Using the Bootstrap in a Two-Stage Nested Complex Sample Design[1]

Steven Kaufman, National Center for Education Statistics
Room 9075,1990 K St. NW, Washington, D.C 20006

Key Words: BHR, Super-Population, Finite Population Correction, Variance Estimation

1.0 Introduction

Replication variance estimation for a two-stage nested sample design is usually implemented by generating replicate samples (weights) that replicate the original first-stage sample selection. Since the second-stage is nested, the second-stage variance can be reflected by associating each second-stage unit with its respective first-stage unit in each first-stage replicate sample. The second-stage sampling can be viewed as being indirectly incorporated into the replicates, because the second-stage sampling is not independently replicated within each replicate. As long as the first-stage sampling rates are not too high or first-stage sampling is done with replacement, this should provide a reasonable variance approximation. This paper investigates whether generating replicates that directly reflect both the first and second stage sampling provide any advantages over replicates that directly reflect only the first-stage sampling.

This paper is particularly interested in the National Center for Education Statistics' (NCES) School and Staffing survey (SASS). In this survey, the first-stage sampling rates can be large. SASS collects data for both the first and second stage units. Because of the large sampling rates, a first-stage finite population correction (FPC) is required for the first-stage data variance estimates. Since estimation frequently requires combining the first and second stage data, the replicate weights for estimates based on first-stage units must be consistent with the replicate weights for estimates based on second-stage units. This implies using the first-stage FPC in both variance estimators. However, when the second-stage variance is indirectly reflected, applying a first-stage FPC can underestimate the second-stage variance component (i.e., a first-stage FPC bias), since the second-stage component is correct without this adjustment. By directly reflecting the second-stage variance in the replicates, this bias can be eliminated.

This paper will present two sets of replicate weights. Both sets will incorporate the same first-stage FPC. One set will indirectly reflect the second-stage variance, while the other set will directly reflect the second-stage variance. Results will be presented using high and low first-stage sampling rates, which will provide a measure of the first-stage FPC bias, for different sized FPCs.

To generate a set of replicate weights that directly reflect the sampling at both selection stages, a bootstrap methodology will be used. In that methodology, a first-stage bootstrap sample is selected of size n_h^*. Within each selected first-stage bootstrap unit i^*, a second-stage bootstrap sample is selected of size m_i^*. n_h^* and m_i^* are chosen to provide unbiased first and second order moments. Sitter (1992) and Kaufman (2000) provide examples how this can be done.

To generate replicate weights that indirectly reflect the second-stage variance, a bootstrap method will also be used. The bootstrap method is the first-stage component of the bootstrap estimator described above (i.e., a first-stage bootstrap sample is selected of size n_h^*, where n_h^* is chosen to provide an unbiased variance estimator for the first-stage sample). The second-stage replicate weight is the product of the first-stage replicate weight, just described, times the conditional second-stage weight given the first-stage unit is in sample.

To compare the variance estimators, a simulation study is performed. Within stratum, the first-stage SASS is selected systematically probability proportional size (PPS). Within each selected first-stage unit, the second-stage sample is selected systematically with equal probability. Kaufman (2001) provides an appropriate systematic PPS FPC under a locally random assumption. By imposing the locally random assumption in the simulation, the appropriate first-stage FPC is known for both variance estimators. The main source of bias is then determined through how this FPC is used in the two sets of replicate weights. Performance is measured by comparing relative errors and coverage rates.

To start, the bootstrap procedures are described.

2.0 Bootstrap Distribution Function

In this discussion, the bootstrap is defined in terms of the sampling process rather than in terms of a specific variable of interest (i.e., the object is to

[1] This paper is intended to promote the exchange of ideas among researchers and policy makers. The views expressed in it are part of ongoing research and analysis and do not necessarily reflect the position of the U.S. Department of Education.

generate a set of bootstrap samples). The advantage of this is that once the bootstrap samples are generated, there is no need to repeat the resampling process for each variable. A set of bootstrap replicate weights can be generated similar to BHR replicate weights. See Kaufman (1999 and 2000).

In this context, let $\mathbf{I}_{n_h^*} = G(\mathbf{I}_{n_h}, \mathbf{A}_h^*)$, where $\mathbf{I}_{n_h^*}$ is a vector representing a bootstrap PSU sample of size n_h^* selected from the original sample $\mathbf{I}_{n_h}$; and $G(\mathbf{I}_{n_h}, \mathbf{A}_h^*)$ is some random mechanism generating $\mathbf{I}_{n_h^*}$. $\mathbf{A}_h^*$ is an appropriate parameter space needed to describe the random mechanism generating $\mathbf{I}_{n_h^*}$ (e.g., n_h^*, the first-stage bootstrap sample sizes is an element of $\mathbf{A}_h^*$). The bootstrap technique is: generate $G(\mathbf{I}_{n_h}, \mathbf{A}_h^*)$, so first-order expectations are preserved for all $\mathbf{A}_h^*$. $\mathbf{A}_{ho}^* \in \mathbf{A}_h^*$ is then determined so that $E^* v^*(\hat{T}_h^*) = v(\hat{T}_h)$, where E^* and v^* represent the bootstrap expectation and bootstrap sample variance, respectively (i.e., second-order expectations are preserved). The choice of $G(\mathbf{I}_{n_h}, \mathbf{A}_h^*)$ and $\mathbf{A}_h^*$ can be flexible.

2.1 Randomized Systematic PPS Sample

When PSUs are placed in a specific order before sample selection, there is no unbiased variance estimator for systematic PPS samples. However, if a small amount of randomization is introduced and it is assumed that the covariance between two selected units is zero then an unbiased variance estimator can be stated (see Kaufman (1999)). This variance estimator can be used to determine n_h^* in a bootstrap variance estimator using a bootstrap-PSU (i^*) frame. Kaufman calls this type of sample, a randomized systematic PPS sample.

The randomized systematic PPS sample variance estimator under a relatively mild super-population model provides an alternative variance model for systematic PPS sampling. See Kaufman (2001) for simulation results measuring the performance of this alternative variance model.

A randomized systematic sample can be chosen in the following way: 1) Order the frame in the desired way for a regular systematic selection. 2) Partition the frame into n_h groups (implicit strata), so each group's total measures of size are equal. 3) Consecutively pair the implicit strata to form variance-strata. 4) Some PSUs may have positive selection probability in two variance-strata. Such PSUs will be split into two new PSUs by assigning a proportionally allocated measure of size to the new PSUs, so that the new PSUs are totally within the respective variance-strata. 5) The PSUs within each variance-stratum are now placed in a random order. Finally, a classical systematic PPS sample is selected within strata.

The randomized systematic sample, as with the classical systematic sample implicitly stratifies the frame according to the original ordering in 1) above. The randomized systematic sample does not control as well as the classical systematic sample, but the control is still strong. For any contiguous group of frame PSUs, the classical systematic procedure will be within one PSU of the expected sample size for that group, while the randomized systematic sampling will be within two PSUs.

One advantage of the randomized systematic sample is that the total covariance, although not necessarily zero, should be expected to be drawn toward zero, because of the randomization. This should reduce the number of extreme total covariances. Since many systematic PPS variance estimators assume these covariances are zero, anything that reduces the number of extremes will reduce the number of extreme over and/or under estimates of variance.

In practice, one does not have to physically randomize the frame to use the randomized systematic PPS sample variance as a model for the nonrandomized systematic sample variance. However, one does need to assume, within variance-strata, the frame is randomized (i.e., locally random). Assuming the frame ordering takes this into consideration, this is not necessarily a difficult assumption to approximate. Kaufman (2001) describes the frame ordering considerations.

Whether one physically randomizes the frame or not, it is necessary to assume the total covariance is zero (e.g., $Cov(\hat{T}) = \sum_i \sum_{j\ (j \neq i)} Cov(\hat{T}_i, \hat{T}_j) = 0$, where $\hat{T}_i$ and $\hat{T}_j$ are the weighted components of the i^{th} and j^{th} selected PSU for some total $\hat{T}$). This may seem like a restrictive assumption; however, many variance estimators, under systematic sampling, make this assumption.

2.2 Directly Reflecting Second -Stage Sampling

In a single stage sample, one can try to choose n_h^* so that $G(\mathbf{I}_{n_h}, n_h^*)$ will produce an unbiased variance estimate. In a two-stage sample, where n_h first-stage and m_i second-stage units are selected, one can try to develop n_h^* from an appropriate $\mathbf{I}_{n_h^*}^{(1)} = G_1(\mathbf{I}_{n_h}^{(1)}, n_h^*)$, and $m_{i^*}^*$ from an appropriate $G_2(\mathbf{I}^{(2)}_{m_{i^*}}, m^*_{i^*} \mid i^* \in \mathbf{I}^{(1)}_{n^*_h}, n^*_h)$, to produce an unbiased variance estimate (See Sitter (1992) when both stages are SRS without replacement).

2.2.1 The Two-Stage Sample Design

To develop a bootstrap variance estimator, it is assumed that the SASS sample design can be approximated by using: 1) the randomized systematic sample, described above to approximate the traditional systematic PPS sample; and 2) a without replacement simple random sample (SRS) to approximate an equal probability systematic sample. Therefore, for the simulation, the simulation sample design is: A stratified randomized systematic PPS sample of schools comprises the first-stage sample. The measure of size is the square root number of teachers in the school. The SASS first-stage frame ordering is slightly altered, by introducing a serpentine ordering to: 1) make the original ordering look more locally random, 2) reduce the number of extreme $Cov(\hat{T})$ and 3) reduce the first-stage FPC. Within each school, the second-stage teacher sample will be selected SRS w/o replacement.

2.2.2 The Estimate of Interest

Let $\hat{T} = \sum_{h\in H} \sum_{i=1}^{n_h} w_i \hat{Y}_i = \sum_{h\in H} \sum_{i=1}^{n_h} \hat{T}_i = \sum_{h\in H} \hat{T}_h$, where w_i is the sampling weight associated with the i^{th} school (i.e., $1/p_i$, p_i being the selection probability for i); and $\hat{Y}_i$ is an unbiased estimator of the teacher total for i. $\hat{Y}_i = \sum_{j=1}^{m_i} M_i y_{ij} / m_i$, where M_i is the number of teachers in school i, y_{ij} is the variable of interest for teacher j in school i, and m_i is number of teachers selected in school i.

2.2.3 Estimating $v(\hat{T})$

Let $\overline{T}_h = \hat{T}_h / \sum_{i=1}^{n_h} w_i M_i = \hat{T}_h / X_h$.

Using a Taylor series approximation, $v(\overline{T}_h) \cong \sum_{i=1}^{n_h} w_i^2 M_i^2 v(\overline{y}_i) / X_h^2$, with $\overline{y}_i = \sum_{j=1}^{m_i} y_{ij} / m_i$.

From Cochran (1977) theorem 11.2, it follows that an unbiased estimator, within the Taylor Series approximation, for $V(\hat{T})$ is:

$$v(\hat{T}) \cong \sum_{h\in H} \left(v_1(\hat{T}_h) + \sum_{i=1}^{n_h} \left(p_i v_{2\,wor}(\hat{T}_i) \right) \right) \qquad \text{(a)}$$

$v_1(\hat{T}_h)$ is an unbiased variance estimator of the first-stage sample evaluated at $\hat{T}_h$. See Kaufman (1999) for $v_1(T_h)$.

$v_{2\,\text{wor}}(\hat{T}_i)$ is the unbiased estimate of the second-stage variance of $\hat{T}_i$. $v_{2\,\text{wor}}(\hat{T}_i) = w_i^2 M_i^2 (1 - f_{2i}) s^2_{2i} / m_i$, where $f_{2i} = m_i / M_i$ and $s^2_{2i} = \sum_{j=1}^{m_i} (y_{ij} - \overline{y}_i)^2 / (m_i - 1)$.

A bootstrap variance estimator is generated by:

1. Using $\mathbf{I}^b_{n^*_h} = G(\mathbf{I}_{n_h}, n^*_h)$ and n^*_h from Kaufman (1999), we have B sets of $\mathbf{I}^b_{n^*_h}$'s, as well as B sets of bootstrap-schools (i^*) weights, $w^*_{i^*b}$, providing an unbiased $v_1(T_h)$. The b^{th} replicate weight, w^*_{ib}, equals $\sum_{i^* \in S_{ib}} w^*_{i^*b}$, where S_{ib} is the set of i^* selected in the b^{th} replicate which were generated from i.

   A solution for n^*_h may not always exist. This can occur in strata where n_h is small and n^*_h =1 is not small enough to sufficiently increase the bootstrap variance. A solution is to combine strata, indexed by C, and sort the combined stratum by original stratum first. This increased n_C in the combined stratum should now allow a solution for n^*_C.
2. Given n^*_h and $\mathbf{I}^b_{n^*_h}$ from step 1, define $\mathbf{I}^b_{m^*_{i^*}} = G_{i^*}(\mathbf{I}_{m_{i^*}}, m^*_{i^*} \mid i^* \in \mathbf{I}^b_{n^*_h}, n^*_h)$ as follows:
   For $i^* \in \mathbf{I}^b_{n^*_h}$, independently select $m^*_{i^*}$ teachers with-replacement from the m_{i^*} originally sampled in school i which generated i^*. The conditional bootstrap replicate weight for the j^{th} teacher is $w^*_j = K^*_j M_{i^*} / m^*_{i^*}$, where K^*_j is the number of times the j^{th} teacher is selected.
3. For the b^{th} school bootstrap sample in step 1, $\mathbf{I}^b_{n^*_h}$, select $\mathbf{I}^b_{m^*_{i^*}}$ for each $i^* \in \mathbf{I}^b_{n^*_h}$, to get a set of conditional teacher bootstrap weights given the i^*'s, $w^*_{i^*jb} = K^*_{jb} M_{i^*} / m^*_{i^*}$ and a set of overall replicate weights, $w^*_{ijb} = \sum_{i^* \in S_{ib}} w^*_{i^*b} w^*_{i^*jb}$.
4. Repeat step $3\,B$ times for each school bootstrap sample, producing B sets of w^*_{ijb}'s.
5. Using the B sets of replicate weights in step 4, compute B estimates $\hat{T}^*_b$. The simple

variance of these B estimates is the bootstrap variance, $v^*(\hat{T}^*)$, where $V^*(\hat{T}^*) = E^* v^*(\hat{T}^*)$.

Now, choosing n_h^* as above and $m_{i^*}^* = (m_i - 1)w_{i^*}^* / (1 - f_{2i})$, it follows that $V^*(\hat{T}^*)$ is an unbiased estimator for $V(\hat{T})$, within the Taylor Series approximation (see Kaufman (2000)).

If $m_{i^*}^*$ is not an integer then it needs to be bracketed between the integer less than $m_{i^*}^*$ (m^L) and the integer greater than $m_{i^*}^*$ (m^U). m^L is selected with probability $m^L(m^U - m_{i^*}^*)/m_{i^*}^*$. If m^L is not selected then m^U is used.

This estimator is denoted by $v_D^*(\hat{T})$.

2.3 Indirectly Reflecting Second- Stage Sampling

Here the bootstrap replicate weight is the first-stage replicate weight, described in section 2.2.3 step 1, times the conditional second-stage weight given the first-stage sample. This estimator applies the first-stage FPC to both the first and second stage variance components.

One should expect this to underestimate the variance, since the first-stage FPC is applied to the second-stage variance component, when it should not be applied. This would be especially true when the first-stage sampling rates are high. This estimator is denoted by $v_I^*(\hat{T})$.

3.1 Simulation Sample Design and Frames

The simulation sample design has been described in section 2.2.1. Now, the school and teacher sampling frames will be described.

NCES does not have a list of all teachers in the elementary/secondary school system. Instead, the 8,600 public school teacher lists collected during the SASS collection, among the 50 states, will be grouped into three simulation states.

The first step in this process is to identify states that have low, medium and high SASS sampling rates and then choosing one per sampling rate category to simulate, each in a different 4 category census region. Within a Census region/SASS stratum, SASS schools providing a teacher list are randomly chosen, along with their reported teacher data, to be included in the simulation for the simulation state corresponding to the respective census region. The number of schools selected within a SASS stratum corresponds to the number of schools actually in the chosen state's school frame. This is the teacher frame and data used to select teacher samples and estimates. The corresponding school data made up the school frame and data to produce school samples and estimates. Any missing data are imputed using a sequential nearest neighbor procedure. The school and teacher sample allocation corresponds to the normal SASS allocation for these three states.

The three simulation states chosen have school sampling rates of 5, 12 and 28%. The teacher sampling rates are chosen to yield an equally weighted teacher sample, given the schools are chosen proportional to the square root of the number of teachers in the school. Additionally, the combined school stratum sampling rate is 31%.

Estimates (26 totals and proportions), based on the variables collected from teacher-listing operation, are computed by state, school level, and urbanicity, each having three categories.

3.2 Performance Statistics

$\bar{v}_D^*(\hat{T})$ and $\bar{v}_I^*(\hat{T})$ will be based on 48 bootstrap samples (B) and 460 simulations. To measure their performance, the following statistics will be compared:

3.2.1 Relative Error

$RE_e = (\bar{v}_e^*(\hat{T}) - V_t(\hat{T}))/V_t(\hat{T})$, where $V_t(\hat{T})$ is the simple variance of the simulation estimates of $\hat{T}, \hat{T}_s$ and $\bar{v}_e^*(\hat{T})$ is the average of one of the bootstrap variance estimators ($e = D$ or I) across the simulation samples.

3.2.2 Coverage Rate

The coverage rate is the percent of the time that the true estimate is within the 95% confidence intervals across the simulation samples.

3.2.3 Total Covariance Term, $Cov(\hat{T})$

Since $v_D^*(\hat{T}^*)$ is unbiased, assuming $Cov(\hat{T})$ =0, an unbiased estimator for $Cov(\hat{T}) = \left[V_t(\hat{T}) - \bar{v}_D^*(\hat{T})\right]$.

Relative $Cov(\hat{T}) = \left[V_t(\hat{T}) - \bar{v}_D^*(\hat{T})\right]/V_t(\hat{T}) = -RE_D$.

4.0 Results

The simulation is designed so that only two sources of error exist for the indirect second-stage method: 1) the use of the incorrect FPC on the second-stage variance component and 2) the assumption that $Cov(\hat{T})$ =0. Similarly, by design, the direct second-stage variance method is unbiased, assuming $Cov(\hat{T})$ =0. So, the only source of error is this assumption.

4.1 Relative Total Covariance

Table 1 provides the distribution of the relative total covariances. For the high sampling rate estimates (estimates with an overall sampling rate>25%), 15.4% of the relative covariances are extremes (< -30%). For the low sampling rate estimates (estimates with an overall sampling rate <25%), 1.0% of the relative covariances are extremes ($\geq 15\%$). This suggests: 1) assuming $Cov(\hat{T})=0$ may not always be appropriate; and 2) the relative covariance may explain a significant share of the variance errors. For this reason, the results control for the covariance whenever possible.

4.2 Relative Error

Table 2 provides the distribution of the relative errors for the two variance methodologies. For the direct second-stage method ($v_D^*(\hat{T})$), there is no variance underestimation for estimates with high sampling rates; all relative error are positive. This indicates that the first and second stage FPCs are appropriately applied. However, 15.4% of the relative errors are on the extreme positive side. For the estimates with low sampling rates, almost all relative errors are in a reasonable range, only 1.0% are in an extreme range. Again, the simulations indicate that the FPCs are appropriately applied. Since the relative covariance is estimated by $-RE_D$, the row '$v_D^*(\hat{T})+\text{cov}(\hat{T})$' in table 2, which adds the Relative $Cov(\hat{T})$ to the relative error, has relative errors of zero for all estimates (i.e., all error is due to $Cov(\hat{T})$ not being incorporated into $v_D^*(\hat{T})$).

For the indirect second-stage method ($v_I^*(\hat{T})$) with high sampling rates, given that the wrong FPC is used in the second-stage variance component, the relative errors do not look too bad, with 9.6% in the extreme negative category. The explanation for this is that the covariances for this group are all negative. With the $Cov(\hat{T})=0$ assumption, these variances should be overestimates, while the incorrect FPC should induce an underestimate. The net affect is a lower error rate than the size of the FPC would indicate. From the '$v_I^*(\hat{T})+\text{cov}(\hat{T})$' row, 65.4% of the relative errors are less than –30%. So, the errors due to the first-stage FPC bias are masked by the negative covariances.

For estimates with low sampling rates, the indirect second-stage method relative errors look very reasonable with only 2.8% in the less than –15% category. After controlling for the covariance, there are no extreme relative errors, although all are negative. When the sampling rates are not high, the first-stage FPC induces a slight variance underestimation.

4.3 Coverage Rates

Table 3 provides the distribution of the coverage rates. For the direct second-stage method with high sampling rate estimates, the coverage rates look good. Only 3.9% of the coverage rates are low (<90%). One might expect the estimates in this category would have positive covariances. However, this is not the case; 100% are negative. However, the negative covariances seems to explain the small number of coverage rates (3.9%) that are very large ($\geq$ 97%). For estimates with low sampling rates, the coverage rate distribution is very reasonable. Only 4.7% are in an extreme category with most seemingly caused by an appropriately signed covariance.

For estimates with high sampling rates, the indirect second-stage method produces a poor coverage rate distribution with 32.7% of the rates in the extremely low category. Since the covariances are all negative, which would imply higher coverage rates, the covariances are not causing this result. Seemingly, the first-stage FPC incorrectly applied to the second-stage variance component is causing the errors. For estimates with low sampling rates, the coverage rate distribution looks good. Where there are extremes, the covariances are mostly consistent with the direction of the error.

5.0 Conclusions

The object of this research is to measure the performance of the indirect variance methodology for the SASS teacher survey, a two-stage design with high first-stage sampling rates. The indirect methodology does not correctly estimate the second-stage variance component. As an alternative, the direct methodology, which correctly estimates the second-stage variance component, is proposed and compared to the indirect method.

Both methodologies assume the total covariance is zero. Results of this analysis show this can be an incorrect assumption. In terms of relative error, both methods perform acceptably, although the indirect method is aided by two biases in different directions reducing the net bias. The indirect method's poor coverage rates provide the overall conclusion that the direct method performs better.

6.0 References

Cochran, W. (1977) *Sampling Techniques*. New York: John Wiley and Sons.

Kaufman, S (2001). "A New Model for Estimating the Variance under Systematic Sampling," *Proceedings for the Section on Survey Methods, American Statistical Association*, Alexandria, Va..

Kaufman, S (2000). "Using the Bootstrap to Estimate the Variance in a Very Complex Sample Design," *Proceedings for the Section on Survey Methods, American Statistical Association*, Alexandria, Va.:ASA.

Kaufman, S (1999). "Using the Bootstrap to Estimate the Variance from a Single Systematic PPS Sample," *Proceedings for the Section on Survey Methods, American Statistical Association*, Alexandria, Va.:ASA.

Sitter, R. R. (1992). A Resampling Procedure for Complex Survey Data, *J. Amer. Statist Assoc.*, **87**.

Table 1 -- %Distribution of the relative total covariance

Design	Relative covariance $< -30\%$	-30% $\le$ Relative covariance<0%	0% $\le$ Relative covariance<15%	Relative covariance $\ge 15\%$	Min.	Max.
High Sampling Rates	15.4	84.6	0.0	0.0	-41.6	-3.0
Low Sampling Rates	0.0	50.9	48.1	1.0	-23.7	16.8

Table 2 -- % Distribution of relative error

Variance Design	Relative error $< -15\%$	-15% $\le$ Relative error<0%	0% $\le$ Relative error<30%	Relative error $\ge 30\%$	Min.	Max.
Direct Second-stage						
High Sampling Rates						
$v_D^*(\hat{T})$	0.0	0.0	84.6	15.4	3.0	41.6
$v_D^*(\hat{T}) + \text{cov}(\hat{T})^1$	0.0	0.0	100	0.0	0.0	0.0
Low Sampling Rates						
$v_D^*(\hat{T})$	1.0	48.1	50.9	0.0	-16.8	23.7
$v_D^*(\hat{T}) + \text{cov}(\hat{T})^1$	0.0	0.0	100	0.0	0.0	0.0
Indirect Second-stage						
High Sampling Rates						
$v_I^*(\hat{T})$	9.6	48.1	36.5	5.8	-25.8	36.7
$v_I^*(\hat{T}) + \text{cov}(\hat{T})^1$	65.4	34.6	0.0	0.0	-34.8	0.0
Low Sampling Rates						
$v_I^*(\hat{T})$	2.8	69.4	27.8	0.0	-20.2	17.5
$v_I^*(\hat{T}) + \text{cov}(\hat{T})^1$	0.0	99.5	0.5	0.0	-14.0	0.1

[1] This row adds the total relative covariance ($-RE_D$) to the relative error.

Table 3 -- % Distribution of Coverage Rates

Variance Design	Coverage <90%	90% $\le$ Coverage <97%	Coverage $\ge 97\%$	Coverage Min.	Coverage Max.
Direct Second-stage					
High Sampling Rates [1]	3.9	92.2	3.9	89.1	97.6
With $\text{cov}(\hat{T}) > 0$	0.0	0.0	0.0	NA	NA
With $\text{cov}(\hat{T}) < 0$	100	100	100	89.1	97.6
Low Sampling Rates [1]	4.7	95.3	0.0	86.9	96.3
With $\text{cov}(\hat{T}) > 0$	80.0	47.5	0.0	86.9	95.0
With $\text{cov}(\hat{T}) < 0$	20.0	52.5	0.0	89.3	96.3
Indirect Second-stage					
High Sampling Rates [1]	32.7	67.3	0.0	85.4	96.9
With $\text{cov}(\hat{T}) > 0$	0.0	0.0	0.0	NA	NA
With $\text{cov}(\hat{T}) < 0$	100	100	0.0	85.4	96.9
Low Sampling Rates [1]	7.1	92.9	0.0	86.5	95.4
With $\text{cov}(\hat{T}) > 0$	80.0	46.7	0.0	86.5	94.6
With $\text{cov}(\hat{T}) < 0$	20.0	53.3	0.0	89.1	95.4

[1] Row percentages sum to 100%