MSE(JPS)}{MSE(RSS)}.$$

A value of  $RE$  greater than 1 indicates that the RSS estimator is more efficient than the JPS estimator. Figure 1 presents the relative efficiencies for the population size  $N = n + 1000$  when samples were generated using the SWR selection procedure. The set sizes and the number of ranking methods are indicated in the legend on each panel. The first panel shows the relative efficiency curves when both RSS and JPS were generated with just one ranking method  $K = 1$ . It is clear in this case that the RSS estimator is more efficient. The efficiency gain is minimal for  $H = 2, 3$ , moderate for  $H = 4, 6$ , and substantial for  $H = 12$ . This intuitively makes sense since large set sizes lead to many judgment groups having no measured observations in a JPS. Empty ranking groups inflate the variance of the JPS estimator. The  $RE$  values are similar to each other for all  $\rho$  values when  $H = 2, 3, 4, 6$ , except for  $\rho$  when  $H = 12$  where it increases.

Figure 1 also presents the relative efficiencies in three different panels when different number of ranking methods ( $K=2, 5, 10$ ) is used. Comparing these panels with panel 1, one can see that the gain in  $RE$  values decreases with the number of ranking methods  $K$ . For example, the  $RE$  values in panel 1 ( $K = 1$ ) are around 1.4 when  $\rho < 0.75$ , and it reduces essentially to 1 in panel 4 ( $K=10$ ). Similar observation can be made in panel 2 ( $K = 2$ ) and panel 3 ( $K = 5$ ). Under perfect ranking, RSS is still superior to the JPS for all set sizes.

Figure 2 presents the efficiency curves for the small population size,  $N = nH + 50$ . In this part of the simulation, both the RSS and JPS were generated under the SWOR selection procedures. The efficiency results are similar to those in Fig. 1, with the key difference being that the  $RE$  curves are higher (lower) in Fig. 2 than in Fig. 1 when  $K = 1$  and  $K = 2$  ( $K = 5$  and  $K = 10$ ).

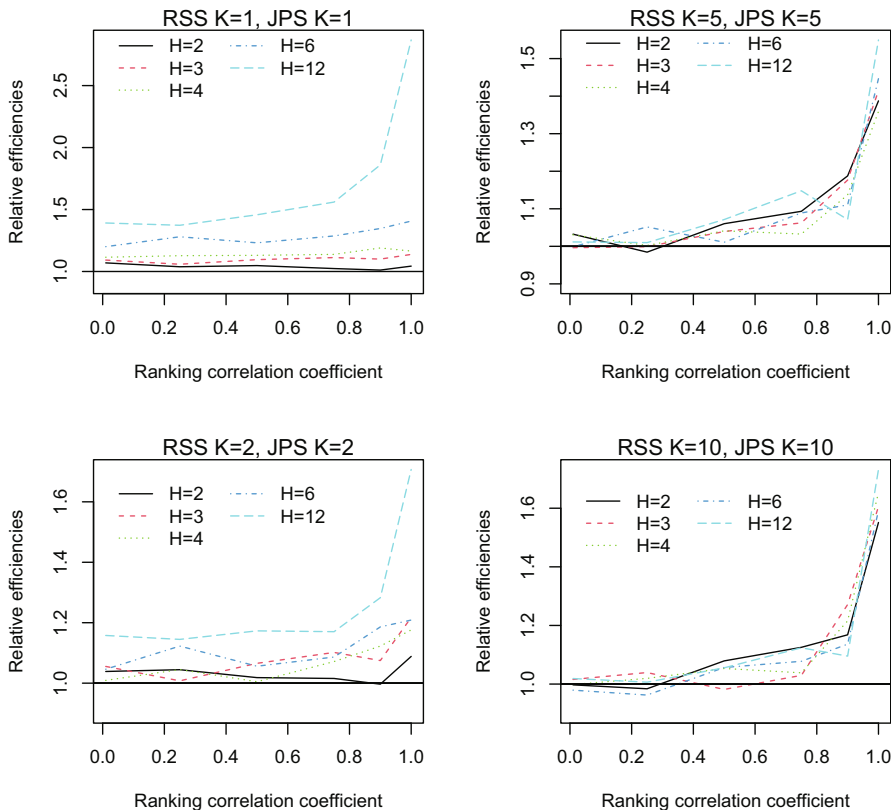
These efficiency results indicate that RSS estimator is more efficient than the JPS estimator when the number of ranking methods is small ( $K = 1, 2$ ). For the larger



**Fig. 1** Efficiency comparison of RSS and JPS designs under SWR selection for large-sized normal distribution population

number of ranking methods ( $K = 5, 10$ ), difference between the efficiency gain of RSS and JPS estimators diminishes.

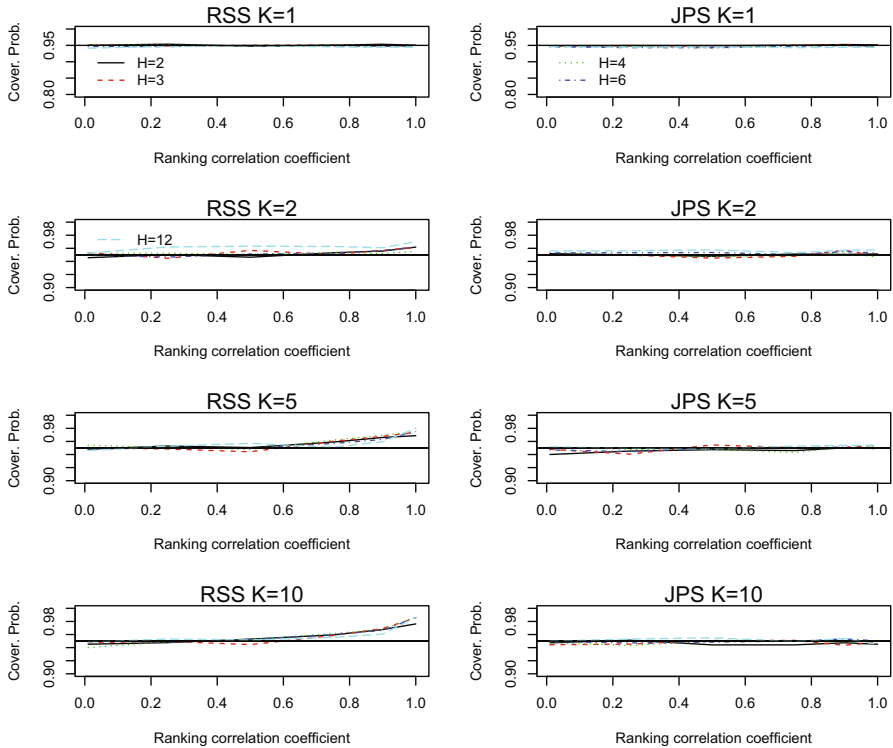
We also investigated the coverage probabilities of the confidence intervals for the population mean. Figure 3 presents the coverage probabilities for the samples constructed with replacement from the population of size  $N = n + 1000$ . We note that confidence intervals are constructed using unbiased variance estimates when  $K = 1$ . For  $K \neq 1$ , we used the jackknife variance estimates. The panels in the first and second columns of Fig. 3 present the coverage probabilities of RSS and JPS confidence intervals for  $K = 1, 2, 5, 10$ , respectively. The coverage probabilities of RSS confidence intervals can be seen to be reasonably close to the nominal coverage probability of 0.95 when  $\rho \leq 0.75$  and  $K = 2, 5, 10$ , but they are slightly larger when  $\rho > 0.75$  and  $K = 2, 5, 10$ . The coverage probabilities in the second column of Fig. 3 are reasonably close to the nominal coverage probability 0.95 for all  $\rho$  and  $K$  values in the simulation study.



**Fig. 2** Efficiency comparison of RSS and JPS designs under SWOR selection for the small-sized normal distribution population

Figure 4 presents the coverage probabilities for the population size  $N = nH + 50$ . In this case, coverage probabilities are again close to nominal coverage probability of 0.95 under imperfect ranking ( $\rho \leq 0.75$ ) for both RSS and JPS and  $K = 1, 2, 5, 10$ . Unlike Fig. 3, coverage probabilities are slightly inflated for both RSS and JPS confidence intervals when  $\rho > 0.75$  and  $K = 2, 5, 10$ . Under perfect ranking ( $\rho = 1$ ), jackknife variance estimator overestimates the variances of the RSS and JPS estimators and leads to a larger coverage probability than the nominal coverage probability of 0.95

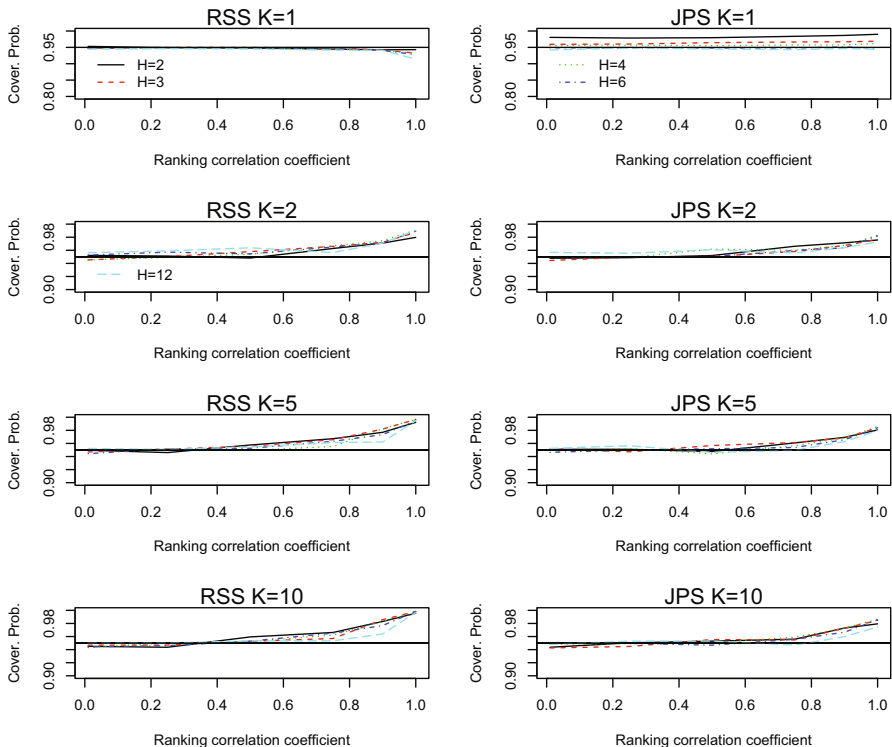
In the second part of the simulation study, we generated samples from the lognormal distribution with the scale parameter  $exp(\mu)$  ( $\mu = 0$ ) and the shape parameter  $\sigma = 1$ . The sample and set sizes were as previously  $n = 48$  and  $H = 2, 3, 4, 6, 12$ . All the other simulation parameters remained the same. The lognormal distribution is strongly positively skewed. For this reason, we increased the sample size from 36 to 48. Figures 5 and 6 present the relative efficiencies of the RSS and JPS estimators for large and small population sizes, respectively. The pattern



**Fig. 3** Coverage probabilities of the jackknife confidence intervals under SWR selection for normal distribution

of the efficiency curves is very similar to that for the normal population. The main difference is in the magnitude of the efficiency gain. The efficiency curves reach to higher values for the normal distribution. This result is consistent with the efficiency results of ranked set samples in (McIntyre, 1952, 2005). McIntyre reported that the efficiencies are higher for symmetric distributions (highest for the uniform distribution) and decrease with skewness. Since the lognormal distribution has strong skewness, the efficiencies are slightly lower than for the normal distribution.

Figures 7 and 8 present the coverage probabilities of the jackknife confidence intervals of the population mean for the SWR and SWOR designs, respectively. It is clear that the coverage probabilities for the lognormal distribution are lower than the nominal coverage probability 0.95. The SWOR selection provides a better coverage probability than the SWR selection. Since a jackknife confidence interval relies on the normal approximation, the sample size  $n = 48$  is not large enough for a good approximation when the underlying population is strongly skewed.

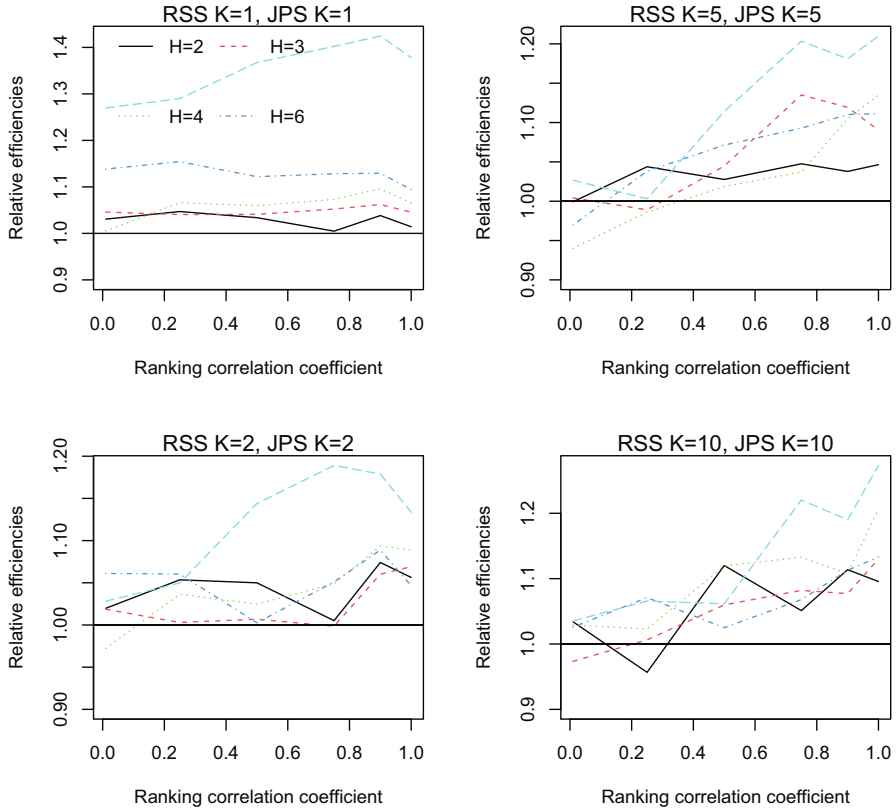


**Fig. 4** Coverage probabilities of the jackknife confidence intervals under SWOR selection for normal distribution

## 5 Application

In this section, we use a real-life finite population example to compare the JPS and RSS estimators. The population consisted of 350 grapevine plants at Coombe vineyard at the University of Adelaide, Waite campus, Australia. The vineyard is used as a research and teaching facility. There are eight different rootstocks originally planted, on which Shiraz is grafted. These rootstocks are popular commercial choices in South Australia. The standard vineyard management of this population requires the monitoring and measuring of certain characteristics of vine plants. In this paper, we consider seven characteristics;  $X_1$ , trunk circumference (cm) in 2018;  $X_2$ , trunk circumference (cm) in 2019;  $X_3$ , shoot counts;  $X_4$ , total shoots;  $X_5$ , pruning weight (kg);  $X_6$ , cordon length (cm); and  $X_7$ , total bunch numbers and  $Y$ , nett fruit weight in 2019 (kg). Our interest was in the estimation of the mean nett fruit yield of this population of grapevines in 2019. The variables  $X_i, i = 1, \dots, 7$ , were used as ranking variables in comparison sets, and hence, the number of ranking methods is  $K = 7$ . There were missing values on some vines, and after removing the plants having missing observations, the population size was reduced to  $N = 309$ . In



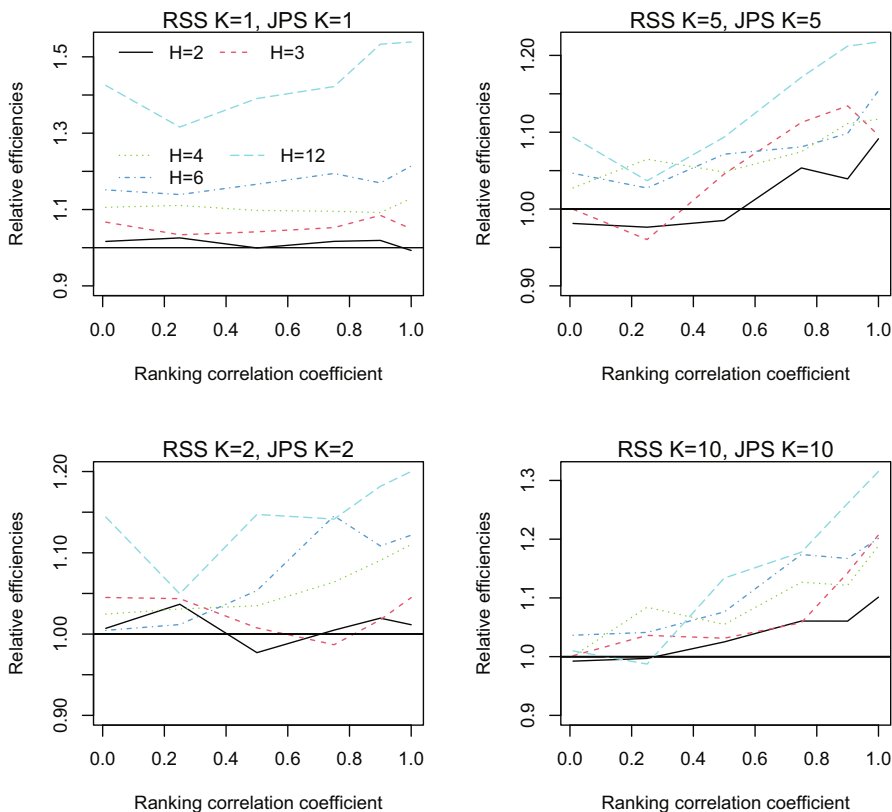


**Fig. 5** Efficiency comparison of RSS and JPS designs under SWR selection for the large-sized lognormal distribution population

this population, the correlation coefficients between  $Y$  and  $X_i$ ,  $\rho_i = cor(Y, X_i)$  are  $\rho_1 = 0.240$ ,  $\rho_2 = 0.191$ ,  $\rho_3 = 0.310$ ,  $\rho_4 = 0.321$ ,  $\rho_5 = 0.172$ ,  $\rho_6 = 0.274$ , and  $\rho_7 = 0.713$ . The mean and standard deviation of the  $Y$  variable are 10.558 kg and 3.855 kg, respectively.

We performed another simulation study using these 309 vine plants. In each replication of the simulation study, we generated the single-ranking judgment post-stratified and ranked set samples with the ranking variable  $X_7$  ( $K = 1$ ), the multi-ranking judgment post-stratified and ranked set samples with  $X_k$ ,  $k = 1, \dots, 7$  ( $K = 7$ ), and a simple random sample. The sample sizes were selected to be  $n = 30$  and 48. For the sample size  $n = 30$ , the set sizes were chosen  $H = 3, 5, 6, 10$ . For the sample sizes  $n = 48$ , the set sizes were  $H = 3, 4, 6$ . Samples were generated using the SWR and SWOR selection procedures. The simulation size was 5000.

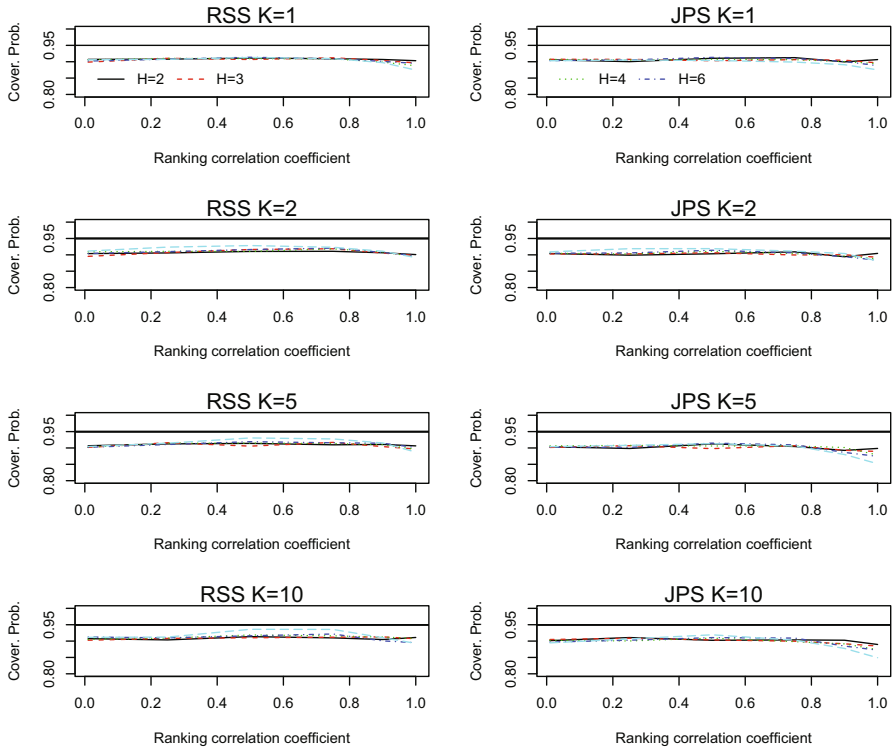
Table 3 presents the relative efficiency of the multi-ranking RSS estimator ( $K = 7$ ) with respect to the other four estimators: the JPS estimator with  $K = 7$  and



**Fig. 6** Efficiency comparison of RSS and JPS designs under SWOR selection for the small-sized lognormal distribution population

$K = 1$ , the SRS estimator, and the RSS estimator with  $K = 1$ . When the entries in Table 3 are greater than one, the multi-ranking RSS estimator with  $K = 7$  was superior. The other efficiency results can be obtained by taking the ratio of any two efficiency columns in Table 3. For example, the efficiency of the JPS estimator with  $K = 1$  relative to the SRS estimator can be obtained by taking the ratio of column 6 and column 5. When  $n = 30$ ,  $H = 3$ , and the replacement is true, this efficiency is calculated  $1.246(1.321/1.060 = 1.246)$ . The other relative efficiencies can be computed in a similar fashion.

All entries in Table 3 are greater than one which indicates that the RSS multi-ranking estimator with  $K = 7$  is more efficient than JPS and SRS estimators. The efficiencies of RSS estimator with  $K = 7$  with respect to JPS and SRS estimators increase with set sizes, but remain relatively constant with RSS estimator with  $K = 1$  (column 7). The reason for this is that the correlation coefficient between ranking variable  $X_7$  and response is 0.729, while the other correlation coefficients are all less than 0.321. Hence, the improvement of ranking quality due to ranking variables with



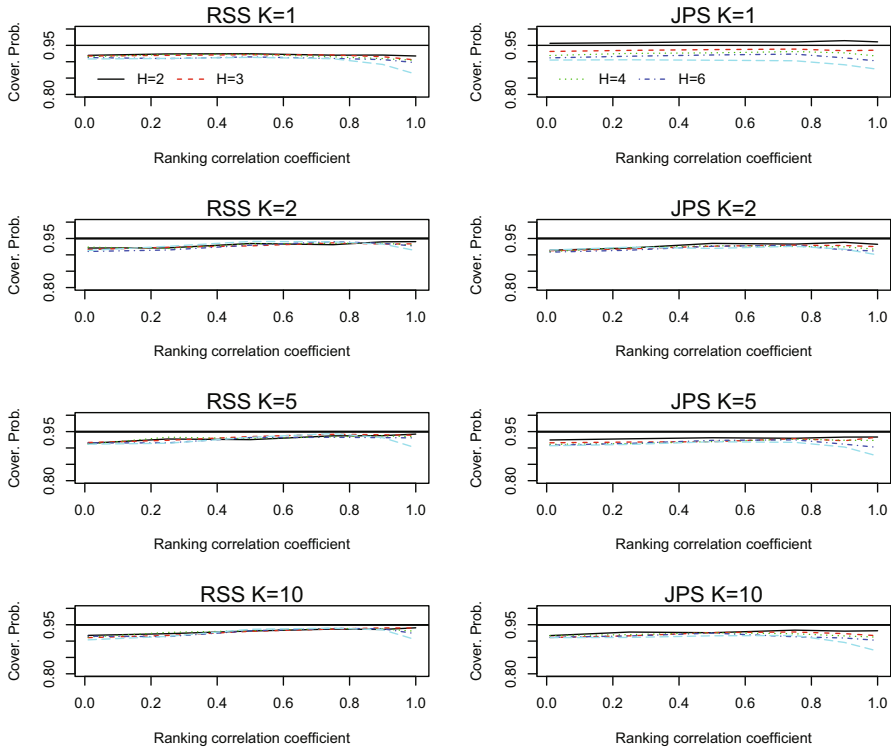
**Fig. 7** Coverage probabilities of the jackknife confidence intervals under SWR selection for large-sized lognormal distribution population

low correlation coefficients is minimal, and the relative efficiency for multi-ranker estimator remains relatively constant. For this particular population and ranking methods, the JPS estimators are more efficient than the SRS estimator and less efficient than multi-ranker RSS estimator.

We also computed the coverage probabilities of the confidence intervals based on the judgment post-stratified, simple random, and ranked set samples for the population mean. All coverage probabilities were reasonably close to the nominal coverage probability of 0.95. Due to space considerations, these empirical coverage probabilities are not reported here.

## 6 Concluding Remarks

Field research is expensive and time-consuming, particularly in natural environments where variables are difficult to control. If auxiliary variables are available, they can be used to account in the analysis for the inherent variation among sampling



**Fig. 8** Coverage probabilities of the jackknife confidence intervals under SWOR selection for small-sized lognormal distribution population

units. These auxiliary variables can be used as blocking variables if they can be evaluated in an objective manner. In certain settings, auxiliary variables may not be assessed accurately. Their assessment may be rough, imprecise, and subjective, but still helpful for ordering the units in a small set independently of knowing the actual values of the variable of interest.

Ranked set and judgment post-stratified sampling designs use this ordering information to construct samples that are more likely to span the full range of values in the population. It has been established in the literature that a ranked set sample is generally more efficient than a judgment post-stratified sample. However, RSS designs induce a strong structure in the sample. Hence, an RSS cannot be analyzed with the inferential procedures developed for an SRS design.

The JPS design may be less efficient than the RSS design, but the sample constructed can be reduced to a simple random sample, allowing the flexibility to perform multiple analyses of various responses on the same data set. This becomes useful if the data set is needed for a multipurpose study. In this paper, we show how to reduce the efficiency loss of a JPS with respect to an RSS by constructing multiple ranks for the response variable on each measured unit. Hence, the JPS

**Table 3** Relative efficiency of the weighted RSS estimators with  $K = 7$  for mean fruit yield of vine plants in Coombe vineyard. Entries greater than one indicate that the RSS estimator with  $K = 7$  is more efficient

Replace	$n$	$H$	JPS		SRS	RSS
			$K = 7$	$K = 1$		$K = 1$
True	30	3	1.209	1.060	1.321	1.018
	30	5	1.283	1.203	1.433	1.020
	30	6	1.335	1.261	1.500	1.016
	30	10	1.508	1.473	1.724	1.011
	48	3	1.227	1.042	1.352	1.022
	48	4	1.270	1.042	1.419	1.026
	48	6	1.403	1.155	1.603	1.030
False	30	3	1.287	1.116	1.407	1.026
	30	5	1.390	1.224	1.560	1.025
	30	6	1.399	1.313	1.575	1.024
	30	10	1.633	1.619	1.860	1.003
	48	3	1.287	1.075	1.426	1.028
	48	4	1.317	1.098	1.475	1.028
	48	6	1.533	1.261	1.757	1.030

design provides the flexibility for multipurpose analysis at the expense of little efficiency loss with respect to a balanced ranked set sample. Another advantage of the JPS design is that it is relatively straightforward to construct a multi-ranking JPS even when there are no additional auxiliary ranking variables, and this can be done by permuting the units selected to form comparison sets. This idea is not easily extended to a ranked set sampling. We would recommend that the JPS design should be considered in field sampling, especially for multipurpose studies.

**Acknowledgments** This research is supported by the Australian Grains Research and Development Corporation (GRDC) as part of the Statistics for the Australian Grains Industry project (UA00164). The data was collected by Bachelor of Agricultural Science students of the School of Agriculture, Food and Wine, University of Adelaide, in 2019 under the supervision of Mr. Peter Kasprzak. The data was kindly made available for our purposes and other publications, as well as to improve the management and research decisions in the Coombe vineyard.

## References

Arslan, G., & Ozturk, O. (2013). Parametric inference based on partially rank ordered set samples. *Indian Journal of Statistics*, 51, 1–24.

Bohn, L. L., & Wolfe, D. A. (1992). Nonparametric two-sample procedures for ranked set samples data. *Journal of the American Statistical Association*, 87(418), 552–561.

Bohn, L. L., & Wolfe, D. A. (1994). The effect of imperfect judgment rankings on properties of procedures based on the ranked set samples analog of the Mann-Whitney-Wilcoxon statistic. *Journal of the American Statistical Association*, 89(425), 168–176.

Bouza, C., & Al-Omari, A. I. (2019). *Ranked set sampling: 65 years improving the accuracy in data gathering*. Elsevier.

- Chen, Z., & Bai, Z. (2000). The optimal ranked-set sampling scheme for parametric families. *Sankhyā: The Indian Journal of Statistics, Series A*, 62(2), 178–192.
- Chen, Z., Bai, Z., & Sinha, B. K. (2003). *Ranked set sampling*. Springer.
- Chen, M., Ahn, S., Wang, X., & Lim, J. (2014). Generalized isotonized mean estimators for judgment post-stratification with multiple rankers. *Journal of Agricultural, Biological, and Environmental Statistics*, 19, 405–418.
- Dell, T. R., & Clutter, J. L. (1972). Ranked-set sampling theory with order statistics background. *Biometrics*, 28, 545–555.
- Deshpande, J. V., Frey, J., & Ozturk, O. (2006). Nonparametric ranked set-sampling confidence intervals for a finite population. *Environmental and Ecological Statistics*, 13, 25–40.
- Fligner, M. A., & MacEachern, S. N. (2006). Nonparametric two-sample methods for ranked set sample data. *Journal of the American Statistical Association*, 101(475), 1107–1118.
- Frey, J. (2011). Recursive computation of inclusion probabilities in ranked set sampling. *Journal of Statistical Planning and Inference*, 141, 3632–3639.
- Frey, J., & Feeman, T. G. (2012). An improved mean estimator for judgment post-stratification. *Computational Statistics & Data Analysis*, 56, 418–426.
- Frey, J., & Feeman, T. G. (2013). Variance estimation using judgment post-stratification. *Annals of the Institute of Statistical Mathematics*, 5, 551–569.
- Frey, J., & Ozturk, O. (2011). Constrained estimation using judgment post-stratification. *Annals of the Institute of Statistical Mathematics*, 3, 769–789.
- Hatefi, A., Jafari Jozani, M., & Ziou, D. (2014). Estimation and classification for finite mixture models under ranked set sampling. *Statistica Sinica*, 24, 675–698.
- Hatefi, A., Jafari Jozani, M., & Ozturk, O. (2015). Mixture model analysis of partially rank-ordered set samples: Age groups of fish from length-frequency data. *Scandinavian Journal of Statistics*, 42, 848–871.
- Hettmansperger, T. P. (1995). The ranked-set sampling sign test. *Nonparametric Statistics*, 4, 263–270.
- Jafari Jozani, M., & Johnson, B. C. (2011). Design based estimation for ranked set sampling in finite populations. *Environmental and Ecological Statistics*, 18, 663–685.
- Koti K. M., & Babu, J. G. (1996). Sign test for ranked-set sampling. *Communications in Statistics - Theory and Methods*, 25(7), 1617–1630.
- MacEachern, S. N., Stasny, E. A., & Wolfe, D. A. (2004). Judgment post-stratification with imprecise rankings. *Biometrics*, 60, 207–215.
- McIntyre, G. (1952). A method for unbiased selective sampling using ranked set sampling. *Australian Journal of Agriculture Research*, 3, 385–390.
- McIntyre, G. A. (2005). A method of unbiased selective sampling using ranked-sets. *The American Statistician*, 59, 230–232.
- Ozturk, O. (1999). Two-sample inference based on one-sample ranked set sample sign statistics. *Nonparametric Statistics*, 10, 197–212.
- Ozturk, O. (2013). Combining multi-observer information in partially rank-ordered judgment post-stratified and ranked set samples. *The Canadian Journal of Statistics*, 41, 304–324.
- Ozturk, O. (2014). Statistical inference for population quantiles and variance in judgment post-stratified samples. *Computational Statistics and Data Analysis*, 77, 188–205.
- Ozturk, O. (2016a). Statistical inference based on judgment post-stratified samples in finite population. *Survey Methodology*, 42, 239–262.
- Ozturk, O. (2016b). Estimation of finite population mean and total using population ranks of sample units. *Journal of Agricultural, Biological, and Environmental Statistics*, 21, 181–202.
- Ozturk, O. (2017). Statistical inference with empty strata in judgment post stratified samples. *Annals of the Institute of Statistical Mathematics*, 69, 1029–1057.
- Ozturk, O. (2019). Statistical inference using rank based post-stratified samples in a finite population. *Test*, 28, 1113–1143.
- Ozturk, O., & Bayramoglu Kavlak, K. (2018). Model based inference using ranked set samples. *Survey Methodology*, 44(1), 1–16, Catalogue No. 12-001-X.

- Ozturk, O., & Bayramoglu Kavlak, K. (2019). Statistical inference using stratified ranked set samples from finite populations. In: C. Bouza & A. I. Al-Omari (Eds.), *Ranked set sampling: 65 years improving the accuracy in data gathering* (pp 157–170). Elsevier.
- Ozturk, O., & Bayramoglu Kavlak, K. (2020). Statistical inference using stratified judgment post-stratified samples from finite populations. *Environmental and Ecological Statistics*, 27, 73–94.
- Ozturk, O., & Demirel, N. (2016). Estimation of population variance from multi-ranker ranked set sampling designs. *Communications in Statistics-Simulation and Computation*, 45(10), 3568–3583.
- Ozturk, O., & Jafari Jozani, M. (2014). Inclusion probabilities in partially rank ordered set sampling. *Computational Statistics and Data Analysis*, 69, 122–132.
- Ozturk, O., & Kravchuk, O. (2021a) Combining ranking information from different sources in ranked set samples. *Canadian Journal of Statistics*. <https://doi.org/10.1002/cjs.11656>
- Ozturk, O., & Kravchuk, O. (2021b). Judgment post-stratified assessment combining ranking information from multiple sources, with a field phenotyping example. *Journal of Agricultural, Biological and Environmental Statistics*. <https://doi.org/10.1007/s13253-021-00439-1>
- Ozturk, O., Rogers, S., Kravchuk, O., & Kasprzak, P. (2021). RankedSetSampling: Easing the application of ranked set sampling in practice. R package version 0.0.1. <https://biometryhub.github.io/RankedSetSampling/>
- Patil, G. P., Sinha, A. K., & Tallie, C. (1995). Finite population corrections for ranked set sampling. *Annals of the Institute of Statistical Mathematics*, 47, 621–636.
- Stokes, L. (1995). Parametric ranked set sampling. *Annals of the Institute of Statistical Mathematics*, 47, 465–482.
- Stokes, S. L., Wang, X., & Chen, M. (2007). Judgment post-stratification with multiple rankers. *Journal of Statistical Theory and Applications*, 6, 344–359.
- Takahasi, K., & Wakimoto, K. (1968). On unbiased estimates of the population mean based on the sample stratified by means of ordering. *Annals of the Institute of Statistical Mathematics*, 20, 1–31.
- Wang, X., Stokes, S. L., Lim, J., & Chen, M. (2006). Concomitants of multivariate order statistics with application to judgment post-stratification. *Journal of the American Statistical Association*, 101, 1693–1704.
- Wang, X., Lim, J., & Stokes, S. L. (2008). A nonparametric mean estimator for judgment post-stratified data. *Biometrics*, 64, 355–363.
- Wang, X., Wang, K., & Lim, J. (2012). Isotonized CDF estimation from judgment poststratification data with empty strata. *Biometrics*, 68, 194–202.