Conservative variance estimation for sampling designs with zero pairwise inclusion probabilities

Peter M. Aronow and Cyrus Samii*
Forthcoming at *Survey Methodology*

Abstract

We consider conservative variance estimation for the Horvitz-Thompson estimator of a population total in sampling designs with zero pairwise inclusion probabilities, known as "non-measurable" designs. We decompose the standard Horvitz-Thompson variance estimator under such designs and characterize the bias precisely. We develop a bias correction that is guaranteed to be weakly conservative (non-negatively biased) regardless of the nature of the non-measurability. The analysis sheds light on conditions under which the standard Horvitz-Thompson variance estimator performs well despite non-measurability and the conservative bias correction may outperform commonly-used approximations.

Keywords: Horvitz-Thompson estimation, non-measurable designs, variance estimation

1 Introduction

Sampling designs sometimes result in pairs of units having zero probability of being jointly included in the sample. Horvitz and Thompson (1952)'s statement of the properties of the finite population total makes clear that general, unbiased variance estimation for estimators of population totals is impossible for such *non-measurable* designs (Särndal et al., 1992, 33). Optimal methods for variance estimation in these cases remains an open problem to this day. This paper analyzes the nature of the biases that non-measurability introduces for the standard Horvitz-Thompson estimator and studies an approach to correct for this bias in a manner that is guaranteed to be conservative. While our results cannot offer a solution to the non-measurability problem for all practical applications, we

*Peter M. Aronow, Department of Political Science, Yale University, 77 Prospect St., New Haven, CT 06520, Email: peter.aronow@yale.edu; Cyrus Samii, Department of Politics, New York University, 19 West 4th St., New York, NY 10012, Email: cds2083@nyu.edu

do clarify conditions under which the standard estimator performs well and where the conservative bias correction outperforms commonly-used approximations.

Despite their theoretical drawbacks, sampling designs with zero pairwise inclusion probabilities are quite common. A common example of a non-measurable design is one that draws only single units or clusters from a set of strata. This may occur if the population or a subpopulation of interest is incidentally sparse over stratification cells. Another instance of a non-measurable design is a systematic sample in which unit indices are sampled from a list in multiples from a random starting value. In these designs, units whose indices are multiples from different starting values have zero joint probability of inclusion.

Approximate methods have been proposed for special cases, as discussed in Hansen et al. (1953, Section 9.15), Särndal et al. (1992, Chapter 3), and Wolter (2007, Chapters 2 and 8). In the single-unit per stratum case, a common approach is to collapse strata and assume units were drawn via a simple random sample from the larger, collapsed stratum. For systematic samples, the standard approach is to use an approximation based on an assumption of simple random sampling with replacement. These approximate methods are generally biased to a degree that cannot be determined from the data. In some cases, it can be shown that the bias will tend to be positive, but such is not the case generally, and especially so when the zero pairwise inclusion probabilities occur in a haphazard manner.

This paper begins by decomposing the bias of the Horvitz-Thompson variance estimator under non-measurability. This exposes precisely how conditions on the underlying data result in more or less bias. We also show how a simple application of Young's inequality yields a bias correction and a class of estimators guaranteed to have weakly positive bias as well as no bias under special conditions. We discuss implications for applied work.

2 Variance estimation for the Horvitz-Thompson estimator

Consider a population U indexed by $1, ..., k, ..., N$ and a sampling design such that the probability of inclusion in the sample for unit k is given by π_k, and the joint inclusion probability for units k and l is given by π_{kl}. Under a measurable design, two conditions obtain: (1) $\pi_k > 0$ and π_k is known for all $k \in U$ and (2) $\pi_{kl} > 0$ and π_{kl} is known for all $k, l \in U$. Non-measurable designs include those for which either of the two conditions for a measurable design do not hold. Failure to meet the former condition precludes unbiased estimation of totals.

The Horvitz-Thompson estimator of a population total is given by

$$\hat{t} = \sum_{k \in s} \frac{y_k}{\pi_k} = \sum_{k \in U} I_k \frac{y_k}{\pi_k},$$

where $I_k \in \{0,1\}$ is unit k's inclusion indicator, the only stochastic component of the expression, with $\mathrm{E}(I_k) = \pi_k$, the inclusion probability, and s and U refer to the sample and the population, respectively. Define $\mathrm{E}(I_k I_l) = \pi_{kl}$, which is the probability that both units k and l from U are included in the sample. Since $I_k I_k = I_k$, $\mathrm{E}(I_k I_k) = \pi_{kk} = \pi_k$ by construction. When condition (1) holds, as we assume throughout, the Horvitz-Thompson estimator is unbiased.

2.1 Properties of the Horvitz-Thompson variance estimator under measurability

By Horvitz and Thompson (1952), the variance of the Horvitz-Thompson estimator for the total,

$$\begin{aligned}\mathrm{Var}(\hat{t}) &= \sum_{k\in U}\sum_{l\in U}\mathrm{Cov}(I_k, I_l)\frac{y_k}{\pi_k}\frac{y_l}{\pi_l} \\ &= \sum_{k\in U}\mathrm{Var}(I_k)\left(\frac{y_k}{\pi_k}\right)^2 + \sum_{k\in U}\sum_{l\in U\setminus k}\mathrm{Cov}(I_k, I_l)\frac{y_k}{\pi_k}\frac{y_l}{\pi_l}.\end{aligned}$$

We label a sample from a measurable design, s^M, and an unbiased estimator for $\mathrm{Var}(\hat{t})$ on s^M is given by,

$$\widehat{\mathrm{Var}}(\hat{t}) = \sum_{k\in s^M}\sum_{l\in s^M}\frac{\mathrm{Cov}(I_k, I_l)}{\pi_{kl}}\frac{y_k}{\pi_k}\frac{y_l}{\pi_l} = \sum_{k\in U}\sum_{l\in U} I_k I_l \frac{\mathrm{Cov}(I_k, I_l)}{\pi_{kl}}\frac{y_k}{\pi_k}\frac{y_l}{\pi_l},$$

where the only stochastic part of the latter expression is $I_k I_l$, and unbiasedness is due to $\mathrm{E}(I_k I_l) = \pi_{kl}$.

2.2 Properties of the Horvitz-Thompson variance estimator under non-measurability

We now examine the case where condition 2 does not hold: $\pi_{kl} = 0$ for some units $k, l \in U$. Because I_k is a Bernoulli random variable with probability π_k, $\mathrm{Cov}(I_k, I_l) = \pi_{kl} - \pi_k\pi_l$ for $k \neq l$, and $\mathrm{Cov}(I_k, I_k) = \mathrm{Var}(I_k) = \pi_k(1-\pi_k)$. Then, we can re-express the variance above as,

$$\begin{aligned}\mathrm{Var}(\hat{t}) &= \sum_{k\in U}\pi_k(1-\pi_k)\left(\frac{y_k}{\pi_k}\right)^2 + \sum_{k\in U}\sum_{l\in U\setminus k}(\pi_{kl} - \pi_k\pi_l)\frac{y_k}{\pi_k}\frac{y_l}{\pi_l} \\ &= \sum_{k\in U}\pi_k(1-\pi_k)\left(\frac{y_k}{\pi_k}\right)^2 + \sum_{k\in U}\sum_{l\in\{U\setminus k:\pi_{kl}>0\}}(\pi_{kl} - \pi_k\pi_l)\frac{y_k}{\pi_k}\frac{y_l}{\pi_l} - \underbrace{\sum_{k\in U}\sum_{l\in\{U\setminus k:\pi_{kl}=0\}} y_k y_l}_{A}.\end{aligned}$$

For k and l such that $\pi_{kl} = 0$, the sampling design will never provide information on the component of the variance labeled as A above, since we will never observe y_k and y_l together. We label a sample from a design where condition 2 fails as s^0. When $\widehat{\text{Var}}\left(\hat{t}\right)$ is applied to s^0, the result is unbiased for $\text{Var}\left(\hat{t}\right) + A$. We state this formally as follows:

Proposition 1. *When s^0 refers to a sample from a design with some $\pi_{kl} = 0$, we have,*

$$\text{E}\left[\widehat{\text{Var}}\left(\hat{t}\right)\right] = \text{Var}\left(\hat{t}\right) + \sum_{k\in U}\sum_{l\in\{U\backslash k:\pi_{kl}=0\}} y_k y_l = \text{Var}\left(\hat{t}\right) + A.$$

Proof. The result follows from,

$$\begin{aligned}
\text{E}\left[\sum_{k\in s^0}\sum_{l\in s^0}\frac{\text{Cov}\left(I_k, I_l\right)}{\pi_{kl}}\frac{y_k}{\pi_k}\frac{y_l}{\pi_l}\right] &= \text{E}\left[\sum_{k\in U}\sum_{l\in\{U:\pi_{kl}>0\}} I_k I_l \frac{\text{Cov}\left(I_k, I_l\right)}{\pi_{kl}}\frac{y_k}{\pi_k}\frac{y_l}{\pi_l}\right] \\
&= \sum_{k\in U}\text{Var}\left(I_k\right)\left(\frac{y_k}{\pi_k}\right)^2 + \sum_{k\in U}\sum_{l\in\{U\backslash k:\pi_{kl}>0\}}\text{Cov}\left(I_k, I_l\right)\frac{y_k}{\pi_k}\frac{y_l}{\pi_l} \\
&= \text{Var}\left(\hat{t}\right) + \sum_{k\in U}\sum_{l\in\{U\backslash k:\pi_{kl}=0\}} y_k y_l \\
&= \text{Var}\left(\hat{t}\right) + A.
\end{aligned}$$

$\square$

The standard Horvitz-Thompson variance estimator, if applied to designs with zero pairwise inclusion probabilities, can therefore have a positive or a negative bias. If the y_k, y_l values are always nonnegative (or always nonpositive), then the bias is always nonnegative. When values may be positive or negative, A is the sum of cross-products of outcomes that never appear together under the design sample. If the jointly exclusive outcomes are centered over zero, then no correlation in these outcomes would tend to result in small bias, positive correlation in positive bias, and negative correlation in negative bias.

3 Conservative bias correction for the Horvitz-Thompson estimator under non-measurability

The case where A may be less than zero for a non-measurable design suggests the need for some adjustment that will guarantee a bias that is weakly bounded below by zero. We first develop a general bias correction that is guaranteed to be conservative, later providing a special simplified case for practical usage.

3.1 General formulation

Consider the following variance estimator:

$$\widehat{\mathrm{Var}}_C(\hat{t}) = \sum_{k\in U}\sum_{l\in\{U:\pi_{kl}>0\}} I_k I_l \frac{\mathrm{Cov}\,(I_k, I_l)}{\pi_{kl}}\frac{y_k}{\pi_k}\frac{y_l}{\pi_l} + \sum_{k\in U}\sum_{l\in\{U\setminus k:\pi_{kl}=0\}}\left(I_k\frac{|y_k|^{a_{kl}}}{a_{kl}\pi_k} + I_l\frac{|y_l|^{b_{kl}}}{b_{kl}\pi_l}\right),$$

where a_{kl}, b_{kl} are positive real numbers such that $\frac{1}{a_{kl}} + \frac{1}{b_{kl}} = 1$ for all pairs k, l with $\pi_{kl} = 0$. The estimator is guaranteed to produce an expected value greater than or equal to the true variance for all designs, and is thus conservative. We state this condition of the estimator formally:

Proposition 2.

$$\mathrm{E}\left[\widehat{\mathrm{Var}}_C(\hat{t})\right] \geq \mathrm{Var}\,(\hat{t}).$$

Proof. By Young's inequality,

$$\frac{|y_k|^{a_{kl}}}{a_{kl}} + \frac{|y_l|^{b_{kl}}}{b_{kl}} \geq |y_k||y_l|,$$

if $\frac{1}{a_{kl}} + \frac{1}{b_{kl}} = 1$. Define A^* such that,

$$A^* = \sum_{k\in U}\sum_{l\in\{U\setminus k:\pi_{kl}=0\}} \frac{|y_k|^{a_{kl}}}{a_{kl}} + \frac{|y_l|^{b_{kl}}}{b_{kl}} \geq \sum_{k\in U}\sum_{l\in\{U\setminus k:\pi_{kl}=0\}} |y_k||y_l| \geq \sum_{k\in U}\sum_{l\in\{U\setminus k:\pi_{kl}=0\}} y_k y_l = A$$

and

$$A^* \geq \sum_{k\in U}\sum_{l\in\{U\setminus k:\pi_{kl}=0\}} |y_k||y_l| \geq \sum_{k\in U}\sum_{l\in\{U\setminus k:\pi_{kl}=0\}} -y_k y_l = -A.$$

Therefore

$$\mathrm{Var}\,(\hat{t}) + A + A^* \geq \mathrm{Var}\,(\hat{t}).$$

The associated Horvitz-Thompson estimator of A^* would be

$$\widehat{A^*} = \sum_{k\in U}\sum_{l\in\{U\setminus k:\pi_{kl}=0\}}\left(I_k\frac{|y_k|^{a_{kl}}}{a_{kl}\pi_k} + I_l\frac{|y_l|^{b_{kl}}}{b_{kl}\pi_l}\right),$$

which is unbiased by $\mathrm{E}\,(I_k) = \pi_k$ and $\mathrm{E}\,(I_l) = \pi_l$.

Since $\mathrm{E}\left[\widehat{A^*}\right] = A^*$, by Proposition 1,

$$\mathrm{E}\left[\sum_{k\in s^0}\sum_{l\in s^0}\frac{\mathrm{Cov}\,(I_k, I_l)}{\pi_{kl}}\frac{y_k}{\pi_k}\frac{y_l}{\pi_l} + \widehat{A^*}\right] = \mathrm{Var}\,(\hat{t}) + A + A^*$$

$$\mathrm{E}\left[\widehat{\mathrm{Var}}_C(\hat{t})\right] \geq \mathrm{Var}\,(\hat{t}). \tag{1}$$

Substituting terms,

$$\mathrm{E}\left[\sum_{k\in U}\sum_{l\in\{U:\pi_{kl}>0\}} I_k I_l \frac{\mathrm{Cov}\,(I_k, I_l)}{\pi_{kl}}\frac{y_k}{\pi_k}\frac{y_l}{\pi_l} + \sum_{k\in U}\sum_{l\in\{U\setminus k:\pi_{kl}=0\}}\left(I_k\frac{|y_k|^{a_{kl}}}{a_{kl}\pi_k} + I_l\frac{|y_l|^{b_{kl}}}{b_{kl}\pi_l}\right)\right] \geq \mathrm{Var}\,(\hat{t}).$$

□

$\widehat{\mathrm{Var}}_C(\hat{t})$ is justified as a conservative estimator for the case when A is not known to be positive. This estimator is unbiased under a special condition:

Corollary 1. *If, for all pairs k, l such that $\pi_{kl} = 0$, (i) $|y_k|^{a_{kl}} = |y_l|^{b_{kl}}$ and (ii) $-y_k y_l = |y_k||y_l|$,*

$$\mathrm{E}\left[\widehat{\mathrm{Var}}_C(\hat{t})\right] = \mathrm{Var}\,(\hat{t}).$$

Proof. By (i), (ii) and Young's inequality,

$$\frac{|y_k|^{a_{kl}}}{a_{kl}} + \frac{|y_l|^{b_{kl}}}{b_{kl}} = |y_k||y_l| = -y_k y_l.$$

Therefore,

$$A^* = \sum_{k\in U}\sum_{l\in\{U\setminus k:\pi_{kl}=0\}} \frac{|y_k|^{a_{kl}}}{a_{kl}} + \frac{|y_l|^{b_{kl}}}{b_{kl}} = \sum_{k\in U}\sum_{l\in\{U\setminus k:\pi_{kl}=0\}} |y_k||y_l| = \sum_{k\in U}\sum_{l\in\{U\setminus k:\pi_{kl}=0\}} -y_k y_l = -A$$

It follows that

$$\mathrm{Var}\,(\hat{t}) + A + A^* = \mathrm{Var}\,(\hat{t})$$

and

$$\mathrm{E}\left[\widehat{\mathrm{Var}}_C(\hat{t})\right] = \mathrm{Var}\,(\hat{t}).$$

□

If any units k, l are in clusters (i.e., $\mathrm{Pr}\,(I_k \neq I_l) = 0$), these units should be totaled into one larger unit before estimation. Combining units will, in general, reduce the bias of the variance estimator because only pairs of cluster-level totals will be included in A^*, as opposed to all constituent pairs.

3.2 Simplified special case

In general, it would be difficult to assign optimal values of a_{kl} and b_{kl} for all pairs k, l such that $\pi_{kl} = 0$. Instead, we examine one intuitive case, assigning all $a_{kl} = b_{kl} = 2$:

$$\widehat{\mathrm{Var}}_{C2}(\hat{t}) = \sum_{k\in U}\sum_{l\in\{U:\pi_{kl}>0\}} I_k I_l \frac{\mathrm{Cov}\,(I_k, I_l)}{\pi_{kl}}\frac{y_k}{\pi_k}\frac{y_l}{\pi_l} + \sum_{k\in U}\sum_{l\in\{U\setminus k:\pi_{kl}=0\}}\left(I_k\frac{y_k^2}{2\pi_k} + I_l\frac{y_l^2}{2\pi_l}\right).$$

As a special case of $\widehat{\mathrm{Var}}_C(\hat{t})$, $\widehat{\mathrm{Var}}_{C2}(\hat{t})$ is also conservative:

Corollary 2.
$$\mathrm{E}\left[\widehat{\mathrm{Var}}_{C2}(\hat{t})\right] \geq \mathrm{Var}\,(\hat{t}).$$

Proof. For all pairs k, l such that $\pi_{kl} = 0$, $\frac{1}{a_{kl}} + \frac{1}{b_{kl}} = \frac{1}{2} + \frac{1}{2} = 1$. Proposition 1 therefore holds. $\square$

The choice to set all $a_{kl} = b_{kl} = 2$ is justified by the fact that it will yield the lowest value of the estimator $\widehat{\mathrm{Var}}_C(\hat{t})$ subject to the constraint that a_{kl} and b_{kl} are fixed as constants a and b over all k, l.

Corollary 3. *Among the class of estimators $\widehat{\mathrm{Var}}_{Cab}(\hat{t})$, defined as the set of estimators $\widehat{\mathrm{Var}}_C(\hat{t})$ such that all $a_{kl} = a$ and all $b_{kl} = b$, $\widehat{\mathrm{Var}}_{C2}(\hat{t}) = \min_{a,b}\left[\widehat{\mathrm{Var}}_{Cab}(\hat{t})\right]$.*

Proof. By simple algebra,

$$\begin{aligned}
\widehat{\mathrm{Var}}_{Cab}(\hat{t}) &= \sum_{k\in U}\sum_{l\in\{U:\pi_{kl}>0\}} I_k I_l \frac{\mathrm{Cov}\,(I_k, I_l)}{\pi_{kl}} \frac{y_k}{\pi_k}\frac{y_l}{\pi_l} + \sum_{k\in U}\sum_{l\in\{U\setminus k:\pi_{kl}=0\}} \left(I_k \frac{|y_k|^a}{a\pi_k} + I_l \frac{|y_l|^b}{b\pi_l}\right) \\
&= \widehat{\mathrm{Var}}\,(\hat{t}) + \sum_{k\in U}\sum_{l\in\{U\setminus k:\pi_{kl}=0\}} \left(I_k \frac{|y_k|^a}{a\pi_k} + I_l \frac{|y_l|^b}{b\pi_l}\right) \\
&= \widehat{\mathrm{Var}}\,(\hat{t}) + \frac{1}{2}\sum_{k\in U}\sum_{l\in\{U\setminus k:\pi_{kl}=0\}} \left(I_k \frac{|y_k|^a}{a\pi_k} + I_l \frac{|y_l|^b}{b\pi_l} + I_k \frac{|y_k|^b}{b\pi_k} + I_l \frac{|y_l|^a}{a\pi_l}\right) \\
&= \widehat{\mathrm{Var}}\,(\hat{t}) + \frac{1}{2}\sum_{k\in U}\sum_{l\in\{U\setminus k:\pi_{kl}=0\}} \left(\frac{I_k}{\pi_k}\left(\frac{|y_k|^a}{a} + \frac{|y_k|^b}{b}\right) + \frac{I_l}{\pi_l}\left(\frac{|y_l|^a}{a} + \frac{|y_l|^b}{b}\right)\right).
\end{aligned}$$

By Young's inequality, given $\frac{1}{a} + \frac{1}{b} = 1$, $\frac{|y_k|^a}{a} + \frac{|y_k|^b}{b} \geq y_k^2$, but equality must hold if $a = b = 2$. Similarly, $\frac{|y_l|^a}{a} + \frac{|y_l|^b}{b} \geq y_l^2$, but equality must hold if $a = b = 2$. Since $\frac{I_k}{\pi_k} \geq 0$ and $\frac{I_k}{\pi_k} \geq 0$, any choice $a \neq b$ can only yield $\widehat{\mathrm{Var}}_{Cab}(\hat{t}) \geq \widehat{\mathrm{Var}}_{C2}(\hat{t})$. $\square$

Given all values of y_k and y_l, it is possible to derive an optimal vector of a_{kl} and b_{kl} values that varies over k, l, but such a derivation may not be of practical value.

4 Applications

Proposition 1 indicates that the bias of the Horvitz-Thompson estimator under non-measurability is given by,
$$A = \sum_{k\in U}\sum_{l\in\{U\setminus k:\pi_{kl}=0\}} y_k y_l.$$
This expression, along with the fact that $A^* \geq A$, makes it evident that the degree of bias in $\widehat{\mathrm{Var}}\,(\hat{t})$ and $\widehat{\mathrm{Var}}_C(\hat{t})$ depends a great deal on the number of pairs with zero pairwise

inclusion probabilities. For designs where this number is small, $\widehat{\text{Var}}\,(\hat{t})$ may provide a reasonable and conservative estimator for cases where y_k takes the same sign for all k, and $\widehat{\text{Var}}_C(\hat{t})$ may provide a reasonable and conservative estimator for cases where y_k may take different signs for some k. An example that arises frequently is stratified sampling where for a relatively small proportion of cases, we have small strata from which we draw only one unit.

For designs that result in many pairs having zero inclusion probabilities, $\widehat{\text{Var}}\,(\hat{t})$ and $\widehat{\text{Var}}_C(\hat{t})$ could be wildly over-conservative and other estimators may be preferred in terms of criteria such as mean square error. A prominent example is systematic sampling. Indeed, Särndal et al. (1992, 76) propose that under systematic sampling, the Horvitz-Thompson variance estimator, $\widehat{\text{Var}}\,(\hat{t})$, can give a "non-sensical result." The expression for A makes it clear why this would be the case. Wolter (2007, Ch. 8) shows that simpler biased estimators, such as the with-replacement (Hansen-Hurwitz) variance estimator, can be reliable, if slightly conservative, in a broad range of data scenarios under equal probability and probability proportional to size (PPS) systematic sampling. Nonetheless, it is known that the with-replacement estimator fails to account adequately for sampling variance when outcome variance within systematic sample clusters is smaller than the between cluster variance. In such cases, $\widehat{\text{Var}}\,(\hat{t})$ would bound this variance in expectation when outcomes are all of the same sign, and $\widehat{\text{Var}}_C(\hat{t})$ would always bound this variance in expectation. Of course, it may still be the case that the bias is too large to be of much use, and so we would not suggest that $\widehat{\text{Var}}\,(\hat{t})$ and $\widehat{\text{Var}}_C(\hat{t})$ provides a full solution to the variance estimation problem for systematic sampling under high intra-cluster correlation.

Results from simulation studies are available in a supplement and $\widehat{\text{Var}}_{C2}(\hat{t})$ perform relative to other, commonly-used approximations in these applied scenarios. The simulations demonstrate situations when these estimators are preferable to commonly used alternatives. In the case of one-unit-per-stratum sampling, we show that these estimators are less biased than the commonly used "collapsed stratum" estimator in a range of scenarios. In the case of PPS systematic sampling, these estimators perform favorably when the population exhibits substantial periodicity, a case when the commonly-used with-replacement estimator may be grossly negatively biased. [These simulation results are appended to the manuscript.]

5 Conclusion

We have characterized precisely the bias of the Horvitz-Thompson estimator under non-measurability and used this characterization to develop a conservative bias correction. These estimators reflect the fundamental uncertainty inherent to non-measurable designs. Compared to available approximate methods, these estimators may sometimes perform better and sometimes worse from a practical perspective. But available approximate methods may be biased in ways that cannot always be evaluated in terms of either magnitude or sign. The estimators developed in this paper may therefore provide an informative measure of sampling variability with which analysts can agree without invoking additional assumptions or resorting to methods that carry the potential for negative bias. The bias term, A, has a simple form that suggests the possibility of refinements to the estimators developed here, something that we leave open for future research.

References

Hansen, M. H., W. N. Hurwitz, and W. G. Madow (1953). *Sample Survey Methods and Theory: Volume 1 Methods and Theory*. New York: Wiley.

Horvitz, D. and D. Thompson (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association 47*(260), 663–685.

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

Wolter, K. (2007). *Introduction to Variance Estimation, Second Edition*. New York: Springer.

Conservative variance estimation for sampling designs with zero pairwise inclusion probabilities

Supplement: Simulation Studies

A Introduction

This supplement contains a simulation study of applications of the estimators discussed in the paper, "Conservative variance estimation for sampling designs with zero pairwise inclusion probabilities." In the simulations, expected values of $\widehat{\mathrm{Var}}\,(\hat{t})$ and $\widehat{\mathrm{Var}}_C(\hat{t})$ were computed exactly using expressions from Propositions 1 and 2 in the main text. Expected values for alternative estimators were computed using exact expressions for bias when available and via simulated sample draws otherwise. All simulations study expected values of estimators for the sampling variance of the Horvitz-Thompson estimator for the population mean,

$$\hat{\mu}_{HT} \equiv \frac{1}{N}\hat{t}.$$

All simulations were carried out in the R statistical computing environment (the code is available from the authors).

B "Lonely" primary sampling units in stratified sampling

The first set of simulation studies examines variance estimation when single units are drawn from strata under stratified sampling. This situation arises in a number of applied contexts, including stratified sampling under proportional allocation (Lohr, 1999, 104-6) as well as subgroup analyses using subgroups that are sparsely distributed over strata. We build our simulations to match the context considered by Wolter (2007, 50-57) in his study of the collapsed stratum estimator. Simplifying that context, we consider two strata each of size M, in which case $N = 2M$. From each stratum a single unit is randomly sampled. Thus, for unit k in stratum s,

$$\pi_{ks} = \pi = \frac{1}{M},$$

and for units k and l in strata 1 or 2, we have

$$\pi_{ks,lt} = \begin{cases} 0 & \text{if } s = t \\ \frac{1}{M^2} & \text{if } s \neq t \end{cases}.$$

In stratum s, outcomes $y_s = (y_{1s}, ..., y_{Ms})$ are drawn as,

$$y_{ks} = \alpha_s + \epsilon_{is},$$

where α_s is a stratum-specific constant and $\epsilon_{ks} \sim \mathrm{N}(0,1)$. For the two strata, 1 and 2, we keep things simple by setting $\alpha_1 = \alpha = -\alpha_2$, in which case the intra-class correlation over strata is given by,

$$\rho_{ICC} = \frac{\alpha^2}{\alpha^2 + 1/2}.$$

We simulated 1,000 sets of outcomes, (y_1, y_2) over varying values of M and ρ_{ICC}. To each set, we applied the single-unit-per-stratum design. Then, we computed three approximations of the sampling variance of the Horvitz-Thompson estimator of the population mean:

1. The Horvitz-Thompson estimator for the sampling variance of the estimator for the population mean, $(1/N^2)\widehat{\mathrm{Var}}(\hat{t})$.
2. The bias-adjusted estimator, $(1/N^2)\widehat{\mathrm{Var}}_{C2}(\hat{t})$.
3. The "collapsed stratum" estimator (Wolter, 2007, 51-2), which for this case, given unit k from stratum 1 and unit l from stratum 2 selected into s^0, equals,

$$\widehat{V}_{CS}(\hat{\mu}_{HT}) = \frac{1}{N^2}\left(\frac{y_{k1}}{\pi_1} - \frac{y_{l2}}{\pi_2}\right)^2 = \frac{1}{4}(y_{k1} - y_{l2})^2.$$

Because of the zero inclusion probabilities within strata, each of these estimators is biased. No generally unbiased estimator exists for the sampling variance of the population mean estimator in this scenario. The collapsed stratum estimator is a commonly used approximation in the applied literature.

Table A1 and Figure A1 display results from these simulations. For each ρ_{ICC} and M value, we used the empirical variance of $\hat{\mu}_{HT}$ as an estimate of the true variance (Var or "True"), and means of the computed $(1/N^2)\widehat{\mathrm{Var}}(\hat{t})$ ($\widehat{\mathrm{Var}}_{HT}$ or "Horvitz-Thompson"), $(1/N^2)\widehat{\mathrm{Var}}_{C2}(\hat{t})$ ($\widehat{\mathrm{Var}}_{C2}$ or "Conservative (A*)"), and $\widehat{V}_{CS}(\hat{\mu}_{HT})$ values as estimates of the expected values of these estimators. In these simulations, $(1/N^2)\widehat{\mathrm{Var}}(\hat{t})$ strictly dominates $(1/N^2)\widehat{\mathrm{Var}}_{C2}(\hat{t})$ and $\widehat{V}_{CS}(\hat{\mu}_{HT})$ as an approximation to the true sampling variance. The $(1/N^2)\widehat{\mathrm{Var}}_{C2}(\hat{t})$ estimator dominates $\widehat{V}_{CS}(\hat{\mu}_{HT})$ for smaller stratum sizes and the two perform similarly for larger stratum sizes. When ρ_{ICC} is close to 1, all approximations perform very badly. The reasons for this differ for the estimators, however. The bias of $\widehat{V}_{CS}(\hat{\mu}_{HT})$ is driven up directly by between stratum variance. The bias of $(1/N^2)\widehat{\mathrm{Var}}(\hat{t})$ and $(1/N^2)\widehat{\mathrm{Var}}_{C2}(\hat{t})$ is driven by the fact that within-stratum values become less symmetric over zero as α increases.

C Systematic probability-proportional-to-size sampling

A second set of simulations is based on systematic probability-proportional-to-size (PPS) sampling, another commonly used design (Lohr, 1999, 185-6; Wolter, 2007, 332-5). The problem that arises for variance estimation under systematic sampling, whether PPS or not, is that it produces zero pairwise inclusion probabilities among population units that are proximate on the list used to take the systematic sample.

For the simulations, a size variable, $x = (x_1, ..., x_N)$, is produced as a vector of N random draws from a negative binomial distribution with mean 6, variance 16, and offset of 1 to avoid zero values. These settings were meant to recreate a data scenario roughly comparable to Wolter (2007, Ch. 8). An outcome variable, y, is created as,

$$y_k = \beta x_k + \epsilon_k,$$

for $k = 1, ..., N$, where $\beta = \rho(\sigma_y/\sigma_x)$, ϵ_k is Gaussian noise with mean zero and standard deviation $\sigma_y\sqrt{1-\rho^2}$, and σ_x and σ_y are the standard deviation of the N values in x and y respectively, where $\mathrm{E}_x(\sigma_x) = 4$ and we fix $\sigma_y = 30$ (again, to mimic data distributed as in Wolter (2007, Ch. 8)). Thus, x and y have a linear relationship through the origin and their correlation is given by ρ.

Systematic PPS sampling was implemented such that each unit's first order inclusion probability is given by,

$$\pi_k = \frac{x_k}{\sum_{l=1}^{N} x_l} m,$$

where m is the size of the systematic sample, and joint inclusion probabilities are computed as per Madow (1949). For the simulations, sampling and computation of inclusion probabilities were carried out using functions provided by Tillé and Matei (2011).

Because y is proportional to x with $x \geq 1$, Proposition 1 in the main text indicates that the Horvitz-Thompson variance estimator for $\hat{\mu}_{HT}$, $(1/N^2)\widehat{\mathrm{Var}}\,(\hat{t})$, provides a conservative approximation of the sampling variance of $\hat{\mu}_{HT}$. For comparison, we also computed (via 1000 simulated sample draws) the expected value of the with-replacement variance estimator for $\hat{\mu}_{HT}$ (Wolter, 2007, expr. 8.7.2),

$$\widehat{\mathrm{Var}}_{WR}(\hat{\mu}_{HT}) = \frac{1}{N^2 m(m-1)} \sum_{k \in s^0} \left(\frac{y_k}{p_k} - \hat{t} \right)^2,$$

where $p_k = \pi_k/m$ is the notional "per-draw" selection probability. $\widehat{\mathrm{Var}}_{WR}(\hat{\mu}_{HT})$ is regularly used for systematic samples in applied settings.

Table A2 and Figure A2 display results from simulations for this scenario. Each graph plots the true variance (Var or "True"), expected value of the $(1/N^2)\widehat{\mathrm{Var}}\,(\hat{t})$ ($\widehat{\mathrm{Var}}_{HT}$ or "Conservative (HT)"), and expected value of $\widehat{\mathrm{Var}}_{WR}(\hat{\mu}_{HT})$ ("With-repl.") over values of ρ and for varying sample sizes (m) and sampling proportions (m/N). As expected,

$\widehat{\mathrm{Var}}(\hat{t})$ tends to produce very over-conservative estimates; this is due to the large number of pairs with zero-inclusion probabilities, each of which contributes to the bias. Given that the size variable is randomly sorted, this data scenario is very favorable to $\widehat{\mathrm{Var}}_{WR}(\hat{\mu}_{HT})$, as is also evident from the results.

For systematic sampling, a more challenging inference scenario arises when the data exhibit periodicity. In such a case, the within-sample variance fails to reflect the total of within and between-sample variance, potentially rendering approximations such as the with-replacement estimator anti-conservative (Särndal et al., 1992, 78-83). To illustrate this point, we performed a second set of simulations where the size variable, x, was drawn as an N-length vector of repeating 2's and 10's, thus maintaining variance values that are comparable to the previous simulation. Also, so that we could provide an illustration of the performance of $\widehat{\mathrm{Var}}_{C2}(\hat{t})$, we produced y_k values in the same manner as above, but then worked with the demeaned value,

$$\tilde{y}_k \equiv y_k - \bar{y},$$

as the variable with which $\hat{\mu}_{HT}$ was computed. By Propositions 1 and 2, $\widehat{\mathrm{Var}}(\hat{t})$ may have either positive or negative bias, while $\widehat{\mathrm{Var}}_{C2}(\hat{t})$ yields a conservative approximation of the sampling variance of $\hat{\mu}_{HT}$.

Table A3 and Figure A3 display the results of these simulations. Each graph plots the true variance (Var or "True"), expected value of $(1/N^2)\widehat{\mathrm{Var}}(\hat{t})$ ($\widehat{\mathrm{Var}}_{HT}$ or "Conservative (HT)"), expected value of $(1/N^2)\widehat{\mathrm{Var}}_{C2}(\hat{t})$ ($\widehat{\mathrm{Var}}_{C2}$ or "Conservative (A*)"), and expected value of $\widehat{\mathrm{Var}}_{WR}(\hat{\mu}_{HT})$ ("With-repl.") over values of ρ and for varying sample sizes (m) and sampling proportions (m/N). As expected, $\widehat{\mathrm{Var}}_{WR}(\hat{\mu}_{HT})$ tends to greatly understate the sampling variance of $\hat{\mu}_{HT}$, while $\widehat{\mathrm{Var}}_{C2}(\hat{t})$ is generally quite conservative. Interestingly, $\widehat{\mathrm{Var}}(\hat{t})$ tends to match the true value of the variance almost perfectly. This occurs because of the special nature of the data generating process, which results in the elements in the bias term, A, canceling each other out, in which case A tends to 0.

References

Lohr, S. L. (1999). *Sampling: Design and Analysis*. Pacific Grove: Brooks/Cole.

Madow, W. G. (1949). "On the theory of systematic sampling, II." *Annals of Mathematical Statistics 20*(3), 333–354.

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

Tillé, Y, and A. Matei. (2011). *sampling: Survey Sampling*. R package, version 2.4. http://cran.R-project.org/package=sampling

Wolter, K. M. (2007) *Introduction to Variance Estimation, Second Edition*. New York: Springer.

Table A1: Simulation results for "lonely" primary sampling units in stratified sampling for different values of ρ_{ICC} and strata sizes (M).

	Sim. Params.		True	Mean from sims.			Relative Bias			Nominal Bias		
	ρ_{ICC}	M	Var	$\widehat{\text{Var}}_{HT}$	$\widehat{\text{Var}}_{C2}$	$\widehat{\text{Var}}_{CS}$	$\widehat{\text{Var}}_{HT}$	$\widehat{\text{Var}}_{C2}$	$\widehat{\text{Var}}_{CS}$	$\widehat{\text{Var}}_{HT}$	$\widehat{\text{Var}}_{C2}$	$\widehat{\text{Var}}_{CS}$
1	0.00	2	0.25	0.25	0.49	0.49	0.00	1.00	1.01	0.00	0.25	0.25
6	0.33	2	0.25	0.30	0.61	0.73	0.24	1.48	1.97	0.06	0.36	0.48
11	0.50	2	0.25	0.38	0.76	1.00	0.50	2.01	2.93	0.13	0.51	0.75
16	0.67	2	0.25	0.50	1.00	1.51	1.04	3.09	5.14	0.26	0.76	1.26
21	0.82	2	0.25	0.82	1.64	2.77	2.33	5.65	10.27	0.57	1.39	2.53
26	0.90	2	0.25	1.40	2.79	5.09	4.60	10.20	19.40	1.15	2.54	4.84
31	0.99	2	0.25	14.30	28.60	56.70	55.77	112.54	224.12	14.05	28.34	56.44
2	0.00	5	0.39	0.39	0.79	0.49	0.00	1.03	0.26	0.00	0.40	0.10
7	0.33	5	0.40	0.50	0.99	0.75	0.24	1.48	0.86	0.10	0.59	0.35
12	0.50	5	0.40	0.60	1.21	1.01	0.52	2.05	1.55	0.21	0.81	0.61
17	0.67	5	0.40	0.79	1.57	1.47	0.98	2.97	2.69	0.39	1.18	1.07
22	0.82	5	0.40	1.30	2.61	2.76	2.28	5.56	5.94	0.91	2.21	2.36
27	0.90	5	0.41	2.26	4.52	5.15	4.58	10.16	11.72	1.86	4.12	4.75
32	0.99	5	0.40	22.94	45.87	56.84	56.59	114.17	141.71	22.54	45.48	56.44
3	0.00	20	0.47	0.47	0.95	0.50	0.00	1.00	0.05	0.00	0.47	0.03
8	0.33	20	0.47	0.59	1.18	0.74	0.25	1.50	0.58	0.12	0.71	0.27
13	0.50	20	0.48	0.72	1.44	1.01	0.50	2.01	1.12	0.24	0.96	0.53
18	0.67	20	0.47	0.94	1.89	1.49	0.99	2.99	2.14	0.47	1.41	1.01
23	0.82	20	0.47	1.53	3.05	2.72	2.26	5.53	4.82	1.06	2.59	2.25
28	0.90	20	0.48	2.67	5.35	5.13	4.63	10.26	9.80	2.20	4.87	4.65
33	0.99	20	0.47	27.23	54.46	56.83	56.74	114.47	119.50	26.76	53.99	56.36
4	0.00	100	0.49	0.50	0.99	0.50	0.00	1.00	0.01	0.00	0.50	0.01
9	0.33	100	0.50	0.62	1.24	0.75	0.25	1.50	0.51	0.12	0.74	0.25
14	0.50	100	0.50	0.74	1.49	1.00	0.50	2.00	1.02	0.25	0.99	0.51
19	0.67	100	0.49	0.99	1.98	1.50	1.00	2.99	2.02	0.49	1.48	1.00
24	0.82	100	0.50	1.61	3.21	2.75	2.24	5.49	4.54	1.11	2.72	2.25
29	0.90	100	0.50	2.79	5.57	5.13	4.61	10.23	9.33	2.29	5.08	4.63
34	0.99	100	0.49	28.36	56.71	56.79	56.47	113.94	114.09	27.86	56.22	56.29
5	0.00	500	0.50	0.50	1.00	0.50	0.00	1.00	0.00	0.00	0.50	0.00
10	0.33	500	0.50	0.62	1.25	0.75	0.25	1.50	0.50	0.13	0.75	0.25
15	0.50	500	0.50	0.75	1.50	1.00	0.50	2.01	1.01	0.25	1.00	0.50
20	0.67	500	0.50	1.00	2.00	1.50	1.00	3.00	2.01	0.50	1.50	1.00
25	0.82	500	0.50	1.62	3.24	2.75	2.25	5.50	4.51	1.12	2.74	2.25
30	0.90	500	0.50	2.80	5.60	5.11	4.62	10.24	9.26	2.30	5.10	4.61
35	0.99	500	0.50	28.57	57.14	56.76	56.15	113.29	112.52	28.07	56.64	56.26

Figure A1: Expected value of estimators of sampling variance of $\hat{\mu}_{HT}$ under "lonely" primary sampling units in stratified sampling for different values of ρ_{ICC} and strata sizes (M).

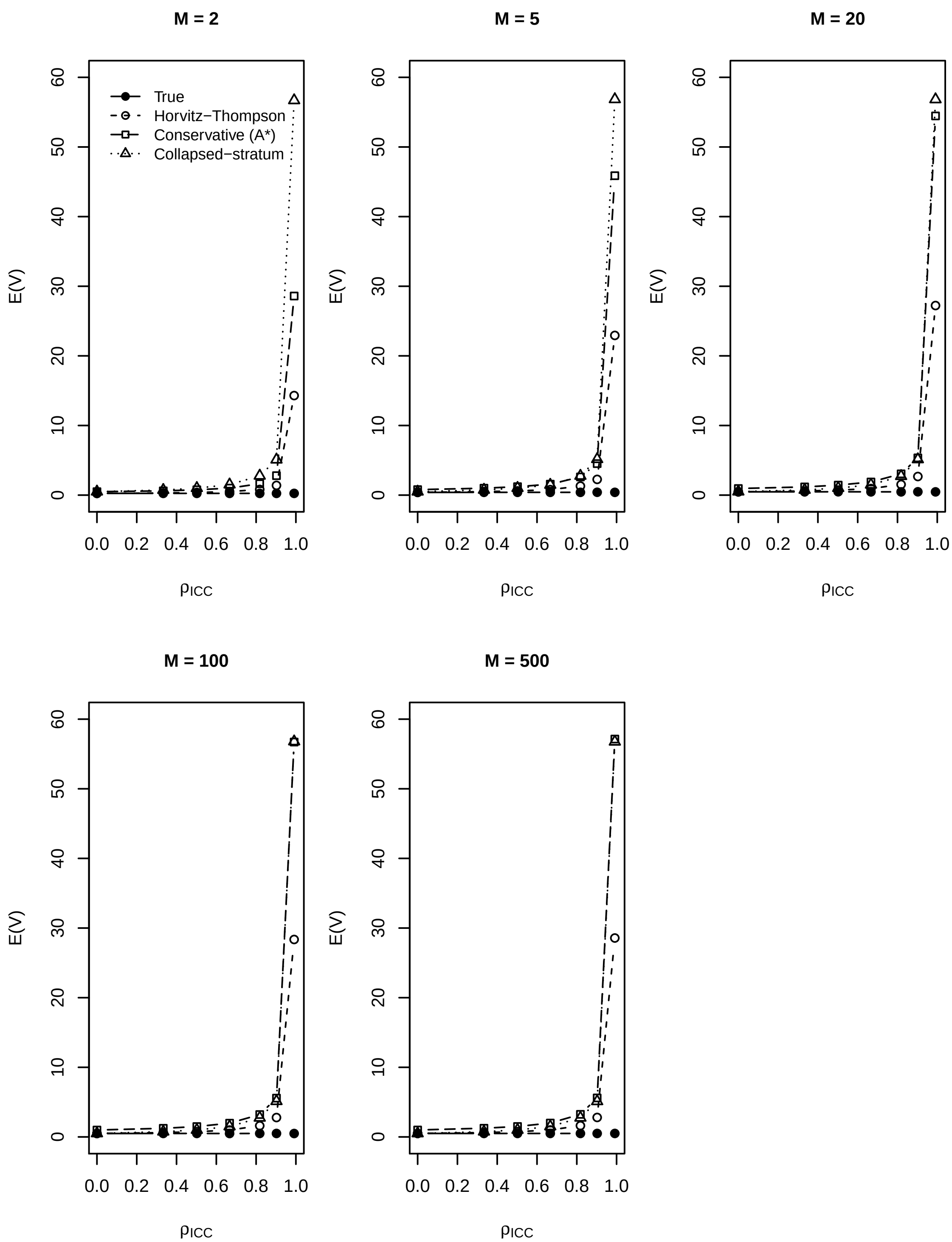

Table A2: Simulation results for PPS systematic sampling: random size case.

	Simulation Parameters				True	Mean from sims.		Relative Bias		Nominal Bias	
	N	m	ρ	$\overline{y}$	Var	$\widehat{\text{Var}}_{HT}$	$\widehat{\text{Var}}_{WR}$	$\widehat{\text{Var}}_{HT}$	$\widehat{\text{Var}}_{WR}$	$\widehat{\text{Var}}_{HT}$	$\widehat{\text{Var}}_{WR}$
1	8	2	0.05	3.29	529.96	544.00	650.20	0.03	0.23	14.04	120.24
2	8	2	0.10	5.87	506.26	529.46	610.87	0.05	0.21	23.20	104.61
3	8	2	0.20	11.59	532.05	614.33	640.17	0.15	0.20	82.28	108.12
4	8	2	0.30	17.82	484.35	671.21	581.15	0.39	0.20	186.86	96.80
5	8	2	0.40	23.72	459.18	811.04	539.63	0.77	0.18	351.86	80.45
6	8	2	0.50	29.55	407.41	941.85	489.26	1.31	0.20	534.44	81.85
7	8	2	0.60	36.19	328.73	1147.58	386.02	2.49	0.17	818.85	57.29
8	8	2	0.80	47.88	189.28	1581.40	227.39	7.35	0.20	1392.12	38.11
9	8	2	0.95	55.97	53.28	1983.83	62.73	36.23	0.18	1930.55	9.45
10	8	2	0.99	60.01	10.71	2236.50	13.31	207.82	0.24	2225.79	2.60
11	40	2	0.05	2.91	639.61	648.53	670.23	0.01	0.05	8.92	30.62
12	40	2	0.10	5.32	639.85	665.15	662.25	0.04	0.04	25.30	22.40
13	40	2	0.20	10.75	615.12	720.87	638.69	0.17	0.04	105.75	23.57
14	40	2	0.30	15.98	577.96	811.99	606.96	0.40	0.05	234.03	29.00
15	40	2	0.40	21.58	538.98	967.71	559.55	0.80	0.04	428.73	20.57
16	40	2	0.50	26.84	476.99	1138.38	495.28	1.39	0.04	661.39	18.29
17	40	2	0.60	32.11	410.38	1356.77	428.89	2.31	0.05	946.39	18.51
18	40	2	0.80	43.00	235.35	1936.15	245.59	7.23	0.04	1700.80	10.24
19	40	2	0.95	51.21	63.69	2473.52	64.90	37.84	0.02	2409.83	1.21
20	40	2	0.99	53.77	12.86	2667.59	13.45	206.43	0.05	2654.73	0.59
21	120	30	0.05	2.81	35.85	39.38	44.17	0.10	0.23	3.53	8.32
22	120	30	0.10	5.15	34.86	44.36	44.08	0.27	0.26	9.50	9.22
23	120	30	0.20	10.65	35.65	77.22	43.42	1.17	0.22	41.57	7.77
24	120	30	0.30	15.94	33.88	128.49	40.70	2.79	0.20	94.61	6.82
25	120	30	0.40	21.44	30.45	199.61	36.76	5.56	0.21	169.16	6.31
26	120	30	0.50	26.51	28.58	285.10	32.94	8.98	0.15	256.52	4.36
27	120	30	0.60	31.83	23.05	394.03	28.47	16.09	0.24	370.98	5.42
28	120	30	0.80	42.53	13.78	678.24	16.47	48.22	0.20	664.46	2.69
29	120	30	0.95	50.04	3.44	920.67	4.35	266.64	0.26	917.23	0.91
30	120	30	0.99	52.26	0.75	1000.50	0.90	1333.00	0.20	999.75	0.15
31	600	30	0.05	2.67	42.20	48.32	44.56	0.15	0.06	6.12	2.36
32	600	30	0.10	5.18	43.10	66.48	44.23	0.54	0.03	23.38	1.13
33	600	30	0.20	10.42	41.99	136.45	42.38	2.25	0.01	94.46	0.39
34	600	30	0.30	15.60	39.41	251.13	41.09	5.37	0.04	211.72	1.68
35	600	30	0.40	21.03	36.99	422.26	37.91	10.42	0.02	385.27	0.92
36	600	30	0.50	26.11	32.60	626.57	33.30	18.22	0.02	593.97	0.70
37	600	30	0.60	31.65	26.96	899.80	28.63	32.38	0.06	872.84	1.67
38	600	30	0.80	42.15	15.49	1563.71	16.15	99.95	0.04	1548.22	0.66
39	600	30	0.95	49.89	4.26	2173.08	4.32	509.11	0.01	2168.82	0.06
40	600	30	0.99	52.29	0.87	2383.53	0.90	2738.69	0.03	2382.66	0.03
41	240	60	0.05	2.73	18.19	21.01	22.57	0.16	0.24	2.82	4.38
42	240	60	0.10	5.14	18.15	27.77	22.27	0.53	0.23	9.62	4.12
43	240	60	0.20	10.56	18.08	58.36	21.47	2.23	0.19	40.28	3.39
44	240	60	0.30	15.97	16.86	108.81	20.42	5.45	0.21	91.95	3.56
45	240	60	0.40	21.19	15.55	176.91	18.50	10.38	0.19	161.36	2.95
46	240	60	0.50	26.44	13.51	264.64	16.77	18.59	0.24	251.13	3.26
47	240	60	0.60	31.82	11.77	373.63	14.46	30.74	0.23	361.86	2.69
48	240	60	0.80	42.18	6.66	646.21	8.01	96.03	0.20	639.55	1.35
49	240	60	0.95	50.23	1.82	909.42	2.18	498.68	0.20	907.60	0.36
50	240	60	0.99	52.24	0.37	980.07	0.44	2647.84	0.19	979.70	0.07
51	1200	60	0.05	2.62	21.61	27.51	22.24	0.27	0.03	5.90	0.63
52	1200	60	0.10	5.27	21.31	45.42	22.10	1.13	0.04	24.11	0.79
53	1200	60	0.20	10.45	20.88	115.68	21.32	4.54	0.02	94.80	0.44
54	1200	60	0.30	15.78	19.42	236.02	20.27	11.15	0.04	216.60	0.85
55	1200	60	0.40	21.15	18.15	407.25	18.66	21.44	0.03	389.10	0.51
56	1200	60	0.50	26.28	16.18	616.64	16.75	37.11	0.04	600.46	0.57
57	1200	60	0.60	31.43	13.81	872.70	14.32	62.19	0.04	858.89	0.51
58	1200	60	0.80	42.05	7.71	1545.25	8.07	199.42	0.05	1537.54	0.36
59	1200	60	0.95	49.95	2.08	2171.46	2.18	1042.97	0.05	2169.38	0.10
60	1200	60	0.99	51.99	0.44	2350.79	0.45	5341.70	0.02	2350.35	0.01

Table A3: Simulation results for PPS systematic sampling: periodic size case.

	Simulation Parameters				True	Mean from sims.			Relative Bias			Nominal Bias		
	N	m	ρ	$\overline{y}$	Var	$\widehat{\text{Var}}_{HT}$	$\widehat{\text{Var}}_{C2}$	$\widehat{\text{Var}}_{WR}$	$\widehat{\text{Var}}_{HT}$	$\widehat{\text{Var}}_{C2}$	$\widehat{\text{Var}}_{WR}$	$\widehat{\text{Var}}_{HT}$	$\widehat{\text{Var}}_{C2}$	$\widehat{\text{Var}}_{WR}$
1	8	2	0.05	-0.00	608.05	527.29	1101.16	819.45	-0.13	0.81	0.35	-80.76	493.11	211.40
2	8	2	0.10	-0.00	644.24	557.15	1146.35	806.80	-0.14	0.78	0.25	-87.09	502.11	162.56
3	8	2	0.20	0.00	625.78	542.86	1099.09	765.79	-0.13	0.76	0.22	-82.92	473.31	140.01
4	8	2	0.30	0.00	661.57	573.37	1123.95	727.74	-0.13	0.70	0.10	-88.20	462.38	66.17
5	8	2	0.40	0.00	726.70	631.77	1199.10	693.81	-0.13	0.65	-0.05	-94.93	472.40	-32.89
6	8	2	0.50	0.00	816.15	713.11	1282.44	611.01	-0.13	0.57	-0.25	-103.04	466.29	-205.14
7	8	2	0.60	-0.00	890.58	776.59	1344.60	523.27	-0.13	0.51	-0.41	-113.99	454.02	-367.31
8	8	2	0.80	-0.00	1145.46	1003.38	1578.73	291.78	-0.12	0.38	-0.75	-142.08	433.27	-853.68
9	8	2	0.95	0.00	1336.81	1173.42	1739.33	80.90	-0.12	0.30	-0.94	-163.39	402.52	-1255.91
10	8	2	0.99	0.00	1397.05	1227.24	1791.48	16.02	-0.12	0.28	-0.99	-169.81	394.43	-1381.03
11	40	2	0.05	0.00	764.97	743.70	1574.64	800.60	-0.03	1.06	0.05	-21.27	809.67	35.63
12	40	2	0.10	-0.00	777.42	755.92	1586.43	803.44	-0.03	1.04	0.03	-21.50	809.01	26.02
13	40	2	0.20	0.00	798.12	775.99	1614.04	799.17	-0.03	1.02	0.00	-22.13	815.92	1.05
14	40	2	0.30	-0.00	852.85	829.37	1667.29	733.92	-0.03	0.95	-0.14	-23.48	814.44	-118.93
15	40	2	0.40	0.00	904.72	879.71	1716.09	670.94	-0.03	0.90	-0.26	-25.01	811.37	-233.78
16	40	2	0.50	0.00	975.72	948.93	1782.24	600.88	-0.03	0.83	-0.38	-26.79	806.52	-374.84
17	40	2	0.60	-0.00	1062.00	1032.97	1865.95	515.88	-0.03	0.76	-0.51	-29.03	803.95	-546.12
18	40	2	0.80	-0.00	1295.01	1259.81	2094.14	291.75	-0.03	0.62	-0.77	-35.20	799.13	-1003.26
19	40	2	0.95	0.00	1503.69	1462.95	2297.02	77.76	-0.03	0.53	-0.95	-40.74	793.33	-1425.93
20	40	2	0.99	0.00	1559.54	1517.30	2347.71	16.14	-0.03	0.51	-0.99	-42.24	788.17	-1543.40
21	120	30	0.05	-0.00	42.55	36.78	653.92	53.70	-0.14	14.37	0.26	-5.77	611.37	11.15
22	120	30	0.10	0.00	58.69	51.57	669.42	53.35	-0.12	10.41	-0.09	-7.12	610.73	-5.34
23	120	30	0.20	-0.00	100.87	89.35	701.66	50.59	-0.11	5.96	-0.50	-11.52	600.79	-50.28
24	120	30	0.30	0.00	181.49	161.43	780.15	49.03	-0.11	3.30	-0.73	-20.06	598.66	-132.46
25	120	30	0.40	-0.00	296.35	264.05	881.44	45.27	-0.11	1.97	-0.85	-32.30	585.09	-251.08
26	120	30	0.50	0.00	431.48	385.26	995.71	40.07	-0.11	1.31	-0.91	-46.22	564.23	-391.41
27	120	30	0.60	0.00	606.48	541.65	1160.33	34.87	-0.11	0.91	-0.94	-64.83	553.85	-571.61
28	120	30	0.80	-0.00	1044.40	933.46	1549.04	19.51	-0.11	0.48	-0.98	-110.94	504.64	-1024.89
29	120	30	0.95	0.00	1450.34	1297.21	1910.48	5.12	-0.11	0.32	-1.00	-153.13	460.14	-1445.22
30	120	30	0.99	0.00	1578.87	1412.15	2029.28	1.07	-0.11	0.29	-1.00	-166.72	450.41	-1577.80
31	600	30	0.05	0.00	56.35	54.90	891.43	54.67	-0.03	14.82	-0.03	-1.45	835.08	-1.68
32	600	30	0.10	-0.00	63.91	62.36	893.22	52.77	-0.02	12.98	-0.17	-1.55	829.31	-11.14
33	600	30	0.20	-0.00	114.48	112.18	949.68	52.71	-0.02	7.30	-0.54	-2.30	835.20	-61.77
34	600	30	0.30	0.00	196.46	193.07	1025.17	48.50	-0.02	4.22	-0.75	-3.39	828.71	-147.96
35	600	30	0.40	-0.00	302.06	297.22	1131.46	46.29	-0.02	2.75	-0.85	-4.84	829.40	-255.77
36	600	30	0.50	-0.00	441.46	434.71	1262.96	40.83	-0.02	1.86	-0.91	-6.75	821.50	-400.63
37	600	30	0.60	0.00	614.43	605.26	1436.63	34.54	-0.01	1.34	-0.94	-9.17	822.20	-579.89
38	600	30	0.80	-0.00	1048.24	1033.01	1861.12	19.59	-0.01	0.78	-0.98	-15.23	812.88	-1028.65
39	600	30	0.95	-0.00	1464.60	1443.65	2275.35	5.36	-0.01	0.55	-1.00	-20.95	810.75	-1459.24
40	600	30	0.99	-0.00	1587.01	1564.34	2396.58	1.08	-0.01	0.51	-1.00	-22.67	809.57	-1585.93
41	240	60	0.05	-0.00	23.14	20.39	622.28	27.02	-0.12	25.89	0.17	-2.75	599.14	3.88
42	240	60	0.10	0.00	34.57	30.43	631.34	27.15	-0.12	17.26	-0.21	-4.14	596.77	-7.42
43	240	60	0.20	0.00	80.14	71.94	671.15	25.81	-0.10	7.37	-0.68	-8.20	591.01	-54.33
44	240	60	0.30	-0.00	166.64	150.20	758.02	24.83	-0.10	3.55	-0.85	-16.44	591.38	-141.81
45	240	60	0.40	0.00	268.72	242.40	839.63	22.60	-0.10	2.12	-0.92	-26.32	570.91	-246.12
46	240	60	0.50	0.00	410.60	370.89	971.00	20.42	-0.10	1.36	-0.95	-39.71	560.40	-390.18
47	240	60	0.60	0.00	587.24	530.72	1132.75	17.26	-0.10	0.93	-0.97	-56.52	545.51	-569.98
48	240	60	0.80	0.00	1046.57	947.17	1551.01	9.70	-0.09	0.48	-0.99	-99.40	504.44	-1036.87
49	240	60	0.95	-0.00	1459.15	1320.20	1922.45	2.62	-0.10	0.32	-1.00	-138.95	463.30	-1456.53
50	240	60	0.99	0.00	1581.13	1430.59	2032.07	0.52	-0.10	0.29	-1.00	-150.54	450.94	-1580.61
51	1200	60	0.05	0.00	28.58	27.82	861.11	26.77	-0.03	29.13	-0.06	-0.76	832.53	-1.81
52	1200	60	0.10	0.00	41.36	40.42	874.69	26.11	-0.02	20.15	-0.37	-0.94	833.33	-15.25
53	1200	60	0.20	0.00	88.89	87.30	918.99	25.90	-0.02	9.34	-0.71	-1.59	830.10	-62.99
54	1200	60	0.30	-0.00	168.73	165.94	997.42	24.28	-0.02	4.91	-0.86	-2.79	828.69	-144.45
55	1200	60	0.40	-0.00	278.60	274.09	1106.83	22.81	-0.02	2.97	-0.92	-4.51	828.23	-255.79
56	1200	60	0.50	-0.00	424.05	417.44	1253.59	20.61	-0.02	1.96	-0.95	-6.61	829.54	-403.44
57	1200	60	0.60	0.00	599.83	590.55	1424.26	17.67	-0.02	1.37	-0.97	-9.28	824.43	-582.16
58	1200	60	0.80	-0.00	1047.24	1031.27	1866.45	9.89	-0.02	0.78	-0.99	-15.97	819.21	-1037.35
59	1200	60	0.95	0.00	1463.81	1441.68	2275.64	2.62	-0.02	0.55	-1.00	-22.13	811.83	-1461.19
60	1200	60	0.99	-0.00	1586.32	1562.38	2395.86	0.53	-0.02	0.51	-1.00	-23.94	809.54	-1585.79

Figure A2: Expected value of estimators of sampling variance of $\hat{\mu}_{HT}$ under PPS systematic sampling for different values of ρ, sample sizes (m), and sampling proportions (m/N): random size case.

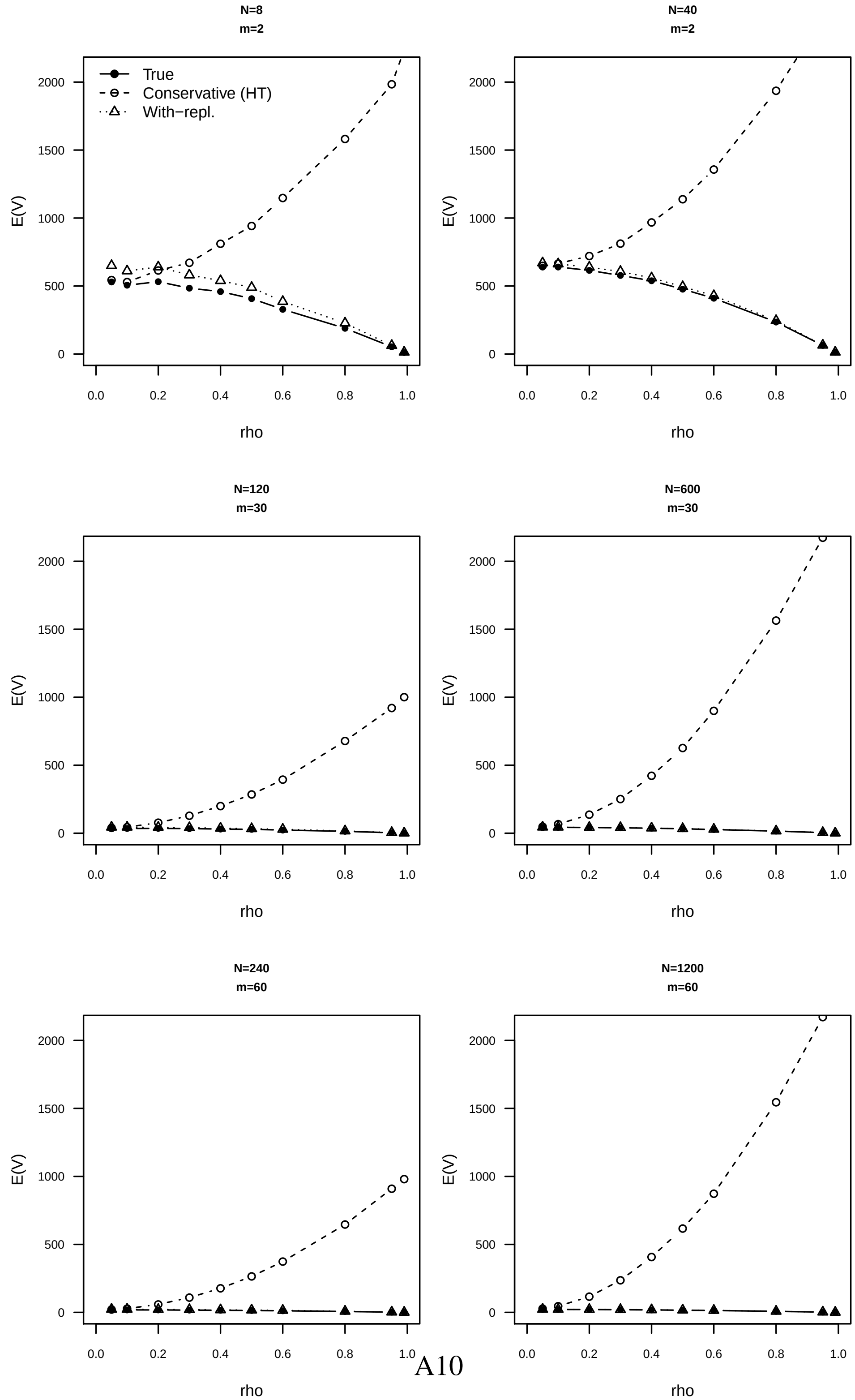

Figure A3: Expected value of estimators of sampling variance of $\hat{\mu}_{HT}$ under PPS systematic sampling, for different values of ρ, sample sizes (m), and sampling proportions (m/N): periodic size case.