Estimation of a Finite Population Mean and Total Using Population Ranks of Sample Units

Omer OZTURK

This paper introduces new estimators for population total and mean in a finite population setting, where ranks (or approximate ranks) of population units are available before selecting sample units. The proposed estimators require selecting a simple random sample and identifying the population ranks of sample units. Selection of the sample can be performed with- or without-replacement. The population ranks of the selected units of with-replacement samples are determined among all population units. On the other hand, the ranks of the sample units of without-replacement samples are identified in two different ways: (1) The rank of a sample unit is determined sequentially among the remaining population units after excluding all previously ranked sample units from the population; (2) The ranks are determined among all units in the population. By conditioning on these population ranks, we construct a set of weighted estimators, develop a bootstrap re-sampling procedure to estimate the variances of the estimators, and construct percentile confidence intervals for the population mean and total. We show that the new estimators provide a substantial amount of efficiency gain over their competitors. We apply the proposed estimators to estimate corn production in one of the counties in Ohio.

Key Words: Finite population; Horvitz–Thompson estimator; Inclusion probabilities; Coefficient of variation; Ranked set sampling; Judgment post stratified sampling.

1. INTRODUCTION

In many survey sampling studies, in addition to the variable of interest, researchers often have additional auxiliary information to improve statistical inference. In many instances, this auxiliary information may not be accurate, cannot be turned into a numerical covariate, or may be even subjective. Even though it contains valuable information, use of this type of information is ignored in practice since it may require strong modeling assumptions. For example, ratio and regression estimators are constructed based on auxiliary variables under strong modeling assumptions. This paper uses rough information to provide ranks

---

Omer Ozturk (✉) Department of Statistics, The Ohio State University, Columbus, OH 43210, USA
(E-mail: *omer@stat.osu.edu*).

(or approximate ranks) for the population units to improve the information content of the sample by avoiding strong modeling assumptions.

The auxiliary variables are very common in survey sampling studies. Hence, finite population settings provide a natural platform to obtain exact or approximate ranks of all population units. Ranking process can be achieved through the use of auxiliary variables, such as, size of sampling units, previous survey outcomes, census tracks, etc. One such setting is given in Husby et al. (2005), namely, The United States Department of Agriculture's (USDA) National Agricultural Statistics (NASS) county crop estimation program. This program samples farms across the United States from the sampling frames that include obvious auxiliary variables, such as acreage in the farm, size of the farm, etc. These auxiliary variables provide a reasonable mechanism to rank the farms based on their crop productions. The detailed description of the USDA/NASS county estimation program can be found in Iwig (1993).

In an infinite population setting, use of subjective information has generated extensive research interests in judgment post stratified (JPS) and ranked set sampling (RSS) designs. Readers can find the recent research activities and detailed description of JPS sampling designs in MacEachern et al. (2004), Frey and Feeman (2012, 2013), Frey and Ozturk (2011), Stokes et al. (2007), Wang et al. (2006, 2008, 2012), and Ozturk (2014a, 2014b). Ranked set sampling design is originally developed to keep the overall cost of data collection minimal in estimating mean pasture yield in agricultural fields in an infinite population setting by McIntyre (1952, reprinted in 2005). In recent years, there has been a surge in research in RSS sampling designs both in finite and infinite population settings. A tiny slice of literature in finite population setting includes Patil et al. (1995), Jafari Jozani and Johnson (2011, 2012), Frey (2011), Gokpinar and Ozdemir (2010), Ozturk and Jafari Jozani (2013), Al-Saleh and Samawi (2007), Ozdemir and Gokpinar (2008) and Ozturk (2014c). A comprehensive up-to-date literature review both in JPS and RSS can be found in a recent review paper in Wolfe (2012).

Both RSS and JPS sampling designs use ranking information from a few units in a set, not from the entire population, to divide the data into homogeneous groups of judgment strata. This ranking process is subjective and does not require strong modeling assumptions. It only needs a consistent ranking scheme to create ranks for the units in a set without requiring an established standard of measurement. On the other hand, RSS and JPS use ranks locally in a set and ignores the global ranking information in the population. This paper, unlike JPS and RSS sampling designs, concentrates on global ranking information in the entire population and creates informative samples.

The paper considers finite population settings and assumes the ranks of population units are available before sampling. It selects a simple random sample (SRS) of n units from a finite population of size N either with- or without-replacement. For each selected unit in the sample, we measure the characteristic of interest along with its population rank. If the sample is selected with-replacement, the rank of the selected unit is determined from the entire population including all population units. If the sample is taken without-replacement, the population ranks can be determined in two different ways. In the first approach, before selecting the sample, ranks are assigned in the entire population including all units, and then a simple random sample without replacement is selected along

with the population ranks of the sample units. In the second approach, ranks are determined sequentially. The rank of a sample unit is determined among the remaining population units by excluding all the previously selected sample units from the population. Even though these two approaches yield without-replacement SRS samples, they create different ranking structures in the sample. Hence, replacement policies and ranking structures lead to three different designs: design-0 , design-1, and design-2. Designs-0 selects the sample with replacement and assigns the ranks in the entire populations; design-1 selects the sample without-replacement and assigns the ranks sequentially; and finally, the design-2 selects the sample without-replacement and assigns the ranks in the entire population.

Ranking structures in these sampling designs provide a lot of information about the exact (or approximate) population location of the sample units. This location information can be used to borrow additional information from other unmeasured population units to improve the information content of the sample. For each measured unit, we consider selecting additional $H - 1$ unmeasured units without replacement from the remaining population units to form a set of size H. The relative position of each measured unit in these sets can be computed through its rank in the set by conditioning on its population rank. The rank of the measured unit in a set of size H yields a discrete conditional probability distribution given the population rank of the same measured unit. These conditional probabilities of within-set ranks further provide a mechanism to compute conditional inclusion probabilities for the population units. The final sample in this process consists of three pieces of information: measured values, conditional ranking probabilities, and the conditional inclusion probabilities. Even though the measured observations form an SRS sample in design-0, design-1, and design-2, the conditional probabilities are different. Hence, these designs show different characteristics.

Section 2 provides detailed developments for the construction of design-0, design-1, and design-2. For each design, we construct a probability distribution for the approximate location of the measured units among the unmeasured population units in a set of size H and compute the first-order conditional inclusion probabilities of the population units given the population ranks of a sample. Section 3 uses ranking information and conditional inclusion probabilities to construct estimators for the population mean and total. Section 4 provides empirical evidence to compare the proposed estimators with its competitors. Section 5 develops a bootstrap re-sampling procedure to estimate the variance of the estimators and to construct percentile confidence intervals for the population mean and total. Section 6 applies the proposed estimators to USDA 1992 Ohio corn data. Finally, Sect. 7 provides a concluding remark.

2. CONSTRUCTION OF SAMPLING DESIGNS

Consider a finite population of N units labeled as $\mathcal{P} = \{u_1, \ldots, u_N\}$. For each population unit u_i, we assume that its population rank s_i, $1 \leq s_i \leq N$, is known. If the population ranks are not known, we assume that there exists an auxiliary variables Y highly correlated with the variable of interest X. We then estimate the population rank of X from the population

rank of the auxiliary variable Y. In the remaining part of the paper, we use s_i to denote the true population rank of the unit u_i whether it is estimated or not.

Design-0 We select a simple random sample, $U = \{u_{s_1}, \ldots, u_{s_n}\}$, with-replacement from $\mathcal{P}$ and measure all of them for the variable of interest X, $X = (X_1, \ldots, X_n)$. We then identify the population ranks of the measured units, $S = \{s_1, \ldots, s_n\}$, where s_j is the population rank of X_j. Our sample then consists of n measurements and population ranks that correspond to these measured units:

$$X_S = \{X_j, s_j\}, j = 1, \ldots, n. \tag{1}$$

We note that since the sample is selected with replacement, some pair (X_i, s_i) may appear more than once in the sample. If we ignore the ranks in Eq. (1), X_S becomes a simple random sample, and the inference can be developed based on standard theory in finite population setting. Since S contains the true (or estimated) population ranks of the sample units, it provides a lot of information about the approximate location (in relative sense) of the sample units in the population. This location indicator allows us to borrow additional information from another $H - 1$, $1 \leq H \leq N - 1$, unobserved population units without a measurement. To achieve this goal, for each selected unit u_{s_j}, we consider selecting $H - 1$ additional units at random without-replacement from the remaining $N - 1$ population units and form a set of size H

$$U_{j,H} = \{u_{s_j}, u_{t_1}, \ldots, u_{t_{H-1}}\}, t_h \neq s_j, h = 1, \ldots, H - 1, \tag{2}$$

where u_{t_h} is the t_h-th smallest unit among N population units. Let R_{s_j} be the rank of X_j, random variable obtained from u_{s_j}, in the set $U_{j,H}$. The conditional probability that R_{s_j} is equal to h given that X_j is the s_j-th smallest unit in the population can be computed by

$$\alpha^{(0)}(h|s_j) = P(R_{s_j} = h|X_j = x_{s_j}) = \frac{\binom{s_j-1}{h-1}\binom{N-s_j}{H-h}}{\binom{N-1}{H-1}}. \tag{3}$$

The above expression shows that the rank of random variable X_j in a set of size H has a conditional probability distribution over integers $(1, \ldots, H)$ given that X_j is the s_j-th smallest unit in the population. This conditional distribution helps us to borrow information from additional $H - 1$ unmeasured units in the population in addition to the information each measured unit has in the sample.

We now look at the problem from a different perspective. Instead of treating X as a simple random sample, we treat it as a sample of independent order statistics by conditioning on rank vector $R = \{R_{s_1}, \ldots, R_{s_n}\}$ generated by Eq. (3). It is clear that the conditional distribution of X_j given that it has a rank $R_{s_j} = h_j$ in a set of size H is the same as the h_j-th order statistics, $X_{(h_j)} \overset{D}{=} X_j|R_{s_j} = h_j$. Let $X_{H|S} = (X_{(h_j)}, \ldots, X_{(h_n)})$ be the n order statistics based on this conditional distribution.

Let $\beta^{(0)}(i, h|s_j)$ be the probability that the h-th order statistics in set $U_{j,H}$ equals to the i-th smallest unit in the population given that $R_{s_j} = h$. This conditional probability can be computed from

$$\beta^{(0)}(i|h, s_j) = P(X_{(R_{s_j})} = x_i | R_{s_j} = h) = \frac{\binom{i-1}{h-1}\binom{N-i}{H-h}}{\binom{N}{H}}.$$

One can interpret $\beta^{(0)}(i|h, s_j)$ as the probability mass function of the h-th order statistics given that the rank of X_j equals to h, $R_{s_j} = h$. By using $\alpha^{(0)}(i, h|s_j)$ and $\beta^{(0)}(i, h|s_j)$, we obtain the conditional probability that random variable X_j equals to the i-th smallest value in the population given the population rank (s_j) of X_j

$$\beta^{(0)}(i|s_j) = P(X_j = x_i|s_j) = \sum_{h=1}^{H} \beta^{(0)}(i|h, s_j)\alpha^{(0)}(h|s_j).$$

Then the conditional inclusion probability of the i-th population unit in the sample $X_{H|S}$ is given by

$$\pi^{(0)}(i|S) = 1 - \prod_{j=1}^{n}(1 - \beta^{(0)}(i|s_j)).$$

Note that $\pi^{(0)}(i|S), i = 1, \ldots, N$, are not the inclusion probabilities of the sample X. They are the inclusion probabilities of the sample $X_{H|S}$. Since i is arbitrary, we can replace it with s_j. In this case, $\pi^{(0)}(s_j|S)$ would be the probability that population unit u_{s_j} in sample X would be included in sample $X_{H|S}$.

Remark 1. If either $H = 1$ or $H > 1$ and $\alpha^{(0)}(h|s_j) = 1/H$ for $h = 1, \ldots, H$, then $\beta^{(0)}(i|s_j) = 1/N$ for $j = 1, \ldots, n$ and $i = 1, \ldots, N$. The inclusion probabilities in these cases reduce to $\pi^{(0)}(i|S) = 1 - (\frac{N-1}{N})^n, i = 1, \ldots, N$.

For design-0, the data structure of the sample will be denoted by

$$\left\{ X_j, \alpha^{(0)}(h|s_j), \pi^{(0)}(s_j|S) \right\}, \quad h = 1, \ldots, H, \quad j = 1, \ldots, n.$$

Design-1 Design-1 selects the sample units without replacement from the population and measures all of them, $X = (X_1, \ldots, X_n)$. Unlike design-0, where ranks in S are computed from all units in the population, the population ranks of the measured units are identified sequentially by removing all of the previously ranked units in the sample from the population. Let s_j^* be the rank of the unit u_j after removing all the previously ranked units from the population. Since the ranks of the selected units are assigned sequentially, we introduce additional notation to accommodate the ranking structure. Let $\mathcal{P}_{-j}$ be the finite population of size $N + 1 - j$ after removing $j - 1$ units from the original population $\mathcal{P}$

$$\mathcal{P}_{-j} = \{\text{All population units excluding the first } j - 1 \text{ units in the sample}\}.$$

Since the population rank of each selected unit is determined after removing all previously selected units from the population, the sample with this new ranking structure becomes

$$U^* = \{u_{s_1^*}, u_{s_2^*}, \ldots, u_{s_n^*}\},$$

where $u_{s_j^*}$ is the j-th selected unit in the sample that has a rank s_j^* in population $\mathcal{P}_{-j}$, and $S^* = \{s_1^*, \ldots, s_n^*\}$ is the set of the ranks obtained from the reduced populations. The expression $s_j^* = s_j - \sum_{k=1}^n I(s_k < s_j)$ provides the connection between s_j^* in the reduced population and s_j in the full population. In this sample, for each selected unit $u_{s_j^*}$, we again construct a set of size H to borrow information from additional $H - 1$ unmeasured units from population $\mathcal{P}_{-j}$:

$$U_{j,H}^* = \{u_{s_j^*}, u_{t_1^*}, \ldots, u_{t_{H-1}^*}\}, \quad j = 1, \ldots, n,$$

where $u_{t_h^*}$ is the t_h^*-th smallest unit in population $\mathcal{P}_{-j}$. Let $R_{s_j^*}$ be the rank of $u_{s_j^*}$ in the set $U_{j,H}^*$. The conditional distribution of $R_{s_j^*}$ given that $s_j^* \in \mathcal{P}_{-j}$ is given by $\alpha^{(1)}(h|s_j^*)$:

$$\alpha^{(1)}(h_j|s_j^*) = P(R_{s_j^*} = h_j|s_j^* \in \mathcal{P}_{-j}) = \frac{\binom{s_j^*-1}{h_j-1}\binom{N-j+1-s_j^*}{H-h_j}}{\binom{N-j}{H-1}}.$$

As in design-0, we again consider a conditional sample $X_{H|S^*}$ given the population ranks S^*. By adopting the notation of design-0, the conditional probability that the h-th order statistics in set $U_{j,H}^*$ equals to the i-th smallest unit in the population $\mathcal{P}_{-j}$ given that $R_{s_j} = h$ is given by

$$\beta^{(1)}(i, |h, s_j^*) = P(X_{(h)} = x_i|R_{s_j^*} = h) = \frac{\binom{i-1}{h-1}\binom{N-j+1-i}{H-h}}{\binom{N-j+1}{H}}, \, x_i \in \mathcal{P}_{-j}.$$

In design 1, the conditional probability that random variable X_j equals to the i-th smallest value in the population $\mathcal{P}_{-j}$ given its population rank (s_j^*) follows from the total probability law over the conditional distribution of rank: $R_{s_j}^*$

$$\beta^{(1)}(i|s_j^*) = P(X_j = x_i|s_j^*) = \sum_{h=1}^H \beta^{(1)}(i|h, s_j^*)\alpha^{(1)}(h|s_j^*), x_i \in \mathcal{P}_{-j}.$$

To compute the conditional inclusion probability of the i-th population unit in the sample $X_{H|S^*}$ given the rank vector S^*, $\pi^{(1)}(i|S^*)$, we use the sequential algorithm given in Frey (2013) in a slightly different context. To compute $\pi^{(1)}(i|S^*)$, we first need to develop some additional notation due to sequential identification of population ranks. For $j = 0, \ldots, i-1$, let $W(j, d)$ be the probability that first d units in the sample include j units smaller than the i-th unit and not the i-th unit. It is obvious that, if $d = 0$, the values of $\{W(j, 0), 0 \le j \le i-1\}$ can be computed from

$$W(j, 0) = \begin{cases} 1, \ j = 0 \\ 0, \ \text{otherwise.} \end{cases}$$

The values of $W(j, d)$ for $d > 0$ can be computed from a recursive relationship between adjacent selection steps. Assume that the values of $\{W(j, d), 0 \le j \le i - 1\}$ are known for

a fixed d. There are then two ways to obtain the values of $\{W(j, d+1), 0 \le j \le i-1\}$ from stage d: (1) There could be j units smaller than the i-th unit among the first d selected units in the sample and the next selected unit in the sample is larger than the i-th unit in the population. (2) There could be $j-1$ units smaller than the i-th unit among the first d selected units in the sample and the next selected unit in the sample is smaller than the i-th unit in the population. These two statements define a recursive equation as follows:

$$W(0, d+1) = T(0, d) \sum_{\lambda=i+1}^{N-d} \beta^{(1)}(\lambda | s_d^*)$$

and for $j = 1, \ldots, i-1$

$$W(j, d+1) = W(j, d) \sum_{\lambda=i+1-j}^{N-d} \beta^{(1)}(\lambda | s_d^*) + W(j-1, d) \sum_{\lambda=1}^{(i-1)-(j-1)} \beta^{(1)}(\lambda | s_d^*).$$

Going through this recursive equation for $d = 1, \ldots, n$, we compute the probability $\{W(j, n), 0 \le j \le i-1\}$. The probability that the i-th unit is not included in the sample is then given by $\sum_{j=0}^{i-1} W(j, n)$. The first-order conditional inclusion probability of the i-th unit given S^* and H is then given by

$$\pi^{(1)}(i | S^*) = 1 - \sum_{j=0}^{i-1} W(j, n), i = 1, \ldots, N$$

Note that even though it is not explicitly stated in the notation, $W(j, d)$ is a conditional probability for given population rank vector S^*. The data structure of design-1 will be denoted by

$$\{X_j, \alpha^{(1)}(h | s_j^*), \pi^{(1)}(s_j^* | S^*)\}, h = 1, \ldots, H, j = 1, \ldots, n$$

Design 2 We select a simple random sample, $X = (X_1, \ldots, X_n)$, of size n without replacement and identify their ranks, $S = (s_1, \ldots, s_n)$, in population $\mathcal{P}$

$$U = \{u_{s_1}, u_{s_2}, \ldots, u_{s_n}\}.$$

To borrow additional information from the unmeasured population units, we select n disjoint sets, each of size $H - 1$. We then randomly match these n sets with selected units in set U to form n sets, each of size H

$$U_{j,H} = \{u_{s_j}, u_{t_1}, \ldots, u_{t_H}\}, j = 1, \ldots, n.$$

The conditional probability distribution of the rank R_{s_j} of X_j in set $U_{j,H}$ given that X_j has the rank s_j in the population $\mathcal{P}$ is given by

$$\alpha^{(2)}(h | s_j) = P(R_{s_j} = h_j | s_j) = \frac{\binom{s_j-1}{h_j-1} \binom{N-s_j}{H-h_j}}{\binom{N-1}{H-1}}, h_j = 1, \ldots, H.$$

In a similar fashion, in the sample $X_{H|S}$, the conditional probability that the h-th order statistics in set $U_{j,H}$ equals to the i-th smallest unit in the population $\mathcal{P}$ given that $R_{s_j} = h$ is given by

$$\beta^{(2)}(i|h, s_j) = P(X_{(h)} = x_i | R_{s_j} = h) = \frac{\binom{i-1}{h-1}\binom{N-i}{H-h}}{\binom{N}{H}}, x_i \in \mathcal{P}.$$

In Eq. (4), the sum over possible values of h yields the conditional probability that X_j equals to i-th smallest unit in the population given the population rank s_j

$$\beta^{(2)}(i|s_j, H) = \sum_{h=1}^{H} \beta^{(2)}(i|h, s_j)\alpha^{(2)}(h|s_j). \tag{4}$$

The first-order conditional inclusion probabilities given the population ranks of the observed measurements then follow from

$$\pi^{(2)}(i|S) = \sum_{j=1}^{n} \beta^{(2)}(i|s_j), i = 1, \dots, N.$$

Finally, the data structure of the sample from design-2 is denoted with

$$\{X_j, \alpha^{(2)}(h|s_j), \pi^{(2)}(s_j|S)\}, h = 1, \dots, H, j = 1, \dots, n.$$

Remark 2. If either $H = 1$ or $H > 1$ and $\alpha^{(2)}(h|s_j) = 1/H$ for $h = 1, \dots, H$, then $\beta^{(2)}(i|s_j) = 1/N$ for $j = 1, \dots, n$ and $i = 1, \dots, N$. The inclusion probabilities in these cases reduce to $\pi^{(2)}(i|S) = n/N, i = 1, \dots, N$.

3. ESTIMATORS FOR POPULATION MEAN AND TOTAL

In this section, we introduce three estimators for population mean and total for each sampling design. The estimators for population total use the data structures established in the previous section:

$$T_1^{(L)} = \begin{cases} \sum_{j=1}^{n} w_1^{(L)}(j|S)X_j & L = 0, 2 \\ \sum_{j=1}^{n} w_1^{(L)}(j|S^*)X_j & L = 1 \end{cases}$$

$$T_2^{(L)} = \begin{cases} \frac{N}{H}\sum_{h=1}^{H}\sum_{j=1}^{n} w_2^{(L)}(h, j|S)X_j & L = 0, 2, \\ \frac{N}{H}\sum_{h=1}^{H}\sum_{j=1}^{n} w_2^{(L)}(h, j|S^*)X_j & L = 1 \end{cases}$$

and

$$T_3^{(L)} = \begin{cases} \frac{N}{H}\sum_{h=1}^{H}\sum_{j=1}^{n} w_3^{(L)}(h, j|S)X_j & L = 0, 2 \\ \frac{N}{H}\sum_{h=1}^{H}\sum_{j=1}^{n} w_3^{(L)}(h, j|S^*)X_j & L = 1 \end{cases}$$

where

$$w_1^{(L)}(j|\boldsymbol{S}) = \frac{1}{\pi^{(L)}(s_j|\boldsymbol{S})}, \quad w_2^{(L)}(h, j|\boldsymbol{S}) = \frac{\alpha^{(L)}(h|s_j)}{\sum_{j=1}^{n}\alpha^{(L)}(h|s_j)},$$

$$w_3^{(L)}(h, j|\boldsymbol{S}) = \frac{\frac{\alpha^{(L)}(h|s_j)}{\pi^{(L)}(s_j|\boldsymbol{S})}}{\sum_{j=1}^{n}\frac{\alpha^{(L)}(h|s_j)}{\pi^{(L)}(s_j|\boldsymbol{S})}}.$$

We note that design-1 estimators always use population ranks $\boldsymbol{S}^*$ in the reduced populations $\mathcal{P}_{-j}$, $j = 1, \ldots, n$, to compute the conditional probabilities $\alpha^{(1)}(h|s_j^*$ and $\pi^{(1)}(s_j^*|s_j^*, H)$. Estimator $T_1^{(L)}$ is motivated from Horvitz–Thompson estimator (Horvitz and Thompson 1952), where units having smaller inclusion probability in the sample is given higher weight. On the other hand, it should be clear that $T_1^{(L)}$ is not a Horvitz–Thompson estimator since $\pi^{(L)}(i|\boldsymbol{S})$ and $\pi^{(L)}(i|\boldsymbol{S}^*)$ are not inclusion probabilities for sample $\boldsymbol{X}$. They are the inclusion probabilities for sample $\boldsymbol{X}_{H|\boldsymbol{S}}$ and $\boldsymbol{X}_{H|\boldsymbol{S}^*}$, respectively.

The estimator $T_2^{(L)}$ is motivated from JPS estimator in MacEachern et al. (2004), where each measured observation is prorated to H ranking classes. The prorate is proportional to the probability that the measured unit has rank h in a set of size H. This prorating process creates H strata. Hence, improvement over simple random sample (SRS) estimator can be anticipated form the theory of stratified sampling design in survey sampling. Even though the estimators $T_2^{(0)}$ and $T_2^{(2)}$ have the same form, they yield different efficiency results since the sample $\boldsymbol{X}$ is constructed with and without replacement in design-0 and design-1, respectively.

The estimator $T_3^{(L)}$ uses the same idea as in estimator $T_2^{(L)}$, but it gives more weight to observations that are less likely to be included in the sample to reduce the variance of the estimator. One then anticipates that the estimator $T_3^{(L)}$ performs better than the other two estimators.

Estimators for the population mean can be obtained by dividing $T_r^{(L)}$ with N

$$\hat{\mu}_r^{(L)} = \frac{1}{N}T_r^{(L)}, r = 1, \ldots, 3, \quad L = 0, 1, 2.$$

4. EMPIRICAL EVIDENCE

In this section, we investigate the efficiency of the estimators. Even though it is theoretically possible to construct the probability distributions of the estimators by computing the weight functions over all possible values of $\boldsymbol{S}$, this would computationally be intensive even for moderate sample and population sizes. Hence, to reduce the computational burden, we use a simulation study to investigate the properties of the estimators.

Simulation study considered two sets of sample (n) and set (H) sizes, $n = 20, 50$ and $H = 2, 5$, respectively. Ranking accuracy is controlled by the correlation coefficient $\rho = 1.00, 0.75, 0.5$ between X and Y. Datasets are generated from discrete normal and exponential distributions of size $N = 300$. Discrete normal and exponential populations are generated by $x_i = Q((i - 0.5)/N)$, $i = 1, \ldots, N$, where Q is the quantile function

Table 1. Biases of the the estimators $\hat{\mu}_r^{(L)}$, $r = 1, 2, 3$ and SRS mean.

n	H	ρ	μ	$\hat{\mu}_1^{(0)}$	$\hat{\mu}_2^{(0)}$	$\hat{\mu}_3^{(0)}$	$\hat{\mu}_1^{(1)}$	$\hat{\mu}_2^{(1)}$	$\hat{\mu}_3^{(1)}$	$\hat{\mu}_1^{(2)}$	$\hat{\mu}_2^{(2)}$	$\hat{\mu}_3^{(2)}$	SRS
20	2	1.00	0	−0.002	−0.002	−0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.000
20	2	1.00	100	2.013	0.001	0.000	−1.126	0.004	0.002	−1.106	0.004	0.003	0.007
20	2	0.75	0	0.006	0.006	0.005	−0.005	−0.006	−0.006	−0.006	−0.006	−0.006	−0.006
20	2	0.75	100	1.901	0.008	0.006	−1.142	−0.003	−0.001	−1.099	−0.003	−0.001	−0.005
20	2	0.50	0	0.003	0.003	0.003	−0.007	−0.008	−0.009	−0.008	−0.008	−0.009	−0.006
20	2	0.50	100	2.070	0.007	0.006	−1.038	−0.001	−0.001	−1.018	−0.001	−0.001	−0.000
20	5	1.00	0	0.001	0.001	0.001	−0.004	−0.005	−0.003	−0.005	−0.005	−0.003	−0.010
20	5	1.00	100	0.469	−0.005	−0.003	−2.463	−0.002	−0.002	−2.356	−0.002	−0.002	−0.006
20	5	0.75	0	−0.004	−0.004	−0.003	−0.002	−0.004	−0.004	−0.003	−0.004	−0.004	−0.002
20	5	0.75	100	0.606	−0.009	−0.010	−2.511	−0.003	−0.003	−2.432	−0.003	−0.003	−0.003
20	5	0.50	0	0.003	0.003	0.003	0.000	−0.001	−0.001	−0.001	−0.001	−0.001	−0.001
20	5	0.50	100	0.462	−0.000	0.000	−2.482	0.013	0.014	−2.404	0.013	0.014	0.011
50	2	1.00	0	0.002	0.002	0.001	0.002	0.001	−0.000	0.001	0.001	−0.000	0.003
50	2	1.00	100	7.265	0.002	0.001	−0.404	−0.001	−0.001	−0.369	−0.001	−0.001	−0.003
50	2	0.75	0	0.004	0.004	0.003	0.001	0.000	0.001	0.000	0.000	0.001	−0.000
50	2	0.75	100	7.289	0.003	0.003	−0.409	−0.001	−0.002	−0.391	−0.001	−0.002	−0.001
50	2	0.50	0	0.006	0.006	0.006	0.001	0.001	0.001	0.001	0.001	0.001	−0.000
50	2	0.50	100	7.267	−0.001	−0.001	−0.401	0.003	0.003	−0.380	0.003	0.003	0.003
50	5	1.00	0	−0.000	−0.000	−0.001	0.001	−0.001	−0.000	−0.001	−0.001	−0.000	−0.003
50	5	1.00	100	6.690	0.000	−0.000	−0.948	0.001	−0.000	−0.888	0.001	0.000	0.002
50	5	0.75	0	0.001	0.001	0.001	0.002	0.000	0.000	0.001	0.001	0.000	0.000
50	5	0.75	100	6.670	0.002	0.002	−0.964	−0.001	−0.000	−0.849	−0.001	0.000	−0.004
50	5	0.50	0	−0.001	−0.000	−0.000	−0.000	−0.000	−0.000	−0.001	−0.001	−0.000	−0.003
50	5	0.50	100	6.653	−0.004	−0.004	−0.973	−0.003	−0.003	−0.867	−0.003	−0.003	−0.005

Datasets are generated from discrete standard normal distribution and shifted by μ. Simulation size is 1000.

of either standard normal or standard exponential distribution depending on the underlying population. Simulation size is taken to be 1000.

Ranking accuracy is simulated by perceived size ranking model in Dell and Clutter (1972). This model, for the population values $\boldsymbol{x} = (x_1, \ldots, x_N)$, selects an N dimensional random vector, $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_N)$, from a normal distribution having mean zero and variance τ^2. These two vectors are added to create a ranking vector $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{\epsilon}$. The ranks of the observations in the vector $\boldsymbol{y}$ are used to predict the ranks of the values ($\boldsymbol{x}$) of population units. The accuracy of ranking is controlled by the correlation coefficient: $\rho = \mathrm{corr}(Y, X) = \frac{1}{\sqrt{1+\tau^2/\sigma^2}}$, or equivalently by the variance τ^2, where σ^2 is the variance of X.

Tables 1 and 2 present the biases of the estimators for discrete normal and exponential distributions. It is clear from these tables that the estimator $\hat{\mu}_1^{(L)}$ has a substantial amount of bias in all sampling designs when the population mean is large, or equivalently when the coefficient of variation, CV$= \sigma/\mu$, is small. For example, the biases of the estimator $\hat{\mu}_1^{(L)}$ are practically zero when $\mu = 0$, but they become very large for $\mu = 100$. The other estimators, $\hat{\mu}_r^{(L)}$, $r = 2, 3$; $L = 0, 1, 2$, appear to be essentially unbiased for all μ and sampling designs. Since the biases of the estimators $\hat{\mu}_1^{(L)}$, $L = 0, 1, 2$, are very large when the coefficient of variation is small, these estimators are not considered any further in this paper.

Table 2. Biases of the the estimators $\hat{\mu}_r^{(L)}$, $r = 1, 2, 3$ and SRS mean.

n	H	ρ	μ	$\hat{\mu}_1^{(0)}$	$\hat{\mu}_2^{(0)}$	$\hat{\mu}_3^{(0)}$	$\hat{\mu}_1^{(1)}$	$\hat{\mu}_2^{(1)}$	$\hat{\mu}_3^{(1)}$	$\hat{\mu}_1^{(2)}$	$\hat{\mu}_2^{(2)}$	$\hat{\mu}_3^{(2)}$	SRS
20	2	1.00	0	0.023	0.001	−0.000	−0.012	−0.004	−0.005	−0.012	−0.004	−0.005	0.002
20	2	1.00	100	2.006	−0.006	−0.008	−1.066	−0.010	−0.010	−1.030	−0.010	−0.010	−0.007
20	2	0.75	0	0.028	0.007	0.004	−0.007	0.001	0.000	−0.007	0.001	0.000	0.005
20	2	0.75	100	1.961	−0.006	−0.008	−1.070	0.000	−0.001	−1.047	0.000	−0.001	0.006
20	2	0.50	0	0.013	−0.008	−0.008	−0.018	−0.008	−0.009	−0.018	−0.008	−0.008	−0.006
20	2	0.50	100	2.012	−0.003	−0.005	−1.094	0.005	0.004	−1.069	0.005	0.004	0.009
20	5	1.00	0	−0.000	−0.008	−0.010	−0.033	−0.014	−0.014	−0.034	−0.014	−0.014	−0.006
20	5	1.00	100	0.369	−0.012	−0.015	−2.616	−0.012	−0.013	−2.527	−0.012	−0.013	−0.000
20	5	0.75	0	0.002	−0.006	−0.008	−0.029	−0.007	−0.008	−0.029	−0.007	−0.008	0.003
20	5	0.75	100	0.594	−0.011	−0.013	−2.687	−0.010	−0.011	−2.591	−0.010	−0.011	0.003
20	5	0.50	0	0.002	−0.005	−0.008	−0.032	−0.011	−0.012	−0.033	−0.011	−0.012	−0.004
20	5	0.50	100	0.334	−0.017	−0.016	−2.575	−0.004	−0.005	−2.528	−0.004	−0.005	0.004
50	2	1.00	0	0.070	−0.004	−0.004	−0.007	−0.004	−0.003	−0.007	−0.004	−0.003	−0.002
50	2	1.00	100	7.343	0.002	0.000	−0.413	−0.005	−0.005	−0.377	−0.005	−0.004	−0.003
50	2	0.75	0	0.066	−0.007	−0.006	−0.008	−0.005	−0.006	−0.008	−0.005	−0.005	−0.003
50	2	0.75	100	7.350	−0.002	−0.002	−0.405	−0.005	−0.005	−0.365	−0.005	−0.005	−0.004
50	2	0.50	0	0.069	−0.004	−0.003	−0.009	−0.005	−0.005	−0.008	−0.005	−0.005	−0.004
50	2	0.50	100	7.378	0.004	0.004	−0.428	−0.008	−0.008	−0.396	−0.008	−0.008	−0.008
50	5	1.00	0	0.062	−0.006	−0.006	−0.016	−0.008	−0.007	−0.016	−0.008	−0.007	−0.008
50	5	1.00	100	6.703	−0.005	−0.005	−0.948	−0.004	−0.003	−0.872	−0.004	−0.003	−0.002
50	5	0.75	0	0.063	−0.004	−0.004	−0.015	−0.009	−0.010	−0.017	−0.009	−0.010	−0.002
50	5	0.75	100	6.798	−0.003	−0.003	−0.878	−0.005	−0.005	−0.838	−0.004	−0.005	−0.001
50	5	0.50	0	0.065	−0.002	−0.002	−0.008	0.001	0.001	−0.007	0.001	0.001	0.001
50	5	0.50	100	6.706	−0.006	−0.006	−0.912	−0.002	−0.003	−0.903	−0.002	−0.003	0.002

Datasets are generated from discrete standard exponential distribution and shifted by μ. Simulation size is 1000.

Tables 3 and 4 present the relative efficiencies of $\hat{\mu}_r^{(L)}$, $r = 2, 3$, and SRS mean with respect to the estimator $\hat{\mu}_3^{(2)}$

$$R_r^{(L)} = \frac{\text{MSE}(\hat{\mu}_r^{(L)})}{\text{MSE}(\hat{\mu}_3^{(2)})}, r = 2, 3; L = 0, 1, 2, \text{ and } R_{\text{SRS}} = \frac{\text{MSE(SRS)}}{\text{MSE}(\hat{\mu}_3^{(2)})}.$$

The values of $R_r^{(L)} > 1$ and $R_{\text{SRS}} > 1$ indicate that the estimator $\hat{\mu}_3^{(2)}$ outperforms $\hat{\mu}_r^{(L)}$, and SRS mean, respectively.

There are several important features in Tables 3 and 4 that need to be discussed. It is clear that $R_{\text{SRS}}/R_r^{(L)} > 1$, for $r = 2, 3$ and $L = 0, 1, 2$, which indicates that all of the proposed estimators have higher efficiencies than SRS mean. The efficiency gain is substantial if the ranking information is accurate and set size is large. For example, in Table 3, the efficiencies of $\hat{\mu}_2^{(0)}$, $\hat{\mu}_3^{(0)}$, $\hat{\mu}_2^{(1)}$, $\hat{\mu}_3^{(1)}$, $\hat{\mu}_2^{(2)}$, and $\hat{\mu}_3^{(2)}$ with respect to SRS mean are 5.480 (13.809/2.52), 11.604 (13.809/1.19), 6.529 (13.809/2.115), 14.310 (13.809/0.965), 5.960 (13.809/2.317), and 13.809, respectively, when $n = 20$, $H = 5$, $\rho = 1$, and $\mu = 0$. Even if $\rho = 0.50$, the new estimators are still better than SRS mean. On the other hand, the efficiency gain is not as high as the ones under perfect ranking.

Tables 3 and 4 also reveal that the relative efficiencies of design-0 estimators are generally lower than the efficiencies of design-1 and design-2 estimators. This is mostly due to the

Table 3. Relative efficiencies of $\hat{\mu}_r^{(L)}$, $r = 2, 3$ and SRS estimators with respect to $\hat{\mu}_3^{(2)}$, $R_r^{(L)} = \text{MSE}(\hat{\mu}_r^{(L)})/\text{MSE}(\hat{\mu}_3^{(2)})$ and $R_{\text{SRS}} = \text{MSE}(\text{SRS})/\text{MSE}(\hat{\mu}_3^{(2)})$.

n	H	ρ	μ	$R_2^{(0)}$	$R_3^{(0)}$	$R_2^{(1)}$	$R_3^{(1)}$	$R_2^{(2)}$	R_{SRS}
20	2	1.00	0	2.044	1.138	1.775	0.989	1.827	3.728
20	2	1.00	100	1.787	1.004	1.798	0.992	1.849	3.770
20	2	0.75	0	1.214	1.039	1.161	0.998	1.172	1.597
20	2	0.75	100	1.294	1.032	1.226	0.998	1.241	1.787
20	2	0.50	0	1.078	1.026	1.058	0.999	1.062	1.216
20	2	0.50	100	1.146	1.084	1.065	0.999	1.069	1.232
20	5	1.00	0	2.520	1.190	2.115	0.965	2.317	13.809
20	5	1.00	100	2.393	1.109	2.244	0.969	2.449	14.688
20	5	0.75	0	1.119	1.033	1.055	1.000	1.068	1.873
20	5	0.75	100	1.145	1.013	1.089	0.999	1.106	2.256
20	5	0.50	0	1.053	1.059	0.999	1.002	1.001	1.185
20	5	0.50	100	1.010	1.012	0.999	1.001	1.002	1.247
50	2	1.00	0	2.309	1.246	1.765	0.960	1.915	3.970
50	2	1.00	100	2.293	1.247	1.776	0.964	1.925	3.997
50	2	0.75	0	1.319	1.071	1.191	0.991	1.228	1.755
50	2	0.75	100	1.474	1.201	1.203	0.991	1.243	1.811
50	2	0.50	0	1.139	1.072	1.052	0.999	1.061	1.198
50	2	0.50	100	1.533	1.421	1.077	0.995	1.093	1.305
50	5	1.00	0	2.887	1.095	2.121	0.926	2.759	17.708
50	5	1.00	100	2.792	1.129	2.107	0.912	2.732	17.356
50	5	0.75	0	1.386	1.221	1.091	0.993	1.144	2.380
50	5	0.75	100	1.202	1.091	1.075	0.992	1.121	2.198
50	5	0.50	0	1.186	1.161	1.017	1.002	1.025	1.328
50	5	0.50	100	1.154	1.135	1.009	1.003	1.017	1.241

Datasets are generated from discrete standard normal distribution and shifted by μ. Simulation size is 1000.

replacement policy of the design-0, where units are selected with replacement. Among these three designs, it appears that design-1 is the most efficient one. For example, $R_2^{(2)}/R_2^{(1)} > 1$. This can be anticipated from the fact that design-1 determines the ranks of sample units sequentially by removing all the previously ranked units in the sample from the population. This sequential ranking provides stronger data structure in the sample in design-1 than the one in design-2, and hence increases the efficiency. On the other hand, design-1 and design-2 are comparable in their efficiency for the estimators $\hat{\mu}_3^{(1)}$ and $\hat{\mu}_3^{(2)}$. They practically have the same efficiency, $R_3^{(1)} \approx 1$. The estimator $\hat{\mu}_3^{(2)}$ is slightly less efficient when $\rho = 1$, but for the other values of ρ, $\rho < 1$, the estimators are essentially equivalent in their efficiencies.

The efficiencies of the estimators in Tables 3 and 4 appear to be increasing function of set size H. To investigate the impact of set size further, we performed another simulation study using discrete normal population ($\mu = 10$ and $\sigma = 10$) of size $N = 100$ with $\rho = 1$. In this part of the simulation study, sample and simulation sizes are taken to be 20 and 1000, respectively. Figure 1 plots the mean square errors (MSEs) of the estimators $\hat{\mu}_r^{(L)}$, $r = 2, 3$; $L = 0, 1, 2$ against set size H. It is clear that MSEs are the decreasing functions of set size H for $H \leq 15$ and almost flat for $H > 15$. The reason that the MSE plots become flat for large H can be anticipated from the behavior of $\alpha^{(L)}(h|s_j)$. For large H, this probability will be very small (or zero) for the values of h that are inconsistent with s_j. For example, if

Table 4. Relative efficiencies of $\hat{\mu}_r^{(L)}$, $r = 2, 3$ and SRS estimators with respect to $\hat{\mu}_3^{(2)}$, $R_r^{(L)} = \text{MSE}(\hat{\mu}_r^{(L)})/\text{MSE}(\hat{\mu}_3^{(2)})$, and $R_{\text{SRS}} = \text{MSE}(\text{SRS})/\text{MSE}(\hat{\mu}_3^{(2)})$.

n	H	ρ	μ	$R_2^{(0)}$	$R_3^{(0)}$	$R_2^{(1)}$	$R_3^{(1)}$	$R_2^{(2)}$	R_{SRS}
20	2	1.00	0	1.354	0.982	1.356	0.996	1.379	2.312
20	2	1.00	100	1.485	1.060	1.390	0.994	1.416	2.400
20	2	0.75	0	1.070	0.965	1.124	0.999	1.132	1.480
20	2	0.75	100	1.046	0.926	1.107	0.999	1.115	1.435
20	2	0.50	0	0.984	0.966	1.014	1.001	1.014	1.073
20	2	0.50	100	1.065	1.000	1.067	0.998	1.072	1.270
20	5	1.00	0	1.684	1.055	1.548	1.023	1.620	5.848
20	5	1.00	100	1.632	1.051	1.547	1.026	1.622	6.109
20	5	0.75	0	1.242	1.094	1.093	1.007	1.107	2.001
20	5	0.75	100	1.092	0.990	1.089	1.012	1.099	1.977
20	5	0.50	0	0.976	0.976	1.006	1.003	1.009	1.242
20	5	0.50	100	1.020	1.025	0.993	1.001	0.996	1.188
50	2	1.00	0	1.548	1.085	1.370	0.980	1.443	2.457
50	2	1.00	100	1.696	1.207	1.370	0.982	1.443	2.477
50	2	0.75	0	1.124	0.995	1.123	0.994	1.147	1.478
50	2	0.75	100	1.200	1.050	1.107	0.994	1.130	1.457
50	2	0.50	0	1.126	1.085	1.035	0.997	1.044	1.155
50	2	0.50	100	1.129	1.085	1.032	0.997	1.039	1.135
50	5	1.00	0	2.034	1.258	1.515	1.075	1.734	6.814
50	5	1.00	100	1.955	1.140	1.510	1.078	1.726	6.653
50	5	0.75	0	1.359	1.235	1.092	1.024	1.124	1.917
50	5	0.75	100	1.215	1.079	1.124	1.025	1.163	2.072
50	5	0.50	0	1.085	1.078	1.003	1.001	1.010	1.168
50	5	0.50	100	1.191	1.156	1.018	1.012	1.024	1.287

Datasets are generated from discrete standard exponential distribution and shifted by μ. Simulation size is 1000.
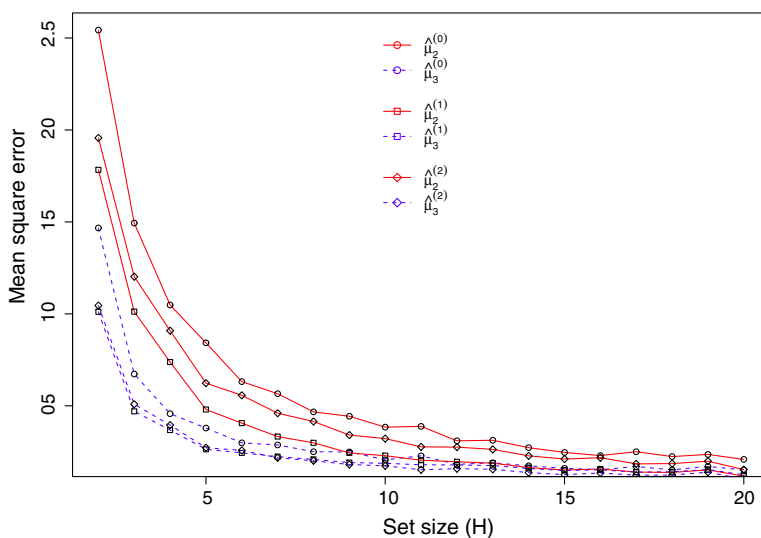

Figure 1. The plots of the simulated mean square errors of the estimators $\hat{\mu}_r^{(L)}$, $r = 2, 3$, $L = 0, 1, 2$, with respect to set size H. Samples ($n = 20$) are generated from a discrete normal distribution of size $N = 100$ with mean $\mu = 10$, $\sigma = 10$ and perfect ranking $\rho = 1$. Simulation size is 1000.

s_j is small, then large values of h yields $\alpha^{(L)}(h|s_j) \approx 0$. Hence, the contribution of these ranking classes would be negligible to reduce the MSE of the estimators.

Figure 1 also indicates that design-1 estimators yield smaller MSEs than the design-0 and design-2 estimators. In design-1, population ranks of the selected units are determined sequentially after removing all the previously selected units. This sequential ranking induces stronger data structure (presumably negative correlations among measured observations). Hence, design-1 yields higher efficiency results than the other designs in finite population setting.

5. BOOTSTRAP VARIANCE ESTIMATE AND CONFIDENCE INTERVAL

In this section, we develop statistical inference for population mean, but similar inference can also be developed for population total with a slight change in the notation. The exact sampling distribution of the estimator is not computationally feasible for reasonable sample and population sizes. Therefore, to reduce the computational burden, we use bootstrap distribution to draw statistical inference.

Discussion in the previous section indicates that design-1 performs slightly better than the other two designs. On the other hand, the computation of the conditional inclusion probabilities in design-1 requires extensive computing time when the population and/or sample sizes are large. Since the estimators $\hat{\mu}_3^{(1)}$ and $\hat{\mu}_3^{(2)}$ are equivalent in their efficiencies for all practical purposes, to increase the bootstrap simulation size, we develop the inference based on design-2 estimator $\hat{\mu}_3^{(2)}$.

Let θ be the parameter of interest. The parameter θ can be considered as a statistical functional $\theta = F(\mathcal{P})$. The estimate of θ then can be obtained from plug-in method by replacing $\mathcal{P}$ with empirical population $\hat{\mathcal{P}}$, $\hat{\theta} = F(\hat{\mathcal{P}})$. The empirical bootstrap population $\hat{\mathcal{P}}$ should preserve without replacement structure of design-0, design-1, and design-2. Let $\boldsymbol{x}_S^{(L)} = \{x_j, s_j\}; j = 1, \ldots, n$, be the measured values of the simple random sample, $U = \{u_{s_1}, \ldots, U_{s_n}\}$, selected from population $\mathcal{P}$ based on design-L, $L = 0, 1, 2$. Let D be the integer part of the ratio N/n. We construct empirical bootstrap population by repeating set $\boldsymbol{x}_S^{(L)}$ D times and selecting $d = N - Dn$ pairs at random from $\boldsymbol{x}_S^{(L)}$ to create an empirical population of size N:

$$\hat{\mathcal{P}}^{(L)} = \{\boldsymbol{x}_S^{(L)}, \ldots, \boldsymbol{x}_S^{(L)}, z_1, \ldots, z_d\},$$

where z_t, $t = 1, \ldots, d$, are randomly selected pairs from $\boldsymbol{x}_S^{(L)}$. It is clear that the size of the empirical population is the same as the original population. The bootstrap samples are then selected without replacement from population $\hat{\mathcal{P}}^{(L)}$. Let $\boldsymbol{x}^* = \{x_j^*, a_j\}; j = 1, \ldots, n$, be a re-sample from population $\hat{\mathcal{P}}^{(L)}$, where $a_i \in S$ for $i = 1, \ldots, n$. For each $b = 1, \ldots, B$, let $\boldsymbol{x}_b^* = \{x_{j,b}^*, a_{j,b}\}; j = 1, \ldots, n$, be a re-sample selected from $\hat{\mathcal{P}}^{(L)}$, we apply our estimator to each one of these bootstrap re-samples to obtain

$$\hat{\mu}_{2,b}^{(L)*} = \frac{N}{H} \sum_{h=1}^{H} \sum_{j=1}^{n} w_2^{(L)}(h, j|A_b, H) x_{j,b}^*, \text{ and}$$

$$\hat{\mu}_{3,b}^{(L)*} = \frac{N}{H} \sum_{h=1}^{H} \sum_{j=1}^{n} w_3^{(L)}(h, j|A_b) x_{j,b}^*, \quad b = 1, \ldots, B,$$

where $A_b = (a_{1,b}, \ldots, a_{n,b})$. The bootstrap variance estimates of the estimators in $\hat{\mu}_r^{(L)}$, $r = 2, 3$, are then given by

$$BV(\hat{\mu}_r^{(L)}) = \frac{1}{B-1} \sum_{b=1}^{B} \left(\hat{\mu}_{r,b}^{(L)*} - \bar{\hat{\mu}}_r^{(L)*} \right)^2, r = 2, 3,$$

where $\bar{\hat{\mu}}_r^{(L)*}$ is the mean of $\hat{\mu}_{r,b}^{(L)*}$, $b = 1, \ldots, B$.

A $(1-\gamma)100\%$ bootstrap percentile confidence interval is constructed by $(Q_r^{\gamma/2}, Q_r^{1-\gamma/2})$, where Q_r^a is the a-th quantile of $\hat{\mu}_r^{(L)}$ satisfying $a = P(\hat{\mu}_r^{(L)} < Q_r^a|\mathcal{P})$. The quantiles $Q_r^{\gamma/2}$ and $Q_r^{1-\gamma/2}$ are obtained from bootstrap distribution of $\hat{\mu}_r^{(L)}$.

In order to investigate the properties of bootstrap variance estimate of $\hat{\mu}_r^{(2)}$, for $r = 2, 3$, and the bootstrap percentile confidence interval of population mean, we performed another simulation study. The simulation parameters are taken to be $n = 30, 50$, $H = 2, 5$, $\rho = 1, 0.75, 0.5$. The shift parameter μ is selected to be $\mu = 0, 100$. Datasets are generated from discrete normal and exponential distributions. Simulation and bootstrap replications are selected to to be 2000 and 1000, respectively.

Table 5 presents 95 % coverage probabilities (C) of the bootstrap percentile confidence intervals based on estimators $\hat{\mu}_r^{(2)}$, $r = 2, 3$, and SRS mean. Table 5 reveals that coverage probabilities are reasonably close to the nominal coverage probability 0.95 for discrete normal distribution. The coverage probabilities for the discrete exponential distribution appear to be slightly lower than the nominal level for small sample sizes. This may be due to the fact that super-population exponential distribution is a skewed distribution. Hence, it may require larger sample sizes to satisfy the regularity conditions of the bootstrap procedure in Booth et al. (1994).

Table 6 presents bootstrap (B) and simulated (S) variance estimates of the estimators $\hat{\mu}_r^{(2)}$, $r = 2, 3$ and SRS mean. It is obvious that bootstrap variance estimates are almost identical to those estimated from simulation study. The simulation results provide a convincing evidence that variance estimates of the proposed estimators can be computed from bootstrap distribution.

6. APPLICATION

In this section, we apply the proposed sampling designs and estimators to 1992 Ohio corn yield data which were used by the Ohio Agricultural Statistics Department in its county estimation program. This dataset includes responses from farms in the USDA's National Quarterly Agricultural Survey and from farms responding to the Ohio supplemental

Table 5.   Coverage probabilities (C) of 95 % bootstrap percentile confidence intervals of population mean based on estimators $\hat{\mu}_r^{(2)}$, $r = 2, 3$.

n	H	ρ	μ	$N(0, 1) + \mu$			$Exp(1) + \mu$		
				$C(\hat{\mu}_2^{(2)})$	$C(\hat{\mu}_3^{(2)})$	$C(SRS)$	$C(\hat{\mu}_2^{(2)})$	$C(\hat{\mu}_3^{(2)})$	$C(SRS)$
30	2	1.00	0	0.931	0.926	0.935	0.914	0.895	0.924
30	2	1.00	100	0.926	0.923	0.929	0.904	0.881	0.916
30	2	0.75	0	0.931	0.928	0.936	0.905	0.903	0.905
30	2	0.75	100	0.930	0.926	0.933	0.914	0.906	0.912
30	2	0.50	0	0.938	0.938	0.940	0.921	0.925	0.924
30	2	0.50	100	0.934	0.933	0.934	0.907	0.908	0.909
30	5	1.00	0	0.909	0.897	0.929	0.880	0.830	0.924
30	5	1.00	100	0.910	0.900	0.927	0.891	0.848	0.928
30	5	0.75	0	0.943	0.945	0.929	0.910	0.910	0.922
30	5	0.75	100	0.934	0.930	0.936	0.904	0.910	0.908
30	5	0.50	0	0.924	0.922	0.935	0.926	0.917	0.926
30	5	0.50	100	0.939	0.933	0.939	0.912	0.910	0.914
50	2	1.00	0	0.945	0.943	0.947	0.920	0.915	0.932
50	2	1.00	100	0.934	0.932	0.938	0.925	0.915	0.927
50	2	0.75	0	0.943	0.943	0.939	0.919	0.917	0.918
50	2	0.75	100	0.937	0.938	0.942	0.924	0.927	0.923
50	2	0.50	0	0.936	0.939	0.938	0.923	0.919	0.922
50	2	0.50	100	0.949	0.948	0.946	0.931	0.933	0.930
50	5	1.00	0	0.925	0.922	0.936	0.900	0.864	0.933
50	5	1.00	100	0.927	0.920	0.938	0.898	0.854	0.930
50	5	0.75	0	0.945	0.945	0.938	0.924	0.923	0.934
50	5	0.75	100	0.941	0.936	0.938	0.921	0.924	0.925
50	5	0.50	0	0.942	0.942	0.942	0.923	0.922	0.924
50	5	0.50	100	0.943	0.941	0.949	0.934	0.935	0.940

Dataset is generated from discrete standard normal and exponential distribution, and shifted by μ. Simulation size is 2000, the bootstrap simulation size is 1000.

survey, Husby et al. (2005). (Also, see Ohio Department of Agriculture, 1993, for published estimates based on these data). The success of the proposed sampling designs depends heavily on accurate ranking information of the population units. To get a reasonably correct rank ordering of the population units, we select our population as one of the counties from the Ohio Corn Data having 202 farms. Hence, the population size is $N = 202$. In this population, there are five variables: corn production (bushels, X), farm size (acreage, Y_1), group size (Y_2), acre planted (Y_3), and acre harvested (Y_4). Our interest lies in estimation of the mean corn production in the county. The constructions of the proposed sampling designs require rank ordering of 202 farms based on their corn production. We use the variables Y_1, Y_2, Y_3, and Y_4 as auxiliary variables to predict rank ordering of X. There is high correlation between X and the other auxiliary variables, $\rho_k = cor(X, Y_k)$, $r = 1, \ldots, 4$. The histogram of X reveals that the population is strongly skewed right. The parameters of this population are given in Ozturk (2014c) and reproduced in Table 7.

The auxiliary variable group size Y_2 is an integer-valued random variable which only takes values 1, 2, and 3. There are also ties in other auxiliary variables Y_1, Y_3, and Y_4. In order to break the ties, we generated a random vector ϵ of size $N = 202$ from a normal distribution with mean 0 and standard deviation 0.001 and constructed $y_j^* = y_j + u$, $j = 1, \ldots, 4$.

Table 6. Bootstrap(B) and simulation (S) variance estimates of the estimators $\hat{\mu}_r^{(2)}$, $r = 2, 3$.

Dist	n	H	ρ	μ	$B(\hat{\mu}_2^{(2)})$	$B(\hat{\mu}_3^{(2)})$	B(SRS)	$S(\hat{\mu}_2^{(2)})$	$S(\hat{\mu}_3^{(2)})$	S(SRS)
$N(0, 1)$	30	2	1.00	0	0.015	0.008	0.030	0.014	0.008	0.029
	30	2	1.00	100	0.015	0.008	0.030	0.014	0.008	0.029
	30	2	0.75	0	0.021	0.017	0.030	0.020	0.017	0.029
	30	2	0.75	100	0.022	0.018	0.030	0.020	0.017	0.029
	30	2	0.50	0	0.025	0.023	0.029	0.024	0.022	0.029
	30	2	0.50	100	0.027	0.025	0.030	0.025	0.023	0.029
	30	5	1.00	0	0.005	0.002	0.031	0.005	0.002	0.029
	30	5	1.00	100	0.005	0.002	0.032	0.005	0.002	0.029
	30	5	0.75	0	0.016	0.015	0.030	0.016	0.015	0.029
	30	5	0.75	100	0.016	0.014	0.030	0.015	0.014	0.029
	30	5	0.50	0	0.025	0.025	0.030	0.023	0.022	0.029
	30	5	0.50	100	0.023	0.023	0.029	0.023	0.023	0.030
	50	2	1.00	0	0.007	0.004	0.016	0.008	0.004	0.016
	50	2	1.00	100	0.008	0.004	0.017	0.008	0.004	0.016
	50	2	0.75	0	0.011	0.009	0.017	0.011	0.009	0.016
	50	2	0.75	100	0.011	0.009	0.016	0.011	0.009	0.016
	50	2	0.50	0	0.015	0.014	0.016	0.015	0.014	0.016
	50	2	0.50	100	0.014	0.013	0.016	0.014	0.013	0.016
	50	5	1.00	0	0.003	0.001	0.017	0.002	0.001	0.016
	50	5	1.00	100	0.003	0.001	0.017	0.003	0.001	0.016
	50	5	0.75	0	0.009	0.008	0.017	0.009	0.008	0.016
	50	5	0.75	100	0.009	0.008	0.017	0.009	0.008	0.016
	50	5	0.50	0	0.013	0.013	0.017	0.013	0.012	0.016
	50	5	0.50	100	0.013	0.012	0.017	0.013	0.012	0.016
$\mathrm{Exp}(1)$	30	2	1.00	0	0.017	0.011	0.029	0.016	0.011	0.029
	30	2	1.00	100	0.017	0.012	0.030	0.016	0.011	0.028
	30	2	0.75	0	0.024	0.022	0.030	0.022	0.020	0.028
	30	2	0.75	100	0.023	0.020	0.030	0.021	0.018	0.029
	30	2	0.50	0	0.025	0.024	0.029	0.024	0.023	0.029
	30	2	0.50	100	0.028	0.027	0.031	0.025	0.024	0.029
	30	5	1.00	0	0.008	0.005	0.030	0.007	0.004	0.029
	30	5	1.00	100	0.007	0.004	0.030	0.007	0.004	0.029
	30	5	0.75	0	0.018	0.017	0.029	0.017	0.015	0.029
	30	5	0.75	100	0.017	0.016	0.030	0.015	0.014	0.028
	30	5	0.50	0	0.023	0.023	0.029	0.022	0.021	0.029
	30	5	0.50	100	0.024	0.024	0.031	0.022	0.021	0.029
	50	2	1.00	0	0.010	0.007	0.017	0.009	0.006	0.016
	50	2	1.00	100	0.009	0.006	0.016	0.009	0.006	0.016
	50	2	0.75	0	0.013	0.012	0.017	0.012	0.011	0.016
	50	2	0.75	100	0.013	0.011	0.016	0.012	0.011	0.016
	50	2	0.50	0	0.014	0.013	0.017	0.014	0.013	0.016
	50	2	0.50	100	0.014	0.014	0.016	0.014	0.013	0.016
	50	5	1.00	0	0.004	0.003	0.016	0.004	0.002	0.016
	50	5	1.00	100	0.004	0.002	0.016	0.004	0.002	0.016
	50	5	0.75	0	0.009	0.008	0.016	0.009	0.008	0.016
	50	5	0.75	100	0.009	0.008	0.016	0.009	0.008	0.016
	50	5	0.50	0	0.014	0.014	0.017	0.013	0.013	0.016
	50	5	0.50	100	0.012	0.012	0.016	0.012	0.011	0.016

Datasets are generated from discrete standard normal and exponential distributions, and shifted by μ. Simulation size is 2000, the bootstrap simulation size is 1000.

Table 7.   Population characteristics of an Ohio county in 1992 Ohio corn data, $\rho_k = \text{corr}(X, Y_k)$.

Parameters	X	Y_1	Y_2	Y_3	Y_4
Mean	16688.089	330.583	1.762	121.960	108.634
Median	9291	200	2	73.5	62.5
St. dev	18622.160	333.265	0.755	130.740	119.216
ρ_k		0.903	0.746	0.940	0.988

Table 8.   Biases (Bi) and relative efficiencies (R) of the estimators $\hat{\mu}_2^{(2)}$, $\hat{\mu}_3^{(2)}$, $\hat{\mu}_{\text{Ra}}$, $\hat{\mu}_{\text{Reg}}$, and SRS mean.

n	H	ρ	Bi($\hat{\mu}_2^{(2)}$)	Bi($\hat{\mu}_3^{(2)}$)	Bi(SRS)	Bi(Ra)	Bi(Reg)	R_2	R_{SRS}	R_{Ra}	R_{Reg}
20	10	0.904	−274.441	−318.453	32.469	−87.861	−281.115	1.133	3.243	1.582	1.527
20	10	0.746	−196.169	−214.271	32.469	−50.931	−253.829	1.037	2.069	1.179	1.209
20	10	0.940	−298.586	−343.230	32.469	−116.539	−334.427	1.268	5.172	2.246	2.052
20	10	0.988	−348.944	−403.992	32.469	−131.604	−367.597	1.403	6.791	2.872	2.592
30	6	0.904	−88.250	−109.436	106.752	−1.544	−138.622	1.242	3.226	1.559	1.451
30	6	0.746	−26.834	−30.915	106.752	47.555	−86.799	1.111	2.118	1.208	1.209
30	6	0.940	−91.972	−110.112	106.752	−15.179	−155.947	1.431	4.778	2.087	1.800
30	6	0.988	−109.835	−133.306	106.752	−21.209	−169.171	1.558	5.760	2.449	2.065
50	4	0.904	−53.277	−50.203	−1.164	−26.087	−82.107	1.315	2.904	1.361	1.223
50	4	0.746	−53.945	−50.456	−1.164	−27.349	−96.345	1.165	2.043	1.143	1.127
50	4	0.940	−53.382	−48.561	−1.164	−30.702	−83.071	1.479	3.826	1.640	1.357
50	4	0.988	−57.896	−54.425	−1.164	−32.935	−86.846	1.553	4.234	1.775	1.435

Simulation and bootstrap replication sizes are 3000 and 2000, respectively.

Since all entries in $\boldsymbol{y}_j^*$ are unique, the rank ordering of vector $\boldsymbol{x}$ with no ties is estimated from $\boldsymbol{y}_j^*$, $j = 1, \ldots, 4$.

By treating these 202 farms as a finite population, we performed another simulation study to investigate the biases and efficiencies of $\hat{\mu}_2^{(2)}$ and $\hat{\mu}_3^{(2)}$. In this part of the simulation, we also included ratio ($\hat{\mu}_{\text{Ra}}$) and regression ($\hat{\mu}_{\text{Reg}}$) estimator of the population mean

$$\hat{\mu}_{\text{Ra}} = \frac{1}{N} \frac{\sum_{j=1}^{n} x_j}{\sum_{j=1}^{n} s_j} \frac{N(N+1)}{2} \quad \text{and} \quad \hat{\mu}_{\text{Reg}} = \hat{B}_0 + \hat{B}_1 (N+1)/2,$$

where $\hat{B}_0$ and $\hat{B}_1$ are the estimated regression coefficients, regressing x_j on s_j. Simulation study also considered the bootstrap estimates of standard deviations of the estimators and coverage probabilities of the percentile confidence intervals. Samples in the simulation are selected with the following sample and set size combination, $(n, H) = (20, 10), (30, 6),$ and $(50, 4)$. Simulation and bootstrap replication sizes are taken to be 3000 and 2000, respectively.

Table 8 presents the biases (Bi) of $\hat{\mu}_2^{(2)}$, $\hat{\mu}_3^{(2)}$, $\hat{\mu}_{\text{Ra}}$, $\hat{\mu}_{\text{Reg}}$, and SRS mean. It also contains the relative efficiencies of the estimator $\hat{\mu}_3^{(2)}$ with respect to $\hat{\mu}_2^{(2)}$, SRS sample mean, $\hat{\mu}_{\text{Ra}}$, and $\hat{\mu}_{\text{Reg}}$

Table 9. Coverage probabilities of bootstrap percentile confidence intervals of population mean based on estimators $\hat{\mu}_r^{(2)}$), $r = 2, 3$, and SRS mean.

n	H	ρ	$\hat{\mu}_2^{(2)}$	$\hat{\mu}_3^{(2)}$	SRS
20	10	0.904	0.852	0.850	0.898
20	10	0.746	0.868	0.870	0.898
20	10	0.940	0.820	0.814	0.898
20	10	0.988	0.803	0.781	0.898
30	6	0.904	0.899	0.893	0.919
30	6	0.746	0.903	0.904	0.919
30	6	0.940	0.881	0.867	0.919
30	6	0.988	0.876	0.854	0.919
50	4	0.904	0.919	0.921	0.928
50	4	0.746	0.920	0.924	0.928
50	4	0.940	0.910	0.913	0.928
50	4	0.988	0.908	0.905	0.928

Simulation and bootstrap replication sizes are 3000 and 2000, respectively.

$$R_2^{(2)} = \frac{\text{MSE}(\hat{\mu}_2^{(2)})}{\text{MSE}(\hat{\mu}_3^{(2)})}, \quad R_{\text{SRS}}^{(2)} = \frac{\text{MSE}(\text{SRS})}{\text{MSE}(\hat{\mu}_3^{(2)})},$$
$$R_{\text{Ra}}^{(2)} = \frac{\text{MSE}(\hat{\mu}_{\text{Ra}})}{\text{MSE}(\hat{\mu}_3^{(2)})} \text{ and } R_{\text{Reg}}^{(2)} = \frac{\text{MSE}(\hat{\mu}_{\text{Reg}})}{\text{MSE}(\hat{\mu}_3^{(2)})}.$$

Again relative efficiencies greater than one indicate that the estimator $\hat{\mu}_3^{(2)}$ has smaller mean square error.

Biases in Table 8 follow a pattern similar to the ones we have observed in Tables 1 and 2. For large set sizes, the estimators have slightly larger negative biases than SRS estimator. This can be anticipated from the fact that the proposed estimators borrow information from unmeasured units in a set of size H. For skew distributions and large set sizes H, effects of extreme observations in the sample are divided into H different strata, Hence the influence of extreme observations on the estimator is reduced. For this reason, for skewed distributions, the estimators provide a slightly under-estimation for the population mean.

The relative efficiencies also follow the similar pattern as in Tables 3 and 4 . The estimator $\hat{\mu}_3^{(2)}$ is the best estimator, and the efficiencies increase with the quality of ranking information, set, and sample sizes. Efficiency gain with respect to SRS sample mean, ratio, and regression estimators is substantial if the ranking information is reasonably accurate.

The same dataset is analyzed in Ozturk (2014c) using RSS design by combining ranking information from different sources. Sampling designs in this paper are different from the RSS designs used in Ozturk (2014c). Our designs select simple random samples and construct weights based on global ranking information in the entire population, whereas Ozturk (2014c) uses local ranking information in a given set of size H. Since we use global ranking information in the population, our estimator performs better than RSS designs in Ozturk (2014c).

Table 9 presents the coverage probability of the bootstrap percentile confidence interval for population mean based on estimators $\hat{\mu}_2^{(2)}$, $\mu_3^{(2)}$ and SRS mean. All these coverage

Table 10.   Bootstrap (B)- and simulation (S)-based estimates of standard deviation of the estimators, $\hat{\mu}_r^{(2)}$, $r =$ 2, 3, and SRS mean.

n	H	ρ	$S(\hat{\mu}_2^{(2)})$	$S(\hat{\mu}_3^{(2)})$	S(SRS)	$B(\hat{\mu}_2^{(2)})$	$B(\hat{\mu}_3^{(2)})$	B(SRS)
20	10	0.904	2310.295	2162.324	3935.808	2045.848	1931.809	3704.009
20	10	0.746	2779.303	2727.653	3935.808	2489.629	2436.527	3704.009
20	10	0.940	1926.164	1696.295	3935.808	1703.761	1507.485	3704.009
20	10	0.988	1754.517	1455.366	3935.808	1555.989	1320.373	3704.009
30	6	0.904	1936.858	1736.645	3123.546	1834.772	1651.334	3032.236
30	6	0.746	2263.906	2147.285	3123.546	2142.044	2028.269	3032.236
30	6	0.940	1708.115	1425.513	3123.546	1608.442	1342.961	3032.236
30	6	0.988	1621.721	1295.439	3123.546	1518.565	1212.606	3032.236
50	4	0.904	1530.260	1334.536	2275.824	1512.304	1326.992	2240.971
50	4	0.746	1717.860	1591.619	2275.824	1690.895	1564.606	2240.971
50	4	0.940	1413.872	1162.421	2275.824	1386.626	1143.265	2240.971
50	4	0.988	1377.160	1104.657	2275.824	1343.104	1074.441	2240.971

Simulation and bootstrap replication sizes are 3000 and 2000, respectively.

probabilities appear to be close to the nominal value 0.95 for large sample sizes. Since the population has strong skewness to right, for small sample sizes, the coverage probabilities are slightly smaller than the nominal coverage probability 0.95. Table 10 presents the standard error estimates of the estimators. Again the bootstrap estimate of standard deviations are close to the estimate of the standard deviation of the estimators from simulation.

7. CONCLUDING REMARK

We have developed three sampling designs to estimate the population total and mean in a finite population setting. The proposed sampling designs select a simple random sample and identify their population ranks. Selection of the sample could be either with- or without-replacement. Replacement policy and the way that population ranks of selected units are identified define the three different sampling designs: design-0, design-1, and design-2. In these designs, population ranks of the measured units provide information about the relative positions of the sample units in the population. This positional information is used to borrow additional information from other unmeasured units in the population to reduce the uncertainty in the sample. We introduced three different estimators for the population mean and total for each one of these sampling designs. We show that the estimators perform better than simple random sample estimator as long as there is meaningful ranking information to rank the population units. The efficiencies of the estimators are an increasing function of the correlation coefficient between the response and auxiliary variable.

The population ranks of the selected units can be considered as covariate. These ranks are not strongly attached to the measured values. They can be ignored completely, and the sample can be analyzed as a simple random sample. Similar to a ranked set sample, where strong tie is established between the rank and measurement, it is possible to attach ranking information to the measured units strongly to induce further stratification in the data, but the resulting sample may not be reduced to a simple random sample. In this case, the properties

of this sample need to be developed further. One of our current project investigates this type of sampling design.

ACKNOWLEDGEMENTS

REFERENCES

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

Booth, J. G., Butler, W. R. and Hall, P. (1994). Bootstrap methods for finite populations. *Journal of American Statisitcal Associaition*, 89, 1282-1289.

Dell, T. R., and Clutter, J. L. (1972). Ranked set sampling theory with order statistics backround. *Biometrics*,28,545-555.

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

Frey, J. and Feeman, T. G. (2012). An improved mean estimator for judgment post-stratification, *Computational Statistics and Data Analysis*, 56, 418-426.

——— (2013). Variance estimation using judgment post-stratification. To appear *Annals of the Institute of Statistical Mathematics.*

Frey, J. and Ozturk, O. (2011). Constrained estimation using judgment post-stratification. *Annals of the Institute of Statistical Mathematics*, 63, 769-789.

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

Horvitz, D. G., and Thompson, D .J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of Amewrican Statistical Association*, 47, 663–685.

Husby, C. E., Stasny, E. A., and Wolfe, D. A. (2005). An application of ranked set sampling for mean and median estimation using USDA crop production data. *Journal of Agricultural, Biological, and Environmental Statisitcs*, 10, 354-373.

Iwig, W. C. (1993). "The National Agricultural Statistics Service County Estimates Program" *in Indirect Estimators in Federal Programs*, Statistical Policy Working Paper 21, Report of the Federal Committee on Statistical Methodology, Subcommittee on Small Area Estimation , Washington DC, 7.1–7.15.

Jafari Jozani, M., and Johnson, B. C. (2011). Design based estimation for ranked set sampling in finite population. *Environmental and Ecological Statistics*, 18, 663–685.

——— (2012). Randomized nomination sampling in finite populations. *Journal of Statistical Planning and inference*, 142, 2103–2115.

MacEachern, S.N., Stasny, E.A., and Wolfe, D.A. (2004). Judgment post-stratification with imprecise ranking. *Biometrics*, 60, 207-215.

McIntyre, G. A. (1952). A method for unbiased selective sampling using ranked-sets. *AustralianJournal of Agricultural Research*, 3, 385–390.

——— (2005). A method of unbiased selective sampling using ranked-sets. *The American Statistician*, 59, 230–232.

Ozdemir,Y. A., and Gokpinar, F. (2008). A new formula for inclusion probablities in median ranked set sampling. *Communications in Statisitcs– Theory and Methods*, 37, 2022-2033.

Ozturk, O. and Jafari Jozani (2013). Inclusion Probabilities in Partially Rank Ordered Set Sampling. In review, *Computational Statisitcs and Data anlysis*.

Ozturk, O. (2014a). Statistical Inference for Population Quantiles and Variance in Judgment Post-Stratified Samples, *Computational Statistics and Data Analysis*,77, 188-205.

—— (2014b). Distribution free two-sample methods for judgment-post stratified data. In review *Statistica Sinica*.

—— (2014c). Estimation of population mean and total in finite population setting using multiple auxiliary variables. *Journal of Agircultural, Biological and Environmental Statistics*, 19, 161-184.

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

Stokes, S. L., Wang, X. and Chen, M. (2007). Judgment post stratification with multiple rankers. *Journal of Statistical Theory and Applications*, 6, 344-359.

Wang, X., Lim, J., and Stokes, S. L. (2008). A nonparametric mean estimator for judgment post-stratified data. *Biometrics*, 64: 355-363.

Wang, X., Wang, K., and Lim, J. (2012). Isotonized CDF estimation from judgment post-stratification data with empty strata. *Biometrics*, 68, 194-202.

Wang, X., Stokes, L., Lim, J., and Chen, M. (2006). Concomitant of multivariate order statistics with application to judgment poststratification. *Journal of American Statistical Association*, 101(476): 1693-1704.

Wolfe, D. A. (2012). Ranked Set Sampling: Its Relevance and Impact on Statistical Inference. *ISRN Probability and Statistics*, doi:10.5402/2012/568385.