Comparison of Variance Estimators for Self-Representing Primary Sample Units November 2024

Stephen Ash
Bureau of Labor Statistics, 2 Massachusetts Ave NE, Washington, DC 20002

Abstract
Many surveys estimate variances with the balance repeated replication (BRR) variance estimator. With the self-representing (SR) Primary Sample Units (PSUs), surveys sometimes split them into parts which are then paired into pseudo strata and then BRR is applied to the pseudo strata. However, there is not much guidance on the number of pseudo strata to split the SR strata into or how (or if) the sort order should be used to split the sample when the sample was selected with systematic random sampling. Our research considered twelve different applications of the BRR variance estimators that varied by the number of pseudo strata formed and by how the sort order of a systematic random sample was used to split the PSU. We also included variations of the delete-a-group jackknife and successive difference replication variance estimators. Using simulations involving data from the Consumer Expenditures Survey, we found that the BRR variance estimator that split the sample of the SR PSUs into the most replicates possible and split the sample using the sort order was the best overall variance estimator for both national-level estimates and individual PSU-level estimates.

Key Words: Variance estimation, self-representing strata, balanced-repeated replication, delete-a-group jackknife, successive difference replication.

1. Introduction

With a two-stage sample design for large household surveys, the first stage sample design often includes self-representing (SR) strata and non-self-representing strata (NSR), where the probability of selection for the Primary Sample Units (PSUs) of the SR strata is equal to 1.0 and the probability of selection for the PSUs of the NSR strata is less than 1.0. This means that the SR strata are comprised of one SR PSU, and it is in the first-stage sample with certainty and the NSR strata are comprised of more than one PSU and not all of them are selected. Then, within the sample PSUs of both the SR and NSR strata, a second-stage sample of households is selected. Within the NSR strata, the sample variance is due to the selection of PSUs and the selection of households, but with the SR strata the variance is only due to the selection of households.

Many surveys estimate the variances of this type of two-stage sample design with the balanced repeated replication (BRR) variance estimator, which is also referred to as balanced half sample variance estimator (McCarthy 1966, 1969a, 1969b). The BRR variance estimator was designed to estimate the variances for NSR strata that select two PSUs per strata, where the two PSUs are referred to as half samples. However, BRR can also be used with sample designs that select one PSU per strata, collapse the strata into pseudo strata with two sample units (or two half samples), and then apply BRR to the pseudo strata as if the pseudo strata were strata with two sample units. See Judkins (1990), Wolter (2007), and Ash (2022) for more background on the adaptation of BRR to collapsed strata.

For the SR strata, there are different replication variance estimators that are applied to estimate the variance that are compatible with BRR as used with the NSR strata. We say that a combination of replicate variance estimators is compatible, if we can produce replicate factors for the SR and NSR strata that are used with the same expression of the variance estimator. Compatibility is an important property of a variance estimator when we want a single expression for the replicate variance estimator that can be used to simultaneously estimate the variance of the SR and NSR strata. If the two replicate variance estimators for the SR and NSR strata are not compatible, the overall variance needs to be estimated with separate expressions, and then the estimates from those separate expressions need to be added together to get the total variance, instead of using a single expression.

One replication variance estimator that is compatible with BRR is BRR itself. To estimate the variance of the SR strata, some surveys apply BRR to pseudo strata that are formed by either splitting the second-stage sample of the SR PSUs into two even half samples or treating a pair of SR strata as a single pseudo stratum, where each SR PSU is a half sample (Lee and Forthofer 2006). Applications of these two strategies are described by several authors. Nixon et al. (1998) describes how the National Health Interview Survey (NHIS) splits the largest SR PSUs into two half samples (one pseudo stratum) and pairs the smallest SR PSUs into pseudo strata. Guciardo et al. (2004) describe how the Bureau of Labor Statistics' National Compensation Survey, a survey of business establishments, splits the establishments selected within their SR strata into two half samples (one pseudo stratum) and then applies BRR. Johnson and Rust (1992) describe how the National Assessment of Educational Process, a survey of schools, assigned the largest SR schools to two pseudo strata and the smallest SR schools treated as one pseudo stratum. The schools were then assigned to one of two half-samples within the pseudo strata, with equal probability and systematically.

Although we found descriptions of applications that either split or paired SR strata into pseudo strata, we did not find any theoretical justification of either strategy. Our paper addresses this gap. We show how splitting the sample into half samples of one or more pseudo strata and using the sort order to split the sample from a systematic random sample from an ordered list (SYS) sample design produces a type of collapsed-strata variance estimator.

We also consider the impact of splitting the SR strata into more than one pseudo strata. Since the number of pseudo strata can range from one to the number of replicates R used by BRR, why not split the SR strata into R pseudo strata? Splitting the sample of the SR strata into an increasing number of pseudo strata should decrease the variance of the variance estimator because it increases the number of replicates for the SR strata, and increasing the number of replicates reduces the variance of the variance estimator. Because we suggest increasing the number of pseudo strata, we do not further consider the strategy of pairing SR strata since it results in fewer replicates than splitting SR PSUs.

We compare splitting the SR strata into a varying number of pseudo strata with a simulation study using data from the Consumer Expenditures Survey (CE). The goal of the simulation is to compare variance estimators with varying number of pseudo strata with respect to bias and variance and determine which is best suited for CE's national estimates and Metropolitan Statistical Area (MSA)-level estimates. Note that all of CE's 23 SR strata are an MSA, which means that we want to know how well the variance estimators perform for the national estimates (all 23 MSA together) and for the 23 separate MSA-level estimates. The simulation study also compares the impact of using the sort order to split the SR strata into pseudo strata with the simpler method of a random assignment.

Our simulation study also includes other replication variance estimators that are compatible with BRR including five variations of the successive difference replication (SDR) estimator as described by Fay and Train (1995) and Ash (2014) and two variations of the delete-a-group jackknife (DAGJK) as described by Kott (1998, 2001).

The simulation study shows that splitting the SR strata into an increasing number of pseudo strata reduces the bias of the variance estimator. This suggests that CE could improve its variance estimation procedures by splitting each of its SR strata into 44 pseudo strata rather than just one, as it currently does. The improvements in the variance of the variance estimator impact both the national and Metropolitan Statistical Area (MSA) estimates but were greater with the MSA-level estimates. The simulation study also shows that the BRR variance estimator that split the sample into 44 pseudo strata and the SDR variance estimator were the two best variance estimators overall.

The paper is organized as follows. Section 2 describes the variance estimators under consideration and further discusses compatible combinations of replication variance estimators. Section 3 reviews the CE sample design and how we produced the complete universe using past CE survey data, section 4 describes the results of our simulation, and section 5 provides our conclusions.

2. Variance Estimators

The first part of this section reviews the replication variance estimators that can be used to estimate the variance of NSR and SR strata and are included in our simulation study. The second part of this section reviews several combinations of compatible variance estimators and specifically, variance estimators for SR strata that are compatible with the BRR variance used with NSR strata.

2.1 Review of Variance Estimators

2.1.1 BRR Variance Estimator for Non-Self-Representing Collapsed Strata

We define the BRR variance estimator as generally as possible using Fay's method of BRR (Dippo, Fay, and Morganstein 1994) and adapted to collapsed strata (Judkins 1990). We define the total of a variable y_k for a two-stage sample design as: $Y = \sum_h \sum_{i \in U_h} Y_{hi}$ and $Y_{hi} = \sum_{k \in U_{hi}} y_k$, where h is an index on the first-stage strata, U_h is the first-stage universe of stratum h, i is an index on the PSUs in first-stage strata h, and U_{hi} is the second-stage universe of PSU i of stratum h. The mean of a variable y_k is a ratio of two totals: $\bar{Y} = Y/N$, where $N = \sum_h \sum_{i \in U_h} N_{hi}$ and $N_{hi} = \sum_{k \in U_{hi}} 1$. For a two-stage sample design, the estimator of Y is: $\hat{Y} = \sum_h \sum_{i \in s_h} w_{hi} \hat{Y}_{hi}$ and the estimator of the PSU total Y_{hi} is: $\hat{Y}_{hi} = \sum_{k \in s_{hi}} w_k y_k$, where s_h is the first-stage sample of PSUs in stratum h, s_{hi} is the second-stage sample of eligible units for PSU i in stratum h, w_{hi} is the first-stage weight for PSU i in stratum h, and w_k is the second-stage weight for unit k. The survey weights are defined as the inverse of the probability of selection for each stage of the sample design or $w_{hi} = \pi_{hi}^{-1}$ and $w_k = \pi_k^{-1}$, where the first- and second-stage probabilities of selection are defined as or $\pi_{hi} = P(i \in s_h)$ and $\pi_k = P(k \in s_{hi})$, respectively. The BRR/CS variance estimator of the variance of $\hat{Y}$ that is adapted for two collapsed strata is:

$$\hat{v}_{BRR/CS}(\hat{Y}) = \frac{1}{R(\kappa-1)^2} \sum_{r=1}^{R} \left(\hat{Y}_r - \hat{Y}\right)^2,$$

where r is an index on the replicates, R is the number of replicates, and the replicate estimator of Y can be alternatively defined as $\hat{Y}_r = \sum_{g=1}^{B} \left(\hat{Y}_{rg} \ + \hat{Y}_{rg} \ \right)$, where the estimator of the stratum total Y_{gh} for the r^{th} replicate is $\hat{Y}_{rgh} = \sum_{k \in s_{ghi}} w_{hi} Y_{rghi}$ and the estimator of the PSU total Y_{ghi} for the r^{th} replicate is $\hat{Y}_{rghi} = \sum_{k \in s_{ghi}} F_{rg} \ w_k y_k$.

With the BRR/CS variance estimator, g is the index on the pseudo strata, and B is the number of pseudo strata. The paired strata within a pseudo stratum are also referred to as half samples and this means that h is still an index on the strata, but it is used interchangeably with the term half sample. Following Judkins (1990), the replicate factor of half sample $h = 1$, pseudo strata g, and replicate r is defined as:

$$F_{rg}^{(BRR/CS)} = 1 + 2a_{rg}(1-\kappa)P_{g2}$$

and the replicate factor of half sample $h = 2$, pseudo strata g, and replicate r is defined as:

$$F_{rg}^{(BRR/CS)} = 1 - 2a_{rg}(1-\kappa)P_{g1},$$

where κ has values $0 \leq \kappa < 1$, $0 < P_{gh} < 1$ and $P_{g1} + P_{g2} = 1$. The variable P_{gh} is customarily defined as $P_{gh} = MOS_{gh}/\left(MOS_{g1} + MOS_{g2}\right)$, where MOS_{gh} is defined as the measures of size for stratum h of pseudo strata g and is used to reduce the bias of the collapsed-strata variance estimator. Ash (2022) provides different solutions of P_{gh} that minimize the bias and the square of the bias of $\hat{v}_{BRR/CS}(\hat{Y})$. The values a_{rg} come from a Hadamard matrix of dimension *R,* where a_{rg} is the value of the r^{th} row and g^{th} column of the Hadamard matrix. The replicate factors are assigned to the units of the sample in PSU i by noting that $F_{ri} = F_{rgh}$, if $i \in s_{gh}$ which means that if PSU i is known, then the pseudo strata g and half sample h are known too because each variance estimator assumes a specific assignment of the PSUs to the pseudo strata and half samples.

CE's application of BRR uses a Hadamard matrix with dimension $R = 44$ which produces 44 sets of replicate weights for variance estimation. Because CE also uses $\kappa = 1$, and $P_{g2} = ½$, CE's variance estimator is the same as BRR as originally suggested by McCarthy (1966, 1969a, 1969b) with replicate weights that are either 0 or 2.

2.1.2 BRR Variance Estimator Applied to Self-Representing Strata

As discussed in the introduction, the BRR variance estimator is sometimes used to estimate the variance of SR strata by splitting the SR strata into two half samples, treating the two half-samples as from one pseudo strata, and then applying BRR to the pseudo strata and half samples. In this section, we define a more general set-up where we split each SR stratum into $2G$ half sample and pseudo strata combinations, where G is the number of pseudo strata ($1 \leq G \leq R$) and each pseudo strata has 2 half samples in it. Each half sample and pseudo strata combination should have an approximately even number of sample units assigned to it. Since the first-stage strata are not functioning as pseudo strata (as with BRR applied to NSR strata), we add an index b for the half sample.

The BRR variance estimator that splits the sample into G pseudo strata is $\hat{Y}_r = \sum_h \sum_{g=1}^{G}\left(\hat{Y}_{rgh1} + \hat{Y}_{rgh2}\right)$, where the estimator of the stratum total for replicate r, strata h, pseudo strata g, and half sample b, is $\hat{Y}_{rghb} = \sum_{k \in s_{ghbi}} w_{hi} Y_{rghb}$. The replicate factor of replicate *r,* strata h, pseudo strata g, and half sample $b = 1$ is:

$$F_{rhg1}^{(BRR/SR)} = 1 + a_{rg}(1 - \kappa)$$

and the replicate factor of replicate *r,* strata h, pseudo strata g, and half sample $b = 2$ is:

$$F_{rhg2}^{(BRR/SR)} = 1 - a_{rg}(1 - \kappa)$$

The term P_{ghb} is not in our expressions for SR replicate factors $F_{rhgb}^{(BRR/SR)}$ because the half samples within a given pseudo stratum have approximately an equal number of second-stage sample units.

Collapsed-strata variance estimator. We now explain how to split the SYS sample from an SR PSU so that it makes a collapsed-strata variance estimator. This idea is not new: similar reasoning has been suggested for other variance estimators of the SYS sample design by Hansen, Hurwitz, and Maddow (1953; Volume 1, Chapter 11, Section 8), Wolter (2007; p. 336), and Megill et al. (1987). First, consider the implicit strata defined by the SYS sample design which are defined by the length of the sampling interval. They can be thought of as implicit strata because one unit is selected within each of the implicit stratum defined by the length of the sampling interval. Figure 1 provides a representation of a SYS sample and the implicit strata with a sampling interval of length 6.

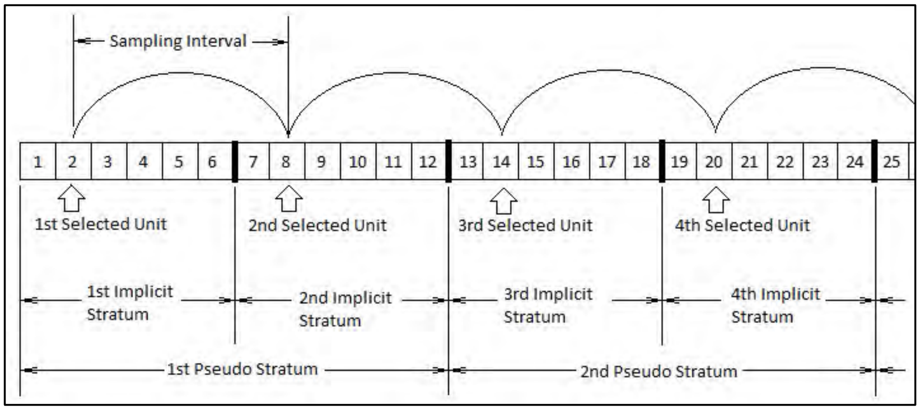

Figure 1: Representation of Systematic Random Sampling and Implicit Strata

In Figure 1, the sampling interval of six selects every 6th unit of the sorted list. This also means that every six units define a different implicit stratum, where one unit is selected within each implicit stratum. The 2nd, 8th, 14th, and 20th units of the universe are one possible sample, where the 2nd unit was randomly selected within the first implicit stratum and the subsequent selections follow from the first selection with each subsequent selection located one sampling interval from the previous selection.

We think of splitting the sample into half samples and pseudo strata as a making a collapsed-strata estimator by collapsing the 1st and 2nd sample units (or the 1st and 2nd implicit strata) into the 1st pseudo stratum, collapsing the 3rd and 4th selected units into the 2nd pseudo strata, etc. The odd numbered selections are assigned to the 1st half sample, and the even numbered selections are assigned to the 2nd half sample. Then, each pseudo stratum has two units allowing us to estimate the variance within each pseudo strata.

Since CE is limited to only having R = 44 replicates and not $n/2$ replicates (for the $n/2$ possible pseudo strata, where n is the second-stage sample size within a SR PSU), and we "reuse" the 44 replicates for multiple pseudo strata: $1 \leq G \leq R$. With our simulation, we used the G = 1, 2, 4, 22, and 44 pseudo strata.

An alternative way of splitting the sample into pseudo strata is to consider the larger implicit strata defined by the variables used in the sort order of the SYS sample design. For example, CE sorts the universe by state/county/STRATUM and then selects the sample using SYS. This ensures that the sample has approximately a proportional number of sample units from each state/county. With this way of thinking, we suggest that each state/county be treated as implicit strata from which we can split into two half samples. With the SYS sample design, we can assign the even number units of the ordered sample to one of the half samples of the pseudo strata and the odd number units to the other half sample. Since CE's sort order uses both state/county and the variable STRATUM, this method of splitting the sample into pseudo strata could also be applied to the variable STRATUM.

2.1.3 SDR Variance Estimator

We include SDR in our simulation study because several large demographic surveys use it to estimate the variance of the SR strata including the Current Population Survey (U.S. Census Bureau 2019), National Crime Victimization Survey (Bureau of Justice Statistics 2014, 2020), and the American Housing Survey (U.S. Census Bureau and Department of Housing and Urban Development 2022). We begin our discussion of SDR by beginning with its motivation – the successive difference (SD) variance estimator, which Yates (1953) and Wolter (2007) expressed as:

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

As discussed by Ash (2014), $\hat{v}_{SD1}(\hat{Y})$ is a collapsed-strata variance estimator, where it collapses all possible pairs of adjacent implicit strata. A second form of the SD variance estimator "connects" the first implicit stratum of the sort order with the last implicit stratum, or:

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

Fay and Train (1995) suggested the SDR variance estimator, a replication form of the SDR, which is expressed as:

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

where the replicate estimator for group g is defined as: $\hat{Y}_r = \sum_h \sum_{k\in s_{hi}} w_{hi} w_k F_{rk}^{(SDR)} y_k$, and the replicate factors are defined as:

$$F_{rk}^{(SDR)} = 1 + 2^{-\frac{3}{2}} a_{b_k,r} - 2^{-\frac{3}{2}} a_{c_k,r},$$

The $a_{b_k,r}$ represents the b_k th row and rth column of a Hadamard matrix and b_k and c_k are an assignment of two rows of the Hadamard matrix to each unit *k*. The key to the row assignment b_k and c_k is that it uses the sort order of the SYS sample and the assignment "connects" adjacent units in the sort order. For example, the row assignments of the first four units of the SYS sample (c_1, d_1), (c_2, d_2), (c_3, d_3), and (c_4, d_4) are connected when $d_1 = c_2$, $d_2 = c_3$, $d_3 = c_4$, and $d_4 = c_1$. Ash (2014) says that this example row assignment is a connected loop because the last unit is linked back to the first thus completing a loop. See Sukasih and Jang (2003) and Ash (2014) for further discussion of the row assignment.

We add that the variance estimator formed by splitting the SR PSUs and the SDR variance estimator are both a type of collapsed-strata variance estimator – see Ash (2014) for an explanation of how successive differences and SDR variance estimators are collapsed-strata variance estimators. With SDR, all possible $(n-1)$ pairs of implicit strata [(1,2), (2,3), (3,4),…] are used by the variance estimator, while the variance estimator that splits the SR strata into pseudo strata only uses $n/2$ pairs of implicit strata [(1,2), (3,4), (5,6),…]. This suggests that SDR could be a better variance estimator since it uses more collapsed strata.

2.1.4 Delete-a-Group Jackknife Variance Estimator

Kott (1998, 2001) expressed the delete-a-group jackknife (DAGJK) variance estimator of $\hat{Y}$ as:

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

The units of the sample are evenly and randomly assigned to the *R* groups and *r* is an index on the groups. The replicate estimator for group *r* is defined as: $\hat{Y}_r = \sum_h \sum_{i\in s_{hi}} w_{hi} w_k F_{rk}^{(DAGJK)} y_k$, the DAGJK replicate factor for group *r* and sample unit *k* is defined as $F_{rk}^{(DAGJK)} = \big(R/(R-1)\big) I_{rk}$ and we define the following indicator variable for sample unit *k* and group *r*, which controls the group to delete – $I_{rk} = 1$, if $k \notin r$, and $I_{rk} = 0$, if $k \in r$.

Although Kott (2001) suggested assigning the groups at random, we consider whether the assignment of groups should use the sort order when a SYS sample is selected, especially with SYS is used with a highly informative sort order. We did this by assigning the 1st unit of the sort order to the 1st group, the 2nd unit to the 2nd group, …, the Rth unit to the R^{th} group, and then repeat the pattern again with the $(R+1)$th unit to the 1st group, the $(R+2)$th unit to the 2nd group, etc.

2.2 Compatible Replication Variance Estimators

We next review three different compatible combinations (CC) of replication variance estimators, which are summarized in Table 1. For each CC, the separate replication variance estimators for the SR and NSR strata are represented in the bottom two rows of Table 1 in terms of their replicate factors and both replicate factors use the replicate variance estimator given at the top of Table 1.

Table 1: Summary of Compatible Combinations of Replication Variance Estimators and their Replicate Factors

Replicate Variance Estimator		1. BRR and BRR/CS	2. SDR and BRR/CS	3. DAGJK and BRR/CS
		$\frac{1}{R(1-\kappa)^2}\sum_{r=1}^{R}(\hat{Y}_r-\hat{Y})^2$	$\frac{4}{R}\sum_{r=1}^{R}(\hat{Y}_r-\hat{Y})^2$	$\frac{R-1}{R(1-\kappa)^2}\sum_{r=1}^{R}(\hat{Y}_r-\hat{Y})^2$
Replicate Factors	SR Strata	BBR by splitting or pairing SR strata $F_{rgh}^{(BRR/SR)} = 1 + a_{rg}(1-\kappa)$ $F_{rgh2}^{(BRR/SR)} = 1 - a_{rg}(1-\kappa)$	SDR $F_{rk}^{(SDR)} = 1 + 2^{-\frac{3}{2}}a_{c_k,r} + 2^{-\frac{3}{2}}a_{d_k,r}$	DAGJK $F_{rk}^{(DAGJK/CC)} = (1-\kappa)\left(\frac{R}{R-1}\right)I_{kr} + \kappa$
	NSR Strata	BRR/CS $F_{rg1}^{(BRR/CS)} = 1 + 2a_{rg}(1-\kappa)P_{g2}$ $F_{rg2}^{(BRR/CS)} = 1 - 2a_{rg}(1-\kappa)P_{g1}$	BRR/CS with $\kappa = 1/2$ $F_{rg1}^{(BRR/CS)} = 1 + a_{rg}P_{g2}$ $F_{rg2}^{(BRR/CS)} = 1 - a_{rg}P_{g1}$	BRR/CS $F_{rg1}^{(BRR/CC3)} = 1 + 2a_{gr}(1-\kappa)(R-1)^{-\frac{1}{2}}P_{g2}$ $F_{rg2}^{(BRR/CC3)} = 1 - 2a_{gr}(1-\kappa)(R-1)^{-\frac{1}{2}}P_{g1}$

BRR/CS has R replicates and DAGJK has R groups

2.2.1 BRR and BRR/CS

The simplest CC of replicate variance estimators the application of BRR to both the NSR and SR strata. This CC uses the expression $\hat{v}_{BRR/CS}(\hat{Y})$, where the NSR strata use replicate factors $F_{rhg}^{(BRR/CS)}$ and the NSR strata use replicate factors $F_{rhgb}^{(BRR/SR)}$.

2.2.2 SDR and BRR/CS

As mentioned earlier, several major household surveys use the combination of BRR with $\kappa = 1/2$ and SDR together to estimate the variance of the NSR and SR strata, respectively. Both the BRR and SDR variance estimators can use the same expression because $\hat{v}_{SDR}(\hat{Y}) = \hat{v}_{BRR/CS}(\hat{Y})$, when $\kappa = 1/2$.

2.2.3 DAGJK and BRR/CS

We derived a BRR version of the DAGJK variance estimator that is compatible with the BRR/CS variance estimator for the NSR strata. The expression of the variance estimator and its replicate factors are provided in Table 1 and Result A1 of the Appendix shows that using the replicate factors $F_{ir}^{(DAGJK/CS)}$ with variance estimator in Table 1 is equivalent to $\hat{v}_{DAGJK}(\hat{Y})$, and using the replicate factors $F_{ir}^{(BRR/CS2)}$ with variance estimator in Table 1 is equivalent to $\hat{v}_{BRR/CS}(\hat{Y})$.

3. Data for the Simulation

This section describes the CE Survey and how we used CE data to create a universe for our simulation study.

3.1 CE Background

CE employs a two-stage sample design. In the first stage, both the Interview and Diary Surveys use the same first-stage sample design, which is selected once every ten years and shortly after the decennial census. In the first stage, the counties of the U.S. are grouped into Primary Sample Units (PSUs), which can be a single county or a group of contiguous counties. In urban areas, PSUs were defined as the Core-Based Statistical Areas (CBSA) as defined by the Office of Management and Budget. In rural areas, where counties are not grouped into CBSAs, PSUs were formed by grouping counties so that each PSU had a minimum of 7,500 or more people and a maximum of 3,000 square miles (King 2012). The PSUs were then grouped into strata that are either self-representing (SR) or non-self-representing (NSR). In the first stage, one PSU was selected within each stratum with probability proportion to the measure of size (MOS), where the MOS is the most recent decennial census population estimates. The current first-stage sample, which we refer to as the 2010 sample design, has a total of 91 sample PSUs, where 23 are SR and 68 are NSR.

In the second stage of the sample design, the Interview and Diary Surveys both select separate equal probability SYS samples of addresses. To do this, the address frame, maintained by the Census Bureau, is sorted and a SYS sample of addresses is selected from the frame within each of the first-stage sample PSUs. The address frame is sorted by state/county and STRATUM, a CE specific stratum code. For more details about CE's overall sample design and estimation methods, see Neiman et al. (2015) and Bureau of Labor Statistics (2024).

3.2 Simulation Universe

The goal of the simulation is to compare alternative variance estimators of the estimate of mean total expenditures with respect to the bias, variance, MSE, and the accuracy of the confidence intervals that are produced with the variance estimators. To measure the bias and the MSE, we need to calculate the actual variance and to do this we needed a universe, where total expenditures was known for every CU in the universe.

To create our universe, we started with the 122,782 eligible CUs available from all 23 SR strata of the 2016-2022 Interview survey and enlarged it by creating 25 CUs for each of the 122,782 eligible CUs. The 25 CUs had the same sort variables, but they had a different modeled value of total expenditures – we added a random term to the model prediction corresponding to the standard error of the model. The model included variables related to tenure, property value/rent, number of people in the CU, and urban/rural status. At the end, we produced a universe of 3M+ CUs, where we know the sort variables and total expenditures for every CU.

4. Simulation Study

4.1 Simulation Details

With the universe described in section 3, we selected 20,000 SYS random samples for the three different sort orders: a random sort, CE's sort order, and sorting by the total expenditures of the CU, where the least informative sort order was a random sort order and the most informative was the sorting by total expenditures. In practice, the total expenditures of a CU is not available when the second-stage sample is selected, but we used it to select our simulation samples in order to examine the extreme case of having a "perfect" sort order, where the universe is sorted by the survey's variable of interest. CE's approximate sort order is in between the extremes in terms of being an informative sort order.

Table 2 lists the alternative variance estimators examined with the simulation that are applied to each of the 20,000 simulated samples within each of the eight sort orders.

Table 2: Variance Estimators of the Simulation

Variance Estimator	Description
$\hat{v}_{srswor}(\hat{Y})$	*1-SRSWOR
$\hat{v}_{SD1}(\hat{Y})$	*2-SD1 | sort order | 0 loops
$\hat{v}_{BBR}(\hat{Y})$	01-BRR | sort order | 1 pseudo stratum
	02-BRR | sort order | 2 p-strata
	03-BRR | sort order | 4 p-strata
	04-BRR | sort order | 22 p-strata
	05-BRR | sort order | 44 p-strata
	06-BRR | random order | 1 p-stratum
	07-BRR | random order | 2 p-strata
	08-BRR | random order | 4 p-strata
	09-BRR | random order | 22 p-strata
	10-BRR | random order | 44 p-strata
	11-BRR | random order | split county
	12-BRR | random order | split strata
$\hat{v}_{SDR}(\hat{Y})$	13-SDR | sort order | 0 loops
	14-SDR | sort order | 4 loops
	15-SDR | sort order | 8 loops
	16-SDR | sort order | 16 loops
	17-SDR | sort order | 44 loops
$\hat{v}_{DAGJK}(\hat{Y})$	18-DAGJK | sort order | 44 groups
	19-DAGJK | random order | 44 groups

The first two variance estimators $\hat{v}_{srswor}(\hat{Y})$ and $\hat{v}_{SD1}(\hat{Y})$ are labeled '*1' and '*2' in Table 2 and are the SRSWOR and SD1 variance estimators, respectively. We include an asterisk in their labels because, as mentioned previously, they are not a replication variance estimator or compatible with BRR, but we included them because they provide simple reference estimates that we can compare with the other replication variance estimators.

The BRR variance estimators $\hat{v}_{BBR}(\hat{Y})$ that split the sample into half samples and pseudo strata are labeled 01 to 12 in Table 2. The BRR variance estimators 01 to 10 vary by how many pseudo strata the SR strata is split into (1, 2, 4, 22, or 44) and by whether the units of the sample are split into half samples either randomly or by using the sort order. The number of pseudo strata ranged from the smallest possible or 1 to the largest possible or 44, the number of replicates that CE uses.

The BRR variance estimators labeled 11 and 12 split the CUs into half samples within state/county or STRATUM, respectively. The idea is that the individual variables of the sort order, like CE's state/county and STRATUM sort variables, define implicit strata when used with the SYS sample design. Therefore, it made sense to estimate the variance within the implicit strata formed by these variables, which means treating the different values of the sort variables as implicit strata (pseudo strata) and splitting them into two half samples. With this strategy, variance estimator 11 treated each state/county as a pseudo stratum, split it into two half samples, and then spread the state/counties across the 44 replicates. Similarly, variance estimator 12 treated the values of STRATUM as pseudo strata, split them into two half samples, and spread the values of STRATUM across 44 replicates. Although splitting the implicit strata seemed reasonable, it was not simple to enact because the sample counts within each PSU by either state/county or STRATUM were sometimes very small or even zero. For this reason, we collapsed some values of the state/county and STRATUM on an ad hoc basis prior to splitting them into half samples.

The SDR variance estimators $\hat{v}_{SDR}(\hat{Y})$ are labeled 13 to 17 in Table 2 and varied by the number of CUs in a connected loop – either 0, 4, 8, 16, or 44. The DAGJK variance estimators $\hat{v}_{DAGJK}(\hat{Y})$ are labeled 18 and 19 in Table 2 and varied by whether the units of the sample were assigned to the 44 groups either randomly or by using the sort order.

With each of the $n_{sim} = 20{,}000$ simulated samples that are indexed by t, we estimated the mean total expenditures for all of the SR PSUs (national estimate) and for each of the SR PSUs (MSA-level estimate) as $\hat{\bar{Y}}_t = \hat{Y}_t / N$, where $\hat{Y}_t$ is the estimator of total expenditures. We also estimated the variance of $\hat{\bar{Y}}_t$ with the 19 variance estimators, which is represented as $\hat{v}\left(\hat{\bar{Y}}_t\right)$.

4.2 Simulation Results

We defined the simulation expectation of the estimator of the mean $\hat{\bar{Y}}$ as $E_{sim}\left(\hat{\bar{Y}}\right) = n_{sim}^{-1} \sum_{t=1}^{n_{sim}} \hat{\bar{Y}}_t$, the simulation variance of $\hat{\bar{Y}}$ as: $v_{sim}\left(\hat{\bar{Y}}\right) = n_{sim}^{-1} \sum_{t=1}^{n_{sim}} \left(\hat{\bar{Y}}_t - \bar{Y}\right)^2$, and simulation standard error of $\hat{\bar{Y}}$ as $se_{sim}\left(\hat{\bar{Y}}\right) = \sqrt{v_{sim}\left(\hat{\bar{Y}}\right)}$. The simulation coefficient of variation is defined as a percent as $cv_{sim}\left(\hat{\bar{Y}}\right) = \left(se_{sim}\left(\hat{\bar{Y}}\right) / \bar{Y}\right) \times 100$. Table 3 presents the simulation standard errors and coefficients of variation for the combination of all SR PSU by the nine different sort orders.

Table 3: Standard Errors of Estimated Mean Total Expenditures for Each Sort Order

Sort Order	All SR PSUs	
	$se_{sim}\left(\hat{\bar{Y}}\right)$	$cv_{sim}\left(\hat{\bar{Y}}\right)$
Random	625	1.0
CE's Sort Order	576	0.9
Expenditures	133	0.2

The standard errors of mean total expenditures of Table 3 are also represented in Figure 2, which presents boxplots to provide a graphical representation of the distribution of $\hat{\bar{Y}}$ for all the SR PSUs.

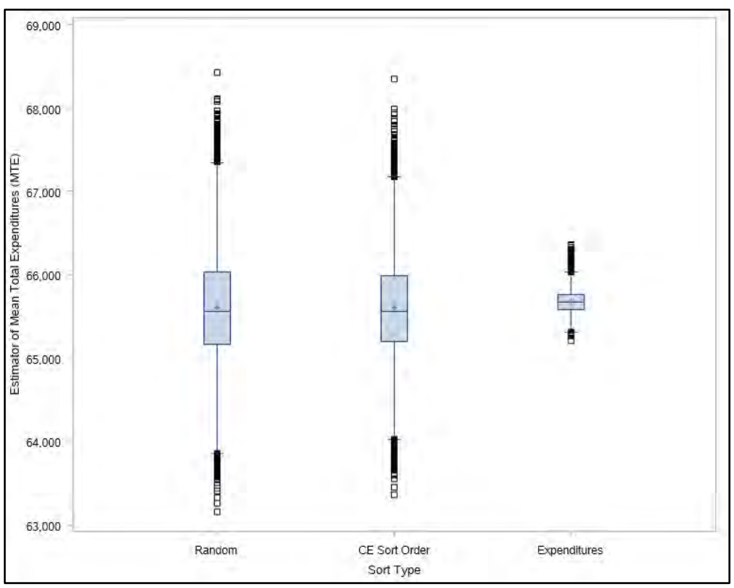

Figure 2: Distributions of the Estimators of Mean Total Expenditures for Varying Sort Orders

In both Table 3 and Figure 2, we see that the variances of $\hat{\bar{Y}}$ get smaller as the sort order becomes more informative: the standard error of the random sort is about five times larger than the standard error with the sort by the value of expenditures.

We considered the simulated standard errors of Table 3 as the target for our variance estimators, and therefore, they are used to measure the bias of the alternative estimators. The bias ratio is a simple measure of the bias and for $\widehat{se}\left(\hat{\hat{Y}}\right)$, we define bias ratio as$Bias\ Ratio\left(\widehat{se}\left(\hat{\hat{Y}}\right)\right) = \sqrt{E_{sim}\left(\hat{v}\left(\hat{\hat{Y}}\right)\right)} / se_{sim}\left(\hat{\hat{Y}}\right)$, where the simulation expectation of $\hat{v}\left(\hat{\hat{Y}}\right)$ is defined as: $E_{sim}\left(\hat{v}\left(\hat{\hat{Y}}\right)\right) = n_{sim}^{-1}\sum_{t=1}^{n_{sim}}\hat{v}\left(\hat{\hat{Y}}_t\right)$. We chose the bias ratios to measure the bias of $\widehat{se}\left(\hat{\hat{Y}}\right)$ because they are simple to interpret: values less than one indicate that $\widehat{se}\left(\hat{\hat{Y}}\right)$ is underestimating $se\left(\hat{\hat{Y}}\right)$ and values greater than one indicate that that $\widehat{se}\left(\hat{\hat{Y}}\right)$ is underestimating $se\left(\hat{\hat{Y}}\right)$. A simple measure of the variance of an estimator is the coefficient of variation. For $\widehat{se}\left(\hat{\hat{Y}}\right)$, the simulation coefficient of variation is defined as a percent as:

$$cv_{sim}\left(\widehat{se}\left(\hat{\hat{Y}}\right)\right) = \frac{\sqrt{v_{sim}\left(\hat{v}\left(\hat{\hat{Y}}\right)\right)}}{se_{sim}\left(\hat{\hat{Y}}\right)} \times 100,$$

where the simulation variance of $\widehat{se}\left(\hat{\hat{Y}}\right)$ is $v_{sim}\left(\widehat{se}\left(\hat{\hat{Y}}\right)\right) = n_{sim}^{-1}\sum_{t=1}^{n_{sim}}\left(\widehat{se}\left(\hat{\hat{Y}}_t\right) - E_{sim}\left(\widehat{se}\left(\hat{\hat{Y}}\right)\right)\right)^2$. We chose to use the coefficient of variation because they are all on the same scale – the coefficient of variation for each variance estimator was divided by the same value $se_{sim}\left(\hat{\hat{Y}}\right)$, and this makes them easy to compare.

Figure 3 presents the distributions of $\widehat{se}\left(\hat{\hat{Y}}_t\right)/\ se_{sim}\left(\hat{\hat{Y}}\right)$ for the random sort or sort order 1.

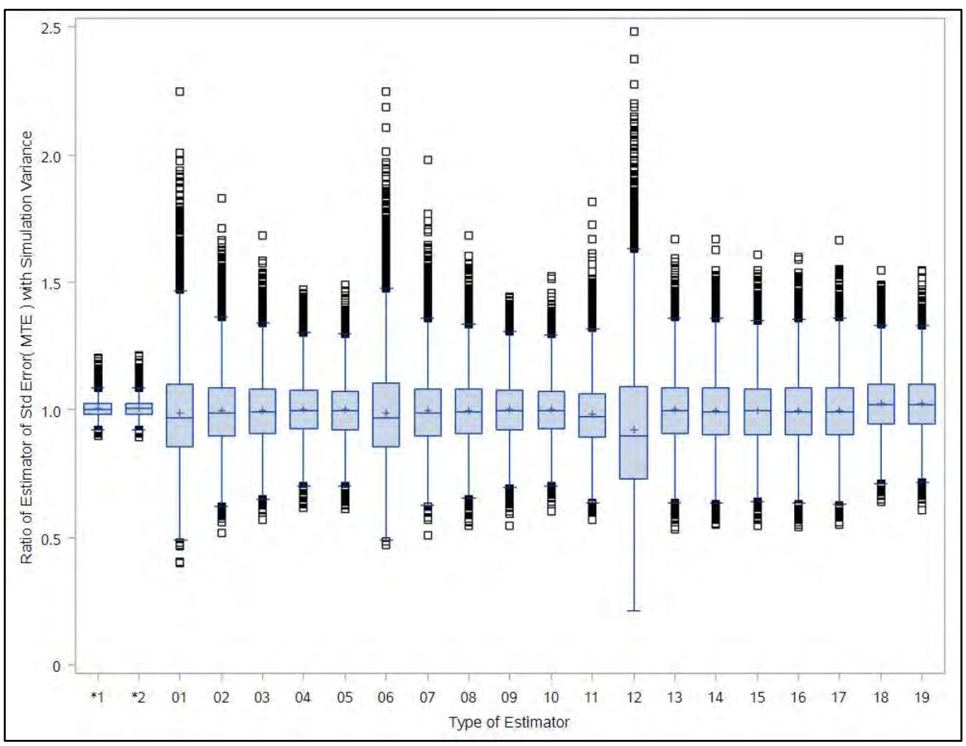

Figure 3: Distributions of the Ratio of the Standard Error Estimators – Random Sort Order

We make the following observations about the bias and variance of the estimators represented in Figure 3:

(1) All the variance estimators are nearly unbiased for the random sort order. A majority of the bias ratios were 1.01, where the bias ratios ranged from 0.96 for the for the BRR variance estimator that split STRATUM (12) to 1.03 for the DAGJK variance estimators (17 and 18).

(2) The variances of the BRR variance estimators that used the sort order (01, 02, 03, and 04) and the BRR variance estimators that randomly split the half samples (05, 06, 07, and 08) all decreased with increasing number of replicates. For the BRR variance estimators that used the sort order, the coefficients of variation decreased from 19.0 percent to 11.1 percent and for the BRR variance estimators that randomly split the half samples, the coefficients of variation decreased from 19.2 percent to 11.0 percent. This is expected because as the number of replicates increases the variance of the variance estimators decreases.

(3) The variance estimators that split STRATUM (12) had the largest variance with a coefficient of variation of 26.9 percent. This may have occurred because we collapsed values of STRATUM with small sample counts in some PSUs. If this is the case, it might be improved by a different collapsing strategy.

(4) With the SDR variance estimators (13, 14, 15, 16, and 17), their biases and variances were all roughly the same. The bias ratios for all five estimators were 1.01 and their coefficients of variation ranged from 13.2 percent to 13.5 percent.

(5) The DAGJK variance estimators (18 and 19) had the largest bias ratios with 1.03. Their variances of 11.5 percent for the DAGJK variance estimator that used the sort order to assign the groups and 11.4 for DAGJK variance estimator that randomized the assignment of the groups were not the smallest, but they were in a group of six variance estimators with coefficients of variation less than 12.0 percent.

Next, Figure 4 presents the distributions of $\widehat{se}\left(\hat{\hat{Y}}_t\right)/\ se_{sim}\left(\hat{\hat{Y}}\right)$ for the approximate CE sort order or sort order 4.

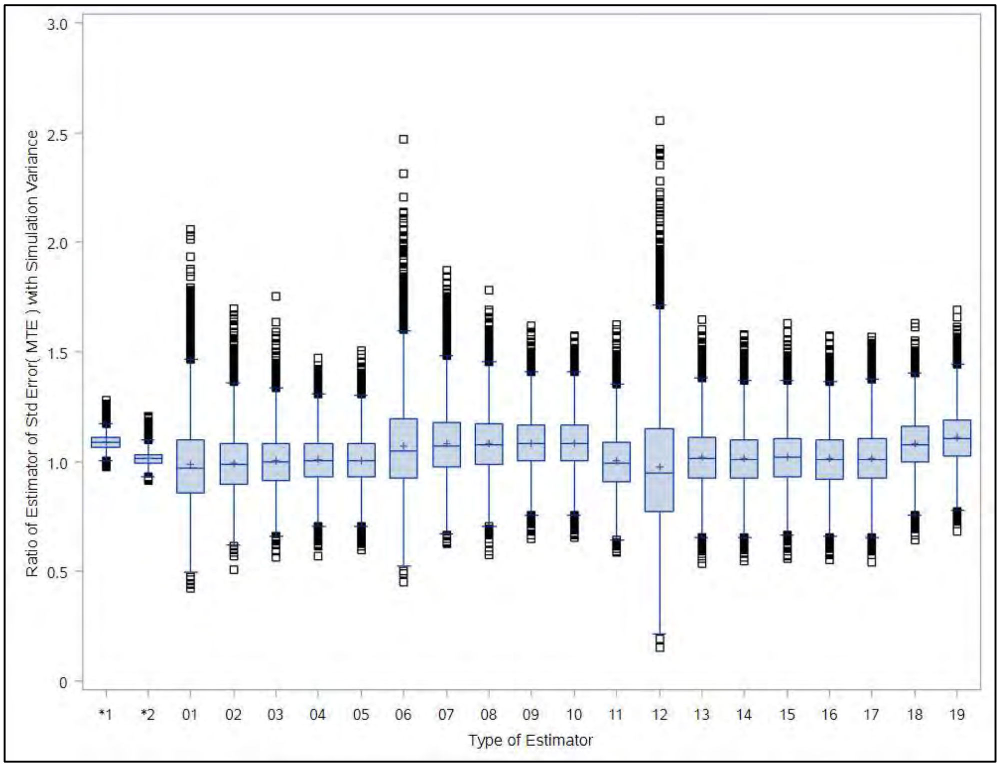

Figure 4: Distributions of the Ratio of the Standard Error Estimators – Approximate CE Sort Order

We remind the reader that we consider sort order 4 to be the closest to CE's current sort order and make the following observations about the bias and variance of the estimators represented in Figure 4:

(1) The BRR variance estimators that used the sort order (01, 02, 03, 04 and 05) were among the variance estimators with the smallest biases with bias ratios ranging from 1.00 to 1.01. We also see that their variances decreased as their number of pseudo strata increased: the coefficients of variation decreased from 18.8 percent for variance estimator 01 to 11.2 percent for variance estimator 05.
(2) The SDR variance estimators (13, 14, 15, 16, and 17) all had similar bias ratios ranging from 1.02 percent to 1.03 percent and they all had similar coefficients of variation ranging from 13.0 to 13.5.
(3) The DAGJK variance estimator that randomly assigned the groups (18) had the largest bias ratio of the 19 replication variance estimators with a bias ratio of 1.12 and the DAGJK variance estimator that used the sort order to assigned groups (17) did better with a bias ratio of 1.09, but both variance estimators were among the variance estimators with the largest values of the bias ratio.

Figure 5 presents the distributions of $\widehat{se}\left(\hat{\hat{Y}}_t\right)/\ se_{sim}\left(\hat{\hat{Y}}\right)$ for the sort by total expenditures or sort order 8.

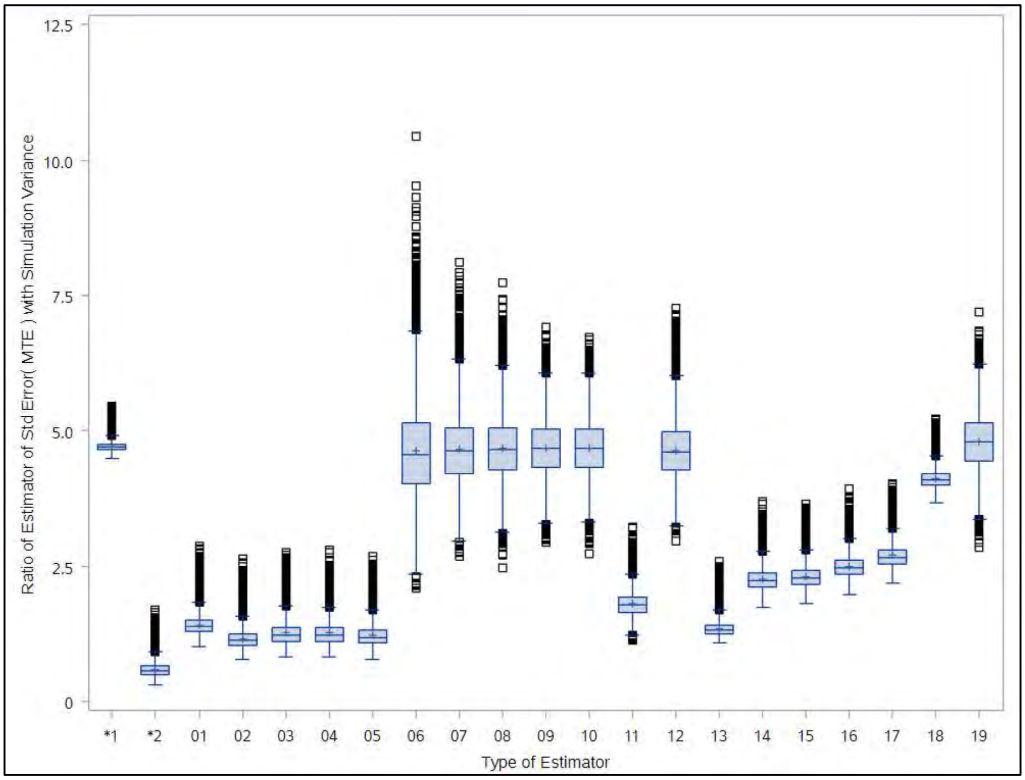

Figure 5: Distributions of the Ratio of the Standard Error Estimators – Sort by Expenditures

We make the following observations about the bias and variance of the estimators represented in Figure 5:

(1) The BRR variance estimators that used the sort order (01, 02, 03, 04 and 05) had both the smallest biases and variances of all the variance estimators. The bias ratios ranged from 1.18 to 1.43 and the coefficients of variation ranged from 19.9 percent to 22.4 percent.

(2) The variance estimators that randomly assigned the sample to half samples, including the BRR variance estimators (06, 07, 08, 09, and 10), the BRR variance estimator that split STRATUM, and the DAGJK variance estimator (19), all overestimated the variance because they estimated the variance as if it came from a SRSWOR sample. We know this because the bias ratios for all of these variance estimators ranged from 4.71 to 4.83, which includes the bias ratio for the SRSWOR variance estimator (*1) with a bias ratio of 4.71.

(3) The biases of the SDR variance estimators (13, 14, 15, 16, and 17) increased as the number of units in a connected loop increased. The SDR variance estimator with no connected loops (13) had the smallest bias ratio with 1.36 and the SDR variance estimator with 44 units in a connected loop (17) had the largest bias ratio with 2.70. The SDR variance estimator with no connected loops (13) had the smallest variance of the SDR variance estimators with a coefficient of variation of 16.4 percent. The coefficients of variation of the other SDR variance estimators (14, 15, 16, and 17) ranged from 21.0 percent to 21.9 percent.

Table 4 brings together the bias, variance, MSE, and coverage ratios for the variance estimators of mean total expenditures for all SR PSUs and includes their rankings for easy comparisons. In Table 12, the top five ranked variance estimators in terms of MSE, bias, variance, and coverage ratios are highlighted in bold.

Table 4: Rankings of the Variance Estimators for CE's Sort Order for All Self-Representing Primary Sampling Units

Variance Estimator	Rankings				Values			
	MSE	Bias (Bias Ratios)	Variance (CVs of SEs)	Coverage Ratio	MSE	Bias (Bias Ratios)	Variance (CVs of SEs)	Coverage Ratio
04-BRR | sort order | 22 p-strata	**1**	6	**1**	**5**	**11.2**	**1.01**	**11.2**	90.5
05-BRR | sort order | 44 p-strata	**2**	**5**	**2**	7	**11.2**	**1.01**	**11.2**	**90.4**
03-BRR | sort order | 4 p-strata	**3**	**3**	7	**2**	**12.8**	**1.01**	12.8	**90.2**
11-BRR | sort order | split county	**4**	**4**	10	**1**	**13.1**	**1.01**	13.1	**90.0**
16-SDR | sort order | 16 loops	**5**	8	9	**4**	**13.1**	1.02	13.1	**90.4**
15-SDR | sort order | 8 loops	6	12	8	10	13.2	1.03	13.0	90.7
17-SDR | sort order | 44 loops	7	10	11	9	13.4	1.03	13.3	90.6
14-SDR | sort order | 4 loops	8	9	12	6	13.4	1.02	13.4	90.5
13-SDR | sort order | 0 loops	9	11	13	8	13.7	1.03	13.5	90.6
02-BRR | sort order | 2 p-strata	10	**1**	14	**3**	13.9	**1.00**	13.9	**89.8**
18-DAGJK | sort order | 44 groups	11	13	**3**	15	14.5	1.09	**12.0**	92.7
10-BRR | rand order | 44 p-strata	12	15	**4**	17	14.7	1.09	**12.0**	92.7
09-BRR | rand order | 22 p-strata	13	17	**5**	16	14.8	1.09	**12.1**	92.7
08-BRR | rand order | 4 p-strata	14	18	15	14	16.3	1.09	14.0	92.5
19-DAGJK | rand order | 44 groups	15	19	6	19	16.7	1.12	12.4	93.4
07-BRR | rand order | 2 p-strata	16	16	16	13	17.4	1.09	15.4	92.3
01-BRR | sort order | 1 p-strata	17	**2**	17	11	18.8	**1.01**	18.8	88.8
06-BRR | rand order | 1 p-strata	18	14	18	12	21.9	1.09	20.7	91.3
12-BRR | sort order | split strata	19	7	19	18	28.6	1.02	28.5	86.8

For the bias, we ranked the absolute value of the difference between the bias ratio and 1.0 and for the coverage ratio, we ranked the absolute value of the difference between the coverage ratio and 90 percent.

From Table 4, we make the following observations about the variance estimators of mean total expenditures of all SR PSUs:

(1) The variance estimators with the smallest five rankings of MSEs were the BRR variance estimators that used the sort order and used 4, 22, 44 pseudo strata (03, 04, and 05), the BRR variance estimator that split by county (11), and the SDR variance estimator that used 16 connected loops (11).

(2) The variance estimators with the smallest five rankings of bias were the BRR estimators that used the sort order and split the sample into 1, 2, 4, and 44 pseudo strata (01, 02, 04, and 05) and the BRR variance estimator that split by state/county (11). We add that all of these variance estimators were close in terms of bias since they and several other variance estimators all had average bias ratios ranging from 1.00 to 1.01.

(3) The variance estimators with the smallest five rankings of variance used 22 or 44 replicates and included the BRR variance estimators with 22 or 44 pseudo strata (04, 05 09, and 10) and the DAGJK variance estimator that used the sort order with 44 groups (18).

(4) The variance estimators with the smallest five rankings of the confidence intervals were the BRR estimators that used the sort order and split the sample into 2, 4, and 22 pseudo strata (02 and 03) or the BRR variance estimator that split by state/county (11), and the SDR variance estimator that used 16 connected loops (16). We add that all of these variance estimators were close in terms of coverage ratios since they and several other variance estimators all had average coverage ratios ranging from 89.8 to 90.5 percent.

Recommendation for CE's national estimates. We recommend estimating variances for the SR PSUs with the BRR estimators that used the sort order and split the sample into 22 or 44 pseudo strata (04 or 05) because they had the smallest MSEs and variances and their rankings for bias, and confidence intervals were all in the top 7.

We also produced the same rankings as shown in Table 5 for each of the 23 MSA-level estimates (or equivalently the 23 SR PSUs) and then averaged the MSA-level rankings and values of the bias, variance, MSE, and coverage ratios across all of the MSAs. Table 5 summarizes the averages for the variance estimators of mean total expenditures applied to CE's sort order. The top five ranked variance estimators in terms of MSE, bias, variance, and coverage ratios are highlighted in bold.

Table 5: Average Rankings of the Metropolitan Statistical Areas by Measures of the Variance Estimators for the Consumer Expenditure Survey's Sort Order

	Average Rankings				Average Values			
Variance Estimator	MSE	Bias (Bias Ratios)	Variance (CVs of SEs)	Coverage Ratio	MSE	Bias (Bias Ratios)	Variance (CVs of SEs)	Coverage Ratio
05-BRR | sort order | 44 p-strata	**1.0**	6.7	**1.1**	**3.4**	**16.8**	1.02	**16.7**	**90.0**
15-SDR | sort order | 8 loops	**3.1**	8.1	**5.0**	**3.6**	**18.0**	1.03	**17.9**	**89.8**
16-SDR | sort order | 16 loops	**3.2**	8.0	5.4	**4.0**	**18.0**	1.02	17.9	**90.1**
17-SDR | sort order | 44 loops	**4.8**	**6.0**	7.1	**4.5**	**18.1**	**1.02**	18.1	**90.1**
14-SDR | sort order | 4 loops	**5.6**	8.2	7.9	4.4	**18.2**	1.02	18.1	90.6
13-SDR | sort order | 0 loops	5.8	8.1	8.1	**4.1**	18.3	1.02	18.2	**91.6**
18-DAGJK | sort order | 44 groups	5.9	14.5	**2.3**	8.4	18.8	1.08	**17.2**	90.2
10-BRR | rand order | 44 p-strata	7.4	15.9	**3.0**	9.5	19.3	1.09	**17.4**	89.2
04-BRR | sort order | 22 p-strata	8.6	6.3	10.0	5.9	19.7	1.02	19.6	93.0
19-DAGJK | rand order | 44 groups	9.6	19.0	**5.1**	11.0	20.8	1.12	**17.8**	93.5
09-BRR | rand order | 22 p-strata	11.0	16.0	11.0	7.7	22.2	1.09	20.8	92.3
12-BRR | sort order | split strata	12.0	**5.0**	12.1	12.3	31.4	**1.02**	31.2	85.0
03-BRR | sort order | 4 p-strata	13.1	6.3	13.1	14.2	36.4	1.02	36.0	83.2
08-BRR | rand order | 4 p-strata	14.3	15.7	14.3	12.5	39.0	1.09	38.8	86.5
11-BRR | sort order | split county	15.1	**5.1**	15.1	15.4	45.1	**1.02**	44.0	81.0
02-BRR | sort order | 2 p-strata	15.7	**5.1**	15.7	16.7	48.7	**1.01**	47.5	77.0
07-BRR | rand order | 2 p-strata	16.8	15.7	16.8	15.7	51.6	1.09	51.5	80.3
01-BRR | sort order | 1 p-strata	17.9	**5.2**	17.9	18.9	64.3	**1.01**	61.2	66.2
06-BRR | rand order | 1 p-strata	19.0	15.3	19.0	17.8	67.7	1.09	66.4	69.5

The results from Table 5 tell a slightly different story as compared with the results from Table 4.

(1) The variance estimators with the smallest five rankings of MSEs were the BRR variance estimators that used 44 replicates (05) and the SDR variance estimators that used 4, 8, 16, and 44 connected loops (14, 15, 16, and 17).
(2) The variance estimators with the smallest five rankings of bias were the BRR variance estimators that split state/county or STRATUM (11 or 12) and the BRR variance estimator that used the sort order and two pseudo strata (02). We add that all of these variance estimators were close in terms of bias since they and several other variance estimators all had average bias ratios ranging from 1.01 to 1.02.
(3) The variance estimators with the smallest rankings of variances were the BRR variance estimators with 44 pseudo strata (05 and 10), the DAGJK variance estimator with 44 groups (18 and 19), and BRR variance estimators that used the sort order and split the sample into 8 pseudo strata (04).
(4) The variance estimators with the smallest five rankings of confidence intervals were the BRR variance estimators that used the sort order and split the sample into 44 pseudo strata (05) and the SDR variance estimators that used 0, 8, 16, and 44 connected loops (13, 15, 16, and 17). We add that all of these variance estimators were close in terms of coverage ratios since they all had average coverage ratios ranging from 89.8 to 90.2 percent.

Recommendation for CE's MSA-level estimates. We recommend the BRR variance estimator that used the sort order and split the sample into 44 pseudo strata (05). This estimator had the smallest average ranking for the MSE, variance, and coverage ratio. Although variance estimators 05 was not the best in terms of bias, it was in the top 5 in terms of bias.

Overall Recommendation for CE: We recommend the BRR variance estimator that used the sort order and split the sample into 44 pseudo strata (05). It was one of the best variance estimators for the national estimates and the best for the MSA-level estimates. The main reason why variance estimator 05 does well is that it uses 44 replicates, which reduces the variance of the variance estimator.

6. Conclusions

The simulation study showed that splitting the SR strata into an increasing number of pseudo strata reduces the bias of the variance estimator. This suggests that CE could improve its variance estimates by splitting each of its SR strata into 44 pseudo strata rather than just one, as it currently does. The improvements in the variance of the variance estimator impacted both the national and MSA estimates but were more impactful to the MSA estimates.

Both the BRR that split the SR strata into pseudo strata and half samples and SDR performed the best overall. We suggest that this happened because both are constructed as collapsed-strata variance estimators. Although we expected SDR to perform better because it includes more pairs of implicit strata, we did not see an appreciable difference, and we think this happened because there might not be much difference between using $n/2$ implicit strata with BRR and $(n-1)$ implicit strata with SDR, when the within PSU sample sizes n are reasonably large as was the case with our simulation study. We do not know what would happen with smaller samples sizes.

7. Disclaimer

This paper provides a summary of research results. The information is being released for statistical purposes, to inform interested parties, and to encourage discussion of work in progress. The presentation does not represent an existing, or a forthcoming new, official BLS statistical data product or production series.

Appendix

Result A1: The replication variance estimator:

$$\hat{v}_{CC3}(\hat{Y}) = \frac{R-1}{R(1-\kappa)^2}\sum_{r=1}^{R}\left(\hat{Y}_r - \hat{Y}\right)^2$$

used with the replicate factors:

$$F_{rg1}^{(BRR/CC3)} = 1 + 2(1-\kappa)(R-1)^{-\frac{1}{2}}P_{g2}a_{gr}$$

$$F_{rg2}^{(BRR/CC3)} = 1 - 2(1-\kappa)(R-1)^{-\frac{1}{2}}P_{g1}a_{gr}$$

is equivalent to $\hat{v}_{BRR}(\hat{Y})$ and when $\hat{v}_{CC3}(\hat{Y})$ is used with replicate factors:

$$F_{rk}^{(DAGJK/CC3)} = (1-\kappa)\left(\frac{R}{R-1}\right)I_{rk} + \kappa$$

is equivalent to $\hat{v}_{DAGJK}(\hat{Y})$. Note that neither set of replicate factors produces negative replicate weights.

With $F_{ghr}^{(BRR/CC3)}$, we start with the difference:

$$\begin{aligned}
\hat{Y}_r - \hat{Y} &= \sum_{g=1}^{B}\left[\sum_{k\in s_{1i}} w_{hi}w_k F_{rg1}^{(BRR/CC3)} y_k + \sum_{k\in s_{2i}} w_{hi}w_k F_{rg2}^{(BRR/CC3)} y_k\right] - \sum_{g=1}^{B}\sum_{h}\sum_{i\in s_h}\sum_{k\in s_{hi}} w_{hi}w_k y_k \\
&= \sum_{g=1}^{B}\left[\sum_{k\in s_{1i}} w_{hi}w_k\left(1 + 2(1-\kappa)(R-1)^{-\frac{1}{2}}P_{g2}a_{gr}\right)y_k \right.\\
&\qquad \left. - \sum_{k\in s_{2i}} w_{hi}w_k\left(1 - 2(1-\kappa)(R-1)^{-\frac{1}{2}}P_{g1}a_{gr}\right)y_k\right] \\
&\qquad - \sum_{g=1}^{B}\left[\sum_{k\in s_{1i}} w_{hi}w_k y_k + \sum_{k\in s_{2i}} w_{hi}w_k y_k\right] \\
&= \sum_{g=1}^{B}\left[\sum_{k\in s_{1i}} w_{hi}w_k\left(2(1-\kappa)(R-1)^{-\frac{1}{2}}P_{g2}a_{gr}\right)y_k \right.\\
&\qquad \left. - \sum_{k\in s_{2i}} w_{hi}w_k\left(2(1-\kappa)(R-1)^{-\frac{1}{2}}P_{g}a_{gr}\right)y_k\right] \\
&= 2(1-\kappa)(R-1)^{-\frac{1}{2}}\sum_{g=1}^{B}\left[a_{gr}\left(\left(P_{g2}\sum_{k\in s_{1i}} w_{hi}w_k y_k\right) - \left(P_{g1}\sum_{k\in s_{2i}} w_{hi}w_k y_k\right)\right)\right] \\
&= 2(1-\kappa)(R-1)^{-\frac{1}{2}}\sum_{g=1}^{B}\left[a_{gr}\left(P_{g2}\hat{Y}_{g1} - P_{g1}\hat{Y}_{g2}\right)\right]
\end{aligned}$$

where a_{gr} is the value from the g^{th} row and r^{th} column of a Hadamard matrix. Then $\hat{v}_{CC3}(\hat{Y})$ can be expressed as:

$$
\begin{aligned}
\hat{v}_{CC3}(\hat{Y}) &= \frac{R-1}{R(1-\kappa)^2}\sum_{r=1}^{R}\left(\hat{Y}_r-\hat{Y}\right)^2 \\
&= \frac{R-1}{R(1-\kappa)^2}\sum_{r=1}^{R}\left(2(1-\kappa)(R-1)^{-\frac{1}{2}}\sum_{g=1}^{B}\left[a_{gr}\left(P_{g2}\hat{Y}_{g1}-P_{g1}\hat{Y}_{g2}\right)\right]\right)^2 \\
&= \frac{4}{R}\sum_{r=1}^{R}\left(\sum_{g=1}^{G}\left[a_{gr}^2\left(P_{g2}\hat{Y}_{g1}-P_{g1}\hat{Y}_{g2}\right)^2\right]\right. \\
&\qquad \left. +\sum_{g=1}^{B}\sum_{\substack{g'=1\\ g'\neq g}}^{B}\left[a_{gr}a_{g'r}\left(P_{g2}\hat{Y}_{g1}-P_{g1}\hat{Y}_{g2}\right)\left(P_{g'2}\hat{Y}_{g'1}-P_{g'1}\hat{Y}_{g'2}\right)\right]\right) \\
&= \frac{4}{R}\left(\sum_{g=1}^{G}\left[\left(\sum_{r=1}^{R}a_{gr}^2\right)\left(P_{g2}\hat{Y}_{g1}-P_{g1}\hat{Y}_{g2}\right)^2\right]\right. \\
&\qquad \left. +\sum_{g=1}^{B}\sum_{\substack{g'=1\\ g'\neq g}}^{B}\left[\left(\sum_{r=1}^{R}a_{gr}a_{g'r}\right)\left(P_{g2}\hat{Y}_{g1}-P_{g1}\hat{Y}_{g2}\right)\left(P_{g'2}\hat{Y}_{g'1}-P_{g'1}\hat{Y}_{g'2}\right)\right]\right) \\
&= 4\sum_{g=1}^{B}\left(P_{g2}\hat{Y}_{g1}-P_{g1}\hat{Y}_{g2}\right)^2 \\
&= \hat{v}_{CS}(\hat{Y}) \\
&= \hat{v}_{BRR}(\hat{Y})
\end{aligned}
$$

where $\sum_{r=1}^{R} a_{gr}a_{g'r} = 0$ and $\sum_{r=1}^{R} a_{gr}^2 = R$ because the rows of the Hadamard matrix are orthogonal. Next with $F_{ir}^{(DAGJK/CC3)}$, we start with the difference:

$$
\begin{aligned}
\hat{Y}_r-\hat{Y} &= \sum_{h}\sum_{k\in s_{hi}} w_{hi}w_k F_{rk}^{(DAGJK/CC\ )} y_k - \sum_{h}\sum_{k\in s_{hi}} w_{hi}w_k y_k \\
&= \sum_{h}\sum_{k\in s_{hi}} w_{hi}w_k\left((1-\kappa)\left(\frac{R}{R-1}\right)I_{rk}+\kappa\right)y_k - \sum_{h}\sum_{k\in s_{hi}} w_{hi}w_k y_k \\
&= \sum_{h}\sum_{k\in s_{hi}} w_{hi}w_k\left((1-\kappa)\left(\frac{R}{R-1}\right)I_{rk}+\left(1-(1-\kappa)\right)\right)y_k - \sum_{h}\sum_{k\in s_{hi}} w_{hi}w_k y_k \\
&= \sum_{h}\sum_{k\in s_{hi}} w_{hi}w_k\left((1-\kappa)\left(\frac{R}{R-1}\right)I_{rk}-(1-\kappa)\right)y_k \\
&= (1-\kappa)\left(\sum_{h}\sum_{k\in s_{hi}} w_{hi}w_k\left(\frac{R}{R-1}\right)I_{rk}y_k - \sum_{h}\sum_{k\in s_{hi}} w_{hi}w_k y_k\right) \\
&= (1-\kappa)\left(\hat{Y}_r-\hat{Y}\right)
\end{aligned}
$$

Then $\hat{v}_{CC\ }(\hat{Y})$ can be expressed as:

$$
\begin{aligned}
\hat{v}_{CC3}(\hat{Y}) &= \frac{R-1}{R(1-\kappa)^2}\sum_{r=1}^{R}\left((1-\kappa)(\hat{Y}_r-\hat{Y})\right)^2 \\
&= \frac{R-1}{R}\sum_{r=1}^{R}(\hat{Y}_r-\hat{Y})^2 \\
&= \hat{v}_{DAGJK}(\hat{Y})
\end{aligned}
$$

References

Ash, S. (2014). Using successive difference replication for estimating variances, Survey Methodology, 40, 47-59. Available at: https://www150.statcan.gc.ca/n1/en/pub/12-001-x/2014001/article/14029-eng.pdf?st=JKP7GRPp (accessed September 2024).

Ash, S. (2022). Evaluating the First-Stage Sample Design for the 2020 Redesign of the Consumer Expenditures Surveys, JSM, Section on Survey Research Methods. Available at: https://ww2.amstat.org/meetings/proceedings/2015/data/assets/pdf/234097.pdf (accessed September 2024).

Bureau of Justice Statistics. (2014). National Crime Victimization Survey Technical Documentation NCJ 247252 September 2014. Available at: https://bjs.ojp.gov/sites/g/files/xyckuh236/files/media/document/ncvstd13.pdf (accessed September 2024).

Bureau of Justice Statistics. (2020). National Crime Victimization Survey, 2016 Technical Documentation NCJ 251442 12/8/2017. Available at: https://bjs.ojp.gov/sites/g/files/xyckuh236/files/media/document/ncvstd16.pdf (accessed September 2024).

Bureau of Labor Statistics. (2023). Handbook of Methods, Available at: https://www.bls.gov/opub/hom/home.htm (accessed September 2024).

Dippo, C.S., R.E. Fay, and R.H. Morganstein. (1994). Computing Variances for Complex Survey Designs with Replication, JSM, Section on Survey Research Methods, 489-494. Available at: http://www.asasrms.org/Proceedings/papers/1984_094.pdf (accessed September 2024).

Fay, R.E. and G.F. Train. (1995). "Aspects of Survey and Model-Based Postcensal Estimation of Income and Poverty Characteristics for States and Counties." In Proceedings of the Section on Government Statistics, American Statistical Association. Available at: https://www.census.gov/content/dam/Census/library/working-papers/1995/demo/faytrain95.pdf (accessed September 2024).

Guciardo, C.J., A.H. Dorfman, L.R. Ernst, and M. Sverchkov. (2004). Variance Estimation for the National Compensation Survey When PSUs are Clustered Prior to the Second Phase of Sampling, JSM, Section on Survey Research Methods. Available at: https://www.bls.gov/osmr/research-papers/2004/pdf/st040230.pdf (accessed September 2024).

Hansen, M.H., W.M. Hurwitz, and W.G. Maddow. (1953). *Sample Survey Methods and Theory*, John Wiley & Sons.

Johnson, E.G. and K.F. Rust. (1992). Population Inferences and Variance Estimation for NAEP Data, Journal of Educational Statistics. 17, 175-190. Available at: https://doi.org/10.2307/1165168 (accessed September 2024).

Judkins, D.R. (1990). Fay's method for Variance Estimation, Journal of Official Statistics, 6, 223-239. Available at: https://www.scb.se/contentassets/ca21efb41fee47d293bbee5bf7be7fb3/fay39s-method-for-variance-estimation.pdf (accessed September 2024).

Kott, P.S. (1998). Using the Delete-a-Group Jackknife Variance Estimator in Practice, JSM, Section on Survey Research Methods. Available at: http://www.asasrms.org/Proceedings/papers/1998_130.pdf (accessed September 2024).

Kott, P.S. (2001). The Delete-a-Group Jackknife, Journal of Official Statistics, 17, 4, 521-526. Available at: https://www.scb.se/contentassets/ca21efb41fee47d293bbee5bf7be7fb3/the-delete-a-group-jackknife.pdf (accessed September 2024).

Lee, E.S. and R.N. Forthofer. (2006). *Analyzing Complex Survey Data*, Sage.

McCarthy, P.J. (1966). Replication: An Approach to the Analysis of Data from Complex Surveys, Vital and Health Statistics: Series 2, No. 14, National Center for Health Statistics, Public Health Service, Washington D.C. Available at: https://stacks.cdc.gov/view/cdc/13018/cdc_13018_DS1.pdf (September 2024).

McCarthy, P.J. (1969a). Pseudoreplication: Further Evaluation and Application of the Balanced Half-Sample Technique, Vital and Health Statistics: Series 2, No. 31, National Center for Health Statistics, Public Health Service, Washington D.C. Available at: https://www.cdc.gov/nchs/data/series/sr_02/sr02_031acc.pdf (accessed September 2024).

McCarthy, P.J. (1969b). Pseudo-Replication: Half Samples, Review of the International Statistical Institute 37(3): 239-264. DOI: https://doi.org/10.2307/1402116 (accessed September 2024).

Megill, D.J., E.E. Gomez, A. Balmaceda, and M. Castillo, Measuring Efficiency of Implicit Stratification in Multistage Sample Surveys, JSM, Section on Survey Research Methods. Available at: http://www.asasrms.org/Proceedings/papers/1987_024.pdf (accessed September 2024).

Neiman, D., S. King, D. Swanson, S. Ash. J. Enriquez, and J. Rosenbaum. (2015). Review of the 2010 Sample Redesign of the Consumer Expenditure Survey, JSM, Section on Survey Research Methods. Available at: http://www.asasrms.org/Proceedings/y2015/files/234097.pdf (accessed September 2024).

Nixon, M.G., J.M. Brick, G. Kalton, and H. Lee. (1998). Alternative Variance Estimation for the NHIS, JSM, Section on Survey Research Methods. Available at: http://www.asasrms.org/Proceedings/papers/1998_052.pdf (accessed September 2024).

Sukasih, A.S. and D. Jang. (2003). Monte Carlo Study on the Successive Difference Replication Method for Non-Linear Statistics, JSM, Section on Survey Research Methods. Available at: http://www.asasrms.org/Proceedings/y2003/Files/JSM2003-000867.pdf (accessed September 2024).

U.S. Census Bureau. (2019). Current Population Survey Design and Methodology Technical Paper 77. Available at: https://www2.census.gov/programs-surveys/cps/methodology/CPS-Tech-Paper-77.pdf (accessed September 2024).

U.S. Census Bureau and Department of Housing and Urban Development. (2022). 2021 AHS Integrated National Sample: Sample Design, Weighting, and Error Estimation. Available at: https://www2.census.gov/programs-surveys/ahs/2021/2021%20AHS%20National%20Sample%20Design,%20Weighting,%20and%20Error%20Estimation.pdf (accessed September 2024).

Wolter, K.M. (2007). *Introduction to Variance Estimation*. 2nd Ed., Springer.

Yates, F. (1953). *Sampling Methods for Censuses and Surveys*, 2nd Edition, Hafner Publishing Company, New York, NY.