Jackknife and Other Replication Methods with a Reduced Number of Replicates

Stephen Ash
U.S. Census Bureau, 4600 Silverhill Road, Washington D.C. 20233

Abstract
Replicate weights used with surveys are becoming more available and expected for public use data. Data users can use replicate weights to estimate their own variances for point estimates or use them with sophisticated software for multivariate analyses.

This paper will discuss how to produce replicate weights with a reduced number replicates for the jackknife and other replication estimators – including using balanced repeated replication and successive difference replication. Our focus will be the use of these estimators with systematic random sampling from an ordered list. The reduced number of replicates is needed because data users require a reasonable number of replicates in their analysis. The paper will compare all of the replication estimators and examine how the reduction of replicates impacts the resultant variance estimators.

Key Words: variance estimation, jackknife estimator, balanced repeated replication, successive difference replication

1. Introduction

Many surveys use a single-stage sample design to select more than two units per strata ($n_h > 2$). The sample is often selected with systematic random sampling from an ordered list or simply *sys*. The *sys* sample design can produce estimates with smaller variances than simple random sample without replacement (*srswor*) when the auxiliary information used with the sort order is associated with the variable of interest. We seek replication variance estimators that achieve the following three goals with respect to estimating the variance of a *sys* sample design.

G1. Replication variance with a varying number of replicates. The number of replicates should be amendable so that we can manage the number according to our needs. We prefer to provide a large number of replicates to reduce the variance of the variance estimators, but understand in practice that we are all constrained by real-world file sizes and run times.
G2. Simple estimator of the variance of a *sys* sample design. We aim to have an expression of the variance estimator that is simple. The variance estimator will be used by an audience with varying levels of sophistication.
G3. Appropriate for *sys*. We also would like estimators that are appropriate for estimating the variance from a systematic random sample and do not overestimate the variance as *srswor* variance estimators generally do.

The replication estimators that we consider include the jackknife (both the regular jackknife and the delete-a-group jackknife), the successive difference replication (SDR) estimator, and an application of the balanced repeated replication (BRR) estimator for *sys*.

In a related variance estimation problem, with two-stage sample designs, we sometimes try to estimate the second-stage variance apart from the overall variance. We show how an application of SDR and BRR can lead to estimators of the second-stage variance where the sample design within the first-stage units is *sys*. Additionally, we also discuss an estimator suggested by Rizzo and Rust (2011).

1.1 Review of Systematic Sampling

For the remainder of our discussion, *sys* will be used as shorthand for systematic random sampling from an ordered list. We abbreviate *sys* this way because systematic sampling from an unordered or randomly ordered list can be shown to be equivalent to simple random sampling (Madow and Madow 1944). For our discussion, we focus solely on equal probability selection and selecting sample in only one dimension. Other names for *sys* include "linear systematic sampling" (Murthy and Rao 1988) and "1-in-*a* sampling" (Gregoire and Valentine 2008).

The sample design *sys* is easy to implement and can be very efficient compared with *srswor*. To implement *sys*, we first sort the universe by a variable that is known for every unit in the universe. With a defined sampling interval $k > 0$, we randomly generate r from a uniform distribution on the interval (0, k]. The units selected are spaced in multiples of the sampling interval from the first selection, i.e, $\lceil r + i * k \rceil$, i = 1, 2, ,.., n and we define $\lceil . \rceil$ as the ceiling function or the next largest integer.

We say that *sys* can be efficient in the sense that the sample design can produce estimates with small sample variances as compared to *srswor*. Cochran (1977) relates the efficiency of *sys* to the intra-cluster correlation. Although the inter-cluster correlation is not the same as a simple correlation, both provide a measure of the association between the variable of interest and variable(s) used to sort the universe prior to sample selection. If the variable of interest is highly associated with the sort variable, the sample design can be very efficient.

The efficiency of *sys* can also be understood in the context of the term *implicit stratification* used by Megill *et al*. (1987). This way of thinking was also discussed by Cochran (1977) where, the universe as a sorted list is divided into $\hat{n} = \lceil N / k \rceil$ implicit strata. The first $\lceil k \rceil$ units are in the first strata, the next $\lceil k \rceil + 1$ to $\lceil 2k \rceil$ units are in the second strata,…, and the last stratum from $\lceil (n-1) \times k \rceil$ to N. The random number r determines the random selection within the first implicit stratum and each of the subsequent strata. Since the universe is sorted, each stratum has units that are similar to each other with respect to the sort variable. This can be efficient when the sort variable and the variable of interest are associated, since the implicit stratification would also group units that are similar to each other with respect to the variable of interest.

Excellent summaries of *sys* and estimating variances from *sys* can be found in Iachan (1982), Wolter (1984), Murthy and Rao (1988), and Bellhouse (1988). In the next section, we review successive differences (SD) in preparation of our discussion of SDR.

1.2 Notation

In most cases we can define the general linear statistic of interest for a single-stage stratified sample design as

$$Y = \sum_{h} \sum_{k \in U_h} y_k \tag{1}$$

and we define a general nonlinear statistic of interest as $\theta = T(Y)$ in terms of linear totals Y where h indexes the set of all possible strata and k indexes the universe of interest U_h for stratum h. For example, if θ was the ratio estimator then $\theta = Y_1/Y_2$ and the totals Y_1 and Y_2 are defined as in (1). An estimator Y is

$$\hat{Y} = \sum_{h} \sum_{k \in s_h} w_k y_k \tag{2}$$

where s_h is the sample in each stratum h and an estimator of θ is $\hat{\theta} = T(\hat{Y})$ that is defined in terms of estimated totals as in (2). The replicate estimate for replicate r is defined as $\hat{\theta}_r = T(\hat{Y}_r)$ and replicate totals are defined as

$$\hat{Y}_r = \sum_{h} \sum_{k \in s_h} w_{rk} y_k \tag{3}$$

where the replicate weight w_{rk} is a function of the regular weight w_k and a replicate factor F_{rk} that we subsequently define for each different replication method. In our paper, we take the simplest case of a replicate weight where the replicate weight is equal to the product of the sample weight and the replicate factor, i.e., $w_{rk} = w_k\ F_{rk}$. Often other weighting adjustments like adjustments for noninterviews or adjustments that use known totals are applied to the w_{rk} and separately for each replicate as a way for the replicate estimator to account for the variability due to the additional weighting adjustments.

2. Replication Variance Estimators for Single-Stage Stratified Sample Designs

In this section, we review several replicate estimators that can be applied to single-stage stratified sample designs. We begin with the jackknife and discuss the original jackknife, the delete-*d* jackknife, and the grouped jackknife. Although we will focus on the grouped jackknife as a replication estimator, we include the original jackknife and the delete-*d* jackknife for comparisons and because we also discuss how they are related to the grouped jackknife.

Additionally, we will examine the SDR and BRR estimators. The SDR estimator is especially suited for estimating variances of *sys*. Developed by Fay and Train (1995), it is a replication estimator that mimics the SD estimator that was discussed by Wolter (1984) as an estimator of variances of *sys* sample designs. The last estimator we examine is BRR which is not normally considered for estimating the variance of a *sys* sample design. We show a way of applying it that piggybacks on the basic ideas of SDR.

Before we begin we note that we already know that the jackknife is not suitable for measuring the variance of the estimates from a *sys* sample design. As noted by Kott (2001) with respect to the delete-a-group jackknife (DAGJK), "[l]et us assume that the sample was selected without replacement but the selection probabilities are all so small, and the joint selection probabilities are such, that using the with-replacement variance estimator is appropriate (this rules out systematic sampling from a purposefullyordered list).” The reason we include the jackknife in our discussion is that we want to include it in our numerical comparisons to show how well it performs with respect to *sys*.

2.1 Delete-1 Jackknife (JK-1)

The original jackknife variance estimator of the variance of $\hat{\theta}$ leaves out one sample unit for each replicate and is defined as

$$\hat{v}_{\text{JK}-1}(\hat{\theta}) = \sum_{h \in H} \frac{n_h - 1}{n_h} \sum_{k \in s_h} \left(\hat{\theta}_{(hk)} - \bar{\hat{\theta}}_{(h)}\right)^2 \tag{4}$$

where

$$\bar{\hat{\theta}}_{(h)} = \sum_{k \in s_h} \hat{\theta}_{(hk)} / n_h.$$

The replicate estimate of a general total of *Y* by leaving out unit k' of strata h' is defined as

$$\hat{Y}_{(h',k')} = \sum_{\substack{h \in H \\ h \neq h'}} \sum_{k \in s_h} w_k y_k + \frac{n_{h'}}{n_{h'} - 1} \sum_{\substack{k \in s_{h'} \\ k \neq k'}} w_k y_k$$

The replicate factors for JK-1 are expressed as

$$F_{k,h} = \begin{cases} 1 & h \neq h' \\ 0 & h = h' \; and \;\; k = k' \\ n_h/(n_h - 1) & h = h' \; and \;\; k \neq k' \end{cases}$$

With respect to our goals, G1: JK-1 cannot vary the number of replicates: the number of replicates for JK-1 is equal to the sample size, so the number of replicates is generally too large, G2: the variance estimator $\hat{v}_{\text{JK}-}\ (\hat{Y})$ is generally complex since the mean of the replicate estimates in (4) only includes replicates from the same stratum, and G3: we know it is not appropriate for the *sys* sample design, but we include it because it is a precursor of the JK-*d* and DAGJK.

2.2 Delete-d Jackknife (JK-*d*)

The JK-*d* leaves out d_h sample units for each replicate and is defined as

$$\hat{v}_{\text{JK}-d}(\hat{\theta}) = \sum_{h \in H} \frac{n_h - d_h}{d_h \cdot m_h} \sum_{g=1}^{m_h} \left(\hat{\theta}_{(hg)} - \bar{\hat{\theta}}_{(h)}\right)^2 \tag{5}$$

where $m_h = \binom{n_h}{d_h}$ sets of size d_h are removed, and the average replicate estimate is defined as

$$\bar{\hat{\theta}}_{(h)} = \sum_{g=1}^{m_h} \hat{\theta}_{(hg)} / m_h.$$

The replicate estimate of a general total of *Y* by leaving out unit k' of strata h' is defined as

$$\hat{Y}_{(h',g)} = \sum_{\substack{h\in H \\ h\neq h'}} \sum_{k\in s_h} w_k y_k + \frac{n_{h'}}{n_{h'} - d_{g(h)}} \sum_{\substack{k\in s_{h'} \\ k\neq D_g}} w_k y_k$$

and replicate factors for JK-*d* are expressed as

$$F_{k,h} = \begin{cases} 1 & h \neq h' \\ 0 & h = h' \; and \;\; k = k' \\ n_h/(n_h - d_h) & h = h' \; and \;\; k \neq k' \end{cases}$$

How to delete groups. An issue that we highlight throughout this paper is how replicates are formed and that question is answered by understanding how sample units are deleted for each replicate. With the JK-*d*, the number of replicates is equal to the number of groups that are deleted or $R = \sum_h m_h$. For the JK-*d* estimator, Shao and Tu (1995) suggests three way to form the groups.

(M1) Use all possible samples.
(M2) Random samples.
(M3) Balanced subsampling (p. 197) where a balanced subsample has the following properties:

1) Every unit *i* appears in the same number of subsamples
2) Every pair (i, j) appears in the same number of subsamples

See also John (1971) for more about balanced incomplete block design (BIBD).

With respect to our goals, G1: DAGJK can vary the number of replicates, G2: the variance estimator $\hat{v}_{\text{JK}-d}(\hat{Y})$ is generally too complex since the mean of the replicate estimates in (5) only includes replicates from the same stratum, and G3: we know it is not appropriate for the *sys* sample design.

2.2 Delete-a-Group Jackknife (DAGJK)

The grouped jackknife (Shao and Wu 1989) or the delete-a-group jackknife (Kott 2001) leaves out group *D_r* (the set of sample units such that $k \in D_r$) for each replicate *r* and the number of replicates is *R*. Additionally, the groups D_r are disjoint and cover the entire sample, i.e., $s = \bigcup_r D_r$ and $D_r \cap D_{r'} = \emptyset$ where $r \neq r'$. We also define d_{hr} as the number of sample units left out for a given stratum *h* and group D_r. Then define the replicate estimate of $\hat{\theta}$ by leaving out group D_r for replicate *r* as

$$\hat{v}_{\text{DAGJK}}(\hat{\theta}) = \frac{R-1}{R} \sum_{r=1}^{R} \left(\hat{\theta}_r - \bar{\hat{\theta}}_r\right)^2$$

and the replicate estimates is defined as

$$\hat{Y}_r = \sum_h \frac{n_h}{n_h - d_{hr}} \sum_{\substack{k\in s_h \\ k\notin D_{hr}}} w_k y_k$$

For the DAGJK, the variance estimator is (3) and the replicate factors are expressed as

$$F_r = \begin{cases} n/(n - d_{hr}) & k \notin D_r \\ 0 & k \in D_r \end{cases}$$

How to delete the groups? Kott (2001) suggests "[i]n order to estimate var(t) with a DAGJK, we first order the strata in some fashion and then order the PSUs within each stratum randomly. The sample is partitioned into *R* systematic samples using the ordered list." Bienias, Kott, and Evans (2003) say that "[t]he technique divides the first-phase sample into mutually exclusive and nearly equal variance groups."

Deleting groups for *sys*. Per forming groups, both DAGJK and JK-*d* with method (M2) forms groups by randomly ordering the original sample. This generally means that the estimator is appropriate for *srswor* and not *sys*. For use with the *sys* sample design, instead of randomly ordering, we deleted groups that were *sys* samples from the original *sys* sample. Our aim of removing groups that were *sys* samples from the original *sys* sample was to "replicate" the *sys* sample design in each replicate.

With respect to our goals, G1: DAGJK can use a reduced set of replicates, G2: the variance estimator $\hat{v}_{DAGJK}(\hat{Y})$ is generally simple, and G3: we know it is not suitable for the *sys* sample design, but we include it in our comparisons. With respect to G2, DAGJK is chosen over the JK-*d* since the DAGJK is generally simpler to implement than the JK-*d* since the JK-*d* requires different averages of replicates for different strata. Most software can handle this, but not all. Second, forming groups with the DAGJK is simpler than with the JK-*d*.

2.4 Successive Difference Replication (SDR)

The SDR estimator as described by Fay and Train (1995) mimics the successive difference (SD) estimator and under conditions given by Ash (2014), the SDR estimator is equivalent to the SD estimator

$$\hat{v}_{\text{SD}}(\hat{Y}) = \frac{1}{2}(1-f)\left[\sum_{k=2}^{n}(y_k - y_{k-1})^2 + (y_n - y_1)^2\right]. \tag{6}$$

where *f* is the sampling fraction. The SDR variance estimator is expressed as

$$\hat{v}_{\text{SDR}}(\hat{Y}) = (1-f)\frac{4}{R}\sum_{r=1}^{R}\left(\hat{Y}_r - \hat{Y}\right)^2 \tag{7}$$

with replicate factors

$$F_{k,r} = 1 + 2^{-\frac{3}{2}}a_{a_k,r} - 2^{-\frac{3}{2}}a_{b_k,r} \tag{8}$$

With respect to our goals, G1: Ash (2014) showed that SDR can use a reduced set of replicates, G2: the variance estimator $\hat{v}_{\text{SDR}}(\hat{Y})$ is not complex, and G3: the SD and SDR estimators are especially suited for *sys* sample designs.

2.5 Application of Balanced Repeated Replication (BRR-SYS) to *sys*

The general BRR variance estimator of McCarthy (1966) works with sample designs that have many strata and a sample size of $n_h = 2$ units per strata. For our discussion, we consider Fay's Method (Dippo, Fay, and Morganstein 1984) due to its improved properties as described by Judkins (1990). The BRR-FAY variance estimator is expressed as

$$\hat{v}_{\mathrm{BBR-F}}\quad\left(\hat{Y}\right) = \frac{1}{R(1-k)^2}\sum_{r=1}^{R}\left(\hat{Y}_r - \hat{Y}\right)^2$$

and the replicate estimates are defined as

$$\hat{Y}_r = \sum_h \left[\frac{\hat{Y}_{h,i=1}}{\pi_{h,i=1}}\left(1 + a_{h,r}(1-k)\right) + \frac{\hat{Y}_{h,i=2}}{\pi_{h,i=2}}\left(1 - a_{h,r}(1-k)\right)\right]$$

and the replicate factors can be expressed as

$$F_{r,h} = \begin{cases} 1 + (1-k)a_{h,r} & i = 1 \\ 1 - (1-k)a_{h,r} & i = 2 \end{cases}$$

In our notation a Hadamard matrix **H** has elements $a_{\text{row,column}}$ which correspond to $a_{\text{strata, replicate}}$.

How to form pseudo strata and half samples for *sys*. We suggest applying BRR to a *sys* sample design using the same rationale and general approach as SDR uses. SDR mimics SD and SD works because it treats a *sys* sample design as a stratified sample design with one unit selected per strata or implicit stratification (Megill *et al*. 1987). The SD estimator uses a collapsed strata approach to estimate the variance of all possible adjacent pairs of strata within the *sys* sample.

With BRR, we apply the same approach by collapsing the implicit strata that have one sample unit into a pseudo stratum also referred to as a variance stratum. BRR refers to each of the two sample units as half samples. We suggest collapsing each consecutive and non-overlapping pair and then estimating the stratum variance with BRR. So SDR and our suggested application of BRR compare with one another in that SDR will use the set of all possible adjacent pairs (1,2), (2,3), (3,4)…and BRR will use one of the two sets of all possible adjacent and non-overlapping pairs, i.e. (1,2), (3,4), (5,6)...

The ordered sample units *k* are assigned to the variance strata h' and half sample i' as described in Table 1

Table 1. Assignment of Variance Strata and Half Sample within each Second-Stage Unit *i*

Sample unit *k*	1, 2, 3, 4, 5, 6,…
Cycle *c*	1, 1, 1, 1, 1, 1,…, 1, 1, 2, 2, 2, 2,…
Variance strata h'	1, 1, 2, 2, 3, 3,…, R, R, 1, 1, 2, 2,…
Half sample i'	1, 2, 1, 2, 1, 2,…, 1, 2, 1, 2, 1, 2,…

Result 2 in the Appendix provides the justification of Fay's Method and Result 3 shows how our application of $\hat{v}_{\mathrm{BBR-F}}\quad\left(\hat{Y}\right)$ is equivalent to

$$\frac{1}{R(1-k)^2}\sum_r\left(\hat{Y}_r - \hat{Y}\right)^2 = \hat{v}_{\mathrm{SRSWR}}\left(\hat{Y}\right) + \hat{A}\left(\hat{Y}\right) \tag{9}$$

and we refer to the variance estimator of (9) as $\hat{v}_{\text{BRR}-\text{SYS}}(\hat{Y})$, where

$$\hat{v}_{\text{SRSWR}}(\hat{Y}) = \sum_{c=1}^{C} \sum_{h'=1}^{R} \frac{1}{2(2-1)} \sum_{i'=1}^{2} \left(\frac{y_{ch'i'}}{p_{ch'i'}} - \hat{\hat{y}}_{ch'} \right)^2$$

is the sum of several $n_h = 2$ *srswr* variance estimators, one for each stratum h' and where

$$\hat{A}(\hat{Y}) = \sum_{c=1}^{C} \sum_{\substack{c'=1 \\ c' \neq c}}^{C} \sum_{h'=1}^{R} \left(w_{c\,h'i'=1} y_{c\,h'i'=1} - w_{c\,h'i'=2} y_{c\,h'i'=2} \right) \left(w_{c'h'i'=1} y_{c'h'i'=1} - w_{c'h'i'=2} y_{c'h'i'=2} \right)$$

and the estimated total of variance stratum h' is

$$\hat{y}_{ch'} = \frac{1}{2} \sum_{i'=1}^{2} \frac{y_{ch'i'}}{p_{ch'i'}}$$

The estimator $\hat{v}_{\text{BRR}-\text{SYS}}(\hat{Y})$ is the sum of n / 2 with replacement variance estimators, one for each pseudo strata h'. The extra term $\hat{A}(\hat{Y})$ has an equal number of positive and negative terms, does not equal zero, but approximately cancels out.

With respect to our goals, G1: Our application of BRR can use a reduced set of replicates, G2: the variance estimator $\hat{v}_{\text{BRR}-\text{SYS}}(\hat{Y})$ is generally simple, and G3: we think it should similarly mimic the SD estimator as SDR does so it should be suited for *sys* sample designs. This will be examined further in our empirical examples.

3. Replication Variance Estimators for Second-Stage Variance

In this section, we discuss a different problem: replication variance estimators for the estimation of the second-stage variance from a two-stage sample design. We consider three different replication methods: an application of SDR and BRR and the estimator suggested by Rizzo and Rust (2011). We begin by reviewing estimation for a two-stage sample design.

3.1 Review of Two-Stage Variance Estimation

Let the overall total of the variable of interest y_k be defined as $Y = \sum_h Y_h$, the total of stratum h as $Y_h = \sum_{i \in U_1} Y_i$, and the total for first-stage unit i as $Y_i = \sum_{k \in U_i} y_k$. We estimate the overall total as $\hat{Y} = \sum_h \hat{Y}_h$, the total of stratum h as $\hat{Y}_h = \sum_{i \in s_h} w_i \hat{Y}_i$, and the total for first-stage unit i as $\hat{Y}_i = \sum_{k \in s_i} w_k y_k$ where $w_{1i} = \pi_{2i}^{-1}$ and w_{2k} are the sample weights for the first and second stages, respectively.

$$\hat{Y} = \sum_{h} \sum_{i \in s_h} \sum_{k \in s_i} w_i w_k y_k \tag{10}$$

Using notation from Särndal *et al*. (1992; p. 137), we can express the variance of $\hat{Y}$ as $v(\hat{Y}) = v_{\text{PSU}}(\hat{Y}) + v_{\text{SSU}}(\hat{Y})$. Our interest lies not in the first-stage variance $v_{\text{PSU}}(\hat{Y})$ but in the second-stage variance which can be expressed as

$$v_{\mathrm{SS}}\ \left(\hat{Y}\right) = \sum_{h}\sum_{i\in U_h}\frac{v\left(\hat{Y}_i\right)}{\pi_i}$$

The estimator of the two-stage variance is as $\hat{v}\left(\hat{Y}\right) = \hat{v}_{\mathrm{PSU}}\left(\hat{Y}\right) + \hat{v}_{\mathrm{SSU}}\left(\hat{Y}\right)$ and the estimator of the second-stage variance is

$$\hat{v}_{\mathrm{SSU}}\left(\hat{Y}\right) = \sum_{h}\sum_{i\in s_h}\frac{\hat{v}\left(\hat{Y}_i\right)}{\pi_i^2}$$

We want to find simple and easy-to-use replication estimator of $v_{\mathrm{SSU}}\left(\hat{Y}\right)$.

3.2 Application of SDR for Second-Stage Variance Estimation

Our proposed estimator is a slight modification of the SDR estimator described by Ash (2014). Here we define the replicate factors for sample unit *k* and replicate *r* as

$$F_{hi} \ = 1 + \sqrt{1 - f_i}\left(2^{-\frac{3}{2}}h_{a_k,r} - 2^{-\frac{3}{2}}h_{b_k,r}\right) \tag{11}$$

The replicate factors (11) are applied to the two-stage weight $w_i\, w_k$ and then the variance estimator (6) is applied. More formally and following the outline of Theorem 1 of Ash (2014):

Theorem 1: Let $n = \sum_h \sum_{i\in s_h} n_{hi}$ be the overall sample size and n_{hi} be the second-stage sample size of first-stage sample unit *i* of stratum *h*. Define the combined vector of the variable of interest as $\breve{\mathbf{y}}' = [\breve{\mathbf{y}}_{h=1}\ \breve{\mathbf{y}}_{h=2} \dots \breve{\mathbf{y}}_{h=L}]$, for stratum *h* as $\breve{\mathbf{y}}'_h = \left[\breve{\mathbf{y}}_{i=1}\ \breve{\mathbf{y}}_{i=2} \dots \breve{\mathbf{y}}_{i=n_h}\right]$, and with each first-stage sample unit $\breve{\mathbf{y}}'_i = \left[\breve{y}_{k=1}\ \breve{y}_{k=2} \dots \breve{y}_{k=n_i}\right]$ as the $n \times 1$ weighted observation vector, where the order of the observations reflects the sort order of *sys*.

(a) Choose a Hadamard matrix of order *k* ($\mathbf{HH}' = k\mathbf{I}$), where $n_{hi} \leq k$.
(b) Choose a RA that assigns two rows (a_i, b_i) to each unit *i* in the sample. Let the RA define C_i connected loops of m_{C_i} units in each connected loop *c*. There may be more than one connected loop within a PSU, but a given connected loop does not cross multiple PSUs.
(c) Choose the $m = n$ rows of **H** corresponding to the RA to make the $m \times k$ matrix **M**. The order of the rows of **M** should correspond to the first row of the RA. For example, the first row of **M** should be row $a_{i=1}$ of **H**, the second row should be row $a_{i=2}$ of **H**, etc. Next define the $m \times m$ shift matrix as $S = block(\mathbf{S}_1, \mathbf{S}_2, \dots \mathbf{S}_C)$ where the $m_c \times m_c$ one row shift matrices $\mathbf{S}_c$ are defined to identify the position of the second row b_i of the RA in **M**. In general, each shift matrix $\mathbf{S}_c$ will be a shift-up, shift-down, or a 2 × 2 shift matrix (see the subsequently defined $\mathbf{S}_4$).

Define the estimator for each replicate total *r* as

$$\hat{Y}_r = \sum_{h}\sum_{i\in s_1}\sum_{k\in s_i}\breve{y}_k F_{hik}$$

where the replication factors are as in (11), $\breve{\breve{y}}_k = w_i w_k y_k$, $\breve{y}_k = w_k y_k$ and the matrix of replicate factors is

$$\mathbf{F} = \mathbf{1}_m \mathbf{1}'_k + \boldsymbol{\gamma} \,\#\left(2^{-\frac{3}{2}}\mathbf{I}_m - 2^{-\frac{3}{2}}\mathbf{S}\right)\mathbf{M}.$$

We also define $\boldsymbol{\gamma} = [\breve{\boldsymbol{\gamma}}_{h=1}\, \breve{\boldsymbol{\gamma}}_{h=2} \cdots \breve{\boldsymbol{\gamma}}_{h=L}]$, $\breve{\boldsymbol{\gamma}}'_h = \left[\breve{\boldsymbol{\gamma}}_{i=1}\, \breve{\boldsymbol{\gamma}}_{i=2} \cdots \breve{\boldsymbol{\gamma}}_{i=n_h}\right]$, and within each first-stage sample unit $\breve{\boldsymbol{\gamma}}'_i = (1 - f_i)^{\frac{1}{2}}\mathbf{1}_{n_i}$. $\mathbf{I}_m$ is a $m \times m$ identity matrix and $\mathbf{1}_m$ is a $m \times 1$ vector of 1s. The individual values of the replication factor within the matrix are defined for each unit *i* (rows of **F**) of replicate *r* (columns of **F**) as in (42). Then the SDR variance estimator for the second-stage variance

$$\hat{v}_{\text{SDR-SSU}}(\hat{Y}) = \frac{4}{R}\sum_{r=1}^{R}\left(\hat{Y}_r - \hat{Y}\right)^2$$

is equivalent to the sum of C_i different SD2 estimators for each first-stage unit *i*.

Result 4 of the Appendix provides the proof for Theorem 1. The estimator works for two reasons:

(a) The connected loops are formed within each first-stage unit *i*. This ensures that we have an estimator of the form of (6) for each first-stage unit.

(b) The first-stage weight w_i is squared by the sum of squares and becomes $1/\pi^2$ in $\hat{v}_{\text{SSU}}(\hat{Y})$.

3.3 Application of Balanced Repeated Replication (BRR-SYS2) for Second-Stage Variance Estimation of *sys*

We propose a replication variance estimator of $v_{\text{SSU}}(\hat{Y})$ that is similar to our BRR estimator of a *sys* sample. The estimator has an extra level of complexity due to the first-stage sample design. Given a two-stage stratified sample design with $n_h = 2$ first-stage units per the first-stage stratum *h*, and n_i second-stage sample units within each first-stage unit *i*, define the estimator for the replicate total of the variable of interest *y* for each replicate *r* as

$$\hat{Y} = \sum_h \sum_{i \in s_h} \sum_{k \in s_i} w_{1i} w_{2k} y_k \qquad (12)$$

where the first-stage weight is $w_{1i} = \pi_i^{-1}$ and the second-stage weight is w_{2k}. We can rewrite (12) as

$$\hat{Y} = \sum_h \sum_{i \in s_h} \sum_{c_i=1}^{C_i} \sum_{h'_i=1}^{R_2} \sum_{i'_i=1}^{2} w_{1i} w_{2 i'_i} y_{i'_i}$$

`

where R_2 is defined in Table 3 below. With the first-stage, define the values of h^* as shown in Table 2.

Table 2: Assignment of h^*

h^*	1, 2, 3, 4, 5, 6,…, 2*L*-1, 2*L*
First-stage strata *h*	1, 1, 2, 2, 3, 3,…, *L*, *L*
First-stage unit *i*	1, 2, 1, 2, 1, 2,…, 1, 2

where L is the total number of first-stage strata?. For our result, we have assumed that $n_h = 2$ for all of the first-stage strata h, but this result works with any number of first-stage units n_h: we only need to assign a unique h^* to each first-stage unit.

Within the first-stage stratum h and first-stage unit i, define the variance strata h_i' and half sample i_i' for the ordered sample units k as shown in Table 3.

Table 3: Assignment of Variance Strata and Half Sample within each First-Stage Unit i

Sample unit k	1, 2, 3, 4, 5, 6,…
Cycle c_i	1, 1, 1, 1, 1, 1,…, 1, 1, 2, 2, 2, 2,…, C_i, C_i
Variance strata h_i'	1, 1, 2, 2, 3, 3,…, R_2, R_2, 1, 1, 2, 2,…, 2, 2
Half sample i_i'	1, 2, 1, 2, 1, 2,…, 1, 2, 1, 2, 1, 2,…, 1, 2

Choose a Hadamard matrix $\mathbf{H}_1$ that is $R_1 \times R_1$ with elements $a_{h_i^*,r_1,}$ and choose a Hadamard matrix $\mathbf{H}_2$ that is $R_2 \times R_2$ with elements b_{h_i',r_2}. We note that $R_1 \geq 2\,L$; that is, the dimension of $\mathbf{H}_1$ must be at least as large as the total number of first stage units but it can be larger since we don't need to use all of the rows of $\mathbf{H}_1$. Let $\mathbf{H} = \mathbf{H}_1 \otimes \mathbf{H}_2$ be a $R \times R$ Hadamard matrix where $R = R_1 \cdot R_2$ with elements $c_{h,r} = a_{h^*,r_1,} b_{h_i',r_2}$ and where ⊗ denotes the Kronecker product.

Next, define the replicate factors as

$$F_{h,r} = \begin{cases} 1 + (1-k) a_{h_i^*,r_1,} b_{h_i',r_2} & i_i' = 1 \\ 1 - (1-k) a_{h_i^*,r_1,} b_{h_i',r_2} & i_i' = 2 \end{cases}$$

Define the estimator for each replicate total as

$$\hat{Y}_r = \sum_{h_1} \sum_{i \in s_h} \sum_{c_i=1}^{C_i} \sum_{h_i'=1}^{R_2} \sum_{i_i'=1}^{2} w_{1i} w_{2i_i'} F_{h,r} y_{i_i'}$$

Then the BRR-FAY variance estimator is equivalent to the sum of n / 2 *srswr* variance estimators for $n_{h'} = 2$ in each stratum h and an extra term $\hat{A}(\hat{Y}_i)$, i.e.,

$$\frac{1}{R(1-k)^2} \sum_r \left(\hat{Y}_r - \hat{Y}\right)^2 = \sum_h \sum_{i \in s_h} \frac{\hat{v}_{\mathrm{SRSWR}}\left(\hat{Y}_i\right) + A\left(\hat{Y}_i\right)}{\pi_i^2}$$

where

$$\hat{v}_{\mathrm{SRSWR}}\left(\hat{Y}_i\right) = \sum_{c_i=1}^{C_i} \sum_{h_i'=1}^{R_2} \frac{1}{2(2-1)} \sum_{i_i'=1}^{2} \left(\frac{y_{ci_i'}}{p_{ci_i'}} - \hat{\bar{y}}_{ch_i'}\right)^2$$

and

$$\hat{A}\left(\hat{Y}_i\right) = \sum_{c_i=1}^{C} \sum_{\substack{c_i'=1 \\ c_i' \neq c_i}}^{C} \sum_{h_i'=1}^{R_2} \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2}\right)\left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2}\right)$$

and the estimated total of variance stratum h' is

$$\hat{y}_{ch'} = \frac{1}{2}\sum_{i'=1}^{2}\frac{y_{ch'_i i'_i}}{p_{ch'_i i'_i}}$$

Result 6 of the Appendix, provides the proof.

3.4. Rizzo and Rust (2011)

Another replication method for estimating $v_{\text{SSU}}(\hat{Y})$ is suggested by Rizzo and Rust (2011) and further examined empirically by Kali *et al*. (2011). First, define the estimator of the total of *y* as in (10) where $w_i = \pi_i^{-1}$. Also, define the estimator for each replicate total *r* as

$$\hat{Y}_r = \sum_h \sum_{i \in s_h} \sum_{k \in s_i} w_i w_k F_{h,k} y_k \tag{13}$$

with replicate factors

$$F_{k,r} = \begin{cases} 1 + a_{h,r}\sqrt{\pi_i} & j \in A_i \\ 1 - a_{h,r}\sqrt{\pi_i} & j \in D_i \\ 1 & \text{otherwise} \end{cases}$$

Divide the second-stage sample units into groups A_i and D_i are such that $A_i \cup D_i = s_i$, $A_i \cup D_i = \emptyset$, and the number of sample units in A_i and D_i is $n_i/2$. How the sample units are divided into groups A_i and D_i is not explained in Rust and Rizzo (2011). For our use with *sys*, and we assigned every other sample unit explicitly from the original sort order into the two groups. We considered this way of dividing the sample into groups A_i and D_i since it mimics the *sys* sample design.

Result 7 of the Appendix shows that

$$\hat{v}_{\text{RR}}(\hat{Y}) = \sum_r (\hat{Y}_r - \hat{Y})^2 = \sum_h \sum_{i \in s_h} \frac{1}{\pi_i}(\hat{Y}_{A_i} - \hat{Y}_{D_i})^2$$

so that the average of the previous expression over all replicates is also equivalent to the *srswr* estimator with groups A_i and D_i.

4. Empirical Examples

To compare the variance estimators from a single-stage *sys* sample design, we used the 3rd grade population given in Valliant, Dorfman and Royall (2000). This population has N = 2,427 students and each student had a variable for a region indicator (with four values), a math score, and a science score. The math and science scores had a correlation of 0.67 and their relationship is shown in Figure 1.

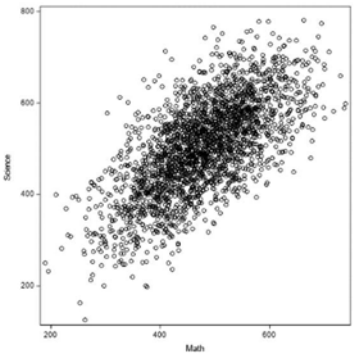

Figure 1. Distribution of Math and Science Scores

Using the region indicator and the science score as sort variables for *sys*, we applied four different sort orders that varied the effectiveness of the sort orders to order the population with respect to the math scores. Figure 2 demonstrates the sort orders where the horizontal axis is the sort order and the vertical axis is the math score. The first sort order in the upper left of Figure 2 represents a random sort. Here we assume that no information is available and therefore no information related to the variable of interest is incorporated into the sample design. The upper right of Figure 2 shows the sort order using the science variable, which results in a sort similar to Figure 1. The bottom left sorts the universe by region and then science within region. If there is a region effect, this sort should be an improvement over the of the sort of science alone. The last sort in the bottom right uses the variable of interest, the math score, as the sort variable. This is the opposite extreme from the random sort where we have information that completely describes the variable math.

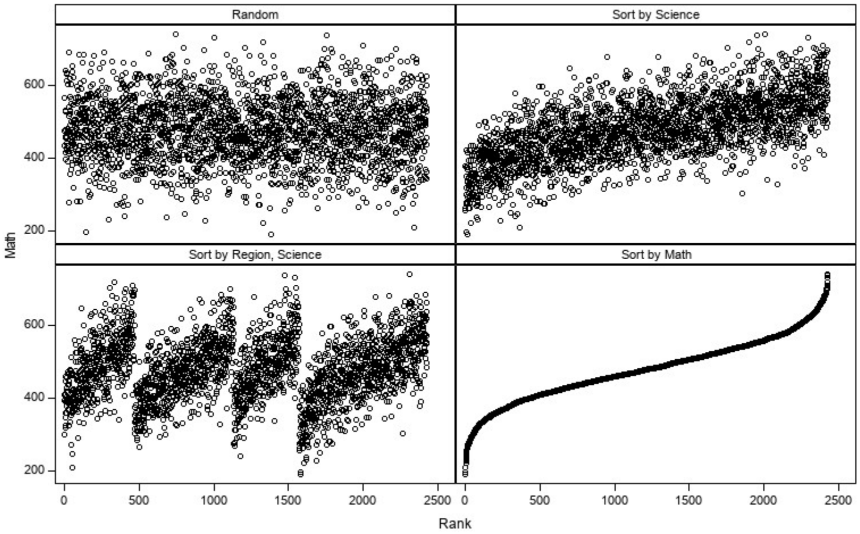

Figure 2. Four Sort Orders of Science Scores for *sys*

For each of the sort orders of Figure 2, we selected n_s = 15,000 *sys* random samples of size n = 100 with equal probability of selection and estimated the total math score, i.e., $\hat{Y} = \sum_{k \in s} w_k y_k$ where the weight is w_k = 2,427 / 100. Figure 3 shows the distribution of the estimated math score totals $\hat{Y}$.

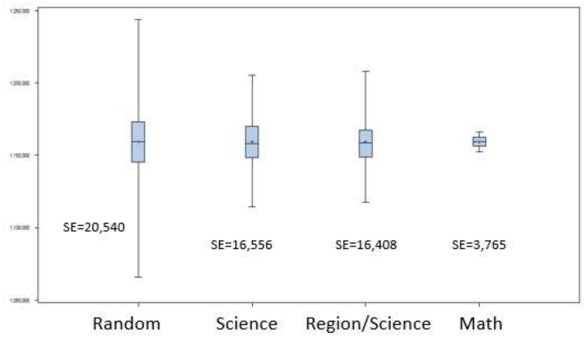

Figure 3. Variances of Estimates of Total Math Scores for the Four Sort Orders of Science Scores

To evaluate the three variance estimators, we produced the measures of Table 2 based on the 15,000 random *sys* samples from each of the four sort orders.

Table 2. Evaluation Measures for Variance Estimators

Measure	Defined as…
Bias Ratio	$Bias\ Ratio\left(\hat{v}(\hat{Y})\right) = \dfrac{E\left(\hat{v}(\hat{Y})\right)}{v(\hat{Y})}$
Mean Squared Error	$MSE\left(\hat{v}(\hat{Y})\right) = E\left(\hat{v}(\hat{Y}) - E\left(\hat{v}(\hat{Y})\right)\right)^2$
Coverage Ratios	The percent of times that the confidence intervals $\hat{Y} \pm z_{\alpha/2}\sqrt{\hat{v}(\hat{Y})}$ included the value of Y

For each of the n_{sim} = 15,000 samples s, we estimated $\hat{Y}_s$ and $\hat{v}(\hat{Y}_s)$. The expectation with respect to the sample design is defined as $E(\hat{Y}) = \sum_{s=1}^{n_{sim}} \hat{Y}_s / n_{sim}$.

Figure 4 presents the bias ratios where the horizontal axis varies by the sort order (0-random, 2-science, 3-region/science, 4-math) and the number of replicates (8, 20, 32, 40). The four bars for each sort order and number of replicate combination represent the four variance estimators that we considered: JK-1, DAGJK, BRR, and SDR. The red horizontal line in Figure 4 identifies where the bias ratio is 1.0 or where there is no bias.

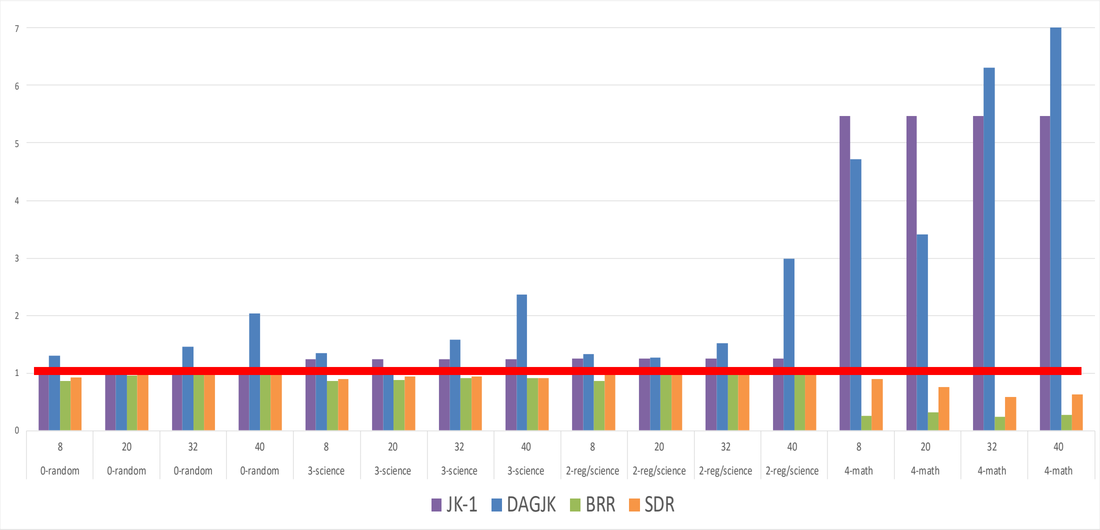

Figure 4. Bias Ratios of Variance Estimators

In Figure 4, we see that both JK-1 and DAGJK overestimate the variance, often to a large degree. Both SDR and BRR underestimate the variance with SDR consistently having bias ratios closer to 1.0 than BRR. Oddly with the DAGJK, the bias increased as the number of replicates increased. With the DAGJK, as the number of replicates increased, the number of sample deleted decreased. We think that as the number of sample units decreased, the sample units left looked less and less like a *sys* sample.

Figure 5 presents the MSEs of the three estimators DAGJK, SDR, and BRR. The horizontal axis varies by the sort order (0-random, 2-science, 3-region/science, 4-math) and the number of replicates (8, 20, 32, 40).

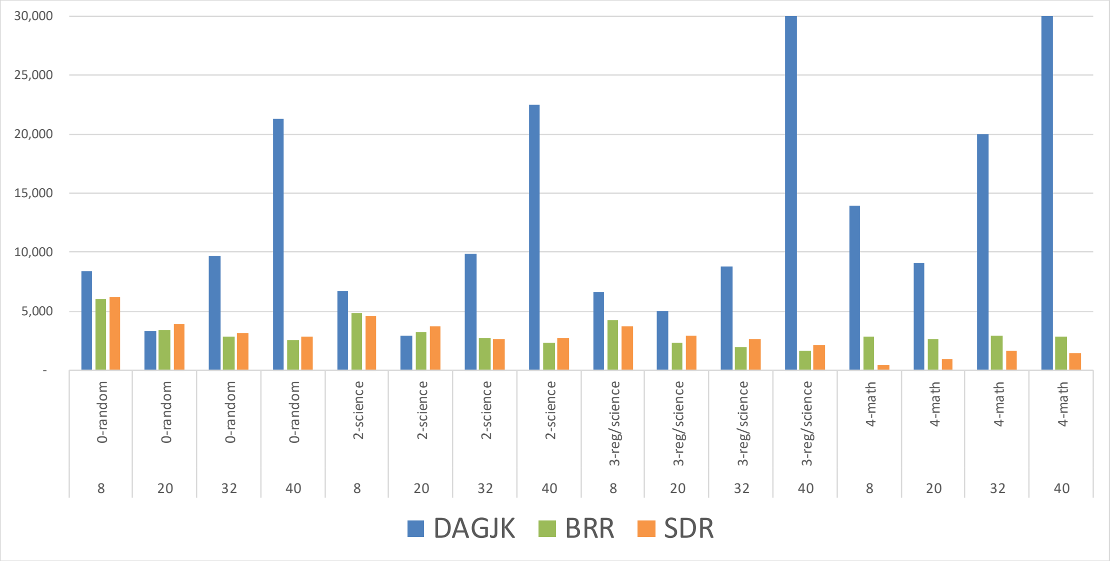

Figure 5. MSEs of Variance Estimators

In Figure 4, we see that the DAGJK has the largest MSEs, BRR the smallest, and SDR are generally just larger than BRR.

Figure 6 presents the coverage ratios of the three estimators DAGJK, SDR, and BRR. The horizontal axis varies by the sort order (0-random, 2-science, 3-region/science, 4-math) and the number of replicates (8, 20, 32, 40). The red horizontal line in Figure 6 identifies the 90% coverage – we expect that 90% of the coverage ratios have 90% coverage.

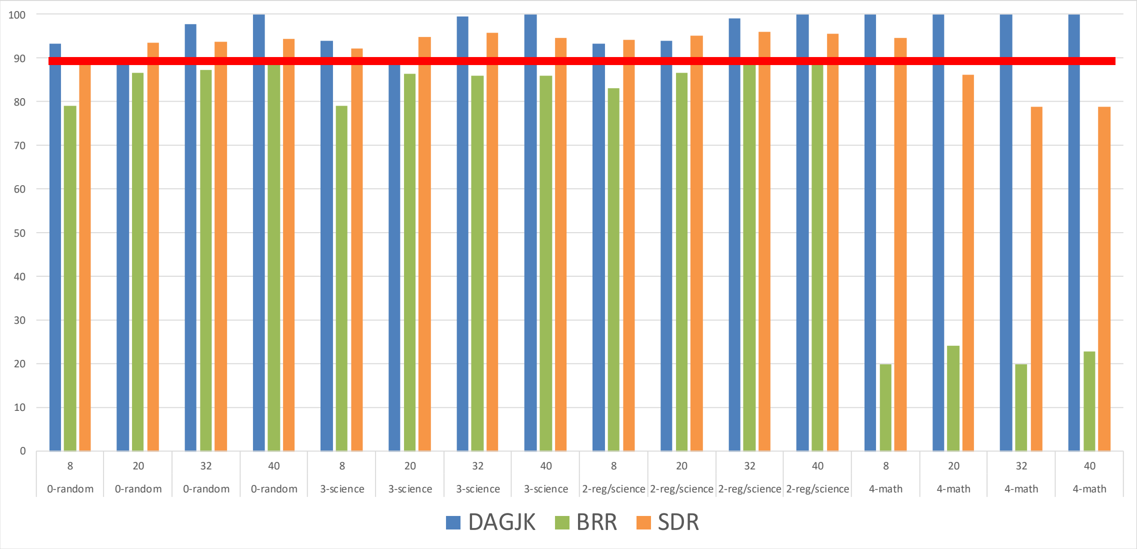

Figure 6. Coverage Intervals

In Figure 6, we see that the DAGJK overestimated the coverage ratios, which was no surprise given the overestimates shown by the bias ratios. BRR regularly underestimates the coverage ratios and SDR regularly overestimates them. With the sort order 2-math, BRR provide a pronounced underestimate which we cannot explain.

5. Conclusions

We have shown that the variance from a *sys* sample design can be estimated with a variety of replication variance estimators that can vary the number of replicates. For a single-stage sample design, our empirical examples provide evidence that SDR and our application of BRR are both reasonable estimators of a *sys* sample design. We think SDR has a slight edge over BRR since the MSEs of the SDR were smaller than BRR and the confidence intervals from SDR were conservative overestimates where the coverage intervals for BRR were under estimates.

We presented replicate variance estimators for $\hat{v}_{\mathrm{SSU}}(\hat{Y})$, but we did not demonstrate them. This is left as future work.

Although it is not discussed in the paper, we think we should consider a slightly different version of our modification of the BRR for *sys* of Results 2 and 6. Instead of only using the *n* / 2 adjacent nonoverlapping pairs of implicit strata, we should have used all (*n* – 1) possible pairs of adjacent implicit strata as done with SDR. The application of this would require an additional term of (*n* / (*n* – 1)) in the replicate factor that would be needed for the same reason it's needed in SDR: to average the *n* – 1 variance strata and then expand that average to all *n* implicit strata. Each sample unit would be used twice in this estimator; this seems odd but it is what is represented in SDR. Results 2 and 6 of the Appendix would require some but not great modifications to show that this would work.

This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. Any views expressed on statistical, methodological, technical, or operational issues are those of the authors and not necessarily those of the U.S. Census Bureau.

References

Ash, S. (2014). "Using successive difference replication for estimating variances" *Survey Methodology*, 40, 47-59.

Bellhouse, D.R. (1988). "Systematic Sampling," excerpt from *Handbook of Statistics, Volume 6*, p. 125-145.

Cochran, W.G. (1977). *Sampling Techniques*, John Wiley and Sons.

Dippo, C.S., Fay, R.E. and Morganstein, D. (1984). "Computing Variances From Complex Samples With Replicate Weights." In *Proceedings of the Survey Research Methods Section*. American Statistical Association: 489–494.

Fay, R.E. and Train, G.F. (1995). "Aspects of Survey and Model-Based Postcensal Estimation of Income and Poverty Characteristics for States and Counties," *Proceedings of the Governments Statistics Section*, American Statistical Association, 154-159.

Gregoire, T.G. and Valentine H.T. (2008). *Sampling Strategies for Natural Resources and the Environment*, Chapman & Hall/CRC.

Iachan, R. (1982). "Systematic sampling: A critical review," *International Statistical Review*, 50, 293-303.

Judkins, D.R. (1990). "Fay's Method for Variance Estimation," *Journal of Official Statistics*, 6, 3, 223-239.

Kali, J., Burke, J., Hicks, L., Rizzo, L., and Rust, K. (2011). "Incorporating a First-Stage Finite Population Correction (FPC) in Variance Estimation for Two-Stage Design in the National Assessment of Educational Progress (NAEP)," Joint Statistical Meetings, *Proceedings of the Section on Survey Research Methods*, 2576-2583.

Kott, P.S. (2001). Delete-a-Group Jackknife," *Journal of Official Statistics*, 17, 4, 521-526.

Madow, W.G. and Madow, L.H. (1944). "On the theory of systematic sampling," *Annuals of Mathematical Statistics*, 15, 1-14.

McCarthy, P.J. (1966). "Pseudo-replication: half-samples," *Review of the International Statistical Institute*, 37, 239-264.

Megill, D.J., Gomez, E.E., Balmaceda, A., and Castillo, M. (1987). “Measuring the efficiency of implicit stratification in multistage sample surveys,” *Proceedings of the Section on Survey Research Methods*, 166-171.

Murthy. M.N. and Rao, T.J. (1988). “Systematic Sampling with Illustrative Examples,” excerpt from *Handbook of Statistics, Vol. 6*, 147-185.

Rizzo, L. and Rust, K. (2011). “Finite Population Correction (FPC) for NAEP Variance Estimation,” Joint Statistical Meetings, *Proceedings of the Section on Survey Research Methods*, 2501-2515.

Särndal, C.-E., Swensson, B., and Wretman, J. (1992). *Model Assisted Survey Sampling*, Spring-Verlag, New York, NY.

Shao, J. and Tu, D. (1995). *The Jackknife and the Bootstrap*, Springer-Verlag.

Vallient, R., Dorfman, A.H. and Royall, R.M. (2000). *Finite Population Sampling and Inference: A Prediction Approach*, John Wiley & Sons.

Wolter, K.M. (1984). “An Investigation of Some Estimators of Variance for Systematic Sampling,” *Journal of the American Statistical Association*, 781-790.

Wolter, K.M. (1985). *Introduction to Variance Estimation*, Springer-Verlag.

Appendix

Result 1. Let the *srswr* estimator of the totals of the variable of interest *y* be defined as

$$\hat{Y} = \frac{1}{n}\sum_{i\in U} X_i \frac{\hat{Y}_i}{p_i}$$

then the following two expressions of its variance estimator are equivalent:

$$\frac{1}{2n^2(n-1)}\sum_{i\in U}\sum_{j\in U} X_i X_j \left(\frac{\hat{Y}_i}{p_i} - \frac{\hat{Y}_j}{p_j}\right)^2 = \frac{1}{n_h(n_h-1)}\sum_{i\in U_h} X_i\left(\frac{\hat{Y}_i}{p_i} - \hat{Y}\right)^2 . \qquad \text{(A1.1)}$$

Proof.

$$\begin{aligned}
\frac{1}{n(n-1)}\sum_{i\in U} X_i\left(\frac{\hat{Y}_i}{p_i} - \hat{Y}\right)^2 &= \frac{1}{n(n-1)}\sum_{i\in U} X_i\left(\left(\frac{\hat{Y}_i}{p_i}\right)^2 - 2\frac{\hat{Y}_i}{p_i}\hat{Y}_h + \hat{Y}^2\right) \\
&= \frac{1}{n(n-1)}\left[\sum_{i\in U} X_i\left(\frac{\hat{Y}_i}{p_i}\right)^2 - 2\hat{Y}\sum_{i\in U} X_i\frac{\hat{Y}_i}{p_i} + \sum_{i\in U} X_i\hat{Y}^2\right] \\
&= \frac{1}{n(n-1)}\left[\sum_{i\in U} X_i\left(\frac{\hat{Y}_i}{p_i}\right)^2 - 2n\hat{Y}\left(\frac{1}{n}\sum_{i\in U} X_i\frac{\hat{Y}_i}{p_i}\right) + \hat{Y}^2\sum_{i\in U} X_i\right] \\
&= \frac{1}{n(n-1)}\left[\sum_{i\in U} X_i\left(\frac{\hat{Y}_i}{p_i}\right)^2 - 2n\hat{Y}^2 - n\hat{Y}^2\right] \\
&= \frac{1}{n(n-1)}\left[\sum_{i\in U} X_i\left(\frac{\hat{Y}_i}{p_i}\right)^2 - n\hat{Y}^2\right] \\
&= \frac{1}{n(n-1)}\left[\sum_{i\in U} X_i\left(\frac{\hat{Y}_i}{p_i}\right)^2 - \frac{1}{n}\sum_{i\in U}\sum_{j\in U} X_i X_j\frac{\hat{Y}_i}{p_i}\frac{\hat{Y}_j}{p_j}\right] \\
&= \frac{1}{2n^2(n-1)}\left[2n\sum_{i\in U_h} X_i\left(\frac{\hat{Y}_i}{p_i}\right)^2 - 2\sum_{i\in U}\sum_{j\in U} X_i X_j\frac{\hat{Y}_i}{p_i}\frac{\hat{Y}_j}{p_j}\right] \\
&= \frac{1}{2n^2(n-1)}\left[2\left(\sum_{j\in U_h} X_j\right)\sum_{i\in U} X_i\left(\frac{\hat{Y}_i}{p_i}\right)^2 - 2\sum_{i\in U}\sum_{j\in U} X_i X_j\frac{\hat{Y}_i}{p_i}\frac{\hat{Y}_j}{p_j}\right] \\
&= \frac{1}{2n^2(n-1)}\left[\sum_{i\in U}\sum_{j\in U} X_i X_j\left(\frac{\hat{Y}_i}{p_i}\right)^2 - 2\sum_{i\in U}\sum_{j\in U} X_i X_j\frac{\hat{Y}_i}{p_i}\frac{\hat{Y}_j}{p_j} + \sum_{i\in U}\sum_{j\in U} X_i X_j\left(\frac{\hat{Y}_j}{p_j}\right)^2\right] \\
&= \frac{1}{2n^2(n-1)}\sum_{i\in U}\sum_{j\in U} X_i X_j\left[\left(\frac{\hat{Y}_i}{p_i}\right)^2 - 2\frac{\hat{Y}_i}{p_i}\frac{\hat{Y}_j}{p_j} + \left(\frac{\hat{Y}_j}{p_j}\right)^2\right] \\
&= \frac{1}{2n^2(n-1)}\sum_{i\in U}\sum_{j\in U} X_i X_j\left(\frac{\hat{Y}_i}{p_i} - \frac{\hat{Y}_j}{p_j}\right)^2
\end{aligned}$$

Note that (A1.1) is stated slightly differently by Tillé (2006; p. 57). In (A1.1), we have explicitly included the random variable X_i that defines the randomization of the *srswr* estimator, where X_i is a multinomial random variable defined for each stratum *h* with parameters p_i and N_h.

Result 2. General Justification of Fay's Method of Balanced Repeated Replication (BRR-FAY). Let the statistic of interest be a total of some variable of interest y_k

$$Y = \sum_{h} \sum_{i \in U_h} \sum_{k \in U_i} y_k$$

Given a two-stage stratified sample design with $n_h = 2$ units per the first-stage stratum h, the estimator of the total of some is

$$\hat{Y} = \sum_{h} \sum_{i \in s_h} \sum_{k \in s_i} w_i w_k y_k = \sum_{h} \left(w_{i=1} \hat{Y}_{i=1} + w_{i=2} \hat{Y}_{i=2} \right)$$

where the first-stage weight is $w_i = \pi_i^{-1}$ and the second-stage weight is w_k. We also define the replicate estimator as

$$\hat{Y}_r = \sum_{h} \sum_{i \in s_h} \sum_{k \in s_i} w_i w_k F_{h,r} y_k$$

with the replicate factors

$$F_{h,r} = \begin{cases} 1 + (1-k) a_{h,r} & i = 1 \\ 1 - (1-k) a_{h,r} & i = 2 \end{cases}$$

the BRR-FAY variance estimator with perturbing factor k and is equivalent to the *srswr* variance estimator for $n_h = 2$, one in each stratum h, i.e.,

$$\hat{v}_{\text{BRR}-\text{FAY}}\left(\hat{Y}\right) = \frac{1}{R(1-k)^2} \sum_{r} \left(\hat{Y}_r - \hat{Y} \right)^2 = \sum_{h} \frac{1}{2(2-1)} \sum_{i \in s_h} \left(\frac{\hat{Y}_i}{p_i} - \hat{Y}_h \right)^2 .$$

Proof. Begin with the difference

$$\begin{aligned} \hat{Y}_r - \hat{Y} &= \sum_{h} \left[\sum_{k \in s_{i=1}} w_{i=1} w_k \left(1 + (1-k) a_{h,r} \right) y_k + \sum_{k \in s_{i=2}} w_{i=2} w_k \left(1 - (1-k) a_{h,r} \right) y_k \right] \\ &\qquad - \sum_{h} \sum_{i \in s_h} \sum_{k \in s_i} w_i w_k y_k \\ &= \sum_{h} \left[\left(\sum_{k \in s_{i=1}} w_{i=1} w_k \left(1 + (1-k) a_{h,r} \right) y_k + \sum_{k \in s_{i=1}} w_{i=2} w_k \left(1 - (1-k) a_{h,r} \right) y_k \right) \right. \\ &\qquad \left. - \left(\sum_{k \in s_{i=1}} w_{i=1} w_k y_k + \sum_{k \in s_{i=2}} w_{i=2} w_k y_k \right) \right] \\ &= \sum_{h} \left[\left(\sum_{k \in s_{i=1}} w_{i=1} w_k \left((1-k) a_{h,r} \right) y_k + \sum_{k \in s_{i=2}} w_{i=2} w_k \left(-(1-k) a_{h,r} \right) y_k \right) \right] \end{aligned}$$

Appendix

$$= (1-k)\sum_h \left[\left(\sum_{k\in s_{i=1}} w_{i=1}w_k a_{h,r} y_k - \sum_{k\in s_{i=2}} w_{i=2}w_k a_{h,r} y_k\right)\right]$$

Define the estimator of the first − stage unit i of stratum h as $\hat{Y}_{h,i} = \sum_{k\in s_i} w_k\, y_k$ and note that when $w_i = \pi_i^{-1}$, then

$$\hat{Y}_r - \hat{Y} = (1-k)\sum_h a_{h,r}\left(\frac{\hat{Y}_{h,i=1}}{\pi_{h,i=1}} - \frac{\hat{Y}_{h,i=2}}{\pi_{h,i=2}}\right)$$

Next, square the difference

$$\left(\hat{Y}_r - \hat{Y}\right)^2 = \left((1-k)\sum_h a_{h,r}\left(\frac{\hat{Y}_{h,i=1}}{\pi_{h,i=1}} - \frac{\hat{Y}_{h,i=2}}{\pi_{h,i=2}}\right)\right)^2$$

$$= (1-k)^2\left[\sum_h a_{r,h}^2\left(\frac{\hat{Y}_{h,i=1}}{\pi_{h,i=1}} - \frac{\hat{Y}_{h,i=2}}{\pi_{h,i=2}}\right)^2 + \sum_h\sum_{h'\neq h} a_{h,r}a_{h,r'}\left(\frac{\hat{Y}_{h,i=1}}{\pi_{h,i=1}} - \frac{\hat{Y}_{h,i=2}}{\pi_{h,i=2}}\right)\left(\frac{\hat{Y}_{h',i=1}}{\pi_{h',i=1}} - \frac{\hat{Y}_{h',i=2}}{\pi_{h',i=2}}\right)\right]$$

Further

$$\frac{1}{R(1-k)^2}\sum_r\left(\hat{Y}_r - \hat{Y}\right)^2$$

$$= \frac{1}{R}\sum_r \frac{(1-k)^2}{R(1-k)^2}\left[\sum_h a_{r,h}^2\left(\frac{\hat{Y}_{h,i=1}}{\pi_{h,i=1}} - \frac{\hat{Y}_{h,i=2}}{\pi_{h,i=2}}\right)^2 + \sum_h\sum_{h'\neq h} a_{h,r}a_{h',r}\left(\frac{\hat{Y}_{h,i=1}}{\pi_{h,i=1}} - \frac{\hat{Y}_{h,i=2}}{\pi_{h,i=2}}\right)\left(\frac{\hat{Y}_{h',i=1}}{\pi_{h',i=1}} - \frac{\hat{Y}_{h',i=2}}{\pi_{h',i=2}}\right)\right]$$

$$= \frac{1}{R}\sum_h\left(\sum_r a_{r,h}^2\right)\left(\frac{\hat{Y}_{h,i=1}}{\pi_{h,i=1}} - \frac{\hat{Y}_{h,i=2}}{\pi_{h,i=2}}\right)^2 + \frac{1}{R}\sum_h\sum_{h'\neq h}\left(\sum_r a_{h,r}a_{h',r}\right)\left(\frac{\hat{Y}_{h,i=1}}{\pi_{h,i=1}} - \frac{\hat{Y}_{h,i=2}}{\pi_{h,i=2}}\right)\left(\frac{\hat{Y}_{h',i=1}}{\pi_{h',i=1}} - \frac{\hat{Y}_{h',i=2}}{\pi_{h',i=2}}\right)$$

Given that the values of $a_{h,r}$ come from a Hadamard matrix, which has rows and columns that are orthogonal, we know that $\sum_r a_{h,r}^2 = R$ and $\sum_r a_{h,r}a_{h',r} = 0$, so we can simplify the previous line as

$$\sum_h\left(\frac{\hat{Y}_{h,i=1}}{\pi_{h,i=1}} - \frac{\hat{Y}_{h,i=2}}{\pi_{h,i=2}}\right)^2. \qquad \text{(A2.1)}$$

Now (A2.1) is equivalent to the variance estimator of *srswr* for each stratum *h*. This result is stated in Wolter (1985; p 123), but we provide a fuller explanation. With our specific case of $n_h = 2$, we have $\pi_i = 2p_{h,i}$ and (A2.1) becomes

$$\sum_{h}\left(\frac{\hat{Y}_{h,i=1}}{\pi_{h,i=1}}-\frac{\hat{Y}_{h,i=2}}{\pi_{h,i=2}}\right)^{2} = \frac{1}{2^{2}(2-1)}\sum_{h}\left(\frac{\hat{Y}_{h,i=1}}{p_{h,i=1}}-\frac{\hat{Y}_{h,i=2}}{p_{h,i=2}}\right)^{2}$$
$$=\frac{1}{2}\,\frac{1}{2^{2}(2-1)}\sum_{h}\left[\left(\frac{\hat{Y}_{h,i=1}}{p_{h,i=1}}-\frac{\hat{Y}_{h,i=2}}{p_{h,i=2}}\right)^{2}+\left(\frac{\hat{Y}_{h,i=2}}{p_{h,i=2}}-\frac{\hat{Y}_{h,i=1}}{p_{h,i=1}}\right)^{2}\right] \qquad \text{(A2.2)}$$

We can see that (A2.1) is the same as the left side of (A2.2) with $n_h = 2$ so that

$$\frac{1}{R(1-k)^{2}}\sum_{r}\left(\hat{Y}_{r}-\hat{Y}\right)^{2}=\sum_{h}\frac{1}{2(2-1)}\sum_{i\in s_{h}}\left(\frac{\hat{Y}_{hi}}{p_{hi}}-\hat{Y}_{h}\right)^{2}$$

and this shows that BRR-FAY variance estimator is equivalent to the sum of several *srswr* variance estimators for $n_h = 2$, with one for each first-stage strata h.

Result 3. Application of Balanced Repeated Replication for estimating the variance of a single-stage *sys* sample design (BRR-SYS). Given a single-stage stratified sample design with $n > 2$ units, define the total of the variable of interest y as

$$\hat{Y}=\sum_{k\in s}w_{k}y_{k}$$

and note that we can rewrite the previous expression as

$$\hat{Y}=\sum_{c\,=1}^{C}\sum_{h'=1}^{R}\sum_{i'=1}^{2}y_{k}$$

Choose a $R \times R$ Hadamard matrix **H** with elements $a_{h',r}$. Then the ordered sample units k are assigned to the variance strata h' and half sample i' as described in Table A3.1.

Table A3.1: Assignment of Variance Strata and Half Sample

Sample unit k	1, 2, 3, 4, 5, 6,…
Cycle c	1, 1, 1, 1, 1, 1,…, 1, 1, 2, 2, 2, 2,…
Variance strata h'	1, 1, 2, 2, 3, 3,…, R, R, 1, 1, 2, 2,…
Half sample i'	1, 2, 1, 2, 1, 2,…, 1, 2, 1, 2, 1, 2,…

The cycles c are needed if $n / 2 > R$ – we repeat the assignment of the of the variance strata h' and half sample i' to the same Hadamard matrix.

Next, define the replicate factors as

$$F_{h_i',r_2}=\begin{cases}1+(1-k)a_{h',r} & i'=1\\ 1-(1-k)a_{h',r} & i'=2\end{cases}$$

Appendix

Then the estimator for each replicate total *r* is

$$\hat{Y}_r = \sum_{c=1}^{C} \sum_{h'=1}^{R} \sum_{i'=1}^{2} w_{r,k} y_k$$

And the BRR-FAY variance estimator is equivalent to the sum of *n* / 2 *srswr* variance estimators with $n_{h'}$ = 2 sample units within each pseudo stratum h' and an extra term $\hat{A}(\hat{Y})$, i.e.,

$$\frac{1}{R(1-k)^2} \sum_r \left(\hat{Y}_r - \hat{Y}\right)^2 = \hat{v}_{\text{SRSWR}}(\hat{Y}) + \hat{A}(\hat{Y}) \qquad \text{(A3.1)}$$

and we refer to the variance estimator of (A3.1) as $\hat{v}_{\text{BRR-}} \quad (\hat{Y})$, where

$$\hat{v}_{\text{SRSWR}}(\hat{Y}) = \sum_{c=1}^{C} \sum_{h'=1}^{R} \frac{1}{2(2-1)} \sum_{i'=1}^{2} \left(\frac{y_{h'_i i'_i}}{p_{h'_i i'_i}} - \hat{\hat{y}}_{h'} \right)^2$$

and

$$\hat{A}(\hat{Y}) = \sum_{c=1}^{C} \sum_{\substack{c'=1 \\ c' \neq c}}^{C} \sum_{h'=1}^{R} \left(w_{i'=1} y_{i'=1} - w_{i'=2} y_{i'=2} \right) \left(w_{i'=1} y_{i'=1} - w_{i'=2} y_{i'=2} \right)$$

or with all of the subscripts

$$\hat{A}(\hat{Y}) = \sum_{c=1}^{C} \sum_{\substack{c'=1 \\ c' \neq c}}^{C} \sum_{h'=1}^{R} \left(w_{c\,h'i'=1} y_{c\,h'i'=1} - w_{c\,h'i'=2} y_{c\,h'i'=2} \right) \left(w_{c'h'i'=1} y_{c'h'i'=1} - w_{c'h'i'=2} y_{c'h'i'=2} \right)$$

and $w_k = \pi_k^{-1}$, $p_k = \pi_k \,/n$ and the estimator of the mean of *y* for each variance stratum is defined as

$$\hat{\hat{y}}_{ch'} = \frac{1}{2} \sum_{i'=1}^{2} \frac{y_{ch'i'}}{p_{ch'i'}}$$

Proof. Begin with the difference

$$\hat{Y}_r - \hat{Y} = \sum_{c=1}^{C} \sum_{h'=1}^{R} \sum_{i'=1}^{2} w_{r,k} y_k - \sum_{c=1}^{C} \sum_{h'=1}^{R} \sum_{i'=1}^{2} w_k y_k$$

We note our abuse of notation: the sample weight w_k for unit *k* and $w_{i'}$ for half sample *i*' are both shorthand for $w_{c\,h'i'k}$ or the sample weight for unit *k* of half sample i' of variance stratum h' of cycle *c*. We choose to include the least amount of subscripting by keeping what is essential at that point of the result, reduce the clutter of the subscripting, and assume the rest of the subscripting is understood and known. We do this similarly with $y_{c\,h'i'k}$.

$$= \sum_{c'=1}^{C} \sum_{h'=1}^{R} \left[w_{i'=1}\left(1 + (1-k)a_{h',r}\right)y_{i'=1} + w_{i'=2}\left(1 - (1-k)a_{h',r}\right)y_{i'=2}\right] - \sum_{c_i'=1}^{C} \sum_{h_i'=1}^{R} \left[w_{i'=1}y_{i'=1} + w_{i'=2}y_{i'=2}\right]$$

$$= \sum_{c'=1}^{C} \sum_{h'=1}^{R} \left[w_{i'=1}\left((1-k)a_{h',r}\right)y_{i'=1} - w_{i'=2}\left((1-k)a_{h',r}\right)y_{i'=2}\right]$$

$$= (1-k) \sum_{c=1}^{C} \sum_{h'=1}^{R} a_{h',r}\left(w_{i'=1}y_{ch'i'=1} - w_{i'=2}y_{i'=2}\right)$$

Next, square the difference

$$\left(\hat{Y}_r - \hat{Y}\right)^2 = \left((1-k) \sum_{c=1}^{C} \sum_{h'=1}^{R} a_{h',r}\left(w_{i'=1}y_{i'=1} - w_{i'=2}y_{i'=2}\right)\right)^2$$

$$= (1-k)^2 \left[\begin{array}{l} \displaystyle\sum_{c=1}^{C} \sum_{h'=1}^{R} a_{h',r}^2\left(w_{i'=1}y_{i'=1} - w_{i'=2}y_{i'=2}\right)^2 \\ \displaystyle + \sum_{c=1}^{C} \sum_{h'=1}^{R} \sum_{\substack{h''=1 \\ h''\neq h'}}^{R} a_{h_i',r}a_{h_i'',r}\left(w_{i'=1}y_{i'=1} - w_{i'=2}y_{i'=2}\right)\left(w_{i'=1}y_{i'=1} - w_{i'=2}y_{i'=2}\right) \\ \displaystyle + \sum_{c=1}^{C} \sum_{\substack{c'=1 \\ c'\neq c}}^{C} \sum_{h'=1}^{R} a_{h',r}^2\left(w_{i'=1}y_{i'=1} - w_{i'=2}y_{i'=2}\right)\left(w_{i'=1}y_{i'=1} - w_{i'=2}y_{i'=2}\right) \\ \displaystyle + \sum_{c=1}^{C} \sum_{\substack{c'=1 \\ c'\neq c}}^{C} \sum_{h'=1}^{R} \sum_{\substack{h''=1 \\ h''\neq h'}}^{R} a_{h',r}a_{h'',r}\left(w_{i'=1}y_{i'=1} - w_{i'=2}y_{i'=2}\right)\left(w_{i'=1}y_{i'=1} - w_{i'=2}y_{i'=2}\right) \end{array}\right]$$

Further, we sum over the replicates and divide by $R(1-k)^2$

$$\frac{1}{R(1-k)^2} \sum_{r} \left(\hat{Y}_r - \hat{Y}\right)^2$$

$$= \frac{1}{R}\sum_{r} \frac{(1-k)^2}{(1-k)^2} \begin{bmatrix} \sum_{c=1}^{C}\sum_{h'=1}^{R} a_{h',r}^2 \left(w_{i'=1}y_{i'=1} - w_{i'=2}y_{i'=2}\right)^2 \\ + \sum_{c=1}^{C}\sum_{h'=1}^{R}\sum_{\substack{h''=1 \\ h''\neq h'}}^{R} a_{h_i',r}a_{h_i'',r}\left(w_{i'=1}y_{i'=1} - w_{i'=2}y_{i'=2}\right)\left(w_{i'=1}y_{i'=1} - w_{i'=2}y_{i'=2}\right) \\ + \sum_{c=1}^{C}\sum_{\substack{c'=1 \\ c'\neq c}}^{C}\sum_{h'=1}^{R} a_{h',r}^2 \left(w_{i'=1}y_{i'=1} - w_{i'=2}y_{i'=2}\right)\left(w_{i'=1}y_{i'=1} - w_{i'=2}y_{i'=2}\right) \\ + \sum_{c=1}^{C}\sum_{\substack{c'=1 \\ c'\neq c}}^{C}\sum_{h'=1}^{R}\sum_{\substack{h''=1 \\ h''\neq h'}}^{R} a_{h',r}a_{h'',r}\left(w_{i'=1}y_{i'=1} - w_{i'=2}y_{i'=2}\right)\left(w_{i'=1}y_{i'=1} - w_{i'=2}y_{i'=2}\right) \end{bmatrix}$$

$$= \frac{1}{R}\sum_{c=1}^{C}\sum_{h'=1}^{R}\left(\sum_{r} a_{h',r}^2\right)\left(w_{i'=1}y_{i'=1} - w_{i'=2}y_{i'=2}\right)^2$$

$$+ \frac{1}{R}\sum_{c=1}^{C}\sum_{h'=1}^{R}\sum_{\substack{h''=1 \\ h''\neq h'}}^{R}\left(\sum_{r} a_{h',r}a_{h'',r}\right)\left(w_{i'=1}y_{i'=1} - w_{i'=2}y_{i'=2}\right)\left(w_{i'=1}y_{i'=1} - w_{i'=2}y_{i'=2}\right)$$

$$+ \frac{1}{R}\sum_{c=1}^{C}\sum_{\substack{c'=1 \\ c'\neq c}}^{C}\sum_{h'=1}^{R}\left(\sum_{r} a_{h',r}^2\right)\left(w_{i'=1}y_{i'=1} - w_{i'=2}y_{i'=2}\right)\left(w_{i'=1}y_{i'=1} - w_{i'=2}y_{i'=2}\right)$$

$$+ \frac{1}{R}\sum_{c=1}^{C}\sum_{\substack{c'=1 \\ c'\neq c}}^{C}\sum_{h'=1}^{R}\sum_{\substack{h''=1 \\ h''\neq h'}}^{R}\left(\sum_{r} a_{h_i',r}a_{h_i'',r}\right)\left(w_{i'=1}y_{i'=1} - w_{i'=2}y_{i'=2}\right)\left(w_{i'=1}y_{i'=1} - w_{i'=2}y_{i'=2}\right)$$

Given that the values of $a_{h',r}$ come from a Hadamard matrix, which has rows and columns that are orthogonal, we know that $\sum_{r} a_{h',r}^2 = R$ and $\sum_{r} a_{h',r}a_{h'',r} = 0$, so we can simplify the previous line as

$$\sum_{c=1}^{C}\sum_{h'=1}^{R}\left(w_{i'=1}y_{i'=1} - w_{i'=2}y_{i'=2}\right)^2 + \sum_{c=1}^{C}\sum_{\substack{c'=1 \\ c'\neq c}}^{C}\sum_{h'=1}^{R}\left(w_{i'=1}y_{i'=1} - w_{i'=2}y_{i'=2}\right)\left(w_{i'=1}y_{i'=1} - w_{i'=2}y_{i'=2}\right)$$

or with all of the subscripts

$$\sum_{c=1}^{C}\sum_{h'=1}^{R}\left(w_{c\,h'i'=1}y_{c\,h'i'=1} - w_{c\,h'i'=2}y_{c\,h'i'=2}\right)^2$$
$$+ \sum_{c=1}^{C}\sum_{\substack{c'=1 \\ c'\neq c}}^{C}\sum_{h'=1}^{R}\left(w_{c\,h'i'=1}y_{c\,h'i'=1} - w_{c\,h'i'=2}y_{c\,h'i'=2}\right)\left(w_{c'h'i'=1}y_{c'h'i'=1} - w_{c'h'i'=2}y_{c'h'i'=2}\right) \text{ (A3.2)}$$

Now the second expression of (A3.2) is $A(\hat{Y})$ and the first expression is equivalent to the sum of *n* / 2 of *srswr* variance estimators for each variance stratum h'. With our specific case of *n* we have $p_{i'} = \pi_{i'}/2$ and the second expression of (A3.2) becomes

$$\sum_{c=1}^{C}\sum_{h'=1}^{R}\left(w_{i'=1}y_{i'=1} - w_{i'=2}y_{i'=2}\right)^2 = \sum_{c'=1}^{C}\sum_{h'=1}^{R}\left(\frac{y_{i'=1}}{\pi_{i'=1}} - \frac{y_{i'=2}}{\pi_{i'=2}}\right)^2$$

$$= \frac{1}{2^2(2-1)}\sum_{c=1}^{C}\sum_{h'=1}^{R}\left(\frac{y_{i'=1}}{p_{i'=1}} - \frac{y_{i'=2}}{p_{i'=2}}\right)^2$$

$$= \frac{1}{2}\,\frac{1}{2^2(2-1)}\sum_{c=1}^{C}\sum_{h'=1}^{R}\left[\left(\frac{y_{i'=1}}{p_{i'=1}} - \frac{y_{i'=2}}{p_{i'=2}}\right)^2 + \left(\frac{y_{i'=1}}{p_{i'=1}} - \frac{y_{i'=2}}{p_{i'=2}}\right)^2\right] \qquad \text{(A3.3)}$$

We can see that (A3.3) is the same as the left side of (A1.1) in each of the $R \cdot C$ variance stratum h' so that

$$\frac{1}{R(1-k)^2}\sum_{r}\left(\hat{Y}_r - \hat{Y}\right)^2 = \sum_{c=1}^{C}\sum_{h'=1}^{R}\frac{1}{2(2-1)}\sum_{i'=1}^{2}\left(\frac{y_{ch'_i i'_i}}{p_{ch'_i i'_i}} - \hat{\hat{y}}_{ch'}\right)^2 + A(\hat{Y})$$

and this shows that our application of the BRR-FAY variance estimator is equivalent to the *srswr* variance estimator in each R · C variance stratum h'_i.

Note that when the number of cycles *C* = 1, $A(\hat{Y}) = 0$. When *C* > 1, $A(\hat{Y})$ is the sum of an equal number of positive and negative terms which do not exactly cancel each other; however, do approximately cancel each other.

Result 4. Application of Successive Difference Replication for estimating $v_{\text{SSU}}(\hat{Y})$ (SDR-SSU). With replicate factors for sample unit *k* and replicate *r*

$$F_{k,r} = 1 + \sqrt{1 - f_i}\left(2^{-\frac{3}{2}}h_{a_k,r} - 2^{-\frac{3}{2}}h_{b_k,r}\right)$$

then

$$\frac{4}{R}\sum_{r=1}^{R}\left(\hat{Y}_r - \hat{Y}\right)^2 \tag{A4.1}$$

is a replication estimator of $v_{\text{SSU}}(\hat{Y})$ where $\hat{v}(\hat{Y}_i)$ is estimated with the SDR estimator.

Proof. We begin by restating the replicate factors in matrix notation as

$$\mathbf{F} = \mathbf{1}_m\mathbf{1}'_k \; + \boldsymbol{\gamma}' \,\#\left(2^{-\frac{3}{2}}\mathbf{I}_m - 2^{-\frac{3}{2}}\mathbf{S}\right)\mathbf{M}$$

The SDR estimator given in (A1) can be expressed as

$$\begin{aligned}
&\frac{4}{R}\left(\breve{\breve{\mathbf{y}}}'\left(\mathbf{1}_m\mathbf{1}'_k + \boldsymbol{\gamma}' \,\#\left(2^{-\frac{3}{2}}\mathbf{I}_m - 2^{-\frac{3}{2}}\mathbf{S}\right)\mathbf{M}\right) - \mathbf{y}'\mathbf{1}_m\mathbf{1}'_k\right)\left(\breve{\breve{\mathbf{y}}}'\left(\mathbf{1}_m\mathbf{1}'_k + \boldsymbol{\gamma}' \,\#\left(2^{-\frac{3}{2}}\mathbf{I}_m - 2^{-\frac{3}{2}}\mathbf{S}\right)\mathbf{M}\right) - \mathbf{y}'\mathbf{1}_m\mathbf{1}'_k\right)'\\
&= \frac{4}{R}\left(\breve{\breve{\mathbf{y}}}'\left(\boldsymbol{\gamma}' \,\#\left(2^{-\frac{3}{2}}\mathbf{I}_m - 2^{-\frac{3}{2}}\mathbf{S}\right)\mathbf{M}\right)\right)\left(\breve{\breve{\mathbf{y}}}'\left(\boldsymbol{\gamma}' \,\#\left(2^{-\frac{3}{2}}\mathbf{I}_m - 2^{-\frac{3}{2}}\mathbf{S}\right)\mathbf{M}\right)\right)'\\
&= \frac{4}{R}(\breve{\breve{\mathbf{y}}} \,\#\, \boldsymbol{\gamma})'\left(2^{-\frac{3}{2}}\mathbf{I}_m - 2^{-\frac{3}{2}}\mathbf{S}\right)\mathbf{M}\mathbf{M}'\left(2^{-\frac{3}{2}}\mathbf{I}_m - 2^{-\frac{3}{2}}\mathbf{S}\right)'(\breve{\breve{\mathbf{y}}} \,\#\, \boldsymbol{\gamma})
\end{aligned}$$

Because $\{rows\ of\ \mathbf{M}\} \subseteq \{rows\ of\ \mathbf{H}\}$, it can be shown that $\mathbf{M}\mathbf{M}' = k\mathbf{I}_m$. With this result, the variance becomes

$$\begin{aligned}
&\frac{4}{R}(\boldsymbol{\gamma} \,\#\breve{\breve{\mathbf{y}}})'(\mathbf{I}_m - \mathbf{S})(R\mathbf{I}_m)(\mathbf{I}_m - \mathbf{S})'(\boldsymbol{\gamma} \,\#\breve{\breve{\mathbf{y}}})\\
&= \frac{1}{2}(\boldsymbol{\gamma} \,\#\breve{\breve{\mathbf{y}}})'(\mathbf{I}_m - \mathbf{S})(\mathbf{I}_m - \mathbf{S})'(\boldsymbol{\gamma} \,\#\breve{\breve{\mathbf{y}}})\\
&= \frac{1}{2}(\boldsymbol{\gamma} \,\#\breve{\breve{\mathbf{y}}})'(\mathbf{I}_m - \mathbf{S} - \mathbf{S}')(\boldsymbol{\gamma} \,\#\breve{\breve{\mathbf{y}}})
\end{aligned}$$

The last line follows from the lemma in Ash (2014) and has a constant value for any choice of **H.** By noting the block diagonal structure of **S** and given that each connected loop was formed within each first-stage unit *i* of stratum *h*, we can write the estimator as

$$\sum_{h}\sum_{i\in s_h}\frac{1}{2}(1 - f_i)\sum_{c=1}^{C_i}\breve{\breve{\mathbf{y}}}'_c(2\mathbf{I}_m - \mathbf{S}_c - \mathbf{S}'_c)\breve{\breve{\mathbf{y}}}_c$$

where the within second-stage variance is estimated with the SDR variance estimator

$$\hat{v}_{\text{SDR}}(\hat{Y}_i) = (1 - f_i)\frac{1}{2}\sum_{c=1}^{C_i}\breve{\breve{\mathbf{y}}}'_c(2\mathbf{I}_m - \mathbf{S}_c - \mathbf{S}'_c)\breve{\breve{\mathbf{y}}}_c$$

$$= (1 - f_i)\frac{1}{2}\frac{1}{w_i^2}\sum_{c=1}^{C_i}\breve{\mathbf{y}}_c'(2\mathbf{I}_m - \mathbf{S}_c - \mathbf{S}_c')\breve{\mathbf{y}}_c$$

where each of the $c = 1$ to C_i connected loops are of the form (22) and

$$\hat{v}_{\text{SSU}}(\hat{Y}) = \sum_h \sum_{i \in s_h} \frac{\hat{v}_{\text{SDR}}(\hat{Y}_i)}{\pi_i^2}$$

Result 5. Let $\mathbf{H}_1$ be a $R_1 \times R_1$ Hadamard matrix with elements $a_{h_1 r_1}$ and Let $\mathbf{H}_2$ be a $R_2 \times R_2$ Hadamard matrix with elements $b_{h_2 r_2}$. Let $\mathbf{H} = \mathbf{H}_1 \otimes \mathbf{H}_2$ be a $R \times R$ Hadamard matrix where $R = R_1 \cdot R_2$ with elements c_{hr}. In our notation any element with a prime, i.e., h', denotes a unit that is different than an element without a prime or simply $h \neq h'$ and $\otimes$ denotes the Kronecker product. Then we know the following are true because $\mathbf{H}_1$, $\mathbf{H}_1$, and $\mathbf{H}$ are all Hadamard matrices.

$$\sum_{i=1}^{R_1} a_{h_1,r_1} a_{h_1' r_1} = 0 \qquad \text{and} \qquad \sum_{i=1}^{R_1} a_{h_1,r_1}^2 = 1$$

$$\sum_{r_2=1}^{R_2} b_{h_2,r_2} b_{h_2' r_2} = 0 \qquad \text{and} \qquad \sum_{r_2=1}^{R_2} b_{h_2,r_2}^2 = 1$$

$$\sum_{r=1}^{R} c_{h,r} c_{h\ ,r'} = 0 \qquad \text{and} \qquad \sum_{r=1}^{R} c_{h,r}^2 = 1$$

Further, we know that

$$\sum_{r=1}^{R} c_{h,r}^2 = 1 \quad \text{iff} \quad \sum_{r_1=1}^{R_1}\sum_{r_2=1}^{R_2} a_{h_1,r_1}^2 b_{h_2,r_2}^2 = 1$$

and

$$\sum_{r=1}^{R} c_{h,r} c_{h',r} = 0 \tag{A5.1}$$

iff one of the following three expressions is true:

$$\sum_{r_1=1}^{R_1}\sum_{r_2=1}^{R_2} a_{h_1,r_1}^2 b_{r_2,r_2} b_{h_2',r_2} = 0 \tag{A5.2}$$

$$\sum_{r_1=1}^{R_1}\sum_{r_2=1}^{R_2} a_{h_1,r_1} a_{h_1',r_1} b_{h_2,r_2}^2 = 0 \tag{A5.3}$$

$$\sum_{r_1=1}^{R_1}\sum_{r_2=1}^{R_2} a_{h_1,r_1} a_{h_1',r_1} b_{h_2,r_2} b_{h_2',r_2} = 0 \qquad \text{(A5.4)}$$

We can relate the subscripts of $c_{h,r}$ to the subscripts of a_{h_1,r_1} and b_{h_2,r_2} as

r_1 = (r div R_1) + 1 and r_2 = r mod R_1

h_1 = (h div R_1) + 1 and h_2 = h mod R_1

The div function is integer portion of simple division and the mod function is the integer remainder from division. For example, 7/2 = 3½ so 7 div 2 = 3 and 7 mod 2 = 1.

Result 6. Application of Balanced Repeated Replication for estimating $v_{\text{SSU}}(\hat{Y})$ (BRR-SSU). Given a two-stage stratified sample design with n_h = 2 first-stage units per the first-stage stratum h, and n_i second-stage sample units within each second-stage unit i, define the estimator for the replicate total of the variable of interest y for each replicate r as

$$\hat{Y} = \sum_{h}\sum_{i\in s_h}\sum_{k\in s_i} w_{1i} w_{2k} y_k \qquad \text{(6.1)}$$

where the first-stage weight is $w_i = \pi_i^{-1}$ and the second-stage weight is w_k. We can rewrite (A6.1) as

$$\hat{Y} = \sum_{h}\sum_{i\in s_h}\sum_{c_i=1}^{C_i}\sum_{h_i'=1}^{R_2}\sum_{i_i'=1}^{2} w_{1i} w_{2i_i'} y_{i_i'}$$

With the first-stage, define the values of h^* as shown in Table A6.1.

Table A6.1: Assignment of h^*

h^*	1, 2, 3, 4, 5, 6,…, 2L-1, 2L
First-stage strata h	1, 1, 2, 2, 3, 3,…, L, L
First-stage unit i	1, 2, 1, 2, 1, 2,…, 1, 2

For our result, we have assumed that n_h = 2 for all of the first-stage strata h, but this result works with any number of first-stage units n_h: we only need to assign a unique h^* to each first-stage unit.

Within the first-stage stratum h and first-stage unit i, define the variance strata h_i' and half sample i_i' for the ordered sample units k as shown in Table A6.2.

Table A6.2: Assignment of Variance Strata and Half Sample within each Second-Stage Unit i

Sample unit k	1, 2, 3, 4, 5, 6,…
Cycle c_i	1, 1, 1, 1, 1, 1,…, 1, 1, 2, 2, 2, 2,…, C_i, C_i
Variance strata h_i'	1, 1, 2, 2, 3, 3,…, R_2, R_2, 1, 1, 2, 2,…, 2, 2
Half sample i_i'	1, 2, 1, 2, 1, 2,…, 1, 2, 1, 2, 1, 2,…, 1, 2

Appendix

Choose a Hadamard matrix **H**$_1$ that is $R_1 \times R_1$ with elements $a_{h_i^*,r_1,}$ and choose a Hadamard matrix **H**$_2$ that is $R_2 \times R_2$ with elements b_{h_i',r_2}. We note that $R_1 \geq 2\,L$ – the dimension of **H**$_1$ must be at least as large as the total number of first stage units but it can be larger since we don't need to use all of the rows of **H**$_1$. Let $\mathbf{H} = \mathbf{H}_1 \otimes \mathbf{H}_2$ be a $R \times R$ Hadamard matrix where $R = R_1 \cdot R_2$ with elements $c_{h,r} = a_{h^*,r_1,} b_{h_i',r_2}$ and where ⊗ denotes the Kronecker product.

Next, define the replicate factors as

$$F_{h,r} = \begin{cases} 1 + (1-k) a_{h_i^*,r_1,} b_{h_i',r_2} & i_i' = 1 \\ 1 - (1-k) a_{h_i^*,r_1,} b_{h_i',r_2} & i_i' = 2 \end{cases}$$

Define the estimator for each replicate total as

$$\hat{Y}_r = \sum_{h_1} \sum_{i \in s_h} \sum_{c_i=1}^{C_i} \sum_{h_i'=1}^{R_2} \sum_{i_i'=1}^{2} w_{1i} w_{2i_i'} F_{h,r} y_{i_i'}$$

Then the BRR-FAY variance estimator is equivalent to the sum of n / 2 *srswr* variance estimators for $n_{h'}$ = 2 in each stratum h and an extra term $\hat{A}(\hat{Y}_i)$, i.e.,

$$\frac{1}{R(1-k)^2} \sum_r \left(\hat{Y}_r - \hat{Y}\right)^2 = \sum_h \sum_{i \in s_h} \frac{\hat{v}_{\text{SRSWR}}(\hat{Y}_i) + A(\hat{Y}_i)}{\pi_i^2}$$

where

$$\hat{v}_{\text{SRSWR}}(\hat{Y}_i) = \sum_{c_i=1}^{C_i} \sum_{h_i'=1}^{R_2} \frac{1}{2(2-1)} \sum_{i_i'=1}^{2} \left(\frac{y_{i_i'}}{p_{i_i'}} - \hat{\hat{y}}_{h_i'} \right)^2$$

and

$$\hat{A}(\hat{Y}_i) = \sum_{c_i=1}^{C} \sum_{\substack{c_i'=1 \\ c_i' \neq c_i}}^{C} \sum_{h_i'=1}^{R_2} \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2} \right) \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2} \right)$$

Proof. Begin with the difference

$$\hat{Y}_r - \hat{Y} = \sum_h \sum_{i \in s_h} \sum_{c_i=1}^{C_i} \sum_{h_i'=1}^{R_2} \sum_{i_i'=1}^{2} w_{1i} w_{2i_i'} F_{h,r} y_{i_i'} - \sum_h \sum_{i \in s_h} \sum_{c_i=1}^{C_i} \sum_{h_i'=1}^{R_2} \sum_{i_i'=1}^{2} w_{1i} w_{2i_i'} y_{i_i'}$$

We note our abuse of notation: the sample weight $w_{2i_i'}$ for half sample i_i' is shorthand for $w_{h\,i\,c\,h'i'k}$ or the sample weight for unit k of half sample i' of variance stratum h' of cycle c of first-stage unit i and first-stage stratum h. We choose to include the least amount of subscripting by keeping what is essential at that point of the result, reduce the clutter of the subscripting, and assume the rest of the subscripting is understood and known. We do this similarly with $y_{h\,i\,c\,h'i'k}$.

$$= \sum_{h} \sum_{i \in s_h} w_{1i} \sum_{c_i=1}^{C_i} \sum_{h_i'=1}^{R_2} \left(w_{2i'_{i=1}} \left(1 + (1-k) a_{h_i^*, r_1} b_{h_i', r_2}\right) y_{i'_{i=1}} + w_{2i'_{i=2}} \left(1 - (1-k) a_{r, h_i^*} b_{h_i', r_2}\right) y_{i'_{i=2}} \right)$$

$$- \sum_{h} \sum_{i \in s_h} w_{1i} \sum_{c_i=1}^{C_i} \sum_{h_i'=1}^{R_2} \left(w_{2i'_{i=1}} y_{i'_{i=1}} + w_{2i'_{i=2}} y_{i'_{i=2}} \right)$$

$$= \sum_{h} \sum_{i \in s_h} w_{1i} \sum_{c_i=1}^{C_i} \sum_{h_i'=1}^{R_2} \left(w_{2i'_{i=1}} (1-k) a_{h_i^*, r_1} b_{h_i', r_2} y_{i'_{i=1}} - w_{2i'_{i=2}} (1-k) a_{h_i^*, r_1} b_{h_i', r_2} y_{i'_{i=2}} \right)$$

$$= (1-k) \sum_{h} \sum_{i \in s_h} w_{1i} \sum_{c_i=1}^{C_i} \sum_{h_i'=1}^{R_2} \left(w_{2i'_{i=1}} a_{h_i^*, r_1} b_{h_i', r_2} y_{i'_{i=1}} - w_{2i'_{i=2}} a_{h_i^*, r_1} b_{h_i', r_2} y_{i'_{i=2}} \right)$$

$$= (1-k) \sum_{h} \sum_{i \in s_h} w_{1i} \sum_{c_i=1}^{C_i} \sum_{h_i'=1}^{R_2} a_{h_i^*, r_1} b_{h_i', r_2} \left(\frac{y_{i'_{i=1}}}{\pi_{2i'_{i=1}}} - \frac{y_{i'_{i=2}}}{\pi_{2i'_{i=2}}} \right)$$

Next, note that we can change the summation over all first-stage strata and first-stage units in the following way

$$\sum_{h} \sum_{i \in s_h} \leftrightarrow \sum_{h^*=1}^{2L} \tag{A6.2}$$

and we can change the sum of all replicates as

$$\sum_{r=1}^{R} \leftrightarrow \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} \tag{A6.3}$$

Next, square the difference

$$\left(\hat{Y}_r - \hat{Y}\right)^2 = \left((1-k) \sum_{h^*=1}^{2L} w_{h^*} \sum_{c_i=1}^{C_i} \sum_{h_i'=1}^{R_2} a_{h_i^*, r_1} b_{h_i', r_2} \left(\frac{y_{i'_{i=1}}}{\pi_{2i'_{i=1}}} - \frac{y_{i'_{i=2}}}{\pi_{2i'_{i=2}}} \right) \right)^2$$

Appendix

$$
= (1-k)^2 \sum_{h^*=1}^{2L} w_{h^*}^2 a_{h_i^*,r_1}^2
\left[
\begin{array}{l}
\displaystyle \sum_{c_i=1}^{C} \sum_{h_i'=1}^{R_2} b_{h_i',r_2}^2 \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2}\right)^2 \\
\displaystyle + \sum_{c_i=1}^{C} \sum_{h_i'=1}^{R} \sum_{\substack{h_i''=1 \\ h_i''\neq h_i'}}^{R_2} b_{h_i',r_2} b_{h_i'',r_2} \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2}\right)\left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2}\right) \\
\displaystyle + \sum_{c_i=1}^{C} \sum_{\substack{c_i'=1 \\ c_i'\neq c_i}}^{C} \sum_{h_i'=1}^{R_2} b_{h_i',r_2}^2 \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2}\right)\left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2}\right) \\
\displaystyle + \sum_{c_i=1}^{C} \sum_{\substack{c_i'=1 \\ c_i'\neq c_i}}^{C} \sum_{h_i'=1}^{R_2} \sum_{\substack{h_i''=1 \\ h_i''\neq h_i'}}^{R_2} b_{h_i',r_2} b_{h_i'',r_2} \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2}\right)\left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2}\right)
\end{array}
\right]
$$

$$
+(1-k)^2 \sum_{h^*=1}^{2L} \sum_{\substack{h^{**}=1 \\ h^{**}\neq h^*}}^{2L} w_{h^*} w_{h^{**}} a_{h_i^*,r_1} a_{h_i^{**},r_1}
\left[
\begin{array}{l}
\displaystyle \sum_{c_i=1}^{C} \sum_{h_i'=1}^{R_2} b_{h_i',r_2}^2 \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2}\right)^2 \\
\displaystyle + \sum_{c_i=1}^{C} \sum_{h_i'=1}^{R_2} \sum_{\substack{h_i''=1 \\ h_i''\neq h_i'}}^{R_2} b_{h_i',r_2} b_{h_i'',r_2} \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2}\right)\left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2}\right) \\
\displaystyle + \sum_{c_i=1}^{C} \sum_{\substack{c_i'=1 \\ c_i'\neq c_i}}^{C} \sum_{h_i'=1}^{R_2} b_{h_i',r_2}^2 \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2}\right)\left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2}\right) \\
\displaystyle + \sum_{c_i=1}^{C} \sum_{\substack{c_i'=1 \\ c_i'\neq c_i}}^{C} \sum_{h_i'=1}^{R_2} \sum_{\substack{h_i^*=1 \\ h_i^*\neq h_i'}}^{R_2} b_{h_i',r_2} b_{h_i'',r_2} \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2}\right)\left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2}\right)
\end{array}
\right]
$$

The previous expression shows how $a_{h_i^*,r_1}$ relates to the first stage and how b_{h_i',r_2} relates to the variance strata of the second stage. The second and fourth terms of the top part and the entire bottom part of the previous expression will be equal to zero due to this construction.

$$
= (1-k)^2 \sum_{h^*=1}^{2L} w_{h^*}^2 \left[\begin{array}{l}
\displaystyle \sum_{c_i=1}^{C_i} \sum_{h_i'=1}^{R_2} a_{h_i^*,r_1}^2 b_{h_i',r_2}^{2} \left(w_{i_i'=1}y_{i_i'=1}-w_{i_i'=2}y_{i_i'=2}\right)^2 \\
\displaystyle + \sum_{c_i=1}^{C_i} \sum_{h_i'=1}^{R_2} \sum_{\substack{h_i''=1 \\ h_i^{**}\neq h_i^*}}^{R_2} a_{h_i^*,r_1}^2 b_{h_i',r_2} b_{h_i'',r_2} \left(w_{i_i'=1}y_{i_i'=1}-w_{i_i'=2}y_{i_i'=2}\right)\left(w_{i_i'=1}y_{i_i'=1}-w_{i_i'=2}y_{i_i'=2}\right) \\
\displaystyle + \sum_{c_i=1}^{C_i} \sum_{\substack{c_i'=1 \\ c_i'\neq c_i}}^{C_i} \sum_{h_i'=1}^{R_2} a_{h_i^*,r_1}^2 b_{h_i',r_2}^{2} \left(w_{i_i'=1}y_{i_i'=1}-w_{i_i'=2}y_{i_i'=2}\right)\left(w_{i_i'=1}y_{i_i'=1}-w_{i_i'=2}y_{i_i'=2}\right) \\
\displaystyle + \sum_{c_i=1}^{C_i} \sum_{\substack{c_i'=1 \\ c_i'\neq c_i}}^{C_i} \sum_{h_i'=1}^{R_2} \sum_{\substack{h_i''=1 \\ h_i''\neq h_i'}}^{R_2} a_{h_i^*,r_1}^2 b_{h_i',r_2} b_{h_i'',r_2} \left(w_{i_i'=1}y_{i_i'=1}-w_{i_i'=2}y_{i_i'=2}\right)\left(w_{i_i'=1}y_{i_i'=1}-w_{i_i'=2}y_{i_i'=2}\right)
\end{array} \right]
$$

$$
+(1-k)^2 \sum_{h^*=1}^{2L} \sum_{\substack{h^{**}=1 \\ h^{**}\neq h^*}}^{2L} w_{h^*} w_{h^{**}} \left[\begin{array}{l}
\displaystyle \sum_{c_i=1}^{C_i} \sum_{h_i'=1}^{R_2} a_{h_i^*,r_1} a_{h_i^{**},r_1} b_{h_i',r_2}^{2} \left(w_{i_i'=1}y_{i_i'=1}-w_{i_i'=2}y_{i_i'=2}\right)^2 \\
\displaystyle + \sum_{c_i=1}^{C_i} \sum_{h_i'=1}^{R_2} \sum_{\substack{h_i''=1 \\ h_i''\neq h_i'}}^{R_2} a_{h_i^*,r_1} a_{h_i^{**},r_1} b_{h_i',r_2} b_{h_i'',r_2} \left(w_{i_i'=1}y_{i_i'=1}-w_{i_i'=2}y_{i_i'=2}\right)\left(w_{i_i'=1}y_{i_i'=1}-w_{i_i'=2}y_{i_i'=2}\right) \\
\displaystyle + \sum_{c_i=1}^{C_i} \sum_{\substack{c_i'=1 \\ c_i'\neq c_i}}^{C_i} \sum_{h_i'=1}^{R_2} a_{h_i^*,r_1} a_{h_i^{**},r_1} b_{r,h_i'}^2 \left(w_{i_i'=1}y_{i_i'=1}-w_{i_i'=2}y_{i_i'=2}\right)\left(w_{i_i'=1}y_{i_i'=1}-w_{i_i'=2}y_{i_i'=2}\right) \\
\displaystyle + \sum_{c_i=1}^{C_i} \sum_{\substack{c_i'=1 \\ c_i'\neq c_i}}^{C_i} \sum_{h_i'=1}^{R_2} \sum_{\substack{h_i''=1 \\ h_i''\neq h_i'}}^{R_2} a_{h_i^*,r_1} a_{h_i^{**},r_1} b_{h_i',r_2} b_{h_i'',r_2} \left(w_{i_i'=1}y_{i_i'=1}-w_{i_i'=2}y_{i_i'=2}\right)\left(w_{i_i'=1}y_{i_i'=1}-w_{i_i'=2}y_{i_i'=2}\right)
\end{array} \right]
$$

Further, sum over the replicates, divide by $R(1-k)^2$, and apply (A6.3).

$$
\frac{1}{R(1-k)^2} \sum_r \left(\hat{Y}_r - \hat{Y}\right)^2 = \frac{1}{R(1-k)^2} \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} \left(\hat{Y}_r - \hat{Y}\right)^2
$$

Appendix

$$= \frac{(1-k)^2}{R(1-k)^2} \sum_{h^*=1}^{2L} w_{h^*}^2 \left[\begin{array}{c} \displaystyle \sum_{c_i'=1}^{C_i} \sum_{h_i'=1}^{R_2} \left(\sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} a_{h_i^*,r_1}^2 b_{h_i',r_2}^2 \right) \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2} \right)^2 \\ \displaystyle + \sum_{c_i=1}^{C_i} \sum_{h_i'=1}^{R_2} \sum_{\substack{h_i''=1 \\ h_i'' \neq h_i'}}^{R_2} \left(\sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} a_{h_i^*,r_1}^2 b_{h_i',r_2} b_{h_i'',r_2} \right) \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2} \right) \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2} \right) \\ \displaystyle + \sum_{c_i=1}^{C_i} \sum_{\substack{c_i'=1 \\ c_i' \neq c_i}}^{C_i} \sum_{h_i'=1}^{R_2} \left(\sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} a_{h_i^*,r_1}^2 b_{h_i',r_2}^2 \right) \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2} \right) \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2} \right) \\ \displaystyle + \sum_{c_i=1}^{C_i} \sum_{\substack{c_i'=1 \\ c_i' \neq c_i}}^{C_i} \sum_{h_i'=1}^{R_2} \sum_{\substack{h_i'=1 \\ h_i'' \neq h_i'}}^{R_2} \left(\sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} a_{h_i^*,r_1}^2 b_{h_i',r_2} b_{h_i'',r_2} \right) \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2} \right) \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2} \right) \end{array} \right]$$

$$+ \frac{(1-k)^2}{R(1-k)^2} \sum_{h^*=1}^{2L} \sum_{\substack{h^{**}=1 \\ h^{**} \neq h^*}}^{2L} w_{h^*} w_{h^{**}} \left[\begin{array}{c} \displaystyle \sum_{c_i=1}^{C_i} \sum_{h_i'=1}^{R_2} \left(\sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} a_{h_i^*,r_1} a_{h_i^{**},r_1} b_{h_i',r_2}^2 \right) \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2} \right)^2 \\ \displaystyle + \sum_{c_i=1}^{C_i} \sum_{h_i'=1}^{R_2} \sum_{\substack{h_i''=1 \\ h_i'' \neq h_i'}}^{R_2} \left(\sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} a_{h_i^*,r_1} a_{h_i^{**},r_1} b_{h_i',r_2} b_{h_i'',r_2} \right) \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2} \right) \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2} \right) \\ \displaystyle + \sum_{c_i=1}^{C_i} \sum_{\substack{c_i'=1 \\ c_i' \neq c_i}}^{C_i} \sum_{h_i'=1}^{R_2} \left(\sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} a_{h_i^*,r_1} a_{h_i^{**},r_1} b_{h_i',r_2}^2 \right) \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2} \right) \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2} \right) \\ \displaystyle + \sum_{c_i=1}^{C_i} \sum_{\substack{c_i'=1 \\ c_i' \neq c_i}}^{C_i} \sum_{h_i'=1}^{R_2} \sum_{\substack{h_i''=1 \\ h_i'' \neq h_i'}}^{R_2} \left(\sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} a_{h_i^*,r_1} a_{h_i^{**},r_1} b_{h_i',r_2} b_{h_i'',r_2} \right) \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2} \right) \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2} \right) \end{array} \right]$$

Given (A5.1) – (A5.4) from Result 5, we can simplify the previous expression as

$$\sum_{h^*=1}^{2L} w_{h^*}^2 \left[\begin{array}{c} \sum_{c_i=1}^{C_i} \sum_{h_i'=1}^{R_2} \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2} \right)^2 \\ + \sum_{c_i=1}^{C_i} \sum_{\substack{c_i'=1 \\ c_i' \neq c_i}}^{C_i} \sum_{h_i'=1}^{R_2} \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2} \right) \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2} \right) \end{array} \right] \tag{A6.4}$$

or the following after reapplying (A6.3) and noting that $w_{h^*}^2 = 1/\pi_i^2$.

$$\sum_h \sum_{i \in s_h} \frac{1}{\pi_i^2} \left[\begin{array}{c} \sum_{c_i=1}^{C_i} \sum_{h_i'=1}^{R_2} \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2} \right)^2 \\ + \sum_{c_i=1}^{C_i} \sum_{\substack{c_i'=1 \\ c_i' \neq c_i}}^{C_i} \sum_{h_i'=1}^{R_2} \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2} \right) \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2} \right) \end{array} \right] \tag{A6.5}$$

Now, apply Result 3 to the top expression in (A6.5) and recognize the bottom expression as $A(\hat{Y}_i)$, which leads to our desired result.

$$\frac{1}{R(1-k)^2} \sum_r \left(\hat{Y}_r - \hat{Y} \right)^2 = \sum_h \sum_{i \in s_h} \frac{\hat{v}_{\text{BRR-SYS}}(\hat{Y}_i) + A(\hat{Y}_i)}{\pi_i^2}.$$

where

$$\hat{v}_{\text{BRR-SYS}}(\hat{Y}_i) = \sum_{c_i=1}^{C_i} \sum_{h_i'=1}^{R_2} \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2} \right)^2$$

and

$$A(\hat{Y}_i) = \sum_{c_i=1}^{C_i} \sum_{\substack{c_i'=1 \\ c_i' \neq c_i}}^{C_i} \sum_{h_i'=1}^{R_2} \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2} \right) \left(w_{i_i'=1} y_{i_i'=1} - w_{i_i'=2} y_{i_i'=2} \right)$$

Appendix

Result 7. Rizzo and Rust (2011). Define the estimator for the replicate total of the variable of interest *y* for each replicate *r* as

$$\hat{Y} = \sum_{h} \sum_{i \in s_h} \sum_{k \in s_i} w_{1i} w_{2k} y_k$$

where the first-stage weight is $w_i = \pi_i^{-1}$ and the second-stage weight is w_k. Also, define the estimator for each replicate total *r* as

$$\hat{Y}_r = \sum_{h} \sum_{i \in s_h} \sum_{k \in s_i} w_i w_k F_{k,r} y_k$$

Divide the second-stage sample s_i for first-stage unit *i* into groups A_i and D_i are such that $A_i \cup D_i = s_i$ and $A_i \cup D_i = \emptyset$. Next, define the replicate factors as

$$F_{k,r} = \begin{cases} 1 + a_{r,h} & j \in A_i \\ 1 - a_{r,h} & j \in D_i \\ 1 & \text{otherwise} \end{cases}$$

Then with the previous replicate factors, the replicate variance estimator is equivalent to a *srswr* variance estimator where the sample units of A_i and D_i are treated as clusters selected with replacement, i.e.,

$$\frac{1}{R} \sum_{r} \left(\hat{Y}_r - \hat{Y} \right)^2 = \sum_{h} \sum_{i \in s_h} \frac{1}{\pi_i^2} \left(\hat{Y}_{A_i} - \hat{Y}_{D_i} \right)^2$$

Proof. Beginning with the difference

$$\begin{aligned}
\hat{Y}_r - \hat{Y} &= \sum_{h} \sum_{i \in s_h} \sum_{s_{2i}} w_{1i} w_k F_{k,r} y_k - \sum_{h} \sum_{i \in s_h} \sum_{s_i} w_i w_k y_k \\
&= \sum_{h} \sum_{i \in s_h} \sum_{k \in s_i} \left(w_i w_k F_{k,r} y_k - w_i w_k y_k \right) \\
&= \sum_{h} \sum_{i \in s_h} \left[\sum_{k \in A_i} \left(w_i w_k F_{k,r} y_k - w_i w_k y_k \right) + \sum_{k \in D_i} \left(w_i w_k F_{k,r} y_k - w_i w_k y_k \right) \right] \\
&= \sum_{h} \sum_{i \in s_h} \left[\sum_{k \in A_i} \left(w_k (1 + a_{h,r}) y_k - w_k y_k \right) + \sum_{k \in D_i} \left(w_k (1 - a_{h,r}) y_k - w_k y_k \right) \right] \\
&= \sum_{h} \sum_{i \in s_h} \left[\sum_{k \in A_i} w_k \left((1 + a_{h,r}) - 1 \right) y_k + \sum_{k \in D_i} w_k \left((1 - a_{h,r}) - 1 \right) y_k \right] \\
&= \sum_{h} \sum_{i \in s_h} \left[\sum_{k \in A_i} w_k a_{h,r} y_k - \sum_{k \in D_i} w_k a_{h,r} \, y_k \right] \\
&= \sum_{h} \sum_{i \in s_h} w_i a_{h,r} \left[\sum_{k \in A_i} w_k y_k - \sum_{k \in D_i} w_k \, y_k \right]
\end{aligned}$$

Appendix

where $\hat{Y}_{A_i} = \sum_{k \in A_i} w_k y_k$, $\hat{Y}_{D_i} = \sum_{k \in D_i} w_k y_k$, and $\hat{Y}_i = \hat{Y}_{A_i} + \hat{Y}_{D_i}$.

then

$$\hat{Y}_r - \hat{Y} = \sum_h \sum_{i \in s_h} \frac{1}{\pi_i} a_{h,r} \left(\hat{Y}_{A_i} - \hat{Y}_{D_i} \right)$$

Next, square the difference

$$\begin{aligned} \left(\hat{Y}_r - \hat{Y} \right)^2 &= \left(\sum_h \sum_{i \in s_h} \frac{1}{\pi_i} a_{rh} \left(\hat{Y}_{A_i} - \hat{Y}_{D_i} \right) \right)^2 \\ &= \sum_h \sum_{i \in s_{1h}} \frac{1}{\pi_i^2} a_{h,r}^2 \left(\hat{Y}_{A_i} - \hat{Y}_{D_i} \right)^2 + \sum_h \sum_{\substack{h' \\ h' \neq h}} \sum_{i \in s_h} \sum_{\substack{i' \in s_h \\ i' \neq i}} \frac{1}{\pi_i} \frac{1}{\pi_{i'}} a_{h,r} a_{h',r} \left(\hat{Y}_{A_{2i}} - \hat{Y}_{D_{2i}} \right) \left(\hat{Y}_{A_{2i'}} - \hat{Y}_{D_{2i'}} \right) \end{aligned}$$

Further

$$\begin{aligned} \frac{1}{R} \sum_r \left(\hat{Y}_r - \hat{Y} \right)^2 &= \frac{1}{R} \sum_r \sum_h \sum_{i \in s_h} \frac{1}{\pi_i^2} a_{h,r}^2 \left(\hat{Y}_{A_i} - \hat{Y}_{D_i} \right)^2 + \frac{1}{R} \sum_r \sum_h \sum_{\substack{h' \\ h' \neq h}} \sum_{i \in s_h} \sum_{\substack{i' \in s_h \\ i' \neq i}} \frac{1}{\pi_i} \frac{1}{\pi_{i'}} a_{h,r} a_{h',r} \left(\hat{Y}_{A_i} - \hat{Y}_{D_i} \right) \left(\hat{Y}_{A_{i'}} - \hat{Y}_{D_{i'}} \right) \\ &= \frac{1}{R} \sum_h \sum_{i \in s_h} \frac{1}{\pi_i^2} \left(\sum_r a_{h,r}^2 \right) \left(\hat{Y}_{A_i} - \hat{Y}_{D_i} \right)^2 + \frac{1}{R} \sum_h \sum_{\substack{h' \\ h' \neq h}} \sum_{i \in s_h} \sum_{\substack{i' \in s_h \\ i' \neq i}} \frac{1}{\pi_i} \frac{1}{\pi_{i'}} \left(\sum_r a_{h,r} a_{h',r} \right) \left(\hat{Y}_{A_i} - \hat{Y}_{D_i} \right) \left(\hat{Y}_{A_{i'}} - \hat{Y}_{D_{i'}} \right) \end{aligned}$$

since $\sum_r a_{h,r}^2 = R$ and $\sum_r a_{h,r} a_{h',r} = 0$, we can simplify the previous line as

$$\frac{1}{R} \sum_r \left(\hat{Y}_r - \hat{Y} \right)^2 = \sum_h \sum_{i \in s_h} \frac{1}{\pi_i^2} \left(\hat{Y}_{A_i} - \hat{Y}_{D_i} \right)^2$$

Now consider $\hat{v}\left(\hat{Y}_i\right) = \left(\hat{Y}_{A_i} - \hat{Y}_{D_i} \right)^2$ as the variance estimator of a *srswr* variance estimator where the sample units of A_i and D_i are treated as clusters selected with *srswr*.