Estimation of Population Mean and Total in a Finite Population Setting Using Multiple Auxiliary Variables

Omer OZTURK

This paper introduces a new sampling design in a finite population setting, where potential sampling units have a wealth of auxiliary information that can be used to rank them into partially ordered sets. The proposed sampling design selects a set of sampling units. These units are judgment ranked without measurement by using available auxiliary information. The ranking process allows ties among ranks whenever units cannot be ranked accurately with high confidence. The ranking information from all sources is combined in a meaningful way to construct strength-of-agreement weights. These weights are then used to select a single sampling unit for full measurement in each set. Three different levels of sampling design, level-0, level-1, and level-2, are investigated. They differ in their replacement policies. Level-0 sampling designs construct the sample by sampling with replacement, level-1 sampling designs constructs the sample without replacement of the fully measured unit in each set, and level-2 sampling designs construct the sample without replacement on the entire set. For these three designs, we estimate the first and second order inclusion probabilities and construct estimators for the population total and mean. We develop a bootstrap resampling procedure to estimate the variances of the estimators and to construct percentile confidence intervals for the population mean and total. We show that the new sampling designs provide a substantial amount of efficiency gain over their competitor designs in the literature.

Key Words: Coefficient of variation; Finite population; Horvitz–Thompson estimator; Partial ranking; Ranked set sampling.

1. INTRODUCTION

In many survey sampling studies, in addition to variable of interest, researchers often have additional auxiliary information to improve statistical inference. In many instances, this information may not be accurate, cannot be turned into a numerical covariate or may be even subjective. Nevertheless, it contains valuable information that can be used at the

---

Omer Ozturk (✉) is Professor (E-mail: *omer@stat.osu.edu*), Department of Statistics, The Ohio State University, 1958 Neil Aven., Columbus, OH 43210, USA.

Journal of Agricultural, Biological, and Environmental Statistics, Volume 19, Number 2, Pages 161–184
DOI: 10.1007/s13253-013-0163-9

design stage of the survey. Ratio and regression estimators, which require strong modeling assumptions, are constructed based on this type of information. In the absence of such strong modeling assumptions, it is not clear how this auxiliary information should be incorporated into a survey sampling design. Hence, it is usually ignored. In an infinite population setting, use of this type of subjective information has generated extensive research interests on ranked set sampling (RSS) which was originally developed to keep the overall cost of data collection minimal in estimating mean pasture yield in agricultural fields (McIntyre 1952 and 2005). Ranked set sampling needs a relative ranking of a few units in the construction of a data set to divide it 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. Finite population settings provide a natural platform for ranked set sampling where the consistency in ranking scheme can be achieved through either the use of a judgment ranking process via visual inspection of the units or auxiliary information, such as previous survey outcomes, census track etc. The auxiliary variables are very common in survey sampling studies and can be very useful for creating judgment strata in RSS data.

One natural setting for the auxiliary information in survey sampling is given in Husby, Stasny, and Wolfe (2005), 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. The detailed description of the USDA/NASS county estimation program can be found in Iwig (1993). Within the USDA/NASS county estimation program, the Ohio Agricultural Statistics Department used the 1992 Ohio corn yield data in its county estimation program. This data set includes responses from farms in the USDA's National Quarterly Agricultural Survey and from farms responding to the Ohio supplemental survey, Husby, Stasny, and Wolfe (2005). (Also, see Ohio Department of Agriculture, 1993, for published estimates based on these data.)

One of the counties in the Ohio corn yield data has 202 farms. The farms in this county will serve as our finite population in this paper. The population has five variables, corn yields (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 yields in the county. We treat the variables Y_1, Y_2, Y_3 and Y_4 as auxiliary variables. The auxiliary variable group size (Y_2) is an integer valued random variable which only takes values $1, 2, 3$. There is high correlation between X and the other auxiliary variables, $\rho_k = \mathrm{cor}(X, Y_k)$, $r = 1, \dots, 4$. The population characteristics of these farms are given in Table 1.

Construction of a ranked set sample in an infinite population setting requires to specify a sample size n, set size H and sampling design D, $D = \{r_1, \dots, r_n\}$, where r_j $(1 \le r_j \le H)$ is the judgment rank of the fully measured unit in set j, $n_h = \sum_{j=1}^{n} I(r_j = h)$ and $I(\cdot)$ is an indicator function. Ranked set sample then selects nH farms at random from the population. These farms are randomly divided into n sets, each of size H. Farms in each set are ranked from smallest to largest without actual measurement of the variable of interest. Ranking may be performed based on visual inspection, available auxiliary information or

Table 1. Population characteristics of one of the farm populations in 1992 Ohio corn data, $\rho_k = \mathrm{corr}(X, Y_k)$.

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

any other means that do not require an actual measurement of the units. Ranking process does not have to be perfectly accurate as long as it assigns ranks for all units in the sets and does not change from set to set. In Ohio corn yield data, the ranking process can use any one of the auxiliary variables, $Y_1, \ldots, Y_4$, to rank the farms in each set. The r_jth smallest ranked farm is then selected for full measurement in set j. The fully measured farms $X_{[r_j]j}$, $j = 1, \ldots, n$, are called a ranked set sample. The design D yields n_h fully measured farms in judgment strata h so that $\sum_{h=1}^{H} n_h = n$. If $n_h \equiv n/H$, ranked set sample is called balanced. In a balanced RSS, each judgment strata has an equal number of observations. The square brackets indicate that ranking process may be in error. If the ranking process is error-free, RSS contains independent order statistics from different sets and creates the biggest separation among the strata of ranking classes. Hence, it yields highest efficiency improvement.

In recent years, research in ranked set sampling drew considerable attention in a finite population setting, e.g., Patil, Sinha, and Taillie (1995), Deshpande, Frey, and Ozturk (2006), Al-Saleh and Samawi (2007), Ozdemir and Gokpinar (2007 and 2008), Jafari-Jozani and Johnson (2011, 2012), Gokpinar and Ozdemir (2010). Deshpande, Frey, and Ozturk (2006) considered three different designs, level-0, level-1 and level-2 sampling designs that differ in their application policies. The level-0 sampling design requires that units in a given set are selected without replacement, but all units in the set, including the measured unit, are replaced back into the population prior to selection of the next set. The level-1 design has the same replacement policy as the level-0 design except that the unit selected for full measurement is not returned into the population. In general, level-1 sampling design is not unique. It depends on the sequence of judgment ranks in design D. Among all possible level-1 balanced designs with C cycles ($C = n/H$), two special cases can be considered: level-1 ascending ($A1$) order and level-1 descending ($D1$) order. A level-1 ascending order design is given by $D = \{(1, 2, \ldots, H), (1, 2, \ldots, H), \ldots, (1, 2, \ldots, H)\}$, where the integers in each round bracket represents a cycle. In this design, judgment order statistics $X_{[r_j]j}$ in each cycle are collected in ascending order, i.e. $X_{[1]1}$ is measured in the first set, $X_{[2]2}$ is measured in the second set and so on. In a descending order level-1 design, $D = \{(H, H-1, \ldots, 1), (H, H-1, \ldots, 1), \ldots, (H, H-1, \ldots, 1)\}$, the judgment ranked order statistics are selected in descending order in each cycle from largest judgment order statistics to smallest judgment order statistics, i.e., $X_{[H]1}$ is measured in the first set, $X_{[H-1]2}$ is measured in the second set and so on. The level-2 design requires that none of the units in a set, regardless of whether they were measured or not, are replaced back into the population. These designs have similar properties for large population sizes, but they have different behaviors when the population size is small. The level-2 design induces a stronger negative correlation among the sample membership indicators of population units than the other two designs and is usually more efficient. The efficiency improvement is the largest when the ranking process is perfect and diminishes as the ranking information becomes poor.

In the context of survey sampling methodology, the use of level-0, level-1 or level-2 designs induces a probability sampling design in which the first order inclusion probability, $\pi_i = P$ (unit i appears in the sample), depends on replacement level. Inclusion probabilities for level-1 sampling design have been considered by Al-Saleh and Samawi (2007), Ozdemir and Gokpinar (2008). These papers use complete enumeration of all possible sequences of draws which is practical only for small sample sizes. Recently, Frey (2011) provided a recursive algorithm that can be used for relatively large sample sizes. Ozturk and Jafari Jozani (2014) extended the results in Frey (2011) to partially rank-ordered set sample designs.

In finite population settings, there are often more than one auxiliary variables available for the study. For example, Ohio corn yield data uses sampling frames that include four auxiliary variables. All of these auxiliary variables can be used to create judgment strata in an RSS design. On the other hand, a standard ranked set sampling design uses the ranking information of a single auxiliary variable (or a single ranker) and ranking information of all auxiliary variables cannot be combined.

In order to reduce the impact of ranking error, Ozturk (2011) introduced partially rank-ordered set (PROS) sampling designs. The PROS design controls the ranking error by increasing the set size H and reducing the number of ranked units in each set. For a given set size H, a PROS design does not require a full ranking of all the units in each set. Instead, it assigns units into subsets of pre-specified sizes. The units within each subset are not ranked, but each unit in subset h is considered to have smaller rank than the rank of each unit in subset h' for all $h < h'$. The PROS design has been used successfully in an infinite population setting to draw inference. Ozturk (2012a, 2012b) and Gao and Ozturk (2012) used this design to develop nonparametric inference for one- and two-sample problems, respectively. Arslan and Ozturk (2013) developed maximum likelihood estimators for the location and scale parameters in a location-scale family of distributions. Recently, Ozturk (2012b) and Frey (2012) relaxed the assumption that the number of subsets needs to be pre-specified. This provides a flexibility in that the ranker is allowed to declare as many subsets as desired depending on his/her ranking ability. They showed that this flexibility further improves the efficiency of PROS design.

In this paper, we use PROS sampling designs in a finite population setting to achieve two objectives. The first objective is to allow ties in within-set ranking process whenever the units cannot be ranked with high confidence. The second objective is to combine ranking information from all available auxiliary variables (or rankers). Section 2 introduces PROS sampling design within the context of K-auxiliary variables. Section 3 estimates the inclusion probabilities of level-0, level-1 and level-2 sampling designs. Section 4 uses the results in Section 3 to construct estimators for population total and mean. Section 5 provides empirical results to evaluate the efficiency of the proposed estimators. Section 6 develops bootstrap resampling procedures to estimate the variance of the estimators and to construct percentile confidence intervals for the population mean and total. Section 7 applies the proposed estimator to USDA 1992 Ohio corn data. Finally, Section 8 provides a concluding remark.

2. COMBINING AUXILIARY INFORMATION

Suppose we have a finite population of N units labeled as $U = \{u_1, \ldots, u_N\}$. Let x_i be the numerical value of the variable of interest X on the population unit u_i. Without loss of generality, we assume that the population units u_i, $i = 1, \ldots, N$, are ordered with respect to the variable of interest X, i.e., $x_{(1)} < x_{(2)} < \cdots < x_{(N)}$ are the ordered values of the variable X on population units. We also assume that either K different auxiliary random variables ($Y_k, k = 1, \ldots, K$) or K different rankers are available for the study. The auxiliary variables could be either discrete or continuous, but they are easy to measure and have relatively high correlation with the variable of interest X. For example, the auxiliary variables could be the census outcomes, previous survey variables, etc.

In order to construct a partially rank-ordered set sample, $X_{[r_j]j}; 1 \leq r_j \leq H;\ j = 1, \ldots, n$, in a finite population setting, one must first determine the design parameter of the ranked set sample, $D = \{r_1, \ldots, r_n\}$ with $1 \leq r_j \leq H$, $j = 1, \ldots, n$ and $n_h = \sum_{j=1}^{n} I(r_j = h)$, where n_h is the number of fully measured judgment order statistics in judgment class h. The fully measured observation $X_{[r_j]j}$ is measured on a single experimental unit in a set of size H. In order to determine which unit should be selected in a set, its judgment rank, $1 \leq r_j \leq H$, must be determined prior to selection of the set. Once we identify r_j, we select a set of H experimental units, $S_j = \{u_{t_1,j}, \ldots, u_{t_H,j}\}$, at random without replacement from the population U, where $u_{t_h,j}$ is the unit u_{t_h} $(1 \leq t_h \leq N)$ selected from U for the set j. On this set, depending on the availability of the auxiliary variables or the rankers, we obtain either the K auxiliary measurements $\boldsymbol{Y}_{k,j} = (Y_{t_1,k,j}, \ldots, Y_{t_H,k,j})$, $k = 1, \ldots, K$, or K rank vectors, one for each ranker. Assume that we have K-auxiliary variables available. We apply the ranking operator to each one of these vectors of auxiliary measurements to construct the rank vectors

$$\boldsymbol{O}_{k,j} = O(Y_{t_1,k,j}, \ldots, Y_{t_H,k,j}) = \{O_{1,k,j}, \ldots, O_{H,k,j}\}, \quad k = 1, \ldots, K,$$

where $O_{h,k,j}$ is the rank assigned to the unit $u_{t_h,j}$ in rank vector $\boldsymbol{O}_{k,j}$ obtained by applying the ranking operator to random vector $\boldsymbol{Y}_{k,j}$. If the auxiliary variable Y_k is discrete, ranking operator may produce tied ranks. In this case, all tied units are assigned the same rank. If the auxiliary variable Y_k and the response variable X have a negative correlation, the ranks of $\boldsymbol{Y}_{k.j}$ are recorded as $O_{r,k,j} = H + 1 - O_{H+1-r,k,j}$, $r = 1, \ldots, H$, to determine the judgment ranks of the X-observations. The rank vector $\boldsymbol{O}_{k,j}$ becomes the judgment rank vector for the set of X-measurements of units in the set S_j from which we wish to obtain r_jth judgment ranked order statistic $X_{[r_j]j}$. If the researcher, instead of K-auxiliary variable, has K-rankers available, the ranking operator is applied to each ranker to produce judgment ranks

$$\boldsymbol{O}_{k,j} = O(\text{Ranker } k) = \{O_{1,k,j}, \ldots, O_{H,k,j}\}, \quad k = 1, \ldots, K.$$

Table 2. Auxiliary measurements of randomly selected five farms (the numbers have been changed to make the farms not identifiable).

t_h	Y_1	Y_2	Y_3	Y_4	X	$\boldsymbol{O}_1$	$\boldsymbol{O}_2$	$\boldsymbol{O}_3$	$\boldsymbol{O}_4$
76	55	1	16	16	x_{76}	1–2	1–4	2	2
147	1280	3	389	389	x_{147}	5	5	5	5
87	55	1	9	9	x_{87}	1–2	1–4	1	1
119	135	1	77	77	x_{119}	3	1–4	4	4
48	855	1	55	55	x_{48}	4	1–4	3	3

In this process, the ranking operator allows rankers to declare ties whenever the ranks cannot be assigned with high confidence. From this point on, the construction of the sample is the same regardless the type of auxiliary information.

In order to determine which unit should be selected to measure $X_{[r_j]j}$, we combine ranking information in K rank vectors $\boldsymbol{O}_{k,j}$, $k = 1, \dots, K$. We first create an H by H weight matrix $\boldsymbol{W}_{k,j}$ for each judgment rank vector $\boldsymbol{O}_{k,j}$. The rows and columns of this matrix identify the units and assigned judgment ranks, respectively. The entries contain the strength-of-weights of the ranking process. If the ranking vector $\boldsymbol{O}_{k,j}$ has no ties, the hth row and the $O_{h,k,j}$th column of the matrix $\boldsymbol{W}_{k,j}$ will be one, and all the other entries on this row (row h) will be zero, $h = 1, \dots, H$. On the other hand, if the ranking vector $\boldsymbol{O}_{k,j}$ has m tied ranks for the hth unit, then all entries in the hth row that corresponds to the tied ranks will have weight $1/m$, all the other entries in this row will be zero. We combine the ranking information obtained from all auxiliary variables (or rankers) by getting the weighted mean of the strength-of-weight matrices

$$\bar{\boldsymbol{W}}_j = \sum_{k=1}^{K} \alpha_k \boldsymbol{W}_{k,j}, \tag{2.1}$$

where $\sum_{k=1}^{K} \alpha_k = 1$. The weights α_k may be chosen to reflect upon the quality of ranking of auxiliary variables or rankers. For example, if the correlation coefficients ρ_k, $k = 1, \dots, K$, between X and auxiliary variable Y_k are available, we select

$$\alpha_k = \frac{|\rho_k|}{\sum_{k=1}^{K} |\rho_k|}.$$

In multi-ranker design, α_k can be chosen based on the experience of the ranker k. If $\rho_k = 0$ for all k, we select $\alpha_k = 1/H$ for all $k = 1, \dots, H$.

To illustrate the construction of the strength-of-weight matrices, we use Ohio corn yield data. The purpose of survey in Ohio corn yield data is to estimate the mean (or total) corn yields (X) in the county. The correlation coefficients between X and the auxiliary variables Y_k, $\rho_k = \text{cor}(X, Y_k)$, $k = 1, \dots, 4$, are available from the survey data in Table 1. Assume that we wish to measure $X_{[r_1]1}$ with $r_1 = 1$ in a set (S_1) of size $H = 5$. We select five farms from the population at random without replacement, $S_1 = \{u_{76,1}, u_{147,1}, u_{87,1}, u_{119,1}, u_{48,1}\}$. The measurements of auxiliary variables on this set along with their ranks are given in Table 2. In this table, the numbers are altered to make the farms anonymous. In rank vector $\boldsymbol{O}_k$, the dashed integers $a - b$ $(a < b)$ indicate that all ranks between a and b are tied for the corresponding units. For example, in rank vector $\boldsymbol{O}_2$, the units (farms) 76, 87, 119, and 48 are tied for ranks $1, 2, 3$, and 4. We note that X variables are not measured yet at this stage. We only use available auxiliary information (variables) for ranking purposes. The strength-of-weight matrices for the rank vectors $\boldsymbol{O}_1$, $\boldsymbol{O}_2$, $\boldsymbol{O}_3$ and $\boldsymbol{O}_4$ are then given by

$$\boldsymbol{W}_{1,1} = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \qquad \boldsymbol{W}_{2,1} = \begin{bmatrix} 1/4 & 1/4 & 1/4 & 1/4 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1/4 & 1/4 & 1/4 & 1/4 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 & 0 \end{bmatrix},$$

$$\boldsymbol{W}_{3,1} = \boldsymbol{W}_{4,1} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

Information of ranking operators needs to be combined in Equation (2.1). From the correlation coefficients between X and auxiliary variables, the weight vector $\boldsymbol{\alpha}$ is given by $\boldsymbol{\alpha} = (0.252, 0.208, 0.263, 0.276)$. Hence, the combined ranking information for this set yields

$$\bar{\boldsymbol{W}}_1 = \begin{bmatrix} 0.178 & 0.717 & 0.052 & 0.052 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.000 & 0.999 \\ 0.717 & 0.178 & 0.052 & 0.052 & 0.000 \\ 0.052 & 0.052 & 0.304 & 0.591 & 0.000 \\ 0.052 & 0.052 & 0.591 & 0.304 & 0.000 \end{bmatrix}.$$

The combined ranking information matrix $\bar{\boldsymbol{W}}_1$ constructed for set S_1 is a doubly stochastic matrix. The hth row of this matrix, $\bar{\boldsymbol{w}}_{h,1} = (\bar{w}_{t_h,1,1}, \ldots, \bar{w}_{t_h,H,1})^\top$, indicates how strongly the ranking operators $O_1, \ldots, O_K$ agree to assign ranks $1, \ldots, H$ to units u_{t_h}. One may also interpret $\bar{\boldsymbol{w}}_{h,1}$ as a probability vector that the ranking operators collectively assign ranks $1, \ldots, H$ to the unit $u_{t_h,1}$. In our example, 17.7 %, 71.7 %, 5.2 %, 5.2 % and 0 % of ranking operators declare that the unit u_{76} has ranks 1, 2, 3, 4, and 5, respectively.

In order to determine the unit on which the judgment order statistics $X_{[r_j]j}$, $1 \le r_j \le H$, will be measured, we use the combined ranking information matrix $\bar{\boldsymbol{W}}_j$. In this matrix, we determine the unit that has the largest agreement probability (weight) in the column r_j. Note that the judgment rank r_j is determined prior to ranking process to avoid any possible bias. If the largest strength-of-agreement weight in column r_j is unique then the unit that corresponds to the largest entry is selected for measurement. If the largest entry is not unique, we then select the unit having the largest concentration of the strength probabilities around the judgment rank r_j. Even if the units having largest concentration is not unique we break the tie at random to select one of the units having equal concentration value. The amount of concentration of agreement weights on the unit h around judgment rank r_j can be measured by

$$\tau_{r_j,h} = \sum_{s=1}^{H}(s - r_j)^2 \bar{w}_{h,s},$$

where $\bar{w}_{h,s}$, $s = 1, \ldots, H$, is the hth row of the matrix $\bar{\boldsymbol{W}}_j$. The expression $\tau_{r_j,h}$ provides information about how strongly the strength-of-agreement weight vector is concentrated around the judgment ranks r_j for unit h.

In Ohio corn yield survey example, we fixed $r_1 = 1$ prior to the ranking process. We then need to identify the maximum entry, 0.717, in column $r_1 = 1$ of matrix $\bar{\boldsymbol{W}}_1$. The location of this maximum entry ($h = 3$) indicates that the judgment ranked observation $X_{[r_1]1}$ must be obtained from unit u_{87} in set S_1. The observed data for this set then becomes

$$\left\{X_{[1]1}, \bar{\boldsymbol{w}}^*_{r_1}\right\} = \left\{X_{[1]1}, (0.717, 0.178, 0.052, 0.052, 0.000)\right\},$$

where $\bar{\boldsymbol{w}}^*_{r_1}$ is the row that contains the largest entry in column r_1 of matrix $\bar{\boldsymbol{W}}_1$. This construction does not only measure $X_{[1]1}$ but also provides a weight vector to relate $X_{[1]1}$ to judgment ranks through the combined ranking information of all auxiliary variables. Using the same principle, one can construct the PROS ranked set sample for a given design D

$$\left\{X_{[r_j]j}, \bar{\boldsymbol{w}}^*_{r_j}\right\}, \quad j = 1, \ldots, n.$$

It is important to realize that $X_{[r_j]j}$ and $\bar{\boldsymbol{w}}^*_{r_j}$ are dependent random variable and vectors, respectively. We note that the selection of sets in the construction of a ranked set sample depends on without replacement level-0, level-1 and level-2 designs.

3. ESTIMATION OF INCLUSION PROBABILITIES

Inclusion probabilities of a PROS sample will depend on the replacement policy and ranking information provided by auxiliary ranking variables. In this section, we construct estimators for the inclusion probabilities of level-0, level-1 and level-2 designs.

Level-0 PROS Design: Note that in level-0 design all units in a set are returned back to the population prior to the selection of the next set. If the ranking process uses only one auxiliary variable with no tie structure, the first order inclusion probabilities are given in Frey (2011). Let $\beta(i, h, H, N)$ be the probability that the ith smallest item in the population is the hth smallest measured unit ($X_{(h)j}$) in set S_j. This probability is given by

$$\beta(i, h, H, N) = \frac{\binom{i-1}{h-1}\binom{N-i}{H-h}}{\binom{N}{H}}.$$

Since our ranking mechanism is not perfect, the strength-of-agreement weights induces a probability distribution on judgment ranks. By conditioning on these strength-of-agreement weights $\bar{\boldsymbol{w}}^*_{r_j}$, the estimate of the probability that the ith smallest unit in the population is the r_jth judgment order statistic in set S_j is given by

$$\beta^*(i, r_j, H, N) = \sum_{h=1}^{H} \bar{w}^*_{r_j,h}\beta(i, h, H, N).$$

The estimate of the first and second order inclusion probabilities then follows from Frey (2011) by using $\beta^*(i, r_j, H, N)$.

Lemma 1. *For level-*0 *PROS sampling design, the estimate of the first order inclusion probability of the i-th unit in the population is given by*

$$\pi_i^{(0)}(K) = 1 - \prod_{j=1}^{n}\bigl\{1 - \beta^*(i, r_j, H, N)\bigr\}, \tag{3.1}$$

where K indicates that estimate depends on the number of the auxiliary variables through the weight vectors $\bar{\mathbf{w}}^*_{r_j}$, $j = 1, \ldots, n$. Also, the estimate of the second order inclusion probability of the ith and i'th units $(i < i')$ is given by

$$\begin{aligned}\pi_{i,i'}^{(0)}(K) = 1 &- \prod_{j=1}^{n}\bigl\{1 - \beta^*(i, r_j, H, N)\bigr\} - \prod_{j=1}^{n}\bigl\{1 - \beta^*\bigl(i', r_j, H, N\bigr)\bigr\} \\ &+ \prod_{j=1}^{n}\bigl\{1 - \beta^*(i, r_j, H, N) - \beta^*\bigl(i', r_j, H, N\bigr)\bigr\}.\end{aligned} \tag{3.2}$$

Level-1 PROS Design: The construction of level-1 design and estimation of the inclusion probabilities strongly depend on the sequence in which the measured units enter the sample. Frey (2011) gives an iterative algorithm to compute the inclusion probabilities of a level-1 design under perfect ranking (or under a ranking mechanism based on a single auxiliary variable). We use the same iterative procedure conditionally for a given strength-of-agreement weight vector $\bar{w}^*_{r_j}$, $j = 1, \ldots, n$. Let $Z_i(t, j)$ be the estimate of the probability that the first j sampled units include t units smaller than the ith unit of the population, but not the ith unit. It is clear that

$$Z_i(t, 0) = \begin{cases} 1, & t = 0, \\ 0, & \text{otherwise}. \end{cases} \tag{3.3}$$

If we know $\{Z_i(t, j), 0 \le t \le i - 1\}$ for a fixed value of $j \ge 0$, we can compute the values $\{Z_i(t, j + 1), 0 \le t \le i - 1\}$ via a simple recursion. First consider the construction of $Z_i(t, j + 1)$ from $\{Z_i(t, j), 0 \le t \le i - 1\}$. The quantity $Z_i(t, j + 1)$ indicates that among the first $j + 1$ selected units from the population, we must have t fully measured units smaller than the ith unit. This can happen in two ways: (1) In the first j selected units, we fully measure t units smaller than the ith unit in the population and select a unit larger than the ith unit at the next selection. (2) In the first j selected units, we fully measure $t - 1$ units smaller than the ith unit in the population and then select a unit smaller than the ith unit on the next draw. By using these two cases, we write

$$Z_i(t, j+1) = Z_i(t, j) \sum_{\lambda=i+1-t}^{N-j} \beta^*(\lambda, r_j, H, N-j) + Z_i(t-1, j) \sum_{\lambda=1}^{i-t} \beta^*(\lambda, r_j, H, N-j). \tag{3.4}$$

When $t = 0$ we set $Z_i(-1, j) = 0$ for all $j = 0, \ldots, n$. Through this recursive equation, after we compute the $\{Z_i(t, n), 0 \leq t \leq i-1\}$, we find that the estimate of the probability that a sample of fully measured units does not include the ith smallest unit in the population is $\sum_{t=0}^{i-1} Z_i(t, n)$. Then the estimate of the first order inclusion probability of level-1 PROS design is given by

$$\pi_i^{(1)}(K) = 1 - \sum_{t=0}^{i-1} Z_i(t, n).$$

To compute the estimate of the first order inclusion probabilities, we use the following algorithm.

Algorithm 1: (Adopted from Frey 2011).
Step 1. Compute $\{Z_i(j, 0), 0 \leq j \leq i-1\}$ using Equation (3.3).
Step 2. For $k = 0, \ldots, n-1$, compute $\{Z_i(j, k+1), 0 \leq j \leq i-1\}$ from $\{Z_i(j, k), 0 \leq j \leq i-1\}$ using Equation (3.4).
Step 3. Compute $\pi_i^{(1)}(K)$ as $\pi_i^{(1)}(K) = 1 - \sum_{j=0}^{i-1} Z_i(j, n)$.

The estimate of the second order inclusion probabilities based on the strength-of-agreement weights can be computed from the recursive equation in Frey (2012) with a slight modification in the notation.

Level-2 PROS Designs: The estimation of the inclusion probability of level-2 designs, with a slight modification, follows from Patil, Sinha, and Taillie (1995) or from Frey (2011). Frey (2011) provides a general version of the results presented in Patil, Sinha, and Taillie (1995). These derivations assume perfect ranking or a single auxiliary variable with no tie-structure. Under our ranking scheme, estimation of the inclusion probabilities conditionally on the strength-of-agreement weights follow from their derivation with mild changes in the notation:

$$\pi_i^{(2)}(K) = \sum_{j=1}^{n} \beta^*(i; r_j, H, N). \tag{3.5}$$

For $i < i'$, the second order inclusion probability can be estimated from

$$\pi_{i,i'}^{(2)}(K) = \sum_{j=1}^{n} \sum_{j'=1}^{n} \sum_{h=1}^{H} \sum_{h'=1}^{H} \bar{w}^*_{r_j,h} \bar{w}^*_{r'_j,h'} \times \sum_{\lambda=0}^{i'-i-1} \frac{\binom{i-1}{h-1}\binom{i'-i-1}{\lambda}\binom{N-i'}{H-\lambda-h}\binom{i'-1-h-\lambda}{h'-1}\binom{N-i'-H+\lambda+h}{H-h'}}{\binom{N}{H}\binom{N-H}{H}}.$$

Marginal distribution of these inclusion probabilities requires marginal distribution of the weight vector $\bar{\boldsymbol{w}}^*_{r_j}$. Since the weight vectors are constructed based on the ranking quality of the auxiliary information, their marginal distributions are not trivial and require strong modeling assumptions on ranking process. To avoid strong modeling assumptions, we first establish the empirical properties of the estimators and then provide bootstrap inference for a given finite population.

4. ESTIMATION OF THE POPULATION MEAN

In a finite population setting, if the sample is selected without replacement, Horvitz–Thompson (1952) estimator provides an unbiased estimator for the population mean (or for the population total). If $\pi_1, \ldots, \pi_N$ are the inclusion probabilities for the N population units and $X_1, \ldots, X_n$ are the simple random sample, the Horvitz–Thompson estimator of the population mean is given by

$$\hat{\mu} = \frac{1}{N} \sum_{i=1}^{N} \frac{I_i x_i}{\pi_i},$$

where I_i is an indicator function taking values 1 or zero depending on whether the unit u_i is included in the sample or not. Estimation of the population total is obtained by multiplying the mean estimate with the population size. Thus, in this paper we only consider the estimation of the population mean. These results can be extended to estimation of the population total with minor modification.

In a ranked set sampling setting, Horvitz–Thompson estimator depends on the sampling designs. Let $\hat{\mu}_L(K)$ be the Horvitz–Thompson estimator based on K auxiliary variables and level-L design, where $L = 0$, $A1$, $D1$, and 2 indicate level-0, level-1 ascending order, level-1 descending order and level-2 sampling designs, respectively. For example, Horvitz–Thompson estimator based on level-0 design can be written as

$$\hat{\mu}_0(K) = \frac{1}{N} \sum_{i=1}^{N} \frac{I_i x_i}{\pi_i^{(0)}(K)}.$$

The other estimators can be written in a similar fashion by using appropriate inclusion probabilities. We note that the Horvitz–Thompson estimator in PROS sampling, unlike simple random sampling, are conditional estimator for given values of weight vector $\bar{w}^*_{r_j}$, $j = 1, \ldots, n$.

The Horvitz–Thompson estimator does not directly use the strength-of-agreement weights provided by auxiliary information in ranking process. We now introduce a new estimator for the population mean (or the population total) based on level-L design, $L = 0, A1, D1, 2$. The new estimator uses the strength-of-agreement weights to incorporate ranking information of all auxiliary variables. Let $V_{j,h} = \bar{w}^*_{r_j,h} X_{[r_j]j}$, $j = 1, \ldots, n$. One can interpret that $V_{j,h}$ pro-rates each measured observation into the hth judgment class distribution proportional to the strength-of-agreement weight $\bar{w}^*_{r_j,h}$. For level-1 ascending order (A1) sampling design, the new estimator is given by

$$\mu^*_{A1}(K) = \frac{1}{H}\sum_{h=1}^{H} \frac{1}{\sum_{j=1}^{n} \frac{\bar{w}^*_{r_j,h}}{\pi_j^{(A1)}(K)}} \sum_{j=1}^{n} \frac{V_{j,h}}{\pi_j^{(A1)}(K)}. \tag{4.1}$$

The estimators for other sampling designs, $\mu^*_0(K)$, $\mu^*_{D1}(K)$, $\mu^*_2(K)$, can be defined in a similar fashion by using the corresponding estimated inclusion probabilities, $\pi_j^{(0)}(K)$, $\pi_j^{(D1)}(K)$ and $\pi_j^{(2)}(K)$ in Equation (4.1).

5. EMPIRICAL RESULTS

In this section, we provide evidence to evaluate the empirical efficiency of the proposed estimators. The quality of ranking information of ranking operators is modeled by using the Dell and Clutter (1972) model. To rank the variable of interest X, the Dell and Clutter model uses its concomitant variable Y (or auxiliary variable Y_k, $k = 1, \ldots, K$). The model assumes that the variable X and its concomitant Y have a correlation coefficient ρ. In a given set of size H, units in the set is first ranked with respect to concomitant variable Y_k and the ranks of Y_k-observations are taken as the judgment rank of the X-observations, for $k = 1, \ldots, K$. The quality of ranking information is controlled by the magnitude of the correlation coefficient between the X and Y_k variables. Dell and Clutter model does not introduce tie structure in the ranking process. Fligner and MacEachern (2006) introduced a class of ranking procedures under stochastic order restriction. In certain cases, this class introduces a tie structure in the ranking mechanism. Most recently, Frey (2012) used discretized tie-structure model to assign ties. In this model, after the set is ranked based on the concomitant variable Y_k, $(Y_{(1),k} < Y_{(2),k} \cdots < Y_{(H),k})$, each value $Y_{(h),k}/\tau + 0.5$ is applied to flooring function $\lfloor Y_{(h),k}/\tau + 0.5 \rfloor$ if $Y_{(h),k} > 0$ and each value $Y_{(h),k}/\tau - 0.5$ is applied to ceiling function $\lceil Y_{(h),k}/\tau - 0.5 \rceil$ if $Y_{(h)} < 0$, where τ is a positive constant that controls the amount of ties in a set. Large values of τ produce larger number of ties in a set.

The first part of the simulation study investigates the properties of the estimated first order inclusion probabilities of level-1 and level-2 designs. Data sets are generated from a discrete normal population of size $N = 50$ with sample size $n = 15$ and set size $H = 3$. Discrete normal population is generated by $x_i = Q((i - 0.5)/50)$, where Q is the quantile function of standard normal distribution. The other simulation parameters are taken to be $K = 2, 5$ and $\rho = 1, 0.75$. Figure 1 presents the simulated (Sim) and estimated (Est) first order inclusion probabilities of level-1 and level-2 designs. Simulated inclusion probabilities are obtained based on one million replications while the estimated inclusion probabilities are computed from Algorithm 1 and Equation (3.5) based on 10,000 simulation replications. Figure 1 indicates that inclusion probabilities in level-1 design show different behaviors depending on whether the sample has ascending or descending order. Under perfect ranking ($\rho = 1$), both simulated and estimated inclusion probabilities are smooth and almost identical. On the other hand, under ranking error ($\rho = 0.75$), simulated inclusion probabilities have a lot of spikes. The estimated inclusion probabilities smooth these

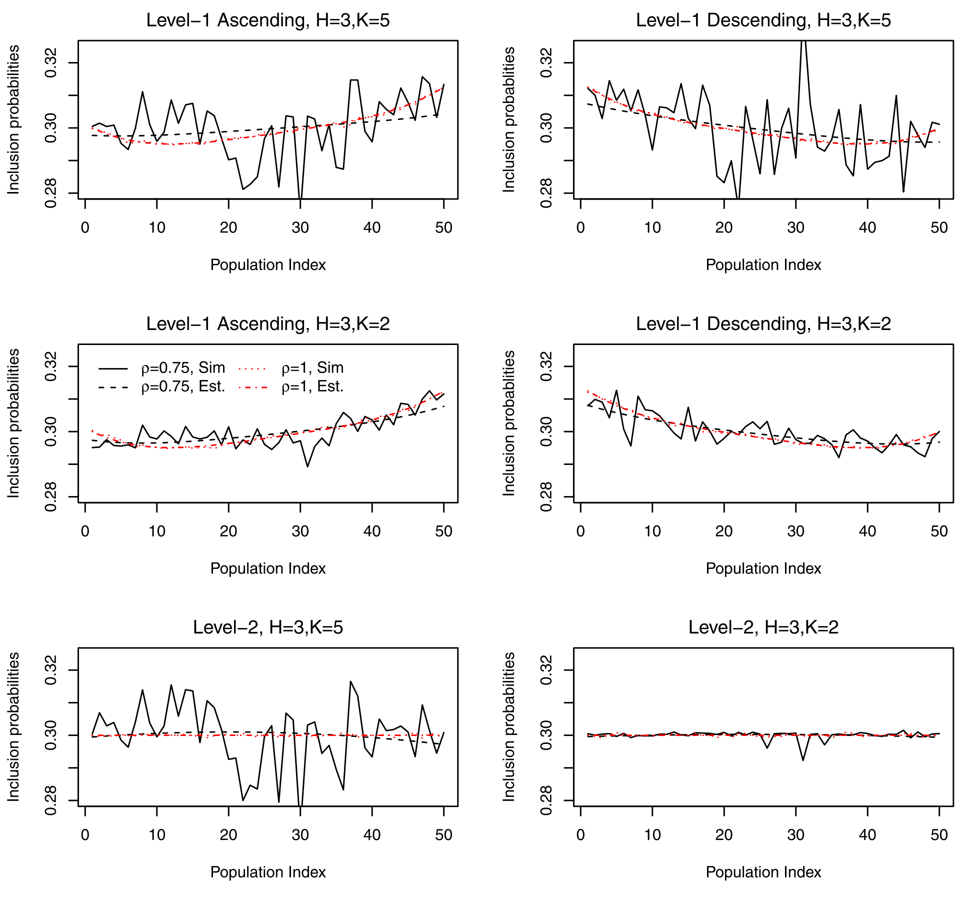

Figure 1. Simulated (Sim) and estimated (Est) inclusion probabilities for level-1 and level-2 designs, sample size $n = 15$. Solid line is for $\rho = 0.75$ and Sim, dashed line is for $\rho = 0.75$ and Est, dotted line is for $\rho = 1$ and Sim, and dashed-dotted line is for $\rho = 1$ and Est. The same ρ is used for all K rankers.

spikes. The simulated inclusion probabilities produce sharper spikes for $K = 5$ than for $K = 2$. Similar results also appear in level-2 design.

We next performed a series of simulation studies to investigate the properties of sampling designs and the estimators. In this part of the simulation, the data sets are generated from a discrete normal population of size $N = 180$ with sample size $n = 30$, set size $H = 3$ and the number of auxiliary variables $K = 3$. Discrete normal population is generated by $x_i = Q((i - 0.5)/180)$, where Q is the quantile function of standard normal distribution. The first simulation compares the sampling designs. The simulation parameters include the set size H ($H = 3$), cycle size C ($C = n/H = 10$), correlation coefficient ρ ($\rho = 0.25, 0.50, 0.75, 1.00$), and discretization parameter τ ($\tau = 0, 1, 1.5$). For notational convenience, $\tau = 0$ is used for no tie structure. Since the computing time for the inclusion probabilities in level-1 design is intensive, for each combination of the simulation parameters, 500 data sets are generated for each sampling design L ($L = 0, A1, D1,$ and 2). The relative efficiencies in mean squared errors (MSE) for Horvitz–Thompson ($R_L(K)$) and the proposed new estimators ($R^*_L(K)$) with respect to simple random sample estimator are

Table 3. Relative efficiencies of the estimators $\hat{\mu}_L$ and μ^*_L with respect to simple random sample mean estimator $\hat{\mu}$, $R_L(K) = MSE(\hat{\mu})/MSE(\hat{\mu}_L(K))$, $R^*_L(K) = MSE(\hat{\mu})/MSE(\mu^*_L(K))$, for level-$L$ sampling designs, $L = 0$, $A1$, $D1$, 2, and $K = 3$.

τ	μ	ρ	$R_0(3)$	$R_{A1}(3)$	$R_{D1}(3)$	$R_2(3)$	$R^*_0(3)$	$R^*_{A1}(3)$	$R^*_{D1}(3)$	$R^*_2(3)$
0.0	0	0.25	1.023	0.982	1.011	1.028	0.952	0.971	0.988	1.012
0.0	0	0.50	1.205	1.433	1.313	1.410	1.184	1.418	1.302	1.402
0.0	0	0.75	1.448	1.850	1.689	1.966	1.413	1.856	1.643	1.968
0.0	0	1.00	1.755	1.997	1.972	2.277	1.662	1.985	1.958	2.277
0.0	100	0.25	0.001	0.034	0.037	0.028	0.854	0.959	1.138	0.996
0.0	100	0.50	0.001	0.045	0.044	0.047	1.371	1.248	1.392	1.751
0.0	100	0.75	0.001	0.061	0.059	0.056	1.349	1.751	1.677	1.728
0.0	100	1.00	0.001	3.825	0.949	2.645	1.573	1.916	1.746	2.645
0.0	1000	0.25	0.000	0.000	0.000	0.000	0.892	1.046	1.209	0.995
0.0	1000	0.50	0.000	0.001	0.001	0.000	1.167	1.441	1.359	1.650
0.0	1000	0.75	0.000	0.001	0.001	0.001	1.545	1.936	1.828	2.397
0.0	1000	1.00	0.000	0.144	0.067	2.288	1.623	2.011	1.863	2.288
1.0	0	0.25	0.994	1.148	1.014	1.234	0.962	1.138	0.994	1.222
1.0	0	0.50	1.169	1.474	1.415	1.549	1.087	1.509	1.420	1.537
1.0	0	0.75	1.458	1.871	1.767	2.278	1.433	1.883	1.831	2.305
1.0	0	1.00	1.605	2.058	2.056	2.756	1.691	2.042	1.998	2.710
1.0	100	0.25	0.001	0.048	0.046	0.037	0.945	1.097	1.039	1.072
1.0	100	0.50	0.001	0.062	0.045	0.042	1.096	1.283	1.312	1.275
1.0	100	0.75	0.001	0.074	0.072	0.071	1.492	1.778	1.791	2.337
1.0	100	1.00	0.001	0.054	0.053	0.049	1.723	2.031	2.190	2.430
1.0	1000	0.25	0.000	0.001	0.001	0.000	0.852	1.040	1.131	1.012
1.0	1000	0.50	0.000	0.001	0.001	0.001	1.374	1.474	1.354	1.656
1.0	1000	0.75	0.000	0.001	0.001	0.001	1.643	1.803	1.624	1.961
1.0	1000	1.00	0.000	0.000	0.001	0.001	1.688	2.034	1.770	2.455
1.5	0	0.25	0.922	1.037	1.095	1.155	0.859	1.016	1.093	1.127
1.5	0	0.50	1.063	1.291	1.393	1.431	1.031	1.292	1.387	1.436
1.5	0	0.75	1.514	1.953	2.082	2.250	1.596	2.034	2.087	2.285
1.5	0	1.00	1.718	2.198	1.988	2.534	1.727	2.224	1.970	2.572
1.5	100	0.25	0.001	0.040	0.039	0.030	0.898	0.974	1.013	1.075
1.5	100	0.50	0.001	0.063	0.061	0.045	1.112	1.487	1.282	1.413
1.5	100	0.75	0.001	0.089	0.081	0.077	1.699	1.952	1.838	2.213
1.5	100	1.00	0.001	0.045	0.040	0.038	1.668	2.017	1.907	2.197
1.5	1000	0.25	0.000	0.000	0.000	0.000	0.853	0.984	1.061	1.043
1.5	1000	0.50	0.000	0.001	0.001	0.001	1.112	1.366	1.325	1.383
1.5	1000	0.75	0.000	0.001	0.001	0.001	1.611	1.845	1.735	2.392
1.5	1000	1.00	0.000	0.000	0.000	0.000	1.801	1.870	1.964	2.531

given by

$$R_L(K) = \frac{MSE(\hat{\mu})}{MSE(\hat{\mu}_L(K))}, \qquad R^*_L(K) = \frac{MSE(\hat{\mu})}{MSE(\mu^*_L(K))}, \quad \text{for } L = 0, A1, D1, 2.$$

The relative efficiencies $R_L(K)$ and $R^*_L(K)$ greater than one indicate that the Horvitz–Thompson (HT) and proposed estimators based on level-L design have higher efficiencies than simple random sample mean estimator, respectively.

Table 3 presents these relative efficiency results. There are several important features in Table 3 that needs to be emphasized. The Horvitz–Thompson estimator based on all sampling designs is very sensitive to the coefficient of variation (σ/μ) when there is randomness in ranking process through either the tie structure ($\tau = 1$ and 1.5) or judgment ranking error ($\rho < 1$). Since the data sets are generated from discretized normal distribution with $\sigma = 1$, Table 3 contains the values of μ instead of coefficient of variation. It is clear from Table 3 that the efficiency of the Horvitz–Thompson estimator with respect to simple random sample mean estimator could be zero for small coefficient of variation. For example, when the coefficient of variation is close to zero (or $\mu = 100$ or $\mu = 1000$) in Table 3, relative efficiencies of the HT-estimator based on all replacement designs (columns 4, 5, 6 and 7) are less than 1. Only exceptions to this are the cases when $\tau = 0$ and $\rho = 1$ where efficiencies R_{A1} and R_2 are 3.825 and 2.645, respectively, when $\mu = 100$ and the efficiency R_2 is 2.288 when $\mu = 1000$. For large values of the coefficient of variation, for example when the coefficient of variation is infinite ($\mu = 0$), the efficiency of HT-estimator for all designs appear to be an increasing function of the correlation coefficient. Also, when the coefficient of variation is infinite ($\mu = 0$), the efficiency of the HT estimator based on level-2 design is larger than the efficiency of the same estimator based on all other replacement designs for $\tau = 0, 1, 1.5$.

On the other hand, the proposed estimator $\mu_L^*(K)$, unlike HT estimator, appears to be invariant with respect to the coefficient of variation for all level-L sampling designs, $L = 0, A1, D1, 2$. The efficiencies are increasing function of the correlation coefficient. It appears that a moderate tie structure improves the efficiency of the new estimators for all sampling designs. Among the four sampling designs, level-2 design provides the highest efficiency. For example the last column in Table 3 is in general as large as or larger than the entries in other columns, indicating that the proposed estimator based on level-2 design outperforms the other estimators in Table 3. In the remaining part of the paper, we thus only focus on the level-2 design.

To evaluate the performance of the estimator $\mu_2^*(K)$ on different populations, we performed another simulation study. In this part of the simulation, data sets are generated from discrete normal, exponential and Cauchy distributions. Since the coefficient of variation is an important factor that affects the efficiency of the HT estimator, we shifted the populations by adding $\mu = 0, 100, 1000$ to each one of the discretized population of size 220 from standard normal, standard exponential and standard Cauchy distributions. Discretized populations are generated by using the quantile function $x_i = Q((i-0.5)/220), i = 1, \ldots, 220$ for normal, exponential and Cauchy distributions. For the simulation parameters we use set size $H = 3$, cycle size $C = 10$, the number of auxiliary variables $K = 1, 4$, discretization parameter $\tau = 0, 1, 1.5$ and correlation coefficient $\rho = \pm 0.5, \pm 0.75, \pm 1.00$. Since simulation results for negative and positive correlations are almost identical up to the simulation variation, we only report the results for positive correlations. The simulation size is selected to be 2000. In this part of the simulation we are primarily interested in the biases and the efficiencies of the Horvitz–Thompson ($\hat{\mu}_2(K)$) and the proposed ($\mu_2^*(K)$) estimators under the ranking of a single auxiliary variable ($K = 1$) and multiple auxiliary variables ($K = 4$). Relative efficiencies of Horvitz–Thompson and the new estimator for level-2 design are given by

$$R_2(K) = \frac{MSE(\hat{\mu})}{MSE(\hat{\mu}_2(K))}, \qquad R_2^*(K) = \frac{MSE(\hat{\mu})}{MSE(\mu_2^*(K))}.$$

Table 4. Biases and relative efficiencies of the level-2 design Horvitz–Thompson ($\hat{\mu}_2(K)$) and the proposed estimator ($\mu_2^*(K)$) based on K auxiliary variables. The efficiencies are given in terms of the ratio of the MSEs of SRS and level-2 design estimators, i.e. $R(K) = MSE(\hat{\mu})/MSE(\hat{\mu}_2(K))$, $R^*(K) = MSE(\hat{\mu})/MSE(\mu_2^*(K))$. Data sets are generated from discretized normal distributions, $N(0, 1) + \mu$.

			Biases				Efficiencies			
τ	μ	ρ	$\hat{\mu}_2(1)$	$\hat{\mu}_2(4)$	$\hat{\mu}_2^*(1)$	$\hat{\mu}_2^*(4)$	$R(1)$	$R(4)$	$R^*(1)$	$R^*(4)$
0.00	0	0.50	0.001	−0.000	0.001	−0.001	1.116	1.471	1.169	1.471
0.00	0	0.75	−0.003	−0.001	−0.003	−0.001	1.357	2.269	1.479	2.348
0.00	0	1.00	−0.002	0.000	−0.002	0.000	2.250	2.153	2.250	2.153
0.00	100	0.50	0.293	−0.025	0.001	0.005	0.029	0.047	1.098	1.486
0.00	100	0.75	0.307	−0.081	0.001	0.003	0.032	0.063	1.363	2.206
0.00	100	1.00	−0.004	−0.002	−0.004	−0.002	2.372	2.255	2.372	2.255
0.00	1000	0.50	2.788	−0.533	−0.000	−0.000	0.000	0.000	1.081	1.388
0.00	1000	0.75	3.062	−0.616	−0.002	0.004	0.000	0.001	1.402	2.110
0.00	1000	1.00	−0.001	0.002	−0.001	0.002	2.241	2.232	2.241	2.232
1.00	0	0.50	−0.002	0.000	−0.002	−0.000	1.079	1.405	1.122	1.415
1.00	0	0.75	0.003	−0.004	0.002	−0.004	1.339	2.173	1.520	2.228
1.00	0	1.00	−0.000	0.003	0.000	0.002	1.605	2.395	2.329	2.349
1.00	100	0.50	0.215	−0.055	−0.004	0.004	0.035	0.056	1.121	1.625
1.00	100	0.75	0.219	−0.118	0.002	0.001	0.044	0.078	1.527	2.437
1.00	100	1.00	0.379	−0.137	−0.001	0.002	0.029	0.047	2.201	2.416
1.00	1000	0.50	2.286	−0.480	0.000	−0.001	0.000	0.001	1.130	1.489
1.00	1000	0.75	2.298	−0.999	−0.003	−0.005	0.000	0.001	1.410	2.241
1.00	1000	1.00	3.427	−1.193	0.001	−0.003	0.000	0.001	2.281	2.446
1.50	0	0.50	0.007	−0.002	0.007	−0.002	1.089	1.632	1.163	1.640
1.50	0	0.75	−0.001	−0.004	−0.001	−0.004	1.254	2.132	1.495	2.214
1.50	0	1.00	0.001	0.000	0.001	0.001	1.303	2.350	2.001	2.364
1.50	100	0.50	0.237	−0.039	0.003	−0.005	0.039	0.055	1.130	1.597
1.50	100	0.75	0.226	−0.131	0.001	−0.004	0.045	0.087	1.523	2.418
1.50	100	1.00	0.470	−0.150	−0.000	−0.002	0.019	0.042	1.985	2.359
1.50	1000	0.50	2.336	−0.155	0.001	−0.000	0.000	0.001	1.091	1.537
1.50	1000	0.75	2.316	−1.208	0.003	0.002	0.000	0.001	1.408	2.391
1.50	1000	1.00	4.670	−1.682	0.007	−0.003	0.000	0.000	2.038	2.333

The empirical estimate of the biases and efficiencies are presented in Tables 4, 5 and 6 for discrete normal, exponential, and Cauchy distributions, respectively. It is clear from these tables that both HT ($\hat{\mu}_2(K)$) and proposed ($\mu_2^*(K)$) estimators appear to be unbiased for large coefficient of variation regardless of the value of τ and ρ. On the other hand, when the coefficient of variation is small, the HT estimator is moderately biased if the quality of ranking information is poor due to either smaller correlation between the auxiliary variables and response (ρ) or the large number of ties in the ranking process (larger τ). The proposed estimator ($\mu_2^*(K)$) is essentially unbiased regardless the values of coefficient of variation, τ and ρ for both single and multi auxiliary variables.

Similar results also hold for the efficiencies. For large coefficient of variation, the HT estimator always outperforms simple random sample estimator $\hat{\mu}$. On the other hand, when the coefficient of variation is small, the efficiency of HT estimator with respect to $\hat{\mu}$ could essentially be zero for both single and multi-auxiliary variables. The proposed estimator

Table 5. Biases and relative efficiencies of the level-2 design Horvitz–Thompson ($\hat{\mu}_2(K)$) and the proposed estimator ($\mu_2^*(K)$) based on K auxiliary variables. The efficiencies are given in terms of the ratio of the MSEs of SRS and level-2 design estimators, i.e., $R(K) = MSE(\hat{\mu})/MSE(\hat{\mu}_2(K))$, $R^*(K) = MSE(\hat{\mu})/MSE(\mu_2^*(K))$. Data sets are generated from discretized exponential distributions, $\text{Exp}(1)+\mu$.

			Biases				Efficiencies			
τ	μ	ρ	$\hat{\mu}_2(1)$	$\hat{\mu}_2(4)$	$\hat{\mu}_2^*(1)$	$\hat{\mu}_2^*(4)$	$R(1)$	$R(4)$	$R^*(1)$	$R^*(4)$
0.00	1	0.50	−0.003	0.005	−0.006	0.003	1.057	1.469	1.096	1.497
0.00	1	0.75	0.009	−0.004	0.005	−0.003	1.194	1.870	1.282	1.945
0.00	1	1.00	−0.001	−0.004	−0.001	−0.004	1.852	1.772	1.852	1.772
0.00	101	0.50	0.260	0.021	0.001	0.001	0.030	0.042	1.122	1.470
0.00	101	0.75	0.297	−0.084	−0.000	0.002	0.029	0.048	1.290	1.814
0.00	101	1.00	−0.002	−0.005	−0.002	−0.005	1.800	1.752	1.800	1.752
0.00	1001	0.50	2.754	−0.464	−0.003	0.007	0.000	0.000	1.130	1.512
0.00	1001	0.75	2.985	−0.408	−0.004	0.000	0.000	0.001	1.399	2.016
0.00	1001	1.00	−0.000	−0.001	−0.000	−0.001	1.911	1.924	1.911	1.924
1.00	1	0.50	−0.001	−0.006	−0.005	−0.006	1.004	1.414	1.074	1.453
1.00	1	0.75	0.004	−0.006	−0.000	−0.006	1.275	1.999	1.448	2.126
1.00	1	1.00	0.004	−0.001	−0.004	−0.002	1.358	1.926	1.892	1.952
1.00	101	0.50	0.279	−0.038	0.004	0.004	0.030	0.049	1.040	1.456
1.00	101	0.75	0.230	−0.086	0.001	−0.001	0.040	0.070	1.367	2.137
1.00	101	1.00	0.445	−0.121	−0.003	−0.002	0.022	0.042	1.870	2.001
1.00	1001	0.50	2.363	−0.318	−0.002	−0.003	0.000	0.001	1.145	1.508
1.00	1001	0.75	2.620	−1.040	0.000	−0.003	0.000	0.001	1.328	1.994
1.00	1001	1.00	3.851	−1.248	−0.007	−0.007	0.000	0.000	1.984	2.073
1.50	1	0.50	−0.001	0.001	−0.004	0.001	1.092	1.459	1.163	1.497
1.50	1	0.75	0.004	0.000	0.001	0.001	1.099	1.713	1.269	1.806
1.50	1	1.00	0.005	0.002	−0.005	0.001	1.206	1.915	1.813	1.942
1.50	101	0.50	0.198	−0.066	−0.006	0.004	0.036	0.054	1.170	1.485
1.50	101	0.75	0.206	−0.100	−0.005	0.003	0.044	0.075	1.428	2.111
1.50	101	1.00	0.450	−0.104	−0.005	−0.008	0.017	0.037	1.736	1.983
1.50	1001	0.50	2.495	−0.112	0.001	0.004	0.000	0.001	1.048	1.469
1.50	1001	0.75	2.338	−0.787	0.007	0.004	0.000	0.001	1.377	2.224
1.50	1001	1.00	4.362	−1.141	−0.004	−0.008	0.000	0.000	1.658	2.140

always outperforms the simple random sample estimator $\hat{\mu}$ and HT estimator $\hat{\mu}_2(K)$ for all coefficient of variation, τ and ρ in Tables 4, 5 and 6.

In general, the efficiency of $\mu_2^*(K)$ increases with the number of auxiliary variables K with possible exceptions when $\rho = 1$ and $\tau = 0$. For larger K, the information content of the strength-of-weight matrix $\bar{\boldsymbol{W}}_j$ increases. Hence, ranking process becomes more accurate. This creates big separation among strata and increase the efficiency. When $\rho = 1$ and $\tau = 0$ ranking is accurate regardless of value of K. In this case, we do not observe an increase in the efficiency. In an infinite population setting, this relationship is also reported in Frey (2012) and Ozturk (2013).

6. STATISTICAL INFERENCE

In the previous section, we established that the level-2 design performs better than all the other three designs. Thus, in this section, we develop statistical inference for the population

Table 6. Biases and relative efficiencies of the level-2 design Horvitz–Thompson ($\hat{\mu}_2(K)$) and the proposed estimator ($\mu_2^*(K)$) based on K auxiliary variables. The efficiencies are given in terms of the ratio of the MSEs of SRS and level-2 design estimators, i.e., $R(K) = MSE(\hat{\mu})/MSE(\hat{\mu}_2(K))$, $R^*(K) = MSE(\hat{\mu})/MSE(\mu_2^*(K))$. Data sets are generated from discretized Cauchy distributions, Cauchy(0, 1) + μ.

			Biases				Efficiencies			
τ	μ	ρ	$\hat{\mu}_2(1)$	$\hat{\mu}_2(4)$	$\hat{\mu}_2^*(1)$	$\hat{\mu}_2^*(4)$	$R(1)$	$R(4)$	$R^*(1)$	$R^*(4)$
0.00	0	0.50	−0.062	−0.082	−0.064	−0.079	1.045	1.149	1.084	1.141
0.00	0	0.75	−0.024	−0.056	−0.027	−0.058	1.014	1.099	1.034	1.071
0.00	0	1.00	−0.028	0.100	−0.028	0.100	1.090	1.031	1.090	1.031
0.00	100	0.50	0.235	−0.115	−0.058	−0.027	0.894	1.029	1.036	1.119
0.00	100	0.75	0.212	−0.088	−0.006	−0.001	0.985	1.069	1.088	1.084
0.00	100	1.00	−0.019	−0.021	−0.019	−0.021	1.068	1.079	1.068	1.079
0.00	1000	0.50	2.968	−1.111	−0.002	0.008	0.071	0.118	1.027	1.133
0.00	1000	0.75	2.174	−0.690	0.014	0.038	0.108	0.174	0.980	1.116
0.00	1000	1.00	−0.001	−0.046	−0.001	−0.046	1.121	1.125	1.121	1.125
1.00	0	0.50	−0.006	−0.046	−0.010	−0.043	0.995	1.143	1.017	1.145
1.00	0	0.75	−0.006	0.037	0.004	0.035	1.015	1.122	1.056	1.111
1.00	0	1.00	−0.009	0.003	−0.014	0.004	1.143	1.067	1.173	1.042
1.00	100	0.50	0.113	−0.114	−0.111	0.008	0.940	1.081	1.074	1.142
1.00	100	0.75	0.204	−0.124	0.016	−0.009	0.945	1.059	1.053	1.134
1.00	100	1.00	0.152	−0.025	−0.122	0.012	0.894	1.053	0.993	1.093
1.00	1000	0.50	2.543	−1.276	−0.027	0.034	0.080	0.131	1.067	1.111
1.00	1000	0.75	2.236	−1.200	−0.003	0.027	0.104	0.161	1.046	1.143
1.00	1000	1.00	2.583	−0.680	0.127	−0.009	0.108	0.154	1.070	1.130
1.50	0	0.50	−0.046	−0.013	−0.036	−0.018	1.080	1.177	1.108	1.184
1.50	0	0.75	−0.031	−0.013	−0.023	−0.017	1.149	1.145	1.180	1.132
1.50	0	1.00	0.065	0.045	0.060	0.050	1.080	1.141	1.097	1.106
1.50	100	0.50	0.171	−0.142	−0.071	−0.027	0.963	1.076	1.084	1.194
1.50	100	0.75	0.277	−0.104	0.064	0.003	0.924	1.059	1.031	1.126
1.50	100	1.00	0.350	−0.099	0.029	−0.023	0.934	1.081	1.128	1.139
1.50	1000	0.50	2.675	−1.101	0.026	−0.020	0.081	0.129	1.016	1.159
1.50	1000	0.75	2.167	−1.219	−0.063	0.052	0.104	0.162	1.041	1.151
1.50	1000	1.00	3.277	−0.633	0.052	0.029	0.077	0.115	1.097	0.995

mean and total based on the level-2 design. The estimated inclusion probabilities under multi-ranker model in the level-2 design, unlike perfect ranking RSS and simple random sampling designs, where inclusion probabilities are constant, are random variables. Thus, the distribution of the estimated inclusion probabilities depends on the underlying ranking structure. The construction of their distributions requires strong modeling assumptions on the ranking mechanism. In order to avoid strong ranking assumptions, we use bootstrap distribution of the estimator $\mu_2^*(K)$ to develop inference for the population mean (and total).

In an infinite population setting, the bootstrap distribution of a statistic can be constructed from a plug-in method. Let the parameter of interest θ be a functional of an unknown distribution $F, \theta = T(F)$. The estimate of θ is then obtained from $\hat{\theta} = T(\hat{F})$, where $\hat{F}$ is the empirical cdf based on a random sample from F. In a finite population setting, the distribution F is replaced with the population $\mathcal{P} = \{x_1, \ldots, x_N\}$ of size N, where x_i is the numerical value of the unit u_i in the population.

Assume that our interest is on the parameter $\theta = T(\mathcal{P})$. We use the natural extension of infinite population bootstrap plug-in rule to estimate the parameter θ, $\hat{\theta} = T(\hat{\mathcal{P}})$, where $\hat{\mathcal{P}}$ is the empirical population constructed based on multi-ranker PROS sample design. Since a ranked set sample consists of H judgment strata, the construction of empirical population $\hat{\mathcal{P}}$ depends on the structure of these judgment strata in the sample. Let $p_h = n_h/n$, $h = 1, \ldots, H$, be the fraction of the observations in judgment class h. The empirical population $\hat{\mathcal{P}}$ then must have $\hat{\mathcal{P}}_{[1]}, \ldots, \hat{\mathcal{P}}_{[H]}$ stratified population with population sizes $N_1, \ldots, N_H$, where $N_h = p_h N$. Let

$$\boldsymbol{V}_{[h]j}^{\top} = \left(\pi_j^{(2)}(K), X_{[h]j}, \boldsymbol{w}^{*\top}_{h,j}\right), \quad j = 1, \ldots, n_h,$$

be the multi-ranker PROS sample observation in judgment class h based on level-2 design. It is clear that $V_{[h]j}$ is a random vector that depends on the underlying ranking mechanism of the multi-ranker model. Let $\boldsymbol{V}_{[h]}$ be an n_h by $H+2$ matrix containing $\boldsymbol{V}_{[h]j}$ in its jth row, $j = 1, \ldots, n_h$. With this notation, our PROS sample combined with the estimated inclusion probabilities of the measured units can be written as $\mathcal{V} = \{\boldsymbol{V}_{[1]}, \ldots, \boldsymbol{V}_{[H]}\}$. In order to construct the empirical population $\hat{\mathcal{P}}_{[h]}$ based on sample $\boldsymbol{V}_{[h]}$ (or $\boldsymbol{V}_{[h]j}$, $j = 1, \ldots, n_h$), we first determine M_h, integer part of N_h/n_h, and $m_h = N_h - M_h$. The empirical population $\hat{\mathcal{P}}_{[h]}$ can be constructed by repeating matrix $\boldsymbol{V}_{[h]}$ M_h times and selecting m_h rows at random from $\boldsymbol{V}_{[h]}$ to create an N_h by $H+2$ matrix

$$\hat{\mathcal{P}}_{[h]} = \begin{bmatrix} \boldsymbol{V}_{[h]} \\ \vdots \\ \boldsymbol{V}_{[h]} \\ V_{[h]t_1} \\ \vdots \\ V_{[h]t_{m_h}} \end{bmatrix}, \quad h = 1, \ldots, H.$$

The empirical bootstrap population then becomes $\hat{\mathcal{P}} = \{\hat{\mathcal{P}}_{[1]}, \ldots, \hat{\mathcal{P}}_{[H]}\}$. The empirical bootstrap populations for other sampling designs can be constructed in a similar fashion.

We now discuss the selection of bootstrap sample from the empirical population $\hat{\mathcal{P}}$. Let $\mathcal{V}^*_{[h]}$ be a sample of size n_h, selected at random without replacement from the hth empirical judgment stratum population $\hat{\mathcal{P}}_{[h]}$. In other words, we select n_h rows from the population matrix $\hat{\mathcal{P}}_{[h]}$. We then call $\mathcal{V}^* = \{\mathcal{V}^*_{[1]}, \ldots, \mathcal{V}^*_{[H]}\}$ a resample from $\hat{\mathcal{P}}$. Let $\hat{\mathcal{P}}_{(1)}, \ldots, \hat{\mathcal{P}}_{(B)}$ be B empirical populations constructed independently from PROS sample $\mathcal{V} = \{\boldsymbol{V}_{[1]}, \ldots, \boldsymbol{V}_{[H]}\}$ as explained above. For each $b = 1, \ldots, B$, let $\mathcal{V}^*_{(b,1)}, \ldots, \mathcal{V}^*_{(b,G)}$, be a collection of G independently selected resamples from $\hat{\mathcal{P}}_{(b)}$. We apply our estimator to each one of these bootstrap resamples to obtain

$$T_{(b-1)G+g}(2) = \mu_2^*\left(K, \mathcal{V}^*_{(b,g)}\right), \quad b = 1, \ldots, B, g = 1, \ldots, G,$$

where $\mu_2^*(K, \mathcal{V}^*_{(b,g)})$ is the estimator in Equation (4.1) applied to resampled data $\mathcal{V}^*_{(b,g)}$. The bootstrap variance estimates of $\mu_2^*(K)$ is then obtained by

$$BV(\mu_2^*) = \frac{1}{BG}\sum_{b=1}^{B}\sum_{g=1}^{G}\big(T_{(b-1)G+g}(2) - \bar{T}(2)\big)^2, \tag{6.1}$$

where $\bar{T}(2)$ is the average of $T_{(b-1)G+g}(2)$, $b = 1, \ldots, B$, $g = 1, \ldots, G$.

A $(1-\gamma)100$ % bootstrap percentile (BP) confidence interval is constructed by $(T^{\gamma/2}, T^{1-\gamma/2})$, where T^a is the ath quantile of T satisfying $a = P(T \le T^a|\mathcal{V})$ for $0 < a < 1$. The quantiles $T^{\gamma/2}$ and $T^{1-\gamma/2}$ are obtained from bootstrap distribution of T.

In order to evaluate properties of the bootstrap variance estimate of the estimator $\mu_2^*(K)$ and the bootstrap percentile confidence interval of the population mean, we performed another simulation study. We used the same simulation parameters as the ones that are used in Section 6, i.e., $H = 3$; $C = 10$; $K = 4$; $\tau = 0, 1, 1.5$; $\rho = 0.5, 0.75, 1.00$ and $\mu = 0, 100, 1000$. Table 7 presents the variance estimates of the estimator $\mu_2^*(K)$ and coverage probabilities of the bootstrap percentile confidence intervals. The heading SV represents the simulated variance estimate of $\mu_2^*(K)$

$$SV = \frac{\sum_{j=1}^{2000}(\mu^*_{2,j}(K) - \bar{\mu}_2^*(K))^2}{1999}, \quad \bar{\mu}_2^*(K) = \frac{\sum_{j=1}^{2000}\mu^*_{2,j}(K)}{2000},$$

where $\mu^*_{2,j}(K)$ is the estimator in Equation (4.1) in the jth iteration of the simulation based on level-2 design. The heading BV represents the bootstrap variance estimate computed from Equation (6.1) with $G = 50$, and $B = 50$. The heading CV represents the coverage probabilities of the bootstrap percentile confidence interval of the population mean.

It is clear from Table 7 that the variance estimates SV and BV are reasonably close to each other for the selected simulation parameters and distributions. The simulation results indicate that bootstrap distribution of the proposed estimator provides a mechanism to estimate the standard error of the estimator. Similar results also hold for the coverage probabilities of the percentile confidence interval. The estimated coverage probabilities are reasonably close to nominal level 95 % for discrete normal and exponential populations. The coverage probabilities for discretized Cauchy distribution appears to be slightly lower than the nominal level. This could be explained from the fact that the super population Cauchy does not have a finite moments and hence violates the regularity conditions of the finite sample bootstrap distribution for finite sample stratified population in Booth, Butler, and Hall (1994).

7. APPLICATION

In this section, we apply the proposed estimator to USDA 1992 Ohio corn data. The ranking information (strength-of-weight matrices in PROS design in Section 2) from these auxiliary variables can be combined by using weights $\alpha_1 = 0.252$, $\alpha_2 = 0.209$, $\alpha_3 = 0.263$ and $\alpha_4 = 0.276$. By treating these 202 farms as a finite population, we performed another simulation study to investigate the biases and efficiencies of $\hat{\mu}_2(1)$, $\hat{\mu}_2(4)$ and $\mu_2^*(1)$,

Table 7. Variance estimate of the estimator $\mu_2^*(4)$ based on simulation (SV) and bootstrap distribution (BV), and coverage probabilities (CP) of the 95-percent bootstrap percentile confidence intervals of population mean for normal, exponential and Cauchy distributions.

			$N(0,1)+\mu$			$\mathrm{Exp}(1)+\mu$			$\mathrm{Cauchy}(0,1)+\mu$		
τ	μ	ρ	SV	BV	CP	SV	BV	CP	SV	BV	CP
0.00	0	0.500	0.020	0.020	0.935	0.019	0.019	0.925	5.544	4.869	0.930
0.00	0	0.750	0.012	0.014	0.957	0.015	0.016	0.927	5.912	5.279	0.921
0.00	0	1.000	0.013	0.014	0.943	0.016	0.016	0.919	6.132	5.548	0.910
0.00	100	0.500	0.019	0.021	0.951	0.019	0.020	0.932	5.658	5.060	0.917
0.00	100	0.750	0.013	0.015	0.949	0.016	0.016	0.934	5.845	5.096	0.907
0.00	100	1.000	0.013	0.014	0.947	0.016	0.016	0.906	5.869	5.451	0.900
0.00	1000	0.500	0.021	0.021	0.937	0.019	0.019	0.930	5.591	5.086	0.921
0.00	1000	0.750	0.014	0.015	0.951	0.014	0.016	0.939	5.671	5.367	0.924
0.00	1000	1.000	0.013	0.014	0.948	0.015	0.016	0.930	5.626	5.170	0.907
1.00	0	0.500	0.020	0.020	0.939	0.019	0.019	0.921	5.529	4.849	0.920
1.00	0	0.750	0.013	0.014	0.952	0.013	0.015	0.936	5.701	5.243	0.919
1.00	0	1.000	0.012	0.013	0.946	0.015	0.014	0.914	6.079	5.298	0.912
1.00	100	0.500	0.018	0.019	0.949	0.019	0.020	0.933	0.019	0.020	0.933
1.00	100	0.750	0.012	0.014	0.959	0.013	0.015	0.942	5.586	4.999	0.923
1.00	100	1.000	0.012	0.013	0.953	0.014	0.015	0.929	5.796	5.224	0.909
1.00	1000	0.500	0.019	0.019	0.941	0.019	0.019	0.924	5.697	5.055	0.925
1.00	1000	0.750	0.013	0.015	0.957	0.014	0.016	0.939	5.542	4.997	0.920
1.00	1000	1.000	0.012	0.013	0.945	0.014	0.015	0.928	5.606	5.280	0.908
1.50	0	0.500	0.018	0.019	0.954	0.019	0.019	0.932	5.349	4.898	0.937
1.50	0	0.750	0.013	0.015	0.947	0.016	0.016	0.931	5.593	4.952	0.919
1.50	0	1.000	0.012	0.014	0.958	0.015	0.015	0.926	5.723	5.250	0.918
1.50	100	0.500	0.018	0.019	0.942	0.019	0.020	0.926	5.302	4.522	0.915
1.50	100	0.750	0.012	0.015	0.964	0.013	0.015	0.942	5.626	4.858	0.917
1.50	100	1.000	0.012	0.014	0.956	0.014	0.016	0.931	5.560	5.201	0.917
1.50	1000	0.500	0.019	0.019	0.936	0.019	0.020	0.926	5.463	4.860	0.937
1.50	1000	0.750	0.012	0.014	0.954	0.013	0.016	0.949	5.499	4.982	0.935
1.50	1000	1.000	0.012	0.014	0.951	0.013	0.015	0.939	6.364	5.589	0.916

$\mu_2^*(4)$. In the simulation, we used cycle size $C = 10, 14, 20, 30,$ and set size $H = 2, 3, 4$. Table 8 presents the biases and relative efficiencies of the HT and the proposed estimators. We note that high correlation between the response and auxiliary variables indicate that ranking process is highly accurate. Thus, as expected, both HT and proposed estimators have very little bias. On the other hand, the new estimator has higher efficiency than the HT estimator.

For the same simulation parameters, Table 9 presents the standard error estimates of the estimators $\hat{\mu}_2(1)$ and $\mu_2^*(4)$, and the coverage probability of a 95 % bootstrap percentile confidence intervals of population mean based on the proposed estimator $\mu_2^*(4)$. It is clear from Table 9 that simulated (S.se) and bootstrap (B.se) estimated standard errors of $\mu_2^*(4)$ are almost identical. The coverage probabilities are also relatively close to the nominal level 95 %.

Table 8. Biases and relative efficiencies of $\hat{\mu}_2(K)$ and $\mu_2^*(K)$ based on 1992 USDA Ohio Survey. Relative efficiencies are computed in terms of MSE of SRS and level-2 sampling design estimators, i.e., $R(K) = MSE(\hat{\mu})/MSE(\hat{\mu}_2(K))$ and $R^*(K) = MSE(\hat{\mu})/MSE(\mu_2^*(K))$ for $K = 1, 4$.

C	H	$B(\hat{\mu}_2(1))$	$B(\hat{\mu}_2(4))$	$B(\mu_2^*(1))$	$B(\mu_2^*(4))$	$R(1)$	$R(4)$	$R^*(1)$	$R^*(4)$
10.00	2	34.349	44.831	16.905	4.858	1.277	1.332	1.291	1.389
14.00	2	−18.944	−14.505	−34.663	−33.607	1.373	1.359	1.382	1.412
20.00	2	−19.742	−40.123	−29.073	−56.726	1.380	1.350	1.384	1.414
30.00	2	−15.042	−8.777	−17.148	−16.149	1.464	1.491	1.467	1.582
10.00	3	13.560	40.813	−1.973	−3.324	1.700	1.719	1.710	1.851
14.00	3	−21.535	−22.636	−36.299	−53.166	1.973	1.799	1.990	1.914
20.00	3	43.980	16.423	34.606	13.300	2.028	2.033	2.037	2.244
10.00	4	76.962	22.470	52.440	4.362	2.164	2.202	2.191	2.392

Table 9. Simulated standard error (S.se) and bootstrap standard error (B.se) estimate of $\mu_2^*(4)$, and coverage probabilities (CP) of the 95 % bootstrap percentile confidence interval of the population mean of USDA 1992 Ohio survey data.

C	H	S.se	B.se	CP
10.00	2	3354.040	3267.418	0.898
14.00	2	2748.577	2733.363	0.911
20.00	2	2216.628	2226.492	0.925
30.00	2	1602.549	1745.553	0.948
10.00	3	2306.067	2360.690	0.911
14.00	3	1847.537	1948.186	0.934
20.00	3	1345.393	1578.769	0.963
10.00	4	1705.066	1837.451	0.945

8. CONCLUSION

We have developed three sampling designs that combines the ranking information from different sources to construct a ranked set sample from a finite population. The samples are constructed with three levels of without-replacement policies, level-0, level-1, and level-2. A level-0 design is constructed with replacement on all the units in a set. A level-1 design is constructed with without-replacement policy on the measured unit in the set. A level-2 design is constructed by using without-replacement policy on all units regardless the measurement status. These sampling designs provide a mechanism to combine the ranking information from different sources to obtain a strength-of-agreement weights. These weights provides a model to construct recursive algorithms to compute the estimate of the first and second order inclusion probabilities. The inclusion probabilities along with the strength of agreement weights are used to construct estimators for the population mean and total. We showed that the new estimator based on level-2 design outperforms Horvitz–Thompson and simple random sample estimators. We construct finite sample bootstrap inference for the population mean. The new sampling design and estimator are applied to 1992 Ohio corn data to show that it provides a viable sampling design and estimator in survey sampling studies.

ACKNOWLEDGEMENT

The author thanks Cheryl Turner at USDA NASS Ohio Field Office for arranging for us to use the Ohio Estimation Data in this study.

[Received April 2013. Accepted October 2013. Published Online November 2013.]

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.

Arslan, G., and Ozturk, O. (2013), "Parametric Inference Based on Partially Ordered Ranked Set Samples," *Journal of Indian Statistical Association*, 51, 1–24.

Booth, J. G., Butler, W. R., and Hall, P. (1994), "Bootstrap Methods for Finite Populations," *Journal of the American Statistical Association*, 89, 1282–1289.

Dell, T. R., and Clutter, J. L. (1972), "Ranked Set Sampling Theory With Order Statistics Background," *Biometrics*, 28, 545–555.

Deshpande, J. V., Frey, J., and Ozturk, O. (2006), "Nonparametric Ranked-Set Sampling Confidence Intervals for a Finite Population," *Environmental and Ecological Statistics*, 13, 25–40.

Fligner, M., and MacEachern, S. N. (2006), "Nonparametric Two-Sample Methods for Ranked Set Sample Data," *Journal of the American Statistical Association*, 475, 1107–1118.

Frey, J. (2011), "Recursive Computation of Inclusion Probabilities in Ranked Set Sampling," *Journal of Statistical Planning and Inference*, 141, 3632–3639.

——— (2012), "Nonparametric Mean Estimation Using Partially Ordered Sets," *Environmental and Ecological Statistics*, 19, 309–326.

Gao, J., and Ozturk, O. (2012), "Two Sample Distribution-Free Inference Based on Partially Rank-Ordered Set Sample," *Statistics & Probability Letters*, 82, 876–884.

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 the American 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 Statistics*, 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.

McIntyre, G. A. (1952), "A Method for Unbiased Selective Sampling Using Ranked-Sets," *Australian Journal of Agricultural Research*, 3, 385–390.

——— (2005), "A Method of Unbiased Selective Sampling Using Ranked-Sets," *American Statistician*, 59, 230–232.

Ozdemir, Y. A., and Gokpinar, F. (2007), "A Generalized Formula for Inclusion Probabilities in Ranked Set Sampling," *Hacettepe Journal of Mathematics and Statistics*, 36, 89–99.

——— (2008), “A New Formula for Inclusion Probabilities in Median Ranked Set Sampling,” *Communications in Statistics. Theory and Methods*, 37, 2022–2033.

Ozturk, O. (2011), “Sampling From Partially Rank-Ordered Sets,” *Environmental and Ecological Statistics*, 18, 757–779.

——— (2012a), “Quantile Inference Based on Partially Rank-Ordered Set Samples,” *Journal of Statistical Planning and Inference*, 142, 2116–2127.

——— (2012b), “Combining Ranking Information in Judgment Post Stratified and Ranked Set Sampling Designs,” *Environmental and Ecological Statistics*, 19, 73–93.

——— (2013), “Combining Multi-observer Information in Partially Rank-Ordered Judgment Post Stratified and Ranked Set Samples,” *Canadian Journal of Statistics*, 41, 304–324.

Ozturk, O., and Jafari Jozani, M. (2014), “Inclusion Probabilities in Partially Rank Ordered Set Sampling,” *Computational Statistics & Data Analysis*, 69, 122–132.

Patil, G. P., Sinha, A. K., and Taillie, C. (1995), “Finite Population Correction for Ranked Set Sampling,” *Annals of the Institute of Statistical Mathematics*, 47, 621–636.