



Mathematical Statistics
Stockholm University

**A simulation study of the performance
of estimators based on estimated
inclusion probabilities**

Lars-Erik Almberg

Examensarbete 2012:1

Postal address:

Mathematical Statistics
Dept. of Mathematics
Stockholm University
SE-106 91 Stockholm
Sweden

Internet:

<http://www.math.su.se/matstat>



Mathematical Statistics
Stockholm University
Examensarbete 2012:1,
<http://www.math.su.se/matstat>

A simulation study of the performance of estimators based on estimated inclusion probabilities

Lars-Erik Almqvist*

Mars 2012

Abstract

With the knowledge of the inclusion probabilities of the elements of a population, with a specific design, we can estimate different parameters of interest. If we know the exact inclusion probabilities, the estimators will be unbiased. In some designs we do not know the inclusion probabilities but we can, with repeated sampling, estimate the inclusion probabilities. In this essay we will do a large simulation study where we investigate the performance of estimators based on estimated inclusion probabilities. We will focus on the point estimator of the total and the estimated variance of the point estimator of the total. The procedure is to first simulate simple random sampling, where all the inclusion probabilities are known, and then to simulate Pareto π ps, where we only have an approximate expression for the first order inclusion probabilities and do not know the second order inclusion probabilities.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden.
E-mail: larserialmqvist@hotmail.com. Supervisor: Gudrun Brattström.

Acknowledgements

I would like to thank my supervisor Gudrun Brattström at the University of Stockholm for her help and support with this essay. I also want to thank Anders Holmberg and Martin Axelson at SCB in Örebro for their help and inspiration.

Contents

1 Introduction	5
2 Theory	7
2.1 Inclusion probabilities	7
2.2 Estimators	7
2.3 Simple random sampling without replacement (SI)	8
2.4 Pareto π ps sampling (PPS)	9
2.5 Bootstrap from a finite population.....	10
3 Implementation	11
3.1 Illustrations	11
3.2 Fattorini´s research	12
3.3 Definitions	14
3.4 Simulations	18
3.4.1 Presentation.....	18
3.4.2 SI simulation, population A, $n=38$	21
3.4.3 PPS simulation, population A, $n=38$	25
3.4.4 PPS simulation, population B, $n=174$	31
3.5 Summary of all simulations	36
4 Conclusions	40
5 References	42
6 Appendix	44
6.1 Summary of the formulas	44
6.2 Figures and Tables referred to in the text	47
6.2.1 SI simulation, population A, $n=38$	47
6.2.2 PPS simulation, population A, $n=38$	50
6.2.3 PPS simulation, population B, $n=174$	53
6.3 SI simulation, population A, $n=100$	56
6.4 PPS simulation, population A, $n=20$	58
6.5 SI simulation, population B, $n=174$	62

6.6 PPS simulation, population B, $n=90$	65
6.7 PPS simulation, population B, $n=45$	68
6.8 PPS simulation, population B, $n=20$	72

1 Introduction

In this essay we will investigate the performance of different estimators based on *estimated inclusion probabilities*. Inclusion probabilities are essential in the concept of *probability sampling*. In this concept we have a finite population that we want to provide information about. We have a sampling procedure, including a random mechanism, which gives us a set of possible samples that can be drawn from the population. The sampling procedure also gives each possible sample the probability to be selected. We call the distribution of these probabilities the *sampling design* or just *design*.

Given the sampling design, we define the probabilities for each element in the finite population to be included in a sample, as the *first order inclusion probabilities*, and we define the probabilities for each combination of two elements to be included in the same sample, as the *second order inclusion probabilities*.

If we know the inclusion probabilities for the elements in the population, we can estimate different parameters of interest. In this essay we will focus on the point estimator of the total and the estimated variance of the point estimator of the total, but we will also include the variance of the point estimator of the total. For the point estimator of the total we only need the first order inclusion probabilities, but for the variances we also need the second order inclusion probabilities. In many designs we know the inclusion probabilities. One of the simplest examples is simple random sampling without replacement (SI), where the sampling is performed by randomly choosing one element at a time until the sample contains the number of elements specified by the sample size. But there are also more complicated designs where we do not know the inclusion probabilities. One such example is Pareto πps Sampling (PPS), where we do not have an exact expression for the inclusion probabilities.

A sampling frame is a material containing information about the elements in the population, in such a way that sampling is possible. The frame can also include auxiliary information needed for stratification and probability proportional-to-size sampling etc. If we have a frame listing every element in the population, we can with a specific design, estimate the inclusion probabilities by repeated sampling.

The procedure of *estimating the inclusion probabilities* is to draw a sample with the specific design, register the elements in the sample and then repeat the procedure many times. The precision of the estimated inclusion probabilities will increase with the number of repetitions. Some research has been done in this field and in this essay we will do a large simulation study where we investigate the performance of the estimator of the total, the variance of the estimated total and the estimated variance of the estimated total, all applied with estimated inclusion probabilities.

We will simulate SI, where all the inclusion probabilities are known, and PPS, where we only have an approximate expression for the first order inclusion probabilities and do not know the second order (there are however numeric methods to compute the inclusion probabilities in PPS). We have chosen PPS because it is of special interest, for further studies, to investigate the situation where we know the first but not the second order inclusion probabilities.

The results of our simulations show that we can get very precise results when applying the estimated inclusion probabilities in the estimators. We show that some of our measures are

inversely proportional to the square root of the number of repetitions we do when estimating the inclusion probabilities.

Section 2 of the essay includes basic theory and formulas. In section 3 we present Fattorini's research, the definitions of our measures and the implementation of the simulations, and in the end of the section we present a summary of all the simulated designs. Conclusions are discussed in section 4 and in the appendix, section 6, we present a summary of the formulas, the figures and tables referred to in the text, and give information about the simulated designs that were not presented in detail in section 3.

2 Theory

2.1 Inclusion probabilities

In *probability sampling* we sample with a random mechanism from a finite population $U = \{1, \dots, k, \dots, N\}$, where N is the size of the population. We define ζ to be the set of all possible samples that can be selected from U . That means ζ is the set of 2^N subsets from U , if we include the empty set, as well as U itself. With a specific sampling procedure, let $p(s)$ be the probability for each possible sample s to be drawn. We then define the distribution of all those probabilities as the *sampling design*, $p(\cdot)$. For a given sampling design $p(\cdot)$, we can regard each possible sample s as the outcome of a random variable S , such that for any $s \in \zeta$

$$P(S = s) = p(s)$$

Depending on the design, the N elements in the finite population can have different probabilities of inclusion in a sample. We define the probability that element k will be in a sample as

$$\pi_k = P(k \in S) = \sum_{s \ni k} p(s)$$

We call π_k the *first order inclusion probability*. In the same way the *second order inclusion probability*, the probability that both k and l will be in the same sample, is given by

$$\pi_{kl} = P(k \& l \in S) = \sum_{s \ni k \& l} p(s)$$

2.2 Estimators

The inclusion probabilities are essential when we want to estimate different parameters of the unknown variable of interest, $\mathbf{y} = (y_1, \dots, y_k, \dots, y_N)$. We define the population total as $\theta = \sum_U y_k$, and in this essay we will only work with the estimator of the population total, the variance of the estimated total and the estimated variance of the estimated total.

When all the first order inclusion probabilities are known, the Horvitz-Thompson estimator (Horvitz and Thompson 1952) of the population total, with a specific sample s , is defined as

$$\hat{\theta} = \sum_{k \in s} \frac{y_k}{\pi_k} \tag{2.1}$$

The estimator is unbiased if $\pi_k > 0$ for all $k \in U$.

The variance of $\hat{\theta}$ is computed for the whole population U , and an estimator of the variance is computed for a specific sample s . For sampling with fixed sample size and if both all the first and all the second order inclusion probabilities are known, the variance of $\hat{\theta}$ for *the whole population* is given by (Yates and Grundy 1953; Sen 1953)

$$V(\hat{\theta}) = -\frac{1}{2} \sum_{k \in U} \sum_{l \in U} \left(\frac{y_k}{\pi_k} - \frac{y_l}{\pi_l} \right)^2 (\pi_{kl} - \pi_k \pi_l) \quad (2.2)$$

If $\pi_{kl} > 0$ for all $k \neq l \in U$, an unbiased *estimator* of the variance of $\hat{\theta}$, with a specific sample s , is

$$\hat{V}(\hat{\theta}) = -\frac{1}{2} \sum_{k \in s} \sum_{l \in s} \left(\frac{y_k}{\pi_k} - \frac{y_l}{\pi_l} \right)^2 \frac{(\pi_{kl} - \pi_k \pi_l)}{\pi_{kl}} \quad (2.3)$$

2.3 Simple random sampling without replacement (SI)

In simple random sampling without replacement (SI), both the first and second order inclusion probabilities are known. The first order inclusion probabilities can be computed by $\pi_k = \frac{n}{N}$ and the second order inclusion probabilities can be computed by $\pi_{kl} = \frac{n(n-1)}{N(N-1)}$. This gives us an exact expression for the variance of $\hat{\theta}$ for *the whole population*

$$V(\hat{\theta}) = N^2 \frac{1-f}{n} S_{yU}^2 \quad (2.4)$$

where $S_{yU}^2 = \frac{1}{(N-1)} \sum_U (y_k - \bar{y}_U)^2$ and $f = \frac{n}{N}$.

and an unbiased *estimator* of the variance of $\hat{\theta}$, with a specific sample s

$$\hat{V}(\hat{\theta}) = N^2 \frac{1-f}{n} S_{y_s}^2, \quad (2.5)$$

where $S_{y_s}^2 = \frac{1}{(n-1)} \sum_s (y_k - \bar{y}_s)^2$.

2.4 Pareto πps sampling (PPS)

When working with sampling with unequal inclusion probabilities, we have an auxiliary variable $\mathbf{x} = (x_1, \dots, x_k, \dots, x_N)$ which is known. The auxiliary variable is related to \mathbf{y} and will be used to increase the efficiency of our estimator. The *target inclusion probabilities* are then given by

$$\lambda_k = \frac{nx_k}{\sum_{l \in U} x_l},$$

where n is the sample size. We will only work with sample sizes where all $\lambda_k < 1$.

Order sampling is a sampling procedure where every unit in the population is assigned a *ranking variable*, Q_k . The sample then consists of the n units with the smallest ranking variables. Ohlsson (1990, 1995) introduced order sampling with uniform ordering distributions, called Sequential Poisson Sampling. Pareto πps sampling (PPS) is an order sampling scheme that was introduced independently by Saavedra (1995) and Rosén (1997b). The PPS scheme is as follows:

1. With the use of an auxiliary variable, compute λ_k for every element in the population.
2. Compute each ranking variable

$$Q_k = \frac{R_k(1 - \lambda_k)}{\lambda_k(1 - R_k)},$$

where R_k is a standard uniform random number, $R_k(0,1)$.

3. Select the n smallest units as the sample.

In PPS we do not have an exact expression for the inclusion probabilities. There are however numeric methods to compute π_k and π_{kl} , see Aires (1999), Matei and Tillé (2007) and Ng and Donadio (2006). Rosén (2000) proved, with a prescribed sample size $n \leq N$ and with all λ_k : $0 < \lambda_k < 1$, that $\pi_k/\lambda_k \rightarrow 1$ as $n \rightarrow \infty$. Rosén and Aires (2005) showed with a simulation study that π_k converge quite rapidly to λ_k with increasing population size N , and if the population is large enough we can use λ_k as an approximation of π_k .

An approximate expression of the Horvitz-Thompson estimator of the total is then defined with the target inclusion probabilities λ_k , as

$$\hat{\theta}_\lambda = \sum_{k \in s} \frac{y_k}{\lambda_k} \tag{2.6}$$

Rosén (1997b) defined an approximate variance of $\hat{\theta}_\lambda$ for *the whole population*

$$V_R(\hat{\theta}_\lambda) = \frac{N}{N-1} \left\{ \sum_{k \in U} \frac{y_k^2}{\lambda_k} (1 - \lambda_k) - \frac{[\sum_{k \in U} y_k (1 - \lambda_k)]^2}{n - \sum_{k \in U} \lambda_k^2} \right\} \quad (2.7)$$

and an approximate *estimator* of the variance of $\hat{\theta}_\lambda$, with a specific sample s

$$\hat{V}_R(\hat{\theta}_\lambda) = \frac{n}{n-1} \sum_{k \in s} \left(\frac{y_k}{\lambda_k} - \frac{\sum_{l \in s} y_l (1 - \lambda_l) / \lambda_l}{\sum_{l \in s} 1 - \lambda_l} \right)^2 (1 - \lambda_k) \quad (2.8)$$

2.5 Bootstrap from a finite population

When we sample without replacement from a finite population, we cannot use the classical bootstrap method, which is based on an identical and independent distribution. Different methods have been proposed that compensate for this. Some methods include rescaling or weighting of the sample, others reconstruct artificial populations from the sample.

Antal and Tillé (2011) have recently proposed a new method that can be applied to sampling designs both with and without replacement. They propose a resampling design that mixes several designs and produces a bootstrap variance that is equal to the estimator of the variance. One part of the resampling design is the one-one resampling design, which produces a sample, or a part of a sample, where the number of times every unit is selected has expectation one and variance one. The one-one resampling design is a mixture of simple random sampling with replacement and simple random sampling with overreplacement. Simple random sampling with replacement can be viewed as a distribution of independent Poisson variables conditioned on their sum. If we instead use a sequence of geometric variables conditioned on their size, we get a sampling design with replacement and with fixed sample size. Antal and Tillé call this design simple random sampling with overreplacement because the repetition of units are more frequent than in simple random sampling with replacement. Simple random sampling with overreplacement also gives a much larger variance than simple random sampling with replacement, a fact that they use in the one-one resampling design.

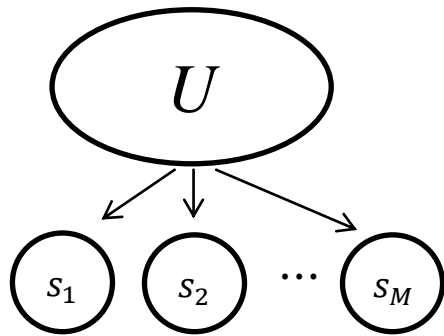
When resampling from a sample selected with unequal probabilities without replacement, the bootstrap algorithm consists of a mixture of the original design and the one-one design.

We will use this method in PPS to estimate the variance of $\hat{\theta}_\lambda$ (2.6), and we will denote this estimator $\hat{V}_{AT}^*(\hat{\theta}_\lambda)$. In this essay the symbol “*” will be used with estimators involving some kind of simulation process.

3 Implementation

3.1 Illustrations

A sampling frame is a material containing information about the elements in the population, in such a way that sampling is possible. The frame can also include auxiliary information needed for stratification and probability proportional-to-size sampling etc. If we have a frame listing all elements in the population, we can estimate the inclusion probabilities.



To estimate the inclusion probabilities of a specific sampling design, we simulate a sample s_1 with that design from the population U . We register the objects in the sample and then repeat the procedure many times. The estimated inclusion probabilities are stored in a matrix and the precision of the estimated inclusion probabilities will increase with the number of repetitions, M . We show the procedure with a simple example.

Say we have a population of 5 people. The sampling design is to draw two persons of five with SI. In the first sample person one and four are drawn and the matrix looks like this

```

1
0 0
0 0 0
1 0 0 1
0 0 0 0 0
  
```

The diagonal contains the estimated first order inclusion probabilities and the rest of the matrix contains the estimated second order inclusion probabilities, the probabilities that the person represented by the row and the person represented by the column will be drawn in the same sample. In the second sample person two and four are drawn and then the matrix looks like this

```

1/2
0 1/2
0 0 0
1/2 1/2 0 1
0 0 0 0 0
  
```

At this stage the matrix includes the information that the estimated inclusion probability for person four, is one with this sampling design. For person three and five, the estimated probability of being drawn is zero and for person one and two, the estimated probability of being drawn is $\frac{1}{2}$. And then we have all the estimated second order inclusion probabilities, the probabilities that both the persons, represented by the row and the column, will be drawn in the same sample.

If we repeat this procedure many times all the estimated first order inclusion probabilities, the diagonal, will converge to their true value: $\frac{n}{N} = \frac{2}{5} = 0.4$ and all the estimated second order inclusion probabilities will converge to: $\frac{n(n-1)}{N(N-1)} = \frac{2 \cdot 1}{5 \cdot 4} = 0.1$. The question is how many repetitions we need and if we can get estimated inclusion probabilities that are precise enough.

3.2 Fattorini's research

Fattorini (2006) has done some research on the subject of estimating inclusion probabilities. To guarantee a positive estimator of π_k in alignment with Fattorini, we define the estimated first order inclusion probabilities $\tilde{\pi}_k$ as

$$\tilde{\pi}_k = \frac{A_k + 1}{M + 1}$$

where A_k is the number of times unit k has been drawn and M is the number of repetitions. The estimated second order inclusion probabilities are defined as

$$\tilde{\pi}_{kl} = \frac{A_{kl} + 1}{M + 1}$$

where A_{kl} is the number of times unit k and l have been drawn together. We can think of it as if we start with a inclusion matrix of ones. This means that we are able to compute the variance even if all combination of units have not yet occurred.

The Horvitz-Thompson estimator of the total with the estimated inclusion probabilities is then

$$\hat{\theta}^* = \sum_{k \in S} \frac{y_k}{\tilde{\pi}_k} \quad (3.1)$$

Fattorini showed that

$$|E(\hat{\theta}^*) - \theta| \leq \frac{\theta}{(M + 2)\pi_0}$$

where π_0 is the π_k with the lowest value, and

$$|MSE(\hat{\theta}^*) - V(\hat{\theta})| \leq \frac{9E(\hat{\theta}^2)}{(M + 2)\pi_0}$$

We denote the variance for *the whole population* with $\tilde{\pi}_k$ and $\tilde{\pi}_{kl}$ as

$$V^*(\hat{\theta}^*) = -\frac{1}{2} \sum_{k \in U} \sum_{l \in U} \left(\frac{y_k}{\tilde{\pi}_k} - \frac{y_l}{\tilde{\pi}_l} \right)^2 (\tilde{\pi}_{kl} - \tilde{\pi}_k \tilde{\pi}_l) \quad (3.2)$$

and its *estimator* for a specific sample s

$$\hat{V}^*(\hat{\theta}^*) = -\frac{1}{2} \sum_{k \in S} \sum_{l \in S} \left(\frac{y_k}{\tilde{\pi}_k} - \frac{y_l}{\tilde{\pi}_l} \right)^2 \frac{\tilde{\pi}_{kl} - \tilde{\pi}_k \tilde{\pi}_l}{\tilde{\pi}_{kl}} \quad (3.3)$$

Fattorini showed that

$$\left| E \left(\hat{V}^*(\hat{\theta}^*) \right) - V(\hat{\theta}^*) \right| \leq \frac{\theta^2}{(M+2)\pi_{00}}$$

where $V(\hat{\theta}^*)$ is the variance of $\hat{\theta}^*$ with the expected estimated inclusion probabilities, and π_{00} is the π_{kl} with the lowest value. Fattorini points out that the bounding expressions indicate that a large number of repetitions, some billion, is needed if we want to guarantee a good precision for $\hat{\theta}^*$ and $\hat{V}^*(\hat{\theta}^*)$. But he then shows with some familiar designs that, in practice, it is enough with one million repetitions to get an acceptable precision.

In an article (2009) Fattorini gives an algorithm to be used when estimating inclusion probabilities, that ensures stability for both $\hat{\theta}^*$ and $\hat{V}^*(\hat{\theta}^*)$ and a required level of accuracy for $\hat{\theta}^*$. The algorithm is based on the Bennet (1962) inequality and the only thing needed is a sample from the current design. You are also able to choose between two different variance estimators and two different precision criteria. With the milder precision criteria and conditioned on the sample s , the algorithm bounds the probability that the absolute relative difference between $\hat{\theta}^*$ (3.1) and $\hat{\theta}$ (2.1), is greater than a desired precision $\epsilon > 0$, as

$$Pr \left\{ \left| \frac{\hat{\theta}^* - \hat{\theta}}{\hat{\theta}} \right| > \epsilon \right\} \leq \gamma$$

Define $c(\epsilon) = \left(\frac{\epsilon}{2+2\epsilon} + 1 \right) \ln \left(\frac{\epsilon}{2+2\epsilon} + 1 \right) - \frac{\epsilon}{2+2\epsilon}$. The value of γ is then computed as $\gamma = 2 \sum_{k \in S} e^{-c(\epsilon)M\tilde{\pi}_k}$. Denote the smallest estimated first order inclusion probability of the sample as $\tilde{\pi}_0$. The minimum of repetitions needed, to ensure the accuracy of γ , is then defined as $M_0 = (2 + 2\epsilon)/(\epsilon \tilde{\pi}_0)$. If we choose the harder precision criteria, see Fattorini's article, more repetitions are needed to bound the probability.

A *step* consists of L repetitions in the process of estimating the inclusion probabilities. After each step i , the estimated total $\hat{\theta}_i^*$ and the estimated variance $\hat{V}_i^*(\hat{\theta}^*)$, both based on the estimated inclusion probabilities, are computed. To guarantee stability of the estimators, the algorithm computes the absolute relative difference between the estimates in consecutive steps. With the precision level $\delta > 0$, $g(i)$ is defined as

$$g(i) = \max \left\{ j: \bigcap_{l=i-j}^i \left(\left| \frac{\hat{\theta}_l^* - \hat{\theta}_{l-j}^*}{\hat{\theta}_{l-j}^*} \right| < \delta, \left| \frac{\hat{V}_l^*(\hat{\theta}^*) - \hat{V}_{l-j}^*(\hat{\theta}^*)}{\hat{V}_{l-j}^*(\hat{\theta}^*)} \right| < \delta \right) \right\}$$

The notation in the formula with “ \cap ” is copied from Fattorini's article. In words, $g(i)$ is the maximum value of j , where every estimate from j steps back up to this step i , has an absolute relative difference with the estimate j steps back, that is less than δ . Let γ_i be the value of γ after step i and define β as the maximum acceptable value for γ and G as the minimum acceptable value for $g(i)$. Then each *stage* is completed when $\gamma_i \leq \beta$, $g(i) \geq G$ and $L \cdot i = M > M_0$. After each stage is completed, new decreased values for ϵ and/or β are entered manually.

We will do this algorithm with the first three designs in our work and we will use the Sen-Yates-Grundy variance estimator with the estimated inclusion probabilities (3.3).

3.3 Definitions

We will do a large simulation study where we investigate the performance of $V^*(\hat{\theta}^*)$, $\hat{V}^*(\hat{\theta}^*)$ and $\hat{\theta}^*$. In the study we will simulate both SI and PPS. We have chosen PPS because it is of special interest, for further studies, to investigate the situation where we know the first but not the second order inclusion probabilities.

The precision of the estimated inclusion probabilities will increase with the number of repetitions. A measure of precision of the estimated first order inclusion probabilities is

$$\varphi_1 = \begin{cases} \frac{1}{N} \sum_{k \in U} \left| \frac{\tilde{\pi}_k - \pi_k}{\pi_k} \right| & \text{in SI} \\ \frac{1}{N} \sum_{k \in U} \left| \frac{\tilde{\pi}_k - \lambda_k}{\lambda_k} \right| & \text{in PPS} \end{cases}$$

In SI it is the mean of the absolute relative differences between the estimated and the exact first order inclusion probabilities, and in PPS we use the target inclusion probabilities λ_k as a reference for the estimated inclusion probabilities. We will use φ_1 as a general measure of the precision of the estimated inclusion probabilities and $\varphi_1 \leq \alpha$ as the limit of the precision. The question is what value for α we should use.

We also compute the maximum absolute relative difference for the first order inclusion probabilities

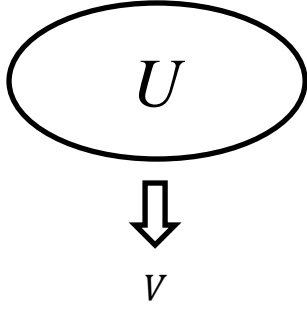
$$\psi_1 = \begin{cases} \max \left| \frac{\tilde{\pi}_k - \pi_k}{\pi_k} \right| & \text{in SI} \\ \max \left| \frac{\tilde{\pi}_k - \lambda_k}{\lambda_k} \right| & \text{in PPS} \end{cases}$$

With SI we can also compute the measures of precision for the second order inclusion probabilities

$$\varphi_2 = \frac{1}{\#\pi_{kl}} \sum_{k \neq l} \sum_U \left| \frac{\tilde{\pi}_{kl} - \pi_{kl}}{\pi_{kl}} \right|$$

where $\#\pi_{kl}$ is the number of π_{kl} , and

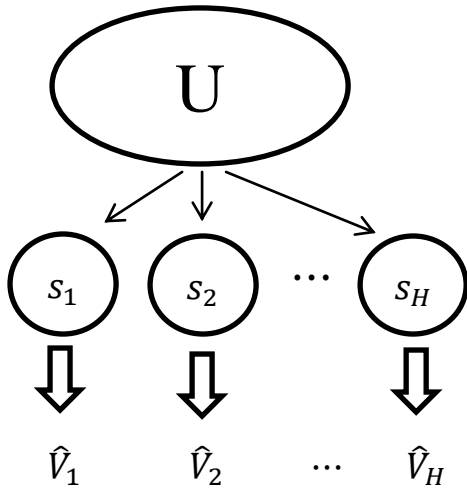
$$\psi_2 = \max \left| \frac{\tilde{\pi}_{kl} - \pi_{kl}}{\pi_{kl}} \right|$$



As we have said, the variance of the estimated total is computed for *the whole population, U*. In SI we have an exact expression for that variance (2.4), which can be used as a reference for the variance with the estimated inclusion probabilities $V^*(\hat{\theta}^*)$ computed from (3.2). A measure of the precision of $V^*(\hat{\theta}^*)$ is then the Relative Bias 1

$$RB.1 = \begin{cases} \frac{V^*(\hat{\theta}^*) - v(\hat{\theta})}{v(\hat{\theta})} & \text{in SI} \\ \frac{V^*(\hat{\theta}^*) - V_{sim}^*(\hat{\theta}_\lambda)}{V_{sim}^*(\hat{\theta}_\lambda)} & \text{in PPS} \end{cases}$$

In the PPS case we do not have an exact expression for $V(\hat{\theta})$ and here we will use a Monte Carlo simulated variance $V_{sim}^*(\hat{\theta}_\lambda) = \frac{1}{99,999,999} \sum_{i=1}^{100,000,000} (\hat{\theta}_{\lambda(i)} - \bar{\theta}_\lambda)^2$, as a reference for $V^*(\hat{\theta}^*)$. The Monte Carlo simulation is done *in advance*, separately from the simulation of the inclusion probabilities, with 100 million repetitions. Here RB.1 is not an exact measure but will give us an idea of the performance of $V^*(\hat{\theta}^*)$.



Next step is to study the empirical performance of the variance *estimator* with the estimated inclusion probabilities $\hat{V}^*(\hat{\theta}^*)$ computed from (3.3). For that sake we do a Monte Carlo simulation study with H samples from which we can compute the variance *estimators*.

In SI we compare the empirical distribution of $\hat{V}^*(\hat{\theta}^*)$ with the unbiased variance estimator $\hat{V}(\hat{\theta})$ computed from (2.5).

In PPS we compare $\hat{V}^*(\hat{\theta}^*)$ with two other variance estimators. One is Rosén's approximate variance estimator $\hat{V}_R(\hat{\theta}_\lambda)$ computed from (2.8), and

the other is Antal's and Tillé's bootstrap variance estimator $\hat{V}_{AT}^*(\hat{\theta}_\lambda)$ described in 2.5. With Antal's and Tillé's bootstrap method we do 1000 bootstrap replications for every sample.

Five measures are used to compare the empirical distributions:

- Lower error rate

$$L = \frac{1}{H} \sum_{i=1}^H I \left[\hat{\theta}_i + 1.96 * \sqrt{\hat{V}_i} < \theta \right]$$

where $\theta = \sum_{k \in U} y_k$ and $I[a] = 1$ if a is true. In SI every $\hat{\theta}_i$ is computed with the exact inclusion probabilities (2.1) and in PPS $\hat{\theta}_i$ is $\hat{\theta}_{\lambda(i)}$ computed from (2.6).

- Upper error rate

$$U = \frac{1}{H} \sum_{i=1}^H I \left[\hat{\theta}_i - 1.96 * \sqrt{\hat{V}_i} > \theta \right]$$

- Total error rate

$$ER = L + U$$

- Relative Bias 2

$$RB.2 = \frac{\bar{\hat{V}} - V}{V} = \frac{B}{V}$$

where $\bar{\hat{V}}$ is the mean of the variance estimates for the H samples compared to the real variance V . In SI we compute V from (2.4) and in PPS we use the variance of the H samples as V ,

$$V = S_{\hat{\theta}_\lambda}^2 = \frac{1}{H-1} \sum_{i=1}^H \left(\hat{\theta}_{\lambda(i)} - \bar{\hat{\theta}}_\lambda \right)^2$$

- Relative Root Mean Squared Error

$$RRMSE = \frac{\sqrt{B^2 + S_{\hat{V}}^2}}{V}$$

here $S_{\hat{V}}^2$ is the variance of the variance estimators for the samples, $S_{\hat{V}}^2 = \frac{1}{H-1} \sum_{i=1}^H \left(\hat{V}_i - \bar{\hat{V}} \right)^2$.

The three measures, L , U and ER , will show us at which degree each of the H variance estimators combined with the estimated total for that sample, covers the true value of the total as a 95% CI. We will use the same estimated total with all variance estimators, in SI the estimated total with the exact inclusion probabilities (2.1) and in PPS the estimated total with the target inclusion probabilities (2.6). The Relative Bias 2 will show us if the estimator has a tendency to over- or underestimate the variance. In the Relative Root Mean Squared Error we get a combination of the bias and the variance of the variance estimator and the optimal is of course to have a small RRMSE.

RB.1 is a measure of precision of the variance with the estimated inclusion probabilities, computed for the whole population and RB.2 for $\hat{V}^*(\hat{\theta}^*)$ is a measure of precision of the mean of the variance estimators with the estimated inclusion probabilities, computed from the H samples. In SI we can also compute RB.2 for $\hat{V}(\hat{\theta})$ computed from (2.5), and this is the value that

RB.2 for $\hat{V}^*(\hat{\theta}^*)$ will converge to. For both SI and PPS, RB.1 and RB.2 show how many repetitions we need to get a stable variance and a stable mean of the variance estimators.

We will also compare how well $\hat{V}^*(\hat{\theta}^*)$ perform when looking at every single one of the H variance estimators. In SI every $\hat{V}_i^*(\hat{\theta}^*)$ is compared with the corresponding $\hat{V}_i(\hat{\theta})$. In PPS we use Rosén's approximate variance estimator $\hat{V}_{R(i)}(\hat{\theta}_\lambda)$ as a reference for $\hat{V}_i^*(\hat{\theta}^*)$. This is not an exact measure but will give us an idea about the development of the precision of $\hat{V}^*(\hat{\theta}^*)$. At least it should work well with large populations. The new measure is the mean of the absolute relative differences

$$\tau = \begin{cases} \frac{1}{H} \sum_{i=1}^H \left| \frac{\hat{V}_i^*(\hat{\theta}^*) - \hat{V}_i(\hat{\theta})}{\hat{V}_i(\hat{\theta})} \right| & \text{in SI} \\ \frac{1}{H} \sum_{i=1}^H \left| \frac{\hat{V}_i^*(\hat{\theta}^*) - \hat{V}_{R(i)}(\hat{\theta}_\lambda)}{\hat{V}_{R(i)}(\hat{\theta}_\lambda)} \right| & \text{in PPS} \end{cases}$$

We will also look at the performance of the estimated total with the estimated inclusion probabilities $\hat{\theta}^*$ computed from (3.1). Fattorini showed (2006) with simulations that the absolute relative bias for $\hat{\theta}^*$ is negligible already at one million repetitions. We will however investigate how every $\hat{\theta}_i^*$ performs in the H samples, compared to the H estimated totals $\hat{\theta}_i$, as

$$v = \begin{cases} \frac{1}{H} \sum_{i=1}^H \left| \frac{\hat{\theta}_i^* - \hat{\theta}_i}{\hat{\theta}_i} \right| & \text{in SI} \\ \frac{1}{H} \sum_{i=1}^H \left| \frac{\hat{\theta}_i^* - \hat{\theta}_{\lambda(i)}}{\hat{\theta}_{\lambda(i)}} \right| & \text{in PPS} \end{cases}$$

In SI we compute every $\hat{\theta}_i$ from (2.1) and in PPS we compute $\hat{\theta}_{\lambda(i)}$ from (2.6).

Storing every single relative difference in τ and v , without absolute value, will allow us to see the empirical distribution of the H relative differences with minimum, mean, maximum and empirical percentiles. The fact that we take the relative differences without the absolute value will allow us to see if the distribution stabilizes around zero. This will be useful specially in PPS.

In the simulation process the computer first generates and stores H samples. After a new initiating of the random number generator, the computer starts simulating the inclusion probabilities. The computer stops regularly and the measures of precision of both $V^*(\hat{\theta}^*)$, $\hat{V}^*(\hat{\theta}^*)$ and $\hat{\theta}^*$ are computed and stored. We can thus see the development of the measures. To get a good precision when comparing the variance estimators, we will use a sample size of $H=100,000$.

3.4 Simulations

3.4.1 Presentation

We will analyze two populations from an administrative register of Swedish business enterprises. The first population, population A, is from the branch 'Manufacturers of optical instruments' with population size $N=323$. The second population, population B, is from the branch 'Manufacturers of metal goods' with $N=2292$. In both populations we use the number of last year's employees as the auxiliary variables x and the number of current year's employees as the variables of interest y .

Population A has no outliers and a very good correlation 0.97 between the two variables, with most of the observation values under 50, see Figure 1 below. The total of the population, $\theta = \sum_{k \in U} y_k$, is 9079 and the maximum sample size with PPS, with all $\lambda_k < 1$, is 38.

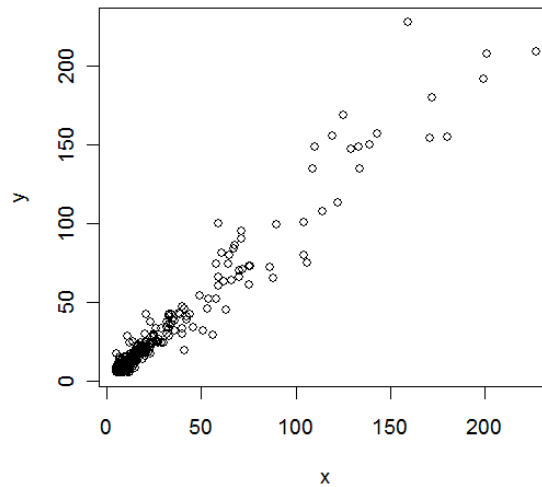


Figure 1: Population A.

Population B has a good correlation between the two variables, 0.75. As we can see in the Figure 2 there is an outlier in the population. The other data from this observation show the same behaviour and there has probably been some kind of restructuring during the year. The correlation without the outlier is 0.96. Here most of the observation values are under 100. The total of the population is 51115 and the maximum sample size with PPS, with all $\lambda_k < 1$, is 174.

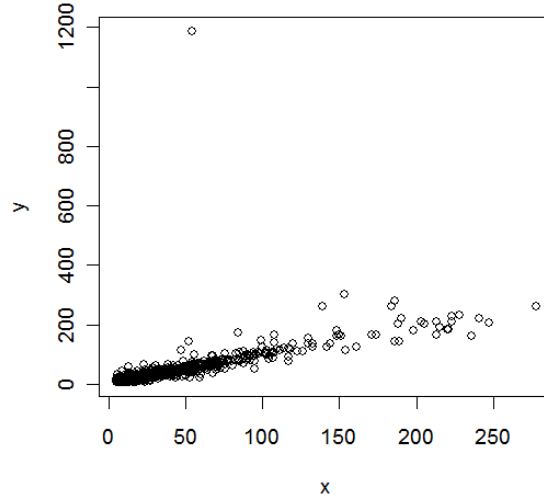


Figure 2: Population B.

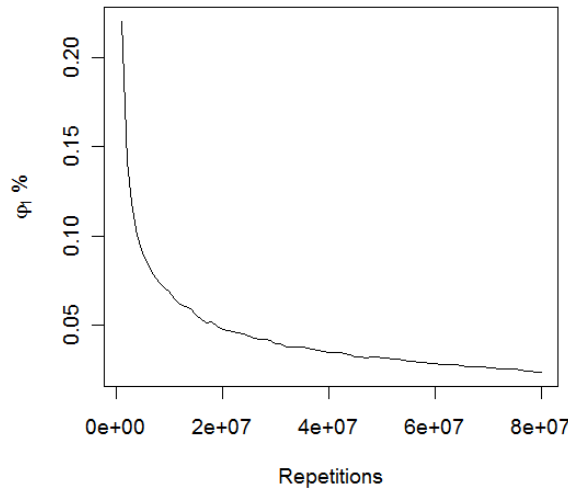


Figure 3: φ_1 , SI simulation with population A, $n=38$.

The question is what value for α , the precision of the estimated first order inclusion probabilities $\tilde{\pi}_k$, we should use. Figure 3 shows φ_1 in a SI simulation with population A, $n=38$. We see a very smooth and regular curve and this curve we also see with all the other designs, both in SI and PPS, when we plot φ_1 . This is the same curve that we see in a plot with $\frac{1}{\sqrt{M}}$, where M is the number of repetitions. If we put different plots of φ_1 on top of each other, we see that they have the same shape. Figure 4 shows φ_1 from seven simulations with seven different designs, both SI and PPS, and $\frac{1}{\sqrt{M}}$, plotted together. The different φ_1 have been rescaled to fit together. In the rescaling process we plotted one of the φ_1 and then we took each of the other φ_1 , one at a time, and tested different values of a constant multiplied to the φ_1 , until the curves coincided.

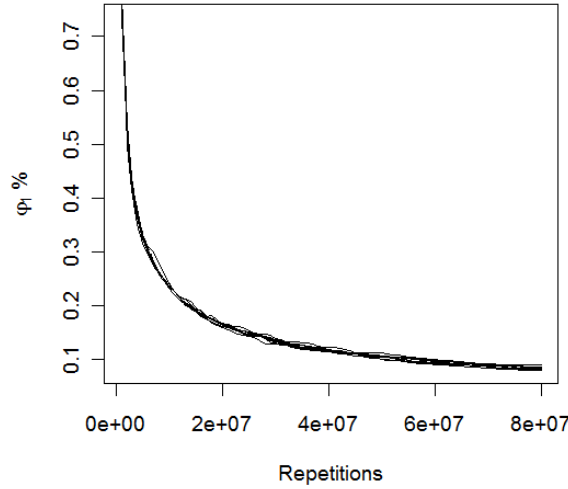


Figure 4: φ_1 for seven different designs and $\frac{1}{\sqrt{M}}$.

This means that if we want to reduce φ_1 to half of its value (at some stage), we have to increase the number of repetitions four times and if we want to reduce φ_1 to a third, we have to increase the number of repetitions nine times.

We will see later that the relationship is the same between the number of repetitions and $\varphi_2, \psi_1, \psi_2, \tau$ and v , for τ however only with SI. The empirical minimum, maximum and percentiles associated with τ and v , also have the same relationship to the number of repetitions, for τ only with SI. The precision of all these measures increases with the same rate and this means that their values will have the same proportion to each other during the whole simulation, with some random variation.

Different designs give different distributions of the estimated inclusion probabilities. The measure φ_1 will however give a basic idea of the precision of the estimated inclusion probabilities and as limit for φ_1 we will use $\alpha=0.05\%$. We will simulate until the limit is reached but the maximum of repetitions will be 320 million. This means for example that we will do about 50 million repetitions when simulating PPS, population A, $n=38$. How many repetitions we need to get estimated inclusion probabilities that are precise enough, is of course always up to the user.

While simulating the inclusion probabilities, the computer stops at 125,000, 250,000 and 500,000 repetitions and computes the measures of precision. Then it stops at every million and from 5 million it stops at every fifth million repetition. From 20 million it stops at every tenth million, from 80 million it stops at every twentieth million and from 160 million it stops at every fortieth million repetition.

To get a fast simulation process, the coding is done in C++ and some parts of the code are also done in parallel programming. We have used the “Mersenne Twister”, mt19937, as the pseudorandom number generator in our simulations. This is a fast generator that is suitable for large simulations.

3.4.2 SI simulation with population A, $n=38$

We start with simulating SI. Here all inclusion probabilities are known and we can get a precise validation of the estimated inclusion probabilities.

At first we analyze the results from sampling with sample size 38, the maximum sample size for PPS with all $\lambda_k < 1$. Here the first order inclusion probabilities are 0.118 and the second order are 0.0135. The Tables and Figures below show the results from the simulation.

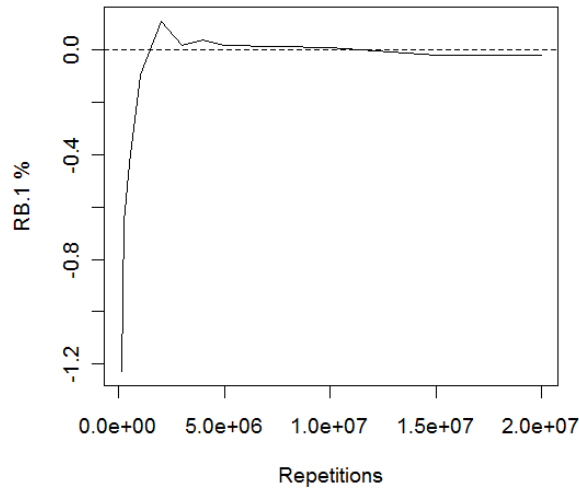


Figure 5: $RB.1$ for $V^*(\hat{\theta}^*)$.

Table A1 in the appendix shows the development of $\varphi_1, \psi_1, \varphi_2, \psi_2$ and $RB.1$ for this design. We can see that all measures, except $RB.1$, have the same relationship to the number of repetitions, described on the previous page, the maximum measures ψ_1 and ψ_2 however with little more random variation. In Figure 5 we see that $RB.1$ stabilizes at about 3-4 million repetitions, in this simulation.

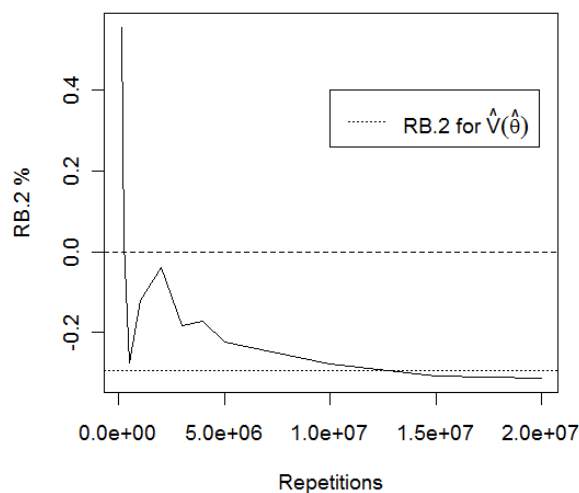


Figure 6: $RB.2$ for $\hat{V}^*(\hat{\theta}^*)$.

Figure 6 shows RB.2 for $\hat{V}^*(\hat{\theta}^*)$. Here in SI this measure will converge to RB.2 for $\hat{V}(\hat{\theta})$ marked in Figure 6, where $\hat{V}(\hat{\theta})$ is computed from (2.5). We will get different RB.2 for $\hat{V}(\hat{\theta})$ in every simulation, depending on the H simulated samples. In this simulation RB.2 for $\hat{V}(\hat{\theta})$ is -0.295 and this is the value that RB.2 for $\hat{V}^*(\hat{\theta}^*)$ will converge to. At 20 million repetitions RB.2 for $\hat{V}^*(\hat{\theta}^*)$ is a bit under -0.295 but if we had continued the simulation we would have seen a stabilization of RB.2 around this value.

Table 1: Empirical distribution measures.

Estimator	L %	U %	ER %	RB.2 %	RRMSE %
$\hat{V}^*(\hat{\theta}^*)$	9.7	0.58	10.3	-0.314	48.9
$\hat{V}(\hat{\theta})$	9.7	0.58	10.3	-0.295	48.9

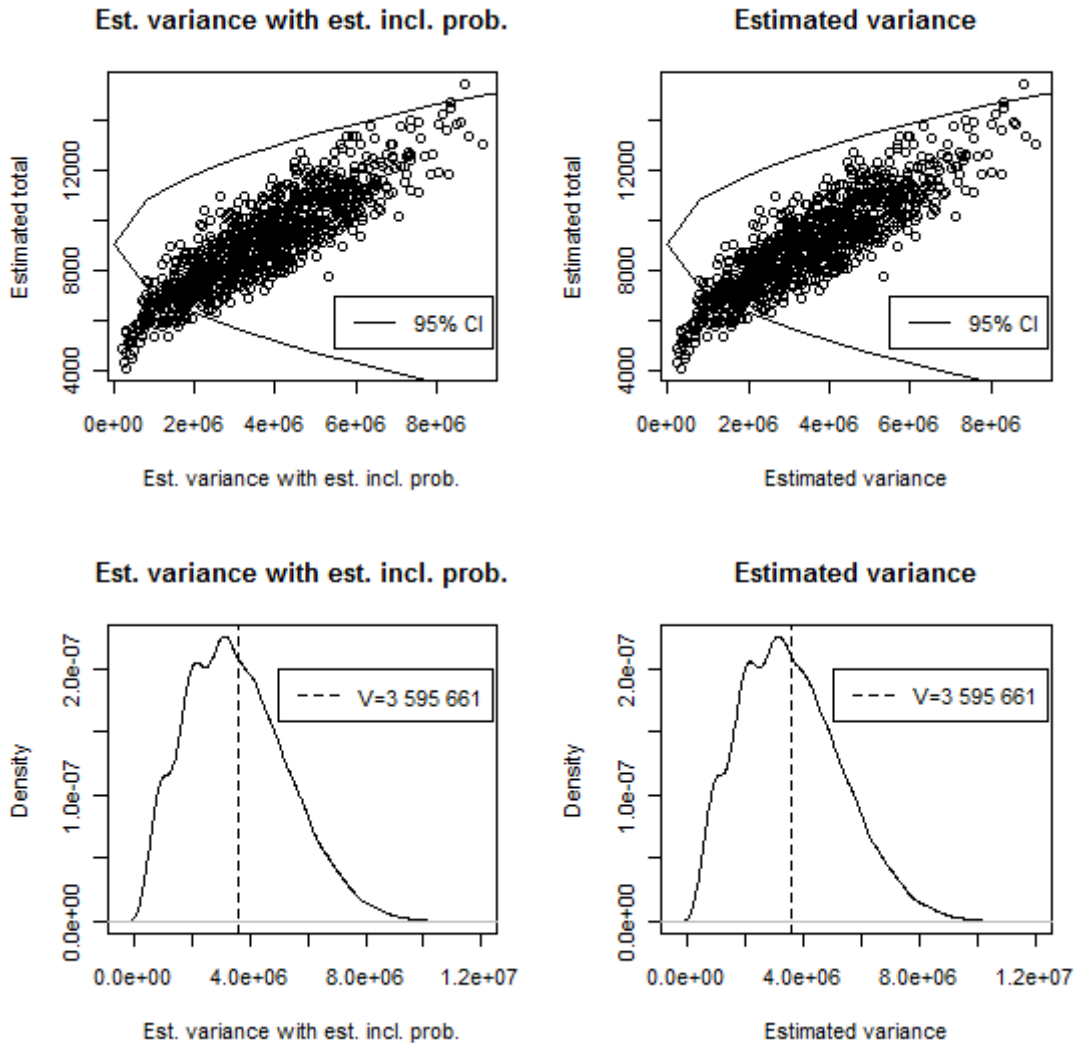


Figure 7: The empirical distributions of $\hat{V}^*(\hat{\theta}^*)$ and $\hat{V}(\hat{\theta})$.

The scatter plots in Figure 7 show the first 1000 values of the estimators. In the figure we also have the 95% confidence interval, which means that the dots inside the lying U-shape are the combination of the estimated total and the estimated variance, whose 95% CI includes the real value of the total.

As we can see in Table 1 and Figure 7, the two variance estimators behave very much the same. The only thing that differs between the two is that $\hat{V}(\hat{\theta})$ has a little better RB.2.

In Table A2 in the appendix we see that the values of the measures L, U and ER are essentially the same in the beginning of the simulation as in the end. With this design these measures do not say so much about the precision of $\hat{V}^*(\hat{\theta}^*)$, the values do not change with increasing number of repetitions. We see however that the RRMSE improves with increasing number of repetitions.

Next we will look at how every single one of the H variance estimators with the estimated inclusion probabilities behaves. As we said, we can take every single relative difference in τ , without absolute value, and get an empirical distribution of these differences. We choose to present the minimum, the 0.5% percentile, the mean, the 99.5% percentile and the maximum. Table 2 shows how τ and the empirical distribution develop with increasing number of repetitions.

Table 2: Development of τ .

Repetitions	Min %	P 0.5 %	Mean %	τ %	P 99.5 %	Max %
125,000	-38	-21.2	0.668	6.55	23.4	49.8
250,000	-28.8	-15.3	0.26	4.61	16.2	33
500,000	-20.8	-11.1	-0.0226	3.29	11.3	22.9
1,000,000	-13.6	-7.62	0.0913	2.32	8.25	14.1
2,000,000	-11	-5.25	0.191	1.63	5.87	10.5
3,000,000	-8.18	-4.42	0.0714	1.33	4.69	9.31
4,000,000	-8.3	-3.76	0.0762	1.15	3.99	7.46
5,000,000	-6.18	-3.44	0.044	1.02	3.51	6.52
10,000,000	-4.25	-2.46	0.0159	0.726	2.48	4.99
15,000,000	-3.81	-2.02	-0.00451	0.592	1.99	3.72
20,000,000	-3.43	-1.77	-0.0127	0.515	1.74	3.19

In Table 2 we can see that the mean rather quickly stabilizes around zero and that τ decreases to half of its value if we increase the repetitions four times. Also the distribution measures follow this pattern, with some random variation. We see for example that the percentiles at 15 million repetitions are $\pm 2\%$. If we then want 99% of the relative differences to be between $\pm 1\%$, we have to increase the number of repetitions to about 60 million. Figure A1 and A2 in the appendix are graphical presentations of Table 2.

In Figure A3 in the appendix we can see that the empirical distribution of the relative differences is stable already at 125,000 repetitions. The shape of the density curves does not change so much with increasing number of repetitions.

Table 3 shows the development of v with empirical distribution of the relative differences in the same manner as with τ . We can see that the measures in this table behave in the same way

as in Table 2, concerning their relationship to the number of repetitions. Figure A4 in the appendix shows this relationship. The pattern in Figure A5 in the appendix is the same as in Figure A3 with little change with increasing number of repetitions.

Table 3: Development of v .

Repetitions	Min %	P 0.5 %	Mean %	v %	P 99.5 %	Max %
125,000	-0.682	-0.426	0.033	0.15	0.483	0.716
250,000	-0.45	-0.291	-0.0252	0.0882	0.251	0.418
500,000	-0.329	-0.232	-0.0291	0.0681	0.172	0.286
1,000,000	-0.283	-0.181	-0.0131	0.056	0.163	0.247
2,000,000	-0.22	-0.136	-0.0134	0.0374	0.0929	0.156
3,000,000	-0.164	-0.108	-0.00813	0.0328	0.0935	0.149
4,000,000	-0.145	-0.0912	-0.00246	0.027	0.0793	0.128
5,000,000	-0.122	-0.0831	-0.00497	0.0246	0.0692	0.107
10,000,000	-0.0748	-0.049	-0.00603	0.0143	0.0362	0.058
15,000,000	-0.0587	-0.041	-0.00247	0.0132	0.041	0.0652
20,000,000	-0.057	-0.0347	-0.0033	0.0112	0.0313	0.0493

Finally we apply Fattorini's algorithm, described on page 13, to this design. We use $L=1,000$, $\delta=0.1\%$ and $G=100$, the values Fattorini used in his article. The maximum of steps is 100,000 (100 million repetitions). At each stage, Table 4 first shows how many steps that was needed until $\gamma_i \leq \beta$ and $L \cdot i = M > M_0$, noted Pr. steps. Then it shows how many steps that was needed until also $g(i) \geq G$. Then we have the values of ϵ and β at that stage, and $g(i)$, $\hat{\theta}_i^*$ and $\hat{V}_i^*(\hat{\theta}^*)$ at the end of that stage.

We can see from Table 4 that the values of ϵ and β are very conservative. In Table 4 we see that in stage 2, after 6,210 steps (6,210,000 repetitions), the measure Pr.steps shows that the probability that the absolute relative difference is larger than 1% (ϵ), is less or equal to 1% (β). Table 3 shows that the maximum and minimum are not absolute higher than 1% (compare with ϵ) already at 125,000 repetitions, and in fact the absolute relative difference is, already here, larger than 0.5% (compare with ϵ) with probability less than 1% (compare with β), the result that Table 4 shows after 24,570 steps. At stage 4 we entered $\epsilon=0.1\%$ and $\beta=1\%$, and with those values it was not enough with 100,000 steps (100 million repetitions) to reach the limit values. In Table 3 we see that already at 4 million repetitions (4,000 steps), the probability that the absolute relative difference is larger than 0.1% (compare with ϵ), is less than 1% (compare with β). The values of Pr. steps do not change more than ± 1 in other simulations with this design. If we would have used the harder precision criteria, even more repetitions would have been needed to reach the limit values, and the results would have been even more misleading. The empirical distribution in Table 3 is based on 100,000 samples. If we increase the number of samples to 1 million, we still get that at 125,000 repetitions, the maximum absolute relative difference is below 1%. Here, at 125,000 repetitions, the probability that the absolute relative difference is larger than 0.5, is just above 1%, but as before, at 4 million repetitions, the probability that the absolute relative difference is larger than 0.1% , is less than 1%.

We can state that Fattorini's algorithm ensures stability for $\hat{\theta}^*$ and $\hat{V}^*(\hat{\theta}^*)$, but that the values of ϵ and β are very conservative.

Table 4: Fattorini's algorithm.

Stage	Pr. steps	Steps	ϵ %	β %	$g(i)$	$\hat{\theta}_i^*$	$\hat{V}_i^*(\hat{\theta}^*)$
1	56	3,505	10	10	100	10,617	5,177,081
2	6,210	6,238	1	1	100	10,620	5,215,153
3	24,570	24,570	0.5	1	940	10,617	5,155,123

With a Intel Core i3-350M 2.27 GHz processor, it takes 7 minutes and 5 seconds to produce a matrix with the inclusion probabilities, based on 20 million repetitions, for this design.

To summarize this design, we can say that we are able to simulate precise estimated inclusion probabilities, within a short period of time. 20 million repetitions only take about 7 minutes and if we multiply the number of repetitions with 16, to 320 million repetitions, we would get very precise values of τ and ν , about 0.13% and 0.0028%, with 99% CI about $\pm 0.44\%$ and $\pm 0.008\%$ for the empirical distributions.

The Relative Biases show that we need at least 3-4 million repetitions to stabilize $V^*(\hat{\theta}^*)$ and $\bar{V}^*(\hat{\theta}^*)$. If we use Fattorini's algorithm to just measure the stability of $\hat{\theta}^*$ and $\hat{V}^*(\hat{\theta}^*)$, the number of steps until $g(i) \geq G$, with $\delta=0.1\%$, varies between about 3,000 to 5,000.

3.4.3 PPS simulation with population A, $n=38$

Next we simulate PPS with the same sample size $n=38$. We see in Figure 8 that φ_1 with the target inclusion probabilities, can be used as a precision measure for the estimated first order inclusion probabilities with this sampling design. The size of the population $N=323$ and the sample size $n=38$, are both big enough for that purpose. We see that the relationship between the number of repetitions and φ_1 , follow the expected pattern. This can also be seen in Table A3 in the appendix.

Rosén's approximate variance for the whole population $V_R(\hat{\theta}_\lambda)$ computed from (2.7), is here 89,170, which is below $V_{sim}^*(\hat{\theta}_\lambda)=89,250$. The relative difference is -0.09%.

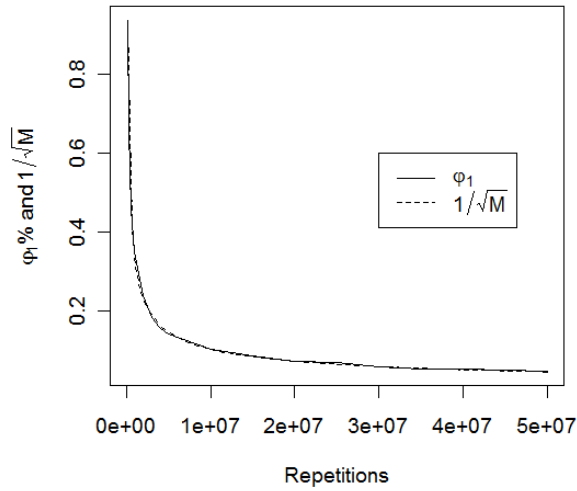


Figure 8: φ_1 and $\frac{1}{\sqrt{M}}$

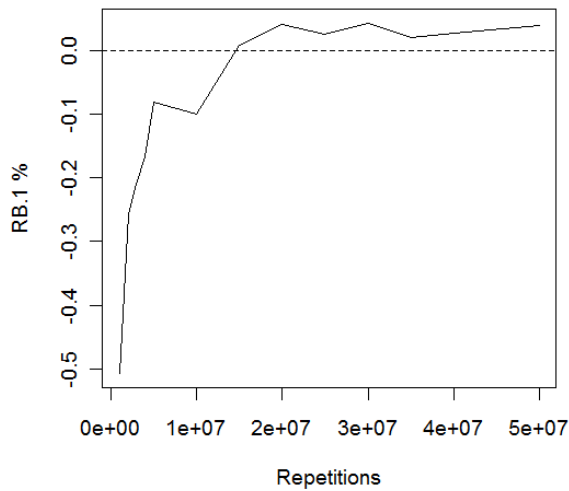


Figure 9: RB.1, 1-50 million repetitions.

With PPS we need more repetitions to get the required precision of φ_1 , compared to SI. The observations with the smallest x values have target inclusion probabilities λ_k that are only about 0.02. As we can see in Figure 9, RB.1 stabilizes at about 15 million repetitions with this design. In this simulation RB.1 is a bit above zero at 50 million repetitions, but as we know, in PPS we use a simulated variance as reference for the variance with the estimated inclusion probabilities. Figure 9 shows the values of RB.1 from one million repetitions, which gives a more precise presentation of the development of the measure in the end of the simulation.

In Figure 10 we see that RB.2 stabilizes at about 15 million repetitions. We also see that, in this simulation, RB.2 converges to a value above zero.

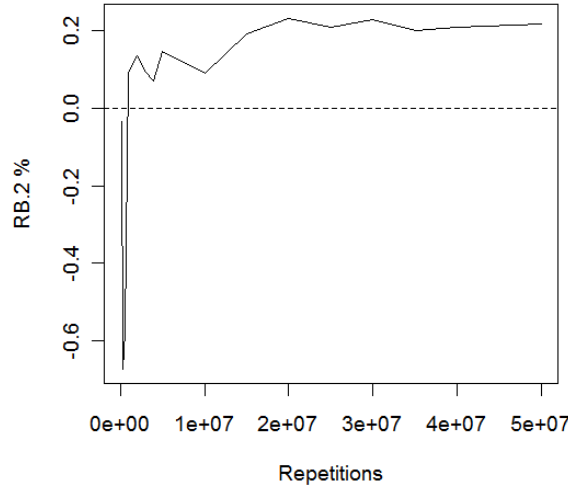


Figure 10: RB.2 for $\hat{V}^*(\hat{\theta}^*)$.

Table 5: Empirical distribution measures.

Estimator	L %	U %	ER %	RB.2 %	RRMSE %
$\hat{V}^*(\hat{\theta}^*)$	4.7	1.6	6.2	0.22	64.1
$\hat{V}_R(\hat{\theta}_\lambda)$	4.9	1.7	6.6	-0.50	60.8
$\hat{V}_{AT}^*(\hat{\theta}_\lambda)$	4.7	1.5	6.2	1.24	66.8

As we saw in the simulation with SI, $\hat{V}^*(\hat{\theta}^*)$ had almost converged to $\hat{V}(\hat{\theta})$ at the end of the simulations and, as seen in Table 5, here we can state that Rosén's variance estimator has a negative Relative Bias and that Antal's and Tillé's variance estimator has a positive Relative Bias. This is a tendency we also see with other simulations with the same design. A clear pattern is also that $\hat{V}_R(\hat{\theta}_\lambda)$ has the lowest RRMSE and that $\hat{V}_{AT}^*(\hat{\theta}_\lambda)$ has the highest. On the other hand, $\hat{V}_R(\hat{\theta}_\lambda)$ has the highest ER, also a tendency we see in other simulations with the same design. An explanation for the high ER is that $\hat{V}_R(\hat{\theta}_\lambda)$ underestimates the variance and as we can see in the scatter plots in the Figure on the next side, this means that the chance to miss the 95% CI increases.

Table A4 in the appendix shows the development of the measures L, U, ER and RRMSE. We see that L, U and ER change a little in the beginning of the simulation but that they stabilize rather quickly at about 2 million repetitions. We also see that RRMSE varies a little before it stabilizes in the middle of the simulation.

Figure 11 shows the empirical distribution of the three variance estimators. In the Figure we see a clear group of observations to the right of the main group. This can also be seen with other PPS designs. Figure A6 in the appendix shows only the distribution of $\hat{V}^*(\hat{\theta}^*)$, with the first 10,000 observations. Here we can see the extra group more clearly and we also see that there are observations overlapping the two groups.

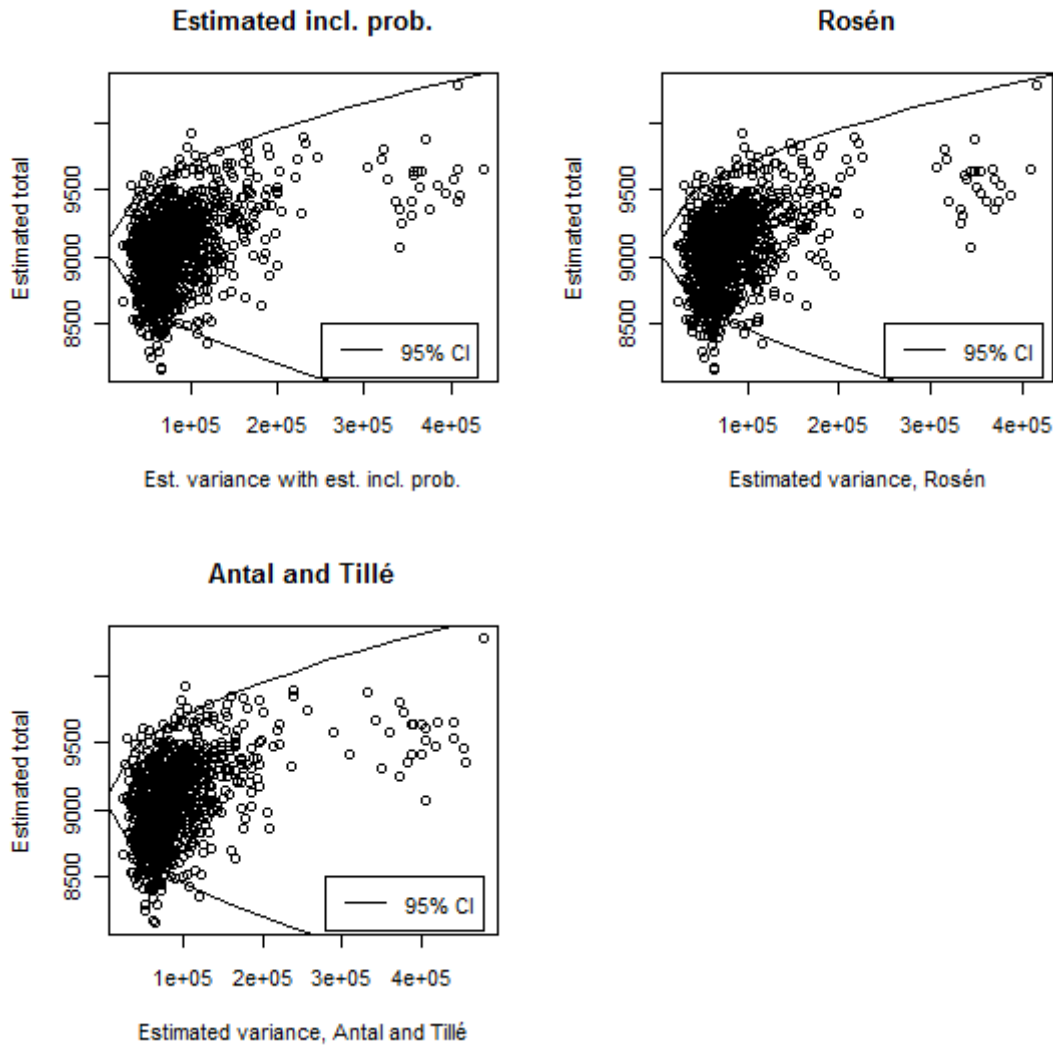


Figure 11: The first 1000 of the variance estimators with 95% CI.

As we can see in Table 6, the mean stabilizes at just above 3%. But the precision of τ does not increase with the number of repetitions in the end of the simulation, after 10 million repetitions nothing much really happens with all the measures. We can see in Table A7 in the appendix (page 61), that τ works much better with sample size $n=20$. Aires and Rosén (2005) showed that the approximation of π_k with λ_k , can be more precise with a smaller sample size and that could be a reason why Rosén's variance estimator (2.8), which is defined with λ_k , performs better as a reference for $\hat{V}^*(\hat{\theta}^*)$ with sample size $n=20$. In SI every value in Table 2 (page 23) converges to zero. With PPS we will see that the mean stabilizes at above zero with every design we simulate, and this means that Rosén's variance estimator more or less has a negative bias with every design. If every one of Rosén's H variance estimators, in each simulation, would have the same bias, every measure in the "tau tables" would converge to that value. This is not the case, at least not with this design. We will subtract the values of τ and the percentiles in the end of each simulation, with the value of the mean, to get an approximate value of τ and the percentiles. At the end of this simulation τ is then 1.07% and we get a 99% CI (-10.02, 10.9) for the empirical distribution. We can state that this values are not the values we would have got if we had compared $\hat{V}^*(\hat{\theta}^*)$ with an unbiased variance estimator. When we look at each of H variance estimators, Rosén's variance estimator has a negative bias at about

3% and it is also not precise enough to be used as a reference for $\hat{V}^*(\hat{\theta}^*)$, with this sample size. If we look at the relative difference between the means of H variance estimators, the relative difference is 0.7%.

Table 6: Development of τ .

Repetitions	Min %	P 0.5 %	Mean %	τ %	P 99.5 %	Max %
125,000	-91.7	-30.9	3.04	9.07	40.5	95.8
250,000	-67.6	-22.3	2.66	6.84	29.2	79.4
500,000	-42.5	-15.6	2.66	5.58	22.5	61.5
1,000,000	-29	-11	3	4.95	19.8	43.3
2,000,000	-25.6	-8.98	3.01	4.55	17.2	45.9
3,000,000	-21.2	-8.33	2.97	4.36	16.1	34.8
4,000,000	-18	-7.95	2.98	4.3	15.7	29.9
5,000,000	-15.4	-7.7	3.02	4.27	15.3	29.8
10,000,000	-13.9	-7.37	2.95	4.13	14.6	26.3
15,000,000	-13.5	-7.17	3.04	4.15	14.6	24.5
20,000,000	-13.2	-7.09	3.06	4.15	14.5	23.6
25,000,000	-13.1	-7.04	3.05	4.13	14.4	23.2
30,000,000	-13.1	-6.98	3.06	4.13	14.4	22.8
35,000,000	-13.1	-7.02	3.02	4.11	14.3	22.8
40,000,000	-13.1	-7	3.03	4.11	14.3	22.6
50,000,000	-13	-6.98	3.04	4.11	14.3	22.7

The empirical densities in Figure A7 in the appendix do not behave as in SI, where the shape of the empirical densities were the same during the whole simulation. Here the shape stabilizes in the third graph at 10 million repetitions. Other simulations show that the empirical density at 50 million repetitions and the measures in the end of Table 6, remain almost the same all the way up to 320 million repetitions.

We can see in Table 7 and in Figure A8 in the appendix, that v behave as expected. The estimated total with the target inclusion probabilities work well as a reference for the estimated total with the estimated inclusion probabilities.

Table 7: Development of v .

Repetitions	Min %	P 0.5 %	Mean %	v %	P 99.5 %	Max %
125,000	-0.578	-0.334	0.0068	0.102	0.341	0.663
250,000	-0.462	-0.265	-0.00326	0.0725	0.239	0.404
500,000	-0.337	-0.187	-0.00402	0.0516	0.162	0.286
1,000,000	-0.202	-0.124	0.00163	0.0377	0.122	0.206
2,000,000	-0.15	-0.0809	0.00299	0.0255	0.0843	0.14
3,000,000	-0.119	-0.0628	0.00131	0.0198	0.0693	0.129
4,000,000	-0.101	-0.0585	0.00107	0.0181	0.0622	0.106
5,000,000	-0.0977	-0.0505	0.00239	0.016	0.0552	0.0886
10,000,000	-0.0695	-0.0398	0.000142	0.0116	0.0373	0.0652
15,000,000	-0.0559	-0.0308	0.000548	0.00963	0.0325	0.0571
20,000,000	-0.0445	-0.026	0.000696	0.00833	0.0294	0.0503

25,000,000	-0.0439	-0.0239	0.000489	0.00775	0.0284	0.048
30,000,000	-0.0357	-0.0212	0.000556	0.00693	0.0261	0.0454
35,000,000	-0.0374	-0.0204	5.53e-005	0.00632	0.0237	0.0401
40,000,000	-0.0324	-0.0194	0.000171	0.00619	0.0235	0.0391
50,000,000	-0.0272	-0.0166	0.000508	0.00559	0.0245	0.0419

Table 8: Fattorini's algorithm.

Stage	Pr. steps	Steps	ϵ %	β %	$g(i)$	$\hat{\theta}_i^*$	$\hat{V}_i^*(\hat{\theta}^*)$
1	108	4,767	10	10	100	8,677	60,998
2	13,740	13,740	1	1	128	8,678	61,247
3	54,407	54,407	0.5	1	1,069	8,679	61,706

If we compare Table 8 with Table 7, we see the same thing as with the last design, the values of ϵ and β are very conservative. At stage 4 we entered $\epsilon=0.1\%$ and $\beta=1\%$, and with those values it was not enough with 100,000 steps (100 million repetitions) to reach the limit values.

It takes 48 minutes and 10 seconds to produce a matrix with the inclusion probabilities, based on 50 million repetitions, for this design.

It is more time consuming to simulate PPS than to simulate SI. We also need more repetitions to get the required precision of φ_1 . But if we compare Table 7 with Table 3, page 24, we see that v has lower values in Table 7, at the same number of repetitions, compared to Table 3. This also applies to the measures of the empirical distribution, where the precision of the measures is higher in Table 7 compared to Table 3, at the same number of repetitions. The reason is that, in PPS, the observations with large λ_k has a high probability to be drawn and the precision of their estimated first order inclusion probabilities quickly reach a high precision. When we compute the estimated totals, the observations with large y_k and $\tilde{\pi}_k$ values have larger contribution to sum, than those observations with small y_k and $\tilde{\pi}_k$ values. This means that we get estimated totals with estimated inclusion probabilities that are more precise in PPS than in OSU, with the same sample size and same number of repetitions.

At the end of this simulation we have a very precise value of v , 0.0056% and a 99% CI (-0.017%, 0.024%) for the empirical distribution. The value of τ , after subtraction, is 1.07% and a 99% CI for the empirical distribution is (-10.02%, 10.9%). To quadruple the number of repetitions to 200 million, would take less than 4 hours and we would then double the precision of τ and v .

Rosén's variance $V_R(\hat{\theta}_\lambda)$ has a negative relative bias, -0.09%, compared to $V_{sim}^*(\hat{\theta})$. We could see that $\hat{V}_R(\hat{\theta}_\lambda)$ has a negative RB.2, 0.72% below RB.2 for $\hat{V}^*(\hat{\theta}^*)$ and that Antal's and Tillé's variance estimator $\hat{V}_{AT}^*(\hat{\theta}_\lambda)$ has a positive RB.2, 1.02% above RB.2 for $\hat{V}^*(\hat{\theta}^*)$. We could also see that $\hat{V}_R(\hat{\theta}_\lambda)$ has the lowest RRMSE and that $\hat{V}_{AT}^*(\hat{\theta}_\lambda)$ has the highest RRMSE.

The Relative Biases show that we need at least about 15 million repetitions to stabilize $V^*(\hat{\theta}^*)$ and $\tilde{V}^*(\hat{\theta}^*)$. If we use Fattorini's algorithm to just measure the stability of $\hat{\theta}^*$ and $\hat{V}^*(\hat{\theta}^*)$, the number of steps, with $\delta=0.1\%$, varies between about 3,000 and 6,000.

3.4.4 PPS simulation with population B, $n=174$

Now we turn to population B and simulate PPS with the maximal sample size with all $\lambda_k < 1$. It took 60 million repetitions until we reached the precision measure for the estimated first order inclusion probabilities.

Rosén's variance for the whole population has a negative relative bias compared to $V_{sim}^*(\hat{\theta}_\lambda)$, -0.07%.

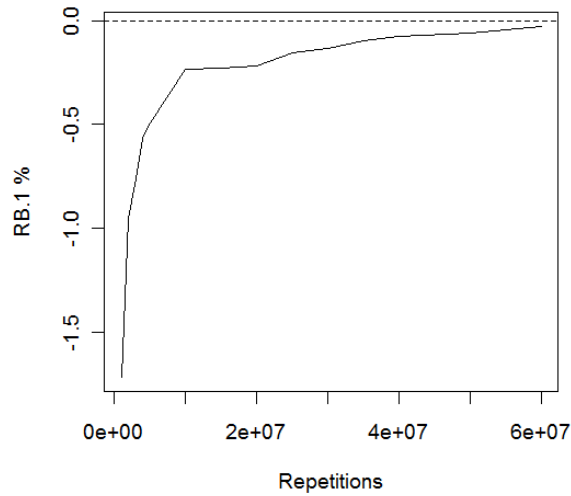


Figure 12: RB.1, 1-60 million repetitions.

Figure 12 shows that RB.1 gets closer to the simulated variance with increasing number of repetitions and maybe we could say that RB.1 stabilizes at about 30 million repetitions. Figure 12 shows the values of RB.1 from one million repetitions, which gives a more precise presentation of the development of the measure in the end of the simulation. Table A5 in the appendix shows the development of φ_1 , ψ_1 and RB.1 for this design.

In Figure 13 we see the same tendency as in with RB.1, RB.2 increases in the end of the simulation. Also here we could say that RB.2 stabilizes at about 30 million repetitions.

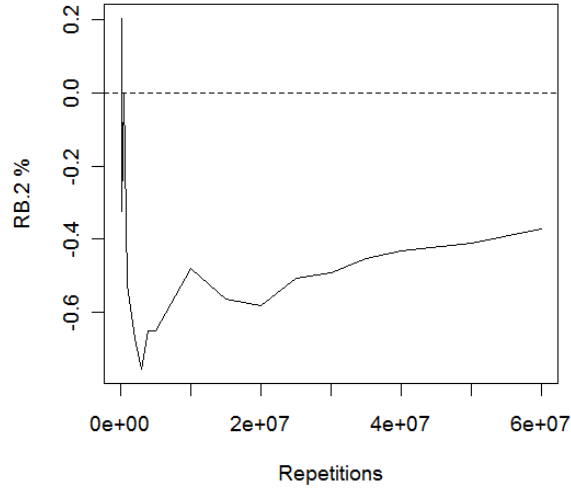


Figure 13: RB.2 for $\hat{V}^*(\hat{\theta}^*)$.

Table 9: Empirical distribution measures.

Estimator	L %	U %	ER %	RB.2 %	RRMSE %
$\hat{V}^*(\hat{\theta}^*)$	25.4	0.02	25.4	-0.37	177
$\hat{V}_R(\hat{\theta}_\lambda)$	25.5	0.02	25.5	-0.42	176
$\hat{V}_{AT}^*(\hat{\theta}_\lambda)$	25.5	0.02	25.5	-0.78	177

We see in Table 9 that Antal's and Tillé's variance estimator $\hat{V}_{AT}^*(\hat{\theta}_\lambda)$ has a little negative Relative Bias compared to the other two. In simulations without the outlier the Relative Bias is even more negative, about -1% compared to the others. With all the other PPS designs with population B, $\hat{V}_{AT}^*(\hat{\theta}_\lambda)$ has a positive Relative Bias which disappears in some of the designs if we simulate without the outlier. Without the outlier, the ER for the variance estimators is just above 6% and the RRMSE is about 47%.

Table A6 in the appendix shows that all the measures L, U, ER and RRMSE stabilize already at 500,000 repetitions in this simulation.

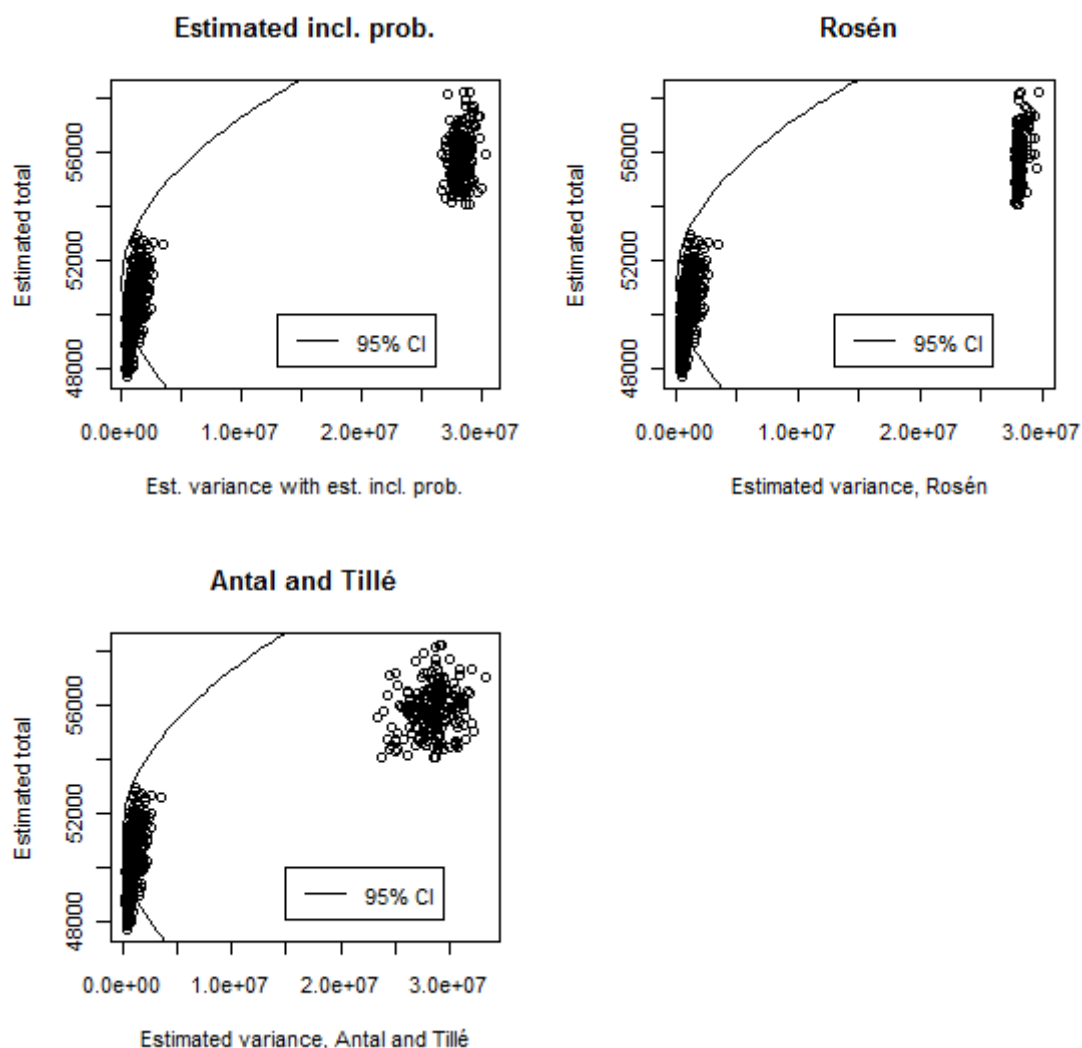


Figure 14: The first 1000 of the variance estimators with 95% CI.

In Figure 14 we can clearly see the group of samples with the outlier. We see that $\hat{V}_{AT}^*(\hat{\theta}_\lambda)$ is most spread and that $\hat{V}_R(\hat{\theta}_\lambda)$ is least spread in the outlier group. Figure A9 in the appendix shows the first thousand observations of $\hat{V}^*(\hat{\theta}^*)$ in the group without the outlier. Here we see the same thing as we could see in PPS with population A, there is a group of observations to the right of the main group.

Table 10: Development of τ .

Repetitions	Min %	P 0.5 %	Mean %	τ %	P 99.5 %	Max %
125,000	-199	-77.4	0.564	17.7	84.7	251
250,000	-136	-55.5	0.246	12.5	58.7	137
500,000	-96.5	-38.5	0.815	8.93	42.7	103
1,000,000	-67.9	-27.7	0.571	6.3	29.7	71.3
2,000,000	-49	-19.7	0.463	4.51	20.4	54.1
3,000,000	-37.7	-16.1	0.452	3.74	16.6	42.5
4,000,000	-35.8	-13.8	0.488	3.3	14.7	39.6

5,000,000	-30.9	-12.6	0.48	3.01	13.3	36.6
10,000,000	-22.9	-8.66	0.587	2.31	10.2	26.8
15,000,000	-19.3	-7.28	0.572	2.02	8.54	20.4
20,000,000	-17.7	-6.42	0.55	1.84	7.65	18.4
25,000,000	-14.1	-5.87	0.562	1.73	7.11	16.1
30,000,000	-13.7	-5.37	0.563	1.65	6.59	14
35,000,000	-12.8	-4.92	0.569	1.59	6.35	13
40,000,000	-13.7	-4.66	0.571	1.54	6.1	11.9
50,000,000	-13.5	-4.29	0.588	1.48	5.7	12.2
60,000,000	-11.1	-3.99	0.601	1.44	5.49	10

Table 10 shows that τ works better as a measure of precision than with the last design with population A. The mean stabilizes at about 0.6 and if we subtract every value with 0.6, τ is 0.84% at the end of the simulation with a 99% CI (-4.59%, 4.89%) for the empirical distribution. When we look at each of H variance estimators, Rosén's variance estimator has a negative bias at about 0.6%. If we look at the relative difference between the means of H variance estimators, the relative difference is 0.04%. When we compare Table 10 with Table A8 in the appendix (page 64), we can see that τ has almost the same values in the beginning of both the tables, at the same number of repetition. But here we can see that the precision of the measures of the empirical distribution is less in Table 10 compared to Table A8. Computation of the estimated variances includes the second order inclusion probabilities, and it looks like we get more extreme values of the estimated variances with the estimated inclusion probabilities in PPS compared to SI, at the same number of repetitions. This is also the case if we compare PPS with SI in other simulations with the same sample size with population B.

We see from Figure A10 in the appendix that the shape of the empirical distribution of the relative differences change with increasing number of repetitions. In the beginning the distribution includes more extreme values and then, with increasing number of repetitions, the distribution gets more centered around its mean.

We see from Table 11 and from Figure A11 in the appendix, that v and the relative differences behave as expected. If we compare Table 11 with Table A9 in the appendix (page 64), we can see the same thing as with the last design, v has lower values in Table 11, at the same number of repetitions, compared to Table A9, and that this also applies to the measures of the empirical distribution, where the precision of the measures is higher in Table 11 compared to Table A9, at the same number of repetitions.

Table 11: Development of v .

Repetitions	Min %	P 0.5 %	Mean %	v %	P 99.5 %	Max %
125,000	-0.357	-0.197	-0.00283	0.0601	0.192	0.302
250,000	-0.236	-0.142	-0.00413	0.0429	0.137	0.227
500,000	-0.169	-0.0961	0.00324	0.0307	0.103	0.169
1,000,000	-0.109	-0.07	-0.000461	0.0217	0.0711	0.13
2,000,000	-0.0871	-0.0538	-0.00234	0.0161	0.0493	0.0889
3,000,000	-0.0759	-0.0466	-0.00334	0.0136	0.0389	0.0724
4,000,000	-0.0698	-0.0375	-0.00214	0.0113	0.0334	0.0586
5,000,000	-0.0593	-0.0342	-0.00202	0.0102	0.0301	0.0563
10,000,000	-0.0357	-0.0212	0.000333	0.00666	0.0218	0.0339

15,000,000	-0.03	-0.019	-0.000539	0.0058	0.018	0.0314
20,000,000	-0.0271	-0.0172	-0.000865	0.00518	0.0154	0.0263
25,000,000	-0.0245	-0.0142	-0.000422	0.00436	0.0136	0.0221
30,000,000	-0.0226	-0.0134	-0.000459	0.00406	0.0126	0.0237
35,000,000	-0.0223	-0.0123	-0.00042	0.00371	0.0115	0.0199
40,000,000	-0.0225	-0.0117	-0.000179	0.00346	0.0109	0.0199
50,000,000	-0.0171	-0.01	0.000159	0.00305	0.0098	0.0177
60,000,000	-0.0159	-0.00893	0.000534	0.0029	0.00955	0.0151

Table12: Fattorini's algorithm.

Stage	Pr. steps	Steps	ϵ %	β %	$g(i)$	$\hat{\theta}_i^*$	$\hat{V}_i^*(\hat{\theta}^*)$
1	269	12,019	10	10	100	56,618	26,262,811
2	32,090	32,116	1	1	100	56,617	27,256,991

Table 12 shows Fattorini's algorithm for this design. At stage 3 we entered $\epsilon=0.5\%$ and $\beta=1\%$, and we did not reach the limit values within 100,000 steps (100 million repetitions). As before, we can state that the values of ϵ and β are very conservative.

It takes 15 hours, 31 minutes and 32 seconds to produce a matrix with the inclusion probabilities, based on 60 million repetitions, for this design.

At the end of this simulation we have a very precise value of v , 0.0029% and a 99% CI (-0.009%, 0.010%) for the empirical distribution. After the subtraction τ is 0.84% at the end of the simulation and a 99% CI for the distribution, is (-4.59, 4.89). Population B is about seven times bigger than population A and the process of estimating the inclusion probabilities is of course much more time consuming. To quadruple the number of repetitions in the end of the simulation with this design, would take about two and a half days.

With this design, Rosén's approximate variance $V_R(\hat{\theta}_\lambda)$ has a negative relative bias, -0.07%, compared to $V_{sim}^*(\hat{\theta})$, but $\hat{V}_R(\hat{\theta}_\lambda)$ has the same RB.2 as $\hat{V}^*(\hat{\theta}^*)$. Here Antal's and Tillé's variance estimator $\hat{V}_{AT}^*(\hat{\theta}_\lambda)$ has a negative RB.2, 0.41% below RB.2 for $\hat{V}^*(\hat{\theta}^*)$. In simulations without the outlier, RB.2 for $\hat{V}_{AT}^*(\hat{\theta}_\lambda)$ is even more negative, about -1% compared to the others. With this design all three variance estimators have the same RRMSE.

The Relative Biases show that we need about 30 million repetitions to stabilize $V^*(\hat{\theta}^*)$ and $\bar{V}^*(\hat{\theta}^*)$. If we use Fattorini's algorithm to just measure the stability of $\hat{\theta}^*$ and $\hat{V}^*(\hat{\theta}^*)$, the number of steps, with $\delta=0.1\%$, varies between about 5,000 to 10,000.

3.5 Summary of all the simulations.

All the measures are in percent.

Population A

SI

Population A, SI, n=38, 20 million repetitions.

Estimator	L	U	ER	RB.2	RRMSE	0.05%	τ	99.5%	0.05%	ν	99.5%
$\hat{V}^*(\hat{\theta}^*)$	9.7	0.58	10.3	-0.31	48.9	-1.77	0.51	1.74	-0.035	0.011	0.031
$\hat{V}(\hat{\theta})$	9.7	0.58	10.3	-0.29	48.9						

Population A, SI, n=100, 10 million repetitions.

Estimator	L	U	ER	RB.2	RRMSE	0.05%	τ	99.5%	0.05%	ν	99.5%
$\hat{V}^*(\hat{\theta}^*)$	5.6	0.94	6.6	0.20	26.6	-0.80	0.25	0.85	-0.014	0.0052	0.017
$\hat{V}(\hat{\theta})$	5.6	0.94	6.6	0.16	26.5						

PPS

Population A, PPS, n=38, 50 million repetitions.

Estimator	L	U	ER	RB.2	RRMSE	0.05%	τ	99.5%	0.05%	ν	99.5%
$\hat{V}^*(\hat{\theta}^*)$	4.7	1.6	6.2	0.22	64.1	-10.0*	1.07*	10.9*	-0.017	0.0056	0.0245
$\hat{V}_R(\hat{\theta}_\lambda)$	4.9	1.7	6.6	-0.50	60.8						
$\hat{V}_{AT}^*(\hat{\theta}_\lambda)$	4.7	1.5	6.2	1.24	66.8						

Population A, PPS, n=20, 100 million repetitions.

Estimator	L	U	ER	RB.2	RRMSE	0.05%	τ	99.5%	0.05%	ν	99.5%
$\hat{V}^*(\hat{\theta}^*)$	5.05	1.7	6.8	-0.13	79.7	-8.8*	0.19*	9.1*	-0.025	0.0074	0.026
$\hat{V}_R(\hat{\theta}_\lambda)$	5.54	1.9	7.5	-0.52	74.9						
$\hat{V}_{AT}^*(\hat{\theta}_\lambda)$	5.04	1.7	6.7	1.56	86.9						

* value has been subtracted with the mean.

Population B

SI

Population B, SI, n=174, 35 million repetitions.

Estimator	L	U	ER	RB.2	RRMSE	0.05%	τ	99.5%	0.05%	ν	99.5%
$\hat{V}^*(\hat{\theta}^*)$	8.9	0.39	9.3	-1.02	141	-3.97	0.95	4.08	-0.024	0.0070	0.020
$\hat{V}(\hat{\theta})$	8.9	0.39	9.3	-1.02	141						

PPS

Population B, PPS, n=174, 60 million repetitions.

Estimator	L	U	ER	RB.2	RRMSE	0.05%	τ	99.5%	0.05%	ν	99.5%
$\hat{V}^*(\hat{\theta}^*)$	25.4	0.02	25.4	-0.37	177	-4.59*	0.84*	4.89*	-0.009	0.0029	0.010
$\hat{V}_R(\hat{\theta}_\lambda)$	25.5	0.02	25.5	-0.42	176						
$\hat{V}_{AT}^*(\hat{\theta}_\lambda)$	25.5	0.02	25.5	-0.78	177						

Population B, PPS, n=90, 120 million repetitions.

Estimator	L	U	ER	RB.2	RRMSE	0.05%	τ	99.5%	0.05%	ν	99.5%
$\hat{V}^*(\hat{\theta}^*)$	19.8	0.05	19.9	-0.25	261	-4.48*	0.45*	4.77*	-0.013	0.0039	0.013
$\hat{V}_R(\hat{\theta}_\lambda)$	20.0	0.06	20.0	-0.25	258						
$\hat{V}_{AT}^*(\hat{\theta}_\lambda)$	19.9	0.05	19.9	0.90	265						

Population B, PPS, n=45, 240 million repetitions.

Estimator	L	U	ER	RB.2	RRMSE	0.05%	τ	99.5%	0.05%	ν	99.5%
$\hat{V}^*(\hat{\theta}^*)$	16.7	0.10	16.8	0.25	377	-4.57*	0.11*	5.01*	-0.020	0.0062	0.020
$\hat{V}_R(\hat{\theta}_\lambda)$	17.1	0.11	17.2	0.30	368						
$\hat{V}_{AT}^*(\hat{\theta}_\lambda)$	16.7	0.10	16.8	1.54	383						

Population B, PPS, n=20, 320 million repetitions.

Estimator	L	U	ER	RB.2	RRMSE	0.05%	τ	99.5%	0.05%	ν	99.5%
$\hat{V}^*(\hat{\theta}^*)$	15.6	0.21	15.9	0.10	571	-6.96*	0.05*	7.45*	-0.042	0.012	0.037
$\hat{V}_R(\hat{\theta}_\lambda)$	16.4	0.25	16.7	0.22	543						
$\hat{V}_{AT}^*(\hat{\theta}_\lambda)$	15.6	0.21	15.8	0.77	575						

* value has been subtracted with the mean.

Figure 15 shows the Relative Biases 2 in percentage for the three variance estimators in PPS, population A. The two designs are: $n=38$ and $n=20$. As we saw in the simulations with SI, the variance estimator with the estimated inclusion probabilities, is almost unbiased at the end of the simulations and we can state that, with these designs, Antal's and Tillé's variance estimator $\hat{V}_{AT}^*(\hat{\theta}_\lambda)$ has a positive bias and Rosén's variance estimator $V_R(\hat{\theta}_\lambda)$ has a negative bias. The bias of $V_R(\hat{\theta}_\lambda)$ was even larger when we looked at the performance of each of the H variance estimators, about -3% with $n=38$ and about -5% with $n=20$. Other simulations show the same pattern concerning the different biases.

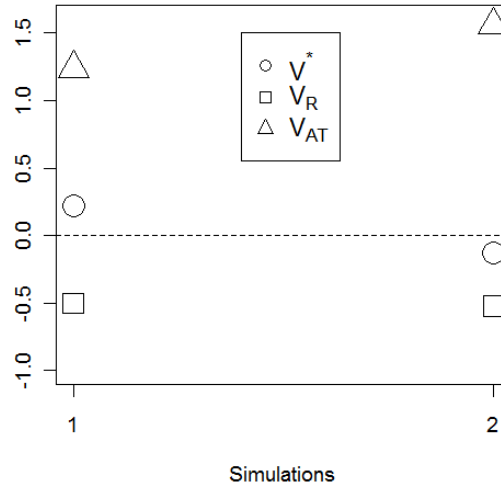


Figure 15: RB.2, population A , $n=38$ and $n=20$.

Figure 16 and 17 show the Relative Biases, RB.2 in percentage, for the three variance estimators in PPS, population B. Figure 16 shows the Relative Biases with the outlier and Figure 17 shows the Relative Biases without the outlier. The four designs are: $n=174$, $n=90$, $n=45$ and $n=20$. We can see that $V_R(\hat{\theta}_\lambda)$ is unbiased with these designs, both with and without the outlier, but the performance of each of the H variance estimators showed that $V_R(\hat{\theta}_\lambda)$ has a negative bias, about -0.6% with $n=174$, about -1% with $n=90$, about -2% with $n=45$ and about -5% with $n=20$. In Figure 16 we see that $\hat{V}_{AT}^*(\hat{\theta}_\lambda)$ has a positive bias with three designs of four, but in Figure 17 we see that, without the outlier, it is unbiased with two designs of four. Also here, other simulations show the same pattern concerning the different biases.

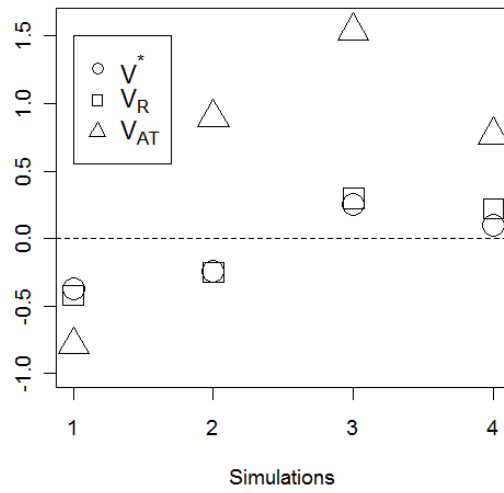


Figure 16: RB.2, pop. B, $n=174$, $n=90$, $n=45$ and $n=20$.

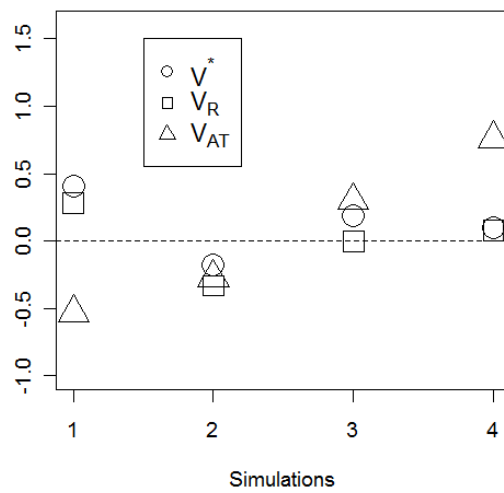


Figure 17: RB.2, population B, without the outlier.

4 Conclusions

The simulations in this essay show that we can get estimators based on estimated inclusion probabilities, that are very precise compared to the unbiased estimators. We also show that some of our measures are inversely proportional to the square root of the number of repetitions. This means that if we want to double the precision of the measures, we have to quadruple the number of repetitions etc. Then the precision is just a matter of time, the time it takes to simulate the inclusion probabilities with required precision.

The simulations with SI show that the estimator of the variance based on estimated inclusion probabilities, has the same performance as the unbiased estimator of the variance. Both the estimators have the same values in our measures L, U, ER and RRMSE, and their values in RB.2 are almost the same at the end of every simulation. If we look at the performance of every single one of the H variance estimators with estimated inclusion probabilities, we can conclude that the values of τ is very precise at the end of the simulations. The performance of every single one of the H estimators of the total with estimated inclusion probabilities, is even better, the values of v at the end of the simulations are all less than 0.02%.

In PPS we get better precision of v and the measures of the empirical distribution, compared to SI with the same sample size and same number of repetitions. The measure τ is not an exact measure in PPS, but it looks like τ behaves almost the same as in SI, at least in the beginning of the simulations. But the precision of the measures of the empirical distribution is less compared to SI with the same sample size and the same number of repetitions.

We can use the stabilization of the measures RB.1 and RB.2, to see how the variance and the estimated variance, with the estimated inclusion probabilities perform. The development of the measures L, U, ER and RRMSE, with increasing number of repetitions, do not say so much about the precision of the estimated variance with the estimated inclusion probabilities. The measures stabilize rather quickly and then remain the same during the simulations.

When comparing the three variance estimators in PPS, we can conclude that Rosén's variance estimator has a negative bias with population A (see page 38), a bias that we can see even clearer with the empirical distributions of the relative differences in τ . We can also conclude that Antal's and Tillé's bootstrap variance estimator has a positive bias with population A. Population B has an outlier and here the results vary for Antal's and Tillé's bootstrap variance estimator, if we include the outlier or not (see page 39). The variance estimator has a positive bias with three designs of four, with the outlier, and is unbiased with two designs of four, without the outlier. Rosén's variance estimator is unbiased, but has a negative bias if we look at the empirical distributions of the relative differences in τ . Rosén's variance for the whole population also has a negative bias in every design we have simulated, if we compare with $V_{sim}^*(\hat{\theta}_\lambda)$.

An interesting fact is that some of our measures are inversely proportional to the square root of the number of repetitions. We have not shown this relationship mathematically and this could be a task for future work.

In this essay we have shown the development of different measures based on the fact that we know the variable of interest, \mathbf{y} , for the whole population. In reality we only have one sample with the variable of interest known only for the elements in that sample. Fattorini's algorithm can be used to guarantee stability for the estimators and some more research could be done

with a modification of the algorithm where we vary the values of G and δ , and only look at the stability of the estimators.

If we are in a situation where we know the first order inclusion probabilities, we can use our measure φ_1 as a measure of precision. Then it would be a good idea to insert the known first order inclusion probabilities in our formulas and only use the estimated second order inclusion probabilities. It would also be interesting to do a simulation study where we compare the estimators based on estimated inclusion probabilities with the estimators based on the known first order inclusion probabilities and the estimated second order inclusion probabilities.

In SI it should also be possible to compute how many repetitions we need so that the probability that the estimated inclusion probabilities differs from the exact ones, is smaller than a specified value.

We have used relatively small populations in our study and we have showed that we can get very precise estimated inclusion probabilities within a relatively short time period. In the time of writing, we are examining how long it will take to simulate with much larger populations. We have access to a server and we are simulating with population sizes up to one hundred thousand.

5 References

- Aires, N. (1999). Algorithms to find exact inclusion probabilities for conditional Poisson sampling and Pareto π ps sampling. *Methodology and Computing in Applied Probability*. 4: 457-469.
- Aires, N. (2000). Comparisons between conditional Poisson sampling and Pareto π ps sampling designs. *Journal of Statistical Planning and Inference*. 88: 133-147.
- Aires, N. and Rosén, B. (2005). On inclusion probabilities and relative estimator bias for Pareto π ps sampling. *Journal of Statistical Planning and Inference*. 128: 543-567.
- Antal, E. and Tillé, Y. (2011). Simple random sampling with over-replacement. *Journal of Statistical Planning and Inference*. 141: 597-601.
- Antal, E. and Tillé, Y. (2011). A direct bootstrap method for complex sampling designs from a finite population. *Journal of the American Statistical Association*. Vol. 106, No. 494: 534-543.
- Bennet, G. (1962). Probability inequalities for the sum of independent random variables. *Journal of the American Statistical Association*. 57: 33-45.
- Fattorini, L. (2006). Applying the Horvitz-Thompson criterion in complex designs: A computer intensive perspective for estimating inclusion probabilities. *Biometrika* 93: 269-278.
- Fattorini, L. (2009). An adaptive algorithm for estimating inclusion probabilities and performing the Horvitz-Thompson criterion in complex designs. *Computational Statistics* 24: 623-639.
- Horvitz, D.G. and Thompson, D.J. (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*. 47: 663-685.
- Matei, A. and Tillé, Y. (2007). Computational aspects in order π ps sampling designs. *Computational Statistics and & Data Analysis*. 51: 3703-3717.
- Ng, P. M. and Donadio, M. (2006). Computing inclusion probabilities for order sampling. *Journal of Statistical Planning and Inference*. 136: 4026-4042.
- Ohlsson, E. (1990). Sequential poisson sampling from a business register and its application to the Swedish consumer price index. Statistics Sweden R&D report. 1990:6
- Ohlsson, E. (1995) Sequential Poisson Sampling. Research Report from Institute of Actuarial Mathematics and Mathematical Statistics at Stockholm University. Report No. 182.
- Rosén, B. (1997a). Asymptotic theory for order sampling. *Journal of Statistical Planning and Inference*. 62: 135-158.
- Rosén, B. (1997b). On sampling with probability proportional to size. *Journal of Statistical Planning and Inference*. 62: 159-191.

Rosén, B. (2000). On inclusion probabilities for order π ps sampling. *Journal of Statistical Planning and Inference*. 90: 117-143.

Saavedra, P. (1995). Fixed sample size PPS approximations with a permanent random number. *1995 Joint Statistical Meetings American Statistical Association*, Orlando, Florida.

Sen, A.R. (1953). On the estimate of the variance in sampling with varying probabilities. *Journal of the Indian Society of Agricultural Statistics*. 5: 119-127.

Särndal, C-E, Swensson, B. and Wretman, J. (1992). *Model assisted survey sampling*. Springer-Verlag, New York.

Yates, F. and Grundy, P.M. (1953). Selection without replacement from within strata with probability proportional to size. *Journal of the Royal Statistical Society*. B 15: 235-261.

6 Appendix

6.1 Summary of the formulas

$$\hat{\theta} = \sum_{k \in S} \frac{y_k}{\pi_k} \quad (2.1)$$

$$V(\hat{\theta}) = -\frac{1}{2} \sum_{k \in U} \sum_{l \in U} \left(\frac{y_k}{\pi_k} - \frac{y_l}{\pi_l} \right)^2 (\pi_{kl} - \pi_k \pi_l) \quad (2.2)$$

$$\hat{V}(\hat{\theta}) = -\frac{1}{2} \sum_{k \in S} \sum_{l \in S} \left(\frac{y_k}{\pi_k} - \frac{y_l}{\pi_l} \right)^2 \frac{(\pi_{kl} - \pi_k \pi_l)}{\pi_{kl}} \quad (2.3)$$

$$V(\hat{\theta}) = N^2 \frac{1-f}{n} S_{yU}^2 \quad (2.4)$$

$$\hat{V}(\hat{\theta}) = N^2 \frac{1-f}{n} S_{ys}^2 \quad (2.5)$$

$$\hat{\theta}_\lambda = \sum_{k \in S} \frac{y_k}{\lambda_k} \quad (2.6)$$

$$V_R(\hat{\theta}_\lambda) = \frac{N}{N-1} \left\{ \sum_{k \in U} \frac{y_k^2}{\lambda_k} (1 - \lambda_k) - \frac{[\sum_{k \in U} y_k (1 - \lambda_k)]^2}{n - \sum_{k \in U} \lambda_k^2} \right\} \quad (2.7)$$

$$\hat{V}_R(\hat{\theta}_\lambda) = \frac{n}{n-1} \sum_{k \in S} \left(\frac{y_k}{\lambda_k} - \frac{\sum_{l \in S} y_l (1 - \lambda_l) / \lambda_l}{\sum_{l \in S} 1 - \lambda_l} \right)^2 (1 - \lambda_k) \quad (2.8)$$

$$\hat{\theta}^* = \sum_{k \in S} \frac{y_k}{\tilde{\pi}_k} \quad (3.1)$$

$$V^*(\hat{\theta}^*) = -\frac{1}{2} \sum_{k \in U} \sum_{l \in U} \left(\frac{y_k}{\tilde{\pi}_k} - \frac{y_l}{\tilde{\pi}_l} \right)^2 (\tilde{\pi}_{kl} - \tilde{\pi}_k \tilde{\pi}_l) \quad (3.2)$$

$$\hat{V}^*(\hat{\theta}^*) = -\frac{1}{2} \sum_{k \in S} \sum_{l \in S} \left(\frac{y_k}{\tilde{\pi}_k} - \frac{y_l}{\tilde{\pi}_l} \right)^2 \frac{\tilde{\pi}_{kl} - \tilde{\pi}_k \tilde{\pi}_l}{\tilde{\pi}_{kl}} \quad (3.3)$$

$$\varphi_1 = \begin{cases} \frac{1}{N} \sum_{k \in U} \left| \frac{\tilde{\pi}_k - \pi_k}{\pi_k} \right| & \text{in SI} \\ \frac{1}{N} \sum_{k \in U} \left| \frac{\tilde{\pi}_k - \lambda_k}{\lambda_k} \right| & \text{in PPS} \end{cases}$$

$$\psi_1 = \begin{cases} \max \left| \frac{\tilde{\pi}_k - \pi_k}{\pi_k} \right| & \text{in SI} \\ \max \left| \frac{\tilde{\pi}_k - \lambda_k}{\lambda_k} \right| & \text{in PPS} \end{cases}$$

$$\varphi_2 = \frac{1}{\#\pi_{kl}} \sum_{k \neq l} \sum_U \left| \frac{\tilde{\pi}_{kl} - \pi_{kl}}{\pi_{kl}} \right|$$

$$\psi_2 = \max \left| \frac{\tilde{\pi}_{kl} - \pi_{kl}}{\pi_{kl}} \right|$$

$$RB. 1 = \begin{cases} \frac{V^*(\hat{\theta}^*) - V(\hat{\theta})}{V(\hat{\theta})} & \text{in SI} \\ \frac{V^*(\hat{\theta}^*) - V_{sim}^*(\hat{\theta}_\lambda)}{V_{sim}^*(\hat{\theta}_\lambda)} & \text{in PPS} \end{cases}$$

$$L = \frac{1}{H} \sum_{i=1}^H I \left[\hat{\theta}_i + 1.96 * \sqrt{\hat{V}_i} < \theta \right]$$

$$U = \frac{1}{H} \sum_{i=1}^H I \left[\hat{\theta}_i - 1.96 * \sqrt{\hat{V}_i} > \theta \right]$$

$$RB.2 = \frac{\bar{\hat{V}} - V}{V} = \frac{B}{V}$$

where $\bar{\hat{V}}$ is the mean of the variance estimates for the H samples compared to the real variance V . In SI we compute V from (4) and in PPS we use the variance of the H samples as V ,

$$V = S_{\hat{\theta}}^2 = \frac{1}{H-1} \sum_{i=1}^H (\hat{\theta}_i - \bar{\hat{\theta}})^2$$

$$RRMSE = \frac{\sqrt{B^2 + S_{\hat{V}}^2}}{V}$$

here $S_{\hat{V}}^2$ is the variance of the variance estimators for the samples, $S_{\hat{V}}^2 = \frac{1}{H-1} \sum_{i=1}^H (\hat{V}_i - \bar{\hat{V}})^2$

$$\tau = \begin{cases} \frac{1}{H} \sum_{i=1}^H \left| \frac{\hat{V}_{(i)}^*(\hat{\theta}^*) - \hat{V}_i(\hat{\theta})}{\hat{V}_i(\hat{\theta})} \right| & \text{in SI} \\ \frac{1}{H} \sum_{i=1}^H \left| \frac{\hat{V}_{(i)}^*(\hat{\theta}^*) - \hat{V}_{R(i)}(\hat{\theta}_{\lambda})}{\hat{V}_{R(i)}(\hat{\theta}_{\lambda})} \right| & \text{in PPS} \end{cases}$$

$$\nu = \begin{cases} \frac{1}{H} \sum_{i=1}^H \left| \frac{\hat{\theta}_{(i)}^* - \hat{\theta}_i}{\hat{\theta}_i} \right| & \text{in SI} \\ \frac{1}{H} \sum_{i=1}^H \left| \frac{\hat{\theta}_{(i)}^* - \hat{\theta}_{\lambda(i)}}{\hat{\theta}_{\lambda(i)}} \right| & \text{in PPS} \end{cases}$$

6.2 Figures and Tables referred to in the text

6.2.1 SI simulation with population A, $n=38$

Table A1: SI simulation from population A, $n=38$.

Repetitions	$\varphi_1\%$	$\psi_1\%$	$\varphi_2\%$	$\psi_2\%$	RB.1 %
125,000	0.63	2.6	1.93	11.6	-1.23
250,000	0.416	1.66	1.35	7.47	-0.654
500,000	0.294	0.987	0.953	5.18	-0.434
1,000,000	0.223	0.945	0.687	3.66	-0.0942
2,000,000	0.157	0.614	0.484	2.71	0.109
3,000,000	0.129	0.538	0.397	2.05	0.0172
4,000,000	0.11	0.477	0.345	1.84	0.0366
5,000,000	0.0993	0.416	0.308	1.67	0.0171
10,000,000	0.0654	0.248	0.214	1.11	0.00735
15,000,000	0.0553	0.231	0.176	1.01	-0.019
20,000,000	0.0468	0.203	0.151	0.882	-0.0201

Table A2: Development of the empirical distribution measures for $\hat{V}^*(\hat{\theta}^*)$.

Repetitions	L %	U %	ER %	RB.2 %	RRMSE %
125,000	9.77	0.589	10.4	0.554	50.5
250,000	9.76	0.584	10.3	0.0212	49.6
500,000	9.76	0.585	10.3	-0.274	49.2
1,000,000	9.75	0.581	10.3	-0.121	49.3
2,000,000	9.74	0.581	10.3	-0.04	49.2
3,000,000	9.76	0.586	10.3	-0.182	49.1
4,000,000	9.76	0.579	10.3	-0.171	49.1
5,000,000	9.75	0.582	10.3	-0.223	49
10,000,000	9.74	0.58	10.3	-0.277	48.9
15,000,000	9.74	0.578	10.3	-0.307	48.9
20,000,000	9.73	0.579	10.3	-0.314	48.9

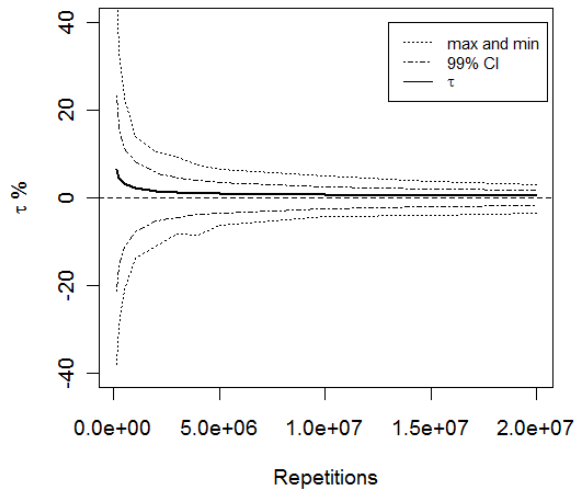


Figure A1: τ

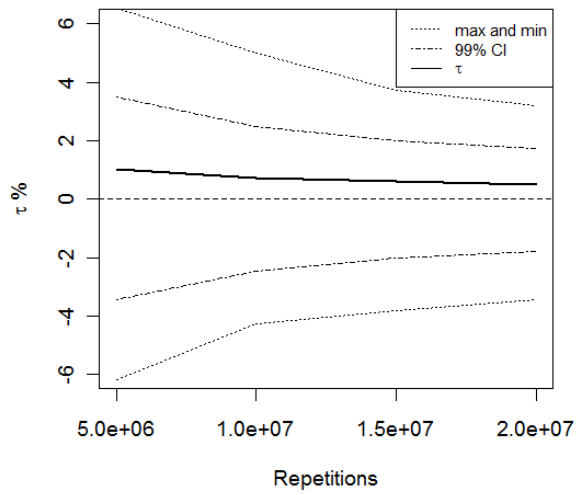


Figure A2: τ , 5-20 million repetitions.

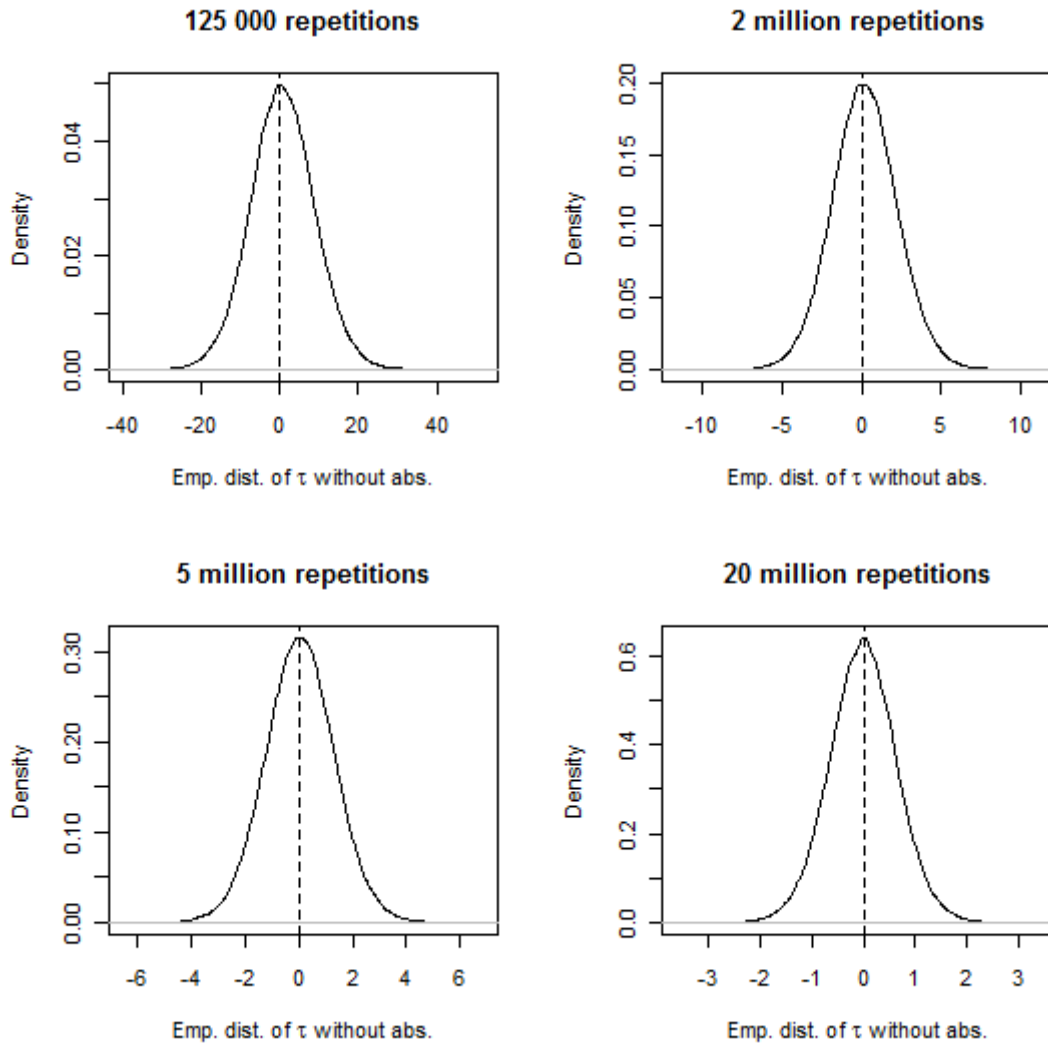


Figure A3: Development of the empirical distribution of the relative differences (τ).

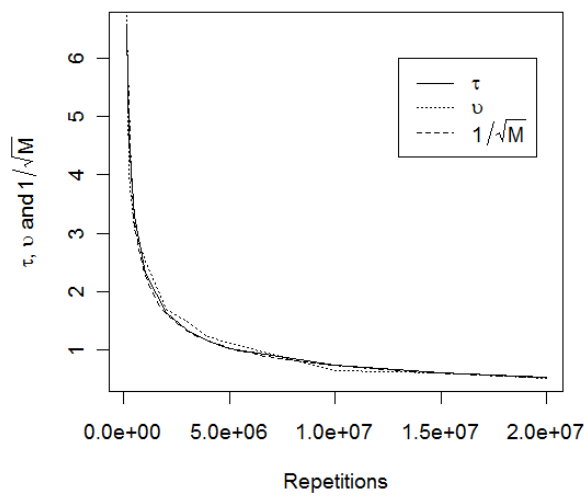


Figure A4: τ , v and $\frac{1}{\sqrt{M}}$.

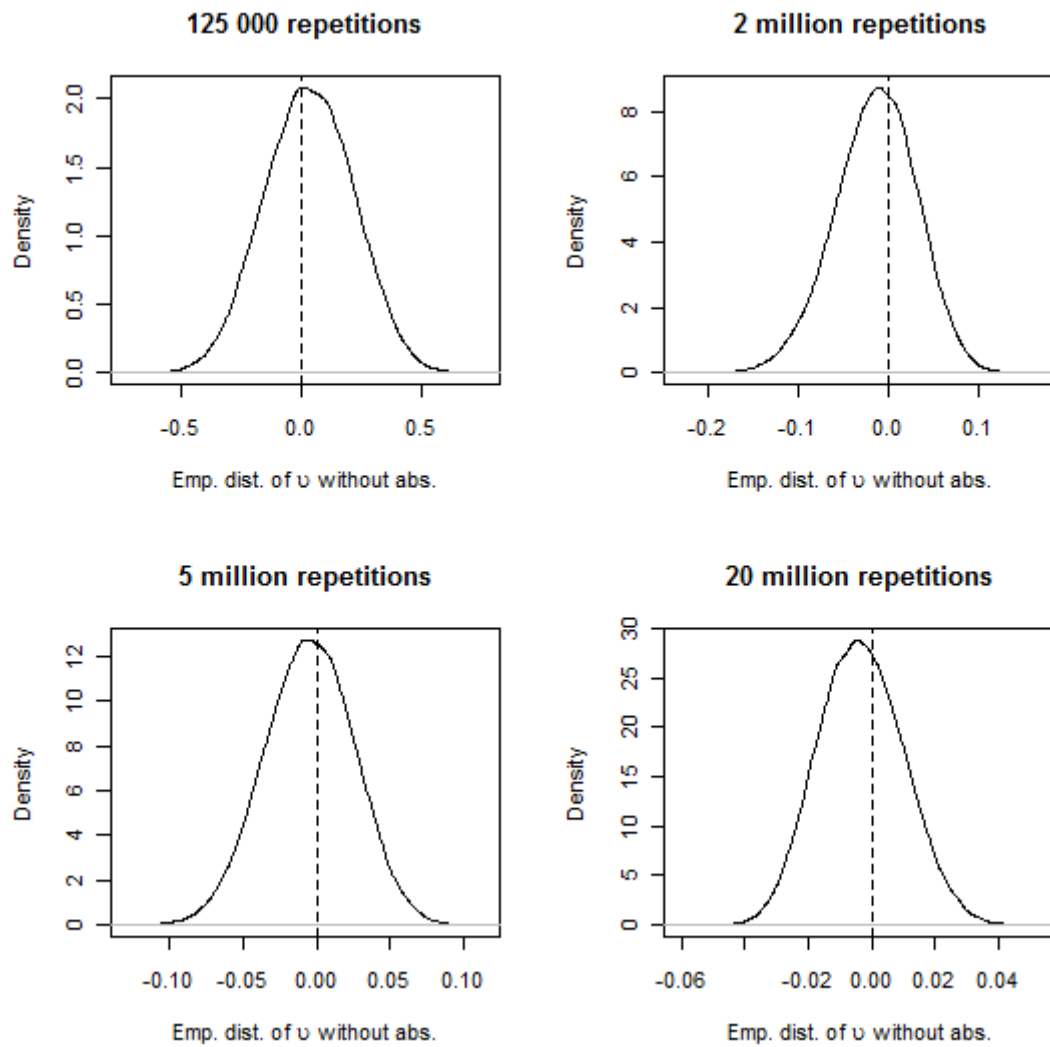


Figure A5: Development of the empirical distribution of the relative differences (v).

6.2.2 PPS simulation with population A, n=38

Table A3: PPS simulation with population A, n=38.

Repetitions	φ_1 %	ψ_1 %	RB.1 %
125,000	0.937	5.53	-4.02
250,000	0.657	3.96	-2.47
500,000	0.459	2.76	-1.47
1,000,000	0.352	1.77	-0.508
2,000,000	0.24	1.11	-0.256
3,000,000	0.185	1.11	-0.21

4,000,000	0.16	0.911	-0.167
5,000,000	0.144	0.862	-0.0823
10,000,000	0.105	0.639	-0.0999
15,000,000	0.0862	0.492	0.00833
20,000,000	0.0747	0.428	0.0408
25,000,000	0.0696	0.365	0.0252
30,000,000	0.0605	0.291	0.0418
35,000,000	0.0548	0.316	0.0201
40,000,000	0.0529	0.29	0.0273
50,000,000	0.0471	0.269	0.0381

Table A4: Development of the empirical distribution measures for $\hat{V}^*(\hat{\theta}^*)$.

Repetitions	L %	U %	ER %	RB.2 %	RRMSE %
125,000	4.75	1.65	6.4	-0.0341	66.5
250,000	4.74	1.62	6.36	-0.676	62.7
500,000	4.69	1.6	6.29	-0.577	62.7
1,000,000	4.69	1.58	6.26	0.0904	64.2
2,000,000	4.68	1.57	6.25	0.137	64
3,000,000	4.69	1.57	6.26	0.0945	63.8
4,000,000	4.68	1.57	6.25	0.0714	63.5
5,000,000	4.68	1.57	6.24	0.148	63.7
10,000,000	4.68	1.57	6.26	0.0901	63.8
15,000,000	4.67	1.57	6.24	0.194	63.9
20,000,000	4.67	1.57	6.25	0.232	64
25,000,000	4.67	1.58	6.24	0.209	64
30,000,000	4.66	1.57	6.23	0.23	64
35,000,000	4.67	1.58	6.24	0.201	64
40,000,000	4.67	1.57	6.24	0.209	64.1
50,000,000	4.67	1.57	6.24	0.217	64.1

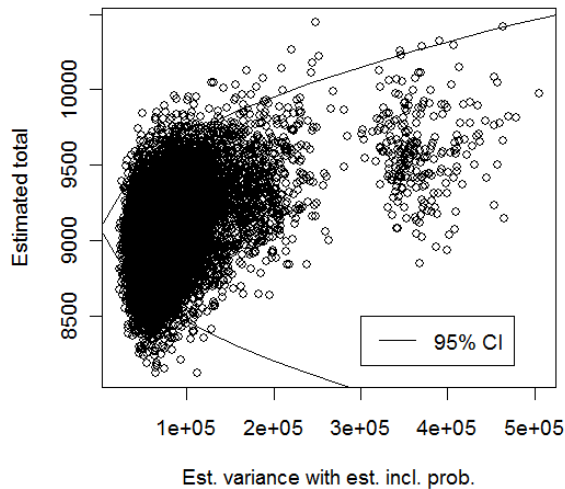


Figure A6: The first 10,000 of $\hat{V}^*(\hat{\theta}^*)$.

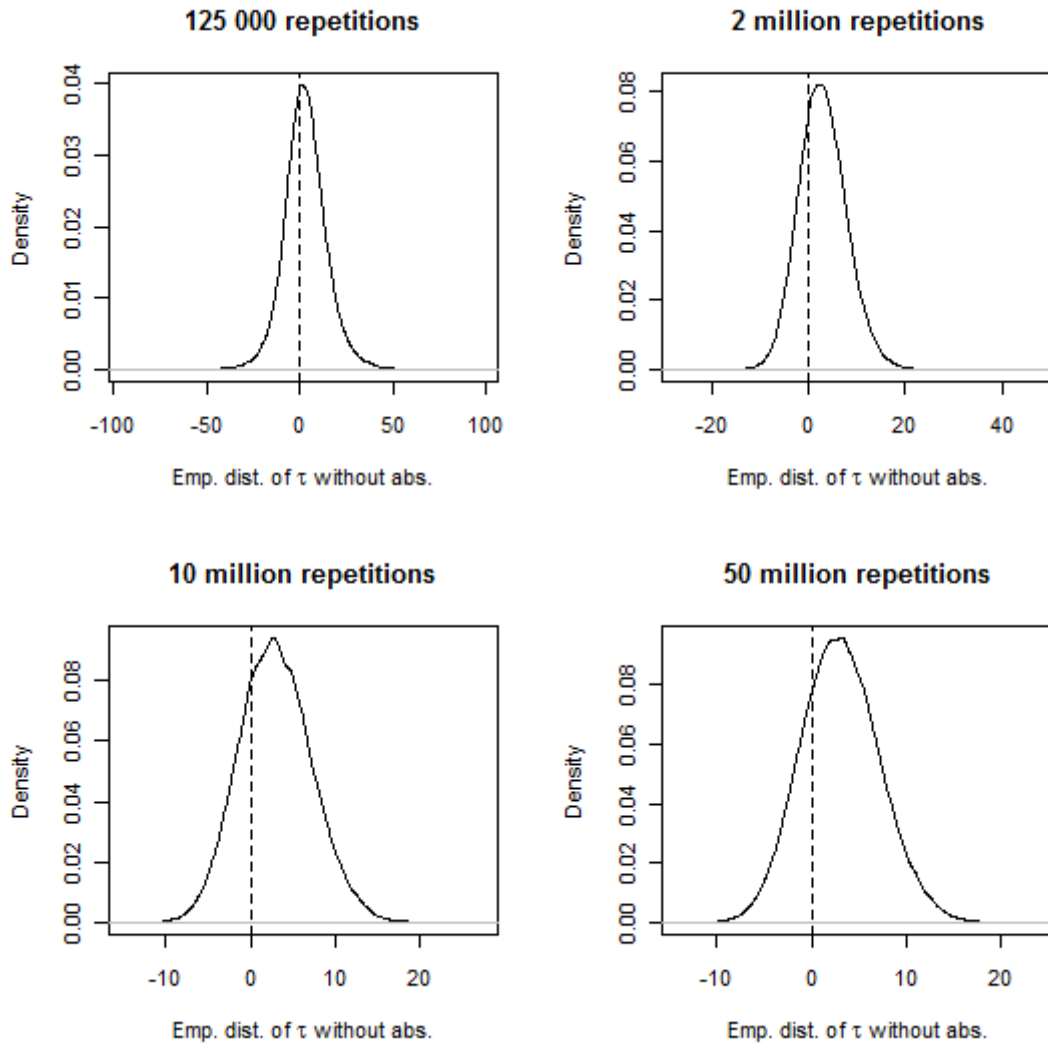


Figure A7: Development of the empirical distribution of the relative differences (τ).

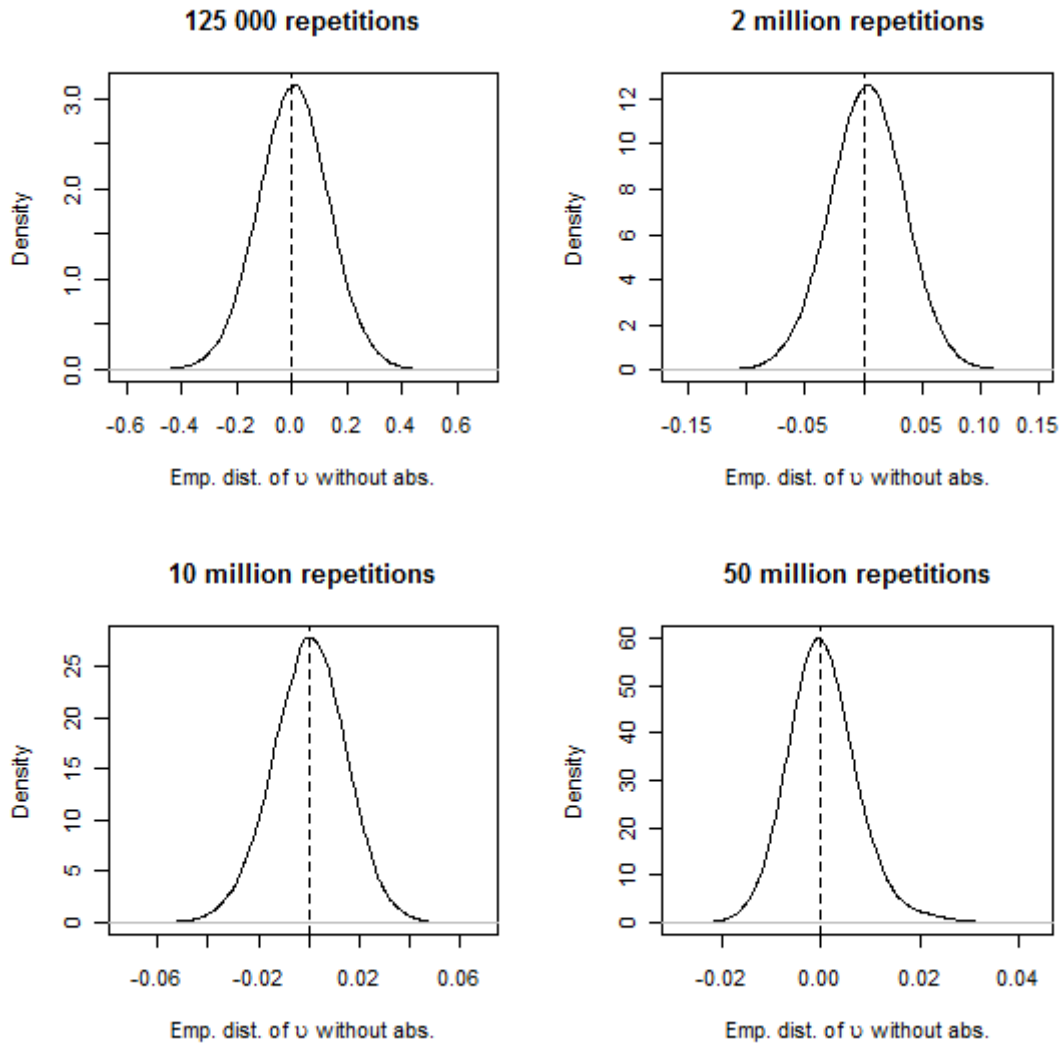


Figure A8: Development of the empirical distribution of the relative differences (v).

6.2.3 PPS simulation with population B, $n=174$

Table A5: PPS simulation with population B, $n=174$.

Repetitions	φ_1 %	ψ_1 %	RB.1 %
125,000	1.04	6.45	-12.4
250,000	0.733	4.13	-6.36
500,000	0.522	3.41	-2.94
1,000,000	0.375	2.09	-1.72
2,000,000	0.265	1.46	-0.954
3,000,000	0.209	1.21	-0.76
4,000,000	0.18	0.986	-0.565

5,000,000	0.162	0.863	-0.495
10,000,000	0.117	0.625	-0.234
15,000,000	0.0957	0.541	-0.227
20,000,000	0.0828	0.459	-0.217
25,000,000	0.0738	0.439	-0.157
30,000,000	0.0681	0.418	-0.133
35,000,000	0.0633	0.35	-0.0975
40,000,000	0.0594	0.351	-0.078
50,000,000	0.0528	0.314	-0.0608
60,000,000	0.0481	0.237	-0.0296

Table A6: Development of the empirical distribution measures for $\hat{V}^*(\hat{\theta}^*)$.

Repetitions	L %	U %	ER %	RB.2 %	RRMSE %
125,000	25.9	0.179	26.1	0.202	190
250,000	25.6	0.048	25.7	-0.323	183
500,000	25.4	0.019	25.4	-0.00037	181
1,000,000	25.4	0.021	25.4	-0.524	178
2,000,000	25.4	0.02	25.4	-0.672	177
3,000,000	25.3	0.02	25.4	-0.756	177
4,000,000	25.4	0.019	25.4	-0.65	177
5,000,000	25.4	0.021	25.4	-0.651	177
10,000,000	25.4	0.02	25.4	-0.478	177
15,000,000	25.4	0.019	25.4	-0.564	177
20,000,000	25.4	0.02	25.4	-0.58	177
25,000,000	25.4	0.02	25.4	-0.506	177
30,000,000	25.4	0.02	25.4	-0.492	177
35,000,000	25.4	0.02	25.4	-0.454	177
40,000,000	25.4	0.02	25.4	-0.432	177
50,000,000	25.4	0.02	25.4	-0.41	177
60,000,000	25.4	0.02	25.4	-0.373	177

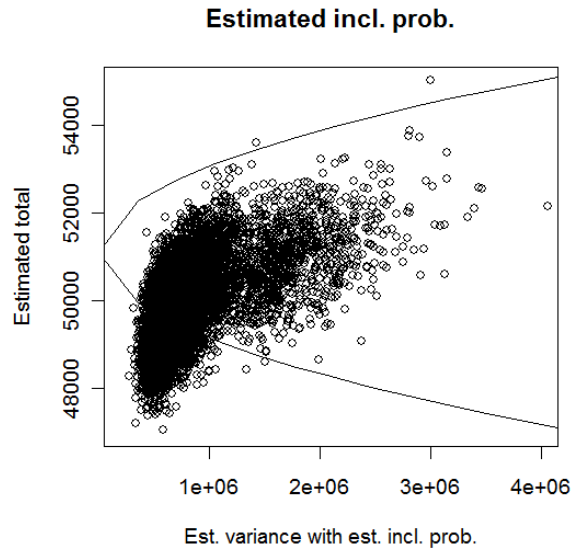


Figure A9: First 1000 observations, “main” group.

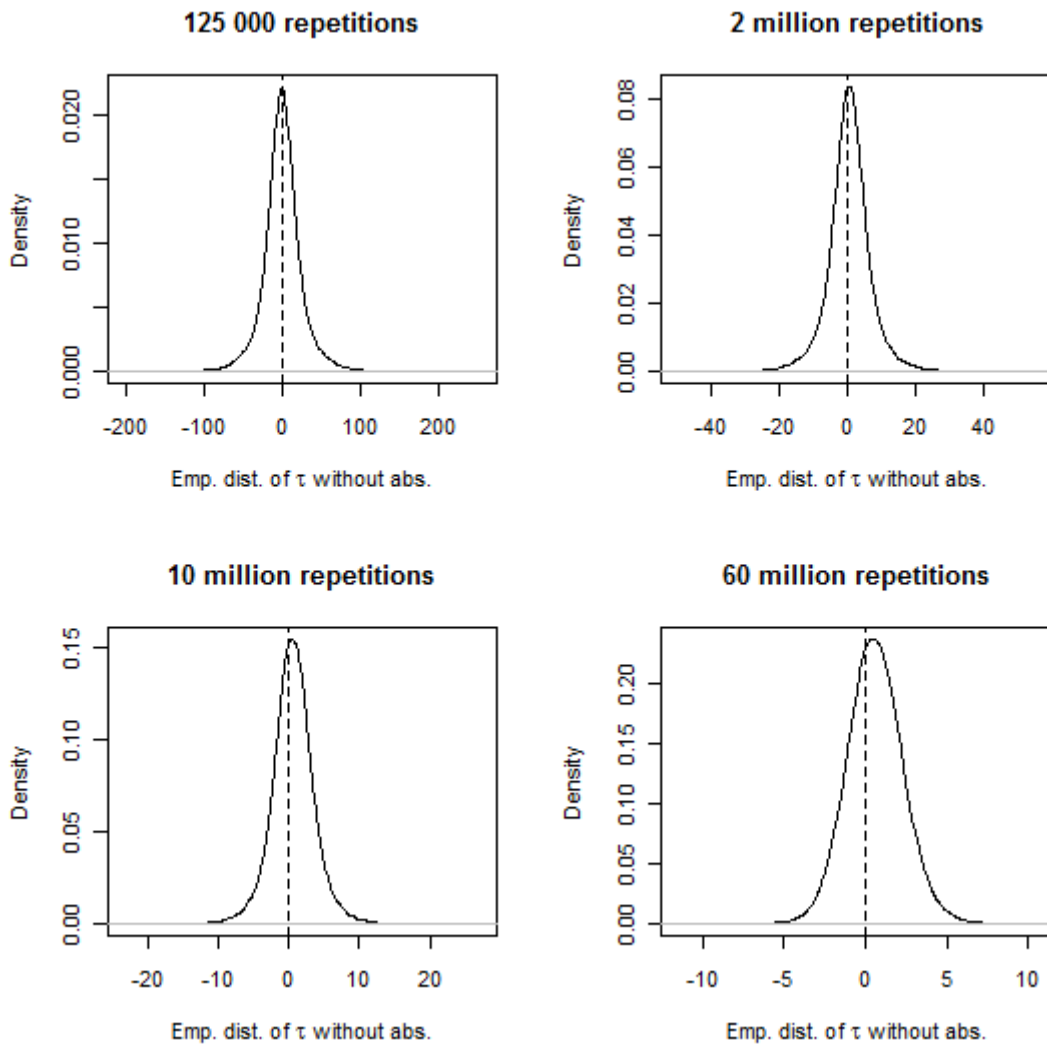


Figure A10: Development of the empirical distribution of the relative differences (τ).

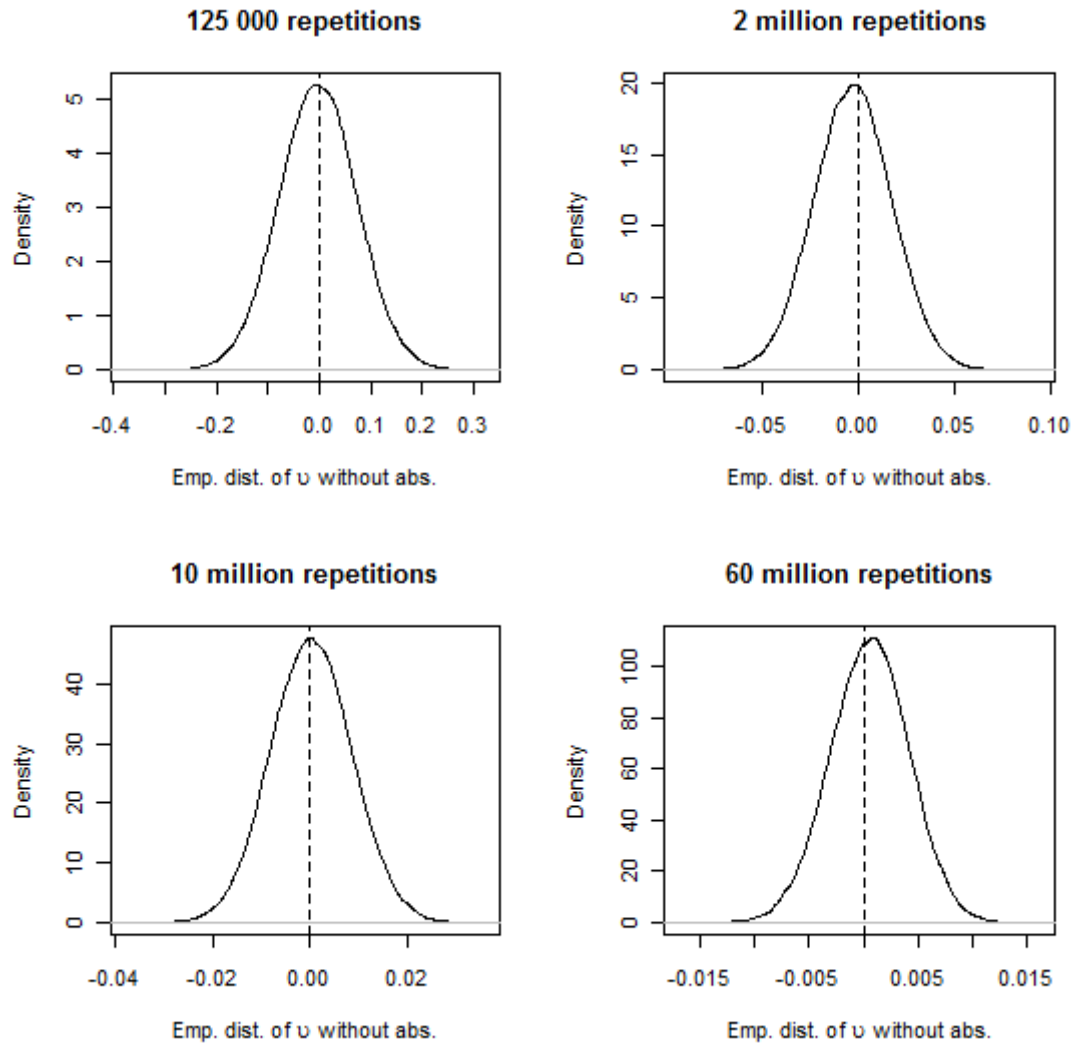


Figure A11: Development of the empirical distribution of the relative differences (v).

6.3 SI simulation with population A, $n=100$

Table: SI simulation from population A, $n=100$.

Repetitions	$\varphi_1\%$	$\psi_1\%$	$\varphi_2\%$	$\psi_2\%$	RB.1 %
125,000	0.34	1.24	0.697	3.63	-0.808
250,000	0.237	0.85	0.488	2.53	-0.419
500,000	0.166	0.634	0.348	1.78	-0.314
1,000,000	0.118	0.491	0.244	1.36	0.0777
2,000,000	0.079	0.332	0.17	0.867	0.149
3,000,000	0.0689	0.264	0.141	0.796	0.0779
4,000,000	0.0596	0.242	0.123	0.654	0.0609
5,000,000	0.0523	0.189	0.109	0.553	0.0183
10,000,000	0.0355	0.11	0.0754	0.419	0.0274

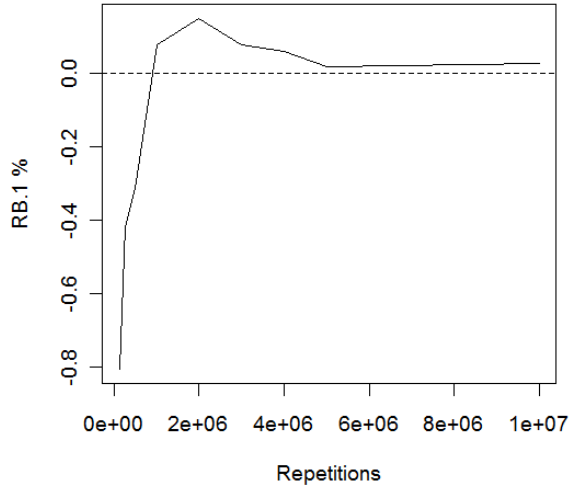


Figure: RB.1

Table: Development of the empirical distribution measures for $\hat{V}^*(\hat{\theta}^*)$.

Repetitions	L %	U %	ER %	RB.2 %	RRMSE %
125,000	5.63	0.951	6.58	-0.0864	26.5
250,000	5.62	0.952	6.57	0.00526	26.5
500,000	5.63	0.945	6.57	0.00787	26.5
1,000,000	5.61	0.93	6.54	0.341	26.6
2,000,000	5.61	0.935	6.54	0.372	26.6
3,000,000	5.61	0.939	6.55	0.276	26.6
4,000,000	5.61	0.936	6.55	0.248	26.6
5,000,000	5.62	0.936	6.55	0.203	26.6
10,000,000	5.61	0.937	6.55	0.199	26.5

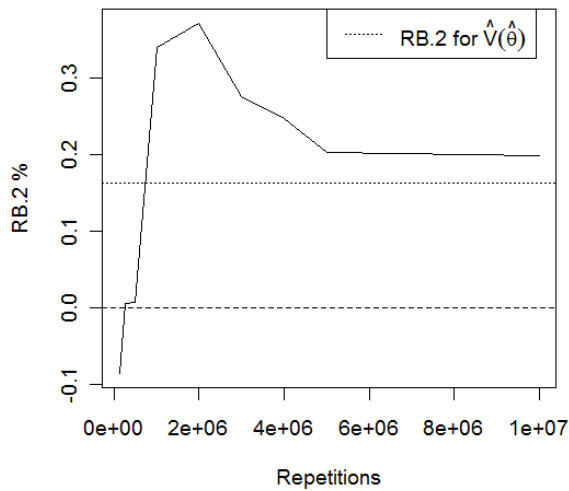


Figure: RB.2 for $\hat{V}^*(\hat{\theta}^*)$.

Table: Development of τ .

Repetitions	Min %	P 0.5 %	Mean %	τ %	P 99.5 %	Max %
125,000	-16.1	-8.02	-0.211	2.32	7.35	12
250,000	-10.9	-5.69	-0.123	1.65	5.23	8.92
500,000	-6.98	-3.96	-0.148	1.15	3.62	6.42
1,000,000	-5.09	-2.61	0.184	0.84	2.85	5.17
2,000,000	-3.34	-1.74	0.2	0.602	2.11	3.42
3,000,000	-2.57	-1.48	0.11	0.48	1.61	2.59
4,000,000	-2.29	-1.24	0.0803	0.4	1.36	2.19
5,000,000	-2.07	-1.15	0.0357	0.358	1.2	2.1
10,000,000	-1.42	-0.797	0.0351	0.251	0.852	1.53

Table: Development of v .

Repetitions	Min %	P 0.5 %	Mean %	v %	P 99.5 %	Max %
125,000	-0.209	-0.134	-0.0115	0.0399	0.112	0.189
250,000	-0.148	-0.0895	-0.000349	0.0285	0.0905	0.138
500,000	-0.0988	-0.0544	0.00597	0.0194	0.0667	0.103
1,000,000	-0.0585	-0.0318	0.012	0.0168	0.0544	0.0865
2,000,000	-0.0433	-0.027	0.00688	0.012	0.0406	0.0616
3,000,000	-0.0369	-0.0212	0.00493	0.00903	0.0305	0.0495
4,000,000	-0.0414	-0.0201	0.0043	0.00849	0.0291	0.0445
5,000,000	-0.0298	-0.0165	0.00563	0.00839	0.0274	0.0378
10,000,000	-0.0228	-0.0138	0.00206	0.00521	0.017	0.0258

Population A, SI, $n=100$, 10 million repetitions.

Estimator	L	U	ER	RB.2	RRMSE	0.05%	τ	99.5%	0.05%	v	99.5%
$\hat{V}^*(\hat{\theta}^*)$	5.6	0.94	6.6	0.20	26.6	-0.80	0.25	0.85	-0.014	0.0052	0.017
$\hat{V}(\hat{\theta})$	5.6	0.94	6.6	0.16	26.5						

It takes 1 hour, 36 minutes and 52 seconds to produce a matrix with the inclusion probabilities, based on 10 million repetitions, for this design.

6.4 PPS simulation with population A, $n=20$

Rosén's variance for the whole population has a negative Relative Bias compared to $V_{sim}^*(\hat{\theta}_\lambda)$, -0.2%.

Table: PPS simulation with population A, $n=20$.

Repetitions	φ_1 %	ψ_1 %	RB.1 %
125,000	1.31	8.82	-6.79
250,000	0.924	5.73	-4.09
500,000	0.674	4.8	-2.43

1,000,000	0.462	2.71	-1.11
2,000,000	0.326	1.7	-0.561
3,000,000	0.253	1.46	-0.318
4,000,000	0.226	1.22	-0.146
5,000,000	0.203	1.16	-0.114
10,000,000	0.141	0.716	0.0359
15,000,000	0.115	0.641	0.0132
20,000,000	0.103	0.462	0.0123
25,000,000	0.0926	0.441	-0.0035
30,000,000	0.0829	0.36	-0.00765
35,000,000	0.076	0.271	-0.00885
40,000,000	0.0707	0.261	-0.0242
50,000,000	0.065	0.285	-0.0244
60,000,000	0.0611	0.269	-0.0106
70,000,000	0.0564	0.291	-0.0393
80,000,000	0.0549	0.256	-0.0463
100,000,000	0.0479	0.232	-0.0479

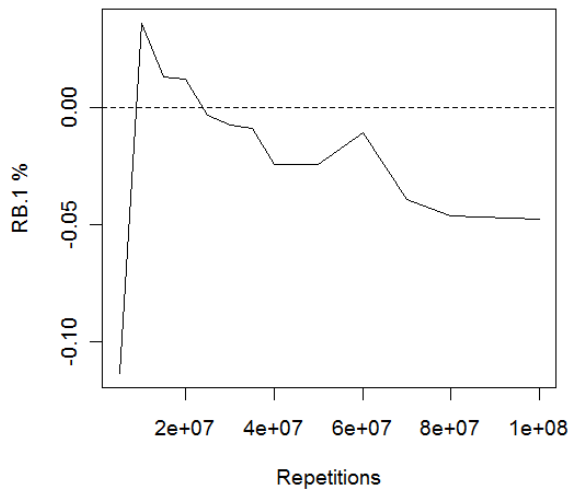


Figure: RB.1, 5-100 million repetitions.

The Figures above show RB.1 from five million repetitions, which gives a more precise presentation of the development of the measure in the end of the simulation.

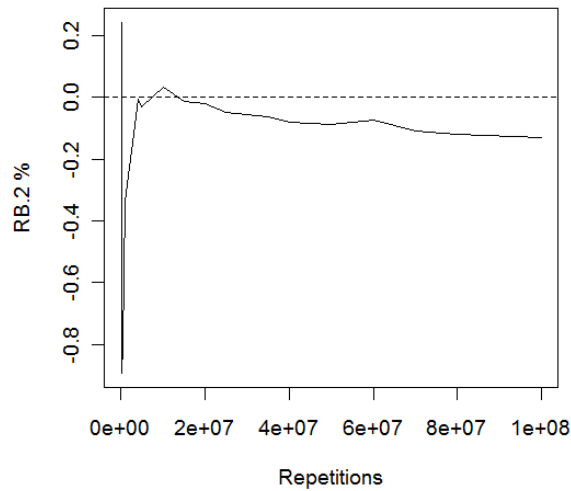


Figure: RB.2 for $\hat{V}^*(\hat{\theta}^*)$.

Table: Development of the empirical distribution measures for $\hat{V}^*(\hat{\theta}^*)$.

Repetitions	L %	U %	ER %	RB.2 %	RRMSE %
125,000	5.19	1.95	7.14	0.242	87.9
250,000	5.19	1.84	7.03	-0.895	80.8
500,000	5.18	1.77	6.95	-0.803	80.5
1,000,000	5.11	1.74	6.85	-0.339	79.1
2,000,000	5.09	1.73	6.82	-0.226	79.5
3,000,000	5.06	1.73	6.79	-0.12	79.7
4,000,000	5.06	1.73	6.79	-0.00453	79.9
5,000,000	5.06	1.74	6.8	-0.0307	79.6
10,000,000	5.05	1.71	6.77	0.0336	79.6
15,000,000	5.05	1.72	6.77	-0.0119	79.8
20,000,000	5.05	1.72	6.77	-0.0189	79.7
25,000,000	5.06	1.71	6.77	-0.0498	79.7
30,000,000	5.06	1.71	6.77	-0.0563	79.7
35,000,000	5.06	1.71	6.77	-0.064	79.7
40,000,000	5.06	1.71	6.78	-0.0822	79.8
50,000,000	5.05	1.72	6.77	-0.0886	79.8
60,000,000	5.06	1.72	6.78	-0.0733	79.9
70,000,000	5.06	1.72	6.78	-0.109	79.8
80,000,000	5.05	1.72	6.77	-0.118	79.8
100,000,000	5.05	1.72	6.77	-0.129	79.7

Table A7: Development of τ .

Repetitions	Min %	P 0.5 %	Mean %	τ %	P 99.5 %	Max %
125,000	-119	-47.3	5.84	13.6	68.5	164
250,000	-108	-32.8	4.85	9.92	49	122
500,000	-61.6	-22.1	4.69	7.72	36.5	94.4
1,000,000	-46.5	-14	5.18	6.72	28	65.6
2,000,000	-30.3	-9.26	5.29	6.1	22.6	48.9
3,000,000	-24.2	-7.15	5.34	5.92	20.7	41.7
4,000,000	-21.1	-6.01	5.44	5.89	19.4	37.9
5,000,000	-19.7	-5.44	5.43	5.82	18.7	33.8
10,000,000	-12.6	-4.24	5.5	5.76	16.7	25.3
15,000,000	-9.72	-3.92	5.43	5.67	16	29.1
20,000,000	-8.86	-3.79	5.43	5.65	15.6	21.8
25,000,000	-8.4	-3.72	5.4	5.61	15.3	21.7
30,000,000	-8.58	-3.71	5.39	5.59	15.2	21
35,000,000	-8.82	-3.67	5.39	5.59	15	21.5
40,000,000	-8.52	-3.62	5.37	5.57	14.9	21.4
50,000,000	-8.66	-3.6	5.36	5.55	14.7	20.1
60,000,000	-8.33	-3.57	5.37	5.56	14.6	20.4
70,000,000	-8.35	-3.56	5.34	5.53	14.6	20.1
80,000,000	-8.24	-3.54	5.34	5.53	14.5	19.6
100,000,000	-8.33	-3.5	5.34	5.53	14.4	19.6

Table: Development of v .

Repetitions	Min %	P 0.5 %	Mean %	v %	P 99.5 %	Max %
125,000	-1.25	-0.706	-0.012	0.199	0.667	1.29
250,000	-1.24	-0.599	-0.0253	0.15	0.472	0.835
500,000	-0.788	-0.408	-0.0166	0.104	0.314	0.65
1,000,000	-0.528	-0.328	-0.00669	0.0723	0.234	0.415
2,000,000	-0.309	-0.178	-0.00264	0.0488	0.164	0.303
3,000,000	-0.268	-0.139	-0.00107	0.0388	0.135	0.244
4,000,000	-0.214	-0.113	0.00143	0.0342	0.119	0.24
5,000,000	-0.241	-0.115	0.00113	0.0316	0.112	0.214
10,000,000	-0.159	-0.0788	0.00182	0.0228	0.0786	0.139
15,000,000	-0.102	-0.0589	0.000438	0.0181	0.0723	0.127
20,000,000	-0.0956	-0.0572	0.000208	0.0164	0.0606	0.104
25,000,000	-0.0901	-0.0498	-0.000409	0.0148	0.0552	0.0959
30,000,000	-0.0885	-0.0462	-0.00046	0.0133	0.0467	0.0852
35,000,000	-0.0819	-0.0443	-0.000699	0.0121	0.0413	0.0784
40,000,000	-0.0581	-0.0363	-0.000918	0.011	0.0384	0.0659
50,000,000	-0.0527	-0.0319	-0.000741	0.00977	0.0344	0.0613
60,000,000	-0.0467	-0.0295	-0.000376	0.00907	0.0319	0.0578
70,000,000	-0.0475	-0.0287	-0.000823	0.00866	0.0294	0.0564
80,000,000	-0.0442	-0.0274	-0.000734	0.00833	0.0287	0.0535
100,000,000	-0.0403	-0.0245	-0.000684	0.00741	0.0256	0.0516

Population A, PPS, $n=20$, 100 million repetitions.

Estimator	L	U	ER	RB.2	RRMSE	0.05%	τ	99.5%	0.05%	v	99.5%
$\hat{V}^*(\hat{\theta}^*)$	5.05	1.7	6.8	-0.13	79.7	-8.8*	0.19*	9.1*	-0.025	0.0074	0.026
$\hat{V}_R(\hat{\theta}_\lambda)$	5.54	1.9	7.5	-0.52	74.9						
$\hat{V}_{AT}^*(\hat{\theta}_\lambda)$	5.04	1.7	6.7	1.56	86.9						

* value has been subtracted with the mean.

It takes 1 hour, 11 minutes and 30 seconds to produce a matrix with the inclusion probabilities, based on 100 million repetitions, for this design.

6.5 SI simulation with population B, $n=174$

Table: SI simulation from population B, $n=174$.

Repetitions	$\varphi_1\%$	$\psi_1\%$	$\varphi_2\%$	$\psi_2\%$	RB.1 %
125,000	0.774	3.08	2.97	19.3	-21.5
250,000	0.556	2.59	2.1	13.9	-10.6
500,000	0.389	1.78	1.48	9.71	-5.18
1,000,000	0.276	1.29	1.05	6.55	-2.75
2,000,000	0.196	0.992	0.743	4.83	-1.47
3,000,000	0.158	0.763	0.606	3.79	-0.962
4,000,000	0.136	0.701	0.524	3.37	-0.694
5,000,000	0.123	0.583	0.47	2.91	-0.492
10,000,000	0.0872	0.374	0.332	2.09	-0.297
15,000,000	0.0732	0.323	0.272	1.85	-0.162
20,000,000	0.0628	0.256	0.235	1.49	-0.15
25,000,000	0.0559	0.236	0.21	1.36	-0.133
30,000,000	0.0511	0.224	0.192	1.17	-0.0872
35,000,000	0.0475	0.206	0.178	1.17	-0.0759

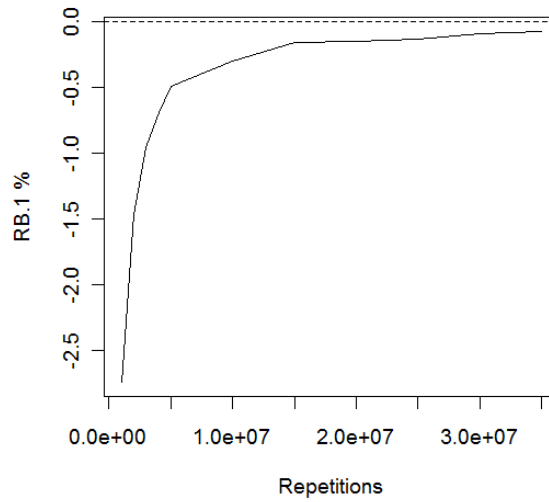


Figure: RB.1, 1-35 million repetitions.

The Figures above show RB.1 from one million repetitions, which gives a more precise presentation of the development of the measure in the end of the simulation.

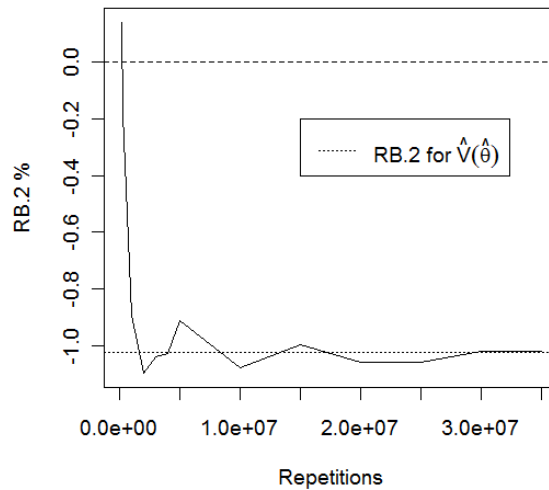


Figure: RB.2 for $\hat{V}^*(\hat{\theta}^*)$.

Table: Development of the empirical distribution measures for $\hat{V}^*(\hat{\theta}^*)$.

Repetitions	L %	U %	ER %	RB.2 %	RRMSE %
125,000	9.15	0.751	9.9	0.14	160
250,000	8.98	0.506	9.49	-0.183	151
500,000	8.95	0.435	9.38	-0.447	147
1,000,000	8.93	0.408	9.34	-0.892	144
2,000,000	8.91	0.39	9.29	-1.1	142
3,000,000	8.9	0.386	9.28	-1.04	142
4,000,000	8.92	0.39	9.31	-1.03	141
5,000,000	8.92	0.39	9.31	-0.913	142

10,000,000	8.9	0.388	9.29	-1.08	141
15,000,000	8.9	0.39	9.29	-0.998	141
20,000,000	8.91	0.39	9.29	-1.06	141
25,000,000	8.91	0.389	9.3	-1.06	141
30,000,000	8.91	0.389	9.3	-1.02	141
35,000,000	8.9	0.388	9.29	-1.02	141

Table: Development of τ .

Repetitions	Min %	P 0.5 %	Mean %	τ %	P 99.5 %	Max %
125,000	-150	-67.1	0.671	15.9	69.3	152
250,000	-125	-46	0.484	11.2	48.4	121
500,000	-78.3	-31.6	0.0577	7.86	34.3	81
1,000,000	-55.1	-23.1	-0.127	5.56	23.9	58.2
2,000,000	-37.1	-16.7	-0.077	3.95	16.4	46.8
3,000,000	-30.4	-13.5	-0.0423	3.21	13.6	32.3
4,000,000	-32.4	-11.8	-0.0389	2.78	11.9	26.7
5,000,000	-29.8	-10.4	0.0167	2.5	10.6	26.2
10,000,000	-20.5	-7.41	-0.047	1.77	7.39	17.6
15,000,000	-13	-5.92	-0.0256	1.44	6.2	15.4
20,000,000	-14.9	-5.37	-0.0458	1.25	5.26	11.6
25,000,000	-11.6	-4.7	-0.0596	1.12	4.73	11.6
30,000,000	-9.94	-4.33	-0.042	1.02	4.4	10.9
35,000,000	-9.55	-3.97	-0.0429	0.945	4.08	11.3

Table: Development of v .

Repetitions	Min %	P 0.5 %	Mean %	v %	P 99.5 %	Max %
125,000	-0.494	-0.322	0.0015	0.114	0.453	0.647
250,000	-0.339	-0.224	-0.00506	0.077	0.275	0.42
500,000	-0.295	-0.182	-0.0106	0.0614	0.249	0.343
1,000,000	-0.187	-0.11	-0.00869	0.0367	0.129	0.184
2,000,000	-0.134	-0.0877	-0.00256	0.0268	0.0818	0.139
3,000,000	-0.0997	-0.0648	-0.00325	0.0206	0.0657	0.109
4,000,000	-0.0815	-0.0538	-0.00246	0.0167	0.054	0.0879
5,000,000	-0.0738	-0.0469	0.00202	0.0166	0.0691	0.101
10,000,000	-0.0621	-0.0392	-0.0035	0.0112	0.0307	0.0579
15,000,000	-0.0496	-0.0316	-0.00188	0.0103	0.0358	0.0513
20,000,000	-0.0468	-0.0292	-0.00428	0.00842	0.0196	0.0406
25,000,000	-0.0407	-0.0277	-0.00395	0.00774	0.0182	0.0348
30,000,000	-0.0371	-0.0254	-0.00188	0.0072	0.0199	0.0328
35,000,000	-0.0361	-0.0236	-0.00227	0.007	0.0196	0.0306

Population B, SI, n=174, 35 million repetitions.

Estimator	L	U	ER	RB.2	RRMSE	0.05%	τ	99.5%	0.05%	v	99.5%
$\hat{V}^*(\hat{\theta}^*)$	8.9	0.39	9.3	-1.02	141	-3.97	0.95	4.08	-0.024	0.0070	0.020
$\hat{V}(\hat{\theta})$	8.9	0.39	9.3	-1.02	141						

It takes 5 hour, 24 minutes and 35 seconds to produce a matrix with the inclusion probabilities, based on 35 million repetitions, for this design.

6.6 PPS simulation with population B, n=90

Rosén's variance for the whole population has a negative Relative Bias compared to $V_{sim}^*(\hat{\theta}_\lambda)$, -0.06%.

Table: PPS simulation with population B, n=90.

Repetitions	$\varphi_1\%$	$\psi_1\%$	RB.1 %
125,000	1.5	8.56	-24.4
250,000	1.07	6.42	-12.1
500,000	0.77	4.78	-6.42
1,000,000	0.527	2.82	-3.26
2,000,000	0.37	2.13	-1.47
3,000,000	0.295	1.78	-1.19
4,000,000	0.264	1.39	-0.969
5,000,000	0.237	1.37	-0.796
10,000,000	0.166	0.911	-0.254
15,000,000	0.136	0.777	-0.166
20,000,000	0.119	0.617	-0.0617
25,000,000	0.106	0.653	-0.0394
30,000,000	0.0965	0.533	-0.044
35,000,000	0.0897	0.531	-0.0562
40,000,000	0.0843	0.522	-0.0476
50,000,000	0.0748	0.416	-0.0384
60,000,000	0.0697	0.352	-0.0413
70,000,000	0.0644	0.353	-0.0526
80,000,000	0.06	0.303	-0.0305
100,000,000	0.0532	0.282	-0.00515
120,000,000	0.0484	0.266	-0.0231

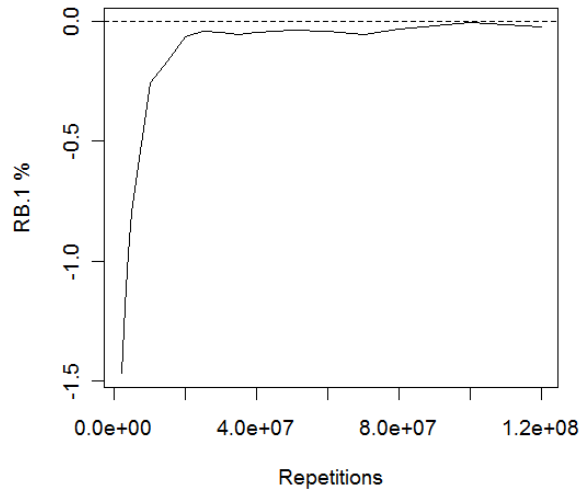


Figure: RB.1, 2-120 million repetitions.

The Figures above show RB.1 from two million repetitions, which gives a more precise presentation of the development of the measure in the end of the simulation.

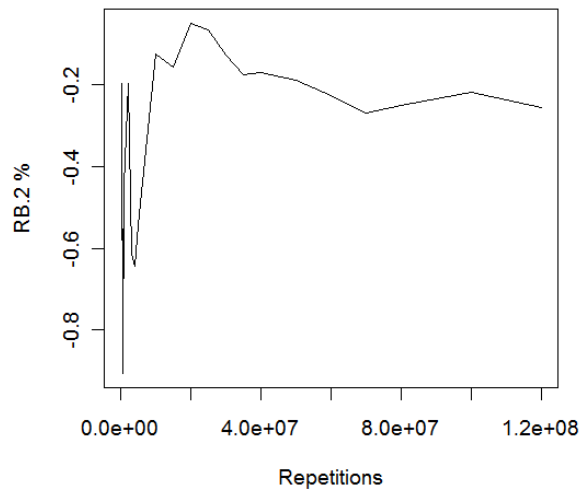


Figure: RB.2 for $\hat{V}^*(\hat{\theta}^*)$.

Table: Development of the empirical distribution measures for $\hat{V}^*(\hat{\theta}^*)$.

Repetitions	L %	U %	ER %	RB.2 %	RRMSE %
125,000	20.8	0.545	21.3	-0.579	286
250,000	20.3	0.252	20.5	-0.196	275
500,000	20.1	0.112	20.2	-0.907	266
1,000,000	19.9	0.07	20	-0.444	264
2,000,000	19.8	0.057	19.9	-0.197	262
3,000,000	19.8	0.053	19.9	-0.615	261
4,000,000	19.8	0.05	19.9	-0.644	260

5,000,000	19.8	0.052	19.9	-0.542	260
10,000,000	19.8	0.052	19.9	-0.125	261
15,000,000	19.8	0.052	19.9	-0.157	261
20,000,000	19.8	0.053	19.9	-0.0494	261
25,000,000	19.8	0.053	19.9	-0.0642	261
30,000,000	19.8	0.053	19.9	-0.126	261
35,000,000	19.8	0.052	19.9	-0.176	261
40,000,000	19.8	0.053	19.9	-0.17	261
50,000,000	19.8	0.052	19.9	-0.189	261
60,000,000	19.8	0.051	19.9	-0.225	261
70,000,000	19.8	0.052	19.9	-0.269	261
80,000,000	19.8	0.051	19.9	-0.249	261
100,000,000	19.8	0.051	19.9	-0.217	261
120,000,000	19.8	0.051	19.9	-0.255	261

Table: Development of τ .

Repetitions	Min %	P 0.5 %	Mean %	τ %	P 99.5 %	Max %
125,000	-370	-118	1.44	26	131	395
250,000	-258	-82.8	1.36	18.3	89.3	295
500,000	-150	-58.8	1.02	12.8	61.3	179
1,000,000	-106	-40.7	1.14	9.11	44	150
2,000,000	-93.4	-28.4	1.21	6.52	32.1	81.9
3,000,000	-61.8	-23	1.16	5.34	25.5	64.5
4,000,000	-60.3	-19.5	1.17	4.67	22.6	58.9
5,000,000	-47	-17.4	1.2	4.24	20.7	52.4
10,000,000	-36.8	-12	1.11	3.13	14.9	38.1
15,000,000	-30.4	-9.74	1.14	2.68	12.4	35.6
20,000,000	-26.9	-8.41	1.14	2.43	11	33.4
25,000,000	-26.9	-7.51	1.14	2.25	9.98	27.7
30,000,000	-21.6	-6.78	1.11	2.11	9.35	23.4
35,000,000	-21	-6.15	1.1	2.02	8.85	20
40,000,000	-18.6	-5.7	1.11	1.95	8.41	21.2
50,000,000	-19.8	-5.23	1.11	1.85	7.83	18.9
60,000,000	-17.6	-4.8	1.09	1.77	7.25	15.9
70,000,000	-14.7	-4.38	1.1	1.71	6.95	15.9
80,000,000	-12.8	-4.07	1.1	1.67	6.65	17.9
100,000,000	-10.5	-3.59	1.11	1.61	6.21	13.2
120,000,000	-9.92	-3.37	1.11	1.56	5.88	13.7

Table: Development of v .

Repetitions	Min %	P 0.5 %	Mean %	v %	P 99.5 %	Max %
125,000	-0.685	-0.42	-0.00505	0.128	0.411	0.672
250,000	-0.466	-0.276	0.00294	0.0865	0.288	0.494
500,000	-0.324	-0.206	-0.00503	0.0635	0.201	0.326
1,000,000	-0.248	-0.138	-0.00018	0.0433	0.143	0.266
2,000,000	-0.174	-0.0975	0.00222	0.0305	0.103	0.171

3,000,000	-0.138	-0.084	-0.00245	0.0259	0.082	0.142
4,000,000	-0.124	-0.0763	-0.00177	0.0234	0.0765	0.174
5,000,000	-0.104	-0.0675	-0.00126	0.0208	0.0682	0.12
10,000,000	-0.0752	-0.044	0.000291	0.0136	0.0444	0.0691
15,000,000	-0.0607	-0.0362	0.00029	0.0111	0.0361	0.061
20,000,000	-0.0528	-0.0317	0.00113	0.0103	0.0339	0.0518
25,000,000	-0.0449	-0.0287	0.000885	0.00924	0.0304	0.0533
30,000,000	-0.0439	-0.0267	0.000428	0.00827	0.0268	0.0448
35,000,000	-0.045	-0.0246	0.000115	0.0075	0.024	0.0399
40,000,000	-0.0392	-0.0226	0.000383	0.00711	0.0233	0.0413
50,000,000	-0.033	-0.0206	0.000222	0.00637	0.021	0.0376
60,000,000	-0.0306	-0.0189	-0.000116	0.00576	0.0188	0.0346
70,000,000	-0.0277	-0.0172	-0.000411	0.00528	0.0173	0.0325
80,000,000	-0.0254	-0.0161	-0.000108	0.00493	0.0163	0.0287
100,000,000	-0.0241	-0.0143	0.000376	0.0045	0.0149	0.0233
120,000,000	-0.0228	-0.0132	-6.49e-005	0.00395	0.0128	0.0221

Population B, PPS, n=90, 120 million repetitions.

Estimator	L	U	ER	RB.2	RRMSE	0.05%	τ	99.5%	0.05%	v	99.5%
$\hat{V}^*(\hat{\theta}^*)$	19.8	0.05	19.9	-0.25	261	-4.48*	0.45*	4.77*	-0.013	0.0039	0.013
$\hat{V}_R(\hat{\theta}_\lambda)$	20.0	0.06	20.0	-0.25	258						
$\hat{V}_{AT}^*(\hat{\theta}_\lambda)$	19.9	0.05	19.9	0.90	265						

* value has been subtracted with the mean.

It takes 15 hours, 20 minutes and 37 seconds to produce a matrix with the inclusion probabilities, based on 120 million repetitions, for this design.

6.7 PPS simulation with population B, n=45

Rosén's variance for the whole population has a negative Relative Bias compared to $V_{sim}^*(\hat{\theta}_\lambda)$, -0.04%.

Table: PPS simulation with population B, n=45.

Repetitions	$\varphi_1\%$	$\psi_1\%$	RB.1 %
125,000	2.11	12.3	-48.4
250,000	1.5	8.86	-23.5
500,000	1.06	6.48	-11.8
1,000,000	0.756	4.18	-6.13
2,000,000	0.529	3.16	-3.04
3,000,000	0.439	2.24	-1.89
4,000,000	0.368	2.54	-1.29
5,000,000	0.333	2.31	-1.08
10,000,000	0.233	1.49	-0.526
15,000,000	0.191	1.27	-0.467

20,000,000	0.163	1.02	-0.35
25,000,000	0.146	0.922	-0.29
30,000,000	0.136	0.817	-0.262
35,000,000	0.125	0.67	-0.24
40,000,000	0.118	0.65	-0.213
50,000,000	0.104	0.569	-0.106
60,000,000	0.0958	0.515	-0.125
70,000,000	0.0889	0.449	-0.0954
80,000,000	0.0823	0.443	-0.0766
100,000,000	0.075	0.455	-0.0813
120,000,000	0.0696	0.38	-0.0773
140,000,000	0.0633	0.391	-0.0687
160,000,000	0.0591	0.347	-0.0621
200,000,000	0.0535	0.287	-0.0451
240,000,000	0.0489	0.274	-0.0339

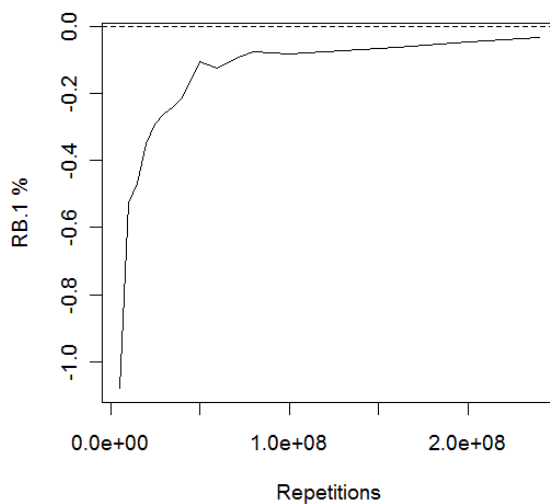


Figure: RB.1, 5-240 million repetitions.

The Figures above show RB.1 from five million repetitions, which gives a more precise presentation of the development of the measure in the end of the simulation.

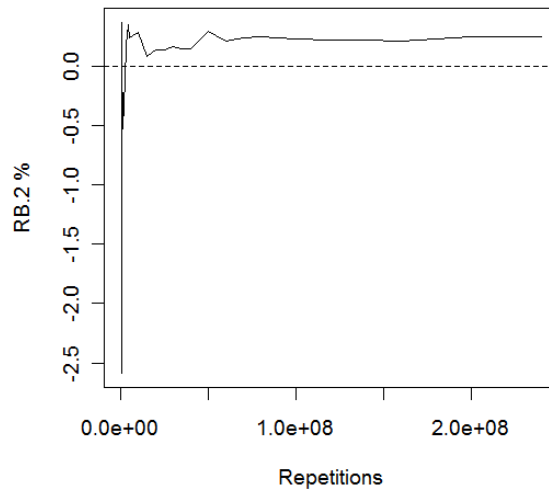


Figure: RB.2 for $\hat{V}^*(\hat{\theta}^*)$.

Table: Development of the empirical distribution measures for $\hat{V}^*(\hat{\theta}^*)$.

Repetitions	L %	U %	ER %	RB.2 %	RRMSE %
125,000	18.6	1.02	19.6	-2.59	435
250,000	18	0.705	18.7	0.365	414
500,000	17.4	0.401	17.8	0.094	394
1,000,000	17	0.235	17.3	-0.529	382
2,000,000	16.9	0.156	17	-0.12	380
3,000,000	16.8	0.135	17	0.181	380
4,000,000	16.8	0.116	16.9	0.344	380
5,000,000	16.8	0.119	16.9	0.235	379
10,000,000	16.7	0.11	16.8	0.282	378
15,000,000	16.7	0.105	16.8	0.0771	376
20,000,000	16.7	0.106	16.8	0.135	377
25,000,000	16.7	0.104	16.8	0.14	376
30,000,000	16.7	0.099	16.8	0.166	377
35,000,000	16.7	0.102	16.8	0.149	376
40,000,000	16.7	0.102	16.8	0.149	376
50,000,000	16.7	0.104	16.8	0.289	377
60,000,000	16.7	0.1	16.8	0.211	377
70,000,000	16.7	0.101	16.8	0.238	377
80,000,000	16.7	0.102	16.8	0.245	377
100,000,000	16.7	0.1	16.8	0.228	377
120,000,000	16.7	0.1	16.8	0.219	377
140,000,000	16.7	0.101	16.8	0.215	377
160,000,000	16.7	0.101	16.8	0.213	377
200,000,000	16.7	0.099	16.8	0.244	377
240,000,000	16.7	0.099	16.8	0.246	377

Table: Development of τ .

Repetitions	Min %	P 0.5 %	Mean %	τ %	P 99.5 %	Max %
125,000	-564	-202	1	45.2	245	1100
250,000	-350	-134	2.17	30.8	158	761
500,000	-286	-92.6	2.24	21.3	105	351
1,000,000	-179	-65	2.33	15	71.7	217
2,000,000	-115	-45.2	2.21	10.6	50.7	175
3,000,000	-99.5	-36.6	2.22	8.75	41.9	125
4,000,000	-74.6	-30.9	2.27	7.63	36.2	102
5,000,000	-70.6	-27.8	2.2	6.87	32.4	85.3
10,000,000	-50.1	-19.1	2.28	5.08	23.6	61.5
15,000,000	-45	-15.2	2.32	4.33	19.9	47.4
20,000,000	-38	-12.9	2.3	3.89	17.7	44.4
25,000,000	-34	-11.4	2.31	3.61	16.4	46.7
30,000,000	-34.7	-10.2	2.3	3.41	15	41.2
35,000,000	-26.9	-9.19	2.28	3.25	14	32.9
40,000,000	-26.2	-8.49	2.27	3.12	13.3	29
50,000,000	-20.6	-7.22	2.29	2.96	12.2	27.3
60,000,000	-18.8	-6.39	2.28	2.84	11.2	27.4
70,000,000	-16.9	-5.83	2.29	2.76	10.8	24.2
80,000,000	-16.1	-5.23	2.3	2.7	10.2	23.7
100,000,000	-13.3	-4.56	2.3	2.61	9.55	22.1
120,000,000	-12.7	-3.94	2.3	2.55	8.98	21.3
140,000,000	-13.3	-3.57	2.29	2.5	8.4	20
160,000,000	-12	-3.21	2.29	2.47	8.11	16.8
200,000,000	-9.14	-2.65	2.3	2.43	7.62	18.9
240,000,000	-8.44	-2.29	2.28	2.39	7.29	17.7

Table: Development of v .

Repetitions	Min %	P 0.5 %	Mean %	v %	P 99.5 %	Max %
125,000	-1.41	-0.806	0.0006	0.245	0.824	1.81
250,000	-1.06	-0.574	0.0142	0.181	0.587	1.19
500,000	-0.736	-0.402	0.00867	0.125	0.42	0.74
1,000,000	-0.574	-0.3	0.00172	0.0908	0.304	0.556
2,000,000	-0.401	-0.212	-0.00474	0.0621	0.199	0.342
3,000,000	-0.292	-0.17	-0.000668	0.05	0.161	0.266
4,000,000	-0.256	-0.149	0.000785	0.0436	0.138	0.235
5,000,000	-0.222	-0.134	-0.000479	0.039	0.124	0.21
10,000,000	-0.15	-0.0902	0.00104	0.0274	0.0881	0.164
15,000,000	-0.14	-0.0748	-0.00104	0.0231	0.0735	0.119
20,000,000	-0.114	-0.0641	-0.00137	0.0197	0.0626	0.103
25,000,000	-0.0993	-0.0585	-0.00149	0.0178	0.0563	0.0913
30,000,000	-0.0942	-0.0552	-0.00149	0.0166	0.0525	0.0937
35,000,000	-0.0874	-0.0536	-0.00184	0.0156	0.0482	0.0921
40,000,000	-0.0773	-0.0524	-0.00186	0.015	0.0452	0.0803
50,000,000	-0.0773	-0.0432	-0.000157	0.0126	0.0417	0.0761
60,000,000	-0.0764	-0.0389	-0.000851	0.0118	0.0379	0.0641
70,000,000	-0.0678	-0.0359	-0.000312	0.011	0.036	0.0654

80,000,000	-0.0634	-0.0333	4.8e-005	0.0102	0.0338	0.0624
100,000,000	-0.0552	-0.0304	-0.000126	0.00933	0.0306	0.0581
120,000,000	-0.0462	-0.0301	-0.000545	0.00885	0.0278	0.0522
140,000,000	-0.042	-0.0272	-0.000561	0.00804	0.0254	0.0465
160,000,000	-0.0475	-0.027	-0.000531	0.00763	0.0237	0.0419
200,000,000	-0.0377	-0.0226	-0.000302	0.00674	0.0209	0.0339
240,000,000	-0.0336	-0.0205	-0.000279	0.00617	0.0196	0.0318

Population B, PPS, n=45, 240 million repetitions.

Estimator	L	U	ER	RB.2	RRMSE	0.05%	τ	99.5%	0.05%	ν	99.5%
$\hat{V}^*(\hat{\theta}^*)$	16.7	0.10	16.8	0.25	377	-4.57*	0.11*	5.01*	-0.020	0.0062	0.020
$\hat{V}_R(\hat{\theta}_\lambda)$	17.1	0.11	17.2	0.30	368						
$\hat{V}_{AT}^*(\hat{\theta}_\lambda)$	16.7	0.10	16.8	1.54	383						

* value has been subtracted with the mean.

It takes 21 hour, 6 minutes and 8 seconds to produce a matrix with the inclusion probabilities, based on 240 million repetitions, for this design.

6.8 PPS simulation with population B, n=20

With this design φ_1 did not reach the limit $\alpha=0.05\%$. At 320 million repetitions we had $\varphi_1=0.063$. We can see from the table below that it takes about 560 million repetitions if we want φ_1 to be less than 0.05%.

Rosén's variance for the whole population has a negative Relative Bias compared to $V_{sim}^*(\hat{\theta}_\lambda)$, -0.1%.

Table: PPS simulation with population B, n=20.

Repetitions	$\varphi_1\%$	$\psi_1\%$	RB.1 %
125,000	3.19	17.9	-109
250,000	2.26	14.4	-53.9
500,000	1.61	10.1	-26.4
1,000,000	1.09	6.09	-13.4
2,000,000	0.782	4.74	-7.08
3,000,000	0.649	5.29	-4.4
4,000,000	0.558	3.92	-3.39
5,000,000	0.502	2.81	-2.8
10,000,000	0.363	2.03	-1.5
15,000,000	0.296	1.97	-1.05
20,000,000	0.251	1.4	-0.798
25,000,000	0.224	1.17	-0.703
30,000,000	0.206	1.12	-0.648
35,000,000	0.191	0.978	-0.499
40,000,000	0.177	1.02	-0.445
50,000,000	0.158	0.801	-0.374
60,000,000	0.146	0.811	-0.35

70,000,000	0.134	0.696	-0.26
80,000,000	0.125	0.633	-0.231
100,000,000	0.114	0.569	-0.2
120,000,000	0.104	0.595	-0.163
140,000,000	0.098	0.51	-0.166
160,000,000	0.0912	0.469	-0.161
200,000,000	0.0805	0.429	-0.171
240,000,000	0.0728	0.395	-0.179
280,000,000	0.0668	0.405	-0.164
320,000,000	0.0627	0.409	-0.169

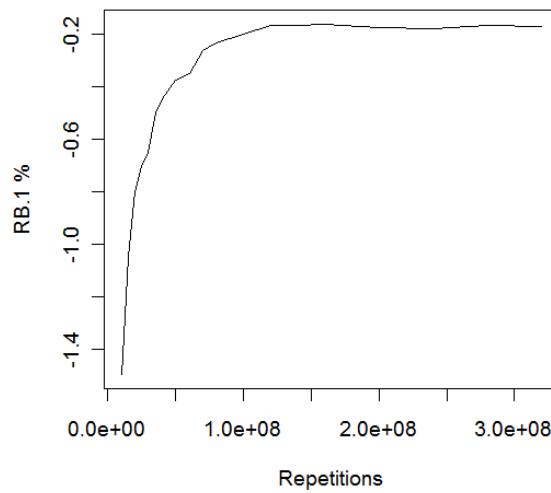


Figure: RB.1, 10-320 million repetitions.

The Figures above show RB.1 from ten million repetitions, which gives a more precise presentation of the development of the measure in the end of the simulation.

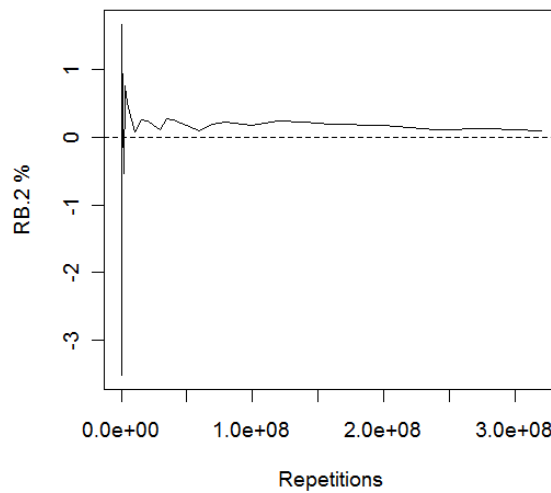


Figure: RB.2 for $\hat{V}^*(\hat{\theta}^*)$.

Table: Development of the empirical distribution measures for $\hat{V}^*(\hat{\theta}^*)$.

Repetitions	L %	U %	ER %	RB.2 %	RRMSE %
125,000	16.1	1.71	17.8	-3.53	912
250,000	17.6	1.53	19.2	-1.19	715
500,000	17.5	1.11	18.6	1.67	652
1,000,000	16.8	0.741	17.5	0.238	605
2,000,000	16.2	0.453	16.6	-0.543	582
3,000,000	16.1	0.348	16.4	0.768	586
4,000,000	16	0.313	16.3	0.612	582
5,000,000	15.9	0.274	16.1	0.431	580
10,000,000	15.7	0.243	16	0.0763	574
15,000,000	15.7	0.231	15.9	0.258	574
20,000,000	15.7	0.222	15.9	0.248	573
25,000,000	15.7	0.218	15.9	0.181	573
30,000,000	15.7	0.209	15.9	0.119	572
35,000,000	15.7	0.209	15.9	0.275	573
40,000,000	15.7	0.213	15.9	0.255	572
50,000,000	15.7	0.212	15.9	0.182	572
60,000,000	15.6	0.219	15.9	0.102	571
70,000,000	15.6	0.212	15.8	0.203	572
80,000,000	15.6	0.216	15.9	0.235	572
100,000,000	15.6	0.22	15.9	0.186	571
120,000,000	15.6	0.216	15.9	0.251	572
140,000,000	15.6	0.211	15.9	0.222	571
160,000,000	15.6	0.212	15.9	0.202	571
200,000,000	15.6	0.212	15.8	0.178	571
240,000,000	15.6	0.213	15.8	0.112	571
280,000,000	15.6	0.215	15.8	0.127	571
320,000,000	15.6	0.215	15.9	0.099	571

Table: Development of τ .

Repetitions	Min %	P 0.5 %	Mean %	τ %	P 99.5 %	Max %
125,000	-822	-402	-40.2	90	325	1320
250,000	-538	-244	-5.54	59.6	295	892
500,000	-390	-164	4.12	42.5	240	1030
1,000,000	-291	-113	5.32	29.4	157	670
2,000,000	-224	-79.3	5.36	20.7	106	401
3,000,000	-179	-65.3	5.36	17	85.2	212
4,000,000	-152	-55.3	5.35	14.9	73.2	206
5,000,000	-140	-48.6	5.35	13.5	65.8	162
10,000,000	-106	-33.3	5.25	10	46.8	112
15,000,000	-77.6	-27.1	5.24	8.59	38.4	88
20,000,000	-72.9	-22.3	5.27	7.8	33.8	79.5
25,000,000	-59.4	-19.6	5.21	7.26	31	64.8
30,000,000	-45.4	-17.4	5.27	6.94	29.1	61.7
35,000,000	-47.9	-15.1	5.27	6.67	27.3	54.2
40,000,000	-52	-14.1	5.27	6.47	25.6	55.7
50,000,000	-38.6	-12.3	5.22	6.16	23.4	55.7

60,000,000	-33.2	-10.8	5.22	5.98	21.8	55.8
70,000,000	-36.6	-9.56	5.21	5.83	20.7	45.5
80,000,000	-38.9	-8.62	5.22	5.72	19.6	42.9
100,000,000	-26.6	-7.01	5.23	5.59	18.1	37.9
120,000,000	-22.2	-5.82	5.24	5.51	17.1	35.9
140,000,000	-22.4	-5.04	5.26	5.47	16.3	33.2
160,000,000	-22.5	-4.53	5.27	5.43	15.5	30.6
200,000,000	-20.1	-3.43	5.27	5.38	14.5	30.6
240,000,000	-17.6	-2.63	5.27	5.34	13.6	28.7
280,000,000	-16.5	-2.08	5.27	5.32	13.1	26
320,000,000	-14.5	-1.71	5.25	5.3	12.7	25.4

Table: Development of v .

Repetitions	Min %	P 0.5 %	Mean %	v %	P 99.5 %	Max %
125,000	-4.12	-2.03	0.0251	0.581	1.93	3.37
250,000	-3.09	-1.41	0.00825	0.404	1.31	2.21
500,000	-1.96	-0.969	0.0149	0.293	1.02	1.83
1,000,000	-1.05	-0.661	0.00503	0.197	0.643	1.1
2,000,000	-0.879	-0.462	0.00162	0.137	0.467	0.827
3,000,000	-0.779	-0.388	0.00737	0.113	0.377	0.692
4,000,000	-0.601	-0.334	0.0044	0.0969	0.327	0.619
5,000,000	-0.487	-0.294	0.00228	0.0867	0.294	0.531
10,000,000	-0.388	-0.211	0.000402	0.0635	0.214	0.434
15,000,000	-0.356	-0.178	-0.000571	0.0519	0.174	0.315
20,000,000	-0.253	-0.147	0.000136	0.0439	0.147	0.271
25,000,000	-0.203	-0.131	-0.00139	0.0395	0.13	0.244
30,000,000	-0.214	-0.121	-0.00133	0.0368	0.123	0.228
35,000,000	-0.195	-0.111	-0.000478	0.0332	0.113	0.215
40,000,000	-0.177	-0.103	-0.00016	0.0308	0.105	0.18
50,000,000	-0.181	-0.0997	-0.000954	0.0278	0.0925	0.173
60,000,000	-0.182	-0.092	-0.00122	0.0258	0.0846	0.142
70,000,000	-0.147	-0.0822	-0.000726	0.0234	0.0772	0.12
80,000,000	-0.142	-0.0791	-0.000521	0.0221	0.072	0.117
100,000,000	-0.129	-0.0694	-0.000522	0.0199	0.0668	0.11
120,000,000	-0.131	-0.0622	-0.000351	0.0182	0.0605	0.111
140,000,000	-0.126	-0.06	-0.000215	0.0172	0.058	0.105
160,000,000	-0.108	-0.0539	-0.000256	0.0159	0.0534	0.0887
200,000,000	-0.0899	-0.049	-0.000264	0.0143	0.0477	0.0874
240,000,000	-0.0844	-0.0449	-0.000619	0.0133	0.0429	0.0737
280,000,000	-0.0798	-0.0425	-0.000569	0.0123	0.0397	0.069
320,000,000	-0.0714	-0.0417	-0.000936	0.0117	0.0371	0.0623

Population B, PPS, n=20, 320 million repetitions.

Estimator	L	U	ER	RB.2	RRMSE	0.05%	τ	99.5%	0.05%	v	99.5%
$\hat{V}^*(\hat{\theta}^*)$	15.6	0.21	15.9	0.10	571	-6.96*	0.05*	7.45*	-0.042	0.012	0.037
$\hat{V}_R(\hat{\theta}_\lambda)$	16.4	0.25	16.7	0.22	543						
$\hat{V}_{AT}^*(\hat{\theta}_\lambda)$	15.6	0.21	15.8	0.77	575						

* value has been subtracted with the mean.

It takes 23 hour, 39 minutes and 40 seconds to produce a matrix with the inclusion probabilities, based on 320 million repetitions, for this design.