Ann Inst Stat Math (2013) 65:551–569
DOI 10.1007/s10463-012-0386-5

Variance estimation using judgment post-stratification

Jesse Frey · Timothy G. Feeman

Received: 19 January 2012 / Revised: 24 August 2012 / Published online: 15 November 2012

Abstract We consider the problem of estimating the variance of a population using judgment post-stratification. By conditioning on the observed vector of ordered in-stratum sample sizes, we develop a conditionally unbiased nonparametric estimator that outperforms the sample variance except when the rankings are very poor. This estimator also outperforms the standard unbiased nonparametric variance estimator from unbalanced ranked-set sampling.

Keywords Conditioning · Imperfect rankings · Judgment ranking · Ranked-set sampling

1 Introduction

Judgment post-stratification (JPS), proposed by MacEachern et al. (2004), is a data collection scheme in which a simple random sample is supplemented with judgment ranking information. It differs from standard post-stratification (see Lohr 1999) in that the strata are based on ranking information rather than covariate information. To draw a JPS sample using set size m and sample size N, one first draws a simple random sample of size N and makes a measurement on each of the N units. One then selects, for each unit in the sample, an additional $m - 1$ independent units, yielding a set of size m. The units in this set are ranked from smallest to largest by judgment, and the rank of the unit from the simple random sample is recorded. This judgment rank need not match the true in-set rank of the unit. The full JPS data set then consists of the measured values for the N units, together with the rank of each unit within its set of size m.

J. Frey (✉) · T. G. Feeman
Department of Mathematics and Statistics, Villanova University, Villanova, PA 19085, USA
e-mail: jesse.frey@villanova.edu

When the collection of measured units includes at least one unit with each rank 1 to m, the standard JPS mean estimator $\bar{X}_{\text{JPS}}$ is simply the average of the post-stratum sample means $\bar{X}_{[1]}, \ldots, \bar{X}_{[m]}$. However, it is possible that some of the post-strata are empty. Indeed, if $(n_1, \ldots, n_m)$ is the vector of post-stratum sample sizes, then $(n_1, \ldots, n_m) \sim \text{Multinomial}(N, (1/m, \ldots, 1/m))$. When there are empty post-strata, the standard JPS mean estimator $\bar{X}_{\text{JPS}}$ is the average of the sample means for the nonempty post-strata. It follows from symmetry considerations that $\bar{X}_{\text{JPS}}$ is unbiased for estimating the population mean μ, and MacEachern et al. (2004) showed that no matter how the judgment rankings are done, $\bar{X}_{\text{JPS}}$ is at least as efficient asymptotically as the simple random sampling (SRS) mean estimator $\bar{X}_{\text{SRS}}$.

JPS uses the same sort of judgment ranking information that is used in ranked-set sampling (RSS), a data collection scheme proposed by McIntyre (1952, 2005). JPS tends to be less efficient than balanced RSS, but it offers advantages in terms of flexibility. One advantage is that users of JPS retain the option of using SRS-based analysis methods if needed, and a second advantage is that rankers may be permitted to declare ties. MacEachern et al. (2004) pointed out these advantages, and they also proposed several methods for estimating the population mean μ in the presence of ties. Like RSS, JPS is preferable to SRS in settings where precise measurements are costly, but reasonably accurate judgment rankings are easily available. Thus, JPS has potential applications in areas such as environmental monitoring (Kvam 2003), forestry (Halls and Dell 1966), medicine (Chen et al. 2005), and entomology (Howard et al. 1982).

Recent work has looked at JPS in several different contexts. Wang et al. (2006) proposed JPS mean estimators for the case where multiple sets of judgment rankings are available, and Wang et al. (2008) proposed an isotonic JPS mean estimator for the case where the post-stratum means are believed to be stochastically ordered. Du and MacEachern (2008) applied JPS in the context of designed experiments, and Frey and Ozturk (2011) obtained alternate JPS mean estimators by deriving relationships between the distributions for the m post-strata. More recently, Frey and Feeman (2012) showed that the standard JPS mean estimator $\bar{X}_{\text{JPS}}$ is inadmissible under squared error loss. They also developed an improved JPS mean estimator by conditioning on the ordered in-stratum sample sizes $s_1 \geq \cdots \geq s_k > 0$, where $k \leq m$ is the number of nonempty post-strata.

In this paper, we consider estimating the population variance σ^2 using JPS. At least two variance estimators can be obtained simply by exploiting the connections between JPS, SRS, and RSS. First, if we ignore the ranking information, then a JPS sample is a simple random sample. Thus, the sample variance s^2 is an unbiased estimator of σ^2. Second, if we condition on the sample sizes $(n_1, \ldots, n_m)$, then we may think of the JPS sample as a balanced or unbalanced ranked-set sample. Thus, variance estimators that work with unbalanced RSS can also be applied with JPS. However, we find in this paper that an alternate approach gives better results.

Following the approach of Frey and Feeman (2012), we condition on the ordered in-stratum sample sizes $s_1 \geq \cdots \geq s_k > 0$. Focusing on a natural class of conditional variance estimators, we obtain both the conditional minimum mean squared error (MSE) and the conditional minimum variance unbiased estimators. Both of these estimators depend on the particular parent distribution and on the type of rankings, but the conditional minimum variance unbiased estimator is relatively robust to these choices. Thus, using the conditionally minimum variance unbiased estimator for the case where the parent distribution is uniform and the rankings are perfect, we obtain an estimator $\hat{\sigma}_C^2$ that performs well across a variety of choices for the parent distribution and the ranking mechanism. In particular, it works well with perfect rankings, and it is only slightly less efficient than the sample variance when the rankings are random.

In Sect. 2, we discuss variance estimation using RSS. In Sect. 3, we derive the conditional minimum MSE estimator and the conditional minimum variance unbiased estimator for JPS. In Sect. 4, we show that the conditional minimum variance unbiased estimator is robust to the choice of the parent distribution and the type of rankings. This leads us to recommend use of $\hat{\sigma}_C^2$. In Sect. 5, we compare the performance of $\hat{\sigma}_C^2$ to that of other potential estimators using a model for imperfect rankings, and in Sect. 6, we compare the performance of the estimators using real data. In Sect. 7, we give conclusions.

2 Variance estimation using ranked-set sampling

RSS differs from JPS in that the ranking step comes before one chooses the units for measurement. To draw a balanced ranked-set sample using set size m and n cycles, one first selects $N = nm$ independent simple random samples (sets) of size m. One then ranks the units in each set from smallest to largest. As in JPS, this judgment ranking need not be perfectly accurate. One then selects a single unit from each set for measurement. Specifically, one selects the unit with rank 1 from each of the first n sets, the unit with rank 2 from each of the next n sets, and so on. This yields a sample of N independent measured values, with n from each possible in-set rank.

To do unbalanced RSS, we simply relax the requirement that each in-set rank be equally represented. Instead, we specify a set size m and a vector $(n_1, \ldots, n_m)$ so that the sample includes n_1 units with rank 1, n_2 units with rank 2, and so on. Whether the sample is balanced or unbalanced, we must employ blinding and appropriate randomization to ensure that the judgment rankings are not affected by knowledge of which ranked unit is to be selected from each set.

For balanced RSS with set size m and n cycles, Stokes (1980) proposed estimating σ^2 using

$$\tilde{\sigma}^2 = \frac{1}{nm-1} \sum_{i=1}^{m} \sum_{r=1}^{n} \left(X_{[i]r} - \hat{\mu} \right)^2 ,$$

where $\hat{\mu} = \sum_i \sum_r X_{[i]r}/(nm)$ and $X_{[i]r}$ is the rth unit with in-set rank i. This estimator is asymptotically unbiased, but it does not perform well for small samples. In particular, it tends to overestimate the true variance.

As an alternative to $\tilde{\sigma}^2$, MacEachern et al. (2002) proposed the estimator

$$\hat{\sigma}_M^2 = \frac{\sum_{i \neq j} \sum_r \sum_s \left(X_{[i]r} - X_{[j]s} \right)^2}{2n^2 m^2} + \frac{\sum_i \sum_r \sum_s \left(X_{[i]r} - X_{[i]s} \right)^2}{2n(n-1)m^2} .$$

This estimator is unbiased, and it tends to be more efficient than $\tilde{\sigma}^2$ when the rankings are good. However, it applies only for balanced RSS with at least two cycles. For unbalanced RSS with in-stratum sample sizes $(n_1, \ldots, n_m)$, MacEachern et al. (2002) proposed the unbiased estimator

$$\hat{\sigma}_M^2 = \frac{1}{2m^2}\sum_{i \neq j} \frac{1}{n_i n_j} \sum_r \sum_s \left(X_{[i]r} - X_{[j]s}\right)^2 + \frac{1}{2m^2}\sum_i \frac{1}{n_i(n_i-1)} \sum_r \sum_s \left(X_{[i]r} - X_{[i]s}\right)^2, \tag{1}$$

which can be applied as long as each of the in-stratum sample sizes is at least two.

Working independently of MacEachern et al. (2002), Perron and Sinha (2004) investigated the class of variance estimators of the form $\sum_i \sum_j \sum_r \sum_s \gamma_{i,j,r,s} X_{[i]r} X_{[j]s}$, where the values $\{\gamma_{i,j,r,s}\}$ are constants that satisfy $\gamma_{i,j,r,s} = \gamma_{j,i,s,r}$. They showed that the estimator (1) has minimum variance among unbiased estimators in this class.

One additional contribution to variance estimation using RSS was made by Yu et al. (1999), who developed RSS-based variance estimators appropriate for use with normal data under perfect rankings. Since our interest here is in nonparametric estimators that perform well regardless of how the rankings are done, we do not consider these estimators further.

3 Optimal variance estimators from a certain class

Consider estimating the population variance σ^2 in the standard JPS set-up. Let $s_1 \geq \cdots \geq s_k > 0$ be the ordered sample sizes $n_1, \ldots, n_m$, where $k \leq m$ is the number of nonempty post-strata. Then, for $i = 1, \ldots, k$, let $Y_{i1}, \ldots, Y_{is_i}$ be the elements of the post-stratum with sample size s_i. Assume that $k \geq 2$ and $s_1 \geq 2$. Motivated by the estimators from Sect. 2, we consider the class of estimators of the form

$$\hat{\sigma}^2 = \sum_{i=1}^{k} w_{ii} \cdot \sum_{r<s} (Y_{ir} - Y_{is})^2 + \sum_{i<j} w_{ij} \cdot \sum_r \sum_s \left(Y_{ir} - Y_{js}\right)^2, \tag{2}$$

where the values $\{w_{ij}\}$ are scalar weights that depend only upon the post-stratum sizes $\{s_i\}$ and not on which judgment stratum is associated with each size. This class includes the sample variance s^2, and provided that each n_i is at least two, the unbalanced RSS estimator (1).

We want estimators that have small MSE. Thus, it is natural to consider the estimator that minimizes the conditional MSE for estimators in the class (2) and the estimator that minimizes the conditional variance among conditionally unbiased estimators in the class (2). We show in what follows that each of these estimators can be obtained using standard minimization techniques and appropriate expressions for the moments of the random variables $\{Y_{ir}\}$. The theoretical results given below prove useful. To ensure that all needed moments of $\hat{\sigma}^2$ exist, we assume that the fourth moment of the parent distribution exists. Proofs of Results 2, 3, and 5 are given in the Appendix.

Result 1 (Dell and Clutter (1972)) Let X be a single draw from the parent distribution, and let $X_{[i]}$ be a single draw from the ith judgment post-stratum. If b is a real number such that $E[X^b]$ exists, then $\frac{1}{m}\sum_{i=1}^{m} E[X_{[i]}^b] = E[X^b]$.

Result 2 For any i and r, $E[Y_{ir}|s_1,\ldots,s_k] = \mu$ and $V(Y_{ir}|s_1,\ldots,s_k) = \sigma^2$.

Result 3 For any i and any $r \neq s$,

$$E\left[(Y_{ir} - Y_{is})^2|s_1,\ldots,s_k\right] = 2\sigma^2 - \frac{2}{m}\sum_{l=1}^{m}(\mu_{[l]} - \mu)^2,$$

and for any $i \neq j$ and any r and s,

$$E\left[(Y_{ir} - Y_{js})^2|s_1,\ldots,s_k\right] = 2\sigma^2 + \frac{2}{m(m-1)}\sum_{l=1}^{m}(\mu_{[l]} - \mu)^2.$$

Applying Result 3 to Eq. (2) gives us the following result.

Result 4 The conditional expected value for an estimator in the class (2) is

$$E\left[\hat{\sigma}^2|s_1,\ldots,s_k\right] = \sum_{i=1}^{k} w_{ii}\binom{s_i}{2}\left(2\sigma^2 - \frac{2}{m}\sum_{l=1}^{m}(\mu_{[l]} - \mu)^2\right) + \sum_{i<j} w_{ij}s_i s_j\left(2\sigma^2 + \frac{2}{m(m-1)}\sum_{l=1}^{m}(\mu_{[l]} - \mu)^2\right)$$

The next result allows us to recognize the conditionally unbiased estimators. The conditionally unbiased estimators are those in which the total weight given to between-stratum comparisons is $m-1$ times the total weight given to within-stratum comparisons.

Result 5 An estimator in the class (2) is conditionally unbiased for σ^2 if and only if

$$\sum_{i=1}^{k} w_{ii}\binom{s_i}{2} = \frac{1}{2m} \quad \text{and} \quad \sum_{i<j} w_{ij}s_i s_j = \frac{m-1}{2m}. \tag{3}$$

Suppose that we wish to find the conditional minimum MSE estimator in a case where the post-stratum distributions are fully known. This requires that we minimize $\text{MSE}(\hat{\sigma}^2|s_1,\ldots,s_k) = E\left[(\hat{\sigma}^2 - \sigma^2)^2|s_1,\ldots,s_k\right]$ over all choices of the weights $\{w_{ij}\}$. Expanding $\text{MSE}(\hat{\sigma}^2|s_1,\ldots,s_k)$ gives

$$\text{MSE}(\hat{\sigma}^2|s_1,\ldots,s_k) = E\left[\hat{\sigma}^4|s_1,\ldots,s_k\right] - 2\sigma^2 E\left[\hat{\sigma}^2|s_1,\ldots,s_k\right] + \sigma^4. \tag{4}$$

Thus, it is sufficient to minimize $E\left[\hat{\sigma}^4|s_1,\ldots,s_k\right]-2\sigma^2E\left[\hat{\sigma}^2|s_1,\ldots,s_k\right]$ over all choices of $\{w_{ij}\}$. Since $E\left[\hat{\sigma}^4|s_1,\ldots,s_k\right]-2\sigma^2E\left[\hat{\sigma}^2\middle|\,s_1,\ldots,s_k\right]$ is a quadratic function of the weights $\{w_{ij}\}$, this minimization can be done using techniques from multivariate calculus. The next result, which follows from squaring the expression for $\hat{\sigma}^2$ given in (2), provides the needed expression for $E\left[\hat{\sigma}^4|s_1,\ldots,s_k\right]$.

Result 6 Let $\hat{\sigma}^2$ be given by (2), and define random variables $\{S_{ij}: 1\le i\le j\le k\}$ by $S_{ii}=\sum_{r<s}(Y_{ir}-Y_{is})^2$ and, for $i<j$, $S_{ij}=\sum_{r=1}^{s_i}\sum_{s=1}^{s_j}(Y_{ir}-Y_{js})^2$. We then have that

$$\begin{aligned}E\left[\hat{\sigma}^4|s_1,\ldots,s_k\right]=&\sum_{i=1}^{k}w_{ii}^2E[S_{ii}^2|s_1,\ldots,s_k]+2\sum_{i\neq j}w_{ii}w_{jj}E[S_{ii}S_{jj}|s_1,\ldots,s_k]\\&+2\sum_{i=1}^{k}\sum_{j<l}w_{ii}w_{jl}E[S_{ii}S_{jl}|s_1,\ldots,s_k]+\sum_{i<j}w_{ij}^2E[S_{ij}^2|s_1,\ldots,s_k]\\&+2\sum_{i<j}\sum_{\substack{l<t\\(l,t)\neq(i,j)}}w_{ij}w_{lt}E[S_{ij}S_{lt}|s_1,\ldots,s_k].\end{aligned}$$

Minimizing the conditional MSE (4) requires setting

$$\frac{\partial}{\partial w_{ii}}\left\{E[\hat{\sigma}^4|s_1,\ldots,s_k]-2\sigma^2E[\hat{\sigma}^2|s_1,\ldots,s_k]\right\}=0\quad\text{for } i=1,\ldots,k \tag{5}$$

and

$$\frac{\partial}{\partial w_{ij}}\left\{E[\hat{\sigma}^4|s_1,\ldots,s_k]-2\sigma^2E[\hat{\sigma}^2|s_1,\ldots,s_k]\right\}=0\quad\text{for } 1\le i<j\le k. \tag{6}$$

Using Results 4 and 6 to simplify the equations in (5), we get that

$$\begin{aligned}&2\sum_{j=1}^{k}w_{jj}E[S_{ii}S_{jj}|s_1,\ldots,s_k]+2\sum_{j<l}w_{jl}E[S_{ii}S_{jl}|s_1,\ldots,s_k]\\&=2\sigma^2\left\{\binom{s_i}{2}\left(2\sigma^2-\frac{2}{m}\sum_{l=1}^{m}(\mu_{[l]}-\mu)^2\right)\right\}\end{aligned}$$

for $i=1,\ldots,k$. Similarly, simplifying the equations in (6) gives us that

$$\begin{aligned}&2\sum_{l=1}^{k}w_{ll}E[S_{ij}S_{ll}|s_1,\ldots,s_k]+2\sum_{l<t}w_{lt}E[S_{ij}S_{lt}|s_1,\ldots,s_k]\\&=2\sigma^2\left\{s_is_j\left(2\sigma^2+\frac{2}{m(m-1)}\sum_{l=1}^{m}(\mu_{[l]}-\mu)^2\right)\right\}\end{aligned}$$

for i and j satisfying $1 \le i < j \le k$. If we substitute known values for the conditional expectations such as $E[S_{ij} S_{ll} | s_1, \ldots, s_k]$, then we obtain a system of linear equations. Solving this system yields the weights $\{w_{ij}\}$ for an estimator that we call the conditional minimum MSE (CMM) estimator. For fixed ordered sample sizes $s_1 \ge \cdots \ge s_k > 0$, the weights for the CMM estimator depend both on the parent distribution and on the ranking mechanism, but only through the first four moments of the distributions for the m judgment post-strata.

Suppose now that we wish to find the estimator that minimizes the conditional variance among conditionally unbiased estimators in the class (2). The conditional variance of $\hat{\sigma}^2$ is $E[\hat{\sigma}^4 | s_1, \ldots, s_k] - E[\hat{\sigma}^2 | s_1, \ldots, s_k]^2$ in general. However, for conditionally unbiased estimators, this reduces to $E[\hat{\sigma}^4 | s_1, \ldots, s_k] - \sigma^4$. As a result, it is sufficient to minimize $E[\hat{\sigma}^4 | s_1, \ldots, s_k]$ under the constraints (3) from Result 5. This minimization can be done using Lagrange multipliers (see Lange 1999, pp. 181–182) with Lagrangian

$$F(\{w_{ij}\}, \lambda_1, \lambda_2) = E[\hat{\sigma}^4 | s_1, \ldots, s_k] + \lambda_1 \left(\sum_{i=1}^{k} w_{ii} \binom{s_i}{2} - \frac{1}{2m} \right) + \lambda_2 \left(\sum_{i<j} w_{ij} s_i s_j - \frac{m-1}{2m} \right).$$

Setting partial derivatives of the Lagrangian equal to 0, we must have that

$$2 \sum_{j=1}^{k} w_{jj} E[S_{ii} S_{jj} | s_1, \ldots, s_k] + 2 \sum_{j<l} w_{jl} E[S_{ii} S_{jl} | s_1, \ldots, s_k] + \lambda_1 \binom{s_i}{2} = 0$$

for $i = 1, \ldots, k$, that

$$2 \sum_{l=1}^{k} w_{ll} E[S_{ij} S_{ll} | s_1, \ldots, s_k] + 2 \sum_{l<t} w_{lt} E[S_{ij} S_{lt} | s_1, \ldots, s_k] + \lambda_2 s_i s_j = 0$$

for i and j satisfying $1 \le i < j \le k$, and that the two constraints (3) hold. Solving the resulting linear system yields the weights $\{w_{ij}\}$ for the estimator that we call the conditional minimum variance unbiased (CMVU) estimator. Like the CMM estimator developed earlier, the CMVU estimator depends both on the parent distribution and on the ranking mechanism.

Computing the CMM and CMVU estimators requires that we compute expected values like $E[S_{ii}^2 | s_1, \ldots, s_k]$ and $E[S_{ij} S_{lt} | s_1, \ldots, s_k]$. Each such value can be expressed in terms of the moments of the $\{Y_{ir}\}$ and the ordered sample sizes $s_1 \ge \cdots \ge s_k > 0$, and the calculations are routine, but lengthy. An `R` function for computing the weights for the CMM and CMVU estimators when the needed moments are given is available from the authors, and the Appendix provides details for one moment-calculating example, namely that of finding $E[S_{11}^2 | s_1, \ldots, s_k]$.

Our work thus far applies as long as $k \geq 2$ and $s_1 \geq 2$. If $k = 1$ or $s_1 = 1$ so that there is either no information about between-stratum variability or no information about within-stratum variability, then we take the CMM and CMVU estimators to be simply the sample variance s^2. Having $k = 1$ becomes progressively less likely as N increases, and $s_1 = 1$ can occur only if $N \leq m$.

4 Choosing a single estimator

Both the CMM and CMVU estimators depend on the particular parent distribution and the ranking mechanism, but only through the first four moments of the distributions for the m judgment post-strata. In practice, these moments are not known, and it is difficult to estimate them accurately from a small sample. Our solution is to always use the CMVU estimator designed for the case where the parent distribution is uniform and the judgment rankings are perfect. To arrive at this solution, we computed conditional MSEs for a variety of CMM and CMVU estimators both in settings where they are optimal by construction and in other settings. In particular, we considered uniform, normal, exponential, and Gamma(5) parent distributions, and we considered both perfect rankings and random rankings. Table 1, which focuses on the case where $m = 3$ and $(s_1, s_2, s_3) = (11, 10, 9)$, gives some representative results.

Each entry in Table 1 is the conditional MSE for a CMM or CMVU estimator relative to the sample variance s^2. Each conditional relative MSE was computed in such a way that values above 1.0 indicate an advantage for the CMM or CMVU estimator. The top four rows in Table 1 show the performance of CMM estimators designed for the normal, uniform, exponential, and Gamma(5) distributions with perfect rankings, and the bottom four rows show the performance of CMVU estimators designed for the same four settings. To obtain the required conditional MSEs, we first generated, for each combination of a parent distribution and a type of rankings, 1,000,000 random sets of size m. We estimated all needed judgment post-stratum moments using the

Table 1 Conditional MSEs relative to that of s^2 when $m = 3$ and $(s_1, s_2, s_3) = (11, 10, 9)$

Estimator	Perfect rankings				Random rankings			
	Norm.	Unif.	Expo.	Gamma	Norm.	Unif.	Expo.	Gamma
Normal CMM	**1.231**	1.106	0.933	1.101	0.404	0.267	0.558	0.467
Uniform CMM	1.173	**1.272**	0.918	1.055	0.499	0.357	0.633	0.557
Exponential CMM	0.826	0.543	**1.315**	0.998	0.381	0.174	1.055	0.556
Gamma(5) CMM	1.158	0.957	1.122	**1.157**	0.910	0.883	0.926	0.918
Normal CMVU	**1.049**	1.070	1.025	1.038	1.000	0.999	1.000	1.000
Uniform CMVU	1.049	**1.072**	1.026	1.039	0.998	0.998	0.998	0.998
Exponential CMVU	1.049	1.071	**1.026**	1.039	0.999	0.999	0.999	0.999
Gamma(5) CMVU	1.049	1.072	1.026	**1.039**	0.999	0.999	0.999	0.999

Values greater than 1.0 indicate that an estimator has outperformed s^2. Bold values indicate estimators that have minimum MSE within a certain class by construction empirical moments over the 1,000,000 sets, and we computed the conditional MSEs using formula (4), together with Results 4 and 6. Bold values in Table 1 indicate relative conditional MSEs that are optimal by construction.

Focusing on the bold values in the top four rows of Table 1, we see that the CMM estimators perform very well in the perfect rankings settings for which they were designed. However, they perform significantly less well in other perfect rankings settings, and their performance is quite poor when the rankings are random. These results, which suggest that CMM estimators are highly nonrobust to changes in the parent distribution and the type of rankings, prompted us to drop CMM estimators from further consideration.

The bottom four rows of Table 1 show the performance of the CMVU estimators. We see that the four different CMVU estimators have virtually identical conditional MSEs in all of the settings considered. They outperform s^2 under perfect rankings, and they are nearly as efficient as s^2 under random rankings. Because of the similarity between the four estimators, there is little performance-based reason to prefer one over another, and the uniform-based CMVU estimator has the advantage that the required moments can be readily computed. Result 7, which follows from the fact that uniform order statistics have Beta distributions (see David and Nagaraja 2003, Example 2.3), gives the moments that, when combined with our work in Sect. 3, make it possible to compute the weights for $\hat{\sigma}_C^2$. Thus, in the remainder of the paper, we consider only the CMVU estimator for the case of a uniform parent distribution and perfect rankings, and we call this estimator $\hat{\sigma}_C^2$.

Result 7 If the parent distribution is uniform on [0, 1], the rankings are perfect, and the set size is m, then the first four moments for a random draw from the ith judgment post-stratum are $\frac{i}{m+1}$, $\frac{i(i+1)}{(m+1)(m+2)}$, $\frac{i(i+1)(i+2)}{(m+1)(m+2)(m+3)}$, and $\frac{i(i+1)(i+2)(i+3)}{(m+1)(m+2)(m+3)(m+4)}$.

To compare the estimator $\hat{\sigma}_C^2$ to the estimator $\hat{\sigma}_M^2$ of MacEachern et al. (2002), we computed Table 2, which shows the weights associated with $\hat{\sigma}_C^2$ and $\hat{\sigma}_M^2$ for all possible choices of the ordered sample sizes when $N = 10$ and $m = 3$. Since $\hat{\sigma}_M^2$ can be computed only when each n_i is at least two, it applies for only about half of the tabled cases.

Under s^2, all of the weights $\{w_{ij}\}$ are equal to $1/(N(N-1))$, but we see from Table 2 that the weights are usually unequal for $\hat{\sigma}_C^2$. The general pattern we see in Table 2 is that w_{ii} is largest when s_i is small and that w_{ij} is largest when both s_i and s_j are small. The same weighting pattern also holds for $\hat{\sigma}_M^2$, but the weights for $\hat{\sigma}_M^2$ are more variable than those for $\hat{\sigma}_C^2$. Note, for example, that in the $(s_1, s_2, s_3) = (5, 3, 2)$ case, the ratio between the maximum and minimum weights is 10 for $\hat{\sigma}_M^2$, but less than 2.4 for $\hat{\sigma}_C^2$.

The three estimators that we compare further in the next two sections all have unbiasedness properties. The estimator s^2 is unconditionally unbiased since $E[s^2] = \sigma^2$, but it is biased when we condition on a particular choice of the sample size vector $(n_1, \ldots, n_m)$ or the ordered sample sizes $s_1 \geq \cdots \geq s_k > 0$. The estimator $\hat{\sigma}_C^2$ is conditionally unbiased given ordered sample sizes $s_1 \geq \cdots \geq s_k > 0$ such that $s_1 \geq 2$ and $k \geq 2$, but it is slightly biased in the unconditional sense because we set $\hat{\sigma}_C^2 = s^2$ when $k = 1$ or $s_1 = 1$. The estimator $\hat{\sigma}_M^2$ is conditionally unbiased given any choice

Table 2 Weights needed for computing $\hat{\sigma}_C^2$ and $\hat{\sigma}_M^2$ for all possible choices of the ordered sample sizes when $N = 10$ and $m = 3$

(s_1, s_2, s_3)	Estimator	w_{11}	w_{12}	w_{13}	w_{22}	w_{23}	w_{33}
(9, 1, 0)	$\hat{\sigma}_C^2$	0.004630	0.037037	NA	NA	NA	NA
(8, 2, 0)	$\hat{\sigma}_C^2$	0.005069	0.020833	NA	0.024728	NA	NA
(8, 1, 1)	$\hat{\sigma}_C^2$	0.005952	0.017299	0.017299	NA	0.056553	NA
(7, 3, 0)	$\hat{\sigma}_C^2$	0.005766	0.015873	NA	0.015192	NA	NA
(7, 2, 1)	$\hat{\sigma}_C^2$	0.006879	0.011520	0.015197	0.022214	0.032840	NA
(6, 4, 0)	$\hat{\sigma}_C^2$	0.006791	0.013889	NA	0.010799	NA	NA
(6, 3, 1)	$\hat{\sigma}_C^2$	0.008009	0.009427	0.015361	0.015510	0.023827	NA
(6, 2, 2)	$\hat{\sigma}_C^2$	0.008244	0.010551	0.010551	0.021505	0.020027	0.021505
	$\hat{\sigma}_M^2$	0.003704	0.009259	0.009259	0.055556	0.027778	0.055556
(5, 5, 0)	$\hat{\sigma}_C^2$	0.008333	0.013333	NA	0.008333	NA	NA
(5, 4, 1)	$\hat{\sigma}_C^2$	0.009560	0.008643	0.016691	0.011845	0.019255	NA
(5, 3, 2)	$\hat{\sigma}_C^2$	0.009874	0.008921	0.010991	0.015588	0.014934	0.021159
	$\hat{\sigma}_M^2$	0.005556	0.007407	0.011111	0.018519	0.018519	0.055556
(4, 4, 2)	$\hat{\sigma}_C^2$	0.012135	0.008484	0.012349	0.012135	0.012349	0.021052
	$\hat{\sigma}_M^2$	0.009259	0.006944	0.013889	0.009259	0.013889	0.055556
(4, 3, 3)	$\hat{\sigma}_C^2$	0.012195	0.009627	0.009627	0.015583	0.011366	0.015583
	$\hat{\sigma}_M^2$	0.009259	0.009259	0.009259	0.018519	0.012346	0.018519

For the sample variance s^2, each weight is $1/90 \approx 0.011111$

of $(n_1, \ldots, n_m)$ where each n_i is at least two. Thus, it is also conditionally unbiased given ordered sample sizes $s_1 \geq \cdots \geq s_m \geq 2$.

5 Performance comparisons using a model

We compared $\hat{\sigma}_C^2$, $\hat{\sigma}_M^2$, and s^2 in terms of MSE. We considered a variety of choices for the set size m, the total sample size N, the parent distribution, and the type of rankings, and we present representative results here. We considered both perfect and imperfect rankings done according to a perceived size as in the Dell and Clutter (1972) model. Specifically, when a random draw X from the parent distribution has mean μ and standard deviation σ, we let the perceived size T be given by $T = \rho \cdot \left(\frac{X-\mu}{\sigma}\right) + \left(\sqrt{1-\rho^2}\right) \cdot Z$, where Z is a standard normal random variable that is independent of X and ρ is the user-chosen correlation between T and X. For values $X_1, \ldots, X_m$ in the same set, the ranking is done according to $T_1, \ldots, T_m$. If we use $\rho = 1$, then we get perfect rankings, while if we use $\rho = 0$, then we get random rankings.

We considered set sizes 2–5, total sample sizes 10–50, and ρ values of 0, 0.8, and 1.0. These ρ values were chosen to give perfect rankings ($\rho = 1$), random rankings ($\rho = 0$), and rankings just good enough that schemes like RSS and JPS tend to be substantially better than SRS for estimating the population mean μ ($\rho = 0.8$). The parent distributions that we considered were the normal, uniform, exponential, and Gamma(5) distributions. Thus, we considered both symmetric and skewed distributions. We also considered both settings that match the setting in which $\hat{\sigma}_C^2$ is optimal by construction (uniform with perfect rankings) and settings where $\hat{\sigma}_C^2$ is not optimal by construction.

To compare $\hat{\sigma}_C^2$ to s^2 under a particular choice of $m,\ N,\ \rho$, and the parent distribution, we first generated 1,000,000 random sets of size m and used the empirical moments for the judgment post-strata to approximate the true moments. We then used Eq. (4) from Sect. 3 to compute the conditional MSE for each estimator for each possible choice of the ordered sample sizes $s_1 \geq \cdots \geq s_k > 0$ for the given m and N. By weighting each conditional MSE by the probability of getting the ordered sample sizes $s_1 \geq \cdots \geq s_k > 0$, we obtained the unconditional MSE for each estimator. We then compared the estimators by computing the relative efficiency $\text{MSE}(s^2)/\text{MSE}(\hat{\sigma}_C^2)$, which is designed so that values above 1.0 indicate an advantage for $\hat{\sigma}_C^2$.

Figures 1 and 2 show relative efficiencies for total sample sizes 10–50 when the set size is $m = 3$ (Fig. 1) and $m = 4$ (Fig. 2). The four plots in each figure correspond

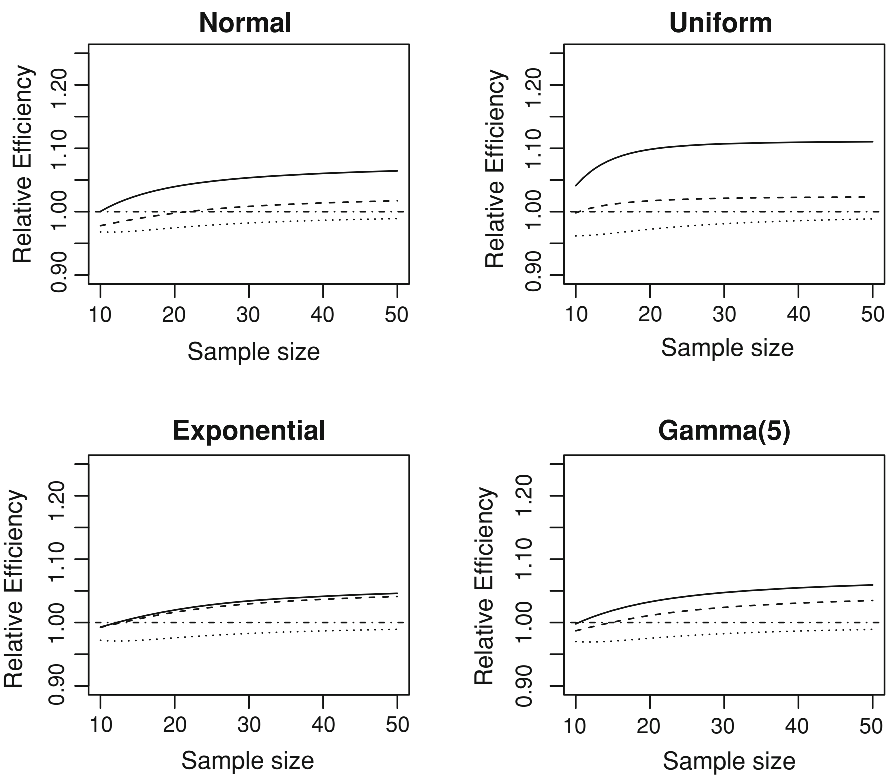

Fig. 1 Efficiency of $\hat{\sigma}_C^2$ relative to s^2 for different choices of the sample size, the parent distribution, and the quality of the rankings when $m = 3$. The rankings were done using a Dell and Clutter (1972) model in which the correlation between the perceived and true values was either 0 (*dotted curves*), 0.8 (*dashed curves*), or 1 (*solid curves*)

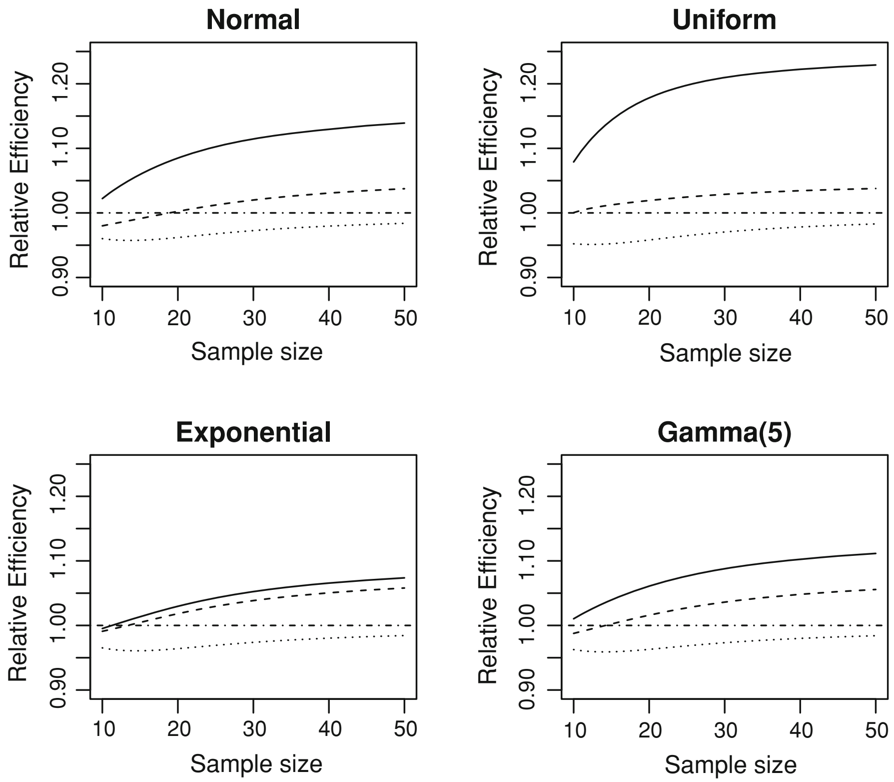

Fig. 2 Efficiency of $\hat{\sigma}^2_C$ relative to s^2 for different choices of the sample size, the parent distribution, and the quality of the rankings when $m = 4$. The rankings were done using a Dell and Clutter (1972) model in which the correlation between the perceived and true values was either 0 (*dotted curves*), 0.8 (*dashed curves*), or 1 (*solid curves*)

to the four choices of parent distribution, and the three curves in each plot correspond to ρ values of 0 (dotted curves), 0.8 (dashed curves), and 1 (solid curves). We see from the individual plots that $\hat{\sigma}^2_C$ is always less efficient than s^2 when $\rho = 0$, though never by much. With $\rho = 0.8$, $\hat{\sigma}^2_C$ is more efficient than s^2 once N exceeds 20, and with $\rho = 1.0$, $\hat{\sigma}^2_C$ is noticeably more efficient than s^2 even for sample sizes as small as $N = 10$. The relative merits of the two estimators differ a bit from one parent distribution to another, with the advantage for $\hat{\sigma}^2_C$ being larger for the symmetric parent distributions (uniform and normal) than for the right-skewed parent distributions (exponential and Gamma(5)). Comparing Fig. 1 to Fig. 2 shows that if other factors are kept fixed, the advantage for $\hat{\sigma}^2_C$ over s^2 tends to increase when the set size increases.

To compare $\hat{\sigma}^2_C$ to $\hat{\sigma}^2_M$ under a particular choice of m, N, ρ, and the parent distribution, we again computed the conditional MSE for each estimator for each possible choice of the ordered sample sizes $s_1 \geq \cdots \geq s_k > 0$ for the given m and N. However, since $\hat{\sigma}^2_M$ only applies when each post-stratum sample size is at least two, we considered only such cases. By weighting each conditional MSE by the probability of getting

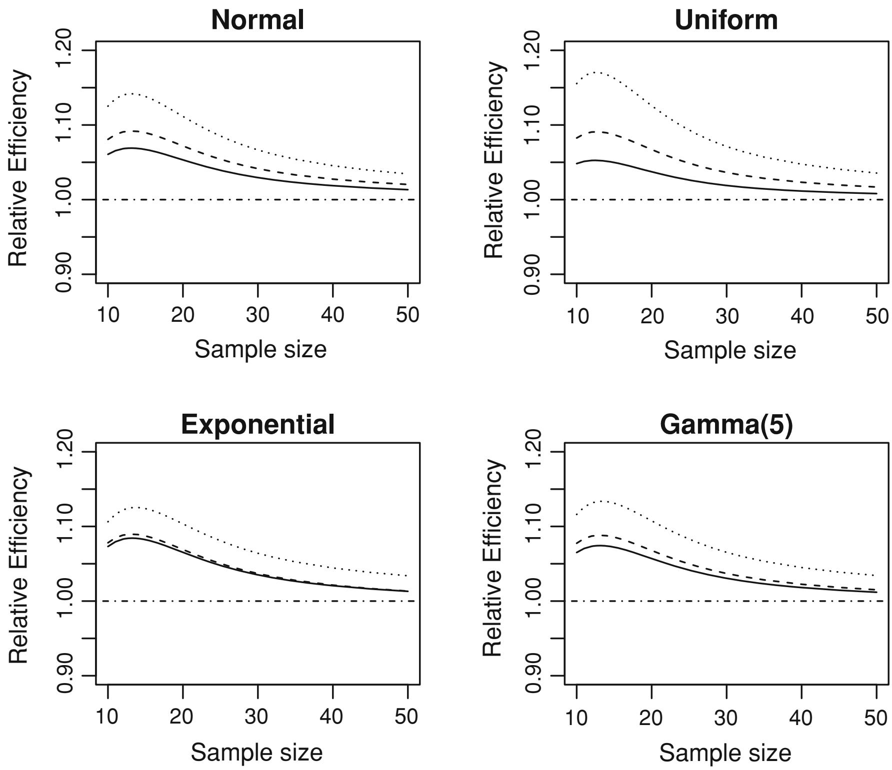

Fig. 3 Efficiency of $\hat{\sigma}_C^2$ relative to $\hat{\sigma}_M^2$ for different choices of the sample size, the parent distribution, and the quality of the rankings when $m = 3$. The rankings were done using a Dell and Clutter (1972) model in which the correlation between the perceived and true values was either 0 (*dotted curves*), 0.8 (*dashed curves*), or 1 (*solid curves*)

the ordered sample sizes $s_1 \geq \cdots \geq s_m \geq 2$, we obtained the MSE for each estimator, conditional on each post-stratum sample size being at least two. We then compared the two estimators by computing the relative efficiency $\text{MSE}(\hat{\sigma}_M^2)/\text{MSE}(\hat{\sigma}_C^2)$, which is designed so that values over 1.0 indicate an advantage for $\hat{\sigma}_C^2$.

Figures 3 and 4 show results for total sample sizes 10–50 when the set size is $m = 3$ (Fig. 3) and $m = 4$ (Fig. 4). The four plots in each figure correspond to the four choices of parent distribution, and the three curves in each plot correspond to ρ values of 0 (dotted curves), 0.8 (dashed curves), and 1 (solid curves). We see from the individual plots that $\hat{\sigma}_C^2$ is at least as efficient as $\hat{\sigma}_M^2$ in all of the settings considered. The advantage for $\hat{\sigma}_C^2$ over $\hat{\sigma}_M^2$ is largest when $\rho = 0$ and smallest when $\rho = 1$. The advantage also tends to be largest for total sample sizes near 15 and smallest for large total sample sizes. In addition, the advantage for $\hat{\sigma}_C^2$ tends to be larger for the symmetric parent distributions (uniform and normal) than for the right-skewed parent distributions (exponential and Gamma(5)). Comparing Fig. 3 to Fig. 4 shows that if other factors are fixed, the advantage for $\hat{\sigma}_C^2$ over $\hat{\sigma}_M^2$ tends to increase when the set size increases.

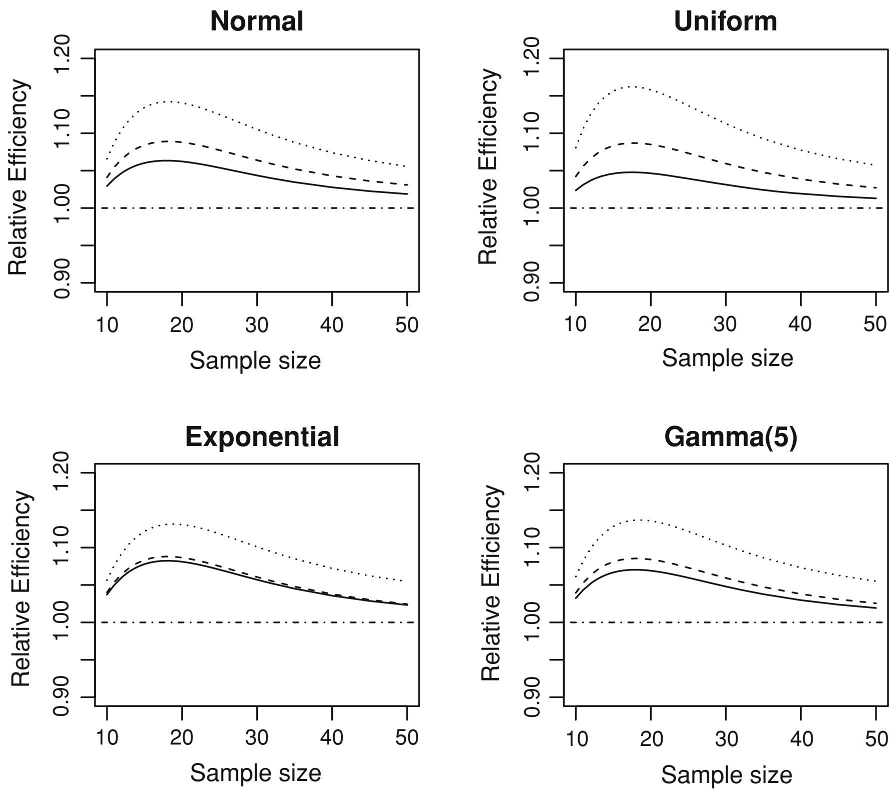

Fig. 4 Efficiency of $\hat{\sigma}_C^2$ relative to $\hat{\sigma}_M^2$ for different choices of the sample size, the parent distribution, and the quality of the rankings when $m = 4$. The rankings were done using a Dell and Clutter (1972) model in which the correlation between the perceived and true values was either 0 (*dotted curves*), 0.8 (*dashed curves*), or 1 (*solid curves*)

6 Performance comparisons using data

In order to assess the performance of $\hat{\sigma}_C^2$ with real judgment rankings as opposed to rankings determined by a model, we repeated the comparisons of Sect. 5 using a data set that is given in Web Table 2 of Wang et al. (2012). A similar data set was used by MacEachern et al. (2004). The variable of interest is the average log adjusted brain weight for a population of animal species, and the data set includes 20 sets of size three, each of which was judgment-ranked by two different rankers. The two rankers worked with the same sets, but did their judgment ranking independently of each other. Proceeding separately for the two rankers and treating the 20 sets as the population of all possible sets of size three, we obtained the first four empirical moments for each of the three judgment order statistics and treated these as the true population moments. We then computed relative efficiencies exactly as in Sect. 5. The results are given in Table 3.

We see from Table 3 that $\hat{\sigma}_C^2$ performed at least as well as s^2 in all of the tabled scenarios except the $N = 10$ case with Ranker #1. The advantage for $\hat{\sigma}_C^2$ over s^2

Table 3 Efficiency of $\hat{\sigma}_C^2$ relative to s^2 and $\hat{\sigma}_M^2$ for different choices of the total sample size when the true distribution is the empirical distribution from the brain weight data set

N	Ranker #1		Ranker #2	
	Rel. to s^2	Rel. to $\hat{\sigma}_M^2$	Rel. to s^2	Rel. to $\hat{\sigma}_M^2$
10	0.994	1.068	1.003	1.072
20	1.029	1.057	1.039	1.057
30	1.045	1.030	1.056	1.027
40	1.053	1.018	1.064	1.014
50	1.058	1.011	1.069	1.007

increased with increasing N. We also see from Table 3 that $\hat{\sigma}_C^2$ outperformed $\hat{\sigma}_M^2$ in all of the tabled scenarios, with the advantage for $\hat{\sigma}_C^2$ decreasing with increasing N. Overall, the relative performance of the estimators here is consistent with what we saw in Sect. 4.

7 Conclusions

By conditioning on the observed vector of ordered in-stratum sample sizes, we have developed a conditionally unbiased nonparametric variance estimator for use with JPS. This estimator outperforms both the standard variance estimator for unbalanced RSS and the sample variance s^2, and its efficiency relative to s^2 tends to increase when either the set size or the total sample size is increased. Frey and Feeman (2012) showed that when using JPS to estimate the population mean, the gain in efficiency from using JPS rather than SRS can be quite large when the judgment rankings are of high quality. When using JPS to estimate the variance, the gain in efficiency is smaller, but goes as high as 20 % in the scenarios considered in Sect. 5.

8 Appendix

Proof of Result 2 Since $(n_1, \ldots, n_m) \sim \text{Multinomial}(N, (1/m, \ldots, 1/m))$, each of the m post-strata is equally likely to be the one with sample size s_i. Thus,

$$E[Y_{ir}|s_1, \ldots, s_k] = \frac{1}{m} \sum_{l=1}^{m} E[X_{[l]}] = \mu,$$

where the last equality follows from Result 1. By similar logic,

$$E[Y_{ir}^2|s_1, \ldots, s_k] = \frac{1}{m} \sum_{l=1}^{m} E[X_{[l]}^2] = E[X^2].$$

Putting these two observations together, we get that

$$V(Y_{ir}|s_1, \ldots, s_k) = E[Y_{ir}^2|s_1, \ldots, s_k] - E[Y_{ir}|s_1, \ldots, s_k]^2 = E[X^2] - \mu^2 = \sigma^2.$$

This proves the result. $\square$

Proof of Result 3 By Result 2, $V(Y_{ir}|s_1,\ldots,s_k) = V(Y_{is}|s_1,\ldots,s_k) = \sigma^2$. Since each post-stratum is equally likely to be the one with sample size s_i, we have that

$$E[Y_{ir}Y_{is}|s_1,\ldots,s_k] = \frac{1}{m}\sum_{l=1}^{m} E[X_{[l]r}X_{[l]s}] = \frac{1}{m}\sum_{l=1}^{m}\mu_{[l]}^2,$$

where the last equality follows from the fact that $X_{[l]r}$ and $X_{[l]s}$ are independent draws from the lth post-stratum. Thus,

$$\begin{aligned}\mathrm{Cov}(Y_{ir}, Y_{is}|s_1,\ldots,s_k) &= \frac{1}{m}\sum_{l=1}^{m}\mu_{[l]}^2 - \mu^2\\ &= \frac{1}{m}\sum_{l=1}^{m}\mu_{[l]}^2 - \left(\frac{1}{m}\sum_{l=1}^{m}\mu_{[l]}\right)^2 = \frac{1}{m}\sum_{l=1}^{m}(\mu_{[l]}-\mu)^2,\end{aligned}$$

which means that

$$\begin{aligned}E\left[(Y_{ir}-Y_{is})^2|s_1,\ldots,s_k\right] &= V(Y_{ir}|s_1,\ldots,s_k) + V(Y_{is}|s_1,\ldots,s_k)\\ &\quad -2\mathrm{Cov}(Y_{ir}, Y_{is}|s_1,\ldots,s_k)\\ &= 2\sigma^2 - \frac{2}{m}\sum_{l=1}^{m}(\mu_{[l]}-\mu)^2.\end{aligned}$$

Since each pair of post-strata is equally likely to be the pair with sample sizes s_i and s_j,

$$\begin{aligned}E[Y_{ir}Y_{js}|s_1,\ldots,s_k] &= \frac{1}{m(m-1)}\sum_{l\neq t} E[X_{[l]r}X_{[t]s}]\\ &= \frac{1}{m(m-1)}\sum_{l\neq t}\mu_{[l]}\mu_{[t]} = \frac{1}{m(m-1)}\left[(m\mu)^2 - \sum_{l=1}^{m}\mu_{[l]}^2\right].\end{aligned}$$

Thus,

$$\begin{aligned}\mathrm{Cov}(Y_{ir}, Y_{js}|s_1,\ldots,s_k) &= \frac{1}{m(m-1)}\left[(m\mu)^2 - \sum_{l=1}^{m}\mu_{[l]}^2\right] - \mu^2\\ &= \frac{1}{m(m-1)}\left[(m^2 - m(m-1))\mu^2 - \sum_{l=1}^{m}\mu_{[l]}^2\right]\\ &= -\frac{1}{m(m-1)}\sum_{l=1}^{m}(\mu_{[l]}-\mu)^2.\end{aligned}$$

Combining this observation with our earlier findings, we have that

$$
\begin{aligned}
E\left[(Y_{ir}-Y_{js})^2|s_1,\ldots,s_k\right] &= V(Y_{ir}|s_1,\ldots,s_k)+V(Y_{js}|s_1,\ldots,s_k)\\
&\quad -2\mathrm{Cov}(Y_{ir},Y_{js}|s_1,\ldots,s_k)\\
&= 2\sigma^2+\frac{2}{m(m-1)}\sum_{l=1}^{m}(\mu_{[l]}-\mu)^2.
\end{aligned}
$$

This proves the result. □

Proof of Result 5 By Result 4, the conditional expectation of the estimator $\hat{\sigma}^2$ is

$$
\begin{aligned}
&\sum_{i=1}^{k} w_{ii}\binom{s_i}{2}\left(2\sigma^2-\frac{2}{m}\sum_{l=1}^{m}(\mu_{[l]}-\mu)^2\right)\\
&\quad+\sum_{i<j} w_{ij}s_i s_j\left(2\sigma^2+\frac{2}{m(m-1)}\sum_{l=1}^{m}(\mu_{[l]}-\mu)^2\right),
\end{aligned}
$$

which simplifies to

$$
\begin{aligned}
&\sigma^2\left\{2\sum_{i=1}^{k} w_{ii}\binom{s_i}{2}+2\sum_{i<j} w_{ij}s_i s_j\right\}\\
&\quad+\frac{2}{m}\sum_{l=1}^{m}(\mu_{[l]}-\mu)^2\left\{\frac{1}{m-1}\sum_{i<j} w_{ij}s_i s_j-\sum_{i=1}^{k} w_{ii}\binom{s_i}{2}\right\}.
\end{aligned}
$$

If the two conditions in the statement of Result 5 are met, then the first bracketed expression is 1, and the second bracketed expression is 0. Thus, the estimator is unbiased for σ^2.

To see that these two conditions are also necessary, note that the sum $\sum_l(\mu_{[l]}-\mu)^2$ takes on different values for different ranking schemes. For example, for random rankings, $\sum_l(\mu_{[l]}-\mu)^2=0$, but for perfect rankings, $\sum_l(\mu_{[l]}-\mu)^2>0$. Thus, in order for the estimator to be unbiased in both cases, the bracketed expressions must be 1 and 0, respectively. This forces the conditions in the result to hold. □

Computing $E[S_{11}^2|s_1,\ldots,s_k]$. Since the random variables $Y_{11},\ldots,Y_{1s_1}$ are exchangeable, we can write $E[S_{11}^2|s_1,\ldots,s_k]$ as

$$
\begin{aligned}
E\left[\left(\sum_{r<s}(Y_{1r}-Y_{1s})^2\right)^2|s_1,\ldots,s_k\right] &= \binom{s_1}{2}E\left[(Y_{11}-Y_{12})^4|s_1,\ldots,s_k\right]\\
&\quad+s_1(s_1-1)(s_1-2)E\left[(Y_{11}-Y_{12})^2(Y_{11}-Y_{13})^2|s_1,\ldots,s_k\right]\\
&\quad+\binom{s_1}{2}\binom{s_1-2}{2}E\left[(Y_{11}-Y_{12})^2(Y_{13}-Y_{14})^2|s_1,\ldots,s_k\right].
\end{aligned}
$$

Expanding the expected values in terms of moments of the $\{Y_{ir}\}$, we obtain the equivalent expression

$$\binom{s_1}{2}\left\{E\left[Y_{11}^4|s_1,\ldots,s_k\right]-8E\left[Y_{11}^3Y_{12}^1|s_1,\ldots,s_k\right]+6E\left[Y_{11}^2Y_{12}^2|s_1,\ldots,s_k\right]\right\}$$
$$+s_1(s_1-1)(s_1-2)\left\{E\left[Y_{11}^4|s_1,\ldots,s_k\right]-4E\left[Y_{11}^3Y_{12}^1|s_1,\ldots,s_k\right]\right.$$
$$\left.+3E\left[Y_{11}^2Y_{12}^2|s_1,\ldots,s_k\right]\right\}+\binom{s_1}{2}\binom{s_1-2}{2}$$
$$\times\left\{4E\left[Y_{11}^2Y_{12}^2|s_1,\ldots,s_k\right]-8E\left[Y_{11}^2Y_{12}Y_{13}|s_1,\ldots,s_k\right]\right.$$
$$\left.+4E\left[Y_{11}Y_{12}Y_{13}Y_{14}|s_1,\ldots,s_k\right]\right\}.$$

We can then find the necessary expected values using the same logic that we used in proving Results 2 and 3. For example, since each of the m post-strata is equally likely to have been the one with sample size s_1, we have that

$$E[Y_{11}^2Y_{12}^2|s_1,\ldots,s_k]=\frac{1}{m}\sum_{i=1}^{m}E[X_{[i]}^2]^2,$$

where $X_{[i]}$ is an ith judgment order statistic from the parent distribution.

Acknowledgments The authors thank the referees for helpful suggestions that have improved the paper.

References

Chen, H., Stasny, E. A., Wolfe, D. A. (2005). Ranked set sampling for efficient estimation of a population proportion. *Statistics in Medicine*, *24*, 3319–3329.

David, H. A., Nagaraja, H. N. (2003). *Order Statistics* (3rd ed.). New York: Wiley.

Dell, T. R., Clutter, J. L. (1972). Ranked set sampling with order statistics background. *Biometrics*, *28*, 545–555.

Du, J., MacEachern, S. N. (2008). Judgement post-stratification for designed experiments. *Biometrics*, *64*, 345–354.

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

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

Halls, L. K., Dell, T. R. (1966). Trial of ranked-set sampling for forage yields. *Forest Science*, *12*, 22–26.

Howard, R. W., Jones, S. C., Mauldin, J. K., Beal, R. H. (1982). Abundance, distribution, and colony size estimates for *Reticulitermes* spp. (Isopter: Rhinotermitidae) in Southern Mississippi. *Environmental Entomology*, *11*, 1290–1293.

Kvam, P. H. (2003). Ranked set sampling based on binary water quality data with covariates. *Journal of Agricultural, Biological, and Environmental Statistics*, *8*, 271–279.

Lange, K. (1999). *Numerical Analysis for Statisticians*. New York: Springer.

Lohr, S. L. (1999). *Sampling: Design and Analysis*. CA, Duxbury: Pacific Grove.

MacEachern, S. N., Ozturk, O., Wolfe, D. A., Stasny, E. A. (2002). A new ranked set sample estimator of variance. *Journal of the Royal Statistical Society, Series B*, *64*, 177–188.

MacEachern, S. N., Stasny, E. A., Wolfe, D. A. (2004). Judgement post-stratification with imprecise rankings. *Biometrics*, *60*, 207–215.

McIntyre, G. A. (1952). A method for unbiased selective sampling, using ranked sets. *Australian Journal of Agricultural Research*, *3*, 385–390.
McIntyre, G. A. (2005). A method for unbiased selective sampling, using ranked sets. *The American Statistician*, *59*, 230–232 (Originally appeared in *Australian Journal of Agricultural Research*, 3, 385–390).
Perron, F., Sinha, B. K. (2004). Estimation of variance based on a ranked set sample. *Journal of Statistical Planning and Inference*, *120*, 21–28.
Stokes, L. (1980). Estimation of variance using judgment ordered ranked set samples. *Biometrics*, *36*, 35–42.
Wang, X., Lim, J., Stokes, L. (2008). A nonparametric mean estimator for judgment poststratified data. *Biometrics*, *64*, 355–363.
Wang, X., Stokes, L., Lim, J., Chen, M. (2006). Concomitants of multivariate order statistics with application to judgment poststratification. *Journal of the American Statistical Association*, *101*, 1693–1704.
Wang, X., Wang, K., Lim, J. (2012). Isotonized CDF estimation from judgment poststratification data with empty strata. *Biometrics*, *68*, 194–202.
Yu, P. L. H., Lam, K., Sinha, B. K. (1999). Estimation of normal variance based on balanced and unbalanced ranked set samples. *Environmental and Ecological Statistics*, *6*, 23–45.