Using Successive Difference Replication for Estimating Variances

Stephen Ash
U.S. Census Bureau 4600 Silver Hill Road, Washington, DC 20233-7400

Abstract
Fay and Train (1995) present a method called successive difference replication that can be used to estimate the variance of an estimated total from a systematic random sample from an ordered list. The estimator uses the general form of a replication variance estimator, where the replicate factors are constructed such that the estimator mimics the successive difference estimator. In the paper, we establish the conditions for successive difference replication to be equivalent to successive difference estimator. We also discuss the impact of using a subset of the replicates as an estimator instead of the full set of replicates that is needed for the successive difference replication estimator to be equivalent to the successive difference estimator.

Key Words: Successive differences, successive difference replication, systematic random sampling

1. Introduction

Fay and Train (1995) present a method called successive difference replication (SDR) that can be used to estimate the variance of an estimated total from a systematic random sample from an ordered list. The estimator uses the general form of a replication variance estimator, where the replicate factors are constructed such that it mimics the successive difference (SD) estimator.

We add to the methodology the idea of a connected loop, which helps us understand how row assignments work. We also provide a method for using SDR with a Hadamard matrix that is smaller than the sample size. The method also helps us understand the consequences of using a reduced set of replicates with SDR.

The paper begins by reviewing systematic random sampling from an ordered list. Next, we review the SD estimator and how it is suited for variance estimation of systematic random samples. The main section of the paper presents two theorems that provide conditions for the SDR estimator to be equivalent to the SD estimator. The paper ends with empirical examples that examine alternative row assignments and the suitability of using of a reduced set of replicates.

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 simple random sample without replacement sample design (*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 $(0, r]$. 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 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). In this way of thinking, 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 $\lceil r + (n-1) * k \rceil$ 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 SD in preparation of our discussion of SDR.

1.2 Review of Successive Differences

Wolter (1984; estimator 2) provides a form of the successive difference estimator of the variance of an estimated mean ($\bar{y}$) for a *sys* sample design as

$$\hat{v}_{SD1}\left(\hat{\bar{y}}\right) = (1-f)\frac{1}{2n(n-1)}\sum_{k=2}^{n}\left(y_k - y_{k-1}\right)^2$$

where

y_k	the variable of interest
k	indexes the units of the ordered sample
Y and $\hat{Y}$	the total of variable y_k and the estimator of the total
N and n	the size of the universe and sample
$\bar{y}$ and $\hat{\bar{y}}$	the mean of the variable of interest y_k and the estimator of the mean
f	the sampling fraction, i.e., $f = n / N$.

This estimator has been described by Yates (1949; p. 229-231) and recommended by Wolter (1984). Murthy and Rao (1988, eq. 32), provide a sketch of why the estimator works. The short version is that since *sys* only selects one unit within each implicit stratum, SD's solution is to collapse adjacent implicit strata. With two units, we can estimate the variance of an implicit stratum. Implicit strata are collapsed and averaged over all possible pairs, and then multiplied by *n*, the number of implicit strata, to give the variance of all the implicit strata.

One SD variance estimator of a total from a *sys* sample is given by Fay and Train (1995), which we will refer to as F&T for the remainder of the paper, as

$$\hat{v}_{SD1}\left(\hat{Y}\right)=\left(1-f\right)\frac{n}{2\left(n-1\right)}\sum_{k=2}^{n}\left(\breve{y}_k-\breve{y}_{k-1}\right)^2 .$$

where we define the weighted variable of interest as $\breve{y}_k=\left(\frac{N}{n}\right)y_k$. F&T also define a second SD estimator, which is made "circular" in that it includes an extra squared difference that links the first and last unit from the sorted list.

$$\hat{v}_{SD2}\left(\hat{Y}\right)=\tfrac{1}{2}\left(1-f\right)\left[\sum_{k=2}^{n}\left(\breve{y}_k-\breve{y}_{k-1}\right)^2+\left(\breve{y}_n-\breve{y}_1\right)^2\right], \tag{1}$$

where the estimator of the total Y is defined as $\hat{Y}=\sum_{k=1}^{n}\breve{y}_k$.

The SD2 estimator can also be expressed in matrix notation quadratic form $\breve{\mathbf{y}}'\mathbf{C}\breve{\mathbf{y}}$, where $\breve{\mathbf{y}}'=\left[\breve{y}_1\ \breve{y}_2\ldots\breve{y}_n\right]$ be defined as the $m \times 1$ weighted observation vector and $\mathbf{C}$ is a square matrix with 2 for each value of the diagonal, -1 for every value of the superdiagonal and subdiagonal, and -1 for the bottom left and top right value. Here the superdiagonals are defined as the diagonals adjacent to the main diagonal. The exception is a 2×2 matrix, which we will see later. Example 1 demonstrates.

Example 1: With $n = 4$ and $k = 4$,

$$\tfrac{1}{2}\left(1-f\right)\left[\sum_{k=2}^{4}\left(\breve{y}_k-\breve{y}_{k-1}\right)^2+\left(\breve{y}_4-\breve{y}_1\right)^2\right]=\tfrac{1}{2}\left(1-f\right)\breve{\mathbf{y}}'\mathbf{C}\breve{\mathbf{y}}$$

where, $\mathbf{C}=\begin{bmatrix}2&-1&0&-1\\-1&2&-1&0\\0&-1&2&-1\\-1&0&-1&2\end{bmatrix}$ and $\breve{\mathbf{y}}=\begin{bmatrix}\breve{y}_1\\\breve{y}_2\\\breve{y}_3\\\breve{y}_4\end{bmatrix}$.

As a preview to the next section, we mention that in a different paper, Fay (1989) shows that if you can represent a variance as a quadratic form $\breve{\mathbf{y}}'\mathbf{C}\breve{\mathbf{y}}$, where $\mathbf{C}$ is positive semi-definite square matrix, then the estimator can also be expressed as a replication variance estimator. Although SDR does not follow the same prescription given in Fay (1989), the next section does show how the SD2 variance estimator can be represented as a replicate variance estimator.

2. Successive Difference Replication

2.1 Definition of Successive Difference Replication

F&T present a method called successive difference replication (SDR) to estimate the variance from a sample selected with *sys*. We show how SDR can be used to produce replicate weights for a general replicate variance estimator that is exactly equivalent to the SD2 estimator. Before we define the SDR estimator in the first theorem, we first establish some terms and provide a lemma that is used by the theorem.

A row assignment scheme, or more simply RA, is an assignment of two rows of a matrix to each unit in the sample. We usually denote the pair of rows as (a_i, b_i) for unit i. A connected loop is a RA that does not repeat any of the rows, i.e., $a_i \neq a_j$ and $b_i \neq b_j$ for all i and j in the connected loop, and is circular, i.e., $b_{i+1} = a_i$ for all $i < n$ and $b_n = a_1$. For example, one possible connected loop for three observations is (1,2), (2,3), (3,1).

A shift matrix **S** can be used to move either the rows or columns of a matrix. We will explain how to move rows because that is useful to us, but note that moving columns is similar. A shift matrix is a square matrix that has all 0s, except a single 1 in each column. If we wanted to move row p to row q, we would put a 1 in the q^{th} row of the p^{th} column and 0s elsewhere. We note that order is important in applying a shift matrix to another matrix. The application of **S** to another square matrix **A** as **AS** shifts the rows of **A** and **SA** shifts the columns of **A**. For example, if we applied the $m = 4$ the shift matrix

$$\mathbf{S}_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

to another square matrix **A** of the same dimension, i.e., **AS** = **B**, the matrix **B** is a copy of **A** with all its rows shifted one position down and the last row made the first.

Lemma: Let $\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_c$ be shift matrices, then $block(\mathbf{S}_1'\mathbf{S}_1, \mathbf{S}_2'\mathbf{S}_2, ..., \mathbf{S}_C'\mathbf{S}_C) = \mathbf{I}$.

Proof. We first define a general block diagonal matrix **A** that is formed by the square matrices $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_C$ as

$$\mathbf{A} = block(\mathbf{A}_1, \mathbf{A}_2, ..., \mathbf{A}_C) = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & \ldots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \ldots & \mathbf{A}_C \end{bmatrix}.$$

It can be shown that if both **A** and **B** are block diagonal matrices and the square matrices $\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_C$ have the same dimensions as $\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_C$, respectively, then $\mathbf{AB} = block(\mathbf{A}_1\mathbf{B}_1, \mathbf{A}_2\mathbf{B}_2, ..., \mathbf{A}_C\mathbf{B}_C)$. For a given shift matrix, we also know that $\mathbf{SS}' = \mathbf{I}$, since **S** and **S'** shift the rows of a given matrix in opposite directions. With the two previous items, the lemma follows.

We also define a one row shift matrix as a shift matrix that either shifts all the rows of another matrix down one row and the last row moved to the first or shifts all the rows of another matrix up one row and the last row moved to the last. If $\mathbf{S}_D$ is a one row shift matrix that moves rows down then it has 1s along the upper superdiagonal and a 1 in the bottom left entry of the matrix, for example $\mathbf{S}_1$. Similarly, if $\mathbf{S}_U$ is one row shift matrix that moves rows up, then it has 1s along the lower superdiagonal and a 1 in the top entry of the matrix, for example the subsequently defined $\mathbf{S}_2$. Note the property that $\mathbf{S}_D = \mathbf{S}_U'$ and $\mathbf{S}_U = \mathbf{S}_D'$, therefore $\mathbf{S}_U + \mathbf{S}_U' = \mathbf{S}_D + \mathbf{S}_D'$.

We now present the main theorem of the paper that establishes the conditions under which SDR is equivalent to SD2.

Theorem 1: Let n be the sample size of a given *sys* sample and $\breve{\mathbf{y}}' = \left[\breve{y}_1 \ \breve{y}_2 \ldots \breve{y}_n\right]$ be defined 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 $n \le k$.
(b) Choose a RA that assigns two rows $\left(a_i, b_i\right)$ to each unit i in the sample. Let the RA define C connected loops of m_c units in each connected loop c.
(c) Choose the $m = n$ rows of $\mathbf{H}$ corresponding to the RA to make the $m \times k$ matrix $\mathbf{M}$. The order of the rows of $\mathbf{M}$ should correspond to the first row of the RA. For example, the first row of $\mathbf{M}$ should be row $a_{i=1}$ of $\mathbf{H}$, the second row should be row $a_{i=2}$ of $\mathbf{H}$, etc. Next define the $m \times m$ shift matrix as $\mathbf{S} = block\left(\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_C\right)$ 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 $\mathbf{M}$. 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 $\hat{Y}_r = \sum_{i=1}^{n} f_{i,r}\breve{y}_i$, where the matrix of replicate factors is $\mathbf{F} = \mathbf{1}_m\mathbf{1}_k{}' + \left(2^{-\frac{3}{2}}\mathbf{I}_m - 2^{-\frac{3}{2}}\mathbf{S}\right)\mathbf{M}$ and individual values within the matrix are defined for each unit i (rows of $\mathbf{F}$) of replicate r (columns of $\mathbf{F}$) as

$$f_{i,r} = 1 + 2^{-\frac{3}{2}} h_{a_i,r} - 2^{-\frac{3}{2}} h_{b_i,r}. \qquad (2)$$

$\mathbf{I}_m$ is a $m \times m$ identity matrix and $\mathbf{1}_m$ is a $m \times 1$ vector of 1s. Then the SDR variance estimator

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

is equivalent to the sum of C different SD2 estimators.

Proof. The SDR estimator can be written in matrix notation as

$$(1-f)\frac{4}{k}\left(\breve{\mathbf{y}}'\left(\mathbf{1}_m\mathbf{1}_k{}' + \left(2^{-\frac{3}{2}}\mathbf{I}_m - 2^{-\frac{3}{2}}\mathbf{S}\right)\mathbf{M}\right) - \breve{\mathbf{y}}'\mathbf{1}_m\mathbf{1}_k{}'\right)\left(\breve{\mathbf{y}}'\left(\mathbf{1}_m\mathbf{1}_k{}' + \left(2^{-\frac{3}{2}}\mathbf{I}_m - 2^{-\frac{3}{2}}\mathbf{S}\right)\mathbf{M}\right) - \breve{\mathbf{y}}'\mathbf{1}_m\mathbf{1}_k{}'\right)'$$
$$= (1-f)\frac{4}{k}\left(2^{-\frac{3}{2}}\right)^2 \breve{\mathbf{y}}'\left(\mathbf{I}_m - \mathbf{S}\right)\mathbf{M}\mathbf{M}'\left(\mathbf{I}_m - \mathbf{S}\right)'\breve{\mathbf{y}}$$

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

$$(1-f)\frac{1}{2k}\breve{\mathbf{y}}'\left(\mathbf{I}_m - \mathbf{S}\right)\left(k\mathbf{I}_m\right)\left(\mathbf{I}_m - \mathbf{S}\right)'\breve{\mathbf{y}}$$
$$= \tfrac{1}{2}(1-f)\breve{\mathbf{y}}'\left(\mathbf{I}_m - \mathbf{S}\right)\left(\mathbf{I}_m - \mathbf{S}\right)'\breve{\mathbf{y}}$$
$$= \tfrac{1}{2}(1-f)\breve{\mathbf{y}}'\left(2\mathbf{I}_m - \mathbf{S} - \mathbf{S}'\right)\breve{\mathbf{y}}$$

The last line follows from the lemma and has a constant value for any choice of **H.** By noting the block diagonal structure of **S,** we can write the estimator as

$$\tfrac{1}{2}(1-f)\sum_{c=1}^{C}\breve{\mathbf{y}}_c'\left(2\mathbf{I}_m-\mathbf{S}_c-\mathbf{S}_c'\right)\breve{\mathbf{y}}_c\,,$$

where $\breve{\mathbf{y}}_c$ corresponds to the vector of the weighted observations in connected loop *c,* which is a result of partitioning the weighted observation vector as $\breve{\mathbf{y}}'=\left[\breve{\mathbf{y}}_{c=1}\,\breve{\mathbf{y}}_{c=2}\ldots\breve{\mathbf{y}}_{c=C}\right]$.

The choice of the RA does not change the result, since we know that $2\mathbf{I}_m-\mathbf{S}_c-\mathbf{S}_c'$ is constant for either an up or down one row shift matrix $\mathbf{S}_c$**.**

Example 2: We can use the normal Hadamard matrix of order $k = 4$,

$$\mathbf{H}_{4a}=\begin{bmatrix}1&1&1&1\\1&-1&1&-1\\1&1&-1&-1\\1&-1&-1&1\end{bmatrix}$$

to estimate the variance a *sys* sample of size $n = 4$. We chose the RA of the second column of Table 1 that forms a single connected loop. With this choice of Hadamard and RA, it follows that $\mathbf{M} = \mathbf{H}_{4a}$ and $\mathbf{S} = \mathbf{S}_1$. Note that a Hadamard matrix is normal if it has 1s in both the first row and first column.

The replicate factor for unit 1 of replicate 3 is $f_{1,3}=1+\frac{1}{\sqrt{2}}\cong 1.7$, since $a_{1,3}$ = -1 and $a_{2,3}$ = -1. Similarly, the replicate factor for unit 2 of replicate 2 is $f_{1,3}=1-\frac{1}{\sqrt{2}}\cong 0.3$, since $a_{1,3}$ = -1 and $a_{2,3}$ = 1. In Table 1 and the remainder of the paper, 0.3 and 1.7 are used as shorthand for $1-\frac{1}{\sqrt{2}}$ and $1+\frac{1}{\sqrt{2}}$, respectively. The last four columns of Table 2 show the resultant replicate factors for the all units of all replicates.

Table 1: Matrix of Replicate Factors ($f_{i,r}$) for Example 1

		Replicate			
Unit #	RA	1	2	3	4
1	(1,2)	1.0	1.7	1.0	1.7
2	(2,3)	1.0	0.3	1.7	1.0
3	(3,4)	1.0	1.7	1.0	0.3
4	(4,1)	1.0	0.3	0.3	1.0

Using the replicate factors of Table 1, the SDR estimator is equivalent to the SD2 estimator, i.e.,

$$(1-f)\frac{4}{k}\sum_{r=1}^{4}\left(\hat{\bar{Y}}_r-\hat{\bar{Y}}\right)^2=\frac{1}{2}(1-f)\left[\sum_{i=1}^{3}\left(\breve{y}_{i+1}-\breve{y}_i\right)^2+\left(\breve{y}_1-\breve{y}_4\right)^2\right]. \qquad (3)$$

For a large sample size, it is not usually practical to use **H**, where $n < k$. The second theorem

shows one way that we can use $\mathbf{H}$ with $k < n$ to produce a larger Hadamard matrix $\widetilde{\mathbf{H}}$ with $k \geq n$ that will result in the SDR estimator being equivalent to the SD2 estimator.

The second theorem also builds upon and clarifies the instructions F&T give for the case of $n > k$. In F&T's instructions, they use the term cycle to denote every $m_d \leq k$ units of the sample. Theorem 2 does not make conditions on the RA, but otherwise it does follow the set-up of F&T.

Theorem 2: Let n be the sample size of a given *sys* sample.

(a) Choose a Hadamard matrix $\mathbf{H}_\mathrm{A}$ of order k_A, where $n > k_\mathrm{A}$.

(b) Choose a RA that assigns rows to $\mathbf{H}_\mathrm{A}$ to the sample. Retaining their original order, split the n sample units into D cycles. Each cycle d has $m_d \leq k_\mathrm{A}$ units. Within each cycle, the RA defines one or more connected loops.

(c) Choose a semi-normal Hadamard matrix $\mathbf{H}_\mathrm{B}$ of order k_B and use it to define a larger Hadamard matrix $\widetilde{\mathbf{H}}$ of order $\widetilde{k}$ that is generated from the original $\mathbf{H}_\mathrm{A}$. This can be done by applying a Welsch construction to $\mathbf{H}_\mathrm{A}$, i.e., $\widetilde{\mathbf{H}} = \mathbf{H}_\mathrm{B} \otimes \mathbf{H}_\mathrm{A}$.

(d) Choose the $m = \sum_{d=1}^{D} m_d$ rows of $\widetilde{\mathbf{H}}$ that correspond to the RA to make the $m \times \widetilde{k}$ matrix $\widetilde{\mathbf{M}}$. The order of the rows of $\widetilde{\mathbf{M}}$ should correspond to the first row of the RA. Next define the $m \times m$ shift matrix as $\mathbf{S} = block(\mathbf{S}_1, \mathbf{S}_2, ..., \mathbf{S}_D)$ where the $m_d \times m_d$ shift matrices $\mathbf{S}_d$ are defined to identify the position of the second row b_i of the RA in $\widetilde{\mathbf{M}}$.

With this prescription, the SDR estimator is defined as $\hat{v}_{SDR}(\hat{Y}) = (1-f)\frac{4}{\widetilde{k}}\sum_{r=1}^{\widetilde{k}}(\hat{Y}_r - \hat{Y})^2$ is equivalent to the sum of at least D SD2 estimators.

Proof. The result follows by applying Theorem 1. The specific value of D follows from the fact that each of the D cycles can have one or more connected loops, so there will be a total of at least D connected loops.

Example 4: Let $n = 14$ and choose the non-normal Hadamard $\mathbf{H}_\mathrm{A} = \mathbf{H}_{4b}$ of order $k_\mathrm{A} = 4$. The number of cycles will be $D = 4$ and the RA within each cycle is given in the second column of Table 3 for each unit. We define $\widetilde{\mathbf{H}}$ of $\widetilde{k} = 16$ using a Welsh construction of the original normal Hadamard matrix as

$$\mathbf{H}_{4b} = \begin{bmatrix} 1 & -1 & 1 & 1 \\ -1 & -1 & -1 & 1 \\ 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & 1 \end{bmatrix} \qquad \mathbf{H}_{16} = \mathbf{H}_{4a} \otimes \mathbf{H}_{4b} = \begin{bmatrix} \mathbf{H}_{4b} & \mathbf{H}_{4b} & \mathbf{H}_{4b} & \mathbf{H}_{4b} \\ \mathbf{H}_{4b} & -\mathbf{H}_{4b} & \mathbf{H}_{4b} & -\mathbf{H}_{4b} \\ \mathbf{H}_{4b} & \mathbf{H}_{4b} & -\mathbf{H}_{4b} & -\mathbf{H}_{4b} \\ \mathbf{H}_{4b} & -\mathbf{H}_{4b} & -\mathbf{H}_{4b} & \mathbf{H}_{4b} \end{bmatrix}.$$

Using $\mathbf{H}_{16}$, we can calculate the replicate factors for 16 replicates as Table 3.

In matrix notation, $\widetilde{\mathbf{M}}$ includes all the rows of $\widetilde{\mathbf{H}} = \mathbf{H}_{16}$ except rows 13 and 16. The rows of $\widetilde{\mathbf{M}}$ are ordered by a_i, the first row assigned in the RA. The shift matrix is defined as $\mathbf{S} = block(\mathbf{S}_1, \mathbf{S}_2, \mathbf{S}_3, \mathbf{S}_4)$, where the shift matrices corresponding to each cycle are

$$\mathbf{S}_2 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad \mathbf{S}_3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad \mathbf{S}_4 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

and $\mathbf{S}_1$ is as previously defined.

Table 3: Matrix of Replicate Factors ($f_{i,r}$) for Example 4

Unit #	RA $\mathbf{H}_A = \mathbf{H}_{4b}$	RA $\widetilde{\mathbf{H}} = \mathbf{H}_{16}$	Cycle	Replicate 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	(1,2)	(1,2)		1.7	1.0	1.7	1.0	1.7	1.0	1.7	1.0	1.7	1.0	1.7	1.0	1.7	1.0	1.7	1.0
2	(2,3)	(2,3)	1	0.3	1.0	1.0	1.7	0.3	1.0	1.0	1.7	0.3	1.0	1.0	1.7	0.3	1.0	1.0	1.7
3	(3,4)	(3,4)		1.0	0.3	1.0	0.3	1.0	0.3	1.0	0.3	1.0	0.3	1.0	0.3	1.0	0.3	1.0	0.3
4	(4,1)	(4,1)		1.0	1.7	0.3	1.0	1.0	1.7	0.3	1.0	1.0	1.7	0.3	1.0	1.0	1.7	0.3	1.0
5	(1,3)	(5,7)		1.0	1.0	1.7	1.7	1.0	1.0	0.3	0.3	1.0	1.0	1.7	1.7	1.0	1.0	0.3	0.3
6	(3,1)	(7,5)	2	1.0	1.0	0.3	0.3	1.0	1.0	1.7	1.7	1.0	1.0	0.3	0.3	1.0	1.0	1.7	1.7
7	(2,4)	(6,8)		0.3	0.3	1.0	1.0	1.7	1.7	1.0	1.0	0.3	0.3	1.0	1.0	1.7	1.7	1.0	1.0
8	(4,2)	(8,6)		1.7	1.7	1.0	1.0	0.3	0.3	1.0	1.0	1.7	1.7	1.0	1.0	0.3	0.3	1.0	1.0
9	(1,4)	(9,12)		1.0	0.3	1.7	1.0	1.0	0.3	1.7	1.0	1.0	1.7	0.3	1.0	1.0	1.7	0.3	1.0
10	(4,3)	(12,11)	3	1.0	1.7	1.0	1.7	1.0	1.7	1.0	1.7	1.0	0.3	1.0	0.3	1.0	0.3	1.0	0.3
11	(3,2)	(11,10)		1.7	1.0	1.0	0.3	1.7	1.0	1.0	0.3	0.3	1.0	1.0	1.7	0.3	1.0	1.0	1.7
12	(2,1)	(10,9)		0.3	1.0	0.3	1.0	0.3	1.0	0.3	1.0	1.7	1.0	1.7	1.0	1.7	1.0	1.7	1.0
13	(2,3)	(14,15)	4	0.3	1.0	1.0	1.7	1.7	1.0	1.0	0.3	1.7	1.0	1.0	0.3	0.3	1.0	1.0	1.7
14	(3,2)	(15,14)		1.7	1.0	1.0	0.3	0.3	1.0	1.0	1.7	0.3	1.0	1.0	1.7	1.7	1.0	1.0	0.3

With the replicate factors of Table 3, the SDR estimator is equivalent to the sum of five different SD2 estimators, one for each connected loop of the RA, i.e. the following expression with $R = 0$,

$$(1-f)\frac{4}{\widetilde{k}}\sum_{r=1}^{\widetilde{k}}\left(\hat{Y}_r - \hat{Y}\right)^2 = \tfrac{1}{2}(1-f)\left[\begin{array}{l}\displaystyle\sum_{i=2}^{4}(y_i - y_{i-1})^2 + (y_4 - y_1)^2 + 2(y_6 - y_5)^2 + 2(y_8 - y_7)^2 \\ \displaystyle + \sum_{i=10}^{12}(y_i - y_{i-1})^2 + (y_{12} - y_9)^2 + 2(y_{13} - y_{13})^2\end{array}\right] + R \cdot \qquad (4)$$

There are a few items to note with Example 4. First, the number of replicates needed is greater than the sample size. This happens when m_d is not constant across all cycles. In example 4, the fourth cycle only had two sample units, but we had to use four replicates from each $\mathbf{H}_{4b}$ because at least one of the cycles used four rows.

To make the example more interesting, we choose a non-normal Hadamard matrix $\mathbf{H}_{4b}$ for $\mathbf{H}_A$. This non-normal Hadamard was generated by starting with the normal Hadamard $\mathbf{H}_{4a}$ and reversing the procedure for finding a normal Hadamard as described by Hedayat and Wallis (1978). Here we simply changed the sign of all units in the second row and then changed all the signs for the second column.

If we would have used the normal Hadamard matrix $\mathbf{H}_{4a}$ for both $\mathbf{H}_A$ and $\mathbf{H}_B$, the replicate factors for replicates 1, 5, 9, and 13 would have all been 1.0. We call a replicate where every unit gets a value of 1.0 and thereby the replicate estimate is equal to the original estimate, a dead replicate. In SDR, there is nothing wrong with dead replicates, it is just the way the replicate factors are distributed by the Hadamard matrix. With a dead replicate, many of the values of 1.0 are in the dead replicate, and the other replicates are more mixed with values of 1.7 and 0.3. All the replicates, even the dead replicates, are needed in the estimation. Section 3.1 will further discuss how the distribution of replicate factors is impacted by the choice of $\mathbf{H}$.

The real value of Theorem 2 is in understanding F&T's original prescription for SDR when $n > k$. In F&T, the RA is applied repeatedly to the $m = k - 1$ rows of $\mathbf{H}_A$ (skipping the first row of $\mathbf{H}_A$), where $\mathbf{H}_A$ is chosen as a normal Hadamard matrix. Replicates are then formed using the k_A columns of $\mathbf{H}_A$. If we apply the larger framework of Theorem 2, we would say that they implicitly used a normal $\mathbf{H}_B$, which results in $\widetilde{\mathbf{H}} = \mathbf{H}_B \otimes \mathbf{H}_A$ and only included the first k_A replicates in the variance estimator. Since a subset of the replicates needed for SDR to be equivalent to SD2 is used, we say that the resultant estimator is an approximation of the SD2 estimator.

Example 4 (continued): If we only used the first four replicates of Table 3, the SDR estimator would be equivalent to (4) with

$$R = \begin{bmatrix} (y_1 - y_2)(y_8 - y_7) + (y_1 - y_2)(y_{11} - y_{12}) + (y_1 - y_2)(-y_{13} + y_{14}) + (y_8 - y_7)(y_{11} - y_{12}) \\ + (y_8 - y_7)(y_{14} - y_{13}) + (y_{11} - y_{12})(y_{14} - y_{13}) + (y_4 - y_3)(y_8 - y_7) + (y_4 - y_3)(y_{10} - y_9) \\ + (y_8 - y_7)(y_{10} - y_9) + (y_1 - y_4)(y_5 - y_6) + (y_1 - y_4)(y_9 - y_{12}) + (y_5 - y_6)(y_9 - y_{12}) \\ + (y_2 - y_3)(y_5 - y_6) + (y_2 - y_3)(y_{10} - y_{11}) + (y_2 - y_3)(y_{13} - y_{14}) + (y_5 - y_6)(y_{10} - y_{11}) \\ + (y_5 - y_6)(y_{13} - y_{14}) + (y_{10} - y_{11})(y_{13} - y_{14}) \end{bmatrix}$$

Note that R includes the same number of positive and negative terms, which do not cancel exactly, but has the result that R is usually close to zero. Similarly, using replicates 1 to $q \times k_A$, where $q = 1, 2, \ldots, k_B$, will result with an R that has an equal number of positive and negative terms. Only with all the replicates of $\widetilde{\mathbf{H}}$ will $R = 0$.

Example 5: The Current Population Survey (CPS) has a monthly sample size of roughly n = 72,000 households per month (U.S. Census Bureau 2006). CPS has a two-stage sample design, where a first-stage sample of Primary Sample Units (PSUs), which are generally counties or groups of counties, are selected and then in the second-stage households are selected within the sample PSUs. Some PSUs, generally the metropolitan areas, are selected with certainty, i.e., their first-stage probability of selection is 1.0. With the certainty PSUs, the *sys* sample can be treated as the first-stage sample design in variance estimation, i.e., SDR is applied to produce replicates. In the non-certainty PSUs, Balanced Repeated Replication (BRR) [McCarthy 1966] is applied to produce replicates. Roughly 75% of the sample or 54,000 units are in SR PSUs, where SDR is applied.

The CPS application of SDR uses a Hadamard matrix with k = 160 and excludes 2 rows, i.e., m = 158. Replicate weights are produced for 160 replicates. Although it may seem like a logical conclusion of the paper, we do not to suggest that CPS should use a Hadamard matrix of order k = 54,000 or produce 54,000 sets of replicate weights. That would result in an unreasonable number of replicates. Instead, we suggest that the subset of 160 replicates used by CPS is large,

therefore provides a reasonable approximation to SD2. Later in the empirical examples, we examine the impact of using a reduced set of replicates.

2.2 Row Assignment when $n > k$

To this point we have assumed a given RA and have not discussed how to make the RA for a given sample, where $n > k$. In this section, we review two RAs, and discuss some considerations about RAs in general. The first RA is similar to the RA described by Sukasih and Jang (2003) and is intended for use with $k < n$ and Theorem 2.

RA1: The RA assigns a pair of rows a_i and b_i to every m_d units of the sample, which we call cycle d, where $m_d \leq k$. After $m_d - 1$ cycles, the RA is repeated until all units of the sample are assigned a pair of rows.

Step 1: Sort the sample in the order in which it was sorted prior to sample selection.
Step 2: Initialize the cycle number as $d = 1$ and the number of connected loops as $c = 1$
Step 3: Start the RA at the beginning of a cycle or a connected loop as $a_1 = c$.
Step 4: Repeat the following RA: $b_i = \text{mod}(a_i + d, k)$ and $a_i = b_i$ until all m_d rows of the cycle have been used or the RA becomes a connected loop. Here, the mod or modulo function is defined as … If all m_d rows of the cycle have been used, start a new cycle: let $d = d + 1$ and go back to step 3. Otherwise (end of a connected loop, but not the end of a cycle) start a new connected loop: let $c = c + 1$ and go back to step 3.
Step 5: At the end of $d = m_d - 1$ cycles, start over with the first cycle – go back to step 2.

RA1 has the following characteristics:

- Each of the cycles $d = 1, 2, \ldots, m_d - 1$ of the RA, assigns m_d pairs of rows. This generates a total of $m_d(m_d - 1)$ pairs of rows.
- The RA repeats itself after $m_d - 1$ cycles. F&T suggest that after 10 cycles, the RA be restarted. We suggest that all $m_d - 1$ cycles be used before restarting the RA.
- The values of a_i and b_i are always c units apart.
- Halfway through the sequence, the pattern repeats itself in reverse order. If m is even, the cycles before and after the $(m_d + 1)/2^{\text{th}}$ cycle repeat themselves in reserve order.

RA1 differs from the RA of Sukasih and Jang (2003) in that we do not suggest that row 1 be skipped, suggest that the RA be repeated after 10 cycles, or require that $k - 1$ be prime. First, a row of all 1s may seem odd, but it is not a problem. Similar to a column of all 1s in **M**, which made a dead replicate, a row of all 1s will only effect the distribution of the replicate factors. The replicate factors for a unit i that is assigned row 1 (either $a_i = 1$ or $b_i = 1$) will have more replicate factors of 1.0 than otherwise. This is not wrong; it is just how the replicate factors are distributed by $\mathbf{H}_A$. The second difference is that we suggest repeating the assignment after m cycles, which is when the pattern repeat, instead of a fixed number of 10 cycles. Lastly, we do not require that k-1 be prime, but note that if $m_d = k$ -1 and k-1 is prime, then every cycle is guaranteed to have only one connected loop.

3. Empirical Examples

The questions of interest for the empirical examples are:

Q1. How well does SDR perform with a subset of all the replicates needed for SDR to be equivalent to SD?
Q2. Which row assignment is better, RA1 or RA2?

Q3. Should we use more or less connected loops?

To address these questions, we applied the SDR variance estimator to several populations. With each population, we selected a *sys* sample of size $n = 64$. Table 6 outlines the three SDR estimators we applied.

Table 6: SDR Estimators for the Empirical Examples

Estimator	k_A	$\mathbf{H}_A$	k_B	$\mathbf{H}_B$
1	4	$\mathbf{H}_{4a}$	16	$\mathbf{H}_{4a} \otimes \mathbf{H}_{4a}$
2	16	$\mathbf{H}_{4a} \otimes \mathbf{H}_{4a}$	4	$\mathbf{H}_{4a}$
3	64	$\mathbf{H}_{4a} \otimes \mathbf{H}_{4a} \otimes \mathbf{H}_{4a}$	1	1

With this construction, the SDR estimators had $k_B = 1$, 4, or 16 cycles, but all used the same $\widetilde{\mathbf{H}} = \mathbf{H}_{4a} \otimes \mathbf{H}_{4a} \otimes \mathbf{H}_{4a}$, which is the normal Hadamard matrix of order $\widetilde{k} = 64$. For the three estimators of Table 6, we also varied the row assignment (RA1 and RA2) and the number of replicates used by each estimator is either 16, 32, 48 or 64. With both RA1 and RA2, there is only one connected loop within each cycle, so estimators 1, 2, and 3 had $k_B = 16$, 4, and 1 connected loops, respectively. The results also include the SD1, SD2, and the *srswor* variance estimators for comparison purposes.

Data sets used. The 'A' populations are borrowed from Wolter's empirical example. For populations A1-A7, we generated 400 finite populations of size $N = 64{,}000$. From each population, there were $b = 100$ possible samples of size $n = 64$. The samples are indexed as $i = 1, 2, \ldots, b = 100$ and the units within each sample are indexed as $j = 1, 2, \ldots, n = 64$. Table 5 summaries how the variable of interest μ_{ij} is generated for each of the 'A' populations.

Table 7: Description of Wolter's Artificial Populations

Population	Description	n	b	μ_{ij}	e_{ij}
A1	Random	20	50	0	e_{ij} iid $N(0, 100)$
A2	Linear Trend	20	50	$i + (j-1)k$	e_{ij} iid $N(0, 100)$
A3	Stratification Effects	20	50	j	e_{ij} iid $N(0, 100)$
A4	Stratification Effects	20	50	$j + 10$	$e_{ij} = \begin{cases} \varepsilon_{ij}, & \text{if } \varepsilon_{ij} \geq -(j+10) \\ -(j+10), & \text{otherwise} \end{cases}$ ε_{ij} iid $N(0, 100)$
A5	Autocorrelated	20	50	0	$e_{ij} = \rho\, e_{i-1,j} + \varepsilon_{ij}$ $e_{i1} \sim N\left(0, 100/\left(1-\rho^2\right)\right)$ ε_{ij} iid $N(0, 100)$, $\rho = 0.8$
A6	Autocorrelated	20	50	0	same as A5 with $\rho = 0.4$
A7	Periodic	20	50	$20 \sin\{2\pi/50 [i + (j-1)k]\}$	e_{ij} iid $N(0, 100)$

In addition to using the seven 'A' populations, we also considered the loblolly pine data from Gregoire and Williams (1992). The variable of interest for the loblolly pine data is tree volume

and the sort variable is tree diameter at breast height (dbh). We used the first 14,336 trees after first sorting by dbh making $b = 224$ possible samples of size $n = 64$.

Evaluation Measures. We evaluated the different variance estimators with the three measures used by Wolter: expected relative bias (ERB), relative mean squared error (RMSE), and coverage ratios. The first measure, ERB, was used to examine the accuracy of the estimators and is defined for a specific estimator θ as

$$\mathrm{ERB}(\hat{v}_\theta) = \frac{\mathrm{E}_m\left(\mathrm{E}_p(\hat{v}_\theta - v)\right)}{\mathrm{E}_m(v)}.$$

In our notation E_p and E_m refer to the design model expectations, respectively. With the Loblolly pine data, only the design expectation applies. To examine the variance of the estimators, we also measured the RMSE, which is defined as

$$\mathrm{RMSE}(\hat{v}_\theta) = \frac{\mathrm{E}_m\left(\mathrm{E}_p(\hat{v}_\theta - v)^2\right)}{\mathrm{E}_m(v)}.$$

Results. With respect to Q1, Table A1 shows that increasing the number of replicates had minimum impact on the bias. Only with the linear trend population (A2) did the SDR estimator with four connected loops show a consistent trend in reduced bias as the number of replicates increased. The other population and estimator combinations showed no significant decreasing or increasing trend as the number of replicates increased. This is a good result, because it indicates that reducing the set of replicates does not increase the bias. As expected, the RMSEs of the estimators in Table A2 did increase as the number of replicates decreased, but surprisingly the increase was relatively minor.

When comparing RA1 and RA2 of Q2, the SDR estimator with four connected loops usually had smaller biases (Table A1) and the variances (Table A2) with RA1 as compared to RA2. With 16 connected loops, both the biases and variances were similar for both RA1 and RA2. The evidence suggests that both the bias and variance are improved, but the impact reduces as the size of the connected loops decreases.

Addressing Q3, Table A1 shows that the biases diminished with increasing number of connected loops. The exception was the periodic population (A7). When the RMSEs of SD1 and SD2 were not similar as in linear trend population (A2) and the loblolly pine data, increasing number of connected loops also reduced the RMSEs. This is not surprising. The estimator with one large connected loop is equivalent to SD2, so it will often have the largest biases and RMSEs due to the term $(\hat{y}_1 - \hat{y}_{64})^2$. In the other direction, more connected loops effectively reduces the impact of the term $(\hat{y}_1 - \hat{y}_{64})^2$, so the estimator acts more like SD1, which generally has less bias and variance than SD2.

5. Concluding Remarks

The paper provided the conditions for SDR to be equivalent to SD2 and showed how to do this when the sample size is both smaller and larger than the chosen Hadamard matrix. When a smaller Hadamard matrix $\mathbf{H}_\mathrm{A}$ is used and replicates are only derived from $\mathbf{H}_\mathrm{A}$, the paper showed how that it results in a reasonable approximation of the SD2 estimator.

The empirical examples also showed that using a reduced set of replicates is reasonable, since decreasing the number of replicates does not increase the bias of the estimates. The examples also showed that using many connected loops reduces the impact of the squared difference between the first and last unit in the sample. Since SD1 usually has larger biases and RMSEs than SD2, SDR estimators that use more rather than less connected loops will have smaller biases and RMSEs than SDR estimators with more connected loops.

References

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.

Fay, R.E. (1989). "Theory and Application of Replicate Weighting for Variance Calculations," Joint Statistical Meetings, Proceedings of the Section on Survey Research Methods, 212-217.

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," Joint Statistical Meetings, Proceedings of the Section on Government Statistics, 154-159.

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

Gregoire, T.G. and Williams, M. (1992). "Identifying and Evaluating the Components of Non-Measurement Error in the Application of Standard Volume Equations," Journal of the Royal Statistical Society, Series D. Vol. 41, No. 5. 509-518.

Hedayat, A. and Wallis, W.D. (1978). "Hadamard Matrices and their Applications," The Annuals of Statistics, 6, 1184-1238.

Iachan, R. (1982). "Systematic Sampling: A Critical Review," International Statistical Review, 50, p. 293-303.

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, 166-171.

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

Sukasih, A.S. and Jang, D. (2003). "Monte Carlo Study on the Successive Difference Replication Method for Non-Linear Statistics," Joint Statistical Meetings, Proceedings of the Section on Survey Research Methods.

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

Yates, F. (1946). *Sampling Methods for Censuses and Surveys*, Hafner Publishing Company, New York, NY.

U.S. Census Bureau (2006). Technical Paper 66, "Design and Methodology: Current Population Survey," October 2006.

Table A1: Expected Relative Bias

Population	SD1	SD2	SRSWOR	k_A	RA	SDR with… 16 replicates	32 replicates	48 replicates	64 replicates
A1	0.009	0.009	-0.001	4	1	0.010	0.009	0.009	0.009
					2	0.010	0.010	0.010	0.009
				16	1	0.009	0.008	0.010	0.009
					2	0.009	0.010	0.010	0.009
				64	1 or 2	0.009	0.009	0.010	0.009
A2	-0.960	1.417	25.317	4	1	-0.696	-0.840	-0.888	-0.907
					2	-0.538	-0.768	-0.845	-0.883
				16	1	0.113	-0.270	-0.500	-0.615
					2	1.302	0.152	-0.231	-0.423
				64	1 or 2	1.302	1.379	1.404	1.417
A3	0.015	0.327	3.462	4	1	0.049	0.031	0.025	0.021
					2	0.070	0.040	0.030	0.025
				16	1	0.155	0.105	0.075	0.060
					2	0.314	0.163	0.112	0.086
				64	1 or 2	0.314	0.324	0.327	0.327
A4	0.006	0.305	3.284	4	1	0.040	0.023	0.017	0.014
					2	0.060	0.030	0.021	0.017
				16	1	0.144	0.095	0.066	0.052
					2	0.291	0.146	0.098	0.075
				64	1 or 2	0.291	0.299	0.303	0.305
A5	0.064	0.065	0.055	4	1	0.063	0.063	0.063	0.065
					2	0.068	0.066	0.066	0.065
				16	1	0.063	0.063	0.063	0.065
					2	0.065	0.067	0.066	0.066
				64	1 or 2	0.065	0.066	0.066	0.065
A6	0.093	0.095	0.084	4	1	0.093	0.092	0.093	0.094
					2	0.092	0.096	0.095	0.094
				16	1	0.099	0.095	0.094	0.094
					2	0.093	0.094	0.094	0.093
				64	1 or 2	0.093	0.096	0.095	0.095
A7	0.112	0.115	20.641	4	1	0.105	0.069	0.112	0.253
					2	0.004	0.004	0.073	0.310
				16	1	0.177	0.168	0.462	0.847
					2	0.002	0.003	0.027	1.248
				64	1 or 2	0.002	0.003	0.030	0.115
Loblolly	-0.294	0.966	5.384	4	1	0.165	0.107	0.111	0.099
					2	0.198	0.123	0.125	0.105
				16	1	0.176	0.139	0.133	0.124
					2	0.960	0.644	0.632	0.535
				64	1 or 2	0.960	0.954	0.964	0.966

Table A2: Relative Mean Squared Errors

Population	SD1	SD2	SRSWOR	k_A	RA	SDR with… 16 replicates	32 replicates	48 replicates	64 replicates
A1	0.049	0.049	0.032	4	1	0.176	0.091	0.066	0.054
					2	0.176	0.095	0.064	0.048
				16	1	0.141	0.080	0.059	0.048
					2	0.194	0.096	0.065	0.049
				64	1 or 2	0.194	0.096	0.064	0.049
A2	0.921	2.008	640.916	4	1	0.485	0.706	0.789	0.823
					2	0.290	0.590	0.714	0.780
				16	1	0.013	0.073	0.250	0.378
					2	1.695	0.023	0.054	0.179
				64	1 or 2	1.695	1.901	1.972	2.008
A3	0.049	0.176	12.203	4	1	0.195	0.095	0.068	0.054
					2	0.222	0.103	0.067	0.050
				16	1	0.207	0.106	0.070	0.055
					2	0.374	0.144	0.085	0.061
				64	1 or 2	0.374	0.245	0.199	0.176
A4	0.057	0.170	11.109	4	1	0.192	0.104	0.077	0.063
					2	0.217	0.110	0.075	0.058
				16	1	0.208	0.109	0.077	0.063
					2	0.357	0.144	0.090	0.067
				64	1 or 2	0.357	0.232	0.191	0.170
A5	0.056	0.056	0.039	4	1	0.192	0.106	0.076	0.063
					2	0.217	0.111	0.075	0.057
				16	1	0.161	0.093	0.068	0.057
					2	0.214	0.111	0.075	0.056
				64	1 or 2	0.214	0.110	0.074	0.056
A6	0.065	0.066	0.046	4	1	0.211	0.117	0.088	0.072
					2	0.229	0.120	0.086	0.067
				16	1	0.185	0.107	0.080	0.067
					2	0.226	0.117	0.085	0.067
				64	1 or 2	0.226	0.118	0.084	0.066
A7	0.063	0.062	427.141	4	1	0.219	0.106	0.091	0.143
					2	0.187	0.098	0.079	0.175
				16	1	0.229	0.137	0.351	0.828
					2	0.187	0.097	0.065	1.689
				64	1 or 2	0.187	0.097	0.065	0.062
Loblolly	0.271	2.477	32.618	4	1	0.977	0.830	0.847	0.827
					2	1.020	0.830	0.862	0.834
				16	1	0.843	0.848	0.829	0.840
					2	2.632	1.700	1.717	1.501
				64	1 or 2	2.632	2.412	2.439	2.477